UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

LEONARDO GOMES DE SOUTELLO VIDEIRA

STRATEGIES FOR DEFENDING THE PRINCIPLE OF IDENTITY OF INDISCERNIBLES: A CRITICAL SURVEY AND A NEW APPROACH

ESTRATÉGIAS PARA DEFENDER O PRINCÍPIO DE IDENTIDADE DOS INDISCERNÍVEIS: UM LEVANTAMENTO CRÍTICO E UMA NOVA ABORDAGEM

CAMPINAS
2023

LEONARDO GOMES DE SOUTELLO VIDEIRA

STRATEGIE FOR DEFENDING THE PRINCIPLE OF IDENTITY OF INDISCERNIBLES: A CRITICAL SURVEY AND A NEW APPROACH

ESTRATÉGIAS PARA DEFENDER O PRINCÍPIO DE IDENTIDADE DOS INDISCERNÍVEIS: UM LEVANTAMENTO CRÍTICO E UMA NOVA ABORDAGEM

Tese apresentada ao Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Filosofia.

Thesis presented to the Institute of Philosophy and Human Sciences of the University of Campinas in partial fulfilment of the requirements for the degree of Doctor.

Supervisor: Marco Antonio Caron Ruffino

ESTE TRABALHO CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELO ALUNO LEONARDO GOMES DE SOUTELLO VIDEIRA, E ORIENTADA PELO PROF. DR. MARCO ANTONIO CARON RUFFINO.

CAMPINAS

2023

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Filosofia e Ciências Humanas
Cecília Maria Jorge Nicolau - CRB 8/3387

V668s

Videira, Leonardo Gomes de Soutello, 1991-
Strategies for defending the principle of identity of indiscernibles : a critical survey and a new approach / Leonardo Gomes de Soutello Videira. – Campinas, SP : [s.n.], 2023.

Orientador: Marco Antonio Caron Ruffino.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas.

1. Max Black, Max, 1909-1988. 2. Identidade. 3. Metafísica. 4. Lógica. 5. Epistemologia. I. Ruffino, Marco, 1963-. II. Universidade Estadual de Campinas. Instituto de Filosofia e Ciências Humanas. III. Título.

Informações Complementares

Título em outro idioma: Estratégias para defender o princípio de identidade dos indiscerníveis : um levantamento crítico e uma nova abordagem
Palavras-chave em inglês:
Identity
Metaphysics
Logic
Epistemology
Área de concentração: Filosofia
Titulação: Doutor em Filosofia
Banca examinadora:
Marco Antonio Caron Ruffino [Orientador]
Décio Krause
Guido Imaguire
Pedro Merlussi
Emiliano Boccardi
Data de defesa: 24-03-2023
Programa de Pós-Graduação: Filosofia

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0001-7176-5658
- Currículo Lattes do autor: http://lattes.cnpq.br/7295267247809666

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

A Comissão Julgadora dos trabalhos de Defesa de Tese de Doutorado, composta pelos Professores Doutores a seguir descritos, em sessão pública realizada em 24/03/2023, considerou o candidato Leonardo Gomes de Soutello Videira aprovado.

Prof. Dr. Marco Antonio Caron Ruffino
Prof. Dr. Décio Krause
Prof. Dr. Guido Imaguire
Dr. Emiliano Boccardi
Dr. Pedro Merlussi

A ata da Defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertações/Teses e na Coordenadoria do Programa de Pós-Graduação em Filosofia do Instituto de Filosofia e Ciências Humanas.

To him, who encouraged this endeavour,

and

to her, who made it possible.

To Roberto

and

Laura.

Acknowledgments

First if all, I would like to thank my father, who recently passed away, for encouraging me to start this philosophical journey more than fourteen years ago and for sparkling the flame of curiosity in me about thirty years ago. I hope that if there is some sort of life after this one, he may be proud.

To the same degree, I must thank my mother and my brother, whose emotional support helped me get through the darkest times. I must also thank them for the financial and mental support in my career choice. Without them I would never be able to finish this Thesis. They were always my crutches when I needed, and I am sure they will be the fuel to help me go further whenever I need.

Talking about financial support, on the one hand, I must thank the São Paulo Research Foundation (FAPESP) for the financial support for the realization of this research (Process #2017/00048-4). *Obrigado*, São Paulo! I must also thank CAPES for funding a month of this research. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

On the other hand, talking about emotional and mental support, I must also thank others who helped me pass through the troubled times of a PhD aggravated by a pandemic period followed by my father's last moments. Those are my long-time friends Tennessee Monteiro, Fernando Mekaru, Erika Mayumi, Diego Sullivan, Barbara Santos, Barbara Daolio, Hugo Vedovatto, Juliana Gallina Bin, Julia Pupin, Camila 'Peipei' Ferezin, Rafael 'Axé' Rossi, Vitor 'Rapunza' Pucciarelli Antloga, 'Landa' Fernandes and the Bar do Márcio's crew: Willy Mesquita, Tulio Dorigan, Tiago Guioti, Gabriel 'Mascote' Rossi, Fabricio 'Fafá' Siansi, Frederico 'Fred' Siansi, Gustavo Wolf, Lucas Lima, Rafael 'Carinha' Batista, Alexandre 'Dimas' Noronha, Eduardo 'Japa' Teixeira Akiyama, Leonardo 'Patak' Hidalgo Batista, Paulo Leandro 'Jay' Cardoso, Felipe 'ipad' Leite; last but not least, my girlfriend Jaqueline Vidotti and her faithful squire Lalá.

I cannot forget to thank my colleagues who helped intellectually in one way or another, suggesting, discussing, revising, or commenting this work. Many helped, but I would like to thank specially, Filipe Martone, Pedro Merlussi, Raoni Arroyo, Edson Bezerra, Daniel

Coimbra, Rhamon Oliveira, Gustavo Lyra, Tiago Ponti, Thainá Demartini, Gilson Olegario and German Mejia. All of them spent considerable amounts of time making me a wee less dumb.

I would like to thank the members of the evaluation committee Décio Krause, Emiliano Boccardi, Guido Imaguire, and Pedro Merlussi, for the time dedicated to read and comment my work and for their valuable insights. They helped me to see the strengths and shortcomings of this research, and that there is still a lot to do in this field. Finally, I would like to thank my supervisor, Marco Ruffino for helping me be a better philosopher and for believing that my – let's say – unorthodox "*gororobic*" ideas were not that dismissible.

"*O ótimo é inimigo do bom.*"

– Roberto M. Videira

"*One hundred idiots make idiotic plans and carry them out.*
All but one justly fail.
The hundredth idiot,
whose plan succeeded through pure luck,
is immediately convinced he's a genius."

– Iain M. Banks

Abstract

The Principle of Identity of Indiscernibles (PII) is the focus of much controversy in the history of Metaphysics and in contemporary Physics. Many questions rover the debate about its truth or falsehood, for example, to which objects the principle applies? Which properties can be counted as discerning properties? Is the principle necessary? In other words, which version of the principle is the correct and is this version true? This thesis aims to answer this questions in order to show that PII is a necessarily true principle of metaphysics. To accomplish this task, the reader will find, in this thesis, an encyclopaedical introduction to the history of PII and to the reasons it matters so much, followed by a presentation of the most famous arguments against it and the defences used against these arguments. Then, the reader finds in-depth discussion of the minutiae involved in postulating the principle as to make clear what is in fact being attacked and defended. With these preliminaries solved, a deeper analysis of these defences is presented aiming to discover which is the most appropriate example to use against the attacks to the principle. This analysis allowed a classification of these defences in four families with different strategies within them. Finally, with these defensive strategies at hand we are able to confront alleged counterexamples to PII in Mathematics with the intention to test these defences.

Resumo

O Princípio de Identidade dos Indiscerníveis (PII) é foco de grande controvérsia ao longo da história da Metafísica e na Física contemporânea. Muitas questões rondam o debate sobre a veracidade ou falsidade deste princípio, por exemplo, a quais objetos o princípio se aplica? Quais propriedades podem ser contadas como propriedades para discernimento? O princípio é necessário? Em outras palavras, qual versão do princípio é a correta e seria ela verdadeira? Esta tese visa responder essas questões a fim de mostrar que o PII é um princípio necessariamente verdadeiro da metafísica. Para realizar esta tarefa, o leitor encontrará, nesta tese, uma introdução enciclopédica à história do PII e às razões pelas quais ele é tão relevante, seguida de uma apresentação dos argumentos mais famosos contra ele e das defesas usadas contra tais argumentos. Em seguida, o leitor encontrará uma discussão aprofundada das minúcias envolvidas na postulação do princípio de forma a deixar claro o que de fato está sendo atacado e defendido. Com essas questões preliminares resolvidas, uma análise mais profunda dessas defesas é apresentada com o objetivo de descobrir qual delas é a mais adequada para ser usada contra os ataques ao princípio. Essa análise permitiu classificar essas defesas em quatro famílias com diferentes estratégias dentro delas. Por fim, com essas estratégias defensivas em mãos, podemos confrontar supostos contraexemplos ao PII na Matemática com a intenção de testar essas defesas.

List of figures

Table of Contents

1. Introduction

This thesis' main target is (A) to defend the Principle of Identity of Indiscernibles from the dispersal argument that finds its more celebrated version in Black's thought experiment known as Black's spheres scenario, but also from derivations of it. My aim is to show that this is a flawed experiment. I will argue that the principle still holds either because (A1) the counterexample does not yield the results claimed by Black and his followers concerning the features of the spheres (as some defenders of the principle argued), or (A2) the experiment is not even conceivable, thereby, it cannot yield evidence for any conclusion whatsoever (as fewer defenders tried). To this end, first, I will present several defensive strategies that favour one possibility or the other, then, I will present my strategy, which contains both possibilities – though I favour the second disjunct.

For a fair, clear, and substantial discussion of the topic, in the rest of chapter 1, I will be firstly concerned with giving some context to the discussion and present the reasons why the debate over the truth or falsehood of PII is so relevant. Secondly, I will put forward the most prominent versions of the dispersal argument, its variations and an argument proposed by Cross (1995) that I called argument *de fictis*, which is not widely discussed, but I felt the need to present it and quickly discredit it, so that the reader is not tempted to go that way when the notion of conceivability get on our ride. Finally, I will sketch the best defending arguments available so far on the field to deal with the dispersal argument, which will be the subjects of deeper analyses in chapter 3. This whole encyclopaedic part can be viewed as (B) the secondary target of this work.

In the second chapter, I shall discuss some fundamental notions for this debate in order to make clear which notions play relevant roles, which notions have nothing to do with it, and, more importantly, how the notions play the roles they do. Furthermore, a substantive part of the solutions to quarrels about PII hinges upon these fundamental notions. Thus, theoretical choices must be made, and I shall argue in favour of these choices. Also, the jargon in this debate is not always in tune between disputants, to the point that it is not always clear that attackers and defenders are playing the same game sometimes. Thus, the tertiary goal of this work is (C) to clean some of the conceptual jungle involved in the debate and to make sure debaters are not engaging in what Brazilians call a conversation between deaf and mute.

In chapter 3, I discuss the defence arguments for PII deepening the objectives (A1), (B) and (C). I followed a somewhat standard taxonomy proposed by Hawley (2009) adding a fourth taxon to it, resulting in the following families of defence arguments: identity defences, discerning defences, summing defences, and – the one I favour – inconceivability defences. The last one is arguably not completely an original idea of mine, one can find it briefly defended by Button (2006), by Hacking (1975), and others; notwithstanding, none of them defended it as thoroughly as I am defending here and certainly not with the tools I am using. Thereby, I believe it is fair to say that, in chapter 4, I am presenting a new defence to PII against Black-like counterexamples, through an underestimated idea. This can be viewed as the deepening of (A2). Having displayed all the defences, in chapter 5, I will analyse under the light provided by those defences the mathematical counterexamples to PII, namely, the cases of imaginary numbers and – allegedly – indiscernible graphs. Finally, I wrap it all up with concluding remarks about this whole exposition in chapter 6.

1.1.Context

The Principle of Identity of Indiscernibles (PII) is, according to Rodriguez-Pereyra (2006), "*one of the most substantive and controversial ideas in metaphysic*" (p. 205-6). It is commonly attributed to Leibniz, since his version of it, in debates concerning the nature of space, time and motion, in correspondence with Clarke, is the most widely spread version; albeit versions of it can be found earlier in Cicero[1] and Nicola of Cusa.[2] The PII can be formulated in many different ways: "if any two objects share each and every property they are – or must be – the same object", or "if *a* and *b* have all the same properties, then $a = b$",[3] or "for all worlds W and all individuals x and y, if x and y are individuals in W and x and y are qualitatively identical,

[1] *Academica*, II, 85, *apud* RODRIGUEZ-PEREYRA, 2014, p. 1; III, 17, 18, *apud* MUGNAI, p. 416.
[2] *De docta ignorantia*, II, 11, *apud* MUGNAI, p. 416; III, 1, *apud* RODRIGUEZ-PEREYRA, 2014, p. 1.
[3] DELLA ROCCA (2005) p. 481.

then x = y",[4] or "There cannot be two individual things in nature which differ in number alone",[5] or "No two things differ *solo numero*",[6] among other variations.

The reader may have noticed that these formulations can be interpreted in slightly different ways which would imply considerable different outcomes in philosophical disputes – disputes that will be explained later in the introduction. To make this clearer, let us see how the principle can be spelled out in second order logic language. Let F be a second-order variable ranging over a set of properties, then, the principle can be seen as:

$$\Box\forall x\forall y\,(\forall F\,(Fx \leftrightarrow Fy) \rightarrow (x=y))$$

or

$$\forall x\forall y\,(\forall F\,(Fx \leftrightarrow Fy) \rightarrow (x=y)).$$

One of the formulations displays the necessity operator saying that the principle must hold in every possible world (whatever that means), whilst the other says nothing about modality and should be presumed to be talking about the actual world. Here the problems begin to surge. The PII is necessitated or is it obtained only in some possible worlds? If the latter is true, in which possible worlds? What about the actual world? Let us set these questions aside for a while. This point will be developed below in the introduction and analysed in depth in chapter 2.

There is a further problem that arises from the interpretation of the statement of the principle, namely, which is the correct interpretation for F? In other words, what does F stands for? Or, plainly put, which properties are the relevant ones here? According to Rodriguez-

---

[4] JESHION (2006) p. 163.

[5] LEIBNIZ (1969) p. 268. Many other formulations by Leibniz can be found in chapter 2 of RODRIGUEZ-PEREYRA (2014).

[6] RODRIGUEZ-PEREYRA (2006) p. 205.

Pereyra (2006) there are a few different possible interpretations for the PII, depending on the meaning ascribed to F,[7] for example:

PII1: No two things share all their intrinsic properties.
PII2: No two things share all their pure properties.
PII3: No two things share all their properties.

Of course, there are also the necessitated versions of these statements. These interpretations vary in strength, being PII3 the weakest and PII1 the strongest,[8] meaning that for PII1 to be false we need to find a scenario where one of the intrinsic properties differs in from *a* to *b*, while PII3 require that any property at all differs from *a* to *b* to be shown wrong. According to Rodriguez-Pereyra, Leibniz proposed the strongest one, a version that cannot be defended, as we shall see. On the other hand, the weakest is trivial and cannot be falsified, thus it is commonly deemed uninteresting. Thus, an interesting version of it must be in between PII1 and PII3, but it is not necessarily PII2. There is, perhaps, another version.

Another issue that will emerge from the interpretation of the principle is the question about which things the principle rules over. According to Rodriguez-Pereyra (2014, p. 200), Leibniz intended PII to apply to all things, however, it seems that some things violate the principle. Things such as *ficta*, *abstracta*, possible individuals and even actual subatomic particles. The question about to which things PII rules over will be dealt mainly in section 2.3., but also in scattered parts of other sections.

The principle can be understood from different angles too. A principle that is equivalent to PII, or, at least, could count as an equivalent formulation of PII is McTaggart's Principle of Dissimilarity of the Diverse, which says that: if any two things, x and y, are diverse, there is at least one property that one has that the other does not.[9] But one thing that should not be confused

[7] These three statements of the PII are from Rodriguez-Pereyra (2006), p. 206.
[8] I am following Adams (1979, p. 11) and Rodriguez-Pereyra (2006, p. 206) in this classification. However, there is a different classification that might be more useful, which we will pressent in following chapters. see SAUNDERS, 2006, p. 57; MULLER, 2015, pp. 205ff..
[9] For more on this principle and its relation to PII, see FORREST, 2016; and BABER, 2019.

with the PII is the converse principle, the Principle of Indiscernibility of Identicals, commonly referred as Leibniz's Law (thus, hereafter LL),[10] which says that for any two identical things, if one presents a property, the other must present it too:

$$\forall x \forall y ((x=y) \rightarrow \forall F (Fx \leftrightarrow Fy))$$

Whilst this principle is very seldomly disputed in the philosophical canon,[11] PII has been thoroughly discredited since Black's dialogue (1952).

Would it not make more sense if we had a principle that says:

$$\forall x \forall y (\forall F (x=y) \leftrightarrow (Fx \leftrightarrow Fy))$$

which is the conjunction of both? Or even:

$$\Box \forall x \forall y (\forall F (Fx \leftrightarrow Fy) \leftrightarrow (x=y)),$$

which is the necessitated conjunction of both? If this were the case, we would have an interesting and elegant definition (or at least an explication) for the notion of identity. This is one of the many motives for saving PII. Let us now take a look at them.

1.2. Motivation; or why is it relevant?

The conundrum over PII is interesting in itself, for it deals with some of the most fundamental notions of metaphysics. For example, our intuitions about identity are intimately linked with it, it is related with how we define our ontologies, i.e., with what we believe exists

[10] There is unclarity concerning what must be referred as Leibniz's Law in this discussion. Some call the Principle of Indiscernibility of Identicals the Leibniz's Law (Cf. HAWTHORNE, 2003; MULLER, 2015), others consider the conjunction of both principles as Leibniz's Law (Cf. FORREST, 2016; COTNOIR, 2014; BAXTER, 2014), others call the Principle of Non-identity of Discernibles the Leibniz's Law (Cf. HAWLEY, 2006). Since this seems to be a merely terminological dispute, here I will use (LL) to denote the Principle of Indiscernibility of Identicals rather than the conjunction of principles, to avoid confusions with the "PII" tag. However, I believe the title Leibniz's Law would be more properly applied to the conjunction of principles.

[11] For disputes over LL, see YABLO, 1987 and BAXTER, 2014.

in the world, and with our theories about what properties are. In other words, this is hardcore metaphysics. Being so, it directly touches many philosophical discussions and throws gasoline at the fire of some scientific discussions too. I present here a brief presentation of some of the most relevant of these discussions.[12]

1.2.1. Quantum Mechanics

Probably the area with most prolific developments for the debate over the truth or falsehood of PII in the last decades is Quantum Mechanics (QM). According to Muller (2015), this debate is relevant for establishing our ontology, i.e., for saying what kinds of things exist and what do not. If the defenders of PII are correct, there is no such thing as *indiscernibles* and there might exist a class of objects called *relationals*, roughly speaking, objects that can be understood only through their relational properties. Notwithstanding, if a particular subset of these defenders is correct, more precisely those who opt for the summing defence (e.g., SAUNDERS, 2006; HAWLEY 2006, 2009), some particles we thought to be individual objects, namely, elementary bosons and – depending on the version of PII being contested – fermions are in fact less than that, because they can only exist in relation to each other. In this case, they are entities called *non-individuals*.

If the critics of PII (e.g., FRENCH, 1989, FRENCH and KRAUSE, 2006) are correct, it follows that there are indiscernibles and there is no need to postulate the existence of relationals or non-individuals. If there are real world counterexamples against PII, it means not only that the principle is not necessary, but outright false. This certainly would have repercussions in other areas of Philosophy and Mathematics (as we shall see below).

There are also those who think that a lot of work must be done before we can reach a proper answer to the questions concerning the PII in QM cases (e.g., LADYMAN and LEITGEIB, 2008). Nonetheless, if one cannot know whether there are real world counterexamples or not to PII, to prove it false by other means would power up the views of the factions that adopt ontologies with indiscernibles. On the other hand, whether or not these

[12] Few reasons did not make the cut to be here, see HAWLEY, 2009, pp. 115-7; MULLER, p. 202.

attacks fail, a lot still must be explained in this field concerning ontology, for the kinds of entities that are needed to explain the phenomena in the spotlight are not the most conservative either.

1.2.2. Space (and time)

PII is central to Leibniz philosophy as was mentioned before, however, it was most notoriously disputed in the context of debates about the nature of space and time, most prominently in his correspondence with Clarke. There were vigorous disputes at the time concerning the nature of space and time where on one side there were the defenders of absolute space (the Newtonians) and, on the other side, the defenders of relational space (the Leibnizeans). The former defended that space was a substance, i.e., something that exists *per se*, and things existed within it. Thus, they defended that there could be empty space, which would be something homogenous composed uniformly by indiscernible parts. The latter defended that space was nothing of the sort, but something that came to existence from the relations objects bear with each other.[13]

Leibniz argued against the absolutist view from an assumption of PII by what Rodriguez-Pereyra calls the *Direct Argument*, which goes something like:

P1) If space is absolute, then its parts are indiscernible.
P2) There cannot be indiscernibles (PII).
C) Thus, space cannot be absolute.

Being so, if one successfully presents situations where there are indiscernibles, the argument fails. Thereby, it would be clear that there is at least the possibility that space is absolute.[14] In fact, one could start to build arguments for absolute space from there. This dispute seems to be

[13] These views can be understood as Primitivisms about Spatial Locations and about Objects, respectively, as Jeshion (2006; p 171) calls them. I will use both terminologies in this text, but I consider them synonymous.
[14] Rodriguez-Pereyra does not believe this is the exact argument, though he acknowledges that the arguments could be interpreted like this. For more on the topic, including readings where Leibniz uses similar arguments that substitute PII for the Principle of Sufficient Reason, see chapter 12 of Rodriguez-Pereyra (2014).

very much alive to this day (See DAITON, 2014; CASATI and VARZI, 1999) and the status about the truth or falsehood of PII is very important territory for the dispute about the nature of space (and time), because it could turn the table for either side.

In this thesis I will not dive in examples and explanations using time as the subject matter, however, the same arguments against absolute space could be transplanted to the dimension of time. Leibniz himself did it, but with less enthusiasm, as Rodriguez-Pereyra reminds us (2014, pp. 164ff., especially pp. 168-9.). Others (e.g., AYER, 1972; SWIMBURNE, 1995) tried to construct counterexamples to PII of cyclical or reversing universes that would have symmetrical time as a component, which for the sake of brevity of this thesis we will not dive in to, but I will just say that the defence arguments that use spatial properties could be transplanted for temporal properties (see MULLER, 2015, pp. 222-6).

1.2.3. Mathematical objects

There is interest in this debate for mathematicians also, more precisely for proponents of *ante rem structuralism*. They believe mathematical objects exist as positions within structures (that are universals), which means that the identity conditions for mathematical objects are given by relational properties and only by them (LADYMAN, 2005, pp. 218-19; MACBRIDE, 2006, pp. 63-64). Button (2006) also proposes this, although he is dealing with another kind of structuralism; in other words, they are nothing more than bundles of universals or parts of a universal (namely, the Structure they exist within). Thus, mathematical objects such as imaginary numbers (e.g., *i* and -*i*), that are considered to present all the same relational properties to each other are thought to be indiscernible (but not identical) within the same structure. However, this is not necessarily so. According to Ladyman (2005), there are irreflexive properties that would solve this issue, saving the PII for mathematical objects. This discussion will be the object of section 4.1.

Curiously, although Ladyman defends PII in the case above, he thinks the same case cannot be made for graph-theoretic entities such as nodes in unlabelled graphs (see LADYMAN & LEITGEB, 2008; LADYMAN, LINNEBO, PETTIGREW, 2012). He entertains the idea that some graphs can contain nodes that present no discernible graph-theoretic properties yet are

distinct entities. This would count them as counterexamples against PII. However, this view on graphs is not without controversy as we shall see in section 5.2.

Additionally, the results of these debates may have implications to the cases of Physics, however, this is not clear yet (see LADYMAN & LEITGEIB, 2008). If PII is proven false in the cases of Mathematics, this could work as a precedent for the cases in Physics, it would at least open a door for counterexamples in Physics. On the other hand, if the counterexamples in Mathematics turn out to be wrong, maybe the alleged counterexamples of Physics will require more thinking about their interpretation of the data. However, if the counterexamples in Physics fail, the cases of graph-theoretic indiscernibles within the same graphs still might come out as true counterexamples to PII. Thus, for the opponent of PII, these cases can be seen as a last resort.

1.2.4. Bundle theory versus substratum *theory debate*

There is also interest on the pure breed metaphysician's part who is debating over the ontological structure of particulars as a bundle of properties (bundle theory champions) or as a *substratum* entity which instantiate the properties (*substratum* theory champions). Briefly explained, bundle theorists believe that particulars are made of groupings of co-present attributes (e.g., the Black's spheres would be the junction of roundness, greyness, X size-ness, etc.), whereas for the *substratum* theorist there is some mysterious entity that is metaphysically prior to these attributes that exemplifies them (e.g., Black's spheres are the junction of something £ that exemplifies roundness, that is grey, that has X size, etc.).

According to Loux (1997), the bundle theorist is committed to two principles, one that he calls Bundle Theory (BT) and the other he calls the Principle of Constituent Identity (PCI). BT says that:

> (BT) "*Necessarily, for any concrete entity* a, *if for any entity* b, b *is a constituent of* a, *then* b *is an attribute"* (1997, p. 98)

and PCI says:

(PCI) "*necessarily, for any complex objects* a *and* b, *and any complex entity* c, *if* c *is a constituent of* a *if and only if* c *is a constituent of* b, *then* a *is numerically identical with* b" (1997, p. 98).

The conjunction of both principles, Loux claims, entails the PII. Thus, if the PII is false, then the *substratum* theorist can build a *modus tollens* against the conjunction of commitments of the bundle theorist. Still, according to Loux, the PCI cannot be false for it is a regulative principle that does nothing more than express a condition for the use of the terms "constituent" and "whole" in situations where wholes share constituents, therefore, BT must be false (p. 99).[15]

However, according to Loux, there are hidden assumptions in this argument that, if not assumed, reduce the conclusion of the argument to an attack on a certain specific breed of bundle theorist, namely, the metaphysical realist bundle theorist, the bundle theorist that believes in universals. For nominalist or tropist bundle theorists this argument shall not work, for they do not believe that the spheres in Black's scenario share even one attribute whatsoever. In the case of tropes, the attributes possessed by the spheres, the more similar they are, the harder it is to discern between one and the other. Also, they cannot be agglomerations of the same tropes, because it would be contradictory with the very definition of trope. If one pays attention to the metaphysical analysis of the notion of trope, one will see that it leads to the conclusion that the spheres are not sharing the same attributes, thus they are not indiscernibles. For the tropist – and presumably for the nominalist too – PII is trivially granted. Yet, this reduction of scope of the argument might go even further, as we shall see in section 2.4.2., for there is a sub-species of realist that adapted and allegedly found a way to avoid Black-like arguments, namely, the immanent realist, such as O'Leary Hawthorne (1996).

Thus, although it is not a decisive blow to bundle theory, it seems in the interest of the *substratum* theorist that the PII turns out false, for it would allow him to show at least that the transcendental realist bundle theorist is wrong. Nevertheless, since the debate between the ontological structure of particulars is intimately linked with the question over the nature of

[15] For a different position on what version of PCI is relevant to this argument and an objection for this argument, see RODRIGUEZ-PEREYRA, 2004, pp. 75-8. In this paper, Rodriguez-Pereyra argues that the connection between PII and BT is not as Loux proposes and BT is compatible with the falsity of PII.

attributes, the results of the debate above could influence the results of the debate about universals. Given that bundle theory is generally linked with tropist and nominalist approaches whilst *substratum* theory is generally linked with realism, the failure in establishing the truth of PII could tilt everything in favour of the *substratum* theorist and the truth of it from an immanent realist perspective could shake and disorganize the board in this dispute resulting in the *substratum* theorist losing some points.

1.2.5. Non-qualitative thisness (haecceity)

One of the most important papers for the opponents of PII is Adams' (1979), where he presents a clear version of the classical dispersal argument against PII and what he calls the argument from almost indiscernible twins (they will be explained in sections 1.3.1. and 1.3.2., respectively). His main motivation for putting forward these arguments is that, by showing that PII is not necessary, he opens a breach to the idea that perhaps thisness, "*the property of being this or that individual*" (*op. cit.*, p. 10),[16] is not the conjunction of all suchnesses (qualitative properties) of an object, as Leibniz understood it, but something else.

This alternative view of thisness was fiercely defended by Swimburne (1995) to be ineffable though ubiquitous. Something that is not simply *possessed* by individuals, but that which permits an individual to be what it is and that permit it to have its suchnessess. He defended the idea of a non-qualitative thisness. This theory is very dear to Christian philosophers who wish to argue for souls and for the holy trinity (see BABER, 2019).

Adams does not go so far. He claims that if there are non-qualitative thisnessess, this does not necessarily commit us to bare-particulars (*substrata*) (1979, p. 7), but given the debate exposed in the last section, one might see how non-qualitative thisnessess can be used to individualize with necessity and without qualities.[17] Of course, *substratum* theory would

---

[16] Adams stated that he is using 'property' "*as light a metaphysical load here as possible*" (p. 6). Baber (2019) also consider thisness a property, more precisely, an impure property. In this dissertation I will follow them, but I am aware that some might not feel comfortable in doing so and prefer to follow the terminology of O'Leary Hawthorne (1995) or Swinbune (1995) and say that thinessess (or haecceities) are not properties.

[17] Adams did not defend the position that necessity is given solely by non-qualitative thinesses, though he opened more doors to it. He defended that there should be some relation between thisnesses and suchnesses, though it is

already be juiced up if PII turns up false – as we just saw – but the access to non-qualitative thisness by the *substratum* theorist would make the dispute practically over for the bundle theorist. Nevertheless, before this the substratum theorist must first prove PII false, for this is a necessary condition to allow non-qualitative thisness. Thus, though it was not Adams' intention, he opened the door for such theories of obscure entities. (To be fair, he just re-opened the door, this debate remounts to the Middle Ages at least. For a general view on such debates, see BABER, 2019).

1.2.6. The Identity Principle grounds

Many believe that identity is a primitive notion, a principle that need not or cannot be grounded without falling in circular reasoning (e.g., BUENO, 2016). This can be understood as the traditional view. However, there are always the unsatisfied philosophers asking: "but why?" or "where did we get this from?" and since it seems that nothing can be considered obvious or be taken for granted anymore in our quantum non-binary days, it is not crazy to pursue this line of thought. If there really is an equivalence between PII and LL, one might wonder what came before, identity or indiscernibility? There are certain cases where identity is putative and evidently came before (e.g., jargon conventions), nevertheless, there are cases in which it might not have this way. We might have learned to identify things because of indiscernibility. If this is the case, instead of accepting that the logical principle of identity is simply an axiom, i.e., something that must be true because of its obviousness or because it is a brute fact, one might want to search for its grounds. This is considered by MacBride (2006) and Hawley (2009), demanded by Button (2006), and it is Della Rocca's main motivation (2005) as well as it is for Keaney (2007).

Thus, an approach to the logical principle of identity through PII, might be an interesting path to follow if one aims to understand logic in a – perhaps – radically naturalized manner (in Quine's sense), or simply to avoid some "*dank metaphysics*" (HAWLEY, 2009, p. 116) such

---

not clear what this relation might be. He envisaged that it might be something more of the scope of conceptual legislation than of metaphysical discovery. He named his position *moderate Haeccetism*.

as transcendental individuation, brute facts or haecceities. If one succeeds in this enterprise, this could be enough to settle the curiosity of some of these Socratic types for a while and avoid some arbitrary imports to our theories.

But even if this is impossible and we assume the view that we cannot ground identity, it is still desirable to explain what we understand by identity somehow at least to enhance our lexicon presenting more fine-grained definitions of the term "identity". This upgrade of our understanding of the notion also might have practical aims for it is desirable to know how to teach machines how to identify and discern stuff while reasoning and interacting with the world, whether we wish to develop strong AI someday. Therefore, even if we fail to ground identity, it would be interesting to enhance our explanations about what it is, in other words, enhance the ways one can introduce the very notion of identity to others.

A couple last things that must be exposed about this topic is that, first, better understanding trans-world identity is something that does not motivate defending PII, for it is not based in PII – given that trans-world identifying is the exercise of "forcing" identity to distinct and differentiable objects – (see HAWLEY, 2009) and it might even not be a necessary identity (see YABLO, 1987). In fact, arguing against PII to enable non-qualitative thisness seem to be a better strategy to stablish the necessity of transworld identity, as Adams points out (1979, p. 23-4). Second, for similar reasons, it is not interesting for a researcher of identity over time to defend PII, given that these notions seem to be different notions (I will expand a bit further on these issues in section 2.2.).

1.2.7. Theoretical virtues

It seems that if we cannot present conditions for individualizations of objects or conditions for differentiation between objects, more precisely spatial delimitations for them, there is no reason to prohibit a proliferation of exactly resembling entities co-located that share all their parts, as we shall see in sections 3.2 and 3.3., when discussing Della Rocca's (2005; 2006) argument. This would be an Ockhamnian nightmare that we hope to avoid by making quantitative parsimony a theoretical virtue in our theories. This is done in a very inelegant way by decree.

Perhaps PII could grant us this theoretical virtue in a non-legislative way. Quantitative parsimony, however, is something that is assumed in many defences of PII (Della Rocca's included).[18] Therefore, advocating for PII as a mean to justify quantitative parsimony seems circular. Nevertheless, it seems to be one of those circularities that we turn a blind eye to for the sake of elegance,[19] another desirable virtue of our theories.

Hawley also raises the point of desirability of quantitative parsimony (2009, pp. 104, 116). However, she starts from the point that such a virtue is accepted. She reminds us, nonetheless, that quantitative parsimony must be assumed together with other theoretical virtues; otherwise, we would be unknowingly walking down the road to ontological monism. Quantitative parsimony must be balanced with explanatory power, cohesion, simplicity, and other virtues; and through this balance PII might be in jeopardy.

By all means, in this dissertation, I will assume quantitative parsimony as a priority virtue. Nonetheless, others in the debate about theoretical virtues may choose otherwise and use PII to stablish quantitative parsimony, instead. Therefore, discussing the truth or falsity of PII might enrich the debate about theoretical virtues providing metaphysical support for some faction or, at least, providing examples of how to put the virtues to work.

1.2.8. Theological reasons

The development of defence of PII as proposed by Leibniz was mainly motivated by the debate about the nature of space and time, but this – as any metaphysical debate of that time – was linked to theological reasons. In this case, also associated with the Principle of Sufficient Reason and the nature of possibility (see RODRIGUEZ-PEREYRA, 2014). Leibniz's views on possibility and necessity were deeply dependent on god's will and capabilities, thus the way in which one is able to describe a possible scenario or not would deeply affect one's theological views.

---

[18] It is arguable that Della Rocca would be using (explanatory) simplicity as a virtue instead of quantitative parsimony in his argument, but at the level of abstraction of this discussion, it is hard to separate one from the other. Thus, I am making the point here thinking of a "worst case scenario".

[19] Elegance is not well defined, in general. But here let us simply assume that it is a mix of simplicity (ontological and procedural) with explanatory power (how much a theory can explain).

But this is only an indirect way PII is relevant for religious debates. There is a topic – already mentioned – in which the method of individualization is much more directly relevant to Christian doctrines, namely, the issue of the holy trinity. According to Baber (2019), there is a relevant problem in Christianity of explaining how God can be father, son and the Holy Ghost, given that these names seem to be addressed to different individuals, who possess different properties, sometimes contradictory properties. There are many different responses to this problem and most of them require a position about PII.

Since this is not a theme of my personal interest and in many aspects runs against the scientific and philosophical consensus, I shall not delve into it and shall avoid as best as I can the use of examples in this field to illustrate the points brought up in the debate over PII. Nevertheless, this dissertation might be of use for people interested in debates of this nature.[20]

1.3. Attacks on PII

As we just saw, the situation of some theories in philosophical or scientifical debates could improve a lot if PII turns out false, therefore, there is a lot of interest in making the principle false; for others, it is just a question of a lack of appeal to the principle, and for some the principle just seems patently wrong on a second thought. However, this second thought must be expressed in form of an argument. Most[21] of these arguments are designed as what Adams (1979)[22] and Jeshion (2006) call the *dispersal argument* against PII:

P1) $\Box\forall x\forall y\,(\forall F\,(Fx \leftrightarrow Fy) \rightarrow (x=y)) \rightarrow \neg\Diamond\exists x\exists y\,(\forall F\,(Fx \leftrightarrow Fy) \wedge (x\neq y))$

P2) $\Diamond\exists x\exists y\,(\forall F\,(Fx \leftrightarrow Fy) \wedge (x\neq y))$

---

[20] For about the religious debates concerning PII, I recommend the reading of Rodriguez-Pereyra (2014), and Jorati (2017) concerning the relations among god, PII, possibility, space, and time in Leibniz; and Baber (2019, specially chapter 4) concerning the holy trinity issue and PII.

[21] I would say all of them, however, the argument *de fictis* can arguably considered one that escapes this mould.

[22] To be precise, Adams separates his arguments against PII between the dispersal argument (section IV) and the argument from almost indiscernible twins (section V). Notwithstanding he admits that the latter also requires an appeal to dispersal features. Therefore, it makes more sense to call the base argument every other argument depends on the dispersal argument and qualify the different versions with the name of whatever gives the different flavour to it, e.g., in the argument named *from almost indiscernibles*, the addition of a modal step from a universe with almost indiscernible spheres to a universe with indiscernible spheres is what makes the job.

C) $\neg\Box\forall x\forall y\ (\forall F\ (Fx \leftrightarrow Fy) \rightarrow (x=y))$

A *modus tollens* that can be read as saying that P1) if PII is true, then, it is impossible to exist an *x* that has all and the same properties of some *y* without *x* and *y* being the same object. However, P2) it is possible that it happens to be such an *x*. Imagine *a scenario such and such*, that in this scenario *x* and *y* present all the same properties but are still distinct individuals. Ergo, C) PII is not necessarily true.

As we saw in section 1.2.1. above, some versions of the argument go even further as to say that in the sub-atomic parts of the actual world there are *x*s, namely, bosons (and, perhaps, entangled fermions).[23] Therefore, PII is not just unnecessary, but altogether false in the actual world. In any case, basically what changes from one argument to the other is the counterexample – the scenario – presented to support P2) and eventually the auxiliary arguments to support such qualitative arrangement.

Hawley (2009) presents some conditions according to which every criticism against PII must follow to be a genuine blow to the principle (equivalent to steps 1 and 3 below). Muller (2015, p. 204) claims that a clarifying step between them is needed if we do not wish to end up with unwarranted conclusions for the argument against PII (step 2). This intermediary step is best applied if subdivided. Thus, he divided this step in three parts and organized the whole procedure in the following steps I reproduced as commandments – the parenthetical content is product of my interpretation:

Step 1: Present a description of a qualitative arrangement.
(The scenario that will be used as a counterexample. I propose that we take description in a very loose sense to make room for depictions, i.e., pictorial representations of the scenario.)

Step 2a: Make clear what items PII meaningfully apply to in the qualitative arrangement.

[23] See Saunders (2006) and Hawley (2006; 2009).

Step 2b: Make clear what sort of features are permitted to discern one item from the other in the arrangement.

Step 2c: Make clear what sort of features are forbidden to discern one item from the other in the arrangement.

Step 3: Present an argument that in this qualitative arrangement we have distinct but indiscernible objects, in the plural.
(For this we already have the dispersal argument framed above)

In step 3, most arguments against PII will present different versions of the dispersal argument or arguments that will boil down to the dispersal argument in the end. The conditions stated in step 2 are the causes of most disagreements concerning PII. Depending on which version of PII is being debated, different features and items could be accepted or denied in the constitution of step 1, which, in turn, would yield different conclusions to the argument on step 3, which, in turn, would require different defences. Step 1 is where the inspiration to name and classify most arguments is drawn. By analysing the different scenarios used in the arguments, I propose that we divide the attacks on PII in five families, namely, classical dispersal arguments, derived dispersal arguments, arguments *de fictis*, arguments from almost indiscernibles, and special dispersal arguments. Let us expand on that.

1.3.1. Classical dispersal arguments

The first dispersal argument was proposed by Kant (2001, A 263 ff./B 319 ff.). He asks the reader whether it is not possible that there could be two droplets that have the same size, colour, etc. in two different places in the world. If the answer to such possibility is positive, this arguably could constitute a valid counterexample against PII. Max Black, in his dialogue, through the lines of character B, puts forward a more robust and actualized version of the counterexample earlier presented by Kant:

> "B: Isn't it logically possible that the universe should have contained nothing but two exactly similar spheres? We might suppose that each was made of chemically pure iron, had a diameter of one mile, that they had the same temperature, colour, and so on, and that nothing else existed. Then every quality and relational characteristic of the one would also be a property of the other. Now if what I am describing is logically possible, it is not impossible for two things to have all their properties in common. This seems to me to refute the Principle." (BLACK, 1952, p. 156)

Later in the dialogue, character B adds that this universe must be perfectly symmetrical with the intention to avoid any relation one sphere could present and the other lack.

If Kant's thought experiment can be used as a counterexample to PII, then this more robust version should also do the trick. Thus, this version of the thought experiment became the basis of the contemporary discussion over PII. Many books and papers have been written on one hand defending the PII, showing that the counterexample proposed by Black is flawed; on the other hand, several other were written enhancing the original thought experiment or creating simpler versions of it to make the case against PII (e.g. Ayer (1972); Strawson (1959), Ladyman and Leitgeib (2008), Hawley (2009), among others dealt in this thesis). Almost every counterexample to PII is somewhat based in Black's scenario – which, in turn, is based on Kant's droplets. Therefore, in this dissertation we shall use Black's scenario as the paradigmatic case of what a champion of PII would have to respond to. Notwithstanding, let us now turn to some of the other families.

1.3.2. Arguments from almost indiscernibles

This family of arguments is due to Adams' (1979), a paper where he deals with many important issues concerning PII. He admits that spatiotemporal dispersal also plays a role in these arguments (p. 17), however, given that these thought experiments start from different intuitions, it makes sense to classify it as a different family of arguments. It could be argued that there are two families here, a family of classic and a family of derived arguments from almost indiscernibles. However, given that it is debatable whether or not these variations are legit, I believe it is best to opt for a lower number of families.

Mainly aiming against Hacking's identity defence of PII, Adams proposes the following thought experiment: imagine a world very similar to Black's world, symmetrical, containing only two spheres made of iron, with the same mass, size, etc. The only difference between them is a small chemical impurity (a scratch, in some versions) on the surface of one of the spheres that the other lacks. Call this world *w*. If such world is possible, then, certainly a world in which this small chemical impurity is absent, is also possible. Call this second world free of impurities *w'*. Adams' point is that it is unreasonable to think that in *w* there were two spheres and in *w'* there is only one sphere just because the impurity has gone. Even more if we think that *w'* can be a future stage of *w*. It should be nonproblematic to think of such pairs of spheres with or without the impurity when they are apart. Why should there be a problem to do it in the same scenario? I can think of no reason for this. In other words, if *w* presented two spheres, there is no good reason to think *w'* contains more or less spheres than *w*. This, in turn, means that *w'* contains two *allegedly* indiscernible objects, given that the only difference the spheres presented in *w* was subtracted and no more differences can be presented.[24]

This argument also has variations. Adams entertains the possibility of two twin minds in which, at a certain point of their lives, present different dreams apart from every other thought that was equal one to the other. In these dreams one mind dreams of a monster with seven horns, whilst the other dreams of a ten-horned monster. According to him, there could be the case that both minds could have dreamt the same monster and, thus, have been indiscernibles.

The reader can see that it is basically the same argument using events instead of simple features as relevant properties and immaterial entities as objects of analysis (i.e., minds). I will not consider arguments of this kind as relevant, for they could not dodge the criticisms of the defences I will present. Bringing temporal features to the special ones (resulting in a different event) adds nothing to the case of the attacker of PII, given that the defences are based in spatial conditions. What could add to her cause is the fact that these are not simple events, they exist only in a mind, they are immaterial, therefore, they are not in space. This is highly questionable, however, because it hinges on a very discredited kind of dualism. The less materialistic credited view on minds today is the one that says that minds emerge from material stuff, viz., even if

[24] This argument is also called the Continuity Argument (see Forrest, 2010)

minds were emergent phenomena, a material source is partially required. Thus, some spatial property must be involved. But even if this attacker claims that one can *imagine* such a scenario, my argument in chapter 4 will take care of it on the lines of "and this images you are forming the scenario with are lacking spatial features by any chance? I believe they do not, sir".

A further derivation in the same line can be drawn if one considers that both minds were indiscernibles until the point they had the dreams with different moments. Adams points out (p. 18-9) that the minds were already distinct beings before the differing dreams occur, therefore, their distinctness (and identities) must have been primitive. In this case, I will argue that this can be seen from another angle, i.e., that the two minds (or spheres) were never indiscernibles, not even in *w'*. Even then, they still present spatiotemporal features that could be presented, but Adams seems not to be aware of them. Muller (2015) points this out, and I agree with him.

