UNIVERSITE PARIS DIDEROT (Paris 7)

Doctorat en Epistémologie, Histoire des Sciences et des Techniques

UNIVERSITAT DE BARCELONA (UB)

Faculty of Philosophy

DANIELE MOLININI

TOWARD A PLURALIST APPROACH TO MATHEMATICAL EXPLANATION OF PHYSICAL PHENOMENA

July 2011

Supervisors: Marco Panza / José Díez

Members of the jury:

José Díez

Jacques Dubucs

Carl Hoefer

Marco Panza

Christopher Pincock

Ivan Smadja

to Silva, Franco, Luca and Bart

Acknowledgements

It was saturday 27 March 2010. We met each other in the area where the escalators reach their topmost point in the lobby of the Gare de l'Est, in Paris. There is a coffee shop there (Pains à la Ligne) and the plan was to have a coffee and talk about philosophy of mathematics and mathematical explanations. At the end of our conversation, before leaving the Pains à la Ligne, he added: 'At the end of a research process, questions are more important than answers'.

I suppose that he was referring to 'good' (or fruitful, say) philosophical questions. Otherwise, my thesis would be a masterpiece. Whether one single good question has been raised during these four years, this question is the result of a long process in which many people have been involved and which has begun a couple of years ago. I've started my training in history and philosophy of science in 2006, when I began my Master LOPHISS at the Université Paris 7 Denis Diderot. There I first learnt what doing history and philosophy of science meant. For this I have to thank my teachers there, especially Régis Morelon, who was my supervisor during my Master Thesis. In October 2007 I began my Ph.D. in philosophy of mathematics, in the same university and under the supervision of Marco Panza. During these four years I had the great opportunity to discuss my ideas with outstanding philosophers, in various situations and places. All these scholars have paid attention to my ideas and I have learned so much from theirs. I have to thank David Rabouin and the Rehseis Group, and also Pascal Crozet and the Chspam Group. The working and friendly atmosphere that I experienced with the members of these groups and the depth of the connections, both professional and personal, that characterizes their activity have been essential to my training. All the researchers of these groups have helped me during my investigations, they have encouraged my research and, more importantly, they have shared with me very nice tea/coffee breaks (someone once told me that it's during the coffee breaks that the most fecund philosophical exchange is carried out). Thank your for all that. I have to thank Virginie Maouchi, of the Laboratoire Sphere, and Sandrine Pellé of the Ecole Doctorale 400 at Paris 7, for your 'administrative stoicism', and for having solved so many bureaucratic questions concerning my Ph.D. and the funding for my participations in conferences and workshops. I also have to thank the LOGOS Group in Barcelona, which I have visited in 2009/2010. During my visiting period I had the occasion to discover another stimulating group, new philosophical topics, great philosophers and very nice persons. I am indebted to all the LOGOS members that I got the chance to meet, and to the LOGOS administrative staff. I am especially indebted to Oscar, Paco, Pablo, Genoveva, Vera, José and Carl for their help, their friendship and their advice during my stay in Barcelona.

My special thanks go to a number of colleagues (and friends) who have commented (with their huge patience!) my paper-drafts and have discussed with me their views (by email, during a conference or just in front of a coffee). With apologies for forgetting anyone, those people are: Paolo Mancosu, Chris Pincock, Henk De Regt, Steven French, Michael Friedman, Dennis Dieks, Mark Steiner, Bas Van Fraassen, Sebastien Maronne, Carl Hoefer, Michael Friedman, José Ferreirós, David Rabouin, Carlos Álvarez Jiménez, Susana Berestovoy, Ken Manders, Sorin Bangu, Davide Rizza, Robert Batterman, Alan Baker, Michel Paty, James Ladyman, Mic Detlefsen, Andrea Sereni, Silvio Bergia, Dan Zeman, Eric Weber, Davide Crippa, Maria Alessandra Mariotti, James McAllister, Annalisa Coliva, the newborn Leonardo Panza (for having provided me with a good argument against the existence of numbers), Sabina Leonelli, Sara Confalonieri and Adán Sus. Furthermore, I have to thank the participants to the following workshops and conferences for their criticisms and for having improved the quality of my work: *Congrés organisé par la Société Française d'Histoire des Sciences et des Techniques et la Societá Italiana di Storia delle Matematiche* (Paris, France), *Segundo Congreso Internacional de Estudios Clásicos en México* (Ciudad de México, México), *19th Novembertagung on the History of Mathematics* (Holbæk, Denmark), *Sixth Conference of European Research in Mathematics Education* (Lyon, France), *Conference on Mathematical and Geometrical Explanations in Physics* (Bristol, UK), *Workshop Explanation, Indispensability of Mathematics, and Scientific Realism* (Leeds, UK), *Workshop Understanding and the Aims of Science* (Leiden, The Netherlands), *Conferencia de Graduados de la SLMFCE* (Granada, Spain).

There are two people that deserve very special thanks. The first is José Diéz, my second supervisor. Without him I wouldn't have discovered the stimulating atmosphere of the LOGOS group. Moreover, his extreme clarity in commenting various parts of this dissertation have significantly improved my work. Especially during the last months, I have profited from his advices and his philosophical skills. José has picked out various weak points of my thesis and his advice has been essential to give to this work a more uniform and coherent structure. I hope to have put into practice his suggestions. For this (and for other more 'pragmatical' reasons) I have to express my gratitude to him. The second person (and probably I should have mentioned him in the first lines of these acknowledgments) I would like to express my gratitude to is Marco Panza, my main supervisor. It is difficult to convey here my thanks to him without mentioning the content of the (more than 450, on Friday 3 June at 3:40 p.m.) mails, conversations, and exchanges of ideas that we had in these four years. His constant support and advice have been crucial for me. Marco has helped me to shape and clarify my ideas, even when these ideas were not solid or they were based on some misconception (or, simply, when they were the product of my philosophical ignorance). His thorough knowledge of history and philosophy of mathematics, and his pedagogical ability to clarify, to the point of obviousness, philosophical issues which were totally new to me, are of incalculable value for my work. He has been confident in my ideas, and he has encouraged me to refine and push further them when I was not confident in myself. I had the sensation to feel free in my investigation, but not lost under his constant supervision.

Working with him I have learnt not only how philosophy should be done, but also how philosophy is done. I am very grateful to him for many other things (among these his friendship), but let me just say that whether there is one good answer to one good question in this dissertation, this is thanks to him.

There are not only philosophers (and philosophies) which have influenced my views on the issues dealt with in this work. Behind this dissertation there are a number of people who have shared their views with me, thus contributing to my work. More importantly, they have supported me and they have shared with me a piece of their life. For me it's a privilege to mention them here. Thank you for having been there, and for still being there: Hank, Salva, Simone, Roby, Ifi, Silvia, Simo, Vale, Ricky, Giá, Peppone and Roberta, Andrea, Fio, Antonio Paci, Ale, Toni, La Caja Sonora, my Cort acoustic guitar (she knows why), Lulú, Corto Maltese (Dani), Marcovino, Paolopablo, Seba, Cico and Marina, Christian, Kiro, Lisa, Diego, Nicoletta and the Bologna crew (physicists and not physicists), GiorgioSax, Irene, Yago, Gaia, Momo, Carlito, Francesco and Alice (and Miró), Gloria, Giacomino and Fede, Fabbietto, Marchetto, Enrico, Romina, Laya, Gabbi, Marco, Equipo Fatti a Mano, Gerardo y Dan, Floriana, Nicco y Meli, Giusi, Guerino, Pietro, Gab, Ale, Tiziana, Federica, Francesco cugino, Elenapoli, Marco and Fede, Sandro and Fusun, Laura, Aysee and Emine, Irina, Davide, Sara, Andrea, Jé, Ghiz, Oscar Scoglio, Ele, Simonetta and Roby, Johanna, James, Nick, Juanca, Laura and Javi. I have to thank my family, to whom this work is dedicated: my mother Silva and my father Franco, my brother Luca and my cat Bart (who passed away some months ago). I have to thank you for so many things, but I will just mention that your strength has been my strength during *all* these years. Finally, I want to thank Giuli, for her patience (especially during the weeks before the submission), for her support and encouragement, for being here and for 'un tema que está en un solo de piano y en el labio más abrazador' (tú sabes). Thank you!

The research process from which this thesis has resulted has been financially supported by a Ph.D. grant from the Ministère de l'Enseignement

Supérieur Français. I also should acknowledge the financial support of the Ecole Doctorale 400 Savoirs Scientifiques, Epistémologie, Histoire des sciences, Didactique des disciplines (Université Paris 7 Denis Diderot), the Equipe Rehseis du Laboratoire Sphere (UMR 7219), the Comité ECOS-NORD at Université Paris 13 (action ECOS M04H01 under the coordination of Marco Panza at Paris 7 and Carlos Álvarez Jiménez at UNAM) and the Ideals of Proof Project (supported by the Agence Nationale de la Recherche).

Contents

Introduction

Scientific explanation

The problem of capturing the notion of explanation in science has a long history that I will not try to reconstruct here. Rather, I will start this dissertation by showing that the urgence to have a valuable notion of explanation can be traced back to ancient philosophy.

As David Hillel Ruben has remarked, Plato's theory of forms can be read as an extended discussion of the requirements for explanation [Ruben, 1990, p. 45]. On the other hand, it is in Aristotle that we have what might be considered as the first account of explanation in science. For Aristotle, scientific knowledge is knowledge of the cause [Aristotle, BWA 1941, p. 111, *Post. An.* I.1, 71b 5-10], where 'cause' (*aitia*) is intended in the sense of his four causes (formal, material, efficient and final)[1]. As observed by Paolo Mancosu, modern translators and commentators of Aristotle prefer to translate the term "aitia" as "explanation", so that "the so called doctrine of the four causes becomes an account of the kinds of explanations that can be used to answer a why-question" [Mancosu, 2000, p. 108][2].

Knowledge, according to Aristotle, is obtained through demonstration.

[1]The account of the four causes (formal, material, efficient and final) is expressed in his *Physics* (II.3) and *Metaphysics* (V.2).

[2]Among the commentators who translate 'aitia' with 'explanation' we find Barnes, Moravcsik, Hocutt, Annas. For instance, Barnes translates from *Posterior Analytics* I.1: "We think we understand a thing simpliciter (and not in the sophistic fashion accidentally) whenever we think we are aware both that the explanation [*aitia*] because of which the object is its explanation [*aitia*], and that it is not possible for this to be otherwise" [Aristotle, CWA 1984, p. 115, *Post. An.* I.1, 71b 5-10].

Nevertheless, only scientific demonstrations are suitable candidates for establish scientific knowledge. There are some necessary conditions for those kinds of demonstrations being scientific: first, the premises of the syllogism must be true; second and third, they have to be primitive and immediate; fourth, they have to be better known than the conclusion and prior to the conclusion, which is further related to them as effect to causes [Aristotle, BWA 1941, p. 112, *Post. An.* I.1, 71b 20-25]. In addition to this, in *Posterior Analytics* I.13 Aristotle distinguishes between "demonstrations of the fact" (*oti* proofs) and demonstrations "of the reasoned fact" (*dioti* proofs). Altough they are both valid, only from the latter we get the required conviction that the result is true plus the convinction of *why* it is true. In fact, while demonstrations of the fact proceed from effects to their causes, demonstrations of the reasoned fact permit us to grasp the causal structure of the phenomena under investigation going from causes (the explanantia) to the effects (the explananda)[3]. Demonstrations of the reasoned fact represent then, for Aristotle, the kind of reasoning which "met en oeuvre la cause" and which produces scientific knowledge. We can call them "explanatory demonstrations", whereas demonstration of the fact are considered as "non-explanatory".

Although Aristotle's deductivist theory of explanation has been extremely influential, especially until the Reinaissance and the seventeenth century [Mancosu, 1996], the first attempt to break with that tradition by providing a fresh perspective on scientific explanation is very recent and is found in Carl Hempel and Paul Oppenheim's paper "Studies in the Logic of Explanation" [Hempel *et al.*, 1948]. In this paper, which is now considered one of the most influential studies of the XXth philosophical literature, the authors proposed to capture the notion of scientific explanation by offering their deductive-nomological model (D-N model).

Hempel and Oppenheim's paper was regarded as a very promising perspective to the study of the notion of scientific explanation. However, the difficulty the D-N model had in solving problems like the so called "asymmetry-

[3]Note that, in this context, 'cause' means any of the four Aristotelian causes: formal, material, efficient, and final.

problem" and the "problem of explanatory irrelevances", which undermined the claim that the D-N model provides sufficient conditions for scientific explanation, soon led some authors to suggest some amendments to the model, or even to develop models of explanation which were based on a different picture of explanation. Among the latter models, the unification models proposed by Michael Friedman [Friedman, 1974] and Philip Kitcher [Kitcher, 1976], the pragmatic approach to explanation put forward by Bas Van Fraassen [Van Fraassen, 1980] and the causal models of scientific explanation such as that proposed by Wesley Salmon [Salmon, 1984a] seemed to offer themselves as thought-provoking alternatives to the Hempelian framework[4].

After the end of the 80's the debate on scientific explanation was far from its conclusion. Various models (for the most part refined versions of the major models which were available yet) were proposed, and the notion of scientific explanation became a central topic to be addressed in philosophy of science. This interest is very well mirrored by the role that the notion of explanation played, and continues to play, in the ontological debate between realist and anti-realists, a role which mainly concerns the so called 'IBE' (Inference to the Best Explanation) and Indispensability Arguments, in the discussion concerning the use of models and idealization in science [Morgan *et al.*, 1999a], and in the debate about the notion of 'understanding' in science [De Regt, 2009]. Furthermore, scientific explanation has been addressed from very distinct perspectives. For instance, two recent papers by Colin McGinn and Tian Yu Cao provide good examples of how this topic could be approached from the philosophy of mind or from an 'ontological' perspective ([McGinn, 2004] and [Cao, 2004]).

[4] I will consider some of these 'classical' models, emerged between the 70's and the 90's, in part I of this dissertation. Perhaps, the best way to grasp the importance that the notion of scientific explanation assumed among philosophers and philosophies of science during the four decades which go from 1948 to 1988, is to have a look at the huge volume XIII of the *Minnesota Studies in the Philosophy of Science* [Kitcher *et al.*, 1989]. This volume, which is consecrated to the subject of scientific explanation, provides a very detailed compendium of the different approaches to scientific explanation proposed during these forty years.

However, as the analysis of the notion of scientific explanation was going on, some philosophers realized that this notion was very general, and that perhaps a further distinction had to be introduced to account for cases where mathematical claims came in as an essential ingredient in the explanation provided. These latter cases correspond to cases of *mathematical explanation*. As clearly expressed by Paolo Mancosu in one of his studies on explanation, we can have two different senses of mathematical explanation:

> In the first sense 'mathematical explanation' refers to explanations in the natural or social sciences where various mathematical facts play an essential role in the explanation provided. The second sense is that of explanation within mathematics itself [Mancosu, 2008b, p. 134]

Mathematical explanation in science

In this dissertation I will consider some classical models of explanation, but the central issue will not be to focus on scientific explanation in general, neither on explanations within mathematics itself (formal and informal proofs within mathematics)[5]. I will turn my attention to situations in the natural sciences where mathematics is supposed to play an essential role in the explanation provided[6]. I will call such explanations, i.e. explanations of physical phenomena where mathematics is supposed to provide such an essential ingredient, 'mathematical explanations of physical phenomena' (henceforth 'MEPP'). MEPP clearly correspond to the first sense of mathematical explanation indicated by Paolo Mancosu (quotation above).

As an illustration of how mathematics might be thought to be offering explanations of physical phenomena, consider the following case. The ex-

[5]The topic of mathematical explanation within mathematics has received a particular attention in the recent philosophical debate. A survey of the relevant literature is provided in [Mancosu, 2008c].

[6]However, where necessary to my discussion, I will consider the topic of mathematical explanation within mathematics. For instance, this will be the case of Mark Steiner's account of mathematical explanation in physics, which relies on his account of explanation in mathematics and which will be presented in chapter 1.

ample concerns a biological phenomenon and is given by Aidan Lyon and Mark Colyvan in their article “The Explanatory Power of Phase Spaces” [Lyon *et al.*, 2008]. Hive-bee honeycombs always have an hexagonal structure. We want an explanation for this, namely we want to know why the honeycomb is always divided up into hexagons and not some other polygon (such as triangles or squares), or any combination of different (concave or convex) polygons. Now, as maintained by the biologists, hive-bees minimise the amount of wax they use to build their combs, since there is an evolutionary advantage in doing so. Nevertheless this gives only a partial answer to our original question ‘Why is the hive-bee honeycomb always divided up into hexagons (instead of some other polygons or some combination thereof)?’. To fill this gap and answer our question we turn to mathematics and we appeal to a theorem proved in 1999 by Thomas C. Hales [Hales, 2001]. The theorem, called the ‘honeycomb theorem’, states that an hexagonal grid represents the best way to divide a surface into regions of equal area with the least total perimeter. This mathematical result, together with the biological remark, is regarded by Lyon and Colyvan as offering an explanation of the physical phenomenon, and this explanation would provide an example of MEPP (the explanation of the biological fact seems to depend essentially on a mathematical fact). They write:

> So the honeycomb conjecture (now the honeycomb theorem), coupled with the evolutionary part of the explanation, explains why the hive-bee divides the honeycomb up into hexagons rather than some other shape, and it is arguably our best explanation for this phenomenon. [Lyon *et al.*, 2008, p. 3]

Another simple example of MEPP has been offered by Peter Lipton:

> There also appear to be physical explanations that are non-causal. Suppose that a bunch of sticks are thrown into the air with a lot of spin so that they twirl and tumble as they fall. We freeze the scene as the sticks are in free fall and find that appreciably more of them are near the horizontal than near the vertical orientation. Why is this?

> The reason is that there are more ways for a stick to be the horizontal than near the vertical. To see this, consider a single stick with a fixed midpoint position. There are many ways this stick could be horizontal (spin it around in the horizontal plane), but only two ways it could be vertical (up or down). This asymmetry remains for positions near horizontal and vertical, as you can see if you think about the full shell traced out by the stick as it takes all possible orientations. This is a beautiful explanation for the physical distribution of the sticks, but what is doing the explaining are broadly geometrical facts that cannot be causes [Lipton, 2004, p. 9-10]

As Lipton observes, this explanation is not causal and it is carried out by essential appeal to mathematical (in this case geometrical) facts. This is a crucial point, because MEPP seem to be counterexamples to the claim that all explanations in the natural science must be causal. For instance, in the case of the explanation of why hive-bee honeycombs always have an hexagonal structure, we do not trace the causal processes and the interactions leading up to an event, but we appeal to a geometrical theorem.

Scientific explanations, of course, make use of mathematics. Furthermore, spheres of mathematical practice and scientific practice frequently overlap. However, there is difference between an explanation in science which is performed by essential appeal to a mathematical fact, such as that concerning the structure of the hive-bee honeycombs, and an explanation which involves the application of mathematics but which does not appeal to mathematics in this essential way (for instance, Newton's explanation for the motion of planets in terms of action at a distance). Unfortunately, the dividing line between MEPP and scientific explanation is not so sharp, and there are situations where it is not easy to separate the empirical part of the explanation from the mathematical part. For instance, if we pass to physics, given the highly mathematized nature of the subject, it becomes difficult to distinguish between the mathematical and the physical components of an explanation. Even in these cases, however, there are situations where mathematics is supposed to play an explanatory (and not purely justificatory) role. As will see in part II of this dissertation, in these situations mathematics plays an explanatory role not through a theorem, as in the hive-bee honeycombs example, but through its internal resources (for instance, operations or structures) which are involved in the process of application of mathematics to the natural sciences.