Notwithstanding, one last argument could be derived from this scenario. It is not clear to me that Adams is trying to push something in these lines in the paper, but it is possible to follow this line. One can think of worlds with almost indiscernible spheres (or minds) in which *x* has an impurity (or a seven-horned monster) and *y* has no impurity (or a ten-horned monster), call it *w* again. Then, think of a scenario where these differences were subtracted, call it *w'* too. But it is also imaginable a scenario *w"* in which *x* has no impurity (or had a dream with a ten-horned monster), whilst *y* is pure (or had a dream with the seven-horned monster). If this is the case, *w* and *w"* are indiscernible, they both present two scenarios with all and the same properties, namely, two objects that present all and the same properties but one of them. These differences are present in one object in one world and in the same object in another world. The dispersal here is given in "logical space" terms, not spatiotemporal terms[25]. For this case too I would appeal to the conceivability of such worlds. Not questioning how the attacker is imagining this world, but how is he describing these worlds. This will become clearer in section 2.1.1. Nonetheless, additionally, one might question whether possible worlds could count as objects, because objects are conceived within possible worlds, if the object is the possible world would

[25] What Adams does in the paper is to argue in favour of transworld identity of individuals that would present some indiscernibility across possible worlds. But for this, he needs PII to be left behind already (see 1979, section V). What I am proposing here is different. I am entertaining the idea of dealing with the whole scenario as the object in the relation of identity with another scenario.

it be conceived within what? This is debatable, but it makes more sense to suppose that possible worlds cannot be objects in the sense relevant for PII.

Just to make it clear for the reader, I will not comment in depth these cases dealing with issues of transworld identity and property attribution across worlds, because they are not required and because the relevant notion of identity for PII is a much more strict notion of identity; although I briefly discuss a case in which transworld identity and properties are concerned, in section 3.1.2.1., where I deal with Zimmerman's response to O'leary Hawthorne's identity defence. What Adams does in this case is trying to reach a scenario that compares only intra-world properties but derives identity through intuitions from transworld relations. This strategy was developed in response to Hacking's identity defence – which is a defence I disapprove, as it shall be clear in section 3.1.1. – with some success. Together with a few other arguments, Adams' arguments were successful in showing why Hacking's view cannot be correct. However, it has its own problems to deal with (e.g., answering which is the correct way to attribute transworld identity) and in the end it turns up into a Black's scenario, which in turn is still vulnerable to the other defences.

1.3.3. Arguments de fictis

This is, in fact, a family of a single argument. Cross (1995), after proposing a (discerning) defence against Black's scenario, argues that although this scenario fails to establish a counterexample to PII, it makes possible to extract two characters – the spheres, not A and B talking – to which we *ascribe* all and the same properties to each. According to him, the subjects in thought experiments should be treated as characters in a fictional work.

So, the properties of characters are delimited by what is settled by the story they are in. This qualitative arrangement, according to Cross, is done by ascription. The ascription of features for fictional entities does not work as simple predication of features. According to him, ascription is "(…) *a* theoretically postulated *relation that holds between a fictional character and each of the properties that specify the character's role* in the story. […]" (p. 352, emphasis added). Thus, the set up established by the author has priority over reality, metaphysics, or even logics, for what regulates its ascriptions is postulated by the author for the story.

If one is to accept this theory, one can say that, although the thought experiment described in B's story is not a description of indiscernible spheres, he ascribed this property to it. Thus, by the laws of fictional discourse, we end up with two fictional entities that disobey PII. In Cross' own words:

> "The heart of my argument can be put very simply: if, according to a given 'story', there exist two things of a given sort, then in 'real life' there exist two *characters* to which the 'story' in question ascribes the property of being of the sort in question, and yet the 'story' can be narrated in such a way that different properties are not ascribed to the two characters. Black's 'story' is precisely of this type." (p. 360. Original emphasis)

If this is correct, a defence against this argument could be based on the very same distinction it relies on. Metaphysics, as we shall see it in chapter 2, is about how the world is or could/can be – of course, also about setting the boundaries of what could/can and could not/cannot be – not about whatever can be expressed. Surely such fictional entities can exist as entities within a story, but not as real entities. At least, not as real as me and you. In a Quinean vocabulary, we could say that indiscernibles are part of our ideology – at best – but never part of our ontology. What can be said to be part of our ontology is the linguistic entity of (a description of) indiscernibles. Therefore, the debate of PII delves about what can be predicated, not about what can be ascribed – if ascription is really a thing. Otherwise, we should import the literature departments to the philosophy departments immediately! Of course, provided that this theory of fictional entities is the correct one.

To be fair, this treatment of subjects of thought experiments as fictional characters makes a lot of sense to me, as it makes sense to think of religious discourse in the same way. However, although it is not clear to me where to draw the line – especially if we take Quine seriously and treat the gods of Olympus as part of the same category as theoretical entities of science –, there is a line. There must be a line, for if there were not, first, there would have no need to differentiate ascription from predication; and second, there are clear cases where one's intention

is talking make-belief, whereas there are clear cases where one's intention is to talk about reality, undoubtedly. There seems to be a clear difference of discourse when one speaks of works of fiction and works of theoretical physics, namely, a claim to truthfulness. Therefore, we should not conflate both topics.

The reader might antecipate that this is not a very popular argument against PII, since it is such a problematic version that relies on the distinction between predication and ascription and a whole debatable metaphysics of *ficta*. Not even Cross revisits this line of thought in his later papers on PII (which I unfortunately will not delve into here). Thus, we will consider it refuted in this very same section and only briefly bring it out of the grave whenever it is convenient with the intention to give the reader confirmation of why this is not a good course of engaging the debate.

1.3.4. Derived dispersal arguments

Adams (1979, p. 13) claims that the classic dispersal arguments are those that use spatial separation as the relevant property that disperses one object from the other. Nevertheless, he reminds us of other kinds of scenarios where the temporal properties are the relevant ones (p. 14). For example, imagine a universe that blasts into existence, expands, then, contracts reaching its end just to blast into existence again and so on repeating the same events. A similar example is presented by Ayer (1972) using infinite similar sequences of sounds with the same volume, timbre, source location, etc.. (e.g., *… ABCD ABCD ABCD…*).

These examples have their peculiarities, and I will not have space to explore them here (pun intended!). Nonetheless, the structure of these arguments is the same as a classic dispersal argument, what differs from classical dispersal arguments are the kinds of properties that are doing the job. Given that it is possible to adapt defences presented to spatial dispersal arguments to temporal dispersal arguments, I will not delve specifically with these cases. I shall only briefly comment on them whenever they are mentioned in the literature leaving the treatment of these other dispersal arguments by analogy for the reader. The reader must only translate imagery-terminology to sound-terminology, spatial-terminology to temporal-terminology, objects-terminology to events-terminology, and so on. If the reader can do this, I believe that

defences against these other dispersal arguments can also be derived from the defences presented below and should work out just as good.

1.3.5. Special dispersal arguments

There are some derived dispersal arguments that deserve their own class, because of the peculiarities involved in the description of the scenarios used in step 1. Let us call this class special dispersal arguments. In such cases, what is disperse between the objects has nothing to do with temporal or spatial features. Spatial features – allegedly – do not even apply to some of them. These entities are found – or created – in special circumstances in which other features justify the understanding that there is not only one individual being presented in the scenario, but two (or more, in some cases) indiscernibles. Developed by philosophers of Physics and of Mathematics, these cases try to derive P2) from scenarios in circumstances of the sub-atomic world, where the otherwise usual rules may be overruled, and the existence of entities involved are still being debated. There are also cases concerning the conceptions of mathematical objects, whose ontological status is not clear at all and depends on heavy metaphysical theories backing them up, such as imaginary numbers and graphs.

In both cases, the more salient special circumstance is the – alleged – lack of spatial properties. Thus, it is expected that a defence that depends on spatial features, such as those commonly used for classical dispersal arguments would not work. However, as it happens with temporal dispersal arguments, there are also features of the scenario that allow the development of analogous defences. These feature, as it will become clear, are the same that are playing the dispersive part in the construction of the scenario.

In the chapter 5, I will deal with more attention to arguments developed in the mathematical realm, namely, the cases of indiscernible imaginary numbers and indiscernible graphs. For now, just as an appetizer, consider that according to some structuralists in mathematics, imaginary numbers, such as *i* and *-i*, lack real world instantiable properties and present symmetrical dispositions within the structure(s) they exist. This means, therefore, that they lack individuating properties, yet they are distinct from each other (see LADYMAN, 2005; MACBRIDE, 2006). Nonetheless, Ladyman (2005) presents some relational properties that

make them weakly discernible, saving PII. I will present his point of view in chapter 5 and I shall entertain alternative defences.

Concerning graphs, on the other hand, Ladyman and Leitgeib (2008) develop arguments against PII based on cases of structuralist graph theory. Their arguments were heavily opposed by a number of papers claiming that there are weakly discernible properties in some of their examples and the ones that are not are inconceivable or are not an appropriate interpretation of what graphs are (see BUTTON, 2006; DE CLERCQ, 2012). This debate will also be covered in chapter 5.

The cases in Mathemathics – mainly the ones concerning graphs – are used as theoretical bases to justify the special dispersal arguments in the cases of Physics, more precisely of Quantum Mechanics, which I will not delve in depth, for the discussion of the minutiae is too long and complex to deal with in this thesis. Also, much more research time and space to write would be needed for this. Thus, I will avoid these cases as much as possible, only dealing with them in section 3.3. when talking about the summing defence, because a good exposition of this defence requires it. Nonetheless, a brief simplified explanation for the reader is welcome. So, here we go!

QM cases deals with objects that lack determinate spatial properties – or, at least, we lack the capacity of determine them –, therefore, we cannot count on spatial properties to individuate them. The best we can grasp is a reading of the state of a system of quantum particles given by a sequence of vectors or a sum of such sequences. We know that there is more than one object in the system, because of the mass of the system, the behaviour of the system within experiments, and, in some cases, there are other dispersal properties such as the spins of the particles. However, these vectors can sometimes coincide within the system, and we end up with indiscernible particles. In this case, we end up with particles that present no distinguishable intrinsic or relational properties, e.g., they present no difference in shape, mass, charge, and no difference in direction of movement.

The cases found in such situation are entangled fermions and entangled elementary bosons.[26] According to Saunders (2006) and Hawley (2006; 2009) fermions present differences in spins, which makes the system antisymmetric, permitting us to capture a discerning feature between the particles, thus, making it in accordance with PII. Thus, the real problem is in the cases of elementary bosons (viz. bosons with no internal fermionic structure, e.g., photons), that displays spins in the same direction, which makes the system symmetric in every respect. In such cases, Saunders and Hawley look for a bold solution claiming that these particles are not even objects, given that they lack individuation conditions. I will explain this solution further in section 3.3.

1.4. Defences

There are several defences for PII. All of them trying to show that P2) of the dispersal argument – namely, $\Diamond\exists x\forall y\ (\forall F\ (Fx \leftrightarrow Fy) \wedge (x \neq y))$ – is false. In other words, the defences aim at showing that such an *x* that shares each and every feature with *y* cannot exist. The strategies for this are of the most diverse kinds. Nonetheless, Hawley (2009) proposes a taxonomy for these defences grouping them under three families, namely, the identity defences, the discerning defences, and the summing defences.[27] I mostly agree with her taxonomy, however, here I add one more family, namely, the inconceivability defence. Also, I believe that she would most likely not agree to my classification of some defences (e.g., Della Rocca's as a discerning defence).[28]

One thing the reader should keep in mind after this exposition is that, although I favour the inconceivability defence, I do not think it is incompatible with the discerning defence, which I also subscribe to. In fact, I believe that they are complementary. In case the inconceivability

[26] For further clarifications of these cases in a resumed manner, see HAWLEY, 2006; for a moderately complex treatment see SAUNDERS, 2006 and MULLER, 2015; for a deep understanding of the whole debate about identity conditions in Physics see FRENCH and KRAUSE 2006.

[27] She first proposed the distinctions in (2006), but the names went only in (2009). Also, she does not call them families.

[28] Some famous defences, such as Russell's analyticity claim (1995, p. 102) and Ayer's infinite repeating sounds (1972, p. 32ff.), are not considered in this taxonomy for they are not explored in depth here (or in the contemporary debate), nonetheless, their classification within such taxonomy is relevant work for future enterprises.

fails, the attacker would find himself trapped in a discerning defence. Additionally, I do not entirely discard the summing defence, though I believe it is best if one can avoid it. Finally, identity defences are the least promising in my eyes and I believe we are better off without them and the metaphysical baggage they bring together.

1.4.1. Identity defences

These are defences that try to show that premiss P2) is false by claiming that the allegedly two objects described in Black-like counterexamples are, in fact, only one object. O'Leary Hawthorne's defence (1995), for example, aims to show that these objects are two aspects of a single bi-located bundle of immanent universals. This is as problematic as it is a bold claim, thereby, it is not a popular one, as we shall see in section 3.1.1. Ian Hacking's defence, in turn, tries to undermine the construction of the scenario itself, which would be enough of a reason to put him in the inconceivability team (see HAWLEY, 2009); however, he claims that Black's description of the proposed scenario can be faithfully described – i.e., interpreted – in a cylindrical non-Euclidean space where the object described is presented as apparently two objects. Thus, he is classified by many as an identity defence champion (see ADAMS, 1979; FORREST, 2002).[29]

1.4.2. Discerning defences

These defences are the ones that fully accept the construction of Black-like scenarios and try to show that in there is in them at least one way to discern the objects involved. More precisely, they argue that the objects are weakly discernible. The discernibility features in most Black-like cases are spatial properties, nevertheless, this is not always the case. Concerning QM cases, which – allegedly – lack such properties, we can pick the *spin* (of a fermion) as a difference, whereas in the cases of mathematical entities, we can pick *being the additive inverse*

[29] I suspect that the correct thing to do is perhaps to separate Hacking's views in Hacking I and Hacking II defences, however, we will do as the literature does and put him in the identity defence team, given that we will not expand what his inconceivability defence could be.

of, in the case of imaginary numbers such as *i* and *-i*.[30] As proponents of defences of this kind we can name Ladyman (2005), Button (2006), for mathematical objects; for QM cases, Saunders (2006) and Muller (2015); and for Black-like regular cases, Muller again, Cross (1995), and – maybe surprisingly – Della Rocca (2005; 2008).[31] Discerning defences are accepted even among proponents of other defences (see HAWLEY, 2009), in fact, they are the most popular ones; and for good reasons: they accept many requirements from both sides of the dispute, and they avoid extravagant metaphysical claims from theories that are not widely accepted. Thus, if the reader is looking for a safe bet in the defending side, this is the one to put the money in.

1.4.3. Summing defences

Discerning defences, however, are not always available, according to some (e.g., SAUNDERS, 2006; HAWLEY, 2009). They believe that there are entities, such as bosons, that are not discernible at all, i.e., no matter how much we try, we can never find properties that one has and the other lacks. Thereby, an alternative defence is required whether one does not want to concede the unnecessity of PII, or even its actual falsity, given that bosons are part of our world. Saunders, then, proposes that such entities that seem to be indiscernible yet non-identical are not really objects, but something else. He proposes that since they lack individuality, they must be less than objects. If this is correct, they are not under the scope of PII, this is why they seem to disobey the principle. Hawley (2006; 2009) further elaborates this idea and names it the summing defence. This proposal is not without problems, but it is nonetheless an interesting account for the unusual world of QM talks.

---

[30] In the cases of graphs, we can still use spatial properties (though of a different kind).

[31] I say maybe, because although he does not develop a clear argument for discerning the objects in the scenario, he left the attacking side in a very inconvenient situation where the only reasonable exit is to accept that there is a discernible feature between the objects.

1.4.4. Inconceivability defences

Last, but not least, I present the inconceivability defence. In a certain sense, all the above were inconceivability defences, because all of them claim that the whole scenario Black puts forth is somehow inconsistent. Nevertheless, all of them accept some part of the story and try to show what is wrong from this accepted part on. What I would like to propose is to not accept any of it from the start, since we cannot accept all of it. The initial settings of the thought experiment are inconsistent and, by showing why, I aim to discourage one to accept the cogency of any argument that might be developed based on such experiment. (see BUTTON, 2006, pp. 219ff.; HAWLEY, 2009, p. 115; arguably JESHION, 2006, pp. 172-3.). Yet, I leave it open the possibility of conceiving a scenario, which would bring the proponent to see that there are discernible properties in this construction.

This might not be a novel idea. Hacking explicitly claimed that it is impossible to construct such scenario (1975, p. 249), although his defence ran in a completely different direction. Later, others vaguely sketched the same idea, e.g., Hawley (2009, p. 115) and Button (2016, pp. 219-20) hint at it; de Clercq (2018) has the same aim but does not fully commit to it; Muller (2015) seems to propose this in many of his analyses of dispersal cases, although he did not bother to unfold it; and arguably Jeshion (2006, pp. 172-3), if we make some adjustments to her terminology, although her aim is to argue against PII. The novelty proposed here, then, is the treatment I give to this idea, in other words, the degree of development I intend to apply to it.

My point can be summarized by saying that whatever is the result of the descriptions or imagery applied to construct a counterexample scenario, it must be in accordance with PII, otherwise it is either a contradiction or an incomplete depiction of the scene.[32] The impression of building a scenario contradicting PII is just this, an impression, a delusion (in the technical sense) generated by a non-holistic approach to the challenge of creating this scenario. Such

[32] Lately I be considering that it might even be a categorical mistake. Yet, more research is needed to defend this idea.

impression is permitted only because we consider building this world step by step and not everything at once. I hope to make these points clearer in chapter 4.

2. PII piece by piece

Hawley reminds us that:

> (…) We can generate different PIIs by varying the entities concerned – concrete or abstract particulars, universals too? – varying the degree of indiscernibility – intrinsic properties, purely qualitative proper ties, all properties? – or varying the modal strength – metaphysical or physical necessity, mere contingency? (Swinburne 1995 and Forrest 2002 offer alternative taxonomies.) Criteria of identity have the same form, ranging over entities of a given kind (e.g. sets), and linking their identity to indiscernibility in a specific respect (co-extensionality for sets). (2009, p. 101)

Thus, to avoid confusions concerning which version of PII I am talking about and to perform a proper analysis of the arguments against PII and their respective defences, in this chapter, I will clarify what each component in the formula "$\Box\forall x\forall y\,(\forall F\,(Fx \leftrightarrow Fy) \rightarrow (x=y))$" stands for. In doing this, I aim to comply with the stages in step 2 from Muller's procedure and go beyond it clarifying what we should understand when we say that the principle is necessary and that two things are identical.

To this end, in the following sections, I shall deal with the discussion concerning the correct way to interpret statements about possibility – or at least the more appropriate approach to deal with this notion in the context of Black's thought experiment –, in other words, what does "$\Box$" and "$\Diamond$" mean in this context. Then, I shall clarify some issues concerning identity, so that I can deal next with questions about the items that are being identified, i.e., what can be considered appropriate to occupy the roles of *x* and *y* in the formalization of the principle. Finally, I shall assess what *F* stands for, what it cannot stand for, and what is the nature of what it stands for; because it seems that most of the disagreements about PII come from disagreements about *F*.

I am aware of the depth that many of the notions discussed here have, though I cannot dive in these discussions for reasons of space. Thus, these notions shall be dealt with just enough

to put the PII discussion in motion. After all, the main objective of this dissertation is not to solve the underlying issues of the debate over PII but to propose a new defence on the game field. Being so, I made some choices so that the text can arrive at the part where the discussions directly concerned with PII are, but I hope that the treatment given to these notions here is good enough to fuel discussions later.

2.1. Logically possible = metaphysically possible?

Through character B's lines, in the passage quoted in the introduction, Black asks the reader whether such a scenario of two indiscernible iron spheres is *logically possible*. But what does it mean to ask such a thing? Or, better put, what did he mean by asking such a thing? Partially answering these questions and the ones they generate is the focus of the following chapter. The completion of these answers will be dealt in chapter 4, where my own views to these problems will be presented.

So, why not start by the general question: what does it mean to ask if something is logically possible? In contemporary philosophy, questions about modalities of possibility and necessity are widely handled in terms of possible worlds, but to answer 'what a possible world is?' is a much more complicated question that is commonly trespassed with very little resistance. The simplest – and most common – way to explain what possible worlds are is to say that they are worlds different than ours displaying ways of how things could have been instead of how they are. Though this is somewhat enlightening and almost intuitive, this cannot be the final answer, for some questions linger: 1) what rules are we assuming we can bend to form a world different from ours? E.g., can we count among them worlds where humans fly? Can we count among them worlds with square circles? In other words, what is the relation between "logical" and "possible" in our original question? 2) Where are these possible worlds? Of which stuff are they made of? Or, in a more philosophically pompous vocabulary, what is their metaphysical natures?

We will dive into those waters. But not too deep, because there are many strains and sub-strains of possible worlds theories (e.g., concretism, abstractionism, fictionalism, combinatorialism, etc.) and it would be futile to try to search in every strain and also resurface with *the* right answers, given that debates are still going on. However, I consider as futile as

such, in this case, to adopt an abstencionist approach, viz., to leave these issues unanswered, because the choices of what kind of possible worlds we are dealing with will influence the kind of the answer we expect to present concerning the dispute over PII, as I hope it will be apparent along the chapter.

The first of our lingering questions can be answered through the classification of possibilities. Some things are *physically possible* (e.g., that I was not sitting while writing this), while others are physically impossible (e.g., that instead of being attracted by Earth's gravitational field, I was being expelled). Given that our world has a physical order governed by laws and that – to a considerable extent – we understand those laws, we can confidently say that we know that we are unable to avoid obeying these laws in our everyday actions; for we are under the scope of those laws. However, we can easily think of situations – such as the last example – where we bypass them and know that in some sense – the physical one – these situations are impossible.

Nevertheless, were the laws that govern the order of our universe different, those physically impossible things could be physically possible, and we can imagine this easily – as can be attested by imagining it or by watching animations or viewing paintings and drawings of it. Those kinds of situations we call *metaphysically possible*, e.g., that I was expelled instead of attracted by the Earth; that reality only takes form when we look at it; that events run backwards. All these things, although in disagreement with our actual physical laws, seem to be in accordance with laws somewhat more basic such that it is hard to imagine situations going against them; laws such as the basic principles of classical logic (namely, laws of identity, excluded middle and non-contradiction), which are since Aristotle, at least, thought to be the basic laws of reality. It seems unimaginable, undrawable, unpaintable and ungraspable the notion of me going towards and away from the ground at the same time, or moving while being static, or any other situation violating those laws.

Thus, for a long time, many philosophers thought that matters of logical necessity and possibility could be subsumed into matters of metaphysical modality.[33] Simplifying, any

[33] Specially the moderns like Leibniz and Kant. See JORATI, 2017; and for more recent evidence, e.g., Kripke collapsing both necessities, see GENDLER and HAWTHORNE, 2002, footnote 7 in p. 4.

situation described with the form F*a*∧¬F*a* seems to be clearly impossible even to conceive, therefore, they must be metaphysically impossible too. This *Principle of contradiction* (or, as I preferer, of no-contradiction) was the guide for modality. Through it, one can introduce the idea of *impossibility* – logical and metaphysical. Which in turn permits one to derive *possibility*, i.e., something that is not a contradiction. Which in turn allows one to derive *necessity*, i.e., something that when negated is not possible – both logically and metaphysically.

However, there seems to be cases of contradictions where the form F*a*∧¬F*a* is not present, e.g., "Harry is an ophthalmologist, but he is not a medic of the eyes". This example has the form F*a*∧¬G*a* and yet it is a contradiction, because G seems to be the essence of Fs, meaning that F=G. We can solve this by saying that a contradiction is anything of the form F*a*∧¬F*a* or reducible to this form through clarifications, viz. through semantic transformations of Fs to Gs, *when appropriate*. Simple enough, though not unproblematic. Logic is intended to be analytic and *a priori*. These metaphysical matters (i.e., the stablishing or discovering of essences and meanings) are not seldomly only solvable *a posteriori* as they are certainly not analytic, as the debate over analyticity started by Quine made clear.[34] It seems a good idea, then, to separate matters of logic, that seem to be purely syntactical, from matters of metaphysics, which have a semantical component.

Besides, given the advances in the field of Logic, we now have the understanding that there are many different logics with different axioms, with different fundamental principles. Some of these logics even tolerate the negations of the basic principles of classical logic (e.g., paraconsistent logics, paracomplete logics). These non-classical logics would permit scenarios where one thing can be not identical with itself or that some events' descriptions have more than one truth value though these would be regarded by most as metaphysically impossible. To think that something disobeys such basic laws as the ones of classical logic is somewhat mind-bending, given that we cannot even imagine such things. This raises the question 'what does logical possibility amount to, then? If we cannot even imagine them, what differentiates them from nonsense? Or in a more philosophical manner, what would be the semantic value of saying

[34] For discussions of this debate see JUHL and LOOMIS (2009) or G. RUSSELL (2008).

that something is logically possible? These are legitimate questions, but this is too deep a dive that we must avoid, otherwise we might drown. Nonetheless, we can present a satisfactorily simple answer for our purposes here by saying that something is logically possible when the syntactic structure of its description is in accordance with a determinate set of axioms and theorems, being all of them purely formal. One can see then that semantical values are not really relevant for the talk about logical possibility, since logical possibility is a merely (with all due respect!) syntactical matter. Thus, it seems that what we understand as the set of things that are logically possible is larger than the set of those that we would consider metaphysically possible. In other words, the realm of metaphysical possibilities can have less rules bended than the realm of logical possibilities.

If this is correct – and here I will assume that it is –, for such a world as described by Black to be logically possible, one only has to choose an underlying logic with principles that allow it. Virtually every rule can be bent to accommodate our desires within possible worlds, thus the truth of a counterexample in such terms should be uninteresting.

However, what is at stake here is a metaphysical principle with semantical content. Thus, we should not take such logic deviations into consideration to talk about metaphysically possible worlds, nor should we think that this is what Black had in mind when he was writing this dialogue (or Leibniz, or Kant, or any of the people that are doing metaphysics about PII nowadays). For if it were the case, we would be also allowing worlds where it is possible to have square circles and wooden iron bars. Most certainly Black would not want that to be the case, for that is possibility for the absurd and anything would follow (unless you explicitly adopt a deviant logic). Black might have some divergent ideas about the application of vague concepts, but to accept the application of contradictions is such a bold move even for him, who can be considered a heterodox for his time. If one is trying to do serious metaphysics, we must assume that one is willing to talk about the structure of reality not about the structure of any made-up absurd world. It seems clear that, in the non-classical sense, it is logically possible that such world as the one described by Black exists, in the same way that it is permissible to say that *virtually any* world is logically possible. One just has to choose an adequate logic to describe it, but, again, that would be uninteresting. We are talking about a metaphysical principle; therefore, this world should have a logic that is in accordance with metaphysical principles at least.

Black was probably, like almost everybody else at the time,[35] collapsing metaphysical possibility with logical possibility. If this was the case – and it probably was –, what Black should really have written in the lines of character B is "Isn't it *metaphysically possible* that the universe should have contained nothing but two exactly similar spheres?". But what is metaphysical possibility after all?

One common answer to this question throughout the history of western philosophy is something in the lines of "whatever God can make" or "whatever is in the power of God to create". This sounds too simplistic and dodgy as a means to avoid difficult discussions about unpleasant topics, to the effect that it would only give birth to the question "but how can god make such-and-such?", which, in effect, would led to the answer "whatever God can think of (desire to)", to which in turn would have to unroll a very unconvincing line of thoughts about the existence and powers of a God, conjunct with another line that explains how his thoughts are linked to possibility.[36] It is wise not to follow this stream, for many have done it before and drowned. However, there are some interesting points in it, and it is not totally detached from what we will defend in chapter 4. Differently from most who follow this stream, I must admit that no one knows the answer. There is no clearly best answer to the question of what metaphysical possibility is yet. Nonetheless, I will do my best to present a satisfactory answer to it – at least in a way that permits us to keep searching for the truth about the PII.

Given that Metaphysics is the subject that deals with fundamental questions about reality, it makes sense to ask of a metaphysically possible event or thing what it needs to be actualizable. We have no clear evidence that there are events or things that exist disobeying the laws of classical logic. There are hypotheses about quantum events or places extremely far away in our universe that *might* not obey classical logic, but nothing widely agreed upon.[37] Additionally,

---

[35] Non-classical logics were already under development at that time, although they were not mainstream yet. Again, see footnote 7 of GENDLER and HAWTHORNE, 2002.

[36] For more on metaphysical possibility as god's will in modernity, see JORATI, 2017; RODRIGUEZ-PEREYRA, 2014.

[37] I do not accept Priestean ideas that there are real contradictions. To my knowledge, his example of (2019) is not a real contradiction, but a contradiction of attitudes, it is psychologism, an epistemic or doxastic attitude. Contradictions of attitudes are perfectly normal, but they are not about contradictory objects in the world, but contradictory about objects in the world. Additionally, they are generally not even synchronic as well stressed by Jago in (2014), for whenever they are synchronic the paradox is evident.

not even the most creative Hollywood designer could even imagine situations where things like this happen. Therefore, there is no good reason to think that principles beyond the ones of classical logic should be taken into consideration as part of metaphysical modalities. Classical logic and its appropriate modal extension[38] seem to be a good starting point for the definition of metaphysical possibility, as past philosophers thought. However, as we have seen, they would not be enough to deal with every case – remember the semantic problem from few pages above.

Thus, we need definitions for our metaphysical vocabulary, we need the definitions of the most fundamental notions of reality, the properties that will instantiate the predicates in the ideology used to talk about reality, e.g., *time*, *space*, *causality*, *property*, *individual*, *essence*, etc. The ideology may vary, according to which theory one is following, for example, although I used "essence" as an example a few pages ago, I do not really believe in what it stands for. I will not pretend I know which are the notions that should compose this fundamental list. However, there are some of them that are (and must be) widely accepted. To avoid too much controversy, I will settle for two, the ones I believe are widely held as the most fundamental, namely, *space* and *time*. Everything that is real, that exists, exists somehow *in* space and time or can be reduced to things in space and time. We can, thus, define what is metaphysically possible as anything that obeys the principles of classical logic – and its appropriate modal extension, whatever it is – and the definitions of the relevant metaphysical notions.

Back to our main problem. Since we are talking about a metaphysical principle, the world described in Black's thought experiment should also be metaphysically possible, not merely logically possible. Otherwise, we must say that this thought experiment is conducted in a world beyond the relevant scope of the subject matter and its results are not relevant to the discussion. Because as I have said, there are logically possible descriptions that do not obey PII, one simply must choose the appropriate logic and place it in a fiction with an ontology tolerant enough for it (as we shall see in 1.3.3.). On the other hand, if the experiment is feasible in a metaphysically

[38] I say its appropriate extension, because it is not clear which system of modal logic is the system of reality. It can be S4 or other not discovered yet, but for our purposes here, it is not important.

possible world, then we must accept that PII is not necessary; and even open ourselves to the possibility of it not being even contingent, unlikely as this sounds.[39]

Nonetheless, the claim that metaphysical possibility is the relevant one here brings up more questions such as "why metaphysical possibility must be imagined?" or "why imagination should be a guide to metaphysical possibility?", which can be better answered through a better understanding about the nature of the things we are attributing possibility to. This brings us to the second lingering question, about the ontology of possible worlds. By making this clear we can have a better grasp on how thoughts of things that distant from our ordinary world can be entertained by our minds and how such things could exist (if they can exist at all). But first let us clearly summarize our answer to the first lingering question.

Summing up the point made so far, the relevant notions of necessity and possibility concerning the discussion over the truth or falsehood about PII cannot be logical possibility/necessity, for this would be too wide a scope, but metaphysical possibility/necessity instead, which are modalities that obey classical extended logical principles and display accordance with the definitions of some fundamental metaphysical notions. Therefore, the counterexample proposed by the opponent of PII must also be metaphysically possible, otherwise we could say that he is describing a situation that is beyond the scope of the principle discussed. In other words, PII is a metaphysical principle that should rule the behaviours of things in the realm of metaphysical possibility. Nonetheless, to avoid begging the question about the truth or falsity of PII, we must allow the example to violate PII, but *only if* it does not violate other metaphysical or logical principles. This would show that PII is applied within the realm of metaphysics but does not range over the whole scope of this realm. There would be at least one metaphysically possible situation where the principle is not obeyed. Thereby, it should not be considered a principle that governs metaphysical stuff.

[39] For arguments against even the contingency of PII, see FRENCH; 1989.

2.1.1. Metaphysical nature of possible worlds

Let us now turn to our second lingering question. What is the nature of possible worlds? There are several different theories about this topic. The positive ones can be divided in two major groups, namely, abstractionists and concretists. The latter, also called modal realism by David Lewis (1986), who championed it, said that possible worlds are real physical existing worlds like ours that bear no causal interaction between us and them, though they can be accessed by thought. This idea is very unorthodox and is only defended by a handful of metaphysicians, which is one of the reasons why we will not follow this path. For a more standardized discussion about PII, we shall follow a more standard path along the lines of abstractionist theories, also called *ersatzsims*. Another reason for not following the concretist path is that although it is a very elegant and ontologically simple theory, anything said within such a framework is completely untestable, i.e., it is unfalsifiable and unverifiable. Given that we are coming from a more naturalistic point of view, this approach is not desirable. Nevertheless, for anything we will say here, I believe we can certainly find a concretist version of it with some conceptual gymnastics.

There are many different versions of abstractionism. Some abstractionists think possible worlds are maximal sets of propositions (e.g., ADAMS (1974)) or worldmaking sentences (e.g., JAGO (2014)); others think of it as maximal states of affairs (e.g., PlANTINGA (1974), ARMSTRONG (1986);[40] others believe they are simply fictions, some useful, some maximal, but not necessarily so (e.g., CROSS (1995)). Nonetheless, most of them believe possible worlds must be maximal things, and that they are dependent of a mind (the one that makes the abstractions), be it a human one or god's one (the exception would be Armstrong, perhaps).

A quick disclaimer about maximality. A maximal thing – be it a set of propositions or world making sentences; a sum of states of affairs; a complete descriptive sentence; or whatever – is a thing that, given any P, contains P or its negation. [41] For example, let P be the proposition

---

[40] I am following Lycan (1990-1) and considering combinatorialist approaches, such as Armstrong's, as a subset of abstractionism, though this is debatable.

[41] For a concretist version where this maximal thing is not a set, but a whole, see LEWIS, 1986, pp. 69 ff. For a remarkably interesting argument against the possibility of worlds as maximal sets of propositions, see MENZEL,

Socrates runs (or the thing 'P' stands for), a possible world must contain P, the running Socrates, or its negation (see JAGO, 2014, p. 159). This means that for any proposition one can think of, it or its negation, must be part of the world.

If we were to pursue the discussions about which view is the correct one, we would lose too much of our breath and hardly reach acceptable answers and probably drown stuck in these argumentative and terminological algae. Then, for the sake of this work, we should accept that the abstractionist version that is most widely accepted is of a strand that takes possible worlds to be maximal mind-dependent entities; thus, made of mental stuff. We will turn to the questions about the nature of this mental stuff which possible worlds are made of, but not yet. This dive will be made in chapter 4, where I will present my views on the topic these views relate to PII more closely. For now, the reader must only have clearly in mind that I believe that possible worlds are mind-dependent maximal things and that I believe that PII must be true for all these things, in other words (quick spoiler!). In other words, I believe that there is no conceivable world in which PII turns out false, which, in turn, means that I believe it is a necessary principle.

2.2. Identity

In the following section, I must give an account of what one means with "$x=y$" within the context of the PII debate, in other words, an account about what should indiscernibility imply, if PII is true, opposed to other senses in which the word "identity" is used. The logical aspects of this notion are widely accepted, whereas its metaphysical ones are widely disputed. Let us now turn to these issues.

The literature on PII is frequently silent about what identity is. Scholars generally assume that the notion of identity involved is clear – at least, clear enough not to discuss it, nor explain what they mean by the term. The only thing the literature is sometimes explicit enough about and never dissonant is saying that it is about a relation between the same number of things (generally two particulars, but we will inquire further in the next section), thus, this relation is

---

2012. For general arguments against abstractionism in general, see KOONS and PICKAVANCE, 2017, pp. 344-48, 360-1.

said to be *numerical identity* (e.g., LEIBNIZ, 1969, p. 268; RODRIGUEZ-PEREYRA, 2006, p. 205). The point of Black-like counterexamples is to show that it is possible to have distinct things with the same features, i.e., they differ in number only, which implies that their mereological sums is two things, instead of one and the same.[42]

Identity is frequently taken to be a primitive and indispensable notion, i.e., it is assumed that it is conceptually impossible to do anything without the notion of identity playing some role. Whenever we think, act, speak, conceive, we are making use of some fundamental notion of identity. Thereby, in principle, it cannot be defined without incurring in circularity[43] – contrary to one of the main motivations for defending PII, as explained in section 1.2.6., it is commonly assumed that it cannot be grounded. Yet, if this is the case, there is still need for some kind of introduction of it into vocabularies and there are many ways of introducing this notion into a vocabulary. In fact, the conjunction of PII and LL is one of the most common ways to introduce identity (see CASATI & VARZI, 1999, p. 38; COTNOIR, 2014, p. 8; BAXTER, 2014, pp. 247).

However, can we introduce it without recuring to PII? Let us see. One can introduce identity through the *extensional* criterion from set theory, which basically says that whenever two sets have all and the same members, they are co-extensional and, thus, are the same set. Nonetheless, one might object that this is not a good strategy to define identity, due to Quine's "creatures with kidneys and creatures with heart" counterexample (QUINE, 1961). It identifies two different classes of things as if they were the same and makes it basically impossible to apply any modal notion to the terms, were they defined that way. However, we will not assume that transworld identity is the same thing we are dealing with. Thus, this objection should not be a problem. What might be a problem for our aim of introducing identity without PII, nevertheless, is the fact brought up by Muller that "*the axiom of extensionality in set theory is just PII for sets*" (MULLER, 2015, p. 226.), if we interpret the elements of a set to be the

[42] Keep in mind that this mereological issue will play an important role in problems within O'Leary Hawthorne's identity defence and it inspires a creative solution displayed in Saunders (2006) and Hawley (2006; 2009) summing defence.

[43] See BUENO (2016). For different views about the primitiveness of identity, see DELLA ROCCA (2006); or KRAUSE and ARENHARDT (2019), who answers Bueno.

properties of an object.[44] Something like the thesis of *composition as identity* (CAI), which says that identity is the relation between the whole of an object and the collective of its parts, i.e., the sum of all the parts of an object is exactly what this (whole) object is.[45] This seems a good method to introduce identity.

The CAI thesis permits us to think identity as a predicate of the form "*… is the same as [*the sum of its parts*]*" or as the dyadic relation where *x* stands for the whole whilst *y* stands for the sum of the parts, or *vice-versa* (the difference between these forms will be explained below). However, this may raise some problems. One concerning a difference in numerosity (i.e., in this relation of identity it is said that one is many, and many is one) and other concerning what the parts of an object are. The first can be solved simply saying that it is a misreading of what these statements say. The confused reader conflates standards of counting when saying that "one is many" or *vice-versa*, i.e., it is one of *F* and many of *G*.[46] Concerning the same counting standards and the same criteria of identity, one is one and is identifiable as one; and many are many and are identifiable as many. Now, concerning the second problem, things might get trickier. If one only considers material parts, no big problems arise, but the scope of debate is not restricted to only material objects and their material parts. Although I am very inclined to physicalism and try to reduce most of the allegedly immaterial objects in the examples within the debate of PII to material objects, there are some objects that do not fit this framework (e.g., graphs, imaginary numbers) and there are some fundamental features of objects that are not reducible to material parts (e.g., some relations). Thus, it seems that it makes more sense to see parts as the properties and relations that form the objects. In this case, if one adopts the bundle theory of substance, which says that objects are the amalgamation of its properties and

---

[44] This is questionable, because to say that a set *has* such and such as elements, is different to say that the set *is* such and such. The set of all visible colours *is not* blue, green, purple, etc., although it *has* all of this colours; on the other hand, saying that an object that *has* all visible colours is blue, green, purple, etc. is the same as saying that it *is* made of all these colours. Nevertheless, there are other ways to introduce identity and we need not to delve further in these issues.

[45] Cotnoir (2014) presents three different versions of CAI, namely, the weak, the moderate and the strong ones. The one I am explaining here is the strong one. See page *ibid*., p. 9 and BAXTER (2014), p. 246.

[46] This is a complex debate, but there are available solutions. Baxter (2014) proposes a *Stranger Composition as Identity* thesis that explains this cross-counting identity. Nevertheless, the examples of PII do not present problems of crossing counting standards, thus, this will not bother us. The scenarios that present problems in counting standards brought up by Hawley (2006; 2009) and O'Leary Hawthorne (1994) and will be addressed, but they are not problems of cross counting.

relations,[47] then the ultimate parts of such objects are these properties and relations. This would be the same as PII and we are trying to introduce identity without appealing to PII, thus, other method of introduction is required.

If one is a substratum theorist, one could appeal to non-qualitative thisness to explain identity without using PII, i.e., one could appeal to a non-qualitative "property" to account for the identity (and the distinction) of individuals. However, I believe it is best to avoid such mysterious entities such as thisnesses. More on this issue will be said in section 2.3.1.

Another way to introduce identity is through its negation, namely, *distinctness*, a notion that seems unproblematic – at least in this debate. When two things are distinct, they are not identical; whereas if they are identical, they cannot be distinct. Now we need to define distinctness. For this task, one can use the Principle of Dissimilarity of the Diverse in conjunction with its converse, which says that two things are diverse (viz., distinct), if, and only if, they are dissimilar (viz., they are discernible); in other words, discernibility implies distinctness and vice-versa; in other language, $\forall x \forall y (\neg \forall F (Fx \leftrightarrow Fy) \leftrightarrow (x \neq y))$. One can see here that there is remarkable resemblance with the conjunction of PII and LL.

The next step, then, is to define discernibility. According to Muller (2015, p. 206ff.) discernibility is a relation between two or more objects that presents grades, i.e., there are grades of discernibility between objects.[48] We may have *absolute discernibility* between objects when they present different properties, which means that they can be easily discerned even conceptually, for one will always present a property *F* that the other will lack, e.g., "one sphere is purple, while the other is not". Simple enough. Alternatively, there is *relational discernibility*, that depends on the relations between the objects.[49] This kind of discernibility comes in two flavours, *relative discernibility* and *weak discernibility*. The former is given by an irreflexive

---

[47] This is a particular reading of the mereology of complex objects adopted by LOUX (2002). For a different account, see RODRIGUEZ-PEREYRA (2004). This strategy for the introduction of identity presupposes bundle theory is correct.

[48] Quine, in (1976), who first noticed the existence of these grades, called them grades of discriminability, but for our purposes here, there will be no distinction between "discriminability", "discernibility" and "dissimilarity".

[49] It will not be an issue for us here, however, I think it might be of some interest to some readers that there is a difference between *relational discernibility* and *absolute extrinsic discernibility*. Muller (*op. cit.*, p. 208) claims that there is a distinction between 1) relationals and 2) extrinsically absolute discernible objects. An example of 1) is a *being taller than* x and an example of 2) is a *being the tallest F* (or *tallest object*).

and asymmetric relation, whereas the latter is given by an irreflexive and symmetric. An example of relatively discernible objects could be found in a description like “one sphere is larger than the other”. We can discern them by their comparative size. On the other hand, an example of weakly discernible objects is described by “one sphere is to the left of the other”. Even in scenarios where we lack directionality, we still can agree that the objects present such relations. If two objects are relatively discernible this means that they can be weakly discernible, for they also present properties of the latter kind, however, the contrary is not the case; there are objects that do not present relatively discernible predicates, e.g., Black’s spheres, doppelgangers, etc.

According to Muller (*op. cit.*, p. 207) every monadic predicate expressing a property is logically equivalent to a dyadic predicate – and we shall follow him on this; reasons will be presented in section 2.4. This permits us to say that:

$$\vdash \text{AbsoluteDiscernible}(a,b) \rightarrow \text{RelativeDiscernible}(a,b) \rightarrow \text{WeaklyDiscernible}(a,b) \rightarrow \text{Distinct}(a,b)$$

This means that if two things are weakly discernible, then they must be distinct (i.e., not identical). Thus, if we have Identical(a,b) (i.e., non-distinct), by *modus tollens*, we have WeaklyIndiscernible(a,b) (i.e., non-WeaklyDiscernible) and so on. We have LL from the Principle of Diversity of the Dissimilars. What we have here is that identity is the negation of distinctness, which in turn is guaranteed by discernibility in any level. We can say, then, that identity – in the relevant sense – guarantees indiscernibility. Additionally, it might be considered logically equivalent to it whether we show that, by analogous steps, if we assume the Principle of Dissimilarity of the Diverse, we can reach PII. We can. However, this is not permitted, for it makes use of PII.

It seems, then, that there is no way to introduce identity in the relevant sense without using PII at some point. This brings us back to the starting point and leaves us with the following dilemma: either identity is primitive, or we need PII to define/introduce it. Notwithstanding, the above attempts to introduce identity without PII have shown the reader how the notion of identity is related to adjacent notions and serve as a motivation as to why saving PII is relevant. Now let us at least refine what we mean by “identity” in this debate. In other words, what we are looking for when we adopt PII – and what we are not looking for.

As a predicate identity can be understood in two ways, namely, as a monadic predicate (e.g., Ix, that reads "x is the *same* as such and such"), sometimes called an I-predicate, that stands for a conjunction of predicates that gives us the *criteria of identity* for some x to be *I*; or as a dyadic predicate (e.g., Ixy, that reads "x is the *same* as y") that expresses the symmetric, transitive, and reflexive relation of sameness between what *x* and *y* stand for. Both understandings are important in this debate, for another way to put PII is to say that whenever two things present the same identity criteria (identity in the first sense), we can say that these things are in a relation of identity (identity in the second sense). However, they are not always equivalent. For example, whenever we say that baby Léo bears identity with adult Léo, we don't mean that baby Léo has a beard and weighs 80+ kg (features that are part of the identity criteria of adult Léo), yet they are the same person. This is the problem of identity through time. In cases of identity through time, there is a violation of the Indiscernibility of Identicals (LL), therefore, a mismatch between the principle of identity and the principle of no-contradiction. Being so, it might be wiser to treat identity over time not as identity, but as something else. It is a different relation that in the literature is called perdurance or endurance depending on which theory one believes is correct.

In these theories of identity over time, identity is granted in different ways. Roughly speaking, in endurantism, identity between the stages of an object is given by a shared non-qualitative thisness (haecceity) of the states, meaning, the same object endures changes and acquires or loses properties. Whereas for perdurantism, roughly speaking, the stages are parts of a four-dimensional entity (the object), and the identity of these parts is somewhat putative, in the same way one can point to my foot and my hand and say that they are the same object (me). In any case, questions concerning these theories and these diachronic identities will not concern us, given that they do not concern PII and are dealing with a different "phenomenon".[50] One might insist and say that it should concern debaters of PII, because endurantist identity is given by a non-qualitative "property" that is stablished before the analysis of change, also, this could be a point in favour of the existence of non-qualitative thisnesses, a unified account of

[50] For an interesting debate about identity beyond PII, in cases of composition and alteration (change) see BAXTER, 2014.

identity for synchronic, diachronic and transworld – the next topic – identities. However, this non-qualitative thisness is controversial. More will be said about it in section 2.3.1.

Summing up this point: the notion of identity in the sense relevant for PII, Baxter says (2014, p. 247), concerns "*eternal, simple objects*", viz., unchanging objects. However, I believe this is an unnecessarily limiting proposition. We can be more flexible and follow Muller (2015, p. 208), saying that it concerns a synchronic relation between objects, i.e., the objects can underlie changes, although these changes must occur within the same interval for both objects *x* and *y*. Thus, the sense in which we are using "identity" here is very restricted. Other senses of identity as the described above engender way more complex debates than the one we are concerned and PII has very little to do with them. Thus, they are beyond the scope of this work.

There is another sense of identity in which similar problems arise. Imagine that baby Léo might have not been a boy, but a girl throughout the whole infancy – an equally delimited interval. In this case, LL is violated again – synchronically this time – yet we are talking about the same person. Well, are we? The problem here is that they are not in the same world (which makes questionable the affirmation that they are at the same time, however, let us just ignore that) and neither PII nor LL talks about *transworld identities*. Both principles are claimed to be true *for every* possible world, not *across* possible worlds.[51] The identity relation we are talking about is – and must be – necessary, viz., something true *in every* possible world. It is not crystal clear, however, that such relation could be held across possible worlds without a long debate over essences, counterparts or other available strategies to establish such relation. Yet, even if this transworld identity were clearly stablished, there would still be plenty of room for arguing for its contingency (see YABLO, 1989). Another reason to not allow transworld identities is the fact that the diagonalization (in Stalnaker's sense) requires naming the particulars involved or describing them in an individualizing way. As it will be clearer in the following sections this cannot be allowed because it would trivialize the dispute over PII.

[51] Hacking (1975, p. 255) claims that PII is not true *in* every possible world, but *about* every possible world, because he sees it as a truth about descriptions. However, it is not clear what he thinks about the principle being applied across worlds, I tend to think he would agree with me saying that it is true *in* every possible world because it is a truth *about* metaphysical possibility and think of it across possible worlds is an unwise stretching of the notion of identity in this sense.

Given that 1) the issue of how transworld identity is established is not clear, it is not clear to me that the objects bearing such relation are really the same or just putatively considered the same (i.e., counterparts); but even if it was clear how to do it, 2) in our best theories about it, transworld identification would require naming or descriptively individualizing the objects involved. In any case, it is clear that this is not the relation of identity we are talking about in the debate over PII.

It may be clear for the reader by now that the notion of identity we are concerned is very strict, therefore, there is too no space for vague identity, i.e., the idea that identity can ignore LL (thereby, the law of no-contradiction), in this debate. This is important, because if we are flexible on this point, we might be inaccurate in establishing the criteria of identity for things, which in turn might incur in wrongful attribution of concepts, which in turn might ultimately incur in wrongful establishment of identity relations between particulars. Only when the criteria of identity are established, one would be able to check whether the objects – or the names, or the variables – one is talking about have numerical identity or distinctness with something else, because only then there is a standard of counting (for more on the relevance of criteria of identity, see MULLER, 2015; BUENO, 2016; and BAXTER, 2014).

So far, we can define identity in the relevant sense as strict (non-vague), intraworld (non-transworld), synchronic (non-dyachronic), and equinumerous. Additionally, it can be expressed as a monadic predicate (I-predicate) that expresses the criteria of identity for some *x* to be such and such thing; or it can be expressed as a dyadic predicate between two bearers, say *x* and *y*, that must be the same particular, in other words, a predicate that expresses a symmetric, transitive, and reflexive relation. This is a pretty restricted notion, then.

To say more about identity I must now turn to the placeholders of this relation.

2.3. Objects, individuals, particulars, entities, items… whatever

As we just saw, it is not always clear whether we are talking about identity as a relation or as a monadic predicate that presents identity conditions. Nevertheless, both predicates are bore by placeholder variables (i.e., *x* and *y*) in debates about identity. However, it is not always clear what can be placed in these variables place. The things they can be substituted for can be physical objects, names, sets, and descriptions, e.g., *x* can be a physical object whilst *y* can stand

for a name in "this (*x*) is (=) Léo (*y*)", in cases of Russellian singular propositions, or can be a description "the author of this thesis (*x*) is (=) Léo (*y*). Combinations with these placeholders are permitted in most languages[52] and are usually not clearly specified in common discourse. This generates a myriad of philosophical problems that raise questions about whether we are talking about the same predicates when we make such different equalization (e.g., Theseus' ship, Frege's Puzzle, etc.).

Concerning the debate over PII, these variables must never be interpreted as names. To name things, acts of baptism must happen by using demonstratives or definite descriptions. In any case, individualization plays a role in it. By individualizing the spheres in a Black-like scenario, one would be breaking the symmetry of the arrangement by presenting at least one difference between the spheres, which would make them absolutely discernible (MULLER, 2015, p.213), or there could be no fact of the matter concerning which name goes with which sphere. In this case, the names would work as variables. In fact, although some philosophers in the literature use names to talk about the objects they are describing (BLACK, 1952; ADAMS, 1979; MULLER, 2015), they are not really naming them or referring to them by names, they are just hypothetically naming them or using the names as rhetoric artifices. In these cases, the names 'Castor' and 'Pollux' frequently used for describing the scenario are only variables in disguise.[53] Genuine naming is completely forbidden in this debate, for this would trivialize the problem. (More will be said about trivializing predicates in the next section.) The same goes for definite descriptions. In Black-like scenarios, one could refer to one or the other sphere by using definite descriptions. Nevertheless, this presents two problems. First, one would have to find different properties in the spheres to put in the descriptions, otherwise there would not be two descriptions, but only one, which means that there should be only one object present. Second, whether one finds such a property, this would also break symmetry and confer different individualities to the spheres, making the thought experiment flawed. This is, again, why transworld identities cannot be used in this debate. Thus, in this debate, the variables must

[52] Exceptions to this could be purely extensional languages, or purely conceptual languages, perhaps.

[53] Muller (p. 213) claims that the names cannot be eliminated. However, they can. They could be substituted by indeterminate pronouns 'one' and the "semi-determinate expressions" 'another' or 'the other', or simply '*x*' and '*y*'. Anyway, 'Castor' and 'Pollux' in such contexts are not genuine names.

always be interpreted as substitutes for objects themselves that bear the qualities given to them by the philosopher running the experiment, through the description of a scenario as a whole.

Concerning the nature of these predicate bearing objects, I must say that they are loosely defined at best, and mostly carelessly used in this debate. Some say that identity is held by individuals (e.g., STRAWSON, 1959; ADAMS, 1979), others use the term "particulars" (e.g., O'LEARY HAWTHORNE, 1994; LOUX, 2002, RODRIGUEZ-PEREYRA, 2004), others say they are objects (e.g., DELLA ROCCA, 2005, 2006; LADYMAN and LEITGEIB, 2008; MULLER, 2015), and there are those who prefer entities (HAWLEY, 2009; also considered by MULLER, 2015). Yet very few take care to define what they mean with these terms. Take "individual" for example. Adams (1979, p. 6) sees individuals basically as particulars, excluding numbers and other universals from this category; on the other hand, Strawson (1959, pp. 226-7) sees individuals not just as particulars, but basically as anything that can be placed as a subject in a proposition, including numbers, properties, classes, events, etc.; while Muller (2015, p. 206) defines individuals as an object that has at least one monadic property that other objects lack, i.e., an absolutely qualitatively discernible. Now, take the notion called "particular", it is not that simple to decide to what it applies. For example, O'Leary Hawthorne (1994, p. 192 ff.) (to a certain extent followed by Hawley (2009)) takes it to be something like an amalgamation of more primitive immanent universals, however, in the same context, Rodriguez-Pereyra (2014) (and to some extent arguably Zimmerman (1995)) argues that every instance – which I will call occurrence – of this amalgamation is to be called the "particular". In any case, they are not clear about the material – or immaterial – extent of these definitions. "Object", in turn, is assumed to apply to anything, from the most abstract things, e.g., graphs and numbers, to only clearly material things, e.g., a molecules and bricks. "Entity" is considered a more general term that could be vague enough to embrace everything we mean by *x* and *y*. However, Muller warns that it might be too vague, to a point where it embraces universals, properties, tropes, and other things considered entities that should be in the realm of *F* not of *x* and *y* in this debate. The problem of assuming properties or other entities under *F*s umbrella in the positions of *x* and *y* is that we would be entering the domain of second order logics, importing its problems to a debate that is already complex enough. Thus, I will avoid the term "entity" as the target of the discussion and use it with the widest scope.

The view I believe is the most adequate to this issue is that proposed by Muller (2015, pp. 205-6) – though I disagree with some of his claims and with his definition of individual. It is a very Fregean view in which he sees "object" as "*a purely logical and metaphysically thin*" notion. He says:

> Anything we can meaningfully quantify over qualifies as an *object* (iron spheres, elementary particles, planets, humans, dreams, novels, tree leaves, numbers, sets, structures, space-time points, etc.) Perhaps *entity* would have been a better word, but that usually also includes universals, properties, tropes, and more, which I shall not address and therefore want to exclude here. (p. 206. Original italics)

Certainly, one could predicate features from these other entities Muller wants to exclude, making them eligible for the positions of *x* and *y* in some propositions, as Strawson considered. Thus, if we are excluding these so-called entities, it means that the notion of object is not purely logical and carries some ontological weight on it. We have a leap from Frege to Quine, then. Another piece of evidence for that is the inclusion of numbers, sets and structures in face of the exclusion of other entities, which implies some ontological commitments concerning those entities that bear a heavy metaphysical load (e.g., numbers and triangles not being universals). Therefore, although the notion of objecthood Muller puts forward is arbitrary and requires some clarification, it seems good enough to employ in this debate, given that most counterexamples against PII are about things that are within the limits he drew. The ones that are not within those limits (or at least, the ones that it is not clear whether are or not), such as minds, imaginary numbers, dreams and graphs, deserve a few words.

I am assuming a (reductionist) physicalist approach to this debate. This means that minds, dreams, possible worlds, etc. are mental representations reducible to material things such as synaptic activations – or things of this sort – in the brain. These physical things, in turn, are contemplated by the defences of PII discussed here. I do not believe in spirits or in the holy ghost, thus, as said in section 1.2.8., I will not enter this field also for a matter of respect for the

ones who do believe in them and might be offended by my lack of acceptance of some premisses involved. Mathematical objects, on the other hand, are trickier. If one adopts a Platonist view, mathematical objects should be interpreted as universals and, thus, should be excluded from the domain concerning PII. If one adopts a nominalist view, the verdicts vary from "there are no mathematical objects" to "they can be reduced to mental representations too". But in the literature concerning PII the main view about mathematical objects is that they are structures. What precisely this means is controversial and is object of intense debate nowadays. If they are structures within nature, that is, within the things themselves (*in re* structuralism), the objects we deal with can be considered mental abstractions (i.e., representations) of relations in nature, therefore, subject to the PII defences presented here. Notwithstanding, the strain of structuralism discussed in the debates over PII says that structures exist as independently as within nature (*ante rem* structuralism), i.e., they exist as a universal and as particulars somehow. I will not delve into these issues, I will just assume that the Structure that pre-exist everything might be considered a particular, thus, an object and can be analysed as such.[54]

Anyway, for the purposes of this dissertation, it is important that the notion of object be malleable enough to make room for most of the things that are commonly used in examples for and against the PII; but even more important than that is that only countable things fit in, because the relevant sense of identity for this discussion requires this, as we saw in the last section. Another relevant point worth stressing is that all the things that fit under the notion of object we adopted are somehow within space-time, otherwise the defences might not work.[55]

The mathematical cases usually are considered not to be under the space-time constraints and Quantum Mechanics cases can be understood escaping these constraints. However, when I talk about them in the final sections of this dissertation, I hope it becomes clear that at least the mathematical cases, if understood as structures or part of structures, they must be under these

---

[54] For more on this debate see; LADYMAN 2005; LADYMAN and LEITGEIB, 2008; LADYMAN et al., 2012; MACBRIDE, 2006.

[55] This space-time constraint may be a weak spot to this whole work. However, it is a very costly spot to reach, for it opens the door to many controversial entities in an ontology.

constraints. Hence, if the QM cases are derived from them, they also must be under these constraints.

Since I am already talking about QM cases, it seems a good idea to anticipate some caveats. These cases present some extra-ordinary characteristics that drive the discussion to minutiae of the features they present, that is, their properties and relations. Muller (2015, pp. 207ff.) presents three classes of objects in which physical objects might fall into, namely, individuals, relationals, and indiscernibles. The first class is widely accepted, it is the class of objects that are absolutely discernible from others, that is, an object that presents at least one property[56] that is not shared with other objects. The second class, relationals, is the class he is arguing for in his paper and the one which the existence is not widely recognized. Those are objects that do not present any absolutely discernible properties but are still weakly discernible by some relational feature that they bear with other objects. He argues that some quantum objects are of this kind, namely, bosons and entangled Fermions. The last class, indiscernibles, whose existence I and him are arguing against, is the class of objects that are neither individuals nor relationals. Indiscernibles are objects that are quantitatively discernible, though qualitatively indiscernible. One should not confuse relationals, i.e., entities that can singly exist but cannot be singled out without the aid of another object, with the non-individuals proposed by Saunders (2006) and Hawley (2006; 2009), which are entities that are not even objects, i.e., they do not exist without other non-individuals that bears a relation with them. In the latter case, the minimal object in question is the system composed of these non-individuals. This notion of object is different from that used by Muller. While his seems to be more epistemically driven, Saunder's and Hawley's is more metaphysically driven. I shall, then, adapt their jargon for matters of simplicity and treat non-individuals as objects, since they can be singled out in discourse by indexical expressions, bear features – more specifically, relations – with each other and other stuff, and are countable.

[56] A monadic and non-trivializing property. These characteristics will be clarified in the next section, but for now just have in mind that they cannot be expressed by predicates using names nor polyadic ones.

Summing up, the entities they are calling “objects” I will call something else, a “thing”, perhaps, and I will keep “object” for a broader class of entities. Individuals, non-individuals, relationals, indiscernibles, they will all be loosely referred as objects here.

Another thing that must be clarified is that objects in some views such as O’Leary Hawthorne’s identity defence and Hawley’s summing defence can be *scattered simples*, i.e., they can occupy two spatial regions at the same time while being a single indivisible object. This is a very controversial thesis and I am not inclined to accept it. The reasons will become clear when such defences are discussed, but in a nutshell: the relation held between these two things is strict identity, the identity that has to do with numerosity. Hence it has to do with counting standards, which in turn, seems very intimately related with mereology. More precisely, to accept that there are perceptible different parts that are analytically inseparable seems to me as absurd as saying that there are two hundred perceptible co-located parts of an object.

Summing up, the objects that can occupy the places of *x* and *y* in PII or that must be used in qualitative arrangements to disprove PII can be of a wide variety of classes, they must only be space-time locatable, so that they can be considered existent, countable and identifiable in the relevant sense of identity, namely, numerical identity. Basically, the only thing they cannot be is be identified by their names or definite descriptions.

2.3.1. Between x*s and* F*s: some remarks*

This section deserves a few closing remarks.

One last possible interpretation for what *x* and *y* stands for are bare particulars. This would be a commitment to a *substratum* theory that I am not happy to make. Mainly for one reason, namely, the adoption of physicalism and a physicalist approach to metaphysical issues do not go well with undetectable entities. If things such as particulars with no image, no sound, no touch, baring no relations to other things, and so on were possible, then what would be the difference between them and a non-existent particular or a Saganian garage dragon? None at all. Unless we accept the existence of non-qualitative thisnesses. This feature would be sufficient condition to confer existence and individuality to *x*s and *y*s, thereby their

numerosities, without the use of descriptive properties that would break the symmetry of the arrangement in a Black-like scenario.

Quick clarifications before we dive into this discussion. Adams (1979, p. 6-7) reminds us that the non-qualitative thisness is not the same as the substratum, and its existence does not necessarily commits us to the existence of these bare particulars. What I am claiming here is quite the opposite, i.e., the commitment to substrata requires the commitment to non-qualitative thisnesses, so that one can individuate and confer existence to them proving that they are two objects instead of one object with different tags. Additionally, Adams treats non-qualitative thisness as a property (1979, p. 6). I do not see a good reason to opt for the obscure view that it is not a property apart from the fact that one wishes to keep the discourse about it surrounded by a mist of mystery. Yet, given that this is how some philosophers like to treat it, this is how I am going to treat it, as a "property" (between inverted comas) – this is an additional reason not to deal with this notion in the following section.

Given that the following section is about properties and related notions, I believe this is the best place for saying a few words on the controversial notion of thisness. Thisness or haecceity is the "property" that makes an object identical with itself, in other words, it is the property that makes it itself. This can be understood at least in two different ways, namely, 1) as something like an ultimate suchness, that is, the sum of each and every qualitative feature of an object; or 2) as something that is present in the existent object, but it is not *possessed* by it and is not a quality, therefore, non-qualitative. 1) is defended by Leibniz and others that defend PII, whilst 2) is defended by Swimburne and others that oppose to PII. In case 1) it seems clear that one can treat it as a legitimate property, a conjunctive property, and it should not be considered a problem neither for the opponent of PII nor the defendant of PII to cast it. Whereas in case 2) it is not clear whether it is a property and can be treated as one. Thus, we shall keep calling it "property".

Anyway, as it was said, in principle, if this "property" really exists in objects, one might try to use it to differentiate them. However, I believe this would not work. Hawley (2009), followed by Muller (2015), claims that a counterexample to PII must consist of a qualitative arrangement. This certainly would exclude non-qualitative thisness. Firstly, because it is not considered a qualitative property; secondly, for the same reason that was used to dismiss bare particulars, viz., we cannot even make sense of them in a physicalist way, that is, it has no

clarifying predicates to bring up when trying to describe it. What I ask of the defendant of non-qualitative thisness is that either she displays its semantical content or, at least, presents some evidence for its existence through demonstrations.

In the case of the former disjunct, whether one tries to spot thisnesses in demonstrative terms, one would have to rely on relations of some sort to use a Dthat operator,[57] failing to be non-qualitative. The latter disjunct, on the other hand, is what Swimburne (1995) tries to do. However, to – allegedly – demonstrate the existence of thisnesses, he uses Black-like scenarios and a dispersal argument to show that PII is false, therefore, there must be thisnesses individuating spheres in the scenario. Well, whether, after that, one uses thisnesses as foundation to defend bare particulars as individuals in a Black-like scenario against PII, one would be incurring in circular reasoning. The denial of PII is required for the defence of non-qualitative thisness (see SWIMBURNE, 1995).

Therefore, I believe that thisness only makes sense if interpreted as the ultimate suchness, as something that is guaranteed that no other entity will display, thus, also will guarantee normal numerosity for *x* and *y*. Thus, I believe that thisnesses and bare particulars cannot help the opponent of PII.

2.4. F is for features

Let us now deal with the most problematic and discussed term on the formula of PII, namely, "F". What kind of thing does "F" stand for? This question has at least two answers and they both present different issues for the debate over PII. One deals with which things count as *F*s and which do not when invoking PII or constructing a counterexample to PII (Muller's steps 2b and 2c, respectively), a question more related to the epistemic side of the dispute; whilst the other deals with a more metaphysical question about the nature or the ontological structure of the objects admitted as referents for "x" and "y" in the last section (Muller's step 2a).

[57] The Dthat operator is a formal device introduced by Kaplan (XXXX) to model our use of demonstratives, such as "this", "that", "those", in debates within Philosophy of Language. The operator has the form "Dthat […]" where inside of the square brackets goes a description of (or a pointing gesture to) the *demonstrata*, a part that is absent in the discourse, whereas the "Dthat" part stands for rigidifying act, analogous to the act of pointing or nodding in the direction the reference.

Nevertheless, both answers have epistemic and ontological/metaphysical implications, as these matters are somehow inseparable. The trick here is just to focus on one aspect more than the other at each time. Let us unfold these answers starting with the latter, which has more straightforward answers.

The literature tends to treat the spheres as amalgamations of attributes of one of these three kinds: tropes, immanent universals, or transcendent universals. There may be nominalist approaches too, however I have not crossed them yet. Nonetheless, for the purposes of this research I believe that the arguments favouring the tropist position should also fit in a nominalist shoe. I am not certain about which one is the correct way to understand the nature of properties – for this is a much broader debate – thus, I will try to briefly address those three that are more prominent in the literature about the PII.

2.4.1. Tropes

If we interpret the attributes of the spheres as tropes, Black's experiment would not lead to his conclusions. Tropes are very complex entities, and their characterization may not be totally consensual in the literature; thus, I will not try to discuss all their characteristics here. However, one thing that seems to be clear about tropes, or even essential to tropes is that they cannot be identical to one another. Probably this is why debates concerning them are about *resemblance*, not identity or sameness. Two tropes of grey can be extremely similar, perhaps indiscernibly similar to the best possible analysis, but there must be something that differentiates them, otherwise they would be the same trope. For the tropist, what the realist would call two different instantiations of the same universal of grey, should be examples of two (different) tropes of grey.

One thing that could account for this difference between very resembling tropes are higher order properties, properties of properties. In this case, the properties of spatial locations of the tropes in question. One grey trope that looks exactly the same as the other trope, concerning every shade of greyness, should be considered a different trope than the first one, because of its spatial – and temporal, if you want – properties. Take, for example, Grey Trope^{1} which has shade^{1}, chroma^{1}, …, spatial location^{1}, whilst Grey Trope^{2} has shade^{1}, chroma^{1}, …, spatial location^{2}. They resemble each other because they have every property that is a higher order

property exactly similar but the spatial one. This might not be the only reason for the two resembling tropes not to be considered the same, this might not even be a reason for it at all; this is just a guess. Nevertheless, there must be one, otherwise trope theory would not be different from immanent realist theory about universals, which shall be analysed bellow. Moreover, if the constituents of Black's spheres were tropes, the conclusion that one wishes to draw from arguments from almost indiscernibles would not follow. Because the spheres present different properties, in other words, different tropes, which in turn, present different properties amongst themselves.

Notwithstanding, if this explanation would not satisfy the reader, we should appeal to definitions. For it is part of the definition of trope that each trope is unique and cannot be shared by other particulars. Therefore, there must be a different amalgamation of tropes in each sphere. We might not know which tropes are involved in the constitution of these particulars, but we can know with certainty that they are different somehow if they are tropes at all.

But what about a situation where we cannot discern between two very resembling spheres consisted of tropes? How do we explain such a situation? In situations like that what happens is that the spheres are just not being *discerned* by the eyes of the beholder. In such cases, they cannot be said to be *indiscernible*, because there must be some conceivable attribute that differentiates them, something that could be perceived from a more privileged point of view. This line of thought will be dealt with in more depth in chapter 4, but for now, we should bear in mind that there is a relevant difference between two undiscerned spheres and two indiscernible spheres.

O'Leary Hawthorne (1995, footnote 1) says that trope theory is less equipped than any other kind of bundle theory to deal with PII. However, I cannot think why someone would think something like that. Trope theory seems to be the one theory in which PII could never be put at risk at all, given that the very definition of "trope" implies that objects in Black-like scenarios are always intrinsically different.

2.4.2. Immanent universals (O'Leary Hawthorne's view)

An alternative is to interpret the spheres as a bundle of universals, more specifically immanent universals. Following the terminology proposed by Armstrong, O'Leary Hawthorne

(1995) proposes that if we assume that properties are immanent universals, i.e., actualized universals that are themselves spatiotemporal parts of the objects they constitute, instead of just being mysterious entities outside of space-time that are merely instantiated in objects (viz., a transcendental universal).

In this view, there is nothing contradictory in saying that a shade of greyness or some roundness is at two kilometres from itself. Whereas in a transcendental view, one would say that an instantiation of a shade of greyness or an instantiation of a form of roundness is instantiated two kilometres from another instantiation of it, in the immanent view, these instantiations are apart from themselves, that is, the roundness is two kilometres from itself as well the shade of grey. An easy way to make sense of this is to think that every occurrence of this shade of grey or of this form of roundness is a mereological part the universal they form. But the easy way is often the wrong way. It is an incorrect explanation – and probably an abuse of mereology jargon – to describe this view in these terms, because the occurrences of greyness or roundness are not part of something larger, different than them, they are the things themselves. The strangeness of this idea will be explored further in section 3.1.2.1., where I present the criticisms to O'Leary Hawthorne's defence of PII.

One last thing that must be said is that, according to O'Leary Hawthorne, at least some relations are also immanent universals. Armstrong, according to O'Leary Hawthorne, sees the world formed by things of three fundamental kinds, namely, immanent universals, 'thin' particulars and states of affairs. A bundle theorist of this kin would believe that in baseline reality there are only immanent universals concatenated in some relations forming states of affairs. From these relations of immanent universals, thin particulars supervene. For example, in base reality we should have the universals *G* (grey), *B* (roundness, ballshapedness) and *R* (compresence) forming the state of affairs *GRB* which would yield the thin particular *a* (a grey ball). The states of affairs will be more complex, the more immanent universals they comprise. In this framework, then, the scenario designed to serve as a counterexample in P2) in the dispersal argument can be seen as a state of affairs formed by a bundle of immanent universals such as roundness, greyness, weight, volume, etc., that allows us to capture particulars, namely, spheres. But given that immanent universals may be in different places, the roundness, the greyness, and the other universals that consist of a sphere could be compresent with each other in different parts of this state of affairs, would not this mean that the allegedly two particulars

that supervene from these universals are in fact only one particular, especially if we count *being at 2km from something* as the spatial property in this state of affairs? This is the point used by O'Leary Hawthorne for his identity defence which will be explained in section 3.1.2.

2.4.3. Traditional universals and nominalist alternatives

The literature says nothing about the nature of the spheres whereas particular instantiations of real universals nor about them being particulars from which we mentally abstract properties. Therefore, I conclude that whatever they are, there is no clear reason to prohibit or to accept that the properties they bear are things of one of these kinds. For reasons of lack of time to investigate it and space to write about it, I will follow the literature and simply leave this issue aside. The only thing that might be unacceptable is to interpret the spheres as universals themselves, for two reasons. First, universals are conceived as values of the *F* variable, not as value of *x*s or *y*s. Second, they would be the same universal or the universal would be locally individuated, and in each case, the universal character evaporates, at least as a traditional universal.

2.4.4. What counts and what does not count concerning discernibility?

The first answer to the question that started section 2.4. says that *F* is a variable that stands for the features displayed by the object (or objects) in *x* and *y* places in the formula of PII. These features are commonly divided into two kinds, namely, properties and relations, the first usually seen as having some kind of ontological priority over the latter. In this view, properties are attributes[58] that an object has independently of anything (see BABER, 2019) and are referred by monadic predicates, e.g., the sphere *is grey*, i.e., the sphere has *the attribute of greyness*; whereas relations are attributes that the possession by one object typically depends on other

[58] Yablo (1987) discerns properties and attributes, but we shall keep it simple and ignore other classifications of properties, as I have already done it in section 1.3.3. with the distinction between properties and ascriptions adopted by Cross (1995), however, this time it is done for reasons of space and time limits to this research. A difference between properties and attributes might very well be relevant for this debate. I believe this will not be source of problems though, since we shall not deal with dispositional properties nor transworld identities.

objects and are referred by polyadic predicates, e.g., the sphere *is 2km from* another sphere, i.e., the sphere has *a distance in the magnitude of 2 km with* another sphere, therefore, is not possessed independently of other objects.

I do not agree with this distinction, and whether it is possible to be drawn is a philosophical issue that would require a whole new dissertation (See MACBRIDE, 2016). As a justification for my position here, take as an example the notion of strict identity. First, as a reflexive relation it is not held by two objects, but only one, thus, it does not make sense to say it has no priority of possession between the objects that hold it. It is possessed by an object. Second, as we saw in the above sections, identity can be understood as a relation *or* as a property depending on how you describe it. In the same spirit, many other features can also be translated from one form into the other. For example, the property of *being the tallest man alive* can only be possessed by an object in relation to other shorter manly objects, therefore, it can be also understood as a polyadic relation of its possessor to every other man e.g., *being taller than* a *and* b *and* c *and* so on. However, it is not clear whether every property can be transformed into a relation. Take for example the property of *having such and such mass* or *having such and such internal structure*, the instantiations of both seem to be utterly independent of other objects – here, I am ignoring theistic considerations, of course.

On the other hand, any instantiated relation can be reduced to a property, e.g., the relation of *being the father of* possessed by Roberto in relation to me and my brother can be understood as the property of *being the father of Leonardo* or *being the father of Arthur* (my brother). They can be translated into impure properties. Muller (2015, p. 206) disagrees that all relations are reducible to properties, yet he believes that at least, binary relations, the relations that will concern us in Black-like scenarios for reasons of "there are only two objects" settings, should be considered in the realm of relevant *F*s dealt in the discussion of PII. Thus, we are on the same side, in this debate at least.

In any case, Black himself (1952) is very explicitly saying that whether his scenario is possible, "(…) *then every quality and relational characteristic of the one* [sphere] *would also be a property of the other* [sphere]" (p. 156). He is followed by Adams, who explicitly says that we should "*reserve the title 'Identity of Indiscernibles' for the doctrine that any two distinct individuals must differ in some suchness, either relational or nonrelational*" (1979, p. 11). Echoing Adams, Rodriguez-Pereyra (2006) agrees that an interesting version of PII that deals

with relational properties. Saunders (2006, p. 53), followed by Hawley (2006; 2009), Ladyman et al (2012), and others involved in the debates of special cases of PII also accept relations as valid *F*s. Therefore, it seems correct to consider relations as appropriate *F*s, even if one does not agree that relations and properties are translatable.

The relevant aspect common to almost every debater in the dispute is that they agree that the features in question must be qualitative, i.e., non-qualitative thisness is not an option (the exceptions are Swimburne and perhaps O'Leary Hawthorne). Nevertheless, they do not agree that these qualities can be of any kind whatsoever. At this point is where generalized trouble begins, where even champions for the same side turn on each other. As one can see from the examples above, the boundaries between the notions of property and relation are blurred, and I shall not try to draw a clear distinction between relations and properties here. Instead, I will just assume that relation is a kind of property, for I assume that all instantiated relations can be described as properties containing the objects with which they hold the relation, but not every property can be transformed into a relation (e.g., mass).[59] Nevertheless, the levels of relationality that properties display seem to engender a taxonomy of properties. Depending on which taxon a property lies under, it might be considered appropriate to be also under the scope of PII relevance or not. In the same spirit, it can be considered to be a valid property for a counterexample to PII or not (Muller's step 2a and step 2b).

There is no shortage of taxonomic divisions for the classification for properties. Some of them are irrelevant for the PII debate, some are of the utmost relevance and others are somewhat relevant but are largely ignored. As an example of the latter, we can name the taxa of primary and secondary properties. Shortly, we can define primary properties are the ones that the objects have independently of any observer, e.g., solidity, numerosity, extension, etc., [60] whereas secondary properties are observer dependent, e.g., colour, sounds, smells, etc. In Black's

[59] To be honest, I tend to believe that every property is relational and can be transformed into a relation. However, I have no space to discuss this issue here and I can only hope for future funding to discuss this idea somewhere else. Thus, for now, I will assume this position that I believe the reader will find at least reasonable where some properties are not translatable into relations.

[60] The list of primary properties is not stable. Properties such as shape, size, and motion can be turned in secondary properties whether one consider a relativistic approach of the space the object in question exists in. This is why these were not used as examples, but they are commonly cited as paradigmatic examples.

excerpt quoted in the introduction, he mentions that "*each* [sphere] *was made of chemically pure iron, had a diameter of one mile, that they had the same temperature, colour, and so on* (…)" (1952, p. 156). It is clear that he – like Kant, Leibniz and everyone else that discussed their examples – thought that the objects in the proposed counterexamples had secondary properties, i.e., properties that are observer dependent.

Nevertheless, I believe secondary properties such as colour, sounds, etc., are wrongfully attributed in counterexamples to PII, for they need other *relata* that would break the symmetry of the example to make sense. First, they need a perceiver (or a conceiver) to perceive (or conceive) these *qualia*, e.g., it makes no sense to say that the sphere looks grey or tastes like iron if there is no being gifted with those senses to instantiate these properties, to make sense of such statements. Second, even if we avoid *qualia* and understand these qualities in a purely physicalistic way, they would not be in the sphere too. Consider the properties of colour and temperature. How could these spheres present any of them in a physicalistic sense, if there is no sun (or other source of light and heat) to cast light upon them and permit a reflection of the light in the wavelength of a colour or to make its molecules to move and generate heat? Such universe would be in absolute zero and complete darkness. Unless we consider the possibility that they have the dispositions to exhibit these properties, i.e., the possibility that they would present the grey *quale* or would reflect light in such and such wavelength if a light source and a perceiver were in the scene. Dispositions do not require the existence of the other *relata* in the scene nor of the relations themselves. But to accept this would lead us to other taxa widely ignored in this debate, namely, the *dispositional properties* and *manifest properties*.

I will be short about this topic: given that we did not accept transworld identity as a relevant form of identity for this debate, I believe we must avoid talking about dispositional properties as much as possible too, for they invoke transworld attributions of features that are not relevant for the comparison of the objects in this discussion. If we do not accept transworld identity, I see no reason for accepting transworld (in)discernibility in the same formula either. Therefore, we shall deal only with manifest properties. This point can be summarized in a short maxim of Scottish wisdom: "if my auntie had balls, she'd be my uncle" – pardon my Gaelic.[61]

---

[61] Again, for the problems of transworld attribution of properties and identity, see YABLO, 1987.

Still, one might reply that the spheres could be suns generating their own heat and light. Still, the observer problem persists. Some retinae still would have to capture this light for its colour to be there. In the case of temperature though, if it is understood as the movement of particles rather than a feeling, it would be there. Fair enough, but it would not be a secondary property anymore, but a primary property of movement of the (particles of the) suns.

In conclusion to the topic of primary/secondary properties, it seems that although Black and others that followed him talked about secondary properties of the objects in their examples, it does not seem to make much sense to do so. A rigorous thought experiment that wishes to serve as a counterexample to PII should avoid secondary properties and deal only with primary properties, because to make sense of the former kind of properties, the postulation of entities that might break the symmetry within the scenario is required (this issue will be further explored in chapter 4). Thus, Muller's step 2 requirements should be only truly met by the proposed mathematical and Quantum Mechanics scenarios. However, given that the most of the debate is done with Black-like scenarios and those are much more intuitive for most readers, the most intuitive thing to do is to keep using Black-like scenarios as the paradigms with the following caveat for the reader: although we shall use secondary properties while describing the objects in the scenario, these properties should not really be considered relevant part of a truly relevant experiment serving in argument against PII. I will engage in explanations with secondary properties only for pedagogical reasons as examples or because my interlocutors have been using them already (when it is the case).[62] By the end of this chapter, I hope that the reader sees that all that matters boils down to spatial properties.

Back to the main track. The main taxonomic groups that will be most relevant for the discussion over PII concerning which properties are under the scope of it and which are permitted to fulfil Muller's step 2 criteria for a counterexample to PII are the *intrinsic properties*, the *extrinsic properties*, the *pure properties*, and the *impure properties*. As the reader can see, they form dichotomies. Let us now analyse what are those properties starting with the later dichotomy.

---

[62] In chapter 4, when I talk about the pictorial conceivability of the counterexample scenario, I will use secondary properties when describing it, but keep in mind that it is to make the exposition simpler to imagine, it is just to help make sense of a non-linguistic conception of the scene.

The pure/impure dichotomy concerns whether the property in question has another particular as part of it or not. A pure property has no particulars as parts of it, whereas an impure one has particulars as parts of it. As examples of pure properties of Roberto, we can list *is extended*, *has such and such mass*, *is white-haired*, and *is a father*, among others. Notice that some of them are non-relational properties, while others are relational, such as *is a father*. To be a father, one must have children, then, *to be a father* is a relational property, although the relata involved are not explicit in nor implied by the predicate that stands for it. On the other hand, *is the father of Leonardo* is an example of impure property, for I, another particular, am part of this relational property, and my name is in the predicate that stands for it. The property *is the father of the author of the current thesis* has no name in it, but is also an impure property, for it contains a description that works as a name singling out a particular in it.

In turn, the intrinsic/extrinsic dichotomy concerns whether the properties in question are possessed by the objects in virtue of relations they have with other things or not. *Roberto's mass* is not a property that he has in virtue of a relation with something else, thereby, it is an intrinsic property he has, whereas his *weight* is due to an interaction of *his mass* and the force of gravity, thereby, it is an extrinsic property. The reader might be tempted now to say that intrinsic properties are those that cannot be expressed by polyadic predicates, the proper properties (pun intended!); whilst extrinsic properties are simply the relational properties, the ones that can be expressed by monadic and polyadic predicates. However, this is not the case. *Having a larger gut than its head* and *being pointier than a perfect circle* are relational properties possessed by Roberto not in virtue of anything but himself (see WEATHERSON, 2018).[63] Therefore, there are some relational intrinsic properties, whereas there are no extrinsic non-relational properties. These kinds of properties will be relevant to establish which version of PII is being defended and which scenarios can be accepted as valid counterexamples to PII.

As mentioned in section 1.1., Rodriguez-Pereyra (2006), following Adams (1979), presents different versions of the PII, namely:

[63] Allegedly. I would argue that his size – consequentially, the size of his parts – and shape depend on external factors such as the point of view of the observer, the external pressure that permits that he does not explode in a vacuum and so on. However, I will follow the mainstream and grant that these properties are internal. Maybe, in the future, somewhere else, I will be able to argue not just that all properties are relational, but also extrinsic. For now, let us be conservative.

PII1: No two things share all their intrinsic properties.

PII2: No two things share all their pure properties.

PII3: No two things share all their properties.

Concerning PII3, if one can rely on properties of identity, such as *is identical to itself* or *is identical to b* (where "*b*" is the name of the object in question), then, PII is trivial. This is considered the weakest[64] version of PII, meaning that it imposes no restrictions to property attribution, thus, being impossible to reject, for it enables no counterexamples to PII at all. However, identity is the very thing the principle is supposed to establish (or at least introduce), thus, it must not depend on it on pain of being uninformative and uninteresting.

Nonetheless, a few words about the triviality and uninterestingness of PII3 should be said here. I must disagree with them for two reasons. First, because in Black's dialogue, character B says: "(…) *Then* every quality and relational characteristic *of the one would also be a property of the other*" (*idem*, emphasis added). He is making a very fragile claim when he postulates that the spheres would be equal in *every* characteristic, even the relational ones (Adams agrees that Black makes no distinction of properties; see ADAMS, 1979, footnote 11). Also, the discussions among people that are not that deep in the debate seems to be about every property, probably because they tend to base their opinions in Black's famous thought experiment. Thus, a very weak version of the PII still seems to render some interesting discussions. Second, the principle should be trivial! As it was said in the introduction, I think that the principle is a reasonable way to ground (or at least, an introductory way) the principle of identity in our experiences, in a naturalized way. Therefore, I believe it is in its own right to be trivial.

Notwithstanding, to say what trivial is is not trivial at all (pun intended again!). If it means that something is always true, it should really be trivial, given that it is intended as necessary. If it means that it is something obviously true like paradigmatic analytic truths, e.g., "oncologist is the medic specialized in cancer" or "Cassius Clay is Muhammad Ali", then, the bottom of

[64] The terms "weak" and "strong" here stand for the how many conditions the interpretation of *F* imposes to the things the principle applies to, not how hard is to find a counterexample to the proposition it expresses. If the difficulty to find counterexamples was the metric, the weakest proposition would be PII1 and the strongest PII3.

this well is further down. Whenever one talks of something being obvious, one must ask obvious to whom and in which conditions, for the debate over analyticity has shown in recent years that this is not that simple. The epistemic component of truisms like the above mentioned are complex and are not always uninformative.[65] Thus, the suppose triviality of PII3 could – and should – be contested.[66] Notwithstanding, let us play along, leave this version of the principle behind, and keep the debate about its interestingness and triviality for future endeavours. Let us interpret, then, that this wide range of features inserted in Black's scenario as relevant for PII by writing "every *quality and relational characteristic*" as a sloppy misplacing in Black's writing.

On the other hand, whether one relies solely on intrinsic properties, as PII1 does, there would be an abundance of counterexamples, given the very restrict number of properties allowed to discern in the counterexample scenario. Fortunately, we have already established that a relevant version of PII (and a counterexample to it) must involve relational features. With this in mind, we can all agree that an interesting version of PII must be in between PII1 and PII3.

PII2, which is proposed by Strawson (1959, p. 120), is a version that is in between a trivial one and an excessively restrictive one. It excludes impure properties from the pool of relevant properties, given that they contain names (or definite descriptions) in its predicates, for they could turn such properties into equivalents to identity properties, i.e., properties that are equivalent to or imply identity. This would make PII trivial.

---

[65] For an advanced discussion on the topic, see G. RUSSELL, 2008. For cases concerning the same proper names or constants flanking the identifying verb, see GLEZAKOS, 2009.

[66] Just to be clear, I am not saying that I am against considering PII3 as trivial and uninteresting or saying that this is the appropriate form that the principle must have. I am simply saying that this might not be that simple to settle given the developments on debates over analyticity, necessity and other kinds of identity. Even identity statements like "$a = a$" can be made false (e.g., Glezako's "Aristotle = Aristotle" case) and statements of the form "$a = b$" can be made trivial, depending on the previous knowledge of the subjects involved (e.g., G. Russell's "Clay = Ali" case. "Cassius Clay is Mohamed Ali" may be trivial for the Iman that baptised Mr. Clay into Islam, but certainly is not trivial for some young teenager that never heard of Mohamad Ali before). RP hints, that a trivial version is something that when denied turns into a contradiction (2006, p. 207), however, this cannot be the case, otherwise every definition would be trivial, and as we just mentioned, this is not the case (and he indirectly agrees with that when he talks about essences, see pp. 210-1). Thus, a criticism that can be made to RP's whole point is that he is not clear about what he understands as *being trivial*. Again, I believe that these points should be enough reason for not considering the PII3 as uninteresting so fast.

However, Rodriguez-Pereyra (RP henceforth) claims that this might be too restrictive. After all, some pure properties such as *being a father*, imply relations with specific relatum sometimes, in other words, it amounts to relations like *being the father of* someone. Also, not every impure property imply identity, for example, *loves Leonardo* is shared by many and implies identity to no one specifically. Thus, RP believes that there is a non-trivial version that allows for some impure properties and argues for another criterium to distinguish between permitted and forbidden properties in de debate for PII (2006, p. 206). He proposes that PII should be understood as:

PII2.5: No two things share all their non-trivializing properties,

then, we just need to discover which properties are trivializing and which are not. This is not an easy question to settle and there seems to be much disagreement among philosophers about that, given the fact that at least three mutually exclusive versions of PII are considered to be correct. Yet, there are some features that everybody – apart from the ones defending PII3 – agrees with, namely, properties of identity or that are tantamount to properties of identity, e.g., *being identical to a*, *=a*, *is an element of the singleton {a}*, *being green and identical to a*, and so one.

RP asks us to consider other kinds of properties that, in the same spirit, would also trivialize PII, namely, properties of difference. Those are any properties that are complements of identity properties, e.g., *being numerically distinct from a*. Again, consider that *a* and *b* share all their properties. It is easy to see that *a* lacks such property like the one in this example. However, if *b* also lacks such property, it is easy to see that "*a = b*". In this case, he claims, the triviality is established because the lack of a property of difference. To lack the property of *being numerically distinct from a* amounts to possess the property of *being identical to a*. More specifically, the properties expressed by the negations of predicated of difference must be out of the scope of PII, for they are equivalent to a predicate of identity. Therefore, properties of difference are not welcomed in the debate over PII.