But does mathematics play this explanatory role in science? The answer, at least for what a number of philosophers suggest, seems to be a positive one. For instance, the following philosophers think mathematics does play such a role in science: Mark Steiner [Steiner, 1978b], Mark Colyvan [Colyvan, 2001], Robert Batterman [Batterman, 2002a], Alan Baker [Baker, 2005], Mary Leng [Leng, 2005], Aidan Lyon and Mark Colyvan [Lyon *et al.*, 2008], Sorin Bangu [Bangu, 2008], Paolo Mancosu [Mancosu, 2008b]. However, it should be noted that there is no general consensus on this point and other philosophers reject the claim that mathematics plays an explanatory role in science. For instance, Joseph Melia, Chris Daly and Simon Langford claim that the role of mathematics is one of "indexing" physical facts, not explaining them ([Melia, 2000]; [Daly *et al.*, 2009]), while Juha Saatsi defends the idea that mathematics does not play any 'genuine' explanatory role but only a representational one [Saatsi, 2011].

Although MEPP *per se* have been subject to an intensive investigation only during the recent years, the attention to this topic (as for the study of scientific explanation) can be traced back to the Greeks [Mancosu, 2008b, p. 134]. For instance, the physics of Aristotle was not mathematized but Aristotle discussed extensively the so-called mixed sciences (optics, harmonics, and mechanics), characterizing them as the more physical of the mathematical sciences (*Posterior Analytics* I.13). Every mixed science was subordinated to an area of pure mathematics (for instance, harmonics to arithmetic and optics to geometry). According to the Greek philosopher, explanatory demonstrations are to be found in the mathematical sciences and therefore he welcomed the idea that there are mathematical explanations of physical phenomena:

> For here it is for the empirical scientist to know the fact and for the

> mathematical to know the reason why; for the latter have the demonstrations of the explanations, and often they do not know the fact, just as those who consider the universal often do not know some of the particulars through lack of observation [Aristotle, CWA 1984, p. 128 (Vol I), *Post. An.* I.13, 79a 1-7]

However, as the domain of applied mathematics grew, the topic of whether mathematics could give explanations of natural phenomena was one on which there was disagreement. The Aristotelian conception of pure mathematics, as abstracting from matter and motion, was evidently difficult to reconcile with the fact that both physics (natural philosophy) and the mixed sciences are all conversant about natural phenomena and thus dependent on matter and motion. For instance, the so-called *Quaestio de Certitudine Mathematicarum*, an important debate which took place in the Renaissance, focused in large part on whether mathematics could play the explanatory role assigned to it by Aristotle. By adducing the argument that mathematics lacks causality, some argued that it cannot play any explanatory role in natural philosophy[7].

As it is natural to think, a real turning point in the use of the notion of explanation in science was marked by the apparition of Newton's *Principia*, where for the first time natural philosophy was subject to a process of 'mathematization'[8]. In his 2001 paper, Yves Gingras has underlined how the "disparition of substances into the acid of mathematics" and the shift in the criteria for explanation are an ontological and an epistemic effect of the process of mathematization started with Newton [Gingras, 2001]. Newton's mathematization which appears in the *Principia*, by preferring an abstract treatment of phenomena, accelerated the disparition of substances like carte-

[7]On the *Quaestio de Certitudine Mathematicarum* and the main issues raised by this debate see chapter 1 of [Mancosu, 1996].

[8]Here mathematization is intended in the following sense: "we should speak of mathematisation only when the object of this science becomes a mathematical object, i.e. mathematics provides a model or a scheme of a natural or social phenomenon and this model or scheme becomes the real object of studying" [Panza, 2002, p. 253-254]. According to this definition, Galileo's law of free fall does not represent a case of mathematization.

sian vortex or luminiferous ether into the mathematical machinery. Moreover, the epistemic effect consists in the fact that the use of mathematics in dynamics (as distinct from the use in kinematics) changed the use of the term 'explanation' with respect to the use which had been made by the philosophers during the 17th century. The case of gravitation shows how the criteria for explanation have shifted from the previous mechanical explanation to an explanation which was legitimated in using mathematics as language. During the 17th century, to 'explain' a phenomenon stood for 'to offer the physical mechanism which is at the base of his production'. This is why Descartes rejected Galileo's explanation of the free fall, claiming that it was not based on a mechanical explanation. However, with the publication of the *Principia* we have the beginning of a shift in the criteria for explanation: "the mathematical explanation begins to be preferable to the mechanical explanation when the latter did not conform to calculations" [Gingras, 2001, p. 398].

Now, even if we recognize that the use of the term 'explanation' has changed in scientific practice and that mathematics has been regarded by some philosophers as playing an explanatory role in science, we are still confronted with the problem of offering a notion, or even a characterization, of MEPP.

The recent exigence of developing specific theories of MEPP is due mainly to the difficulty in accounting for mathematical explanation of physical facts starting from general theories of scientific explanation. In fact, it is often observed that the leading contemporary theories of scientific explanation are in trouble when faced with MEPP ([Batterman, 2002a], [Baker, 2005], [Mancosu, 2008b]). In some cases these theories left apart mathematical explanations and they did not accept pure mathematical statements within their structure. This is the case of Hempel's D-N model, in which the explanans must have empirical content (it must capable, at least in principle, of being tested by means of experiments and observations). In other cases these accounts have been regarded as insufficient for the treatment of specific cases of MEPP. This is what happened, for instance, with Kitcher's unification model[9]. Finally, if mathematical objects are acausal, MEPP represent counterexamples to causal models of scientific explanation, such as Salmon's, which consider explanation in the natural sciences as esssentially causal. By assuming that mathematical objects do not play any essential role in the explanation provided, causal models miss MEPP simply because they rule out the possibility of having such a kind of explanations. Of course, all the previous considerations point to some substantial impediments the major theories of scientific explanation have when confronted with MEPP. Despite the great interest in the linkage scientific explanation-MEPP, however, an extensive discussion of models of scientific explanation in the context of MEPP has not been offered and work is just beginning [Mancosu, 2008b].

But the attention to the specific topic of mathematical explanations in mathematics and in science is not only a consequence of the difficulty that some traditional models of scientific explanation have in capturing the notion of mathematical explanation. Two main factors contributed to the increased study of MEPP in the area of philosophy of science, and more particularly in the context of philosophy of mathematics:

1. The increased interactions between mathematics and natural sciences during the second half of the XXth century ([Urquhart, 2008a] and [Urquhart, 2008b])

2. The emergence of "new directions" in the philosophy of mathematics which gave more attention to the mathematical practice ([Aspray *et al.*, 1988] and [Tymoczko, 1998])

Note that when I am stressing the first point I am not saying that there were no interactions before (something that would be evidently false), or that these interactions were not studied. I am just observing that the boundary between mathematics and natural sciences has become less definite over the last years, and this process has called for a specific attention to the role

[9]The fact that Kitcher's theory cannot account for a particular case of MEPP has been pointed out by Batterman [Batterman, 2002a, p. 35]. Let me note, however, that Batterman provides only a general discussion and not a detailed analysis.

played by mathematics in science. Furthermore, to this process of interaction have corresponded a new interest in the role of empirical procedures in mathematics itself[10].

It is worth observing how John Von Neuman claimed for a injection of empirism in mathematics as a vital condition for this discipline, when in 1947 wrote:

> At a great distance from its empirical source, or after much "abstract" imbreeding, a mathematical subject is in danger of degeneration. At the inception the style is usually classical; when it shows signs of becoming baroque, then the danger signal is up. [...] Whenever this stage is reached, the only remedy seems to me to be the rejuvenating return to the source: the reinjection of more or less empirical ideas. I am convinced that this was a necessary condition to conserve the freshness and the vitality of the subject and that this will remain equally true in the future [Von Neumann, 1947, p. 196]

The fact that the interaction between the methodology of physics and that of mathematics should be considered as beneficial to mathematics was not only Von Neuman's opinion. Even Pierre Cartier, who was a member of the Bourbaki group in the 50's, has highlighted the positive aspect of this interaction:

> The implicit philosophical belief of the working mathematician today is the Hilbert Bourbaki formalism. Ideally, one works within a closed system: basic principles are clearly enunciated once for all, including (that is an addition of the twentieth century science) the formal rules of logical reasoning clothed in mathematical form. The basic principles include precise definitions of all mathematical objects [...] My thesis is: there is another way of doing mathematics, equally successful, and the two methods should supplement each other and do not fight. This other bears various names: symbolic method, operational calculus,

[10] Perhaps the most well-known icon of this interest is Imre Lakatos' famous book *Proofs and Refutations* [Lakatos, 1976].

> operator theory [...] Euler was the first to use such methods in his extensive study of infinite series, convergent as well as divergent [...] But the modern master was R. Feynman who used his diagrams, his disentangling of operators, his path integrals [...] The method consists in stretching the formulas to their extreme consequences, resorting to some internal feeling of coherence and harmony [Cartier, 2000, p. 6].

On the other hand, the discussion of the irruption of methods of physicists into pure mathematics has given birth to a series of controversies, as for instance that which recently took place on the *Bulletin of the American Mathematical Society*. This dispute confronted some theoretical mathematicians, on one side, and on the other mathematicians who pointed to the dangers of using speculative methods in mathematics [Urquhart, 2008b].

Nevertheless, if "there is another way to do mathematics", then it is important for the philosopher of mathematics to consider it and answer the question 'How is mathematics done?'. Furthermore, and perhaps more important for the present study, to a different way of doing mathematics there corresponds a different way to use mathematics in science. It is then natural to consider that the interaction between mathematics and science had strong repercussions not only on the methodology of mathematics, but also on the philosophical study of topics which are related to that methodology. In particular, the emergence of a variety of mathematical procedures in physics offered new material for philosophical thought and opened the way to specific studies concerning the applicability of mathematics, the structure of mathematical explanation in physics, the use of mathematical models in science, the new methodology of mathematics and the strategies of assimilation between the two domains [Urquhart, 2008b].