Thereby, according to RP, there are cases of generally non-trivializing properties that, when instantiated together with properties of difference or properties of identity, entail trivializing properties such as *being identical to a or being green* together with *being identical*

to a or not being green; and *being numerically distinct from a and not being green* together with *being numerically distinct from a and being green*. These conjunctive and disjunctive properties taken separately are not trivializing, however, when instantiated together with some others, they may amount to properties of identity or complements of it. Thus, these properties should count as trivializing properties too. However, at this point, RP notices that there is a very serious problem. Properties like *being green* would also count as trivializing properties, then. Because they are equivalent to the property of *being (identical to a or not being green) and being green*, which amounts to a property of identity too. It seems that what must be avoided are properties that lead to properties of identity. According to the endnote 5 of RP's work (*op. cit.*, p. 220), these are equivalent to properties that have abstractions with the λ-operator that lead to something like (λx)(x = a), which is an identity property.

Notwithstanding, RP raises a very important objection to this view. He considers that if one is an essentialist, i.e., if one believes that everything has pure individual essences, this is, a cluster of only pure properties, one is committed to properties of identity of some kind. For individual essences are nothing more than properties of identity or properties that entail properties of identity. He presents the example of Plato having as its essence *being the greatest philosopher*,[67] which amounts to the property of *being identical to Plato*. But the property of *being the greatest philosopher*, according to him, should not trivialize PII. He argues that, for Leibniz, everything presents a qualitative essence (*op. cit.*, p. 210) and his version of PII certainly was not a trivial one. Therefore, the triviality of such version of PII might not reside in the properties of identity. Thus, we must agree, because otherwise even PII2 would be a trivial formulation.

RP, then, claims that "*if what one proves is that numerically different things must have different pure individual essences, then one has established that every numerical difference goes accompanied by a qualitative difference – and this is no triviality*" (*op. cit.* p. 210), which looks somewhat to the Principle of Dissimilarity of Diverse which, as we acknowledged, is equivalent to PII. It seems, then, that the trivialization aspect resides in the fact that numerical

[67] See RODRIGUEZ-PEREYRA, 2006, p. 210. Also, in endnote 13, p. 211, he stresses that his point does not need to invoke pure individual essence, the impure ones would do the job the same way.

difference is not accompanied with some other qualitative difference. He reminds us, that PII is the thesis that is supposed to show that numerical identity is entailed by qualitative identity (*op. cit.* p. 214). The problem with the arguments presented above showing that "*a = b*" through properties of identity or their complements is that they lack the spirit of PII, i.e., to show that it is impossible to have qualitative identity without numerical identity. They only say that two qualitatively identical things that are numerically identical (that is the only thing the properties expressed amount to) are numerical identical, which is trivial.

He is saying, then, that the properties of identity are properties of numerical identity and differing concerning them is differing in numerical identity and nothing more. Additionally, he claims that even when differing regarding such property entails a qualitative difference, this does not establish automatically that there is a qualitative difference, the qualitative distinguishing fact must be invoked (*op.cit.* p. 215). This also applies to their complements and properties of difference, since lacking a property of difference is the same as having a property of identity and vice-versa. Thus, differing regarding properties of difference is differing regarding properties of identity, which, in turn, is differing numerically.

But not only properties of identity and properties of difference are trivializing properties. RP claims that properties that *may* consist in differing solely numerically, when differing concerning them, should be considered trivializing too. For example, conjunctive and disjunctive properties such as *being identical to a and being green* or *being numerically distinct from a or not being square*. When differing concerning these properties, the individuals *might* be differing only numerically. For they may disagree whilst being green and not green or concerning the property of identity. Thus, almost as a safety clause, we must add this kind of property among the trivializing ones. On the other hand, properties such as *being green, being square, being the greatest philosopher*, should not be considered trivializing properties on their own – without properties of identity or their complements. For differing concerning them clearly consists in differing in more than only numerically, it consists in differing in colour, in shape and in skill with words, respectively. Additionally, RP interestingly claims, that for the same reason *being the mother of a, hating b* and *being in the same place as d* are not trivializing properties – keep an eye in the last one. When differing with respect to them, individuals differ with respect to the motherhood of a, the feelings towards b and spatial locations.

Thus, according to RP, the definition of trivializing property should be settled as:

"*F is a trivializing property* $=_{def.}$ *Differing with respect to F is or may be differing numerically*" (*op. cit.*, p. 219)

However, this cannot be the case, otherwise any property such as *being (all over) green (at time T1)* or *being round*, would be trivializing properties, for differing with respect to any of them would imply that the objects in question have different coloured surfaces or different shapes, which in turn implies different bodies, which in turn imply distinction, or in RP's preferred jargon numerical difference. This must be a formulation error. What he must be trying to say and what is in accordance with all his argumentation is:

F is a trivializing property $=_{def.}$ Differing with respect to F is or may be *only* differing numerically *and nothing else.*

In this formulation one can see the trivializing informational aspect of the property in question. Whenever the display of a property tells us nothing more than that the objects which differ about it are distinct, this display is trivial, for it says nothing that we already knew when the syntactical subjects of the proposition where presented. With this definition, RP presents an intensional definition for trivialising properties, allowing us to settle PII2.5. as the weakest version that is still worthy of attention. Just to make clear, let me spell it out in a more explanatory manner:

PII2.5: For any two objects and for any property but the ones that when differing with respect to them means differing numerically only or may be so, the sharing of those properties entails the numerical identity of the objects.

The reformed version of RP's trivializing criterium for deciding which properties should be within the scope of PII seems appropriate and it seems uncontroversial that it generates an interesting version of PII. What is controversial is whether it is the correct definition of trivializing property and which properties fall under this category. For example, an issue that requires further analysis concerning trivializing properties is addressed by RP in note 17. He

deems interesting the case of the property *having all parts in common with a*. He considers this property as a trivializing property, because it is possible that *b* shares all the proper parts with *a* but differs from *a* in having a different improper part, namely, *b* and *a* themselves; therefore, differing numerically. However, he remembers, there are some that do not believe that two things can share all their proper parts. In that case, differing with regard to *having all their proper parts in common with a* is not a trivializing property, for it means that they differ with regard to something else than numerical identity. Nonetheless, this is a substantive metaphysical issue that is further explored in the debate over PII, specially addressed in the Della Rocca-Jeshion debate (see section 3.2.1.ff.),[68] but needs more attention.

Thus, the issues of what is a trivializing property, and which are the trivializing properties are not clearly settled among PII debaters yet. To my knowledge, this is not widely discussed in the literature and is not widely accepted either (see MULLER, 2015, p. 211 and footnote 13). Nevertheless, there are some properties that are widely accepted as trivializing. For example, properties of identity or that are equivalent to them such as *being identical to x*, *=x*, *is element of the singleton {x}*, *is x*, and others of this sort which would make circular any attempt to establish identity (or difference), being *x* a name or a descriptive name. To be more precise, the literature generally agrees (RP being an exception) that names and properties using names should be used to affirm or to question PII, for they can be used in a way that amounts to identity properties when one affirms PII; and cannot be successfully established in the counterexamples against it, because there is no way of performing the act of baptism without ruining the agreed initial settings of the thought experiment, i.e., it would break symmetry. So, although there might be non-trivializing properties containing names, it is best, for the aims of this thesis, that we do not accept them in the set of properties discussing PII. This is certainly worst news for defendants of PII from a discerning strain than for the opposers, for there are less non-trivializing properties available to discern the spheres in a Black-like scenario. Notwithstanding, PII defendants of the discerning strain are happy enough to agree with leaving these properties apart (see DELLA ROCCA, 2005, pp. 481-2; MULLER, 2015, pp. 211-2 and

[68] For a good introduction to the subject see CASATI and VARZI, 1999, chapter 3. For an interesting discussion about how the issue affects identity and distinctness see also chapter 6 of the same book.

227). Perhaps because this should not be a problem for good defences. Therefore, I shall follow them and not accept properties using names nor equivalent to identity properties as relevant to defend PII, despite the fact that RP's criterium seems more appropriate.

2.4.4.1. *Another taxonomy*

Nevertheless, there is another taxonomy of properties available for us to discuss the issue of PII proposed prototypically by Quine (1979), revived by Saunders (2006), perfected by Ladyman et al. (2012), and adopted by Muller (2015) and others more worried about the discussion in the Quantum Mechanics context. This taxonomy was already partially discussed in latter sections where I displayed a hierarchy of discernibility implying distinction (i.e., numerical discernibility) and when I introduced relationals (see p. XX above).[69] It is a taxonomy based in the discernibility levels of features instead of the classic metaphysical taxa. It divides properties between absolutely discerning and relationally discerning. The latter can be subdivided between relatively discerning and weakly discerning. This division seems to be the most appropriate to discuss discerning defences against Black-like counterexamples.

An absolutely discerning property is any property expressed by a monadic predicate that would make objects differ numerically, i.e., be distinct, whenever one has it, and others lacks it. It does not have to be an intrinsic property. Extrinsic properties are welcome to discern two objects absolutely, it only must be expressed by a monadic predicate though. For example, *being outside the Parthenon*, or *being made of bronze*, both differentiates the statue of Athena from the statue of the Artemision God in different ways; or *being 2 meters from the red shack* which is displayed by one but not the other of two snowflakes that present exactly alike internal structures. Now, Muller (2015, pp. 207ff.) argues that predicates that single out absolutely discerning properties can be translated into logically equivalent dyadic predicates that express

[69] These distinctions are presented in (MULLER, 2015) not focusing on the features, but on the objects. Whenever he speaks of absolutely discernible objects, I interpret them whereas displaying an absolutely discerning property/feature. The analogous goes for relationally discernible objects, relatively discernible objects and weakly discernible objects.

relationally discerning properties. For example, *being made of bronze*, Bx, can be also expressed as the relation R_B: $Bx \wedge (By \vee \neg By)$. Thus,

$$\vdash \forall x\ (Bx \leftrightarrow \forall y/\ R_B(x,y)).$$

Then, he claims (*op. cit*. p. 208) that absolute discernibility between objects imply relational discernibility. Thus, it seems possible and simpler to put every feature presented in this debate in terms of relations. Here we have an additional point in favour PII versions weaker than PII1.

This brings us to relational discernibility which, as said, can be of two kinds, namely relative discernibility and weak discernibility. A relatively discerning relation is irreflexive and asymmetric, whereas a weakly discerning relation is irreflexive and symmetric. A relatively discerning feature is a relation, thus, expressed by a polyadic predicate, that is held by an object to others, whilst other objects do not hold it to the first permitting them to be discerned in this way. In other words, any asymmetric relation, such as, *being larger than*, *being to the south of*, *being asymmetrically opposed to* and *being after* (provided that time has only one direction), are examples of relatively discerning features. Finally, there are weakly discerning features which are relations, thus, also expressed by polyadic predicates, mutually held by objects to each other but not to themselves, in other words, they are symmetric and irreflexive relations. As examples we can list *being symmetrically opposed to*, *being 2km from*, *being a pair with*, *facing off*, *spins in the opposite direction than*, or any with similar structure. For matters of simplicity let us restrict ourselves to binary relations expressed by dyadic predicates.

Muller, following Ladyman, Linnebo and Pettigrew (2012), claims that every case of absolute discernibility implies relative discernibility, which in turn imply weak discernibility, which, finally, imply distinction (see p. 56 of this thesis). Whether distinct objects fail to present even weakly discerning properties, the bare minimum to discern objects, we should call them indiscernibles. PII opponents believe Black-like scenarios are examples of situations where we find indiscernibles within, i.e., distinct objects that fail to present discernibility in any level. One of the most effective lines of defence, then, would be to show that at least weakly discerning features are present in the scenario. This is precisely what defendants of the discerning stripe will do.

Moreover, connecting RP's taxonomy with this one, the lengthy research provided by Ladyman et al. (2012, especially section 6) demonstrates that weakly discernible relations (weakly discerning properties) are the most discernible non-trivializing ones. Thus, we can assume that at the end of the day the features that will decide the quarrel for one side or the other are the weakly discerning ones. This might not be clear at first glance, because relations are commonly left behind in most descriptions of objects and whenever we conceive them, perhaps because they are not seen as part of the objects themselves since they are not intrinsic to them, but of the situations they are in, in other words, the scenarios. However, could the objects be what they are, *as they are*, in the absence of the relations they held? I do not think so. We can, then, update our understanding of PII2.5 from "No two things share all their non-trivializing properties" to:

PII2.5: No two things share all their weakly discerning features.

This being said, we must now specify which weakly discernible features are useful for defendants to look for in Black-like scenarios. They are generally spatial features or features that are somehow intimately related to spatial features, although in the special cases of mathematics and QM, where spatial features are allegedly not present, other weakly discerning features are available to do the discerning job. In QM cases, defendants use particles' spins, whilst in the mathematical cases of imaginary numbers and graphs, complementary features (e.g., *summed with x equals to zero*) and graph edges can be used to discern respectively. Whether these cases are really spaceless is not completely clear to me – in the sections 3.3.2. and 5.2., I will discuss the problem in the very singular case of unnamed graphs. However, I will follow the literature and consider them cases with no spatial features. Notwithstanding, analogous arguments can be made for these cases, provided the above-mentioned discerning features are present.

Being spatial features the paradigmatic examples of weakly discerning properties, let us take a closer look at what more precisely are these spatial features and ponder whether any spatial features can be used as a discerning feature or whether there are some spatial properties that are prohibited.

2.4.4.2. spatial features

As it was mentioned in the beginning of section 2.4.4., an analysis of a Black-like scenario is better off leaving behind secondary properties and sticking to primary properties. The latter can be viewed as fundamentally spatial (spatiotemporal, if you prefer) features[70] or features intimately related to spatial ones, e.g., motion, figure, extension, etc. All of them, apart from mass and spins (of particles) perhaps, are spatial in nature or depend on spatial features to exist (see CASATI and VARZI, 1999). Thus, it seems that the whole conundrum about PII can be boiled down to the question of whether two objects can synchronically share each and every spatial feature they present. But what are spatial features more precisely?

Spatial features can be conceived as absolutely discerning properties, viz., expressed by monadic predicates, e.g., *being outside the Parthenon* or *occupy region xyz*, but can also be translated into relations (see MULLER, 2015, p. 207); or can be conceived as relationally discerning features expressed by polyadic predicates (though we shall limit ourselves to dyadic predicates), e.g., *being 2km from* or *is diametrically opposed to*. These two ways of expressing spatial properties reflect two conceptions of space itself. According to Dainton (2014), these two conceptions of space are named the substantivalist conception and the relationist conception. The former proposes that the complete inventory of the universe contains the material objects *and* the entities of space and time which are ontologically prior to these objects, whereas the latter proposes that the universe is composed only of these material objects and their relations which include spatial relations and temporal relations – which will be avoided for matters of simplicity.

More precisely, the substantivalist view proposes that space is the largest thing that exists, and all other things exist within it. Space is, then, a pervasive medium that comprises every material thing. Although we cannot see or feel space, it is there and will always be, otherwise we would not. The relationist perspective, on the other hand, is ontologically more economic, for it does not require the existence of such a pervasive medium. This does not mean

[70] Locke counts solidity as a primary property/feature. However, I believe he means penetrability instead of a quality related to the feeling of things or to their grip. In this case, it seems to be a spatial mereological feature rather than a perception dependent attribute.

that a relationist advocates that there are no spatial locations or that statements about them are false, somewhat like an error thesis would, rather the contrary, i.e., such statements are true, but they are not about a medium, an independent object, they are about relations that supervene (or even emerge) from the collective of objects that exist. In other words, the substantivalist believes that space is the canvas in which the picture of reality is painted, whereas for the relationist space is the relations happening between objects.[71]

Whether we embrace substantivalism, we would have to change a little bit the description of Black's scenario, for it would be inaccurate from the beginning. The scenario describes that there are only two exactly alike spheres and nothing else. However, for the substantivalist there also must be space. This means that we would have to count space locations as objects – or at least parts of a larger object, viz., space itself – which might break the symmetry of the scenario, given that the locations in it might need be identified for a correct description of the scenario. But even if symmetry is maintained another issue arises. The spatial locations should be formed by spatial points which are also physical objects independent of other material objects such as the spheres that exist in them. Space should, then, be thought as similar to a cartesian plane where the nodes are pre-established, while the lines, curves, figures, etc. that we postulate exist *on* them.[72] In this case, we might as well drop the spheres and use space points instead. These space points share all intrinsic properties and seem to be indiscernible from each other (if we dispense the spheres from the scenario).

In this case, the defendant could argue that the spacetime points (and locations formed by them) could be differentiated by coordinates, e.g., point (2, 2, 2) is a different point than (2, 2, -2), such as in a cartesian coordinate system. This would also allow one to present different attributes to the spheres in Black's scenario, namely, the spatial locations of the spheres, e.g., *occupies the location xyz*. This is discussed further in sections 3.2.1.ff. This seems like an interesting solution, for it would permit us to discern absolutely the objects in question. However, Muller warns us, this would be cheating. Using such coordinate system would be the

[71] These conceptions will be identified, respectively, with the views named Primitiveness About Locations (PAL) and Primitiveness About Objects (PAO), in sections 3.2.1.ff., where they will play key roles in the debate about the challenge posed to the opponent of PII proposed by Della Rocca.

[72] It is not clear to me whether we should say that things exist *on* them, *in* them or *with* (as an attribute) them, but I hope the meaning is clear.

same as using names in features (see MULLER, 2015, p. 212). Something we had agreed not counting as discerning, albeit others, such as Rodriguez-Pereyra, would not see a problem in using these "names" within the properties presented. Muller solves this problem using other spatial features (see *ibid*. p. 210ff. specifically to spacetime points see pp. 225-6).

Nevertheless, I would argue that there is a way of using coordinates to differentiate between spatial locations that does not determinate, thus they would not be named. Take the cases of spheres again. One could say that a sphere has the property of *being located at x, y, z*, while the other has the property of *being located at -x, -y, -z*., where *x, y*, and *z* are never disclosed. Although one may not know which are the values of the variables in the coordinates naming the positions, one can be sure that the spheres are in different places and occupy different spatial locations, thus, having different spatial properties. These spatial properties would account for their difference in number as well as prohibit any attempt to identify one sphere with the other – unless one is an immanent realist like O'Leary-Hawthone (see sections 3.1.2. ff.) In other words, one would know that the spheres present different spatial features, thus, being discernible; despite not knowing exactly which are these features.

However, it is not clear though that this solution would work for space points, given that their only features are these spatial features which are also their names. In other words, they do not *occupy location xyz*, they *are xyz*. Therefore, this might be a trivializing property for them. One solution for that might be saying that these spatial features are not features of the points but of the scenario itself, or perhaps of the entity space itself, which has mereological parts.[73] In this case, the spatial difference is an intrinsic property of the scenario. The spatial properties of parts of the scenario cannot be the same, unless they occupy the same locations, which would intuitively count as the same parts.

In any case, Muller suggests a different approach that avoids these problems, namely, the appeal to relational spatial features. Such features can be spotted in a substantivalist view, but they are the essential spatial features of a relational view. Additionally, some complex

[73] Casati and Varzi differentiate spatial entities (e.g., material objects) from spatial items (e.g., points, lines and locations) (1999, p. 2). For matters of simplicity, I shall treat everything as an object in the sense explained in section 2.3. This should yield no loss to the points made, given that the relevant characteristics of spatial features are present in any of these theories. This might seem relevant for Jeshion's point in (2006), but Della Rocca's (2008) seems to settle the matter in favour of equal treatment.

oddities about space points, such as space possessing ontological priority to objects, viz., locations having objects as their attributes, instead of objects having spatial properties, do not rise. In a relational perspective, we can adapt the explanation above from locations having spheres in them to spheres holding relations between themselves unifying our understanding of features as something objects present be it a colour, a taste, a shape, or a location. One last reason to adopt this view is that it seems to be the most accepted view nowadays given its natural proximity with relativist view, i.e., the view that spatial properties are relative to points of view, the predominant view in contemporary physics; whereas substantivalism being more akin to an absolutist view of space, i.e., the view that spatial properties are fixed and immutable, which was generally abandoned since Einstein's theories were experimentally corroborated.

In this view spatial features will always present the form of a relation. For example, spatial location of an object a, which is between b and c can be expressed as *L*(a, b, c), where *L* can be read as *is between*. Another example is the spatial feature adjacency. It is said to be held by a in relation to b and can be expressed as *A*(a, b), or simply a *is adjacent to* b. Other relevant spatial feature that will be considered in the following chapters is distance. *D*(a, b) can be read as *is at 2km from*, and is held by a to b. Notice that adjacency and distance are weakly discerning features. They are symmetrical and irreflexive relations.

Anyway, as Casati and Varzi explain: "*thinking about space is, first and foremost thinking about spatial things* (…). *Spatial thinking, whether actual or hypothetical, is typically thinking about spatial entities of some sort*" (1999, p. 1). Entities – that I shall treat as objects in the sense presented in section 2.3. – that hold relations to other entities. This is what matters. Whether a Black-like scenario is conceived relationally, the spatial features of the objects (spheres, locations, or spacetime points) would depend on the relations that each object has with other objects, in this case with the only other object within the scenario. Whether it is conceived substantivally, these relations still exist in addition to other metaphysically prior spatial features expressed by monadic predicates (e.g., *is located at xyz*), thus the point of the discerning defender can still be made (see MULLER, 2015, p. 220).

One last thing that deserves attention concerning spatial features is that if RP's characterization of trivializing properties without our addition were correct, perhaps spatial properties would be trivializing properties, for they might be seen as properties of identity or of difference. We can identify an object through its positioning in a map of reality as we can

differentiate it from other objects in the same manner. However, its spatial features do not tell us only that. They also tell us something about other properties, for example, the degree of proximity of an object to a light source will tell us something about the colouring we will attribute to it. In Black-like scenarios it might not say much more about the objects actually, but it would certainly say more about its dispositions – which should not play any role in this debate, as previously stated. Nevertheless, spatial features are qualitative features of the objects and in most scenarios, different spatial features would imply other different features and certainly say more about the objects than simply what is what (or what is not what). Therefore, they can be understood as non-trivializing features. Furthermore, remember that RP himself assumed that at least some spatial properties are not trivializing, namely, properties such as *being in the same place as d*. This property should be avoided because of the use of names in it, however, if there are other ways to describe it, using just permitted relations, then, no problem should arise.

o o o

Summing up section 2.4.4., in completion to steps 2b and 2c in Muller's list of steps for debating over PII (the steps about what is permitted and what is not permitted to count as discernible in counterexamples) we must say that dispositional features, features that say nothing more about the scenario where the objects are conceived than the identity or distinctness of the objects in it (which we loosely call identity and difference properties), properties using names or descriptive names and any other property that would break the symmetry of the scenario are not allowed to be used. Any other kind of feature is permitted. However, I argued that those that are widely known as secondary properties are useless. Additionally, I argued that the really relevant properties are primary properties, which are spatial properties or are intimately linked to those. The most relevant among spatial properties being those conceived relationally. Although I have not completely shut the door for non-relational spatial features, i.e., I argued in favour of the use of coordinates and accepted Rodriguez-Pereyra's arguments for the non-triviality of extrinsic spatial properties, I choose to adopt a more conservative approach and exclude those, for they could be understood as name using features.

3. Defences

In this chapter I will present some of the most influential defences of PII divided in families of defences, respectively, the identity family, the discerning family and the summing family, followed by some objections to them. I shall present my views about these defences throughout the expositions but as an advancement let me say that identity defences should be avoided, discerning defences should be used, and summing defences should be kept in our back pockets in case something goes wrong with discerning defences.

3.1. Identity Defences

Identity defences are those that claim that "*there are not two (or more) objects, but only one object of the same kind of the alleged two (or more) objects belong*" (MULLER, 2015, p. 205) in Black-like scenarios in the dispersal argument. Ian Hacking (1975) and John O'Leary-Hawthorne (1974) are champions of defences in this family. In the following sections, I shall critically discuss these defences. Hacking's defence received great amount of attention and criticism throughout the years. In the concerned section, I shall present a mishmash of objections contained mostly in Adams (1979) and French (1995), Hawley (2009) and Muller (2015). Whereas O'Leary-Hawthorne's defence was not that popular, probably because it was drenched in what Hawley (2009) called *dank metaphysics*, it received replies from her and other metaphysicians such as Zimmerman (1997) and Rodriguez-Pereyra (2004).

3.1.1. Hacking the experiment: Ian Hacking's identity defence

In his (1975), Ian Hacking proposes a defence of PII that later came to be categorized as an identity defence (see HAWLEY, 2009; MULLER, 2015). He is perhaps the first of this kind if we choose not to put Russell's or Ayer's defences in this family. His main thesis in this paper is that any attempt to prove PII false through examples using space and time as relevant properties to the enterprise should be considered inconclusive, for there is no scenario that *must* be conceived as violating PII, although they *may* be done in such way (1975, p. 249, reaffirmed in 255-6). He does not use the conceivability vocabulary, but what he is saying, in other words,

is that although one can describe a scenario where PII seems to fail, it is not the only appropriate way to describe *that* scenario. There are alternatives descriptions of the same scenario where PII is respected. Thus, at the end, it is a matter of adequacy between the description and the phenomenon it describes.

Hacking's line of thought is based on the idea of underdetermination of theories by facts. Roughly speaking, this is the idea that there is more than one appropriate description for a set of phenomena, whereas these descriptions are also incompatible with one another. He starts his point remembering a famous debate about the nature of space between Newton and Mach. Hacking reconstructs Newton's spinning bucket thought experiment for absolute space and Mach's reply to it. As Hacking presents it, Newton asks his reader to imagine a universe with nothing but a bucket half-filled with water spinning continuously. He claims that the water would accumulate and rise in the limits of the bucket while the contrary would happen in the centre. This could not be explained by a relativist approach, given that there would not be another object to reference the spinning bucket or the bucket at rest. Mach replies, centuries later, saying that to claim that the water in the bucket within such abstraction would behave as the water in a bucket in the actual world is simply arbitrary. Nothing guarantees that such possible world obeys the same laws of physics that govern our world. A completely different set of laws could be proposed to explain the movement of the water in the bucket without appealing to absolute rest (thereby, absolute space). To claim that such possible scenarios would follow this or that set of laws, Hacking claims additionally, is a matter of choice. Thus, arguments for that are needed.

The same goes for the adoption of one or another notion of space as the metaphysical base for such scenario, i.e., to describe a phenomenon within a Euclidean space or a non-Euclidean one is also a matter of choice. Hacking claims that it is possible to describe the same scenario proposed by the PII opponent within a Euclidean framework as well as within a non-Euclidean one. As evidence for that he tells us the story of a debate between K(antian) and L(eibnizian).[74] K champions Kant's dispersal argument of the indiscernible drops of water, i.e.,

[74] In the paper, Hacking sometimes calls them Kant and Leibniz, sometimes just K and L, but it seems fairer to the modern philosophers to interpret these characters as different from them, for they are based on interpretations Hacking makes of them, which may be loaded with interpretations from others.

a Black-like scenario. On the other hand, L claims that apart from the question begging proposition "there are two drops", his description of the same phenomena in the scenario containing only one drop of water is as appropriate a list of true propositions as its rival's description. If L's claim turns out to be true, this means that every property attributed to supposed object *y* is also attributed to supposed object *x*, as the description of the scenario proposes, but *x* and *y* do not differ in number, as expected. This should be enough to hold back the truth evaluation of P2) in the dispersal argument. In other words, this means that P2) can be made true under some interpretation and false under another equally appropriated interpretation. Thus, concluding that the dispersal argument against PII is, in fact, inconclusive.

In the story, K insists that "every drop in the scenario *is ten diameters away from a drop*"[75] is a true proposition in his description, which in turn could not be true in L's one drop description. To this claim, L replies that it can, if we consider that the space in this scenario is non-Euclidean. If the space where the droplets are conceived in is considered in a cylindrical shape, this would permit that every drop in it is ten diameters away from a drop. The catch is that it would be ten diameters from itself in this case. If one leaves from one side of the surface of a droplet and floats ten diameters away from it, in a cylindrical space as conceived by Hacking, one will arrive at the other side of the surface of the same droplet. This is a perfectly adequate description of a possible world containing the scenario proposed by PII's opponent. Thus, to favour K's description of the scenario would require showing that there are two droplets, something Hacking believes is to beg the question, or that space is absolute (and Euclidean) (*ibidem*, pp. 251-2).

The dialogue goes on with K's appeals to other properties such as *reflects a drop in its surface* – which can be described as one droplet reflecting its other side – and relations that the droplets might have previously to when and where they are now (i.e., the void) – which would present discernible properties –, but to every attempt, L replies with a charge of question begging, because the descriptions of such properties imply the number of drops, or with an equivalent description with an one droplet scenario (*ibid.*, pp. 252-3). It seems that, as Hacking

---

[75] Notice that he cannot say "… from *another* drop" for this would settle the matter against his interpretation, because it would imply that there are two drops.

notices, the appeal to higher levels of complexity of the scenario, such as the addition of movement and causal relations between the droplets, could have helped K. However, these additions could always be matched by descriptions presented by L, given that the laws that rule these scenarios can be different from the laws that dictate the movements and relations in our world. Adding temporal constants (by the addition of movement) also does not help K's case, for L matches them through the addition of cyclical time structure to the scenario (see MULLER, 2015, pp. 222ff.). Issues with time properties can be dealt with in similar fashion as space properties issues. Thus, whatever relational property K could add to his description of the phenomena can be matched by an equivalent one by L (1975, pp. 254-5).

The morals Hacking intends us to draw from this story is that "*whatever God might create, we are clever enough to describe it in such a way that the identity of indiscernibles is preserved. This is a fact not about God but about description, space, time, and the laws that we ascribe to nature.*" (*ibid.*, pp. 255-6). For Leibniz, the possible is determined by God's will (see RODRIGUEZ-PEREYRA, 2014; JORATI, 2017). Thus, from the Leibnizian point of view, PII is not necessarily true because it is true *in* all possible worlds, but because it is true *about* every possible world, i.e., God cannot wish what He cannot express and if whatever is expressible is translatable in terms compatible with PII, then, PII is necessary. In other words, PII "*is a metaprinciple about possible descriptions. This is part of the force of saying that it is a metaphysical principle.*" (1975, p. 255). The conclusion we can draw for our purposes is that, according to Hacking, relational spatiotemporal properties are not unequivocal providers of predicates used in descriptions of counterexamples that could play the role of truth makers for P2) in the dispersal argument.

3.1.1.1. Objections to Hacking's identity defence

The first objection one could raise to Hacking's point is that, contrary to what he says, nothing requires that the description of the scenario proposed as counterexample to PII omits the number of spheres involved in it as Hacking claims. Most philosophers agree that it is unproblematic to describe the scenario with two spheres as part of the initial setting (e.g., ADAMS, 1979; HAWLEY, 2009; MULLER, 2015, DENKEL *apud* French, 1995; LANDINI and FOSTER *apud* French, *ibid*.). After all, it is a perfectly legitimate way to set up a scenario

starting with: “imagine that there are two spheres and nothing more. These two spheres are…”, if you are precisely willing to show that there are two objects that do not differ in any other characteristic than numerosity. The problem must lie within the description that follows this setting, namely, the features referred by the terms that will fill the rest of the sentence. In this set of features, there must be no conflicting features with the previously settled fact that in this world there are two objects. This is the starting point for discerning defences, as we shall see in the next chapter.

Adams’ raises an adjacent point. He claims that, given that possible worlds are basically collections of objects, Hacking’s descriptions of a scenario with one sphere in a non-Euclidean space and a scenario with two spheres in a Euclidean space would differ in the numbers of objects within, thus, they cannot be descriptions of the same world.[76] This leads us to the same conclusion than the next objection in line – which is probably the most common (found in ADAMS, 1979; FRENCH, 1995; HOY *apud* French, 1995; FORREST 2010; MULLER, 2015), namely, that Hacking counts two possible worlds as one. In a realist framework, if you imagine a scenario in a Euclidean space, this counts as one possible world; if you imagine it in a non-Euclidean space, it counts as another possible world. The same goes for the ersatzist, if you describe the world-story happening in a Euclidean space is one world if it happens in a non-Euclidean space, it is another world. In any case, he cannot omit whether the scene is in a Euclidean space or not, because possible worlds must be *maximal*, be them real worlds or sets of propositions, or whatever. Therefore, the descriptions Hacking claims to be underdetermined by the phenomena that constitute the scenario they are supposedly describing are not descriptions of the same possible world. Each describes a different possible world, given that one contains the propositions “objects are in a Euclidean space” or “there are x number of objects in the scenario” whereas the other does not.

Muller (2015, pp. 218-20) and French (1995, sections IV, V, and VI) present the first and second objections in a tightly intertwined way. I chose to present them separately for the sake of clarity of the essential points of the objections and to show that they may stand alone if the other falls. But the whole point is more or less that for a conventionalist, as Hacking is, the

[76] This is my reconstruction of the point avoiding Adams’ whole talk about thisnesses.

kind of space where the scenario is constituted can be reconstructed as a relativist space or an absolutist space. This depends on something else than the scenario itself and, following Leibniz, he chose a relationist description of the scenario. This means that the space within which the objects exist is not prior to them but exists as relations among them. Thus, the number of objects is a requirement to settle the kind of space the scenario exists in, which makes the difference between scenarios explicit, if one is following the standard possible world building rules.[77]

Nevertheless, a case could be made for Hacking's defence. In fact, it is made by French (1995). He reminds us that Hacking's main point is that whatever "*could or could not be* [the case within the scenario] *cannot be "blandly" asserted*" (FRENCH, 1995, p. 458), there must be reasons for or against one's acceptance of the possibility of either scenario as the referenced scenario. This being established, he continues, the history of physics teaches us that one cannot successfully describe the nature of space without a set of laws attached to it. So, the claims that there are two spheres in a scenario or that the space where the scenario exists in is Euclidean, according to Hacking, must be accompanied with an explanation about how space and the things in it are related. And the proper way to do it, from Hacking's point of view, is from "*our current understanding of the relation between space and the things in it*" (FRENCH, 1995, p. 459). This is because Hacking, according to French, is a natural realist, i.e., he believes that universals (properties and relations) are dependent on the causal structure of the world they came from (our world), thus, subject to those natural laws. In this case, our laws of physics. The same goes for our laws (or principles) of logic, i.e., the laws of logic are also underdetermined by the phenomena and are on equal footing with the laws of physics. Therefore, given that our current laws of physics prescribe a non-Euclidean space for the existence of the properties and relations that we know, the "correct" description of a Black-like scenario would be the one proposed by Hacking, namely, a scenario displaying such and such features and with one particular in it – in agreement with this view about universals and logic.

Hacking insists in the use of the expression "correct description" (HACKING, 1975, pp. 251, 255. French follows him in his description of the case). However, he is not clear about

[77] Their points are more complex than this, but this is enough to illustrate the objection. Nevertheless, for a more robust understanding of the objection, I recommend the reading of the original sections.

what the term “correct” concerns. The use of “correct” gives the reader the idea that there is a clear standard of correctness. Nevertheless, what his critics – and I – show is that there is not, or at least, there is not just one standard. The description may be correct within physical possibility standards, metaphysical possibility standards or logical possibility standards. His whole point depends on a pairing of these standards, which, as we discussed in the previous chapter, is not correct. This confusion of physical (nomological) possibility with logical possibility, and ultimately to metaphysical possibility is another serious problem with Hacking’s defence. It ignores completely the evident difference between the levels of possibility. Certainly, this natural realism is an interesting and theoretically simple explanation to the notion of possibility, but it is, to put it mildly, controversial. The fact that one cannot truthfully state, imagine or point to a contradiction or to an instance of a violation of the principle of the excluded middle, whereas the fact that we can imagine and sometimes even see supposed physical laws being broken, which force us to reach for new laws, seems to say something very meaningful about the rules that govern the actual world and the rules that could be thought to govern other possible worlds. Most philosophers in this debate would disagree with his equivalence. The question about the truth of PII is not whether it is physically necessary, i.e., if it is true in every possible world with our physical laws; but whether it is true in all possible worlds that we could meaningfully conceive. Otherwise, we would be discussing a nomological principle which Hacking himself admits is not the case, as we already saw. Remember, for him, PII “*is a metaprinciple about possible descriptions*” (1975, p. 255). Therefore, to put the different kinds of possibility at the same level is missing the mark.

Additionally, French shows us that Hacking’s view about logic is not as theoretically simple as he thinks. Remind that his views on logic and universals are supposed to be correct (or at least a better view on logic and universals) because they are rooted on empirical facts, such as natural laws (from a descriptivist view), instead of “*cold naked stipulation*”. These learned natural laws and other empirical facts are, nonetheless, subject to the underdetermination thesis, which is not itself an empirical truth, but an assumed one, as cold

and naked as any other stipulation of logic (1995, p. 463).[78] French, then, draws an interesting conclusion: "[…] *you get only as much metaphysics out of a physical theory as you put in and pulling metaphysical rabbits out of physical hats does indeed involve a certain amount of philosophical sleight of hand*" (*ibid.*, p. 466). The philosophical trick here would be clear whether Hacking considered the difference of scope of the laws in question.

Summing up the criticisms, one might say that Hacking's defence is not suited to defend PII, for a) it departs from erroneous descriptions of the scenarios, concerning the number of spheres or the kind of space they are in; and b) it depends on an erroneous view of modalities, mingling nomological, metaphysical and logical possibilities without a good reason for it.

3.1.2. A bi-located sphere: O'Leary-Hawthorne's identity defence

According to O'Leary Hawthorne (1974), Black's scenario used to undermine PII is also used as evidence that bundle theory is false, for it would show that particulars are more than simply an amalgamation of properties. In other words, it would show that two distinct exactly alike bundles of properties co-exist forming two distinct objects. O'Leary-Hawthorne's main target is to show that the scenario does not defeat bundle theory. To this end, he argues that not only the scenario descripted by Black is compatible with bundle theory of immanent universals, but also the interpretation of the scenario by this strain of bundle theory can be used to save PII.[79]

As previously stated in section 2.4.2, in the theory of immanent universals, one has universals as groundfloor entities, and amalgamations of these universals supervene into (thin) particulars. Also, it was stated that there is nothing contradictory for this strain of bundle theorist to accept that a universal is fully present at different locations, e.g., a shade of grey is present at p1 and at 2km from p1. Therefore, it should be as unproblematic that a bundle of universals is as fully present at different locations, e.g., a metallic, grey, sphere-shaped, (…)

---

[78] French's point goes further questioning a bunch of other physical principles assumed by Hacking, but for me the most pressing point is the one concerning the underdetermination thesis, because he explicitly invokes it.
[79] For an alternative view, in which bundle theory does not imply PII and is compatible with the falsity of PII, see RODRIGUEZ-PEREYRA (2004).

mass is present at p1 and at 2km from p1. If this is granted, then, we should expect that from this amalgamation of universals supervene a particular that is 2km from itself. In other words, what we have in Black's scenario is not two diverse objects that present all the same properties, but one bi-located object. Notice that even concerning their spatial properties "they" match, viz., the sphere *is at 2km from an object*, or the sphere *is 2km from a sphere*. Unless one assumes a monadic coordinated substantival description of space, which would yield a difference in spatial properties (e.g., *is at location xyz* or *is at location -xyz*). But, for reasons stated in section 2.4.4.2, this will not be permitted, thus we will stay with the relational description of spatial features.

Furthermore, according to O'Leary-Hawthorne, the same result can be stated for every other relational property one might come up with in this case. This can be understood as a perk of this defence. No further explanation is needed to account for the object within this scenario. Every property is described and listed, and the scenario appears as it was intended by the PII opponent. It looks like two objects, the list of features displayed by the object presents spatial features such as *is at 2km from a sphere* or any other, but the list of objects contained in the scenario contains a single object. The explanation for this is found in the nature of the notions of property and objecthood themselves. The distanced co-existence of the two bundles of properties thought to mean numerical distinction is, in fact, just a bizarre case of bi-location of a particular. Something that is acceptable for the immanent realist. Thus, it seems that Black's scenario does not entail the falsity of bundle theory – at least not of this strain – and, if this view is correct, then it saves PII from any kind of dispersal argument, since P2) cannot be exemplified by a scenario of this kind.

Nevertheless, consider the rebuttal from the PII opponent: what if we consider the different appearances of this bundle of universals as the objects we are counting? One might say that the whiteness in the frame of my window is one thing and the whiteness in my neighbour's window frame is another thing. It makes sense, O'Leary-Hawthorne admits, but it only makes sense in loose uncompromised daily talk, the same way one calls shellfish and

whales fish (depending on the country, I guess) or calls T-shirts shirts.[80] Metaphysician talk would not be this sloppy. The metaphysician who is concerned with a true description and understanding of the world would not do this, she would consider both whiteness as the same universal. Unless she is a tropist, who believes that each of the appearances are in fact different properties – or a nominalist, I would add. This is up for debate, but not here. In this case, we accepted immanent universals bundle theory. Further reasons to adopt a different view on the nature of properties and objecthood are required before one could accept that this scenario presents a counterexample against PII, or, at least, reasons not to accept this immanent realist bundle theory.

Another point brought up by O'Leary-Hawthorne is that spatial distance is usually considered to be a sufficient condition for numerical diversity (*ibid.*, p. 195). He himself replies to this objection by saying that in common talk this might be the case, however, with a more perspicuous analysis one can see that the rules for attributing identity or diversity are more complex than that. In common discourse spatial distance normally entails a whole bundle of differences (e.g., spatial relations to other objects), which imply distinction. However, in this unusual case, distance does not imply any difference in the bundle. Even the relational spatial property the bundle displays is the same. Therefore, no object distinction arises, as the PII attacker would be content to agree. The author illustrates this point with yet another analogy contraposing layman with specialist discourse: by changing some rule of chess, the common man may still say that they are playing chess, however, a specialist would say that it is a whole different game. Thus, in some vague sense, one might say that there are two spheres in the scenario without lacking with the truth, as one can say that a whaler is a fisherman. But in the relevant sense, which is the metaphysical one, there are not two objects in the scenario; provided bundle theory of immanent universals is true, of course.

O'Leary-Hawthorne summarizes his point by saying that it might be that immanent universals bundle theory is absurd, since it is such an unorthodox view. This is yet to be decided.

---

[80] I am enriching his point with different and further examples. For his original unfolding of the point see (1995) pp. 194-6.

However, as long as it is not proven false, PII can be plausibly saved from a dispersal argument using a Blackean scenario by the adoption of such theory of universals.

3.1.2.1. Objections to O'Leary Hawthorne's identity defence

Rodriguez-Pereyra (2004) points out that O'Leary-Hawthorne does not specify which conception of space he adopts, thus it is not completely clear what it means to say that a bundle is a bi-located particular. Whether he adopts an absolutist conception of space – which we already said is not going to be favoured – the conclusion would be absurd, Rodriguez-Pereyra claims. To the absolutist, regions are sets of points and being in a certain place is occupying certain points in this conception of space. However, for the immanent realist, points are really a single universal entity, i.e., pointhood, that bears infinite spatial relations to itself. Rodriguez-Pereyra concludes that for the immanent realist: "(…) *being in a place is occupying the single universal parthood. But if so, since all spatial things occupy the same entity, namely the universal pointhood, all spatial things occupy the same place. This is absurd.*" (2004, p. 75). Thus, one cannot be a bundle theorist of the immanent realist train and an absolutist about space at the same time. This is not the most charitable interpretation though. It seems that to occupy a place in this conception would be better understood as having a set of these relations bore by pointhood.[81] Therefore, to save immanent realism one would have to abandon absolutism.

If, on the other hand, O'Leary-Hawthorne adopts a relationist approach, that it would not save him either, Rodriguez-Pereyra claims. O'Leary-Hawthorne would not be able to distinguish between a world with a bi-located sphere from another world with a tri-located sphere where the spheres form an equilateral triangle. According to him, in both worlds the relation that instates the spatial locations of the spheres *being at 2 km from* (itself) is equally satisfied. In other words, the bundle of immanent universals in question possesses the spatial

[81] Later in the paper (*op. cit.*, p. 80), RP presents a more charitable interpretation, closer to the one I would favour. Nevertheless, he conditions this interpretation of spatial absolutism for bundle theory to the acceptance of his interpretation of bundle theory, which is not immanent realism. In this view, different sets are applied to different object, which means that they have different spatial properties, thus, they would not be indiscernible, although RP is not disposed to argue in defence of PII in this paper.

relations in any of its locations. Therefore, a world with a bi-located particular is the same as a world with a tri-located particular, which seems intuitively wrong.

One might claim that in the bi-location world it is a dyadic relation in play, e.g., *being at 2 km from* x, while in the tri-location world it is a triadic relation that is at play, e.g., *being at 2 km from* x *and* y – where "x" and "y" can be substituted by the same particular. The problem in this case, according to Rodriguez-Pereyra, is that this triadic relation is satisfied in the bi-location world too, i.e., the bi-located particular is always at 2 km from itself *and* at 2 km from itself. The problem here seems to be analogous with Hacking's problem, but in the other direction. Hacking claimed that different worlds could understood as descriptions of the same world, whereas O'Leary Hawthorne has no means to differentiate clearly distinct worlds with his worldmaking tools. Therefore, it seems that this defence is problematic from the starting point, its worldbuilding tools do not work.

However, let me play as Devil's advocate for a minute and claim that maybe RP is not contemplating all the properties available to differentiate these worlds. He failed to consider the property of *being at 2 km from something also being 2 km from something at a 60° angle from the other distance* which, unless there is a constraint prohibiting one to attribute angular positioning together with distance attribution in spatial relations, seems to be instantiated by the tri-located world whereas the bi-location world could not instantiate it – nor could an eventual tetra-location world. Therefore, RP's objection is incorrect and another objection to immanent realist bundle theory allied with relationalism must be presented.

Fair enough. In her (2009, pp. 106-8), Hawley assumes the possibility of a single bi-located sphere in Black's scenario and brings up some puzzling mereological points that a bundle theorist of the immanent strain must explain. Suppose that this bundle of universals has as constituents the properties *1kg mass* and *1m^3 volume*. This should be unproblematic, after all the spheres must have a mass and a volume since they are made of matter and are in space. However, what should we expect of the mereological sum of these bundle *appearances*?[82] In other words, what should we expect of the sum of one appearance of grey metallic spheric

[82] What I am calling appearances here might be called occurrence by other or even instances. I opted for "appearance" to invoke the idea of phenomena.

voluminous mass with another appearance of it? According to her, there are two ways the bundle theorist could unfold this question. One is saying that the mereological sum of the two appearances is a sphere itself with 1kg and $1m^3$ of volume in spherical configuration, whereas the alternative is to say that the mereological sum is a 2kg, $2m^3$, non-spherical object. The latter must be true, because certainly there is something contained in this scenario with the features of non-spherical shape, 2kg and $2m^3$. We can describe it and we can picture it in the scenario, and there is no other thing in the scenario aside from two appearances of a bi-located object. Additionally, the former alternative escapes our basic notions of arithmetic and geometry. Opting for the former alternative would be like saying that the mereological sum of me and a copy of myself is just the same as me (1 person plus 1 person equals 1 person). Hawley even considers the possibility that the object has no mereological sum, but dismiss it, for even "*the sum of 0 plus 0 is guaranteed to exist if 0 does, for it is just 0 itself*" (*op. cit.*, p. 107).

If the object in question is not a metaphysical simple, it has parts. If it has parts, then these parts have a mereological sum. If the object is a metaphysical simple, then which is the simple in the scenario? Would it be the scattered appearances, the junction of them, or both? Any of these options are very hard to swallow. In any case, she considers a way these sums and divisions could work: what if the object presents the properties of *being $1m^3$ in volume* and *being $2m^3$ in volume*? Given that the object occupies two distinct $1m^3$ space regions, in some sense it must occupy $1m^3$ and $2m^3$. Ok, but it does not occupy *exactly* both. This would be paradoxical, e.g., the referenced object presents the property of *being exactly $1m^3$ in volume* and the property of *being exactly $2m^3$ in volume*. The same goes for its mass and, in a different way, for its shape too, i.e., how could it have the property *being exactly spherical shaped* and *being exactly two-sphere-shaped* (which entails *being non-spherical shaped*) at the same time? Well, this is starting to look more and more like a series of categorial mistakes not just concerning numerosity but concerning other features too. Hawley, then, is puzzled about how property-exclusion would work in a world like that. So am I. Is every property a part has transplanted to the whole and vice-versa? Questions of this kind must be answered by the immanent realist in order to make sense of the property mechanics in this world, otherwise we cannot make sense of it.

Nevertheless, Hawley considers another route that the bundle theorist could take. Maybe these properties must be relativized by the regions they occupy, e.g., the object has 1kg and $1m^3$

relative to one of the spherical regions it occupies, and 2kg and $2m^3$ relative to the two-sphere-shape region it occupies. This generates several problems, according to her. First, it would reduce drastically the number of intrinsic properties of the object. Second, it imports a conceptual machinery designed to deal with variation into a problem that has no variation (indiscernibility). Finally, she believes that it cannot apply to quantum cases, which deal with co-located entities.

I believe that the first problem is not a problem *per se*. I personally favour a more holistic and relationist approach to features, i.e., I do not see why things cannot be mainly made of relations. I would not say totally, because there seems to be some things that are not relational, nor relative, e.g., arithmetic end, perhaps, mass. But if relativizing the properties to their spatial locations would not violate the arithmetic and there is a feasible topology for this world, why cannot we say that there are very few intrinsic properties?[83]

The second problem does not strike me as a problem *per se* too. A knife was built to cut but can also be used to spread butter and as a screwdriver if needed. Her point, I believe, is that most of the notions and values would be theoretically idle, making it a non-economic theory. However, these expenses might be justified by their explanatory power for other scenarios. If it explains this case with no loss to explanations of ordinary examples, then we would just be trading theoretical virtues.

The last problem seems to be the most serious. If the trick was to relativize attaching the values of mass and volume to different spatial locations, co-location would disable this strategy. But to be fair, co-location is a problem for other defences too. The way out she and Saunders propose is to adopt a new category of entities, namely, non-individuals, as we shall see in section 3.3. The other way is to reject co-location, as Della Rocca (2005, 2006) does, as we shall also see in section 3.2.1. This would be a fair exit for the immanent realist, though he would have to find a way to reconcile Della Rocca's challenge with the immanent realist extravagant mereology and property mechanics he still owes us. In any case, the bundle theorist lost me at the mereology problems.

---

[83] To be fair I would not discard a more radical view such as relationalism, viz., the view that every property is relational. This is up to debate, nonetheless, here it is not the place for this, thus, I am content with the less radical view of relationism.

Nonetheless, my point concerning these objections is basically the same Hawley's:

> *This is not a* reductio *of the identity defence. But these considerations do show that taking the defence seriously will require us to completely rethink the relationships between what properties are instantiated by the objects in a world and the way in which regions of that world appear to be filled and qualified.* (2009, p. 108.)

However, I believe that a *reductio* could have been called out when the first mereological problems arouse or even earlier. The idea of a bi-located particular is one that I would not be inclined to accept *ab initio*, for it strikes me that it would be very difficult to explain interactions with these particulars, which, in turn, would somehow put a wedge between the idea of particular and individual, something that, from my point of view, should be intimately intertwined. I any case, Zimmerman points out a similar problem for this defence that should be enough to discourage pursuing this road.

In his (1997) dialogue, Zimmerman highlights some problems concerning modal properties of the object(s) in the scenario. One of his characters, namely, A, is supposed to represent O'Leary Hawthorne, while the other, B, represents himself – or so I suppose. As soon as A finishes his description of how Black's scenario should be understood in the light of bundle theory of immanent universals, character B brings up the contradiction of something being at a distance from itself – in a Euclidean space let us assume. A acknowledges it and even assumes the possibility that the whole immanent universal idea is absurd, as O'Leary Hawthorne did. However, since B allowed the existence of immanent universals and its ontology, the possibility of self-distancing as bi-location is entailed.[84]

This being settled, they proceeded to re-descript Black's scenario in terms that do not beg the question for the bundle theorist. Remember that Black's description states from the start

[84] I do not see self-distancing as a crucial problem. There might be cylindrical universes, such as Hacking's where this can happen. The major problem, as I previously said, is bi-location.

that there are two spheres. But to assume this would be unfair to the bundle theorist, A claims. We must, then, assume that there *seems* to be two spheres or that the phenomena descripted in the scenario is compatible with there being two spheres, though there is actually only one sphere. They reached the point where A claims that the scenario consists of two appearances of a bi-located bundle.

Then, B asks us to think of a Black-like universe with only two electrons (or a bi-located bundle of electronish properties) and from this redescription he puts forth his objections. He asks his public to imagine that such universe obeys indeterministic laws, which is a very reasonable demand. In such scenario, it is possible that at some point these electrons present different behaviours, e.g., at some point one electron starts spinning up, whilst the other keeps spinning down.

Forecasting objections of the kind "but we could also imagine an eternally symmetric universe where these changes never really happen", he acknowledges this possibility and claims that even in this case the physical, metaphysical, and logical possibilities of these changes happening are still there, i.e., these electrons would still have these *modal properties*, in other words, they would still have these potentialities – if one prefers a more Aristotelian scent – (*op. cit.*, p. 306). I argued against the dispositional properties and aim to avoid them. B would not have to appeal to them, he could simply use the same reply used against Hacking: of course, there could be worlds where nothing of the sort happens, however, there is the possibility of worlds in which these differences come up and PII is meant to be true even in these worlds.

Anyway, back to Zimmerman's dialogue, B foresees A's objection to this claim, saying that A could never acknowledge it because he conceives these two electrons in fact as just one bundle of properties in a way such that whatever property one has, the other must also have. In other words, there is nothing that could be true of one and false of the other. A promptly agrees and recognizes that it is an impasse that stems from another disagreement they had about how to correctly interpret modal statements. A believes in some sort of counterpart theory, that says that modal *de re* statements must be interpreted as saying not what an object can do in other possible worlds, but what other object (or objects) very similar to it in some relevant aspects and different in other less relevant aspects does. In this theory, A claims that he could interpret B's electronic scenario as having one bi-located electronish bundle that in another world has

two counterparts that differ from it in the way required by B, namely, presenting a different behaviour than their counterpart.

B does not share this belief in counterpart theory. But he points out that even if he accepts A's description of the scenario, a problem still arises, because, according to B, if the single bi-located particular could have been two electrons (this is what A implied, when he allowed that the electronish entity could have two counterparts), something else would be possible:

> (…) *the electron on the one side could have developed differently while the one on the other side did not. But if "they" are identical, "they" must have the same counterparts in every possible situation – and so there's no possible world in which the one* but not the other *has a counterpart with a particular future.* (p. 307, original emphasis)

This passage is somewhat confusing because B talks about the bi-located particular as if it was two electrons and he describes a situation where these two appearances of the same particular evolve to different things. In this future scenario two twin particulars arise that have as counterparts the bi-located predecessor. These twins are not identical in their world, viz., in their time-slice, given that they have different properties. One of them presents a property G after a certain point, while the other does not. They both have a bi-located past counterpart that presented no G. Thus, it seems that this past counterpart has future counterparts with contradictory properties.

A might have replied that this is no bigger problem than reproduction by budding. Until a certain point the appearances were one bi-located particular and, as soon as property G appeared in one of them, it turned out to become a new individual, while the appearance that kept ¬G would be a counterpart of the bi-located particular. This is not obvious at first glance, because the way B describes the situation is a way in which the two appearances are already assumed to be two different individuals. In the case where the past bi-located particular had G and its parts changed into other individuals with no G (let us say that one evolved to present ¬G, ¬H, and F, whilst the other ¬G, H, and ¬F) it might be trickier to say which is the

counterpart, given that the difference to the past stage is in some sense symmetrical. Maybe both are counterparts as it was already stablished as possible.

However, there seems to be two problems for this possible reply from A. The first is that there are two counting standards in play. One for A's particular, which is bi-located, and other for B's appearances of this particular, which are understood as two electrons. Thus, we are talking about different entities in each case, i.e., the counting standard changes with the scenario's evolution. The second problem concerns the appropriate way to attribute counterparts, that is, how can we say which object in one world (or stage) is a counterpart of another object in another world (or stage)? It seems that it is a problem of identity trough time, which in turn requires a diachronic analysis of identity, and that is beyond the scope of PII and of this thesis.[85]

In any case, A opts for another line of replies. He questions the claim that assuming both appearances of the electronish bundle would enable changing their states differently from one another, given they are the same thing. B, in turn, counters saying that it follows from A's view that these electronish bundles would have to act like bosons, i.e., that though being different things, they are linked in such a way that they are also necessarily forever synchronized, something A had denied before when he admitted the possibility of a world with two counterparts for his bi-located particular. A, then, finally recognizes that there is something odd in his theory.

Thus, he proceeds to argue for a version of bundle theory exclusively for the actual world, ignoring symmetrical worlds such as the Blackean one. B points out that this move is *ad hoc* and serves A the Adams' argument from almost indiscernibles, in which a non-symmetrical world turns into a symmetrical one by the subtraction of the different property that distinguished the two spheres (or bundle of spherical, metallic, etc. properties), to show why this would not be acceptable. The dialogue, then, ends with A retreating from biting the bullet of the *ad hoc* contingent bundle theory to the idea that it is not clear that the scattered appearances of the

[85] For a possible solution to these problems, see the treatment of the problem of discernibility of identicals, especially the amoeba example in BAXTER, 2014, p. 247-8ff.

eletronish bundle must not behave like bosons. To which B replies that it would be obvious if A was not such a devotee of the bundle of immanent universals theory.

Adam's argument in this context would work as follows: It is possible that there is a symmetrical world with almost indiscernible particulars that – allegedly – only differ by the possession of property *G* by one particular and the lack of it by the other. Something happens and both particulars lack *G* from this point onwards. According to A's theory these two particulars would become one single bi-located particular or the *G*ed particular would cease to exist and another ¬*G*ed instance would appear? Any alternative seems farfetched, but I believe the reply mostly tunned with this strain of bundle theory is the latter. It presents a strange mechanics of properties and an even stranger cardinality in this ontology, but it is what it is. The best way to avoid this situation, in my point of view, is to give up bi-location and claim that there is no way to go from a scenario with almost indiscernibles to a scenario with indiscernibles just giving up a difference property *G*. For almost indiscernibles to become indiscernibles, I believe the particulars would have to accord concerning *G* and somehow become co-located, i.e., they would have to perform some reverse budding, because it seems that spatial features are intimately linked with numerosity. This will become clearer in the next section. Summing up this point, although I do not believe that Adam's argument works to prove the existence of indiscernibles, I believe it works to show how bizarre the idea of bi-location is.

One last criticism that was already mentioned but should be stressed further is that there is a confusion about what should be counted as numerically identical in this scenario, the particular or the occurrences of co-present universals? The immanent realist would argue that the particular is the object in question. On the other hand, to me and to others, such as Rodriguez-Pereyra (2004, p. 78)[86] and Zimmerman (remember that character B treats the objects as separated electrons), the correct way to describe the scenario in Muller's step 1 seems to be making the distinction between the bundle and the occurrences (RP calls it instances) of this particular. The bundle is the amalgamation (co-presence) of universals. The particular, understood as the bundle of universals, is one, yet the occurrences, which in fact constitute the

[86] My terminology is not in tune with RP's. To him the "particular" are the instances (which I called occurrences, to avoid import of previous discussions concerning instantiations), while I used "particular" to capture the bundle of universals, following Hawley (2009). Notwithstanding, our aims are tuned – until certain point.

scenario are two. The appropriate descriptions of objects in the scenario seems to me to be counting the appearances as the objects. It seems to be the most linguistically adequate alternative and a more faithful description of the phenomena.

We can sum up the criticisms against O'Leary Hawthorne's defence by saying that it depends on the truth of the immanent universals bundle theory. This metaphysical view, in turn, implies the strange unpalatable notion of bi-located particulars; an unexplained mechanics of how properties interact (property addition and subtraction), that results in a unexplained cardinality of its ontology, both probably unexplaineable; obscure results of mereological operations, such as those highlighted by Hawley; and, finally, it is performed in a confuse vocabulary about objects that ended up just moving the distinctness from particulars to appearances/instances. These reasons seem enough to avoid any defence that comes with so many malfunctioning pieces.

o o o

The adoption of identity defences turns out to be a bad strategy to defend PII after all. Notwithstanding, instead of considering that they are completely useless, Muller (2015, pp. 223-4) presents it as a relevant defence for purely auditive scenarios such as the one proposed by Ayer (1972), in which one claims that, in a cyclical universe containing only one sound token that repeats itself every cycle, these repetitions would be indiscernibles. Well in this case, it is the *same* token that is heard within this universe, thus, no violation of PII has happened. There we have a clear case of an identity in action. In case the opponent insists that it is the same token, but with different occurrences, Muller claims that this hearer must be counting this from outside this cyclical universe, i.e., in a universe where he can *discern* these occurrences by their relative or weakly discernible properties. In this case, he considers, one would have to appeal to a discerning defence, which shall be the object of our analysis in the next section.

3.2. Discerning Defences

Discerning defences are those that claim that "*there is, on closer inspection, some qualitative difference in the two (or more) distinct objects*" (MULLER, 2015, p. 205) in

counterexamples used as instances for P2) in the dispersal argument. These defences are clearly espoused by Muller (2015) and Ladyman (2005). Nonetheless, I believe that Della Rocca's (2005; 2008) defence, though it adopts a very singular strategy, is best classified as a discerning defence too, for it forces the opponent to acknowledge that there is a discerning feature playing a role in Black-like examples. Being so, in the following sections I shall present Della Rocca's defence followed by a section with criticisms mainly from Jeshion (2006) but also from Hawley (2009) and Muller (*op. cit.*).[87] Then, I shall present Muller's defence as the best representative of the alternative and apparently more popular genus of strategy within this family of defences, followed by a mix of criticisms by French and Krause (2006), Saunders (2006), and Hawley (2006; 2009).

3.2.1. Forcing discernibility: Della Rocca's challenge

Della Rocca (2005; 2008) provides an interesting and peculiar strategy of defence for PII. He is sceptical about the possibility of a scenario such as the described by Black, because any such scenario appeals to brute facts about identity – something he is open to, but only as a last resort to make sense of things (2005, p. 487). He aims to challenge Black-like experiments on their individuation processes. According to Della Rocca, the proponents of Black-like counterexamples to PII accept that there is a primitive individualization or there are brute facts about the non-identity of objects; in other words, the individualization (or the non-identity) of objects ontologically precedes any of their qualitative properties. He challenges this idea by asking how can we say for sure that there are two and only two spheres, if they have all the same qualities? To which the scenario proponent would have to present some discerning property or to stubbornly say that it is a given fact. Then, his strategy is basically to put the opponent of the PII, who accepts such brute factualization of diversity (thus, of identity also), in an awkward situation where he must answer a question like "how can you say for sure that

[87] Initially other objections more closely related to problems in philosophy of language, such as direct refence and transmundane identity that can be found in Cross (2009; 2011) would also be in this section, however they had to be cut from this version for lack of space and time.

there are no other twenty spheres co-located with one of those spheres, given that in this case they would also have all the same properties too?"

The idea he is sneaking in is that given that there are no facts grounding the distinctness (or the identity) of the spheres, then, in principle, there are no facts impelling us to think that there are only two spheres. There could be twenty spheres with all the same qualities occupying the same region of space as any of those we thought to be the original spheres. Unless the opponent presents a criterion that involves some qualitative properties to account for the individualization and distinction of the spheres that exist in such universe, there is no good reason to reject the hypothesis that there are twenty other spheres co-locating the regions of space we previously thought were occupied by two spheres.

Let us make the case more explicit. Della Rocca asks us to imagine that, right in front of us, there are twenty spheres with the same weight, size, shape, etc. at the same place, in fact, sharing all the same parts. Also, the spheres do not move and all of them came into existence at the same time and will exist for the same amount of time. So, even the temporal parts are shared. These spheres present all the same properties, yet they are not the same, because they were primitively individuated. This situation looks somewhat weird and philosophically unpalatable, for nothing forbids us from saying that there are two thousand or two hundred thousand spheres in the same situation as the twenty spheres described. This is an Ockhamnian nightmare for any world view!

A point worth stressing here is that Della Rocca is not saying that there cannot be objects of the same kind in the same place, e.g., an iron sphere and an aggregate of iron molecules. No. In fact, this would be philosophically tasty and a commonly accepted suggestion by those who see identity as composition, a subject Della Rocca and I shall remain neutral during the discussion about PII.[88] The Ockhamnian nightmare is something else. It requires a proliferation of entities of the same kind. If this is clear, we can, then, describe Della Rocca's challenge as an account of a Black-like scenario in which there are two indiscernibles without unleashing the Ockhaminian nightmare.

---

[88] For interesting views on the subject see BAXTER (2014) and COTNOIR (2014).

Della Rocca entertains a few escapes for the PII opponent. One could say that what individuates each sphere is the spatial location where they reside, e.g., "One sphere is at region *a*, whereas the other is at *b*". Notice that here Della Roca is assuming that the spatial location of the spheres are not features of the sphere and he is allowing the naming (or something similar) of the spatial locations, something not allowed for the defender. In this case, Della Rocca would move the same challenge a step further: how those regions of space are themselves individuated? How can we know that there are not twenty overlapping regions of space? Those answers would depend on which concept of space one holds, viz. substantival or relational. Della Rocca considers the alternative where space is relational and argues that if the opponent were to individuate the space regions through the spheres, the argument would be circular, for the spheres should be individuated through the space regions. Therefore, if space is to be understood relationally, it must be relative to something else. If space is understood substantivally, on the other hand, a reason for the individuation of its regions must be provided. However, as it was previously discussed, this would involve naming or labelling space points, which is forbidden in the act of discerning. Does this mean that space points are indiscernibles? No. They are relationally discernible. This is perfectly fine from a substantivalist point of view. Substantivalism says only that space points are ontologically prior – viz., independent – from objects, not that they are intrinsically determined (see MULLER, 2015, p. 220).

Another alternative entertained by Della Rocca is an appeal to simplicity. One could argue that it is simpler to think that there is only one sphere at each place. The mere appeal to simplicity is not a good strategy, for it would be way simpler to think that the Blackean scenario has only one sphere at all, i.e., Hacking's cylindrical universe. Thus, an identity defence could be invoked and PII would be safe. Jeshion (2006), who argues against the need for accepting PII to avoid the twenty spheres case, agrees with Della Rocca that simplicity could not be effectively invoked here. She claims she is "(…) *with him in remaining extremely doubtful that any principle of simplicity will be found that will suffice to do the job of admitting bi-located*

and ruling out co-located indiscernibles that share all their parts" (pp. 165-6).[89] Thus, I see no reason to pursue further this route.

Della Rocca, then, sketches three more ways the opponent of the PII could try to debunk the twenty spheres case while saving the two spheres case, though he considers these ways to be defeatists. The first of them is to accept that there is no explanation for the conundrum. He is willing to accept that there are brute facts about individuation (viz., about identity and distinctness). But only as a last resort. If there is an alternative explanatory route, we must take it. But the PII opponent has not made the effort to prove that there is no other alternative so far. The second way is to claim that the two spheres case is legitimate because of our individuating practices, i.e., "*our practices of individuating objects embodies the view that the former* [two-spheres] *case is legitimate and the latter* [twenty spheres case] *not*" (2005, p. 487). He claims, however, that this would be a case of the banana eating the monkey, because we want our practices to fit the metaphysical facts, not the other way around, at least ideally. The third answer is, according to him, remarkably similar to the previous. One might claim that the twenty spheres case is not legitimate, because there is no way to know how many spheres there are in that region of space, whether there are only twenty or twenty thousand spheres in that place would be undecidable, for it always looks as one sphere; whereas in a Black-like scenario there is a way one could attest the multiplicity of objects. This means that there is an epistemic difference between the cases. Notwithstanding, according to Della Rocca, this answer presents the same problem as the second one, that is, it takes things backwards. Epistemic facts should not dictate metaphysical facts, but the other way around, ideally.

Only ideally. Jeshion agreed with him so far (2006, p. 165-6), but parted ways later. I part ways with them now (although I will reunite with Della Rocca down the way). Given that the context of PII is intimately linked with Quantum Mechanics' talk, we should at least take more seriously the possibility that epistemic facts (viz., observation) might influence metaphysical facts. Moreover, given the growing attention enactivist theories are receiving in

[89] What she means by "bi-located indiscernibles" is different than what O'Leary-Hawthorne meant. For her, the two spheres proposed in Black's experiment are what she is calling bi-located. They are not really bi-located since they are different individuals. We should interpret "bi-located" here as characterizing the distinguishability. Thus, these bi-located indiscernibles are only absolutely indiscernible.

Philosophy, it seems reasonable to take more seriously the idea that our practices might somehow place limits to our ontology; especially if we take seriously debates of metametaphysics, or philosophy of philosophy. If we are favouring a more naturalistic – in the Quinean sense – approach, as I claimed to be doing here, it seems inevitable that our metaphysics depends – at least partially – on epistemology. I will argue latter on that our metaphysics is intimately linked with our conceivability faculties. I believe that it is precisely because of this epistemic dependence that we are able to derive a difference between the spheres in the Black-like scenario. Nevertheless, the point Della Rocca is aiming at is very close to this. He aims to make the PII opponent question this epistemic difference he sees in Black-like cases whereas it is absent in the twenty spheres case. The answer, as one would see, is a metaphysical fact, namely, the existence of different spatial features between the spheres in the two spheres case.

There is still one way the opponent could get rid of the twenty spheres case, which is appealing to what Jeshion calls the *Weak Lockean Principle* (WLP):

(WLP): There cannot be two or more indiscernible things with all the same parts in precisely the same place at the same time.

Della Rocca is happy to accept this principle. After all, it seems absurd to deny it. In fact, he goes beyond Jeshion's considerations and ponders whether it would not be a conceptual truth regarding the concept of objecthood (2008, p. 30). Nonetheless, as all who followed the debate over analyticity know, relying on conceptual truths may be building up in quicksand. To be consistent, however, although Della Rocca considers the possibility of WLP being a conceptual primitive truth, he must apply the same scepticism he applied to the primitiveness of individuation to WLP, i.e., he must ask whether there could not be something beyond WLP. Could not it be grounded the case that WLP is also grounded in something else?

The reason for this questioning is very compelling. He asks the reader to consider cases of partially overlapping objects such as Siamese twins, statues made of smaller versions of it, branching trees, branching objects in general. Consider the example of two Siamese brothers who share the same pair of legs and most of the torso. Are they the same individual? Clearly not. Each has his own pair of arms and his own head, which would confer them their

individuality. But what if they share every limb, but in their head, they have two brains? Well, in this case, they would still be twins that share a whole body but their brains. Their brains individuate each of them. Fair enough, but what if they share everything, every single cell? Well, in this case it seems that "they" are the same individual, whatever one feels or does, goes for the other too. It seems that partial overlap of objects is acceptable, but complete overlap is not. But why is so? According to Della Rocca (2005, p. 489), it is so because partial overlap allows for an explanation of non-identity, while complete overlap does not. More precisely the reason seems to be that the non-individuation in the partial overlap case is given by a *difference* in parthood, which is not present in the case of complete overlap. In the complete overlap case, the non-identity is gratuitous. The case of twenty spheres only presents gratuitous – viz., trivializing – properties such as *being sphere A* and *being sphere B* as criteria for non-identity. Such properties present no explicative value whatsoever to account for the non-identity or individuation of these spheres. Now, if the spheres would have at least one unshared part, one *difference*, we could use it as a criterion for explaining the non-identity between the spheres and the problem would be solved.

Della Rocca invokes the notion of explicability to account for the acceptance of WLP, which is a reasonable move. This means that, unless there is another account for the WLP, it seems that the opponent of the PII would have to accept this account and the unpalatable consequence that comes with it, viz., a commitment to a mereological explanation for non-identity. In other words, Della Rocca's point is that, in order to avoid the absurd twenty spheres case, the opponent of the PII must appeal to WLP, which in turn is grounded in mereological facts that can also be used to distinguish the spheres in a Black-like scenario.[90]

[90] It is worth to stress that, according to him, although this argument is directed to Black-like putative counterexamples that deal with spatial material particulars, it is applicable to examples of non-physical objects, such as minds (in reference to Adams (1979) counterexample) and – presumably – non-individuals, the entities postulated in QM by some like Saunders (2006) and Hawley (2006; 2009). See DELLA ROCCA, 2005, pp. 490-1.

3.2.1.1. Objections to Della Rocca's argument

Della Rocca's requirement for an explanation of why one should accept WLP is reasonable. However, Jeshion argues that accepting PII is not necessary for accepting WLP. According to her, there are at least three ways to avoid Della Rocca's requirement for WLP's explanation. She argues that if there might be brute metaphysical facts, WLP would be one of them and, thereby, this explains why it is impossible to have co-located indiscernible individuals. Hawley, on the other hand, subscribes to Della Rocca's point of questioning the reasons for denying the absurd twenty spheres case on the one hand, but accepting Black-like cases on the other. However, she argues that there is a difference between a twenty co-located spheres case and a case of two spheres placed in different locations, namely, their additive properties. The treatment of these additive properties will require a different defence. Muller, on the other hand, stresses that though Della Rocca's point is interesting, nonetheless he neglects a kind of property that could end the debate more easily, viz., relational properties. In other words, Della Rocca's whole trap system for the PII opponent is unnecessary, given that the discernible feature between the Blackean spheres is evident from the start.

The first way Jeshion sees for the PII opponent to get out of the awkward position created by Della Rocca's challenge is the acceptance of the primitive metaphysical truth of WLP. She champions this view by making an analogy with the cases of acceptance of primitive mathematical truths by mathematicians. The paradigmatic case that she brings up, unfortunately, is Frege's postulation of Basic Law V in the *Basic Laws of Arithmetic*. According to her, Frege "*felt that theorizing within the system he proposed the* Basic Laws of *Arithmetic would help one recognize the self-evidence and primitive standing of Basic Law V* (…)" (2006, p.168). This theorizing would not ground the law, it would only help its understanding as stating a brute fact.

The problem here is that in such cases it is hard to talk about primitiveness. Consider analytical truths, for example, which were once considered primitive (in a different way, of course), but since the debate over their reliability started with Quine, philosophers know that the primitiveness stated by some analytic statements was not the case. The same can be said to occur with Frege in that case. As Jeshion points out, Basic Law V turned out not to be a logical principle, hence, not primitive in the sense intended by Frege. It seems that her own example

works against the acceptance of such line of argumentation. What seems to occur in cases of mathematics is that no counterexamples could even be conceived. Whenever counterexamples are conceived it becomes clear that such principles were never axioms or universal truths to begin with, i.e., they were not primitive or brute facts. This is what Russell's paradox made with Basic Law V, what non-Euclidean geometries made to Euclid's axioms and what Black's scenario intend to make with PII. The abundance of examples of truths thought to be primitive universal and immutable truths should be enough to discourage anyone to follow this line of thought.

An alternative way to see the primitiveness of WLP, according to her, is by noticing that there is no need for symmetrical treatment of the explanations for the partial overlap case and the complete overlap case. It is perfectly possible that there is an explanation for the acceptability of the non-identity in partially overlapping objects, namely, the existence of non-overlapping parts, whereas the explanation for the non-acceptability of cases of totally overlapping objects is absent or is simply the existence of WLP as a brute fact. In other words, she sees partial overlap and complete overlap as different phenomena, different kinds of non-identity, therefore there is no need to match explanations for them.

This alternative, however, presents an unnecessary double standard. Della Rocca insists that both phenomena are of the same kind only differing in degree and I must agree with him. The adoption of such double standard seems merely *ad hoc* for the acceptance of the primitiveness of WLP, since Della Rocca's alternative is explicatively simpler (or uniform). Della Rocca stressed that we *could* accept primitive facts, but only if there is no explanation available. I would go further and add that we *should* accept them, but only if it is demonstrated that no further answer can be presented. However, there is clearly an alternative available, which Jeshion completely glosses over in favour of a more complex explanation. It seems that she does that just to avoid answering a fairly legitimate question posed by Della Rocca: why should WLP be considered a conceptual truth? Or, in other words, could not it be the case that WLP is analogous to Basic Law V, Euclid's axioms and a myriad of analytical truths with the same destiny?

The second way to avoid the twenty spheres case without relying on the primitiveness of WLP proposed by Jeshion is to consider the non-identity in the twenty co-located spheres case gratuitous. Gratuitous in the sense that there is no – and cannot be any – explanation for

accepting such uncomfortable possibility. Thus, if something is gratuitous, this implies that it is impossible (see JESHION, 2006, endnote 14). This is perfectly in line with Della Rocca's claim. However, they diverge as she does not think that we deserve an explanation for every non-identity in terms of qualitative properties while claiming that the twenty spheres case is gratuitous. In other words, according to her, total overlapping cases are unacceptable because of gratuitousness, i.e., because they need an explanation that cannot be given. But this does not mean that every non-identity (and identity) case has an explanation or needs an explanation, i.e., it does not mean that there cannot be primitive metaphysical facts about non-identity in other cases. It might be that cases such as Black's spheres are cases where no explanation in terms of qualitative properties is needed. Again, a double standard in play.

To support this claim, she puts forward some examples of some identities between partially overlapping objects that are gratuitous but accepted, highlighting that the existence of this double standard should be unproblematic.[91] Consider the following cases: a compact steel sphere that presents thousands of nested concentric steel spheres with outer layers one millimetre smaller in diameter than the immediately larger, all placed inside one another, and a cat with twenty whiskers. Now, ask yourself if there are other nineteen cats co-located with it, each lacking just a single whisker with respect to the previously considered cat. A similar question can be raised concerning the nesting spheres inside the largest sphere. She asks the reader, then, to consider whether there is one sphere or a thousand spheres in the latter scenario and whether there is one cat or twenty cats in the former. She claims that a sane answer would be saying that there is only one sphere and only one cat, i.e., they are all the same individuals despite *differences* being provided by the outer layers and the surplus whiskers, respectively. In both examples, one can find criteria for the non-identity of these *individuals,* yet most people would say that this is unacceptable. She claims that, although no one knows how to account for their identity, we all accept it. Thus, their identity seems to be as gratuitous as the non-identity of the twenty spheres case proposed by Della Rocca. Her point is that if we can sustain the

[91] The framing of the discussion I am making here is slightly different of the original debate. Nonetheless, I believe it is more enlightening than the original discussion. Jeshion talks only of non-identities, but what is really in question is the gratuitousness of the identity in cases where there is an explanation for non-identity. As an additional reason for this framing is the fact that in (2008) Della Rocca does not give Jeshion's point the most charitable reading.

numerical identity of such cases even though it would be gratuitous, then, it is not the gratuitousness that is making the non-identity of the twenty co-located spheres case unacceptable, but something else. This something else, she hints, is the primitive identity and non-identity facts.

From an ontological point of view, the correct thing to do is to accept this internally sectioning cases as different *particulars*, indeed. But, from a pragmatical point of view, we should treat them as the same *individual*. The smaller concentric spheres inside the larger sphere and the whiskers-lacking cats are *sets of parts* of the original ones, which can be understood as sets of instantiated universals. Therefore, they are not under the same counting standard as the object that counts as an individual.[92] What we commonly understand as the counting standard here is the *maximal aggregate*,[93] which I coarsely define here as the largest slice of the object, or, more precisely, as the slice of the object that only has boundaries with other entities of the same kind from the inside, or, alternatively, the object that only shares external boundaries with entities of other sorts. The nineteen whiskers cat "inside" the twenty whiskers cat have a whiskery boundary with the cat outside of it, whilst the sphere (2r – 1mm) shares all its external borders with the 2r sphere. That is, there are multiple entities, indeed, but there is only one maximal aggregate, which is the relevant counting standard in everyday (and scientific, if so) discourse. Thus, there is only one individual in the scene but not only one particular. Furthermore, it is worth stressing that the entities in her examples are finite and differentiable. At some point there will be no more sub-entities to subtract parts from, which at least, presents us with metaphysical limits to search for epistemic differences. On the other hand, the entities in the Ockhamnian nightmare are, in principle, infinitely multipliable and – apparently – impossible to present epistemically detectable differences. Therefore, the double standard for the treatment of identity and non-identity based on gratuitousness that Jeshion claims to exist, does not exist. Those are different phenomena, thus, require different accounts. The identity

[92] This is a thorny issue. For a deeper discussion on counting standards and cross-count identity related with discernibility (and indiscernibility) see BAXTER (2014). The following account is inspired by Baxter's questions but is my own. It might not be a satisfactory answer to end the debate, but it points to the correct direction and is enough to show that there is an explanation involved, thus, it is not gratuitous.

[93] She points out something of the sort, in endnote 15, and claims that her objection can easily dodge this kind of issue. However, she does not elaborate it further. Thus, I believe the following account still stands.

cases she presented are explicable, as I just argued, whereas the non-identity presented by Della Rocca are – so far – inexplicable. Finally, any position one adopts here is anything but gratuitous.

The third way out for the PII opponent is to appeal to the primitiveness of non-identity of spatial locations, instead of primitiveness of objects' non-identity. Della Rocca poses his challenge in terms of objects. In these terms, Jeshion believes that the challenge is a considerable threat to the PII opponent, but if the challenge is translated into spatial terms, it can no longer be a threat. Della Rocca himself claims that it is acceptable to present the dispersal argument with other scenarios as counterexamples in P2), such as the ones using minds and space *regions*.

If primitive individuation of things is unacceptable, the primitive individuation of space will be unacceptable too, since space is a thing, right? Not according to Jeshion. For her, space is not an ordinary thing and, thus, should not be treated as an ordinary object. It must operate with different rules, which must allow primitive individuation. According to her, the understanding of space as a different thing would allow us to see how misbegotten Della Rocca's challenge is conceptually. It would show that there are primitive metaphysical facts about identity and non-identity, which would rule out the case of twenty overlapping spatial regions, which in turn could be used to rule out the twenty spheres case.

Her point hinges on a couple of distinctions, namely, the distinction between *primitiveness about objects* (PAO) and *primitiveness about locations* (PAL); and between *spatial locations* and *spatial regions*. PAO says that identity conditions for objects are not reducible to anything more basic, not even qualitative properties. Whilst PAL says that identity conditions for objects are reducible to space locations (relative to a time), which in turn are not reducible to anything, not even qualitative properties (see 2006, p. 171). Notice that PAL is roughly equivalent to saying that space is substantival. On the other hand, PAO can be assumed to be roughly equivalent to saying that space is relational. The other distinction concerns portions of space. A *spatial location* is the equivalent of a unit of space, it is a point. Space is comprised of all spatial locations. Whereas a *spatial region* is the sum of space locations comprised within borders, which are sets of locations that are parts of the region – in the same spirit that lines that mark the end of a sports court are also part of the playing court – (*ibid.*, p. 172). Differently from spatial locations, spatial regions are not primitives. Thus, two regions,

say, *x* and *y* can overlap if they share their parts, namely, spatial locations. With this assumption, Della Rocca's challenge concerning twenty co-located objects, can be translated into regions of space: if we can have partially overlapping regions of space, why cannot we have twenty completely overlapping regions of space?

Jeshion claims that, although it is impossible for twenty objects to be co-located, it is not incoherent to think about this situation, in other words, it is conceivable. It is so, for we can imagine twenty qualitatively indiscernible objects in different locations and then unite them in a single location. Thus, impossible, but not incoherent. Whereas it is incoherent to express a situation like this concerning space regions, in other words, it is unimaginable (*ibid.*, p. 173) – in my words, it is *inconceivable*. One cannot imagine that twenty regions of space are at the same place. For the sake of simplicity, imagine a region that contains only one location, viz., one point. To say that twenty locations are contained in one location is absurd! Jeshion points out that Euclid's first axiom, the one that says that between two points there is only one straight line that crosses both, is evidence of that. If the twenty overlapping locations exist, infinite straight lines could cross them.[94] Her point is basically that if two regions completely overlap sharing all their locations this means that they are co-extensional, viz., they must be the same region. In other words, they are not two things at the same place but the same thing, which is a single place. To think that there are two things there is conceptually incoherent. The very nature of space makes it not only impossible, but incoherent. Such scenario is inconceivable. On the other hand, it is perfectly conceivable that two regions of space that share no locations with each other exist separately. And, of course, we cannot distinguish these regions without appealing to their names, or the coordinates (which are name-like tags) of their boundary locations – she assumes.

She believes that Della Rocca would not be content with this and would inquire for a further reason as to why the boundaries uniquely delimit one and only one space region. Which is exactly what he does (2008, p. 35-6). But she claims that no more explanations should be

[94] In 2006, p. 173, she claims that twenty lines could pass through these regions that shares the same locations (i.e., points), though this is not quite correct. Maybe because she is not considering single location regions, thus, thinking about two different locations within the regions that could be crossed twenty times. My example seems more radical, given that these single location regions can be crossed from virtually infinite different angles.

required and insisting on such explanations would be tantamount to insisting dogmatically on the truth of a principle that says that non-identity is not primitive (be it PII or something else), which is what Della Rocca should be arguing for. And I am inclined to agree with her. If she has shown that to think otherwise is inconceivable, I agree that further inquiring for support of the claim that twenty co-located in space are impossible is unrequired and I would add: ineffable. But has she shown that?

This is certainly the most interesting of Jeshion's responses, for it deals with central points on the matter, nonetheless it also fails to work against PII. The first problem one might raise is that she is framing her reply in conceptual terms (which, as we saw on her first response, is not adequate, for concepts might change), when instead she could be doing it in terms of conceivability. I took the liberty to reconstruct her point in terms of conceivability and inconceivability here, so this problem is fixed. It is inconceivable to think of twenty spheres occupying one single location.

However, contrary to what she said, it is also inconceivable to think of the same situation in terms of objects. If it were possible to do it, this would disobey the Principle of Dissimilarity of the Diverse, which is not in trial. Also, a situation of fusion of twenty distinct indiscernible objects previously conceived in different regions into a single region would breed inconsistencies akin to the ones discussed in Hawley's response to O'Leary-Hawthorne: These objects are formed by smaller constituents; if they were co-located, would they share the same constituents, or would their constituents fuse? If they fuse, they would be something else from those twenty previously conceived objects. E.g., twenty distinct spheres that present all the same properties, including constituent parts and mass. The parts that occupy the nuclei of the spheres, when fuse will occupy same locations? Twenty particles will occupy the same location? Let us say each sphere weighs 1kg each, when they overlap would they also present 1kg or 20 kg? If they share the same constituents, then they must be the same object, perhaps with different names. Her argument against the capacity of conceiving co-located space regions must cut for co-located objects too. This is more or less the same objection raised by Hawley.

She claims that twenty spheres' cases and the Black-like cases of supposedly indiscernibles are different. The former present qualitatively insignificant duplication, viz., no matter how much entities we ascribe to the scenario, no difference can be found. In other words, concerning the qualities of the scenario as a whole, nothing changes; On the other hand, the

latter presents qualitatively significant duplication, i.e., the quantity of entities entails differences in the qualities of the scenario as a whole (see 2009, pp. 103 ff.).

When qualitatively insignificant duplication happens, we are better off assuming that there is only one object described in the scenario, due to quantitative parsimony, a theoretical virtue we should assume. However, as Hawley reminds us, this virtue must be applied alongside with other virtues, such as explanatory power, simplicity, elegance, and so on, as to avoid one to become an ontological monist. In the twenty spheres case we gain nothing violating the maxim of quantitative parsimony. Nothing is better explained in scenarios like that, on the contrary, inconsistencies might surface with closer inspection. Thus, we would lose in ontological and explanatory simplicity.

Nevertheless, Black-like (and Quantum Mechanics) cases, are different according to Hawley. These cases involve qualitatively significant duplication, i.e., the number of entities involved in the scenario causes a difference in the scenario, as can be seen in the examples presented above (in QM cases, we can point to mass and charge, if it is the case that these entities present no shape or other spatial properties). More precisely, by comparing the properties of the individuals alone and the same properties of the system, one will see that there are empirically detectable differences, thereby, the number of entities considered in the system plays a role in the property ascription for the system. Hawley calls these properties *additive properties*.

When qualitatively significant duplication is involved, which is the case for Black-like scenarios (and QM cases), things in such scenario might be better explained or more elegantly described sometimes if we assume more entities in the scenario. Thus, qualitative parsimony can justifiably be dropped *if needed* for the sake of a better explanation. The opponent of PII can claim that this is the case in QM and Black-like scenarios. In this case, a PII defender of the Dellaroccan strain would have to drop his weapons and leave the game, according to Hawley. However, as she remarks, this player must not feel dishearten, for she has a better defence, which we shall analyse in section 3.3.

From my point of view, this is a misinterpretation of what is happening. Again, Della Rocca's point was not to defend PII from co-located counterexamples, but to undermine the basis of counterexamples, namely, primitive (indiscernible) non-identity. He succeeded in doing so, for he shows that it is a difference of parthood between objects or regions of space

that sponsors the distinctness (non-identity) between them. These parts can be understood as the features possessed by the objects in phenomenological terms (also in relational spatial terms) or in terms of space locations (points) in substantivalist spatial terms. If one maintains explanatory uniformity, one shall see that, in Black-like cases, the spatial features are not the same in both spheres, which is even more evident when one thinks in substantivalist terms, viz., the spheres occupy different regions. In other words, they present distinct sets of locations (more on this topic will be said later). Thus, Hawley's additive properties and Jeshion's appeal to conceptual incoherence are, in fact, evidence for a defence of PII.

More objections concerning Jeshion's conception of space on her third point could be raised. For example, she explicitly says that her point relies on the nature of space itself (2006, p. 173). However, her point assumes that the nature of space is Euclidean, which might not be the case. If space is non-Euclidean, let us say, it is spherical, then, there are infinite lines crossing the polar regions of it. Her point for the conceptual incoherence of twenty co-located spheres would have to change. Yet, one might, as I have made in Hacking's case, argue that it does not matter whether space is Euclidean or non-Euclidean, but whether it is possible that it was.

Another point that could be raised is that if one adopts PAO (equivalent to relationism), which seems to be the option more suited to our contemporary relativistic conception of space, regions of space would have to be treated as feature, as *F*s, instead of *x*s and *y*s. Thus, being out of the scope of *objects* PII applies to. However, Jeshion choose to take PAL (which is equivalent to substantivalism) as her preferred conception of space. This makes it really difficult not to see space as an object, as portions of a substance, contrary to what she claims. One evidence for that is that she is following Casati and Varzi (1999), who explicitly says that spatial thinking is thinking about spatial things;[95] another evidence is the mereological treatment that she gives to spatial regions. Thus, although she is not treating space as a material object, it certainly

[95] In chapter 2, I have made clear that I am assuming "things" as Casati and Varzi put, as synonymous with "objects" in this thesis. What she seems to be doing is applying "objects" only to material objects. In any case, PII applies to a larger class of things that include space. Casati and Varzi reserve "objects" for "material objects" (see 1999, p. 9)

classifies as an object in the sense I am assuming in this thesis and, thereby, it is subject to the same treatment as other objects.