We can say than that the rise of the investigation of the interface between physics and mathematics is connected to the second factor listed above: the emergence during the sixties of a strong opposition to the classical foundational programs in philosophy of mathematics (logicism, Hilbert program and intuitionism). This reaction against the "dogmas of foundationalism"

[Tymoczko, 1998, p. 95], which started with Imre Lakatos and which has been pursued by what Aspray and Kitcher defined as "a maverick group of philosophers" [Aspray *et al.*, 1988], gave a central importance to the history of mathematics and assumed mathematical practice as a driving force in the philosophical research[11]. The research questions posed by those philosophers were thus more oriented on the heuristics of mathematics, the way in which mathematics grows, the notion of explanation in mathematics, the distinction between formal and informal proofs and reasonings in mathematics. For instance, in their *History and Philosophy of Modern Mathematics* [Aspray *et al.*, 1988], Aspray and Kitcher write:

> Philosophers should pose such questions as: How does mathematical knowledge grow? What is mathematical progress? What makes some mathematical ideas (or theories) better than others? What is mathematical explanation? [Aspray *et al.*, 1988, p. 17]

A similar claim about the importance of scientifc practice in philosophical investigation comes from the general philosophy of science:

> Nowadays few philosophers of science will contest that they should take account of scientific practice, both past and present. Any general characteristic of actual scientific activity is in principle relevant to the philosophical analysis of science [De Regt *et al.*, 2005, p. 139]

The previous considerations about the importance to focus on mathematical and scientific practice in a philosophical investigation will be essential to my study. In particular, I will show how the attention to mathematical and scientific practice is now regarded as a crucial factor to the emergence, the rejection and the refinement of a philosophical model of MEPP.

[11] It should be observed that a decisive step toward the possibility of a theoretical analysis of mathematics in line with natural science came before, from Quine and his refusal of a distinction between analytic and synthetic. Quine wrote: "Total science, mathematical and natural and human, is similarly but more extremely undetermined by experience. The edge of the system must be kept squared with experience; the rest, with all its elaborate myths or frictions, has as its objective the simplicity of laws. Ontological questions, under this view, are on a par with questions of natural science" [Quine, 1951, p. 42].

The outline of this dissertation

The previous paragraphs show how the philosophical interest concerning the notion of scientific explanation has a long history, and the attention to this topic has now large ramifications in different areas of philosophy. Moreover, I drew the reader's attention to the fact that classical models of scientific explanation have difficulties in accounting for MEPP, and I put forward some remarks which illustrate how even the investigation of MEPP can be traced back to the ancients. Only in the recent years philosophers addressed the notion of MEPP with the intention of capture it through a model. This interest, as I suggested, is the result of various factors and of a renewed consideration of mathematical and scientific practice. The problem of capturing the notion of MEPP, however, stands as yet in need of a detailed analysis and there is no general consensus on the fact that mathematics plays an explanatory role in science.

It is now time to illustrate the general project of my work. I will present this by summarizing the contents of the three parts which compose this dissertation.

The traditional tendency toward scientific explanation has been to capture the nature of explanation by providing a single model, i.e. a model to which the variety of explanations can be reduced. Call this approach the *winner-take-all* approach to explanation (WTA)[12]. In the first part I shall present the WTA views on MEPP. I shall take as representative of this view three major accounts of explanation: Mark Steiner's account, Bas Van Fraassen's pragmatic account and Philip Kitcher's model of explanation in

[12]Note that my use of the expression "winner-take-all conception of explanation" is different from that James Woodward made of it [Woodward, 2003, p. 367]. Woodward used the same expression to indicate an account of explanation which considers as only options for a theory or a derivation to be explanatory/unexplanatory. According to him, such a type of account automatically rules out a less explanatory theory (or derivation) as unexplanatory and does not leave room for a judgement of more/less explanatoriness on a continuum. This is the case of Kitcher's and Friedman's unification accounts, where only the more unifying theory (or systematization) is considered as explanatory, and less unifying theories (or systematizations) are not qualified as providing less explanatoriness but are marked as nonexplanatory.

terms of unification. Every model will be discussed in the relative chapter and, at the end of each chapter, I shall report the major criticisms which were addressed to each account.

While Mark Steiner has developed an explicit model of mathematical explanation of physical phenomena, which is connected to his theory of mathematical explanation of mathematical facts, Van Fraassen's and Kitcher's accounts were built to cover the notion of scientific explanation. However, these two models have been proposed as good candidates to cover MEPP as well, and this is why I will include them in my study. In particular, the discussion of Van Fraassen's account as a potential model for MEPP will be based on a paper by Alan Baker [Baker, 2005], in which such an idea is sketched (i.e. the idea that the pragmatic model can deal with mathematical explanation in science). In his paper, Baker implicitly suggests an extension of two traditional models of scientific explanation (the D-N model and the pragmatic account) as to treat cases in which mathematics is recognized to provide an essential ingredient in the explanation of a physical phenomenon. I will thus discuss the possibility of extending the pramatic model to MEPP and, *en passant*, I will also present the classical D-N model of explanation and some classical problems which were not solved by that account. With respect to Kitcher's account of explanation as unification, let me note here that this model has been considered by Kitcher himself as an encompassing model for explanation in mathematics and empirical science. This makes it relevant to the topic of MEPP. Furthermore, the choice to offer a detailed presentation of this account, in chapter 3, is worth for the general strategy of my dissertation. Although very long, in fact, my presentation of Kitcher's account will be instrumental in introducing some characterizations of the WTA approaches (chapter 4) and in constrasting the WTA approach with what I will call a 'pluralist' approach to MEPP. Finally, in the last part of my study, the details of Kitcher's theory will come out as essential to accomplish two main tasks. First, I will assess this model on a case of MEPP coming from the scientific practice, thus providing a testing of the unification model in the context of MEPP. In order to achieve this task, I will need a comprehensive picture of Kitcher's theory, namely Kitcher's unification picture of explanation given in chapter 3. Second, in the last chapter I will propose the idea that my own approach to MEPP can be generalized, and that its basic notions might be extended to scientific explanation as well. And, again, I will turn to some aspects of Kitcher's model.

The philosophical discussion about MEPP has inherited the same sort of connection that the topic of scientific explanation had, and continues to have, with the ontological arena. This is why, in presenting the three WTA models, I will devote some attention to the role that MEPP play (or are supposed to play) in the platonist-nominalist debate concerning the so called 'Enhanced Indispensability Argument'[13].

In the conclusive chapter of this first part, chapter 4, I will propose some characterizations in order to distinguish between some essential features of the three models and the ontological view which is associated with each account.

In the second part I will focus on the *pluralist* view on MEPP. As the name itself suggests, the authors who endorse this attitude towards explanation do not welcome the idea that there exists a single model of MEPP and consider that MEPP are heterogeneous. In other words, a pluralist considers that what makes something a good explanation can vary from case to case and that we cannot design a single model able to capture all these instances of MEPP[14]. As representative of this pluralist attitude toward MEPP I shall

[13]The Enhanced Indispensability Argument (EIA) belongs to the terrain of the ontological debate in philosophy of mathematics. The realist-partisans of the EIA refer to the indispensable explanatory power of mathematics in scientific theories as an instrument to support the claim that some mathematical objects exist. I will show how to this platonist attitude there corresponds a strong criticism intended to block the possibility of making such a realist inference in EIA.

[14]The reader may perhaps be surprised at reading that I am going to consider Van Fraassen's model among the WTA models, and not among the pluralist views of this second part. I will give a motivation for this choice in chapter 2, when presenting the criticisms which have been leveled against the pragmatic account. Nevertheless, let me note here that Van Fraassen accepts the idea that there are different kinds of explanation (and this is a core idea of his model), and this is perfectly in line with a pluralist view on

take into consideration two recent views, that of Christopher Pincock and Robert Batterman. As in the first part, I shall also discuss the ontological commitment which results from the adoption of such positions. These positions are extremely important because they represent fresh perspectives to the study of MEPP. Furthermore, to introduce these views will highlight the entanglement of the notion of MEPP in the contemporary debates on modelling, application of mathematics and idealization in science.

In the final part of my dissertation, part III, I will propose my own approach to MEPP. In particular, my aim will be twofold: I will defend the idea that pluralism is the best alternative to the study of MEPP (at least for what the scientific practice seems to suggest us); I will show how through the introduction of the categories of *intellectual tools* and *conceptual resources* we are able to account for MEPP which have been considered as genuine in scientific practice, and this without losing the pluralist principle as guide.

In proposing my own approach to MEPP I will take into consideration a general moral which emerges from the criticisms to the models presented throughout this dissertation: in order to capture a notion of MEPP which accords with our scientific and mathematical practice we have to consider some *qualitative* factors into our philosophical model of explanation (instead of purely *quantitative* factors). In general, I will refer to quantitative factors as those factors which can be captured through a formal scheme or analysis. On the other hand, qualitative factors are pragmatic factors, which cannot be captured through such a formal scheme or analysis. To analyze the nature of these qualitative ingredients, I will adopt a 'bottom-up' methodological approach to MEPP: I will take the case studies themselves, i.e. cases of MEPP recognized as such in the scientific practice, as starting point for philosophical analysis. This is why, in chapter 7, I will assess the three WTA models on a case of MEPP coming from the scientific practice. The choice to perform this assessment in this chapter, and not during my discussion of the models

explanation. However, he gives a single encompassing model for explanation (his 'why-question' account). In this sense, Van Fraassen is therefore not pluralist (at least according to my definition of pluralism), and his approach is a WTA approach to explanation.

in Part I, is motivated by the fact that I will use this testing to introduce my approach to MEPP in the final chapter. This strategy gives, I think, a more uniform character to the final part and to the general structure of the dissertation.

The picture of explanation I am going to sketch in the last part is based on a paradigm totally different from that which stands behind the majority of the accounts presented in this dissertation. With the only exception of Batterman's and Van Fraassen's models, in fact, all these models are based on the idea that the feature which contributes to the genuineness of an explanation is an objective feature, i.e. a feature which does not depend upon the observer performing the explanation. According to this view, the task of a theory of explanation is to individuate this particular feature (a particular quality of the mathematical formalism, a particular state of affairs, a fact, a relation which holds in the world or in mathematics). On the other hand, I will consider that it is the *way* in which we identify a particular state of affairs that contributes to the genuineness of a MEPP. In my view, a genuine explanation does not result from the identification of a particular state of affairs or property of the world or mathematics, but rather from the fact that we can look at that property or state of affairs in a specific way. I am going to argue that when we can reconceptualize a particular state of affairs (through particular mathematical concepts which I call conceptual resources), and this reconceptualization permits to use particular abilities to reason (intellectual tools), we do have a genuine explanation.

As I have observed, Batterman's and Van Fraassen's models are different from the majority of the models considered in this study because they are not based on the idea that the feature which contributes to the genuineness of an explanation is an objective feature. And therefore they seem to be compatible to my approach in some respect. However, let me anticipate that there is some essential difference between my approach and these accounts. First, Van Fraassen's model is a WTA model, and my approach will be based on a pluralist principle. Second, in the details, my proposal will considerably depart from Van Fraassen's and Batterman's.

I will present the notions of conceptual resources and intellectual tools in the final chapter. In the same chapter, in order to illustrate how my ideas work, I will discuss my approach in the context of the example of MEPP introduced in chapter 7. A specific section will be devoted to the possibility of generalizing my framework. Finally, I will point to the payoff of adopting such an approach. In particular, how it might provide insights into the ontological dispute in philosophy of mathematics (Enhanced Indispensability Argument) and into the debate on the notion of scientific understanding.

Far from offering a solution to the problem of what a MEPP *is*, my study will suggest, I hope, new directions which are yet to be explored.

Part I

The winner-take-all approach to explanation

Chapter 1

Steiner's approach to MEPP

The first explicit attempt of giving an account of MEPP in analytic philosophy has been made by Mark Steiner in his paper "Mathematics, explanation and scientific knowledge" [Steiner, 1978b]. His account, which is illustrated by taking into consideration a single test-case from the realm of kinematics of rigid body motion, relies on his theory of explanation in mathematics, which is presented in his other paper "Mathematical explanation" [Steiner, 1978a].

In Steiner, the possibility of grounding an account of mathematical explanation in physics on an account of mathematical explanation within mathematics comes from the assumption that there exists some kind of continuity between natural sciences and mathematics. As he observes, however, this continuity refers only to methodological similarities and not to the possibility of interactions between the two worlds. Although these methodological similarities are discussed in a more comprehensive way in his books *The Applicability of Mathematics as a Philophical Problem* [Steiner, 1998] and *Mathematical Knowledge* [Steiner, 1975], the point is stressed in both his 1978 papers on explanation:

> I myself have argued for continuity between the natural and mathematical sciences in *Mathematical Knowledge*. But such continuity begins and ends with methodological alikeness: both describe an objective world of entities, and (I argue) the methods used in exploring

> the two worlds are, despite common opinion to the contrary, remarkably similar. There may even be a power of observing mathematical truth akin to physical perception. But the foregoing considerations preclude any interaction between the two worlds. [Steiner, 1978b, p. 27].

> The growing acceptance, however, of continuity between the natural and mathematical sciences – urged by Quine, Putnam and the present author – has prepared the way for what follows here. [Steiner, 1978a, p. 135]

I am going to examine his account of MEPP and therefore I will concentrate on his [Steiner, 1978b]. However, since Steiner's account of MEPP is based on his theory of mathematical explanations within mathematics, I will also need to discuss his [Steiner, 1978a].

Steiner's [Steiner, 1978b] is divided into three parts. In the first part, he presents his account of mathematical explanation in physics. As a test case he takes an example from the kinematics of rigid body motion: the Euler's theorem for the existence of an instantaneous axis of rotation. As we will see, the fact that this example is considered by Steiner as a genuine case of MEPP is parasitic on the fact that, for him, such a theorem is based on an explanatory mathematical proof. This is in line with Paolo Mancosu's assertion that "whatever account we will end up giving of mathematical explanations of scientific phenomena, it won't be completely independent of mathematical explanation of mathematical facts" [Mancosu, 2008b, p. 192-193]. I will present Steiner's model of MEPP in the next sections. Before that, however, let me shortly consider the second and the third part of Steiner's paper.

The second part of [Steiner, 1978b] is devoted to a general discussion of the causal theory of knowledge under the forms:

ω One cannot know anything about F's unless this knowledge is caused by at least one event in which one F participates.

θ It is impossible to know anything about an entity x unless x itself participates in the cause of that knowledge[1]

This discussion is important because it introduces Steiner's platonist position in philosophy of mathematics, which is something that (as we will see throughout this chapter) is often discussed by him in connection with his view on explanation. In particular, Steiner argues that the causal theory of knowledge in its forms ω and θ is not scientific, i.e. it is unable to account for our scientific knowledge. He points out that a correct account of scientific inference could not be obtained starting from a theory which considers perception as the only way to obtain knowledge. To substantiate his claim, he takes an example from nuclear physics, and specifically the beta-decay of a stationary nucleous of Lithium 6:

> Now we do not learn about the neutrino by transmission of energy from the neutrino to us –the neutrino is very difficult to detect by direct interaction. Indeed, as far as is known, beta decay is noncausal –no anterior event causes the breakup of the unstable lithium 6 nucleus. Nor does the neutrino participate in any event which causes the other particles' motion-through which we infer the existence of the neutrino. Beta decay "just happens" in accordance with the law of conservation of momentum, enabling us to infer a new particle. Laws of conservation are simply not causal laws. What they provide are constraints on what is allowed to happen [Steiner, 1978b, p. 22].

According to Steiner, in this case the inference (from which we deduce the existence of the neutrino) is given only by the law of conservation of momentum. But this law violates "the spirit" of the causal theory of knowledge because: a) the considered law does not provide the sort of knowledge required by ω and θ (there is no anterior event which causes the decay; the neutrino does not participate in any particle-interaction which is used to infer its ex-

[1]Observe that θ is a version of ω more suitable for attacking the platonist position, as Steiner rightly observes [Steiner, 1978b, p. 21].

istence); b) the neutrino does not exert any causal influence on the knower[2]. To adopt the causal theory of knowledge in the forms above amounts to considering that empirical evidence for numbers is impossible because of the non-material qualities attributed to mathematical entities. However, as the example shows, the causal theory of knowledge under these forms is not the epistemology required by modern science. Therefore Steiner's conclusion is that the causal theory of knowledge must be considered as unable to block mathematical knowledge.

The previous considerations, contained in the second part of [Steiner, 1978b], are intended to show that the causal theory of knowledge cannot be used to deny the existence of numbers. In the third and last part of his paper Steiner makes a different move. He assumes the existence of numbers and asks if empirical results can be used to refuse their existence. He starts his discussion by taking into account Benacerraf's structuralists considerations about arithmetics as the science of progressions:

> It was pointed out above that any system of objects, whether sets or not, that forms a recursive progression must be adequate. But this is odd, for any recursive set can be arranged in a recursive progression. So what matters, really, is not any condition on the objects (that is, on the set) but rather a condition on the relation under which they form a progression. [Benacerraf, 1965, p. 69]

> Any object can play the role of 3; that is, any object can be the third element in some progression. What is peculiar to 3 is that it defines that role -not by being a paradigm of any object which plays it, but by representing the relation that any third member of a progression bears

[2]The latter point is a response to W. D. Hart's view of θ [Hart, 1977]. Hart argues that when learning something about the world (an object, the neutrino, etc.), every learner changes materially and the change results from an energy absorption from the environment. This is how the object participates in the cause of our knowledge of it (according to θ). With the example of the beta-decay of Lithium 6, Steiner is pointing out that the transfer of information needs not involve the transfer of energy. In particular, the neutrino interacts with us but this interaction is particularly difficult to detect, therefore in this case the transfer of energy does not participate in the cause of the knowledge of the neutrino. This, according to Steiner, would provide a lever on Hart's interpretation of θ.

> to the rest of the progression. [...] Arithmetic is therefore the science that elaborates the abstract structure that all progressions have in common merely in virtue of being progressions. It is not a science concerned with particular objects – the numbers. The search for which independently identifiable particular objects the numbers really are (sets? Julius Caesars?) is a misguided one. [Benacerraf, 1965, p. 70]

Steiner observes that, once we assume the existence of a single infinite progression, empirical results cannot falsify arithmetics. They can only state its inapplicability. Contrary to Quine's and Goodman's arguments [Goodman *et al.*, 1947][3], the existence of an infinite arithmetical progression "cannot be experimentally demonstrated".

To sum up, according to Steiner the causal theory of knowledge (in its forms ω and θ) cannot be used to make inferential claims about numbers, and empirical results cannot falsify them. Furthermore, as we are going to see, Steiner argues that mathematical explanation of physical phenomena do exist but this does not conflict with the impossibility of showing the existence of an infinite progression using empirical results [Steiner, 1978b, p. 27]. In presenting his account of mathematical explanations in physics, in fact, Steiner also considers the existential implications of assuming the existence of such kind of explanations.

In the following section, I will concentrate on the first part of Steiner's [Steiner, 1978b], where his account of mathematical explanation in physics is offered (together with the ontological considerations which follow). Next, I will skip to his account of mathematical explanation in mathematics, which is essential in order to fully understand the former. After that I will come back to Steiner's account of MEPP and I will show how the previous consid-

[3]Steiner considers Quine's argument for platonism in mathematics as a transcendental argument: "But Goodman and Quine pointed out thirty years ago the apparent impossibility of describing the world without reference to numbers. I would put their point thus: to describe the experience of diversity and change requires mathematicatical entities. [...] We cannot say what the world would be like without numbers, because describing any thinkable experience (except for utter emptiness) presupposes their existence" [Steiner, 1978b, p. 19-20]. In passing, let me observe that Steiner's reading of Quine's argument does not reflect the common interpretation of that argument [Panza *et al.*, 2010].

erations concretely fit the example of MEPP given by Steiner. Finally, I will report some criticisms.