Nevertheless, one could raise the objection that although regions of space can be given the same treatment given to objects, space locations cannot. They are space points that present no distinguishable property from each other. Their individuation can only be done by naming/tagging, which, as we previously saw, cannot be used for discerning objects on the penalty of trivialization. Therefore, they are indiscernible. This is not quite correct either. They can be distinguished and individuated relationally as Muller demonstrates in (2015, pp. 225-6). The fact that they are ontologically prior to other entities (material objects, properties, etc.) does not imply that spatial relations do not supervene on (perhaps, emerge from) them. We shall see how this works in the next section where Muller's discerning defence will be presented.

Speaking of the devil, the last criticism to Della Rocca's challenge I will entertain in this section is a mild one presented by Muller. He points out that the twenty spheres scenario from Della Rocca's challenge is an arguably nomic impossibility. Nevertheless, he is not convinced that it is also a metaphysical impossibility, for it certainly is a logical possibility. He claims that Della Rocca's argument is a *reductio ad absurdum*, not a *reductio ad contraditionem*, which readers might interpret as saying that it gives a good reason to accept PII based on the nomological absurdity of this scenario, yet it does not force one to accept it, on pain of a contradiction. If the reader requires for a contradiction, this would not be enough (see *ibid.*, footnote 18, p. 217). Therefore, although this might be a suited defence for an unnecessitated version of PII, it might not be the best defence if one is dealing with possible worlds that have different physics than ours.

In chapter 2, I provided good reasons to believe that what PII opponents meant when they wrote "would it not be logically possible for a scenario with only two spheres…?" was rather whether it was not metaphysically possible that there is a world contradicting the metaphysical principle of PII. Also, I argued that matters of logical possibility or impossibility were so wide that they could not be what was at stake in this debate. Thus, to look for a logical contradiction might not be much more elucidative than to settle for an absurdity. Given the understanding of metaphysical impossibility proposed in that very same chapter, Della Rocca's twenty spheres scenario configures a metaphysical impossibility, which is an absurdity. Hence, I side with Della Rocca and against Muller on this point.

On the other hand, I agree with Muller when he points out that Della Rocca's view on the nature of spatial properties is poor. Della Rocca seems to be a relationist about space tying the identities of space locations to the objects that occupy them, e.g., region b *is the space region occupied by sphere b*, whereas sphere b *is in region b*. This would prevent one from discerning one sphere from the other by means of their spatial properties, because these properties are grounded in the spheres themselves. Circular reasoning. Nevertheless, these are not the only spatial features available to differentiate the spheres. Della Rocca is thinking about spatial properties only as extrinsic absolute properties of the spheres and ignoring relational features such as distance. Muller will right from the start explore this feature to establish the discernibility between the spheres reaching Della Rocca's goal quicker.

Nevertheless, I disagree with Muller when he says that Della Rocca oversees step 2a, presumably because he believes Della Rocca does not deal directly with the possibility of posing the challenge with space as the subject of identity and indiscernibility, but only talks of space as features of the spheres (this criticism was also veiled in Jeshion, by the way). Of course, Della Rocca does this. He focuses on the spheres, because they are the objects in question in Black-like scenarios. However, he explicitly considers the possibility of individuating spatial regions as objects, even though he does this to use them as features of the spheres later.

3.2.2. Naturally discerning: Muller's discerning defence

The discerning defence is the most intuitive defence for PII. Here I will present Muller's (2015), for it was the most complete and the most discussed I could find in the literature. Most people that think PII is intuitively true first reach out for a discerning defence using some spatial feature as the difference between the objects in question in a dispersal argument's scenario. Probably because of the intuitiveness of WLP. The most intuitive spatial property available is the location of the objects, i.e., the predicate that gives us the region of space it occupies. Nonetheless, it is not a reliable property, because it requires naming/tagging, which is not allowed for discerning. Thus, Muller's version of the discerning defence appears as a superior alternative because it reaches for the *distance relation* between the objects as the relevant discerning feature.

Muller describes the thought experiment scenario in a Euclidean space (E^3) with only two spheres that share every absolutely discerning property (i.e., those that are expressed by monadic predicates, e.g., intrinsic properties and monadic extrinsic properties). He calls these spheres Castor and Pollux for sake of convenience in completion of step 1 (see section 1.3). [96] Exactly like Black's own setting. Then, starting step 2, he asks us to consider the *features*[97] *of the scenario* instead of features of the spheres when looking for discerning properties. Given that the absolutely discerning properties are shared by both spheres, he asks the reader to consider the relatively discerning properties that exist in the context where the spheres exist. In this case, given that there is nothing else in the scenario, the only relations bore by the objects that could be used as relatively discerning features are spatial relations, namely, the distances they bear with themselves or with each other.

Just to be clear, there are other discerning features available, but they are not allowed to discern on pain of triviality. As examples of non-permitted discerning features Muller presents N(*a*), namely, the property that says that "*a* = Castor", the property that says that *a* is identified by the name "Castor". Were this property permitted in the experiment, the spheres would be absolutely discernible, for N(Castor) and ¬N(Pollux). However, N is an identity property, it has a name as its constituent. So, N is basically the utmost trivializing property! Its use is absolutely prohibited. Another kind of discerning feature not permitted to the defendant is that of spatial properties established by coordinates, which amounts to a different way of describing the space of the scenario within a relational framework. Muller reminds us that since E^3 (Euclidean space) implies R^3 (real coordinate space), spatial locations can be given by a set of coordinates (x, y, z) of values in the axes x, y, and z. One could conceive Castor located at the origin *(0, 0, 0)*, i.e., the location crossed by every axis; and Pollux crossed only by the Y-axis, lying in (0, 2, 0). Considering the following predicates O(*a*) and Y(*a*), respectively meaning *sphere* a *lies in the origin* (0, 0, 0) and *sphere* a *lies on the Y-axis*, the spheres would be absolutely discernibles,

[96] He follows Black (1952), who permits that an observer baptizes the spheres and exits the scenario. This is unnecessary. Given that the names are being used as bounded variables, not as names, we could just keep using variables. However, he claims that they are ineliminable because one cannot have a definite description of each sphere, otherwise they would be absolutely discernible (MULLER, 2015, p.213). I disagree. One could obtain the same results using terms such as "one" and "the other" or "another".

[97] He does not use the term "features", I am adapting to the jargon I am proposing.

because the scenario would present the following distribution of properties: O(Castor), ¬O(Pollux), Y(Castor) and ¬Y(Pollux). However, according to Muller, not every spatial feature available can be a permitted discerning feature. He claims that the use of these properties is unacceptable for two reasons. First, because these triples of numbers, when put together forming a coordinate, work as names or tags arbitrarily assigned. Second, the use of these coordinates would break the symmetry of the scenario. None of these things are permitted. Nonetheless, both these objections are questionable, and I shall come back to them in chapter 4.

After carefully searching, the reader will conclude that the only permitted discerning feature ends up being the distance between the spheres. Muller, then, asks us to consider "*a*", "*b*" and "*c*" as sphere-variables having as range the set {Castor, Pollux} in the scenario so that no naming takes place, and defines distance D(*a*, *b*) precisely in this context as the spatial relation between two spheres, if and only if, sphere a *is 2 miles apart from* sphere b (*ibid.*, p. 210). Being so, in the description of the scenario we have the following relations:

D(Castor, Pollux), D(Pollux, Castor), ¬D(Castor, Castor), ¬D(Pollux, Pollux)

As one can see, this relation held by the spheres is symmetric and irreflexive, so, it is weakly discerning (see section 2.2); and grounded in the structure of space (E^3), according to Muller. Thus, it is not a feature that break any of the constraints previously established (see section 2.4.4). We can say, then, according to Muller, that the spheres in this scenario are absolutely indiscernible, but relationally discernible. More specifically, they are weakly discernible. This means that they are not indiscernibles, but another kind of entity, which Muller calls relational. Black's spheres are relationals.

Summing up Muller's point: From a consistent qualitative description of a scenario with quantitative diversity (a feature of the scenario assumed by the proponent of the thought experiment), we can find qualitative diversity of a permitted kind that can be used to discern the objects said to be indiscernibles (the weakly discernible feature of distance). Thereby, showing that they are not indiscernibles, but relationally discernible, i.e., they are relationals. Therefore, this scenario cannot be used to instantiate P2) in the dispersal argument. In other words, PII is safe.

3.2.2. Objections to Muller's discerning defence

PII discontents still have a card in their hands, though. Muller claims that French and Krause (2006), and Hawley (2009) accuse such defence of being circular, because it *only* shows that the spheres are different (viz., distinct), which is the very fact that is assumed to begin the description of the scenario, i.e., that there are two spheres. According to Muller (*op. cit.*, p.210-1), French, Krause and Hawley understand the core of the discerning argument as something like:

P1) All sides agree that the question whether there is quantitative identity or diversity is the same as whether spheres are distinct or identical with each other, which in turn is the same as whether there is one or there are two spheres in the scenario.

(Numerical diversity/distinctness definition – accepted)

P2) Opponents of PII claim that there is a possible world where there are two, and only two, distinct objects, that is, Black's scenario with Castor and Pollux.

(Black's scenario – accepted)

P3) Castor and Pollux are distinct.

(Established *ab initio* – accepted)[98]

P4) If Castor and Pollux are distinct, then they must present some feature that distinguishes them.

(Dissimilarity of the diverse (equivalent to PII), claimed by the defender)

---

[98] If this was not established, there would be no way to say that there were two objects in the scenario, claims Muller (*ibid.* p. 210), because if "Castor = Pollux" were the case, then, there would be only one object with two names; and if it were not established whether one or the other is the case it would be impossible to deduce whether there were two objects in the scenario. I would add that in the last situation we would find ourselves in Della Rocca's challenge, i.e., there would be no way to say how many objects are there. In either case, Muller reminds us, an argument against PII could not take off.

C1) There is a difference between Castor and Pollux, namely, distance (D).
(From P3) and P4) by *modus ponens*)

P5) If there is a feature (namely, D) distinguishing Castor and Pollux, then Castor and Pollux are distinct.
(Diversity of the dissimilars)

C2) Therefore, Castor is Distinct from Pollux.
(From C1 and P5 by *modus ponens*)

Thus, the discerning defendant would be concluding something already known from the start, namely, that Castor and Pollux are distinct. There is no need for arguments from discernibility to reach C2). (That is not exactly what the accusers are claiming, however. We shall come back to this point latter.)

3.2.2.1. Muller's replies

Muller claims that the opponents of the discerning defence confuse epistemological questions with metaphysical ones. He claims that this circularity charge confuses what the precise target of the argument in the discerning defence is, because it mixes up three questions involved in setting up the scenario for the dispersal argument. There is an epistemic and two metaphysical questions. The first metaphysical question is:

(Q1) In this qualitative arrangement (Black's scenario), is there one object or there are more objects?

It basically asks whether there is quantitative identity or diversity. The epistemic question is:

(Q2) How can we find out whether in this qualitative arrangement there is quantitative identity or diversity?

Which simply asks about how to answer to confirm the answer for (Q1). Which may lead to (Q3).

> (Q3) Given that there is quantitative diversity in this qualitative arrangement, is there also qualitative diversity?

The second metaphysical question, which enquires about the presence of features. (These questions can be found in MULLER, 2015, p. 209, written slightly differently).

Muller answers (Q2) rather simply: just read the description of the scenario! The description of the thought experiment says explicitly that there are two spheres (or droplets, or whatever). Although this might not be the only way to conceive the scenario[99] and one might come up with steelman versions of Black's scenario, we will leave this issue aside in order to focus on the answers to the other questions, since we must answer (Q2) to answer (Q1). The latter, as we have already established, is answered in favour of quantitative diversity, i.e., it is in the description of the scenario that there are two spheres disposed in such and such ways. Whereas the question Muller tries to answer with his argument is (Q3), a question that requires (Q1) to be answered positively in order to be meaningfully made. The answer to (Q3) is about features, not about numerosity, in other words, it is about quality, not quantity.

In Muller's understanding, the circularity charge is that his argument is answering (Q1), already assuming it as a premiss. But this is not the case, he assures. His argument is answering (Q3). He is not merely concluding that there is numerical diversity. After all, there is no need to. It is granted from the set up: the argument against PII establishes that there are two objects in the scenario. Otherwise, there would be no argument against PII, for the scenario would not present distinct supposedly indiscernibles. The proponent of the dispersal argument is the one who should be worried about making clear that there are two objects in his scenario. Muller, the defendant, is worried about stating *why* there is numerical diversity. By doing so, something

[99] One might think that a description without the number of objects could be presented such as Hacking's cylindrical world, but this would lead to objections we have already discussed in section 3.1.1.1.

I believe worth stressing is that he is also giving another answer to (Q2). He is presenting another description – or at least directing us to another part of the total description making this world – that is not concerned with quantitative identity or diversity, but with the qualitative features of the scenario that allow us to infer quantitative diversity by means of weak discernibility (Remember: weak discernibility implies distinctness, see section 2.2). In other words, he is not arguing for the distinctness of the spheres, he is rather *showing* that, when they are distinct, they are also discernible in the scenarios the opponent drew. More than that, he is showing *how* they are discernible, i.e., through relational property D, namely, distance.[100]

Thus, the discerning defendant's argument's core really is:

P1) All sides agree that the question whether there is quantitative identity or diversity is the same as whether one sphere is distinct or identical with the other, which in turn is the same as whether there is one or there are two spheres in the scenario.

(Numerical diversity/distinctness definition – accepted)

P2) Opponents of PII claim that there is a possible world where there are two, and only two, distinct objects, namely, Black's scenario with Castor and Pollux.

(Black's scenario – accepted)

P3) Castor and Pollux are distinct.

(Established by the accusers, accepted by defendants)[101]

P4) If Castor and Pollux are distinct, then they must present some feature that distinguishes them.

---

[100] In other cases, such as QM and mathematical cases, other relational properties are summoned, such as spins, edges and scale relations to zero.

[101] If this was not established, there would be no way to say that there were two objects in the scenario, claims Muller (*op. cit.*, p. 210), because if "Castor = Pollux" were the case, then, there would be only one object with two names; on the other hand, if it were not established whether one or the other is the case it would be impossible to deduce whether there were one or two objects in the scenario. I would add that in the last situation we would find ourselves in Della Rocca's challenge, i.e., there would be no way to say how many objects are there. In any case, Muller's point is that an argument against PII could not take off.

(Dissimilarity of the diverse (logically equivalent to PII, but not PII))

C1/P5) There must be a distinguishing feature between Castor and Pollux. (Hypothesis from P3 and P4 by *modus ponens*. Remember that Distinctness is forbidden to discern!)

P6) There is a difference between Castor and Pollux, namely, distance (D).
(Demonstrated through the analysis of spatial features of the described arrangement (E^3 space))

P7) The distinguishing feature D is a permitted (non-trivializing) weakly discerning feature.
(Lemma supporting P6 – found through the analysis of permitted features (step 2b and 2c of building the dispersal argument))

P8) Weak Discernibility implies distinctness/numerical diversity.
(accepted rule (see section 2.2))

C2) There is a permitted distinguishing feature between Castor and Pollux, namely, distance (D); thereby, they are distinct.
(Conjunction of P6, P7 and P8)

With this argument, Muller is not concluding that (i) Castor and Pollux are distinct, something already established by the description of the scenario and accepted by the defendant. Concluding this would be answering (Q1), which is not the question he was aiming at, and which has to be answered affirmatively so that (Q3) – the answer he is aiming at – can answered. (Q1) is answered affirmatively by P1) and reaffirmed in P3). Thus, the answer to (Q3) can be pursued.

The above argument shows that (Q3) must be answered affirmatively too. He concludes that (ii) there is also qualitative diversity in the scenario described by the proponents of the dispersal argument. More than that, he shows that (iii) there is a permitted feature that discerns the spheres, that (iv) the feature is of very special kind, namely, a weakly discerning relation,

and that (v) exactly which feature it is, namely, the distance (D). Even more than that, he concludes that (vi) the distinctness between the spheres is due to this relation: "(…) *we begin with two distinct objects and we end with their distinctness being grounded by a permitted and weakly discerning relation:* that *was not* assumed, *but rather* demonstrated." (MULLER, 2015, p. 213. Original emphasis).

There are at least five conclusions that can be drawn from the discerning defence that were not previously known, thus the charge that this defence is a *petitio principii* cannot be correct, as Muller himself reminds us: "*when the conclusion is not the same as the premiss, the connecting argument cannot be a* petitio principii." (MULLER, 2015, p. 213. Original emphasis). Fair enough. However, Krause, French and Hawley are not charging him exactly with a *petitio principii*. Their issue is not with the conclusion, but with some corollaries within the defendant's argument used as premisses in the core argument presented above.

3.2.2.3. The charge from another point of view

The reconstruction of the circularity charge by Muller is not exactly what the accusers are claiming.[102] What they are claiming is not that the conclusion is contained in the premisses of the discerning defence, but that some corollaries used as premisses in the core argument above require the conclusion of the core argument to be put to use. French and Krause's point is that individuality (thereby, distinctness also) must be secured before distinguishability comes into play. This is done by P4). Notwithstanding, PII is aiming to establish individuality through discernibility (2006, pp. 170-1), which would be redundant. Of course, we must assume that no party in this debate is taking discernibility as something different than distinguishability (though Saunders (2006) hints at a useful distinction of use in the context of QM, this distinction is not adopted in the rest of the literature).

[102] To be fair, French and Krause's argument is quite different from what Muller exposed, but Muller's formulation is a possible reading given that they are very brief in their exposition of the circularity charge. In their book they are dealing with the principle in a very specific context of QM, in which it is not clear whether there is one or two or many objects in question. So, one of the things they were discussing in those pages was how to establish that there are two objects. This might have directed Muller to this interpretation.

Hawley repeats this point in a slightly different and clearer way in her later work (2009, pp. 108-12). She objects that the feature used to discern one sphere from the other requires previously the identity (thus, also distinctness) of the spheres involved in it, namely, Castor and Pollux, and the criteria of identity for the relation *being 2km from*. She claims that the weakly discerning property Castor has that discerns it from Pollux is *being 2km from Pollux* – and vice-versa. Which, in turn, is grounded in the dyadic relation *being 2km from* and in the identity of Pollux (and the identity of Castor for Pollux). But then, D cannot be said to ground the distinctness of Castor and Pollux alone, precisely because it depends on their identities and distinctness to be applied.[103] If the discerning feature between Castor and Pollux is dependent on their very identities, then she claims that the discerning defence cannot be used to ground identity. Given that one of the main reasons to defend PII is the grounds of the identity notion, for her, this identity is to be avoided in favour of the summing defence.

In other words, both French and Krause's, and Hawley's points can be summed in our reconstruction of Muller's argument as: for the defendant to be able to claim that the distinguishing relations D(C,P) and D(P,C) are grounds for the distinctness of Castor and Pollux (Muller's conclusion C2), one must *know* previously that both objects are distinct P3), which in turn is grounded in P5/C1), that was granted by a *modus ponens* using P3), which requires the truth of C1/P5), and so on. The reader must have gotten the point by now, but to make sure, here is another formulation: *to imply* that the spheres are distinct *because* they are distinguishable (discernible), is circular precisely because *to imply* that they are distinguishable, one had to assume their distinctness.

Muller avoids this criticism in two ways. First, by arguing that the distinctness between the spheres is settled by the opponents of PII. Without this assumption there would have no case against PII. If the description of the counterexample does not say that there are two spheres, i.e., that Castor is numerically distinct than Pollux, one of two things could be the case: there

---

[103] Hawley focuses her point on the property *being 2 miles from b* which I adapted to Kilometres and to a dyadic relation. Both Hawley and French/Krause are dealing with QM cases using the discerning relation *has a different spin direction than* as the relevant feature. Nevertheless, it seems that their points can be translated into arguments using *distance* (D) as an equivalent feature and points made about D can be translated for QM cases. Hawley seems to assume this since her whole explanation is made in spatial terms although she makes clear that there are cases where no spatial feature is available, thus, I will assume this too.

would be only one sphere with two names, then there is no counterexample and PII is safe – this is what is brought by Muller (*op. cit.*, p. 210) – or the proponent of the scenario is faced with in Della Rocca's challenge and has to proof that there are not two hundred spheres instead of just two – this I am adding to Muller's point. Second, by arguing that the discerning feature is not the monadic extrinsic property presented by Hawley, but the dyadic relation of D that she claimed was grounding the property. Two reasons for that: One is that, if the monadic properties were doing the discerning, then the spheres would be extrinsically absolutely discernible. The other is that the very reason to appeal to a relational approach to things is the fact that monadic features (properties) were not properly accounting for the scenario, thus, her argument should be aimed at relation D, not at the monadic properties she used to discern the objects. Therefore, the grounding of identity facts by discerning relations is still on the menu.

Summing up Muller's reply: the identity/distinctness are given with the description of the scenario, but the description does not say in what this distinction is grounded. This is something that must be discovered. The metaphysical part played by D in grounding the distinctness of Castor and Pollux was always present, although it was not explicit in the description offered by the PII opponent. What Muller is displaying is the epistemic part, that is, the part where we can now *know why* the metaphysical part is the way it is. Remember: PII is a metaphysical principle with a great epistemic load. Moreover, Muller argues that this relation is not grounded in the identity of the spheres, but in the structure of space itself. Therefore, there is no circularity in the discerning defence. The identity and the distinctness of the spheres are grounded in the permitted relational feature D, which in turn is grounded in the very nature of space (assuming that it is E^3).

Certainly, this seems a compelling defence. Nevertheless, one further objection that I considered during those years is that the spatial feature of D is a property of E3, not of the spheres, and, thus, is not within the scope of PII. In other words, it is an F that is not applied to x or y, but to a z that contains x and y. Well, there are two shortcomings for this objection. First, the way in which we describe this feature being applied is *about* the objects, that is, the distance is held *between* the objects; in other words, it emerges from (or supervenes on) the objects, not from (or on) space. Actually – and here is the second shortcoming – it might as well be said that this distance *is* the space. If one holds a relatival conception of space (PAO), space (E3) must be understood as these distances that emerge (or supervene) from the objects. Thus, E3 is

grounded in the spheres ultimately. However, in that case, one might press further and say that because space itself is grounded in the distinctness (thus, somewhat also on the identity conditions) of the spheres, the very relation that discerns the spheres already presupposes that their distinctness is grounded in their previous distinction. Hawley, French and Krause have a point, then.

One way out might be abandoning the relational conception for a substantival approach. Notwithstanding, one could just change the objects of the counterexample to spacetime points instead of spheres and the discerning defendant would have nowhere to run then. Another strategy – and this is the one I adopt – is to bite the bullet and argue that this might not be a vicious circle, but a virtuous one. This might mean that we have reached rock bottom in grounding. We can see that distinction is indissociable from spatial differing and (strict) identity[104] is indissociable from spatial co-location because they simply are the same thing. We cannot even conceive counterexamples to this. Or at least this is what I will argue in the conceivability defence. Therefore, one might interpret the conceivability defence as a further step in the discerning defence. Furthermore, this could be used to explain WLP and rule out bi-location.

But before we finally go for the inconceivability defence, let us consider one last serious objection presented by proponents of the summing defence such as Saunders (2006) and Hawley (2006; 2009). It has been said that there are cases, such as those of entangled fermions and bosons, which are absolutely indiscernibles, to which a discerning defence might not be available because there are no discerning spatial features available to do the job. Nevertheless, Saunders claims that they could be weakly discerned by their spin. Thus, a discerning defence can still be applied, the defender only has to choose other kind of feature to do the job. However, he also claims that there still is an exception to this strategy, namely, elementary bosons (*ibid*., p.57-60). In this case, he and Hawley argue that a summing defence is needed.

---

[104] Although I am talking about a very particular kind of identity, one could adapt this answer to other kinds of identity. For example, identity over time can be thought of co-existence within the same location in space *and* time. Whenever two stages of the same thing are considered, they are temporally identical no more. Putative identities such as baptisms, can be thought as conjuring the co-location of a material object to a linguistic entity from now on. In other words, wherever and whenever this aggregate of mater is found, such and such name is to be found too.

Muller (*op. cit.*, sections 5 and 6) replies to this criticism by claiming that the unavailability of a discerning feature only comes up if one interprets these bosons within Quantum Field Theory (QFT), which has a different ontology than QM. In QFT bosons are not interpreted as particles but as modes of a quantum field (see SAUNDERS, 2006, p. 60). In any case, whatever theory turns out to be the correct theory, there seems to be an appropriate defence available. If QM is the theory applied, then a discerning defence is available; if QFT is in play, the defender can appeal to the summing defence.

3.3. Summing defences

Summing defences are those that claim that, in the counterexamples used as instances for P2) in the dispersal argument, "*there are not two (or more) objects* [in the scenario], *there is one object of a kind that is* different *from the kind that the alleged two (or more) objects belong to, and that one object has no parts* […] *it is a simple*" (MULLER, 2015, p. 205). The proponents of this defence are Saunders (2006) and Hawley (2006; 2009). Here we shall analyse just Hawley's version, because it is the most updated of them and its scope encompasses Saunders' version in it. After presenting her view, I shall present some criticisms raised by Muller and by myself.

3.3.1. Hawley's summing defence

As said before, the main reason to reach for a summing defence is the allegedly undiscerning objects of QM. The first version of it was presented by Saunders (2006) to deal with *some* counterexamples to PII in QM in which discerning features, such as mass, charge and spatial properties, were nowhere to be found in a symmetrical scenario (e.g., elementary bosons). Thus, he proposed that maybe some of the entities in question were better understood as non-individuals, that is, as entities that lack conditions to exist as independent objects. Another way to put it is that such entities cannot be understood as wholes, only as parts of something larger. This implies that they cannot figure as an x or an y in the PII formula, thereby, they cannot be under the scope of PII.

Saunders is a proponent of the discerning defence who champions the summing defence only when the former is not appliable. Hawley, on the other hand, argues for a uniformity of defences. She claims that the summing defence is available for Black-like cases, and in praise for theoretical virtues (viz., uniformity and parsimony) we should adopt it, even though the summing defence is also available. This is a relevant point to stress. She does not dismiss the discerning defence, she just champions the view that the summing defence has a wider range of explanation associated with other theoretical virtues than the discerning defence, thereby, it is preferable. Besides, she does not believe that a strategy such as Della Rocca's challenge can accomplish anything, given that there are reasons for one to accept significant duplication cases such as Black-like cases and there is no reason to accept insignificant duplication such as his co-located counterexamples. The reasons are the inexplicability of additive properties and the priority of uniformity when confronted with ontological simplicity (parsimony) involved in the former case which is not involved in the later.

Every PII defence we saw so far accepts some version of Black-like scenarios developed as counterexamples to PII. What they attempt to do is to reinterpret some aspects of what Black proposed in a way that PII and the described scenario can fit together either exploring the lack of maximality in Black's design of the thought experiment or changing one or other detail here or there. The summing defence, like the identity defence, tries to change the cardinality of the scenario. Hawley's main point is that we do not have to describe Black's scenario with two objects weighing 1kg and measuring 1L in spherical-shape each, we may say that there is just a single object with 2kg, 2L in a scattered two-sphere shape (*op. cit.* p. 111). The main difference between these defences is that in the identity defence, there was a single bi-located object instantiating twice sphere-shapeness, whereas in the summing defence there is a single scattered object that does not instantiate sphere-shapeness at all. In other words, the whole system of spheres – the sum of the two spheres – can be understood as an indissoluble object. In this case, according to her, the separated spheres *need not* count as objects.

According to Hawley, it is reasonable to believe that, in such scenarios, there is a larger object (e.g., the system of spheres) that is the sum of two smaller objects (e.g., the spheres). On the other hand, it is disputed in metaphysics whether some objects have proper parts. As support for the more controversial claim that some objects have no proper parts, she brings up debates within metaphysics that might direct our intuitions to this such as the Doctrine of Arbitrary

Undetached Parts, Inverse Special Composition Question (*ibid.* p. 111), then, asks the reader to consider two things: First, "*whether there are features of the larger object which are best explained by the existence and features of the smaller objects*", and second "*whether, considered separately, each of the smaller objects is a good candidate for existence*" reminding us that both considerations may overlap (*ibid.*, p. 111). She is appealing to theoretical virtues here, namely, explanatory power and ontological simplicity (parsimony).

The last consideration might sound strange, because one might think that if we are talking about the smaller component objects and they are playing some role in the existence of the larger object, these smaller objects must exist. Nevertheless, she draws a distinction between *objects* and *existing objects*. The latter is a sub-group of the former. In her terminology, theoretical entities are objects, but this does not mean that they are existing objects. Ideas, universals, minds, and all those almost-mystical metaphysical entities are considered objects but not necessarily existing objects, I suppose.[105] Thus, the PII we are interested in here would apply only to what she means by existing objects; other kinds of objects are not under the scope of PII.

According to Hawley, the existence of these objects would be justified only if they were to present clear identity conditions (viz., individuality), which they do not. Remember that these objects are intrinsically indiscernibles (absolute indiscernibles) and are only weakly discernible, i.e., their individuality depends on the relation held with each other. So, for her, they can only exist in comparison with each other. This means that, would one cease to exist, the other would follow. Thus, why should we believe that they exist as objects at all? She claims that, in the light of the first consideration, the answer to this question is that it makes more sense to think that the minimal existing object in the scenario is the whole system, which has clear identity conditions, to which the identity conditions of the smaller objects supervene.

On the other hand, the existence of the whole does not need explanation in terms of the smaller objects, that is, not all properties of the system are reducible to properties of the parts,

[105] It is not clear whether this distinction is equivalent to Quine's ontological entity/ideological entity although it seems to be. Therefore, I chose not to make this translation, but I it is a translation that may be worth considering. Minds, universals, ideas, etc. being part of our ideology but not of our ontology, thus being outside the scope of PII.

i.e., some relations are non-supervenient.[106] Of course, there might be qualities of the system that may be better explained by the parts, e.g., mass: the system has a 2kg mass, because it has two parts with 1kg each. The same goes for charge, shape, additive qualities in general. Hawley admits that there *may* be those cases where these qualities of the system are better explained by the properties of their parts. This would depend on how much of a reductionist one is though (*ibid.*, p. 112). Her point is that one does not need to do so. Perhaps, one should not do so, because this theory would lose in uniformity, given that it would need to present an ontology that has the smaller objects as minimal object to explain these properties, but would have to present another ontology, which has the system as a minimal object, to account for PII.

Additionally, in cases of QM in which the particles are co-located, the appeal to these properties bears very little explanatory power and are better explained assuming the system as the minimal object and the particles as putative parts than vice-versa. More precisely, Saunders (2006, p. 60) proposes – and Hawley endorses – that whenever in states that present no discerning features elementary bosons can be (and might be better) understood not as a pair of particles, but as modes of a single quantum field. The quantum field is the only object existing in that scenario and bosons of excitations in this field. The imagined cardinality of particles in the scenario are in fact levels of excitation in this field. Which makes these bosons better understood as *F*s rather as *x*s or *y*s. In other words, we must assume that these modes are not objects (given that they cannot sustain existence by themselves), therefore, they are not subject to PII. It seems that the summing defence is mandatory in this case.

In other words, Hawley's point is that these smaller objects are better understood as theoretical posits playing an explicative role in a theoretical model instead of really existent entities. According to her, the fact that they are theoretical objects does not imply that they are existing objects, objects with real existence. Thus, we can excuse our theory of the ontological commitments with the smaller objects, if this is the most virtuous thing to do. Then, she proposes that we must look for the best equilibrium between parsimony, explanatory power and theoretical uniformity. On the one hand, we can postulate the existence of the smaller parts as

[106] This involves a debate about the supervenience or non-supervenience of the properties of such particles that I avoided here. But Hawley does not develop the discussion further, she just assumes the spatiotemporal relations are non-supervenient (2006; 2009) and follows through.

objects and lose on parsimony and on uniformity by adopting a defence that works for Black-like cases but not for special QM cases and gain the power to explain why there are two objects in such scenario; or, on the other hand, gain in parsimony by committing to a single scattered object and in uniformity using just the summing defence for any presented counterexample to PII, but losing in explanatory power by biting the bullet of assuming that there is not an entire object occupying a spherical region of space in Black-like scenarios.

Hawley is not crystal clear about this, but it is worth steelman her point. It seems that it is not the case that the stuffing of those spherical regions within the Black-like scenarios are not explainable at all. They are. Her point seems to be that they are just unexplainable independently. One would still be able to explain the occupation of those regions as parts of a larger scattered region in two-sphere shape. Thus, one might still say that the summing defence does not imply less explanatory power, just less economy or theoretical simplicity.

In any case, it seems clear that the summing defence presents a very little price to pay for QM cases, since the objects in question must not be interpreted as scattered independent objects, but as co-located objects (or not in space at all). Nonetheless, in Black-like scenarios, it seems too high of a price to pay. Yet, she claims it is worth paying if one wishes to ground identity facts on discernibility facts, which she sees as one of the main reasons to defend PII and which is not accomplishable by the discerning defence – or so she believes, as we saw in the last section.

Finally, her views about the summing defence can be summarized through the following questions: must we postulate the existence of two smaller objects (e.g., each sphere, each boson) forming the larger one (e.g., the system of spheres, the bosonic system) in the scenario? She answers negatively. Thus, given that we do not have to do it, would we gain something by doing it? She answers positively: we gain uniformity and parsimony to our physical theory, and a non-circular way to ground identity facts. We must only give up the independent explicability of some spatial facts in Black-like scenarios.

3.3.2. Objections to Hawley's defence

Concerning Black-like cases, there is an evident problem to the summing defence and that Hawley is very aware of, namely, the existence of scattered *simples*. We do accept many

scattered objects in our ontology (e.g., archipelagos, institutions), however, these objects are divisible in smaller parts of matter that are somewhat attached to separated spatial regions. Scattered simples, on the other hand, are made of supposedly indivisible smaller parts of matter that are also somewhat attached to separated spatial regions. But there is no good reason to forbid the division of these simples into their parts, given that they are attached to the spatial regions they occupy. This conclusion could be derived from Adams' (1979) argument from almost indiscernible spheres. Remember: two almost indiscernible spheres exist occupying different spatial regions, if one of them were slightly different from how it is now and turned out to be intrinsically indiscernible from the other would it cease to exist or magically turn into a scattered simple with the other? None of these answers seems reasonable. The most reasonable answer would be that two spheres still exist but with no intrinsic difference.

Hawley admits that this is a problem for her view. She entertains the possibility of assuming that even in the almost-indiscernible-spheres scenario there is only one scattered object, however, she also admits that this would be the first step on the downhill road to monism, thereby, not an acceptable alternative. In such case, the discerning defence seems to be more appropriate, she concludes. But only if one is willing to give up grounding identity. Thus, she proceeds to conclude that even though the summing defender would have to accept scattered simples, it is a price worth paying in name of theoretical virtues and of the possibility to ground identity, as we saw previously.

Nevertheless, Muller (2015) replies that this price does not have to be paid at all. As we saw in section 3.2.2.3, he argued that Hawley's worries about adopting the discerning defence are unfounded since, to elaborate her circularity charge, Hawley used the wrong discerning features, namely, the extrinsic monadic spatial features of the spheres instead of relation D (distance), which is actually doing the job and is a permitted discerning relation. Therefore, the discerning defence can also be used to ground identity.

However, what about the QM cases? These would not suffer the problem of scattered simples, given that the entities in question are co-located objects (or lack spatial properties at all), but some of them, namely, elementary bosons, still present the problem of not displaying features that could be used to distinguish the objects in question. Thus, the discerning defence still cannot deal with all QM cases whereas the summing defence can. So, it seems that the

summing defence has a broader explanatory range than the discerning defence and should still be considered more virtuous than the discerning defence.

Muller claims that this might not be the case if we consider a larger scope of issues faced by the ontology implicit in the summing defence. The ontology required to use the summing defence presents two problems according to him. I shall skip the technical minutiae and just present one problem raised by him that might solve the whole quarrel (for the whole point see MULLER, 2015, pp. 229-31). When one conceives the thought experiment about QM particles, one has to explain how the measurements are performed. The way in which one might do it is by simultaneously measuring two entangled intrinsically indistinguishable particles with two spatially separated apparatus. If done this way, this implies that these particles can be found in two separated portions of space generating a Black-like scenario. From this point onwards one can invoke relation D and use a discerning defence. So, it seems that the discerning defence is available for QM cases, after all. One just needed a little bit more of imagination to look for discerning properties. Unless one can find another method of conceiving this thought experiment. In any case, the *onus* is with the proponent of the thought experiment.

It turns out that the discerning defence has the upper hand concerning the theoretical virtues, then. A quick recap: the explanatory power of the summing defence seemed to be wider, given that it explained every case the discerning defence did and the cases of elementary bosons. A mixed strategy (proposed by Saunders) would present the same explanatory reach, but it would lose in explanatory uniformity, for it would require two different strategies. Additionally, both would lose in ontological parsimony to the summing defence. This is why Hawley thought that the summing defence is preferable. Muller has shown that, although the discerning defence still loses in parsimony (it would still have to claim three objects, the spheres and the pair of them, to work), it presents, at least, the same explanatory power, for it also takes care of the elementary bosons within QM. In fact, the explanatory power might be wider, given that the summing defender might not be able to explain the mereology and some mathematics present at the core of QM and in the phenomena captured in the conceived scenario added to the blatant case of scattered simples in Black's scenario. If the summing defender manages to explain those things, then it certainly would lose in economy, for the gymnastics required to explain them would certainly use a lot of ink and presumably would require non-uniformity in other areas. But more importantly than this, it would lose in coherence and conservativeness, for it would

break a continuity with the previously gathered knowledge (2015, p. 229).[107] Muller considers coherence and conservativeness as more relevant virtues than parsimony, though this is hard to decide. In the aftermath, then, it seems that the preference for the summing defence is unjustified.

Furthermore, another problem with the summing defence in quantum cases noticed by Muller is that Saunders moves from an initial description of the scenario within QM to a description within QFT, which is a different theory with a different ontology than the former, to present a solution to the problem. Some might consider this mixing of ontologies as cheating, unless one proposes that QM is wrong and QFT is the correct theory to interpret reality. Though, it might turn out that QFT is indeed the correct physical theory or, at least, its ontology is the most appropriate to deal with the quantum world. In that case, one might still insist that the cases of bosons may be a genuine problem for the discerning defence. In this case, Muller concedes that we should adopt Saunders' position of a mixed defence (2015, pp. 232-3). I do not see any problem with this suggestion. Given that the entities involved are of a completely different kind of an ordinary object, perhaps this is the correct course of action. Moreover, it is often said that QM requires its own rules. Thus, we should accept that QM cases have their own explanation to their own problems.

Nevertheless, one might object that since I am accepting different rules for QM cases, why not accept giving up PII? After all, this would be QM accepting its own rules. Fair enough. I could accept that. However, there must be a reason to do so and there must be order to give up on old rules and adopt new rules. The reason presented by PII opponents is related to the explicability of phenomena (actual or possible). To which I would reply that the mixed strategy presented by Saunders is enough to explain that. This certainly would raise the objection of lack of explicative uniformity, and that can be answered invoking another theoretical virtue, namely, conservativeness. The order required is due to conservativeness. In a Quinean vocabulary, we wish to disturb our web of belief as little as possible. Thus, we want to move as little as possible the central nodes of it that would cause a chain reaction of movement towards the outskirts of

[107] Muller talks only about coherence and conservativeness. I am adding the part of explanatory power, economy and loss of uniformity, which might be equivalent to what he is saying. In that case, I beg the reader excuses for the redundancy.

it. It is always preferable to move a more peripheral node. PII as a ubiquitous metaphysical principle is way closer to the centre than a strategy of defence of PII. If one does not subscribe to Quinean jargon, in a Lakatosian way, one could say that PII is way closer to the nucleus of the research project for reality than any of the alternatives presented to defend it. Thus, if really needed, Saunders alternative is to be preferred rather than abandoning PII.

Great! PII seems safe enough. But can we do better? Perhaps we can find a defence that is ontology-swap proof, proving more uniform. Maybe we can find a defence that is swifter to establish that denying the relevant version of PII is impossible. This is what I shall attempt in the next chapter. But first, to finish this section, I would like to raise a couple of small objections that could have discourage a summing defence from the beginning. First, the reader probably noticed that Hawley's point relies on what we *need* or do *not need* for a theory to work smoothly. Nevertheless, to ask if we need to posit the existence of smaller objects in the scenario is the wrong question. Perhaps it is a pertinent question concerning QM cases or a non-necessitated version of PII. Given that we are discussing a necessitated version of PII, the question posed should be whether we *can* posit a scenario with such existences. If one believes that the scenario proposed by Black is metaphysically possible, it seems we can think of such universe. Therefore, this line of defence seems only available for an unnecessitated version of PII, in worlds like the one described in which co-located particles may exist. The reader might have noticed that this is basically the same objection she and others raised against Hacking's defence. It is only fair that her defence stands up for the same criteria. Second, Hawley's metaontological distinction between objects and existing objects seems misplaced in this discussion, since the notion of object used by everybody must be loose enough to contain possible entities (i.e., non-existent objects), not only actual ones. Or, even better, as Muller put it, it should be "*metaphysically thin* (…) [so that we can put] *anything we can meaningfully quantify over* (…)" (*op. cit.* 206).

4. My strategy: the inconceivability defence

My defence of PII should be viewed as a metametaphysical objection to the supposed counterexamples exemplifying premiss 2 in the dispersal argument. It is an inconceivability argument, structurally very similar to those widely accepted objections to Chalmers' argument

for philosophical zombies (e.g., DENNETT, 1995). Like Chalmers, I believe that conceivability is a guide to metaphysical possibility. Nevertheless, I believe that no one is capable of meaningfully express the idea of two metaphysically indiscernible objects.[108] Whenever distinct objects are conceived, a discerning feature must be conceived with them, usually it is a spatial feature, but temporal ones certainly could also play this role and presumably others, such as *being the additive inverse* and *spin*. Thus, it is not absurd to think that it is not an entirely new defence, but a complementary step of the discerning defence. Especially because it places the PII denier in an awkward situation similar to that of Della Rocca's challenge. However, given the nature of the point being raised as not merely about what is contained in the possible world used as a counterexample or how a Black-like scenario is depicted, but about the very capabilities of building such scenarios, I believe this defence is better classified as a different kind of defence. Thus, it should be seen as a defence that contains the discerning defence as part of it or that has the discerning defence as a complementary step.

4.1. Worldmaking: the relation between conceivability and possible worlds

What is being set whenever one thinks of a scenario with alleged indiscernibles in the relevant sense is something very close to what is needed to instantiate P2) of the dispersal argument, but not quite that, because what is needed for such ends is not really conceivable – or so I will argue. The description of Black's experiment has the apparent illusion of a consistent description because the separate chunks of descriptions form perfectly sensical aspects of the scenario of the thought experiment, but the chunks put altogether do not. This situation resembles somewhat the famous Chomskyan example "colourless green ideas sleep furiously", where the words have consistent semantic meanings separately, but fail to do so when put together. "Two distinct objects" and "completely indiscernible amalgamations of properties" are incompatible chunks of the scenario description, as the discerning argument have shown.

[108] The reader might be thinking I am reaching for a logical positivist position where I am about to say that Black's thought experiment is nonsensical. Well, not necessarily. I am trying to argue that it is not meaningful, but "meaningsome", i.e., its meaning is not full or saturated, but it means something; and that is the problem. This partial meaning links us to an unsaturated aspect of an impossible world, and this gives us the idea that one is talking about a possible world.

The illusion of compatibility emerges because we are not properly entertaining all the aspects of the formation of the scenario being built. Either the description is made in a way that the contradicting parts are never faced together explicitly, or the description lacks some part of the possible world it is describing.

Remember that this scenario is meant to be a possible world in which PII does not hold, thus, showing that the principle cannot be necessary. Assuming an ersatzist view of possible worlds, there are three theories of worldmaking, namely, pictorially, linguistically and hybrid, which are, according to Berto and Schoonen (2017), historically the main theories of how we think about things. Further understanding them might help us to get a clearer picture of the problem with conceiving the scenario.

In the linguistic approach, according to them, thoughts are conceived in our minds in the form of sentences. Not as propositions – whatever they might be – but as sentences themselves; strings of written words or sounds within a natural or artificial language that signify in our minds. In other words, we can entertain[109] thoughts about things in our minds without the use of auxiliary imagery.

Good evidence for the existence of this faculty is that one does not need to have in mind a complete thought with a saturated meaning to entertain thoughts, it can be something unsaturated, e.g., one could entertain a thought of "cow" or "the largest prime number" alone, without thinking about a particular cow or about the largest prime number.

Another pressing evidence in favour of this theory is that it seems that this kind of mental representation is the one involved in general thoughts that could be associated with a large number of images, such as "mammals have mammary glands", which you can *imagine* associated with a big diversity of animals with teats and suddenly catch yourself wondering where they are located in whales instead of imagining whales nursing.