1.1 Steiner on MEPP

In the opening section of his [Steiner, 1978b], Steiner considers two questions:

- Do physical phenomena have mathematical explanations?
- If so, what existential conclusions follows? Do such explanations make reasonable the existence of mathematical entities?

In order to answer the first question, he discusses a single example from mechanics, and in particular from kinematics: the general displacement of a rigid body with one point fixed. This motion can always be obtained by rotating the body of a certain angle about a fixed axis. The axis, called the "instantaneous axis of rotation", passes through the fixed point. The result was proved for the first time by Euler in his *Decouverte d'un Nouveau Principe de Mécanique* [Euler, 1750] and is known in physical and mathematical textbooks under the name of "Euler's theorem" for rigid body motion[4]. As formulated in a classical texbook of mechanics, Euler's theorem states that

Theorem 1.1. *The general displacement of a rigid body with one point fixed is a rotation about some axis [Goldstein, 1957, p. 118].*

Although we can prove the theorem by geometry alone[5], a simple algebraic proof shows the existence of the axis. Steiner points out that we have a mathematical *explanation* for the physical fact (i.e. the existence of

[4]Observe that, in his presentation of the theorem, Steiner does not report Euler's original proof. For a reconstruction of Euler's original argument see [Koetsier, 2007, p. 184-185]. Euler does not use the word "instantaneous axis". He refers to it simply as "axe de rotation" [Euler, 1750, p. 95].

[5]See [Whittaker, 1904, p. 2], or [Targ, 1987, p. 221-222].

such a fixed axis in real space) because we obtain it from such "explanatory" algebraic proof plus the following two physical assumptions (or "bridge principles", as Steiner calls them):

- Space is 3-dimensional euclidean
- The rotation of a rigid body around a point generates an orthogonal, real, proper transformation[6]

How then are the foregoing remarks linked to an "account" of MEPP? According to Steiner, we have a mathematical explanation of a physical fact when, removing the physical assumptions (such as those above), what we are left with is a mathematical explanation of a mathematical fact. In other words, if we delete the physics we remain with an explanatory proof of a theorem (where the import given by Steiner to the expression "explanatory proof" will be clarified the next section). He writes:

> I shall not reproduce my analysis of mathematical explanation here, but assume that mathematical explanation of mathematical truth exists. The difference between mathematical and physical explanations of physical phenomena is now amenable to analysis. In the former, as in the latter, physical and mathematical truths operate. But only in mathematical explanation is this the case: when we remove the physics, we remain with a mathematical explanation of a mathematical truth! In our example, the "bridge" between physics and mathematics is the assumptions that space is three-dimensional Euclidean, and that the rotation of a rigid body around a point generates an orthogonal, real, proper transformation (to use the lingo). Deleting these assumptions, we obtain an explanatory proof of a theorem concerning transformations and eigenvectors. In standard scientific explanations, after deleting the physics nothing remains [Steiner, 1978b, p. 19].

[6]A *proper* tranformation is a tranformation whose representative matrix has determinant $+1$. Tranformations whose representative matrix have determinant -1 are called *improper*. I will come back to these notions in section 1.3, where I will also present the algebraic proof of Euler's theorem.

Hence, as it is clear from the quotation, Steiner's response to the first question "Do physical phenomena have mathematical explanations?" is positive: there exist mathematical explanations of physical phenomena. Moreover, these explanations rely on the 'explanatoriness' of proofs in mathematics ("when we remove the physics we remain with a mathematical explanation of a mathematical truth"). Thus, in order to appreciate his account of MEPP, it is natural to turn to what Steiner considers as a mathematical explanation within mathematics. This is why, in the next section, I will move to Steiner's account of explanation in mathematics.

Before going thorugh his theory of mathematical explanation in mathematics, however, let me mention as an aside Steiner's answer to the second question reported at the beginning of this section: "Do such explanations make reasonable the existence of mathematical entities?". Although supporting Quine and Goodman's argument about the necessity of mathematical entities in order to "describe the experience of change and diversity", Steiner denies that mathematical explanation of physical phenomena could be used to infer the existence of mathematical entities [Steiner, 1978b, p. 19]. And this because the existence of mathematical entities is presupposed in the description of the phenomena to be explained. As we will see in the next chapter, this position is controversial. For instance, Alan Baker, another philosopher interested in the role that MEPP play in the ontological dispute in philosophy of mathematics, will argue for an opposite claim (he will maintain that MEPP *do* make reasonable the existence of mathematical entities).

1.2 Steiner's account of explanation in mathematics

In his [Steiner, 1978a], after having discussed (and rejected) four criteria of explanatoriness of a mathematical proof (abstracteness; generality; discoverability and visualizability[7]), Steiner gives his own account of explanation within mathematics in order to distinguish between an explanatory and a non-explanatory proof.

The criteria of abstracteness, generality and discoverability are discussed comparing different proofs of a same theorem (respectively, the sum of the first n integers equals $\frac{n(n+1)}{2}$, the irrationality of $\sqrt{2}$, the Pythagorean theorem and the Eulerian identity $(1+x)(1+x^3)(1+x^5) = 1+\frac{x^2}{1-x^2}+\frac{x^4}{(1-x^2)(1-x^4)}+\frac{x^9}{(1-x^2)(1-x^4)(1-x^6)}$), while the suggestion that there exists a link between explanation and the ability to visualize a proof is dismissed because of the subjective character of the latter. Indeed, Steiner's general idea is that a theory of mathematical explanation for mathematical proofs must show the plausibility of those four criteria.

The starting point of his discussion of explanation in mathematics is the following observation:

> [...] to explain the behaviour of an entity, one deduces the behavior from the essence or nature of the entity. [Steiner, 1978a, p. 143].

The previous remark is aimed to face a well-known problem: mathematical truths are commonly regarded as necessary, then it is meaningless to speak of essential properties of a mathematical entity. Thus, in order to escape all the difficulties related to the definition of an essential property of a mathematical entity x, i.e. a property x enjoys in all possible worlds, Steiner introduces the relative notion of *characterizing property*:

> Instead of "essence", I shall speak of characterizing property, by which I mean a property unique to a given entity or structure within a *family* or domain of such entities or structures. (I take the notion of a family or domain undefined in this paper; examples will follow shortly.) We thus have a relative notion, since a given entity can be part of a number

[7]If we regard each criterion: the more explanatory proof is the more abstract (abstracteness); the more explanatory proof is the more general (generality); the more explanatory proof is the proof which can be used to determine –and not to verify!– the result (discoverability); the more explanatory proof is that which can be visualized (visualizability).

> of different domains or families. Even in a single domain, entities may be characterized multiply. [Steiner, 1978a, p. 143]

For Steiner, an explanatory proof depends on such a property, while a non-explanatory proof does not. In particular,

> an explanatory proof makes reference to a characterizing property of an entity mentioned in the theorem, such that from the proof it is evident that the result depends on that property. [Steiner, 1978a, p. 143]

The dependence characterizing property-result comes from the fact that if we try to manipulate the proof, by substituting in it a differerent object of the same domain, the theorem collapses. This introduces us to Steiner's second core-notion about explanation by proofs: *generalizability* – through the variation of a characterizing property. If we deforme the proof varying a certain characterizing property of a related entity, what we obtain in response is a change of the theorem. To every deformation of the proof there corresponds a deformation in the theorem, i.e. to an array of proofs there corresponds an array of theorems. The theorems obtained are proved and explained by the deformations of the original proof. This is what Steiner takes for an explanatory proof to be generalizable. Observe, however, that although Steiner offers some examples, the notion of 'deformation' is left undefined in his discussion. He writes: "Deformation is similarly undefined – it implies not just mechanical substitution, but reworking the proof, holding constant the proof idea" [Steiner, 1978a, p. 147].

To sum up, Steiner offers two criteria for a proof to be considered as explanatory:

C_1 The proof depends on a characterizing property mentioned in the theorem (dependence criterion)

C_2 It is possible to deform the proof "substituting the characterizing property of a related entity" and getting "a related theorem" (generalizability criterion)

He analyzes some examples in order to show how his approach allows us to distinguish an explanatory proof from a non-explanatory one, and how it is also suitable to account for the plausibility of the four previous criteria (abstracteness; generality; discoverability and visualizability). Since we are interested in MEPP, I will not examine all his cases but only what is essential to illustrate his two criteria C_1 and C_2 at work.

As first example he takes the following theorem: the sum of the first n integers is equal to $\frac{n(n+1)}{2}$. Steiner claims that the classical inductive proof of this theorem is not explanatory because it does not characterize anything mentioned in the theorem. Thus the impossibility of deforming the theorem by varying the inductive procedure[8]. Two explanatory proofs (explanatory in Steiner's sense) are given. The following proof:

$$\begin{array}{lllllllllll}
1 & + & 2 & + & 3 & + & ... & + & n & = & S \\
n & + & n-1 & + & n-2 & + & ... & + & 1 & = & S' = S \\
\hline
n+1 & + & n+1 & + & n+1 & + & ... & + & n+1 & = & n(n+1)
\end{array}$$

and a geometrical argument based on the diagram of Figure 1.1. Concerning the geometrical argument, the proof is very simple. Consider the diagram in Figure 1.1: we divide a square of dots, n to a side, along its diagonal; what we get is an isoscele triangle containing $S(n) = 1 + 2 + 3 + ...n$ dots (triangle ABC in the diagram); the square of n^2 dots is composed of two triangles $S(n)$ (ABC and DBC), but the diagonal BC (composed by n dots) is counted twice in the sum of the triangles; thus we have the result: $S(n) + S(n) = n^2 + n$.