One last piece of evidence I would like to bring is that there might be thoughts in which we only partially understand the terms involved. For example, if you approach my circle of

[109] I will use "entertain" for the vaguest possible kind of thought. Whenever one entertains the thought expressed by "cow", one *conceives* an image of a cow or a description of a cow (or of cows in general, with an incomplete description). "Conceiving" will become clearer latter. For now, just accept that "entertain" is something very vague.

friends in a party and hear me saying that "Bora Horza felt really welcomed in *The Hand of God* circled by Idirans", unless you are also an Iain M. Banks fan, you will have no clue about what I am talking about. However, you could have a hint of what was meant because there are terms you understand disposed within a syntactical structure you recognize. You could even try to look cool and join the conversation saying something like "Namaste, brothers! I also agree that Ms. Horza felt comfortable as she was held by God in the spiritual ritual of the Idiran tribe". A line to which everybody would react like "what the hell are you talking about?". At this point, it would be clear to me that you were not *imagining* the scene of a human-like male alien circled by three-legged monstruous aliens inside a partially sentient ship called *The Hand of God 137*. Nonetheless, you managed to pull off a comprehensible – wrong, though! – reply, based only on the presumably little knowledge you had about the topic, which was basically the syntactical information and some ideas (in the Fregean sense of *Vorstellungen*) associated with the sounds and the unsaturated meanings of the words.

On the other hand, it seems perfectly plausible that fans of Banks' book series would be able to imagine Bora Horza and Idiran soldiers based on the descriptions of the books in vivid details of imagery. This pictorial alternative is described by Berto and Schoonen as mentally picturing a scene of the episode described, somewhat like a memory, which is commonly thought to be like an actual perceptual experience but lacking the external *stimuli* part. Notice that although the chosen name for the theory (i.e., pictorial) points to a visual experience, it does not have to be necessarily an image, it can be a sound, a smell a feeling of touch, or whatever is analogous to what we perceive through our senses. To which I must add that it also does not mean that it must be a static picture. It can be something like a paced up little movie, a succession of pictorial images (viz., a GIF), sounds and other sensations that has some duration. All these modalities are ingrained with spatiotemporal aspects to it.

It seems clear that our minds can engender such scenes, e.g., one can fantasise about a contingent future, imagine whether some episode could have happened differently in the past or even create completely original scenes in richness of phenomenological details. In fact, if one pays attention to how Black describes characters A and B conceiving the scenario in which they perform the thought experiment with the spheres, one will see that this is the conceivability in play in the (1952) dialogue. At that time the tools to discuss modalities were in their earliest

form and the possible world talk was commonly put as "ways in which we *imagine* our world could be"

However, according to Berto and Schoonen, there are some situations that require a different kind of imaginative power, e.g., one can purely pictorially imagine that "John kisses Mary", but one cannot imagine that "John kisses Mary, *his second cousin*" in a purely pictorial way. In this case, there seems to be a component that is tagged into this picture to complete the scene. The relation of *being a second cousin of* is not perceivable in the mental GIF of John and Mary kissing. Berto and Schoonen are prone to claim that these kinds of relations must be thought in a linguistic way through some kind of labelling. I would like, however, to suggest that for simple scenario of this kind, there may be a purely pictorial alternative to explain this kind of thought entertaining. Perhaps these relations are *felt* by us (in the lack of a better word), not perceived. One could imagine more than one scene associated with the kissing scene; perhaps a simultaneous collage of other scenes of stereotypical actions performed by cousins, whatever they are, or, in worst case, a simultaneous collage of their siblings that are related, followed by the scenes of their conceptions and the scenes of their births leading up to the kissing scene. Just to make clear, I am not insinuating that we normally conceive scenarios like this, but only that it would be possible to do so, unless one is plagued with aphantasia.[110] Certainly, a godlike mind could do it from every angle and without editing – keep that in mind.

It seems that both capacities play relevant roles in our mind's conceiving routines. At this point, it is safe to say that a hybrid theory that encompasses both capacities is more adequate to describe how we think. I suppose that hybrid imagination sometimes works more pictorially with pinches of linguistic functions such as the tagging of titles (e.g. 'second cousin of John' or 'the best president') or of imperceptible things such as relations (e.g. 'being a second cousin of' or 'being elected by… to…') and sometimes more linguistically with pinches of pictoriality, similar in structure to what Russell meant by singular propositions, a sentence that instead of a

[110] Aphantasia is the condition in which people cannot voluntarily create images in their minds. The existence of people with this condition is evidence that some other way of thinking must be available to them. For though they are able to remember and dream pictorially, they cannot do it voluntarily, they cannot deliberately form images, so they must be able to think in another way, *viz*. linguistically, otherwise they could not have normal intellectual lives, which apparently, they do.

name in it has a representation of the thing.[111] However, we can consider that some events are conceived purely linguistically, given that there are individuals with aphantasia who can perform mental operations with hypothetical scenarios and can understand scenarios of science fiction when they read them, as well as we can consider that some scenarios can be thought purely pictorially.

In any case, this seems to be the most adequate way to explain our possible-world-making faculties. Since here we are committed with ersatzsism, one might think that we are committed only with a linguistic worldmaking way. But this need not be the case. Adopting a hybrid theory approach to worldmaking permits us to create possible worlds in pure linguistic ways, of course, but we can translate them in pictorial or hybrid ways too. Let us see how this can be accomplished.

If we choose to understand possible worlds as purely linguistic constructions or as hybrid constructions of the Russellian-singular-proposition-like kind, such as the ones described above, we could define possible worlds as *maximal sets of worldmaking sentences*.[112] Remember: a maximal set of worldmaking sentences is a set whose members are only sentences of the worldmaking language, such that for any sentence "P", either P or its negation is a member of it (see section 2.1.1. or JAGO, *op. cit.*, p. 159). Thus, there is no space for incomplete descriptions when talking about possible worlds. Furthermore, if we are talking about metaphysical principles, non-vague descriptions are a must, unless we allow the possibility that reality might be vague itself – which I am open to accept, nonetheless argumentation would be needed to back this up.

---

[111] Although Russell said that Mont Blanc itself was part of the proposition "Mont Blanc is more than 4,000 metres", this cannot be the case, if propositions are mental constructs. It makes more sense to say that our pictorial representations of "Mont Blanc" are part of the proposition This would also help to explain how two individuals could disagree about analytic sentences that have "Mont Blanc" as subjects without appealing to Fregean senses, i.e., they imagine pictorial representations. Note that I said *similar* to what Russell meant. I want to avoid any commitment to an almost platonic entity such as proposition in the sense he uses it, and I do not want to say that the individuals themselves are parts of the propositions. Pictorial representations of them are. It could be defended that this linguistic hybrid I am describing are propositions, but I do not see the necessity to argue for this here, and this would leave the term available to denote other things such as the set of possible worlds where such and such obtains (see JAGO, 2017, pp. 27ff)

[112] One could construct it as a maximal set of *propositions*, but, again, I do not wish to commit myself with such entities. Therefore, I am going *pace* Jago (2014) consider it a set of worldmaking sentences, whatever they are.

Notice that if we allow worldmaking sentences to be Russelian-singular-proposition-like, we are admitting that pictorial participation in our worldmaking. Thus, if we turn the knob of pictoriality in our possible world to the maximum, we can in principle make a fully pictorially conceived possible world. Thus, let us expand the meaning of "worldmaking sentence" to capture imagined scenes too. I know that it would be best to apply other term here, but for the sake of simplicity, lack of creativity for new jargon, and desire to avoid the thorny notion of proposition, I beg the reader to accept this expansion. In the worst case, one could say that these scenes are translatable (or de-renderable) into purely linguistic descriptions and, thus, one could say that they are equivalent to worldmaking sentences.

Back to the previous subject, purely linguistically conceived ersatz worlds have enormous descriptive power because its limits are basically settled by the syntax and vocabulary of the language they are built from. If one chooses to *express* a world with a deviant logic, one can do it. There are no *rendering* constraints for purely linguistic ersatz worlds, which are very present in most of the hybrid and all of the pictorial alternatives (I will explain this below). However, these worlds should not be considered *metaphysically* possible, but *logically* possible (relative to some logic, of course) at best. Metaphysically possible worlds require that we adopt classical logic constraints (see section 2.1.). Thus, no contradictions are allowed in this set, i.e., it is a set that, concerning any sentence P, must have P or its negation *and never both*.

Notice that I reserved the term "express" to talk about the possible worlds with deviant logics, e.g., those with contradictions, instead of the term "conceive". Our minds entertain thoughts *about* contradictions, nonsensical sentences, unsaturated sentences, etc., but it never entertains thoughts *of* these things. We can linguistically express them, but we cannot conceive them. If we made an analogy of our mind with a computer with a very potent graphics card, we must say that it would never render these lines, nor anything that mimics or models them. Which we certainly can do with other lines. Those that can be rendered, we shall call metaphysically possible.

I am borrowing the terms *renderable* and *unrenderable* from Computer-Generated Imagery (CGI) engineering vocabulary. Rendering is the term used to describe the image synthesis generated by some algorithms from linguistic inputs given by the digital artist. Analogously, I believe that our mental imagery is generated by some (yet mysterious) algorithm-like mechanisms that enable us to create the pictorial scenarios from the sense data

previously acquired by our senses and – most times – by some kind of linguistic assembly. What I have in mind is nothing too different from the typical modern British empiricist, just with a few more senses, such as enteroception, proprioception, etc., that were discovered later.

Unrenderable worlds are ersatz worlds that escape this rendering capacity because they contain information that is beyond the capability of our machinery. It is beyond for at least one of the following reasons: it clearly disobeys the principles of classical logic (e.g. "the square is a circle", "a wooden iron bar"), it clearly contains terms that lack meaning (e.g. "Chinforinfola", "Xolofompila"), it contains semantically empty contents (e.g. "Green ideas sleep furiously"), it contains more than three spatial dimensions,[113] or it contains more than one temporal dimension (whatever this could mean).[114] This means that they are expressible *only* linguistically, which means that the set of unrenderable worlds does not overlap with the set of the renderable ones at all. Incomplete descriptions of scenarios are also unrenderable, for they lack directions for some of its parts. But *strictu sensu* those are not worlds. The lack of some relevant pieces of information that exclude them from this class, for they are not maximal. Notwithstanding, for the sake of simplicity, we shall say that they are unrenderable worlds too. One last thing worth mentioning is that, although unrenderable worlds may not be useful for metaphysics, they are useful for fiction and epistemology.

Renderable worlds, on the other hand, are those we can conceive purely or majorly pictorially. I say majorly, because as it will be seen later, there might be cases in which pure pictorial conceivability cannot be achieved. Thus, hybrid world can also be renderable, although many of them are not. Rendering is a more limited worldmaking tool exactly because it depends on our capacity for pictorially conceiving and we cannot express meaninglessness, nor unsaturated thoughts, nor contradictions through it. Rendering is constrained by the senses in our minds. In other words, what I am here calling renderable worlds are usually described as *scenarios* describing ways worlds could be (whereas the unrenderable worlds would be descriptions of how the world could not be). Therefore, I believe rendering is a good guide to

---

[113] This does not mean that we cannot develop four-dimensional thinking. We can do it through purely linguistic ways or like 4-D engines for games, where you see three dimensions at a time, but the whole context gives you a fourth dimension – thus, in a hybrid way –, but it is just an "illusion" of a four-dimensional world.

[114] This might not be an exhaustive list.

learn what is metaphysically possible: If it is unrenderable, it certainly is not metaphysically possible. If it is *prima facie* renderable, let us try to render in a broader picture from a broader description and see if it is really renderable.

Nevertheless, one could object that there are some cases which we can depict with our minds' eyes – or even see in drawings and pictures – but certainly could not be the case in reality (e.g., Escher drawings and Reutersvärd's impossible triangles). Thus, there must be something else involved in the construction of metaphysically possible worlds. A metaphysically possible world cannot just be a scene *simpliciter*. This conceived scene must have some other information attached to it, be it by tags of information such as second cousin John example, or this scene must be more complex, it must have different senses working together or different points of view. In other words, there must be methods to avoid mistaking an impossible world as a possible world. There must be tests the conceiver can apply to avoid this situation. The method I propose is reaching for a Godlike perspective.

4.1.1. Godlike perspective, or ultimate perspective, or total perspective

Yablo said that "*Metaphysics aspires to understand reality as it is in itself, independently of the conceptual apparatus observers bring to bear on it*" (1987, p. 307), however, I am afraid this can never be the case. Although the debates about this thought experiment seem to describe it in an aperspectival way, i.e., a description from no point of view, we can never understand the reality without a *perspective*. Not just because a spatially determined point of view is required to imagine it, but even in a looser sense, even conceptually. There is no way to describe a scenario from an aperspectival point of view, for the very language we use to describe it is contaminated with our human perspective to the core. The words we use to do it are based on our ontological (an ideological) commitments, which themselves bear some prior dependency on our senses.

The best we can do – and what we should do – is to extrapolate our limited human point of view and see things from other different – and sometimes – more complex points of view. This we can do. Thus, we should not limit ourselves to such limited, restricted points of view. I propose that in order to test the renderability of complex scenarios and make a fair setting for the thought experiment the PII opponent is building, one must seek for an ultimate – total –

point of view, a godlike perspective of this allegedly possible world – or at least the closest to this that one can reach. However, we should seek this always keeping in mind that the baseline conceptual, sensory, and logical apparatuses for this extrapolation will always be human.

Take a perspective to be a pictorial conception of something from a determined spatiotemporal point of view loaded with the best collection of contents available to our mind's eye. Thus, what I mean by a godlike perspective is definitely not something like being inside a mind that has the content of every predicate possible and knows everything beforehand, but a perspective that *enables* one to know everything within a possible world if one sufficiently engages with the imaginative resources available. To be more precise, it is not *a* perspective, but *a sum* of perspectives, ideally the sum of all possible baseline human perspectives this world can provide.

I believe that this is what Nagel was reaching for at the end of his (1974) celebrated paper. We certainly cannot know how bats conscientiously experience the world, but we can at least have an approximation by analogy to our own senses and by behavioural tests, a translation to a human conscience, if you will. Concerning a godlike point of view, the latter is not available to us – for obvious reasons – nonetheless the former is. One can imagine a sculpture from different, previously unseen, points of view, based on what is already seen from familiar points of view. Or even, depending on how rich the description is, one can imagine this sculpture even without having seen it – in the same spirit that history of art was made for centuries before photography. Nowadays, with the rich vocabulary toolbox of the disciplines such as Graphic Design and Cinema, one can describe many ways in which we can visualize a sculpture (e.g., zoom in, zoom out, saturation of colours for temperature capture, bullet time capture, etc.). By imagining such a sculpture from every possible angle, from every possible zoom distance, with every possible filter (e.g., one that turns temperatures into colours, other that turns magnetic fields into colours) applied at the same temporal slice, one can have a total perspective of the sculpture and its parts.

Of course, one cannot do it all at once, for our limited human brain does not have the computational power to deal with such a massive amount of information. But there are ways we could use as allegories to explain how a godlike being would deal with such information. For example, the way the Architect sees everything that happens inside the Matrix and in the past versions of it, at the end of Matrix Reloaded. For those who never watched: he is in a room

full of screens that play everything happening inside the Matrix at his will. Another possibility is to think that one can control time – forward, backward, freeze –, select a temporal slice and may access every corner of this world as an FPS player in spectator mode.[115] This spectator does not interact causally with the particulars of the world, whereas he can perceive everything in it while he flows around summing particular perspectives in his mind. A third possibility that may be not that simple to entertain, for it has no use of visual experiences, although it might be easier to apply to Black's scenario because it is a simpler scenario, is what I call the *interoceptive proprioceptive universe*.

Fortunately, we do not have to base everything in visual perception only, we have other senses that could assist in this imagining or could even substitute vision. For example, Strawson in his (1956, Chap. 2) entertains the possibility of a world conceived *only* by sounds. Thus, I propose to conceive the world by other senses too. Interoception is the sense that allows humans – and probably other living beings – to feel their inner organs. Proprioception, on the other hand, is the sense that allows humans – and probably other living beings – to perceive the position and motion of their bodies. Together with the exteroception – which is not a sense but the sum of our five senses – and the vestibular system (inner ear) they allow the individual to understand its spatial positioning in the world.

Since we are not dealing with things outside the universe that Black invites us to think about, we will ignore exteroception. Now, suppose that this universe is sentient, and that it is the conceiver of the thought experiment. It perceives its limits – if it has them – and its contents. This universe feels through enteroception the spheres inside it, like we can feel stones in our kidneys. Thereby it can know how many objects there are inside it and, more relevantly for our ends here, *where* they are. We can conceive being in this universe's perspective.

One interesting feature of such allegories that is important to stress for the purposes of Black's thought experiment is that the viewer in these scenarios does not interfere with the physical settings of the experiment. In other words, she does not bear any causal relation to the world she is watching. Nonetheless, one could raise some objections to this strategy. First, one

---

[115] For the older reader, that is not acquainted with First Person Shooter games, the best substitute would be a character in movies like any 1980's Christmas ghost film, where the character is in a ghostly form where he cannot interact with the people or the objects of his surroundings, but he can see, hear and feel everything.

might object that the observer bears some metaphysical relations to the scenario, which might be understood as contaminating the experiment, because it adds more relations to the description of the scenario than its intrinsic features and they break the symmetry of the scenario descripted (e.g., *it is perceived in such and such way* or *is viewed from such and such point of view* (angle)). Second, one could argue that, though this godlike perspective captures many aspects of reality, there may be properties that escape it. Let us, then, start the next section analysing some possible objections to the view proposed.

4.2. Relations, contents, references… objections.

Let me separate the first objection in two. The first is the accusation of creating the relations between the conceiver and the scenario, thus, relations with the objects, that are not part of the scenario themselves. The second is the claim that these relations would break the symmetry of the experiment. Well, could we not create those relations? Not at all. By asking "would it not be possible for such and such a universe to exists containing such and such particulars in it?" is not the proponent of the thought experiment creating a relation of some sort with those objects? For example, the relation of *referring to* which the conceiver holds with the conceived scenario or to its parts or the relation *being conceived by* held by its parts to their conceiver. These relations are conveyed to the reader/listener and without them one could not make sense of the thought experiment. It seems that without these relations, the descriptions of the experiment would be meaningless. One cannot run the thought experiment without relations of some kind of the experimenter with the contents of the experiment, be it a relation of reference or of conceptualization, or even understanding. Thus, they cannot be excluded, otherwise there would be no experiment to begin with. These relations between the experimenter and the experiment are unavoidable.[116]

[116] One might think that my argument is the same as Berkeley's Master Argument. But this is not quite right. Berkley's argument is concerned with the actual world, mine concerns only possible worlds, which are dependent on the mind anyway (ersatz worlds). My argument, however, can be seen as a modal spin off of Berkeley's if you will. Nevertheless, the criticisms of Berkeley's might not apply as smoothly to mine's.

However, we can argue that they are not metaphysical relations, but metametaphysical relations (epistemic or semantic relations, if the reader feels more comfortable). The observers/conceivers bear no causal powers to the contents of the worldmaking sentences nor to the rendered objects they are conceiving. Thus, it is not that clear whether this should count as breaking the symmetry of the scenario or not. I would argue that they do not since they interfere in nothing in the scenario's mechanics.

The PII opponent may argue that they do not break the symmetry of the scenario's physics, but it breaks the symmetry analysis of the scenario, thus, they break some sort of epistemic symmetry that must be maintained for it to work given that it is testing the epistemic part of PII.[117] Fair enough. It is comprehensible that, from some points of view, the relations to the observer are asymmetric. But we cannot give up these relations, otherwise we might be caught amidst a categorial mistake. The principle has an evident epistemic part, namely, the (in)discernibility part. The experiment must test whether there is a situation in which the objects cannot be discerned. To claim that there are indiscernibles, there must be an epistemic agent performing (or failing to perform) the discernibility part. If we do not allow a point of view as part of this experiment there is no point in saying that there is a thought experiment about (in)discernibility in the first place, since it would not be testing the capacity of discerning objects within the scenario. In other words, if no epistemic relations are permitted, it seems like a categorial mistake to question the (in)discernibility of a scenario, for if there is no access to the objects discerned a cat would be as indiscernible from a stool in an inaccessible room as one sphere from the other in Black's scenario.

So, since we cannot discard these relations, on of conceiving an inconceivable world, maybe we could find a way to make these relations symmetrically integrated with the scenario. Character A in Black's (1952) dialogue tries that making the observer symmetrically related to the spheres and distinguish them by the relations *being to the left of* and *being to the right of* that would enter the scenario with the observer. However, character B does not accept that, because there would be neither left nor right if there was no observer. B accepts the entrance of

[117] Remember: PII is a metaphysical principle because it deals with identity, but it has an important epistemic component, namely, the discerning part.

the observer, but he would put the relations with the observer on the list of forbidden features since they break the symmetry of the scenario. This seems to make the task impossible for the PII champion then.

However, there are more complex ways to conceive the experiment that do not affect the symmetry at all. They require unusual ways of conceiving but if we limit our experiment to these perspectives, symmetry might be maintained yet the spheres can be discerned. One alternative was already described. We could conceive this world from the point of view of itself, that is, a sentient universe endowed with enteroception. From this point of view, it seems that the scenario can *only* be pictorially conceived symmetrically, since there would be no *being to the left of* or *being to the right of*, etc., since the observer is everywhere. Thus, the experiment could be performed. Another way to do it would be from a split screen visual conception in the style of 90's FPS videogames from two (or more, if the symmetry is radial instead of bilateral) symmetrically opposed points of view. Both perspectives respect the symmetry of the scenario in a way that the conceiver will always know that one sphere is itself and also know that another is not the former, she will also know that the latter is itself and not the first one. This would be known by perception of the spatial features displayed by each of them. Even if one does not know which sphere is on its left or on its right, or cannot locate the spheres using names, one will always know I) that they are at different spatial locations and II) that they hold spatial relation D with one another but not with themselves. This can be linguistically conceived in a very Brazilianish sentence: "one is not the other and the other is not the one". The conceiver would know that there is *this* sphere and there is *that* sphere.

Probably character B would not accept this too. He would argue that these conceptions would break the symmetry of the scenario because they use some kind of demonstrative, e.g., *this* sphere opposed to *that* sphere, is *here* opposed to is *there*. This point can be enhanced to capture other indefinite articles such as "one" and "(an)other" in some chronological order that would break the symmetry. I would still appeal to the metaphysician's court that none of these relations are in the scenario, they are only in the conceiver's mind, and they create no new relations inside the scenario. Yet, my plead might not be accepted. It might be that my view about the relation between metaphysical possibility and renderability is wrong. In this case, it might be wiser for me to assess whether the spheres could be discerned linguistically, then.

Remember that the possible world in which the experiment is conceived must be maximal, i.e., for any worldmaking sentence, it must present it or its negation. So, we can ask the proponent of the thought experiment whether the scenario and its constituents present some feature or not. To avoid unwillingly breaking the symmetry of the scenario by pictorially conceiving it, let us play a quiz with the proponent of the experiment in which she can only answer "yes" or "no". If the opponent refuses to answer any of these questions, we can conclude that this description is incomplete; thus, this scenario is incomplete, and this possible world is not maximal. Therefore, we can claim that the experiment is not just unrenderable, but worst, it is inconceivable and cannot yield any results about anything whatsoever.

We should start asking whether there is numerical distinction of objects in the scenario (Muller's Q1), to which she would answer positively. Then she would answer positively to the question about existence of two objects. If we asked whether these two objects occupy the same spatial locations, fearing hurting WLP, she would reply negatively again. Would she answer that one positively, we would find ourselves in a situation where the Ockhamnian nightmare of unreasonable proliferation of co-located objects that Della Rocca threatened us with is upon us. This is not acceptable. Besides, it is clear in Black's description that they share the property of being at some distance from each other (but there is no need to rely on it). Thus, we can imply that these objects occupy different locations. If this is the case, then, we can ask her whether they have different spatial properties, or even more precisely, whether they occupy different locations, to which she will be obliged to say "yes". In this case, then, we can conclude that the objects in this scenario do not share spatial properties, viz., that they present different spatial properties. I might not know exactly which are these properties, because were I to acquire this knowledge I might be breaking the symmetry or be using coordinates (considered equivalent to names) in the process, but I know that the different properties are there. There is at least one *F* that an object (*x*) has that the other (*y*) lacks. Even though I do not know who *x* is, nor who *y* is, nor which *F* is not shared by them, I know that there is a *spatial property* that distinguishes these objects. This should be enough to demonstrate that PII is safe (without relying on relation D, as Muller proposed, in case someone might raise an issue about it). I am not discerning the objects within the scenario, notwithstanding, I know they are discernible.

Nonetheless, let us see whether we can go further and discover more about this property without spoiling the experiment. From the corollary that these objects do not occupy the same

spatial locations, we can conclude that the coordinates for the spheres would not be the same, although we do not know which are they. We also know that they are somehow symmetric, for this was also conceded by the opponent when asked. Thus, whatever kind of symmetry this scenario presents (radial, bilateral, or whatever) the spatial coordinates of the objects must be presented in this form: whenever an object occupies (x, y, z), the other will occupy (-x, -y, -z). From this, we know that the objects differ in spatial locations and that there is a relation of opposition held between these locations (if one is a substantivalist), which implies the same opposition being held by the objects (with one is a relationist).

One might object that this defence could not be used by a relationist, however. For the spatial features of the objects are settled relationally with the other object, i.e., they do not present coordinates for spatial locations. Well, they might not present coordinates for spatial locations *at first*, according to relationism, but one could derive them from the relations that are very well present in the scenario. For example, from the relation …*is at 2km from*…, which is part of at least one worldmaking sentence in forming this scenario, we could derive that its bearer is located at somewhere, thus, that it bears *is located at spatial region* [ ] which I might not know which one is, but I know that it is 2km away from some other object that bears the same property with a different content in the brackets. Furthermore, we can do better, given that this is a symmetrical scenario, we can know that their positions are additive inverse of each other in relation to the axes of symmetry of the scenario. Thus, *occupying a location that is the additive inverse of the occupied by* is another feature that is not shared by the objects in the scenario.[118] Thus, we can even know that whatever value fills a bracket, the other would be filled with the same value with the opposite sign. These arguments might be enough to avoid Muller's worries about trivializing spatial features by the use of coordinates (2015, p. 212-3).

On the other hand, if one is an absolutist about space (a particular kind of substantivalist), a defence might even be facilitated. Whether one is considering spatial points as objects, she could even defend a stronger version of PII, namely, PII1 which says that no two objects share all their intrinsic properties. However, I will not do this here. In any case, if one ignores the

[118] This relation is brought up by Ladyman (2005) to solve the problem with imaginary numbers. See section 5.1. below.

epistemic part of PII or, at least, severely limits its scope, then to show that distinctness comes necessarily accompanied with a difference in (non-trivializing) properties between objects must be enough for establishing a metaphysical principle.

One objection that might be passing through the reader's mind right now is that given that pictorial conceivability is less expressive than its alternative and that most points of view within it are not permitted because they break the symmetry of the experiment, then, it is irrelevant, because matters will be settled through a linguistic analysis in the end. Well, that is the correct understanding of my proposal. The test of renderability is required to tell us whether the world is metaphysically possible or not. If the scenario does not pass this test, we could safely say that it is metaphysically impossible, for the world in which it should be placed is not a possible world. The linguistic search is aimed at finding information about the permitted discerning features. Furthermore, there seems to be a small group of special cases that are supposedly linguistically inconceivable whereas pictorially conceivable, namely, unlabelled graphs, which are the subject of section 5.2. below.

Still on the subject of the expressive power of pictorial conceivability, the second original objection from last section was that the godlike perspective might not cover every feature the world could be made of. Let us also divide this objection in two. The first part, concerns cases that seem to escape our enhanced sensorial/pictorial mechanisms, for example, there is a micro-level where zooming in would not be possible anymore because it would not even make sense to say that images are being formed, given that the zoom is at the level of photons themselves; four-dimensional beings, which escape our three-dimensional constrains; colours that other animals can see that we can never see or other *qualia* for which we lack the sensorial apparatus to do so. The second part of the objection concerns cases in which even our linguistic conceivability capacities fail to reach, e.g., contents of contradictions, Lovecraftian creatures that cannot even be put in words – dark matter, perhaps.

Concerning the former group of cases, the first thing that might be stressed about the zooming in example is that it might be the case that at this level we are not dealing with objects anymore (see sections 3.3.ff.). But glossing over this possibility, assuming there is a continuity in reality. By zooming in, we can treat this example as the others and say that we can do as we always do when we are unable to have a clear grasp on how things are presented, namely, build models that would represent those situations. In the micro cases we can build geometrical

models of how the particles interact. The resulting structures of these models, which are perfectly perceivable through the usual means, would resemble structures of things in macro-levels. In the case of four-dimensional beings, we can picture them three sides at a time. This is how 4-D videogames are made. The senses of animals such as the "electric nose" of sharks, for example, can be understood through an analogy with our other senses, as one could explain colour to a blind person, and the colours beyond our spectrum can be understood modulating the intervals between the colours being compared to relations within our visible spectrum (see NAGEL, 1974, pp. 449-50).

It might be further objected that in all these cases some kind of linguistic component must be present to perform the analogy. Thus, a purely pictorial conception is not always available, or worst, a purely pictorial conception is never available if one adds enough complexity to the situation. Perhaps we cannot escape the linguistic component of conceivability. Fair enough. The possible world as a whole might not be completely renderable at the same time, for we have a humanly limited graphics card. One could expect that only God's point of view would do this job. But for our purposes it is enough to grant that part of this world is pictorially conceivable at a time, namely, the scenario, i.e., the aspects that are being assessed together. The sum of these aspects can work as linguistic hinges linking all the scenarios composing the possible world. Thus, we might still maintain that any object that is said to be a metaphysically possible object can be translated into a pictorial framework or into a model that represents it pictorially at least in its most fundamental metaphysical components, namely, space and time (see section 2.1.). Well, can we? Let us now consider the second group of objections.

It seems that there are cases in which what is being described escapes even our linguistic conceivability faculties. Take, for example, the contents of contradictions, the content of the term "void", Lovecraftian creatures that supposedly cannot be described, the Abrahamic god – dark matter, perhaps. All these things might be ineffable entities, that is, it may be that they exist somehow, but the way in which they exist is ungraspable for us. Each of them has a different nature (supposedly), thus, the verdict about what is the case may vary, however, most of them perform as significative placeholders in sentences with comprehensible meaning – at least to a certain point.

Some claim that these terms are meaningless (e.g., logical positivists); others, that they all share the same referent, something like the False, or an empty set, or the Void (e.g.,

Fregeans); others, like Jago (2014), claim that they all refer to impossible worlds, i.e., sets of worldmaking sentences containing contradictions. The fact of the matter is that they cannot be pictorially conceived and cannot be linguistically further explained – at least beyond a certain point. In any case, these things must be outside the realm of metaphysics if metaphysics concerns reality and what there is and what could be. The Void certainly is non-existent, the latter is precisely the absence of stuff of any kind. Empty sets and impossible worlds are mental entities that only exist abstractly in ways that are not actualizable. In the latter case, we can express thoughts *about* these things, and they refer to these ersatz worlds, thus they can be conceived linguistically. Therefore, they might still be, depending on which logic is considered, logically – and epistemically – possible worlds (see JAGO, *op. cit.*, chapters 2, 4 and 5).

Here, for the sake of brevity again, I would like to extend the concept of impossible (ersatz) worlds to include open incomplete worlds i.e., non-maximal sets of worldmaking sentences. Incomplete worlds are even beyond what Jago proposes as impossible worlds, concerning their absurdity, however, for our purposes here we can lump them in the same class without loss. The point is that these worlds are unrenderable, thus they are not metaphysically possible. They are only epistemically possible, i.e., we can have beliefs *about* them and treat them as objects of knowledge, though we cannot entertain their contents (see *ibid.*, chapters 6 and 7). So, one would never be able to make this world come into existence, because they defy the laws of (classical) logic (and its extensions) or the concepts of space and time. They are metaphysically impossible.

One last objection that was briefly raised in the end of section 4.1., but not explicitly addressed is that there are some scenarios that are pictorially conceivable, but seem to be metaphysically impossible, e.g., Escher drawings and the Reutersvärd's impossible triangle. These scenes are illusions. On closer inspection one would see that there is nothing impossible in them, but in our presuppositions or implications about them. They are two-dimensional images. They render a description of something as two dimensional. The problem arises when we assume that they represent a description of something three-dimensional, which cannot be made. In this case, when one has the description of what is seen added with the presupposed clause of being a three-dimensional object, then, contradiction arises, and it becomes unrenderable. Therefore, they are metaphysically impossible scenarios. (For more on the pictorial conceivability of impossibilities, see SORENSEN, 2002).

4.3. Final remarks about my strategy

Summing up the defence: its main point is that one must accept that there is a relation of some sort between a conceived scenario (or, more precisely, the objects within it) and the conceiver of this scenario. Otherwise, the thought experiment could not run. It is a thought experiment, after all! This relation may be thought as an epistemic relation or even a semantic relation. There is no need to discuss its nature here. However, granted that this relation must exist, there are two ways in which could exist, linguistically or pictorially.

The best test one could run for the metaphysical impossibility of a scenario is to render the description of the scenario, i.e., to conceive it meticulously and exhaustively in pictorial fashion. If one finds an unrenderable line within the description, i.e., if there is an angle that cannot be pictorially conceived, then the scenario is metaphysically impossible, for it is running in an impossible world (in the extended sense). The scenario described by the PII opponent could be either renderable and display discerning features for the objects within them or be unrenderable. However, some critics might not accept the pictorial conceivability to yield a non-trivializing feature to the experiment. In this case, one might linguistically conceive the scenario to search for this feature in permitted ways.

Conceiving the scenario linguistically permits us to express unrenderable scenarios, though. Therefore, it is riskier to take one of those as the scenario where the experiment runs. Nevertheless, this technique presents an advantage for the PII champion, i.e., one can deduce the existence of discerning features without picking them out (viz., through the quiz). So, given that we can imply the existence of discerning (non-trivializing) features (in black-like cases, the spatial locations) by the description of other features of the scenario, we have a warranty that the spheres are discernible, even if by conceiving them I trivialize the properties. According to Muller (2015), to present the coordinates of the spheres is to use names, thus it should not be allowed. According to Black (1952), one cannot pictorially conceive the scenario without breaking the symmetry; however, none of them can deny that there are spatial features playing a distinctive role. This leaves the PII dissident with a trilemma: either 1) the experiment is not feasible, because the world in which it should run is a non-maximal world, an incomplete description; or 2) it is feasible, because the maximal world is linguistically conceivable, but the

spheres are discernible because they have different spatial locations; or 3) the experiment is feasible because the maximal world is pictorially conceivable, but the undeniable spatial relations between the spheres and the conceiver jumps to her eyes (or other senses). In any case, PII is safe.

Before we pass to the analysis of these defences in the cases of mathematics, I believe it is worth making a few scattered remarks about some topics explored in this chapter.

The first issue worth bringing up is that Muller (2015, p. 213) claims that the spheres cannot be definitively described, otherwise they would be absolutely discernible. Well, maybe they are absolutely discernible and this whole debate is a big misunderstanding caused by our lack of attention to some preconditions for such thought experiment. Another way to express this concern is to claim that we find ourselves in a situation similar to the following: we are conducting an experiment in an inaccessible room, where we have a cat and a stool, with the same mass, the same number of atoms, the same overall temperature, the same height, the same distribution of colour in the surface, four legs, and a handful of other equal features. We, then, tell the subjects of the experiment that there are two objects in the closed room with these features of the objects keeping from them any of the features that could be used to discern the cat and the stool and ask based on the information conveyed which is a cat, and which is a stool. The subjects will be unable to tell which is the cat and which is the chair, or even that one is a cat and the other a chair, but we should not conclude that they are unable to distinguish a cat from a chair. We can just conclude that in that situation, from that point of view, with that limitation of information, they cannot. This also does not permit us to conclude that there is no catish and stoolish properties in the room. From a point of view where they could access the relevant information one should be perfectly able to distinguish the cat from the chair. The cat and the chair are clearly not indiscernible, in this case, they are being made indiscernible for the subjects, in a very biased experiment.

The difference from that biased experiment to the case of Black's spheres is that in the latter the experimenters grant us access to a greater number of features of the objects, more precisely, every characteristic of the spheres but the relevant ones, namely, the spatial features. Therefore, my diagnosis of the conundrum is that the prohibition to access spatial features is unreasonable, since they are within the scope of PII, and they are essential to the conceivability of the scenario. What is really happening in Black's thought experiment is simply that the

spheres presented in it are not being discerned, instead of a presentation of a scenario with indiscernible spheres. Perhaps it would be more accurate to say that Black's spheres are indiscerned, rather than indiscernible.

The second issue I would like to bring up is that it is not clear to me whether this experiment is an antinomy concerning rules of identity and the concept of space or a categorial mistake in which we are displaying things that do not belong together, which might be one of the following: indiscernibles and spatial dissonance, or distinctness and indiscernibility. If a categorical mistake is the case, it is comprehensible why this debate has going on for so long. Apparently, we are not very good to spot categorial mistakes when categories that are not clearly defined in our vocabularies are in play. If an antinomy is the case, we should try to describe the scenario as best as we can in order to make this clear outright. Nevertheless, perhaps both are the case.

One more thing that should be stressed, is that, although we cannot properly conceive or completely describe the scenario proposed by Black in a consistent description, that does not mean we cannot talk about or refer to it. As we saw, Black's universe can be interpreted as an impossible world with a semantic value different from an evident contradiction such as *P and not-P* if we adopt a hyperintensional theory. Nevertheless, I believe PII would still be true because impossible worlds are devices to talk about epistemic possibilities or possibilities within deviant logics and, thus, should be excluded from talks of metaphysical possibilities. Thus, if we take at face value what character B in the dialogue wants to know, that is, whether it is *logically* possible, the answer is: depends on which logic. However, I believe that the relevant sense of the phrase for this debate is metaphysical possibility, otherwise we should also count as logical possible that there is a universe where "a frozen inexistent invisible green unicorn ride freely in trios", and certainly this sounds preposterous.

One last minor thing that might worth noticing is that if we consider mathematical objects as objects in this picture, it is impossible to have a scenario with only two things. From the point you say these two things are spatially related you have at least one more thing in this scenario, namely, an axis of symmetry (bilateral or radial). Black entertained the impossibility of having a universe with only two objects. Character A noticed that the axis of symmetry is as an object that cannot be excluded from the scenario. However, the problems character A puts on the table are immediately put aside when character B changes from a bilateral axis to a radial

axis. Maybe the interactions of these three entities are not that simple and further investigation is welcome. Another possibility is that there are, at least, two more things, the axis, and the number of objects there. This last hypothesis is less plausible, since numbers might have a different nature (they might be *F*s), but the axis of symmetry and other geometrical (or topological) features might be an issue for the proponent of the experiment, especially if the objects considered for the experiment are points instead of material objects. This issue is partially dealt with in the cases of imaginary numbers as we shall see below, but there might be more to discuss in these scenarios.

5. Putting the defences to work against mathematical cases

In section 1.2.3., we briefly touched the subject of PII for mathematical objects. This is a discussion way more delicate than any other case of counterexamples to PII because, given the abstract nature of the definitions in this debate, every view held can be drastically changed by any small change in the understanding of things in any part of the theory, which would include a change in answers for questions about PII. As it was said in that section, it is a broad and still ongoing debate, therefore, I will outline some of the counterexamples used in arguments against PII with very little technicalities, only enough to test the possible defences against them. It is worth stressing, however, that the answers shown here might very well be overruled in the light of better explanations. In any case though, I believe they might be useful at least to test our understanding of what was discussed so far.

Given that we are talking about what is metaphysically possible or impossible and mathematical possibility has the wider scope within the set of metaphysical possibilities, if we find something that is mathematically impossible, this would be an indicative that no other counterexamples of this kind could be made. Thus, in this chapter, I intend to analyse some of the alleged counterexamples to PII in Mathematics and present the defences that better deal with each of them.

5.1. The case of imaginary numbers

Mathematical objects are very controversial entities. The problem of the nature of numbers pesters philosophers and mathematicians alike since antiquity. There are a handful of accounts for the nature of numbers. The most classic is, perhaps, Platonism. Roughly and simplistically putting, Platonists believe that mathematical objects that exist immaterially inhabiting another realm of things apart from our physical world, namely, the realm of forms. These mysterious entities have some mysterious intrinsic properties that make them be what they are (e.g., haecceities).

When imaginary numbers where discovered (or invented), they added a further layer of mystery to the debates about the nature of numbers, for they present properties that real numbers lack. We struggle to make sense of imaginary numbers in the broad picture, e.g., make sense of the square roots of negative numbers. Thus, they seem to be entities of a different kind than your average family friendly real number.

On the other hand, a more contemporary approach would be structuralism. For *ante rem structuralists*, a subset of structuralists, mathematical objects are existing points in a structure that have their identity conditions defined by their positions within this structure, i.e., by their relational features to the other objects in the structure (LADYMAN, 2005, p. 219; MACBRIDE, 2006). In turn, according to Ladyman *et al*. (2012, p. 166), structures are entities that consist of a set of objects, viz., a domain; a collection of distinguished elements from the domain called constant elements; relations held among the elements of the domain for each n, a collection of n-ary relations; and for each n, a collection of n-ary functions on the domain. Additionally, every structure comes with a signature, containing constants for each constant element, relation symbols for each relation and function symbols for each function.

This means that, in this view, mathematical objects have no intrinsic properties. Thus, any distinctness can *only* be provided by relational properties – if they are to be given at all. However, the imaginary numbers i and $-i$ present a non-trivial automorphism which implies that to any substitution of one by the other in any formula that represents a complex plane, results in the same structure. For example, $4 + 5i$ has the same structure as $4 – 5i$ or $8 – 10i$

presents the same structure as $8 - (-10)i$ (which is equivalent to $8 + 10i$).[119] If this is correct, both complex numbers present the same relational properties concerning the rest of the structure.

This puts the structuralist in a dilemma. Either she must assume that $i = -i$, which is not correct; or she can give up on her thesis and assume that what distinguishes these numbers is some form of haecceity, which means that structuralism presents nothing that old platonism had not presented before. As MacBride (2006, p. 64) puts it: "*it is either bad news or old news*".

However, Ladyman presents an interesting exit to this dilemma. He claims that although they are structurally indiscernibles, i and $-i$ present irreflexive relations that distinguishes one from the other, e.g., the property of *being the additive inverse of*, which i holds of $-i$ but not of itself, and *vice-versa*. The result of the sum of i and $-i$, that is, zero (0) is different than the sum each of them with itself, which will result in $2i$ and $-2i$, respectively. Thus, those numbers are weakly discernible amongst themselves. Therefore, putative cases against PII using imaginary numbers can be disarmed by discerning defences.

Nonetheless, there is an objection to this response pointed out by MacBride (2006) due to an argument by Russell. It roughly says that such irreflexive relations imply numerical diversity, because for objects to bear such relations they must already be numerically diverse. In other words, numerical diversity is *prior* to these relations. The particulars that would bear the relational predicate must be there *before* (in some sense) holding the relation. This is what Russell would call *bare particulars*. It looks very much akin to the objection raised by Hawley (2009) and French and Krause (2006) against the discerning defence in QM. If this is true, it means that there is something that makes them diverse that is not a relation, it must be some kind of haecceity. This is bad news – again – for the structuralist and for the PII champion. Notwithstanding, according to MacBride, this is not necessarily so. Although he does not agree with Ladyman, he also disagrees with Russell. He believes that the question is unsettled and the answer hinges on the question of how objects are constituted. According to MacBride, the objection to Ladyman's response relies on a specific view of object constitution, namely, a *predicative* view; whilst the structuralist's view is impredicative. That is, the objection requires

[119] Both examples are presented in LADYMAN, 2005, p. 219.

the prior existence of an object to predicate its features, whereas the structuralist does not. In the structuralist view one simply has the structure which does not hinge on prior existing objects.

I am inclined to accept an impredicative view to avoid bare particulars. Although this might sound controversial, one might champion the view that numerosity supervenes (or emerges) from these relations forming the structure. My two cents on the matter are that the structuralist view is preferable, because of theoretical virtues, namely, simplicity and ontological and ideological economy. Instead of importing bare particulars, haecceity and other mysterious entities to our theory, one can competently explain mathematical objects just with structures and the positions within them. Additionally, I am inclined to favour a holistic relationalist view instead of a view with intrinsic properties. But this is theme for another thesis.

Another solution still worth considering is proposed by Hawley where she argues in favour of "*neither falsely identifying nor controversially distinguishing* i *and* -i*, but denying their very existence in favour of the system of which they are putative elements*" (2006, p. 303); in other words, treating them as non-individuals and denying their presence within the scope of PII. In this view, I believe one would have to interpret these imaginary numbers as non-individuals, but part of complex numbers within a structure. In this case, what must be understood under the concept of number is a complex number with a real part and an imaginary part that cannot exist alone. That is, operations done with these parts are illusions that resemble what is done with operations with both parts together as a whole. Well, this is the summing defence applied to imaginary numbers. It comes with all the problems listed in section 3.3.2, therefore, I would advise to use it only as a last resort.

Finally, an inconceivability defence for this case would be somewhat more controversial, since it would require conceiving of these numbers in a cartesian plane of some sort, and this might be interpreted as not being about numbers, but about object with numerosity. Nonetheless, this objection could be halted by a counter-objection that says: a) either this is permitted, and the distinction can be seen as it was in the discerning defence just mentioned or through the numerical distance a number has with the other in the plane; or b) these numbers are only logically possible and not metaphysically possible objects, thus being better interpreted as *F*s instead of *x*s in PII. This would put the inconceivability defence closer to the summing

defence rather than to the discerning defence, differently from what happened in regular Black-like cases.