Are the two criteria C_1 and C_2 fulfilled in the previous proofs? In both of the cases, according to Steiner, the proof involves a characterizing property. In the first case, the characterizing property is given by the symmetry

[8]Against Steiner, Hafner and Mancosu pointed out that the inductive proof of 'For all n, $1+2+...+n = \frac{n(n+1)}{2}$' *does* allow for deformation. See [Hafner *et al.*, 2005, p. 234-237] for their argument.

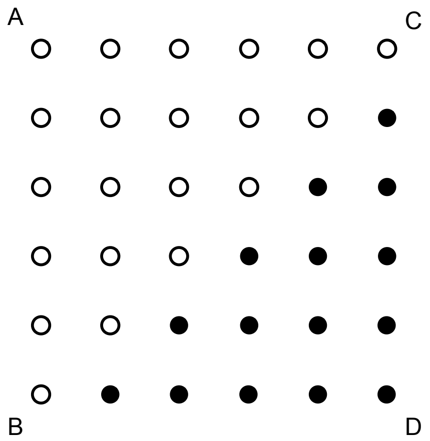

Figure 1.1: A geometrical argument for $1 + 2 + ... + n = \frac{n(n+1)}{2}$.

properties of the sum $1 + 2 + ... + n$, while in the second case by its geometrical properties[9]. Concerning C_2, he suggests that the generalizability is given by the fact that if we vary the symmetry or the geometry we obtain new results[10].

Let's consider his second example. Pointing to the irrationality of $\sqrt{2}$, Steiner claims that using the fundamental theorem of arithmetic (i.e. each number has a unique prime power expansion) we can "argue for the irrationality of the square root of two too swiftly and decisively" [Steiner, 1978a, p. 138] than in the case of the traditional Pythagorean proof.

The Pythagorean proof for the impossibility of $a^2 = 2b^2$ (with a and b positive integers) is considered by Steiner non-explanatory because depending upon the crucial lemma L_c that 'a^2 is divisible by 2 only if a is'. The proof, as it is well-known, runs as follows: we assume that a and b are relatively prime; then, if the equation $a^2 = 2b^2$ is true, a^2 must be divisible by 2; by lemma L_c, if a^2 must be divisible by 2, so must be a; consequently, we

[9] What exactly are these geometrical properties to be considered as characterizing properties? Steiner does not say anything about [Steiner, 1978a, p. 145]. The same point (i.e. "the need for precise definitions here") has been raised by Hafner and Mancosu [Hafner *et al.*, 2005, p. 233].

[10] Again, Steiner is vague and he does not provide any detailed evaluation of how the proof can be generalizable according to his criterion C_2.

obtain the new equation $(2a')^2 = 2b^2$; if we divide by 2 the left and right terms of the equation we obtain that $2a'^2 = b^2$; the last equation shows that also b^2 is divisible by 2, but this contradicts our assumption that a and b are relatively prime (q.e.d.). To generalize this result for other numbers than 2 is to verify lemma L_c for other numbers. This can be made by squaring an arbitrary odd number $2q+1$ and showing that the result must be odd. For each prime p it can be verified that if p divides a^2 it must divide a also, but "the proofs become more and more complex" [Steiner, 1978a, p. 138][11]. Therefore Steiner considers this proof non-explanatory because if we want to generalize the result we have to reprove every time the crucial lemma L_c on which it depends, thus increasing the complexity of the proof. In other words, the fact that the crucial lemma must be proved again every time makes Steiner not accepting this proof as a generalizable proof[12].

Let's now focus on what Steiner considers as an explanatory proof of the irrationality of $\sqrt{2}$, i. e. the proof which uses the fundamental theorem of arithmetic. The proof is the following: consider the prime expansion of a^2 in the equation $a^2 = 2b^2$; focus on the 2's which appear in the prime expansion of the right and the left term of the equation; b^2 is multiplied by 2, thus the prime 2 will appear with an *odd* exponent on the right side of the equation; but, on the left side, the prime 2 which appear in the expansion of a^2 will appear with an *even* exponent (the exponent is doubled in the expansion of a); we can then conclude, because of the uniqueness of the expansion, that a^2 never equals $2b^2$ and the equation is not true. In this case Steiner considers as characterizing property the prime expansion of a number, i.e. the property a number has to have a prime expansion. This uniquely determines a number within the domain of all natural numbers. But what about C_2? The generalizability of the proof is given by the fact that if we vary the object (and thus the characterizing property) by introducing n we get a general theorem:

[11]If $p = 5$, for instance, we have to square $5q + 1$, $5q + 2$, $5q + 3$, $5q + 4$ and show that the result is not divisible by 5.

[12]Steiner's motto is: "It is not the general proof which explains; it is the *generalizable* proof " .

for positive integers n, a, and b, $a^2 = nb^2$ only if n is a perfect square; so the square root of n is either an integer or an irrational[13].

Referring to the four criteria (abstracteness; generality; discoverability and visualizability), Steiner is now able to make some remarks. Concerning generality as a criterion for explanatoriness, he underlines that it is the *generalizable* proof which explains, and not the more general proof[14]. Generality is necessary "for capturing the essence of a particular. [...] To characterize the primes may take the full resources of complex analysis" [Steiner, 1978a, p. 146]. The same holds for abstraction. For instance, in the case of induction, abstraction is considered by Steiner as a useful tool to find a characterizing property and have a better comprehension of what is going on in the proof[15]. In the case of the sum of the first n integers, referring to the proof which makes use of symmetrical considerations, he points out how abstraction is what permits us to formalize the pictorial proof and highlight the symmetry properties of the sum $1 + 2 + 3 + ...n$ [Steiner, 1978a, p. 136, 145]. The same raise in abstraction (quantification over sequences of natural numbers rather than on numbers themselves) makes possible to visualize the geometrical properties of the sum $1 + 2 + ... + n$ in the picture-proof of the sum

[13]How do we arrive at this generalization? Substitute to n any prime 3, 5, 7, ... in the equation $a^2 = nb^2$. Count the occurrences of the considered prime (for instance, count the occurrences of 3 rather than that of 2). The conclusion follows as in the example for 2, because the occurrences will be odd. Now, if we substitute to n a number which is not a prime (for instance: 6, 8, 15, 20, etc..), by considering its prime power expansion ($6 = 2 \cdot 3$, $8 = 2^3$, $15 = 2^0 \cdot 3^1 \cdot 5^1$, $20 = 2^2 \cdot 5^1$) there will be again an odd number of some factor ('2' and '3' in the first case, '2' in the second case, '3' and '5' in the case of $n = 15$, '5' in the case of $n = 20$). Finally, we observe that in order for the equality to be true all the exponents in the prime power expansion of n must be even. In other words, the number to be substituted in the place of 2 must be a square. Steiner considers these remarks (leading to a new theorem) as something evident which emerges from the original proof-idea concerning $\sqrt{2}$; however, other authors do not [Resnik *et al.*, 1987, p. 145].

[14]Note that this contrasts with some remarks by Polya about the importance of generality in proving theorems ([Polya, 1954] and [Polya, 1968]).

[15]Observe that, in order to fix a criterion of abstraction, in his example concerning the sum of the first n integers, Steiner states that the inductive proof is less abstract than the proof which uses symmetry because, when formalized, the former has to do with numbers themselves while the latter quantifies over sequences of natural numbers.

of the first n integers (*Ibid.*)[16]. He discards, however, the suggestion that something counts as a mathematical explanation only if it can be visualized, on the grounds that such a condition would make mathematical explanation "subjective" [Steiner, 1978a, p. 143].

Finally, Steiner rejects "discoverability" as a criterion for explanation and inverts the usual connection which is considered between the two concepts: discoverability, as in the case of the proof of the irrationality of $\sqrt{2}$ using the fundamental theorem of arithmetic, is for him "at best a *symptom* of explanation in mathematics, not a criterion" [Steiner, 1978a, p. 140][17].

1.3 Steiner's test-case of MEPP (revisited)

Now, with Steiner's account of mathematical explanation within mathematics in our hands, we can come back to Steiner's account of MEPP.

As we have seen in section 1.1, for Steiner we have a mathematical explanation in physics when removing the physics (physical assumptions or bridge principles) we rest with a mathematical proof which depends on a characterizing property and could be generalized by varying that property. Or, which is equivalent, we have a mathematical explanation of a physical phenomena when the following criterion is satisfied:

C_{MEPP} If we remove the physics (physical assumptions or bridge principles) we rest with a mathematical proof which satisfies criteria C_1 and C_2

where C_1 and C_2 are the criteria for explanatory proofs introduced in the previous section:

[16]Let me note, again, that Steiner identifies the "geometrical properties [of the sum $1+2+...+n$]" with the characterizing property on which depends the picture-proof based on diagram 1.1 [Steiner, 1978a, p. 145]. However, he does not push further his discussion here, and it is not clear to what geometrical properties he is referring to.

[17]For a discussion on discoverability, proof and explanation in mathematics see [Giaquinto, 2005] and [Giaquinto, 2008]. In particular, Giaquinto offers an example where, using a diagram, we are intuitively led to the discovery of the Pythagora's theorem; however, he does not consider this as a way of proving the theorem. The example shows that there may be discovery without proof.

C_1 The proof depends on a characterizing property mentioned in the theorem (dependence criterion)

C_2 It is possible to deform the proof "substituting the characterizing property of a related entity" and getting "a related theorem" (generalizability criterion)

We can now reconsider Steiner's example of MEPP, given in his [Steiner, 1978b] and introduced in section 1.1, in order to see how criterion C_{MEPP} holds for that case.

Consider again the statement of Euler's theorem 1.1: "The general displacement of a rigid body with one point fixed is a rotation about some axis" [Goldstein, 1957, p. 118]. The statement is very clear. It says that when a rigid body is moved around a fixed point (for instance, its center) it is always possible to find an axis, passing through the fixed point, whose position is the same as before the motion. This is, of course, a physical fact. In this section I am going to show why Steiner considers that criterion C_{MEPP} holds for this case. To do that, I will adopt the following 2-step strategy:

Δ I will show how it is possible to remove from theorem 1.1 the physical assumptions singled out by Steiner, thus obtaining a "pure" mathematical theorem.