In any case, there seems to be no uncontroversial alternative. The answer to PII here depends on the nature of mathematical objects, which is still an unsettled debate that could be theme for a whole new doctorate thesis. Therefore, I will not dive into it. However, the important thing to stress is that there are defences that can be used.

5.2. The cases of graphs

Similar to the case of imaginary numbers, it seems that – depending on how one interprets what constitutes a graph – some graphs may be presented as legitimate cases against PII. Ladyman – on the offense team this time – and Leitgeib (2008) present an argument against PII (reaffirmed in LADYMAN *et al.* 2012). The defence team is composed by de Clercq (2012), who defends the view that this menace relies on an unorthodox – though plausible – and unnecessary interpretation of what a graph is, and Button (2006), who actually is one of the authors Ladyman and Leitgeib are replying to.

Graphs, according to Ladyman and Leitgeib, "*are mathematical structures that contain only two kinds of objects, namely nodes and edges between nodes*" (*op. cit.*, p. 390), which can be labelled or unlabelled. Nodes in an unlabelled graph are indistinguishable when isolated, whilst in labelled graphs they are distinguished by their labels even across graphs (trans-graph identity). Graphs can be undirected, when they are edgeless, or directed, when their nodes are connected by an edge. They can be simple, when there is only one edge connecting the nodes; they are called multigraphs when nodes are linked by two or more edges; and when an edge connects more than two nodes, it is called a hypergraph. Finally, the graphs can be symmetric or asymmetric, respectively, depending on whether there is a function that permits a reconfiguration of the nodes and leaves the structure unchanged or not, i.e., on whether they present a non-trivial automorphism or not.

Although graphs are presented by Ladyman and Leitgeib as "officially" set-theoretic entities, it is clear by the practical use of graphs and by the use of templates that they are intimately related to topology. Graph-theoretic properties, then, are basically topological properties and relations (for the *ante rem* structuralist, ultimately, they are relations). Thus,

graph-theoretic objects are the ones that instantiate those properties and relations, and only them.

According to de Clercq (2012, p. 663), a graph-theoretic counterexample to PII must, at least, work against a graph-theoretic version of PII. This is a reasonable demand. Thus, he presents two graph-theoretic versions of the principle (*ibid.* italics on the original):

> "***PIIgraph1:*** *Graph-theoretic objects are identical if they are indiscernible with respect to all (qualitative) graph-theoretic properties*",

and

> "***PIIgraph2:*** *Graph-theoretic objects are identical if they are indiscernible with respect to all (qualitative) graph-theoretic properties they have* within some graph G".

Now let us consider the example of dumbbell graph G1:[120]

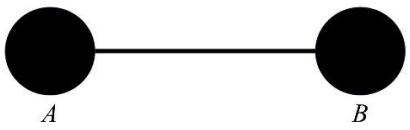

Fig. 1

One might think, as Button (2006, pp. 218-9) points out, that this graph presents a counterexample to PIIgraph2, given that the *distinct* nodes in this graph present the same relations, i.e., both *a* and *b* are only directed by one edge {*a*, *b*}; thus, are *indiscernible*.[121]

---

[120] All graphs presented here are also presented in Button, 2006, p. 218; de Clercq, 2012, p. 663; Ladyman and Leitgeib, 2008, p. 391; Ladyman *et al.*, 2012, p. 170.

[121] Button uses the term "indistinguishable" which he defines as objects present all the same properties and relations including to themselves (p. 218). Since no one but Ladyman and Leitgeib (2008) followed this terminology and considered it synonymous to "indiscernibles", we will put the term in the same basket with the

A dissenter, then, might argue that i) there are other non-explicit relations in play beyond those instantiated by the edges and that ii) they are discernible by their labels or by their edge's labels. Labeling, though, would be equivalent to naming, and we have stablished that this would be trivializing the question. Nonetheless, de Clercq claims that this is not the real reason why they are distinct. They are distinct because ii*) they are discernible by their appearance in other graphs with different relations (which is one kind of non-explicit relations). Consider the graph G2, yielded by the *addition* of the edge {*a*, *a*} (which – together with *removal* – is a legitimate operation in graph-theory practice for edges and nodes):

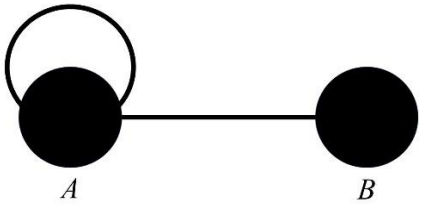

Fig. 2

Thus, *a* is discernible from *b* when considered in other graphs, i.e., *a* has three edges whereas *b* has only two if considered in G1 and G2 together.

This alternative presents two main problems. It relies on inter-structural relations – which would not properly reply to a counterexample for PIIGraph2 – and cross-graph identity. According to de Clercq, the former is only a problem if we opt for a certain reading of the structuralist beliefs.[122] He claims that it is not necessary for a structuralist to account for the individuation of mathematical objects of a determinate structure through the consideration of the properties in *any* structure, only in *some* structure. If we accept that, we may individuate *a* or *b* in graph G1 and find (viz. putatively transfer) them in graph Gn. This would allow us to find a discernible relation to make nodes distinct. This would be enough to make structuralism and cross-graph identity compatible.

---

other terms of the "discriminability" family. For a slightly different use of "indistinguishable" see Saunders (2006) p. 52.

[122] Ladyman and Leiteib see no problem with that. They concede that identity facts might be primitive, ungrounded, non-qualitative relations of identity or distinctness between nodes and edges.

This is, however, a way of saying that PIIgraph2 makes no sense, because what individuates an object in graph theory is the sum of all its properties in an existent higher structure, which is the relevant structure for this analysis. Thus, making PIIgraph1 the relevant principle to be disputed. De Clercq thinks this changing of scope from a single graph as a structure to a graph as a part of a larger structure – graph-theory Structure, with a capital "S", let us say – is innocent in the present discussion (*op. cit.*, p. 665). I beg to disagree, for this seems somewhat *ad hoc*. The relevant notion of structure must either be an ultimate graph-theory Structure, to which each graph is a portion or a sub-graph of, or each graph should be considered as a structure by itself. De Clercq seems to opt for the former, but his reasons for doing so are unclear. Reasons for not considering the single graph as the relevant structure for analysis should be presented. In any case, what he should not do is to jump from one notion to the other bending the rules of identity attribution to claim that intra-graph identity is the same as cross-graph identity (for a similar criticism, see DUGUID, 2016).

Furthermore, this inter-structural move is a double-edged sword. Consider graph G2*:

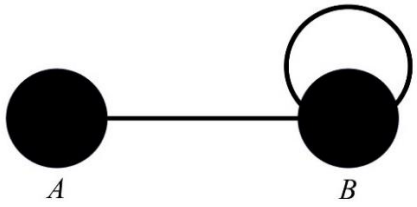

Fig. 3

Which is constructed from G1, but instead of the additional edge {*a*,*a*}, it has the additional edge {*b*,*b*}. The relations are equivalent to the ones in G2 (if we ignore the names, which we are not allowed to use anyway). Therefore, cross-graph, under the same Structure, we end up with the same relations for *a* and for *b*, namely, two irreflexive relations and one reflexive for each. This would be even worse, for we would have two indiscernible graphs within the (ultimate) graph-theory Structure, namely, G2 and G2*. This is a very unfortunate rebound[123]

[123] Interpret this as the rebound of the most preposterous weapon Hollywood makes us believe it worked as kids, the nunchaku. As soon as the attacker hits its opponent full force with the nunchaku, it bounces back at her just to hit her in the face. Totally useless attack.

effect that the appeal to cross-graph identity has for the PII defender. Furthermore, the above explanation relies on the very suspicious idea that by adding (or subtracting) an edge to (from) a graph, results in another graph with the same nodes. That is, the same nodes appear in different graphs, but under the same structure, i.e., graph-theory Structure, in de Clercq's interpretation. However, the notions of trans-graph identity – which is remarkably similar to cross-world identity in modal metaphysics – is not the kind of identity we are looking for (see section 2.3.). De Clercq's defence seems to be plausibly effective to PIIGraph1, but not to PIIGraph2.

Nevertheless, Ladyman and Leitgeib claim that the real problem appears when we consider unlabelled graphs. Consider the graph G:

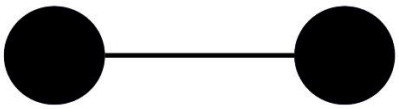

Fig. 4

In graph theory, according to them, the substitution of one unlabelled node by another unlabelled node, i.e., subtracting of one of the nodes then adding another, yields *the same graph G* – keep this in mind. Differently from the cases of labelled graphs, it is impossible to select any exclusive relation held by one node or the other that substituted it, given that we could not cross-graph identify them and reach for the distinguishing relations. Also, given that it is a symmetric graph with no irreflexive relations, we cannot even pull off an explanation similar to the account for imaginary numbers.

Yet, some PII champion might try to pull off a weak discernible property from the single edge available in the graph. To strangle this impulse, they ask us to consider the unlabelled graph G0:

Fig. 5

According to Ladyman and Leitgeib (p.392n7), the graph above is an even worse challenge for the PII champion. It presents distinct absolutely indiscernible and unconnected (thereby, unrelated) nodes. So, there is no way of pulling off a discerning relation, for there are no relations at all apart from the – according to them – primitive, non-qualitative, purely logical relation of distinctness (p. 392-3); an idea metaphysically suspicious, as Button (2006, p. 220) puts it, though logically possible, and which Ladyman and Leitgeib take to serve as a perfect representation for the cases of – supposedly – indiscernibles in QM. This seems to commit them to the existence of haecceities of nodes, but it does not, for, as they point out, substituting nodes results in the same graph. Were it the case that the nodes present haecceities, the graphs would differ. Clever move! They avoided the tag of being Platonists. At least, of a strong strand of Platonism, for they consider a weaker strand of Platonism which is – according to them – compatible with *ante rem* structuralism. So, Ladyman and Leitgeib assume a neutral position about this tag. Given that they do not develop further this point, I am not doing it either.

Nonetheless, they still must present an account for the distinctness (or identity) of these nodes. How can we know that there are two nodes rather than one, given that they present no distinguishable feature? A challenge posed by Button (2006) for the defendant of indiscernibles.[124] Ladyman and Leitgeib met this challenge by slightly changing the question to: how can we know that there is exactly a graph with two nodes and no edge? Their answer is somewhat disappointing. They claim that it is *established* that the graph has exactly two nodes and no edges. Therefore, by knowing the properties that were established for the graph,

[124] In (2006, p. 219), he makes a remarkably interesting point inquiring whether one would have to deny the identity of *m* and *n* when they present all the same properties and relations to everything (including colocation), until one is presented to the primitive identity facts that make them identical.

we are able to know what is established for the nodes. They claim that we could establish a one node graph and by legitimate operations, such as adding a node, obtain a two-node graph.

Although this answer seems to be acceptable, because since there are identities *by fiat*, such as the baptism of Cassius Clay as Mohammed Ali, surely there might be distinctness by fiat (e.g., something like a reverse baptism). But there seems to be an adhocness to this (remember Della Rocca's challenge). Someone could ask: why cannot we just, then, establish that a = b in G2 or say that *this* text is different from the text you are now reading? Because there seems to be rules to identity/distinction attribution beyond simple fiat, as the hundreds of pages created every year by researchers on necessity, analyticity and the a priori can attest.[125] One of these rules says that whenever there is a difference in features there must be distinctness (e.g., G2). It is expected that symmetrically there is a rule that says that whenever there is no difference to be found, there must be identity.

What might have engendered the impression that indiscernible graph cases are unproblematic in this regard is the use of unlabelled templates (drawings) representing the graphs by graph-theorists, instead of sets. Ladyman and Leitgeib claim that their position is tuned in with graph-theoretical practice, thus, it must be correct. This, however, is not a good explanation, de Clercq (2012) claims (followed by Duguid (2016)). According to de Clercq, although the Ladyman and Leitgeib's argument that unlabelled graphs are graphs that violate PIIgraph1 is based on passages of graph theorist William Thomas Tutte, that they count as an example of graph-theory practice, they ignore a vast literature that does not agree with that (see the list DE CLERCQ 2012, p. 668, especially the selected quote of an *introduction to Graph Theory* by Douglas West). Also, he argues, Tutte's view is that pure graph theory is concerned with isomorphism classes (or abstract graphs), not graphs as objects themselves (which are "officially" set-theoretical entities – as Ladymand and Leitgeib themselves (2008, p. 390)).

This seems to be the correct take on the issue: labelled graphs are set-theoretic entities, viz., mathematical objects that can be individuated and, thus, are within the scope of PIIgraph1. Whereas unlabelled graphs are isomorphism *classes* of these set-theoretic entities. They cannot

[125] Ladyman and Leitgeib's claim about identity and distinctness by fiat can be justified by those who think mathematical objects are created speech acts. Nevertheless, this must be argued for. Given the publishing date, I believe that they did not have this in mind though.

be counted as individuals. They are more appropriately understood as universal-like entities, therefore, lay outside the scope of PIIgraph1 (and probably any relevant version of PII). De Clercq's angle is less incisive. He claims that, although Ladyman and Leitgeib's is a plausible way of understanding graph-theory from its practice, they have not excluded interpretations such as the one above. If the above interpretation is the correct one, it seems that a summing defence is applicable in.

The problem with Ladyman and Leitgeib's view should be evident if one paid attention to the conceivability constraints of this counterexample, given that, when unlabelled, graphs lose the capacity of being described in set-theoretic language! Thus, they should cease to be graphs. However, they do not see this as a problem, but as something assumed as a presupposition for graph-theory practice. In their defence, it seems that there is malpractice going on among graph-theorists, perhaps a widespread abuse of notation. Although the templates help us visualizing the graphs, they are not the graphs, they are drawings that represent the graphs; or in a more optimistic view, they are instances of the graphs.[126] Nonetheless, Ladyman and Leitgeib acknowledge that there is an epistemological problem of explaining how we can know things from these templates (*op. cit.*, p. 395). They, however, dismiss it as a general problem beyond the scope of their paper, which I believe is exactly missing the point. They seem to be neglecting the epistemic part of PII.

Let us now look at Button's (2006) hybrid solution that is "*very similar in spirit*" to de Clercq's proposal (in his own words, 2012, p. 671).[127] Button proposes the distinction between basic structures (**B**) and construed structures (**C**). **B**s are realistically interpreted as the mathematical objects, whilst **C**s are "*treated eliminativistically as follows: talk about **C** is read paraphrastically as universal generalizations over the systems of positions in **B** which exemplify the structure **C***" (2006, p 220). In other words, the nodes and edges in structures of a **C** are not (mathematical) objects, but generalizations of (mathematical) objects – classes of objects, better put. Thus, they are not under the scope of PII. Although this solution is, on one

[126] This view, however, would necessarily make them manifest non-graph-properties (e.g., sizes, distances, colours, etc.). Then, PII would be safe.

[127] Button attributes the first consideration of this solution to Parsons and Keränen, though both rejected its viability. According to him (2006, p. 221), the solution is viable if considered in a much lower level than they considered.

hand, elegant; on the other, it lacks simplicity, since it postulates a whole new level of structures. Furthermore, it chooses to interpret this new structural level as something without objecthood. Why cannot classes be understood as mathematical objects of higher levels? Why stop at level 1? This seems somewhat *ad hoc*, as much as de Clercq's changing the interpretation of graphs from viewing them as structures to seeing them as parts of structures or not even that, since the difference between Button's solution and his – in his own words – is that the latter is not committed to the interpretation of graphs, labelled or unlabelled, as structures or even part of structures. This is unsettling. It seems that there is something incomplete in this picture. I intend to present a more complete picture by other means.

However, let us look at one last objection from Button against the view that there are distinct indiscernibles. Button claims that accepting such entities requires an unusual notion of objecthood that still must be provided by the proponents of such idea (*op. cit.*, p. 219-20). The nodes in those examples hold no relation to each other although they exist in the same template (viz., the same drawing) without bearing any spatial relation to each other. Button notices that this is inconceivable. The irreflexive relation of distance should be held by them were they spatial objects. Also, any analogy to conventional number systems would fail, since they also bear irreflexive relations, as saw in the previous section.

There is a challenge of providing criteria for the objecthood of such mathematical objects that must be addressed. Ladyman and Leitgeib unsatisfactorily addressed it by saying that we should not expect objects such as these (or quantum particles) to present familiar criteria of objecthood, their objecthood can be attested by mathematical practice: "*Graph theorists quantify over these objects; they consider identity claims for them; they count them; in short, they regard them as objects proper*" (2008, p. 395). Also, as Button himself admits (2006, p. 220), the usage of distinct indiscernibles is logically coherent (though metaphysically suspicious). Ladyman and Leitgeib insist in their point: why should these objects present more than mathematical and logic conformity to a notion of objecthood?

Well, first because, in the way it was presented by them, this treatment is not perfectly in tune with the mathematical practice, as de Clercq (2012) showed. Second, it is not as if mathematicians were never wrong, there might be a case of malpractice with the notations going on or even mathematicians might be working with something they do not clearly understand yet (assuming there is no consensus about what a mathematical object is). Third, there must be

some minimal common content between mathematical objecthood and ordinary objecthood, this seems to be what is missing.

Although Button and de Clercq's reasoning are cogent enough, I believe there is a better and more complete way of defending PII for graphs. I believe we can go a firmer step further pushing in the same direction that Button is pushing if we do it with a different technique. My suggestion is to use a strategy similar to the Identity Defence proposed by Della Rocca. At the same time, *circa* 2005, parallel to this discussion about graphs, Della Rocca developed an argument against the opponent of PII who relies on brute facts of identity and distinctness of objects such as the spheres in Black's example. This argument applied to the context of graphs would run as follows: The proponent of the PII counterexample in graphs requires that the distinctness between the nodes to be primitive. An alternative would be to derive it from a one node graph through the addition of a node resulting in a two nodes graph, but it would still depend on the primitive identity of at least one node (the same outcome would come by deriving it by subtraction from a three nodes graph too, requiring two primitive identities instead). If this is granted, how can he assure that there are only two nodes in the graph in question? Could not it be that, instead of two nodes, there are actually twenty nodes? In other words, could not there be ten nodes with all the same features (including those pictorially conceived) in the place we thought there was only one node? This seems absurd. Let me stress that we are not dealing with a case in which a node bear ten different names, like a Frege's Puzzle on steroids, or cases such as the statue and the piece of clay, in which there are different kinds of things occupying the same space. In this case there are ten things of the same kind bearing numerical distinction to each other whilst occupying the same place in the template, possessing the same features that the other has. We are back at Button's inquiry "why is there not only one node instead of two, then?", but in a different proportion. The strategy of Ladyman and Leitgeib was changing the question to "How can we know that there are exactly two nodes?" and answering it with a "because we are telling you so". To be fair, they said:

> *So the primary question to pose is the one about how are we able to know graphs like* [G0] *to exist. But that question can be answered in more or less standard ways: because graph theory postulates it and we have every*

> *reason to believe that the basic principles of graph theory are coherent;* because we can generate graphical templates that indicate so; and so forth." (2008, p. 394. Emphasis added)

First, as we previously saw, Ladyman and Leitgeib's views are not quite in tune with graph theory principles. Second – and more important in this context – they are relying on the templates, which are drawings, i.e., graphic (maybe topological) representations of what nodes are. The intuition that G0 *has two and only two nodes* hinges on the fact that we are looking at two points drawn on the template. This means that it somewhat relies on spatial features of some sort displayed in the template. Otherwise, it would be unintelligible, that is, inconceivable. We can apply an inconceivability defence here. If one does not appeal to templates, the scenario is inconceivable, since it cannot be linguistically (set-theoretically) constructed, given that the nodes are unlabelled. Thus, the inconceivability defence reduces from a trilemma to a dilemma: either the scenario is inconceivable, or it is pictorially conceived. If it is pictorially conceived, it must display spatial features. Then, the discerning features surface and a discerning step can be applied.

Nevertheless, the opponent might not accept that there are permitted discerning features. Let us assume that this is the case. In this case, if these nodes present no properties nor relations, nothing prohibits that there are twenty or two hundred nodes represented by these drawings. Della Rocca's challenge is on. The opponent might appeal to simplicity. She may argue that there is no reason to believe that there are more than two nodes being represented in G0. To which I would reply that there are no reasons to believe that there is not just one node drawn twice either; this would be even simpler. In the case of G1, nothing prohibits it from being a representation of a self-connected node represented in such a way as to avoid loops, a form of drawing economy (remember the Muller-Hawley debate, see section 3.3.2.). Unless there is a further reason for not assuming that there is more than one object there, I see no reason to believe there are.

She may, then, appeal to our individuating practices. It is common and reasonable to think that, when there are two available representations, there are two and only two objects represented. However, as we just saw, it is also reasonable (viz., more theoretically virtuous) to

think otherwise. Furthermore, this would be interpreting things backwards, as Della Rocca (2008) has shown. Our individuating practices should be subject to the facts, not the other way around. Thus, we should be able to correctly grasp the metaphysical facts before jump to conclusions about individualization (see the Della Rocca-Jeshion debate in section 3.2.1.1.).

One might appeal to an epistemic difference: in the twenty nodes case, there is no way to tell them apart, whereas in the two nodes case there is (remember the additive properties of HAWLEY, 2009). However, this would resort to different properties, namely, graphical ones, which boil down to spatial ones (see section 2.4.4.). Thus, the discerning defence is in play. She might insist, however, that this is not permitted, because graphs bear no spatial properties; actually, the nodes in G0 bear no properties at all apart from logical ones like existence, identity, and distinctness, according to Ladyman and Leitgeib. If they lack any feature whatsoever, then they can be neither imagined nor communicated to me. Thereby, why should I believe in their existence? It seems we are back to the inconceivability stalemate.

Della Rocca, as we already saw in section 3.2.1., proposes that the PII opponent could differentiate the twenty objects case from the two objects case by appealing to a conceptual truth that everyone should accept, what Jeshion (2006) would later call the Weak Lockean Principle (WLP): *There cannot be two or more indiscernible things with all the same parts in precisely the same location at the same time*. Everyone should accept it, but there must be a reason for that, apart from the opinion of Della Rocca and the apparent obviousness of it. The reason seems to be intimately linked with relations of parthood. Della Rocca makes this reason clear by means of a challenge: why is it ok for two objects to partially overlap, but not totally overlap? For example, why Siamese twins, who share considerable portion of their bodies, are considered two distinct individuals and I am not considered a different individual from my copy that shares all properties places and times with me?[128] Or, even more appropriate for this discussion, why two sets that partially overlap are considered two distinct sets, whilst sets that completely overlap are not? The answer clearly relies on the parts that they do not share, be it spatial regions or elements. Thus, there is some relations playing roles in this distinction. If the

[128] If the reader is thinking about Shyamalan's *Split* film as a counterexample, I must remind you that the different personalities do not manifest in the same body at all times, thus, they do not share temporal features.

nodes lack relation, there seems to be no reason to adopt WLP. Besides, this discussion concerning graphs might be even worse for the PII opponent than the cases of physics. Hawley (2009) claims that, if there were more than one sphere co-located, we could apply mental operations of division or sum and see that it would have controversial results, e.g., if we divide these twenty spheres across the scenario, how would their mass be distributed? This kind of question would also emerge in graph cases but there would be no controversial results since these cases lack these additive properties. However, this is controversial in itself, because then the nodes could be seen as an infinite source of new nodes in a mathematical Ockhamnian nightmare.

It seems, then, that to avoid accepting the absurd case of twenty node graph, the opponent of PII must admit that spatial relations play some role in the attribution of identity or distinctness of nodes. At least for unlabelled graphs, since they solely rely on the templates. By admitting this, the opponent opens the possibility for the defender to pick weakly discerning properties, thus, solving the quarrel. Again, the inconceivability defence works, the discerning defence works, and even a summing defence can be applied if necessary.

o o o

Summing up the view about the mathematical counterexamples to PII, I might say that, to present more solid answers to these questions, mathematicians would have to reach some agreement about what mathematical objects are in the first place. Here I dealt with a very narrow set of possible interpretations of these objects, basically just the (*ante rem*) structuralist view, which was the most discussed in the literature. Thus, it is not at all warranted that these defences would work for every conception of Mathematics. There is still work to be done, but I believe that the pro PII side has a better case from the point of view of structuralism.

6. Conclusion

In this tome, I hope to have fulfilled the initially established objectives. In chapter 1, an encyclopaedical presentation of the Principle of Identity of Indiscernibles was displayed. I briefly discussed its different versions, the historical context in which the debate over the

principle gained momentum, the relevance of the verdict for its truth or falsehood in metaphysics and other areas, and the implications of this verdict for these areas. Also, I presented a more complete classification of the arguments against PII than the one available in the contemporary literature. I classify them in the following classes: the classical dispersal arguments, the arguments from almost indiscernibles, the arguments *de fictis*, the derived dispersal arguments, and the special dispersal arguments. Then, I did the same with the defences used to preserve PII. I classified them using Hawley's original division of defences that contained the identity defence, the discerning defence and the summing defence, then I added one more taxon, the inconceivability defence, which includes my own proposal of defence. I divided these taxa in families of defences to allow further classification of lower taxa. Thus, I hope that I have mostly concluded my secondary goal, that is, (B) to present an encyclopaedic introduction that could map the issues involved with the topic to a sufficiently philosophically literate reader, so as she can follow the discussion. This task could only be completed in latter chapters, where goal (C), namely, to clarify the complex notions involved with PII, is achieved.

In chapter 2, most of the work to complete (C) was done. I first dealt with issues concerning which kind of modality is in play when discussing PII. I distinguished between nomical, metaphysical and logical modalities and settled that what is being contested with most arguments against PII is its metaphysical necessity instead of its logical necessity, as it is commonly spelled (the exceptions are QM cases, which questions its actual truth and its physical necessity).

Then, I presented the view I adopted to talk about possible worlds, namely, ersatzism, just as far as it was necessary to introduce my views about the PII debate. From this point onwards, I clarified which kind of identity is supposed to be established by PII separating it from other adjacent notions, such as transtemporal identity, transmundane identity, and criteria of identity, that are not relevant for this debate, explaining why these are not relevant for the debate. Thus, I settled that the relevant notion of identity in this debate must be an intraworld relation between two objects.

Establishing objects as the entities involved in this relation required from me a clarification of what counts as an object, since this was not clear at all in the literature. Some authors talked about PII applied to particulars, others to individuals, and some to any entity whatsoever, at all. I settled for the most contemporary term "object" because it carries with it

very little metaphysical load differently from “individual”, for instance. However, it does carry some load. It is not free of metaphysical weight like Muller (2015) intended, for example, nor purely logical as Ladyman and Letgeib (2008) think. These objects must present some basic metaphysical characteristics such as spatial and chronological constrains and obey classical logic principles.

With objecthood sufficiently defined I advanced to talk about the thorniest issue in the debate, namely, properties. The core of most disagreements about PII is due to a disagreement about which properties are being analysed to establish identity and which properties are allowed to count as discerning properties in the putative counterexamples to PII. One of the most unclear issues in the literature up to the 21st century was whether relations (and relational properties) should be under the scope of PII. The literature has already decided that they should, but by clarifying what objects are (in this discussion), it became clearer that relations must be under the scope of PII, given the relational nature of space. Furthermore, a careful analysis of the scenarios’ descriptions supposed to work as counterexamples revealed that they require relations to be established. Thus, to leave the terms “properties” and “relations” to be used in specific contexts, respectively, when talking about intrinsic characteristics of objects usually expressed by monadic predicates and extrinsic characteristics held among other objects usually expressed by polyadic predicates, I used the term “features” to talk about a more general class of characteristics displayed by objects containing both properties and relations.

Another aspect of the features that had to be clarified is related to their discernibility potentials. Whenever two (or more) objects display a difference in properties, one can say that they are absolutely discernible, thus, they are also clearly distinct. However, they can fail to display such features, yet be discerned by the relations they bear, thus being relationally discernible, which in turn, implies that they are distinct. This relational discernibility can be further divided into relative discernibility and weak discernibility. The latter is the minimum level of discernibility that two objects must display to be considered distinct. Two objects are weakly discernible when they present symmetrical non-reflexive relations to each other. The most ubiquitous weakly discerning relations one can find are spatial features, thus, I argued that the whole debate about PII could basically be reduced to a debate about what spatial features (whether interpreted as relations or properties). Of course, some mathematical cases and QM cases would escape this prediction, that is why I called them special cases.

With all of this settled, I could turn to the debate about which version of PII is worth saving. If only properties were permitted, we would have a very strict notion of PII that could only be said to concern individuals (in Hawley's sense, see section 3.3.), not objects (or possible objects). This would result in an uninteresting metaphysical principle – or at least not as metaphysically interesting as a version about objects. Yet, it is worth stressing that if this is the version of PII in debate, there is a vast number of counterexamples. Thus, let it be clear, a PII about individuals is doomed by any of these counterexamples. On the other hand, if we permit any feature to be under the scope of the principle, this would trivialize it, for there are some features such as identity features and distinction features. Therefore, a balance must be established. Most philosophers would be content with a version of PII that allows only pure properties, that is, it permits relations but only if they are not held between external objects. I argued, in accordance with others like Rodriguez-Pereyra and Muller, that this would still be too restrictive. Rodriguez-Pereyra (2006) argued for a more permissive version which he called a non-trivializing version. Non-trivializing properties (here features) turned out to be difficult class to define and I settled for a more restrictive view than that Rodriguez-Pereyra argued for and in line with the one proposed by Muller (2015). This version excluded identity features, features using names, and features that could break the symmetry of the thought experiments designed to serve as counterexamples. If a PII in this version can be saved, certainly a version accepted by Rodriguez-Pereyra can.

In this chapter, I also briefly discussed the metaphysical nature of the features concluding that, if they are tropes, questioning PII makes no sense; if they are immanent universals, there is an identity defence that can explain Black-like counterexamples to it; and that there is very little to say about them if they are instantiations of classic universals or they are universals in the nominalist reinterpretation of them.

In chapter 3, I dealt with vestigial questions of (C) and aimed at my primary objective, namely, (A) defend PII from the putative counterexamples. First, I presented defences from the Identity family, namely, Hacking's defence and O'Leary-Hawthorne's defence. Basically, both defences argued that the phenomena interpreted as two spheres in Black-like scenarios are, in fact, only one object perceived twice, which can be viewed as an attempt to fulfil objective (A2), i.e., to show that the scenario described is impossible to be build. The former justified this claim by arguing that the space the spheres exist in is, in fact, cylindrical, whereas the latter

argued that they are actually a bi-location of the amalgamation of immanent universals. I presented objections to these claims and concluded that they are inadequate defences. In short, because Hacking's only works for a limited number of possible scenarios, whereas O'Leary-Hawthorne's metaphysical assumptions about what properties are present obscure and implausible mechanics in the interactions among them.

Then, I went for discerning defences, namely, Della Rocca's challenge and Muller's version of a discerning defence. Both try to complete objective (A1), i.e., showing that the scenario presented does not yield the intended conclusions concerning its features. I argued that both are adequate defences and that their respective objections I collected – and enhanced when possible – are not effective after closer inspection. While Della Rocca adopts a passive strategy to show that there must be spatial discerning features in Black-like scenarios, leaving to the opponent the decision as to whether he wants to bite the bullet and accept that there are indiscernibles in the scenario, but also importing the possibility of multiple co-located indiscernibles, or to give up the scenario; Muller, on the other hand, adopts a more active approach showing that there is a permitted weakly discerning feature, viz., the relation of distance, that the spheres in any Black-like scenario display. Moreover, he goes further and also shows that, in QM cases, which are thought to be different from Black-like scenarios, either present other weakly discerning relations, viz., spins, or can be reimagined as Black-like scenarios with no loss to their existence conditions.

Nevertheless, since this Quantum talk is always uncertain, alternative defences are always welcome. Therefore, I reached for the summing defence, a defence elaborated to firstly deal with those cases. More precisely, I went for Hawley's version of it. She argues that, in both QM cases and Black-like cases, the alleged two objects in the scenario are in fact only one object. Not like the Identity defences that saw the two phenomena as the same thing – aiming for (A2) –, but rather as both being parts of one minimal object lacking individuality themselves – aiming for (A1). Hawley argued that this defence is preferable to the ones in the Discerning family, because it is more theoretically virtuous, since Discerning defences are not always available. Analysing further the Discerning defences we saw that they are available, at least, in the cases that Hawley judged them not to be and adopting them turned out to be more virtuous. Thus, I concluded that the summing defence is not appropriate to defend PII against Black-like cases

(or cases that can be seen as Black-like scenarios), though it might be useful to have them as a last resort for cases that might lack spatial features (or any discerning features at all).

At last, but not least – actually, the contrary of least – the inconceivability defence was presented and developed here. This defence aimed to show that a counterexample to PII is not possible because it depends on the way we make sense of modalities. To make sense of the possibility of such scenario we need to construct a possible world that contains it. However, the means to conceive this possible world – linguistically or pictorially – make it clear that there are inevitable inconsistencies being described. Thus, either (A1) the scenario conceived presents discerning features or (A2) the scenario is inconceivable. This defence seems to be the best, for it naturally contains the discerning defence as a step and might even contain the summing defence, whenever the understanding of what an object is escapes the ordinary, whilst it opens a door to the possibility that the opponent of PII is committing a categorial error when claiming that spatial things are not discernible.

To test this defences, in chapter 5, I considered some putative counterexamples from mathematics, namely, the case of imaginary numbers and the cases of graphs. The analysis of these cases – as it was with the QM cases – turned out to show that they might not be so special after all, for the inconceivability defence and the discerning defence worked with them as well as for ordinary Black-like cases. Additionally, the summing defence may come handy if the understanding of the objects in question is not quite favourable to the discerning defence.

Thus, as a general conclusion to this dissertation, we can say that, in the battle for PII, the champions of the principle have four strategies at their disposal. Hacking's identity defence would be equivalent to engage in battle with the opponent army in their terrain with all troops charging in one direction straight to the enemy commander, killing all the troops in this line – solving the problem in one possible world. However, this would leave the flanks open to attacks of the opponents – there are possible worlds in which the counterexamples are not the ones that the identity defender thinks they are. O'Leary-Hawthorne's identity defence would be the same as letting loose the elephants on the enemy. This may destroy the enemy lines, but it is very probable that the beasts would run amok and destroy your lines too and the domains you are trying to defend too – the chaos generated in the metaphysical system by the additive properties.

Muller's discerning defence would be equivalent to engage in battle in the opponent terrain with extreme care, surrounding it, pinching, piercing their lines looking for the best

breach. Adopt Della Rocca's challenge is to circle the opponents, taunting them to do the first move just to respond more fiercely in the flanks they uncover. In any case, when the breach is found – the existence of weakly discerning features – the whole defensive line breaks and the battle is won.

The summing defence should not be viewed as an engagement in battle, but as a change of terrain – interpret the supposed two objects as features of a larger single object. By adopting it we would be retreating in search of a higher ground to engage whenever the conditions are favourable for PII. It works, but it must be viewed as it is, a retreat and should be used only if necessary. For we would defeat the opponents but only with great cost to our domains.

Finally, the inconceivability defence should be viewed as a projectile strategy, akin to the roman *pilla* throwing strategy. While marching towards the enemy lines, the legionaries throwed projectiles (*pilla*), just before engaging in battle shield to shield. Sometimes, this was effective enough to make the enemy lines break and disband their troops – the scenario is shown to be inconceivable. If the troops maintained their grounds, they would do so in disarray and weakened, making it easier for our troops to look for a breach in their lines – the discerning defence – or, in the worst scenario, leaving the opportunity to retreat for a higher ground – adopting the summing defence – without the risk of being chased when doing so. The inconceivability strategy is the best course of action.

References:

ADAMS, R. M. ‘Primitive Thisness and Primitive Identity’. *Journal of Philosophy*, 76, p. 5-26. 1979.

_____. “Theories of Actuality”. *Noûs*, v. 8, n. 3, p. 211-231, set. 1974.

ARMSTRONG, D. M. “The nature of possibility.” *The Canadian Journal of Philosophy*, v. 16, n. 4, p. 575-594, 1986.

AYER, A. J. The Identity of Indiscernibles. In: _____. *Philosophical essays*, London: Macmillan. 1972.

BABER, H. E. *The Trinity*: A philosophical investigation. London: SCM Press, 2019.

BAXTER, D. L. M. Identity, Discernibility and Composition. In: Cotnoir, A. J.; _____ (Ed.) *Composition as identity*, Oxford: Oxford University Press. 2014. p. 244-54.

BERTO, F.; SCHOONEN, T. ‘Conceivability and Possibility: some dilemmas for Humeans’. *Synthese*, 195, p. 2697-715. 2018.

BLACK, M. ‘The identity of indiscernibles’. *Mind*, 61, p. 153–64. 1952.

BUENO, O. ‘Identity in physics and elsewhere’. *Cadernos de História e Filosofia da Ciência*. Série 4, v. 2, n. 1, p. 93-105. 2016.

BUTTON, T. ‘Realistic structuralism’s identity crisis: a hybrid solution’. *Analysis*, 66, 3, p. 216-22. 2006.

COTNOIR, A. J. Composition as Identity. In: _____; Baxter, D. L. M. (Ed.) *Composition as identity*, Oxford: Oxford University Press. 2014. p. 3-23.

CROSS, C. ‘Max Black on the Identity of Indiscernibles’, *Philosophical Quarterly*, 45, p. 350-60. 1995.

_____. ‘Causal Independence, the Identity of Indiscernibles, and the Essentiality of Origins’, *The Journal of Philosophy*, 106, 5, p. 277-291. 2009.

_____. ‘Brute facts, the necessity of identity, and the Identity of Indiscernibles. *Pacific Philosophical Quarterly*, 92, p. 1-10. 2011.

DE CLERCQ, R. ‘On some putative graph-theoretic counterexamples to the Principle of Identity of Indiscernibles’, *Synthese*, 187, p. 661-72. 2012.

DAINTON, B. F. *Time and Space*. New York: Routledge. 2014.

DELLA ROCCA, M. ‘Two Spheres, Twenty Spheres, and the Identity of Indiscernibles’. *Pacific Philosophical Quarterly*, 86, p. 480–492. 2005.

_____. ‘The Identity of Indiscernibles and the Articulability of Concepts’. *Linguistics and Philosophical Investigations*, 7, p. 29-36. 2008.

DENNETT, D. C. ‘The unimagined preposterousness of zombies’, *Journal of Consciousness Studies*, 2, 4, p. 322-6. 1995.

DUGUID, C. ‘Graph theory and the identity of indiscernibles’. *Dialectica*, 7, 3, p. 463-74. 2016.

FORREST, P. "The Identity of Indiscernibles", *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/win2016/entries/identity-indiscernible/

FRENCH, S. ‘Hacking Away at the Identity of Indiscernibles: Possible Worlds and Epstein’s Principle of Equivalence’. *The Journal of Philosophy*, 92, p. 455-66. 1995.

_____. ‘Why the Principle of the Identity of Indiscernibles is not Contingently True Either’. *Synthese*, 78, p. 141-66. 1989.

FRENCH, S.; KRAUSE, D. *Identity in Physics*: A historical, philosophical and formal Analysis. Oxford: Oxford University Press. 2006.

GENDLER, T. S.; HAWTHORNE, J. Introduction: conceivability and possibility. In: _____. *Conceivability and Possibility*. Oxford: Oxford University Press. 2002. p. 1-70.

GLEZAKOS, S. Can Frege pise Frege’s puzzle?. In: Almog, J.; Leonardi, P. (Ed.) *The philosophy of David Kaplan*. Oxford: Oxford University Press. 2009. p. 202-8.

HACKING, I. ‘The Identity of Indiscernibles’, *Journal of Philosophy*, 72 (9), p. 249-256. 1975.

HAWLEY, K. ‘Identity and Indiscernibility’, *Mind*, 118, p. 101-9. 2009.

HAWLEY, K. ‘Weak Discenibility’, *Analysis*, 66, p. 300-3. 2006.

HAWTHORNE, J. Identity. In: *The Oxford Handbook of Metaphysics*. Loux, M. J. and Zimmerman, D. W. (eds.). Oxford: Oxford University Press. 2003.

JAGO, M. *The Impossible*: An Essay on Hyperintensionality. Oxford: Oxford University Press. 2014.

JESHION, R. ‘The Identity of Indiscernibles and the co-location problem’. *Pacific Philosophical Quarterly*, 87. p. 163-76. 2006.

JORATI, J. ‘The contingency of Leibniz’s principle of identity of indiscernibles’. *Ergo*, 31, 4, p. 899-929. 2017.

JUHL, C; LOOMIS, E. Analiticity. London; New York: Routledge, 2009. New problems in Philosophy series.

KANT, I. *Crítica da Razão Pura*. Trad. Manoela Pinto dos Santos e Alexandre Fradique Morujão. Lisboa: Fundação Calouste Goulbekian, 2001.

KAPLAN, D. Demonstratives. In: Almog, J; Perry, J.; Wettstein, H. (Ed.) *Themes from Kaplan*. Oxford: Oxford University Press. 1989. p. 481-563.

KOONS R. C.; PICKAVANCE, T. *The Atlas of Reality*: A comprehensive guide to metaphysics. Oxford: Wiley-Blackwell, 2017.

KRAUSE, D.; ARENHARDT, J. R. B. 'Is Identity Really So Fundamental?'. *Foundations of Science*, 24. p. 51-71. 2019.

LADYMAN, J.; LEITGEIB, H. 'Criteria of identity and structuralist ontology'. *Philosophia Mathematica*, 16, III, p. 388-96. 2008.

LADYMAN, J.; LINNEBO, O.; PETTIGREW, R. 'Identity and discernibility in philosophy and logic'. *The Review of Symbolic Logic*, 5, 1, p. 162-86. 2012.

LADYMAN, J. 'Mathematical Structuralism and the Identity of Indiscernibles'. *Analysis*, 65, p. 218–21. 2005.

LEIBNIZ, G. W. Letters to Clarke. In: Philosophical Papers and Letters, ed. L. E. Loemker. Dodrecht: Kluwer Academic Publishers. 1969.

LEWIS, D. K. *On the Plurality of Worlds.* Oxford: Blackwell, 1986.

LOUX, M. J. *Metaphysics*: A contemporary introduction. New York: Routledge. 2002.

MACBRIDE, F. 'What constitutes the numerical diversity of mathematical objects?'. *Analysis*. 66, 1, p. 63-9. 2005.

_____. "Relations", *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.) URL = https://plato.stanford.edu/archives/win2016/entries/relations/

MENZEL, C. 'Sets and worlds again'. *Analysis*, 72, 2, p. 304-9. 2012.

MUGNAI, M. "Identidade dos Indiscerníveis". In: Abbagnano, N. *Dicionário de Filosofia*. Translated to Portuguese by Alfredo Bosi and Ivone Castilho Benedetti 5th edition. São Paulo: Martins Fontes. 2007. p. 614

MULLER, F. A. 'The rise of relationals'. *Mind*, 124, p. 200-37. 2015.

NAGEL, T. 'How it is to be a bat?'. *The Philosophical Review*, 83, 4, p. 435-450. 1974.

O'LEARY-HAWTHORNE, J. 'The Bundle Theory of Substance and the Identity of Indiscernibles', *Analysis*, 55, p. 191–196. 1995.

PLANTINGA, A. The Nature of Necessity. Oxford: Clarendon Press, 1974.

QUINE, W. V. Two dogmas of empiricism. In: _____. From a logical point of view. Cambridge: Harvard University Press, 1961.

_____. 'Grades of discriminability'. *Journal of Philosophy*, 73, 5, 113-6. 1976.

RODRUIGUEZ-PEREYRA, G. *Leibniz's Principle of Identity of Indiscernibles*. Oxford: Oxford University Press. 2014.

_____. How not to trivialize the identity of indiscernibles. In: ed. Strawson, P. F.; Chakrabarti, A. *Universals, Concepts and Qualities*: New Essays on the Meaning of Predicates. Aldershot: Ashgate. 2006. p. 205-23.

_____. ‘The bundle theory is compatible with distinct but indiscernible particulars’. *Analysis*, 64, p. 72-81. 2004.

RUSSELL, B. *An Inquiry into Meaning and Truth*. New York: Routledge, 1995.

RUSSELL, G. *Truth in Virtue of Meaning*: a defense of the analytic/synthetic distinction. Oxford: Oxford University Press, 2008.

SAUNDERS, S. ‘Are quantum particles objects?’. *Analysis,* 66, p. 52–63. 2006.

SORENSEN, R. The art of the impossible. In: GENDLER, T. S.; HAWTHORNE, J. (Ed.) *Conceivability and Possibility*. Oxford: Oxford University Press. 2002. p. 337-68.

STRAWSON, P. F. *Individuals*: An essay on descriptive metaphysics. Norfolk: University Paperbacks, 1959.

SWIMBURNE, R. ‘Thisness’, Australasian Journal of Philosophy, 73, 3, p. 389-400. 1995.

WEATHERSON, B. “Intrinsic vs. Extrinsic Properties”, *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.) URL = https://plato.stanford.edu/archives/fall2008/entries/intrinsic-extrinsic/ 2006.

ZIMMERMAN, D. ‘Distinct Indiscernibles and the Bundle Theory’, *Mind*, 106, p. 305-09. 1997.