Σ I will concentrate on a particular algebraic proof of this mathematical theorem, considered explanatory by Steiner himself, and I will report Steiner's indications about the validity of criteria C_1 and C_2 in the proof.

[**Step** Δ] Since this step requires some technicalities, I am going to proceed gradually. I will report first some kinematical considerations which will be essential to the comprehension of the physical problem. Next, I will show that we can obtain a pure mathematical version of Euler's theorem if we eliminate the physical assumption which are behind the formulation 1.1 given in section 1.1. This mathematical version of the theorem, whose proof will be discussed in Step Σ, will make use of the language of linear algebra and group theory.

In kinematics, the orientation of a rigid body in motion can be described at any instant by three independent parameters (Euler angles θ, ϕ, ψ are commonly used). The Euler angles are defined as three successive angles of rotations. If we consider two coordinate systems with the same origin (not necessarily the center of mass of the object), a Cartesian system of axis fixed in space and another Cartesian system fixed with the rigid body, the motion is described as a transformation from the former system to the latter by means of the three successive rotations performed in a specific sequence (the displacement of the rigid body involves no translation of the body axes and the only change is in orientation). We can identify this orthogonal transformation (a transformation which preserves distances and angles) with an orthogonal matrix the elements of which are espressed in terms of this suitable set of parameters (Euler angles). An orthogonal matrix $\mathbf{A}$ is a matrix with the property that its transpose matrix $\mathbf{A^t}$ is equal to its inverse $\mathbf{A^{-1}}$: $\mathbf{A^t} = \mathbf{A^{-1}}$. It is easy to see that, if a matrix is orthogonal, its determinant is $+1$ or -1[18].

Assume now that at time $t = 0$ the body axes, i.e. the axes fixed with the rigid body, are chosen coincidents with the space axes, i.e. the system of axes parallel to the coordinate axes of external space. With the progression of time the orientation of the rigid body will change and the matrix of the tranformation will evolve continuousily from the identity tranformation $\mathbf{A(0)} = \mathbf{I}$ to the general 3×3 matrix $\mathbf{A}$ with real entries. If the evolution of the matrix is continuous from the unit matrix, which has determinant $+1$, then the determinant of $\mathbf{A}$ cannot change and must be $+1$[19]. Euler's theorem

[18] Since $\mathbf{AA^t} = \mathbf{I}$, we have: $|\mathbf{AA^t}| = |\mathbf{A}||\mathbf{A}^t| = |\mathbf{A}||\mathbf{A}| = |\mathbf{A}|^2 = |\mathbf{I}| = 1$. Here I used two well-known properties of determinants: the fact that, for any square matrix $\mathbf{A}$, $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$, and the property $|\mathbf{A^t}| = |\mathbf{A}|$.

[19] Another method of reaching this conclusion is to observe that a matrix $\mathbf{A}$ with determinant -1 amounts to a transformation in which a right-hand coordinate system is changed into a left-handed system. This transformation, knows as "inversion" or "reflection" of the coordinate axis, cannot be accomplished by any *rigid* change in the orientation of the coordinate axis. In other words, this transformation never corresponds to a physical

states that the body set of axes at any time t can always be obtained by a *single* rotation of the initial set of axes (taken as coincident with the space set). In other words, the operation implied in the matrix $\mathbf{A}$ describing the physical motion of the rigid body is a rotation.

To make the assumption that the operation implied in the matrix $\mathbf{A}$ describing the physical motion of the rigid body is a rotation assures that one direction (the axis of rotation) remains unaffected in the operation, and the magnitude of the vectors is unaffected too. This is to say that, if there exists an axis of rotation, any vector lying along this axis must have the same components in both the initial and the final frame of reference. Observe, however, that the fact that the magnitude of the vectors be unaffected is automatically provided by the orthogonality conditions. Consequently, to prove Euler's theorem 1.1 is to show that there exists a vector $\mathbf{R}$ that has the same components in both the space and the body axis system. This particular vector, called an *eigenvector* of the transformation α, is a non-null vector v that is changed into itself by the transformation: $\alpha(v) = \lambda v$, where $\lambda = +1$. The scalar λ, to which the eigenvector corresponds, is called *eigenvalue*[20]. We can thus restate Euler's theorem in the following form:

Theorem 1.2. *The real orthogonal matrix specifying the physical motion of a rigid body with one point fixed always has the eigenvalue* $+1$ *[Goldstein, 1957, p. 119]*

If we indicate with $\mathbf{R}$ a generic vector and with $\mathbf{R}'$ its transformed vector, to prove the theorem we have to demonstrate that there exists an $\mathbf{R}$ which has the same components in both the space and the body axis system. Such

displacement of a rigid body. This is why the transformations representing rigid body motion must be restricted to matrices having the determinant $+1$. See [Goldstein, 1957, p. 122] for a discussion.

[20] In general, an *eigenvector* of a transformation α (a linear mapping) is every non-null vector v that is transformed in a scalar multiple of itself: $\alpha(v) = \lambda v$. The scalar λ is called *eigenvalue*. To say that there exists an eigenvector corresponding to the eigenvalue $\lambda = +1$ is to say that the transformation changes a vector into itself (the vector has the same components in both the systems and also the same magnitude), i.e. there exists an axis which is left unchanged by the transformation.

a vector **R** will be unchanged by the transformation. In matrix notation, the problem can be formulated as follows:

$$\mathbf{R}' = \mathbf{A}\mathbf{R} = \mathbf{R} \tag{1.1}$$

which is a particular case (for $\lambda = 1$) of the general eigenvalue problem

$$(\mathbf{A} - \lambda\mathbf{1})\mathbf{R} = 0 \tag{1.2}$$

From the theory of linear algebra, we know that we have non trivial solutions for the eigenvalue equation 1.2 when the following determinant vanishes:

$$|\mathbf{A} - \lambda\mathbf{1}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0 \tag{1.3}$$

A proof of Euler's theorem, then, must show that the *characteristic* (or *secular*) equation 1.3 has at least one root $\lambda = +1$.

Now, as we have seen in section 1.1, Steiner claimed that by separating the physical part from the mathematical part in the example, we are left with an explanatory proof of a mathematical theorem concerning transformations and eigenvectors. The previous considerations show that the first statement of Euler's theorem ("The general displacement of a rigid body with one point fixed is a rotation about some axis"), given in section 1.1, is equivalent to the statement 1.2 above ("The real orthogonal matrix specifying the physical motion of a rigid body with one point fixed always has the eigenvalue +1"). However, this is true under the assumption that the real orthogonal matrix *specifies* the physical motion of a rigid body with one point fixed. As we have seen in the kinematical considerations above, in order for this to be true the matrix **A** must have the following properties: it must be 3×3, it must have real entries, it must be orthogonal and it must be proper (i.e. its determinant must be +1). These properties are important because they map some physical ingredients into the mathematical formalism: the fact that the matrix must be 3×3 maps the 3-dimensionality of the real space; the orthogonality condition is required in order for the rigidity of the body to be conserved in the motion (because of the orthogonality condition, in fact, the angles and the magnitude of the vectors are unaffected in the transformation); the fact that the determinant of the matrix be $+1$ eliminates reflections which cannot be accomplished by a rigid motion. Furthermore, our orthogonal 3×3 matrix is a representation of an orthogonal mapping α of the 3-dimensional Euclidean space ($\alpha : V \longrightarrow V$)[21]. Hence, in addition to the 3-dimensionality of the real space, we are also assuming that the real space has an Euclidean structure. To sum up, behind Euler's theorem there are two physical assumptions:

- Space is 3-dimensional euclidean
- The rotation of a rigid body around a point generates an orthogonal, real, proper transformation

Without these assumptions in play, as Steiner observes, we rest with a pure mathematical theorem about transformations and eigenvalues. The theorem says that the real proper 3×3 orthogonal matrix $\mathbf{A}$ has always an eigenvalue $+1$.

As we have seen above, an orthogonal matrix $\mathbf{A}$ is a matrix with the following property: $\mathbf{A^t} = \mathbf{A^{-1}}$ (the transpose coincides with the inverse). Observe now that the class of $n \times n$ orthogonal matrices is a group under matrix multiplication[22]. The group of real orthogonal $n \times n$ matrices is called the *orthogonal group*, and it is denoted by $O(n)$. The property $\mathbf{A^t} = \mathbf{A^{-1}}$

[21] To be precise, the matrix $\mathbf{A}$ represents the mapping with respect to an orthonormal basis, for instance the standard basis $e_1, e_2, ..., e_n$ of R^n. This matrix is orthogonal, and thus $\mathbf{A}^{-1} = \mathbf{A}^t$ holds. An orthogonal mapping requires stronger conditions than a linear one. The condition for transformations of an Euclidean space of dimension n to be orthogonals is that they preserve scalar product. Scalar product preserves angles and distances between points.

[22] A *group* is a set G, together with a binary operation $*$ on G, which has the following properties: 1) for all g and h in G, $g*h \in G$; 2) for all f, g and h in G, $f*(g*h) = (f*g)*h$; 3) there is unique e in G such that for all g in G, $g * e = g = e * g$; 4) if $g \in G$ there is some h in G such that $g * h = e = h * g$.

clearly holds for every real orthogonal matrix which belongs to $O(n)$. Furthermore, as we have seen above, the determinant of an orthogonal matrix can be only $+1$ or -1, and then also $|\mathbf{A}| = \pm 1$ holds for every member of $O(n)$. The subgroup of matrices with determinant $+1$ is called the *special orthogonal group*, and it is denoted by $SO(n)$.

Keep in mind that we regard a $n \times n$ matrix as representation of a linear mapping α of the Euclidean space ($\alpha : V \longrightarrow V$), and n is the dimension of the space ($n = dim(V)$). The characteristic polynomial $P_\alpha(\lambda)$ of the linear map $\alpha : V \longrightarrow V$ is the polynomial $|\mathbf{A} - \lambda \mathbf{I}|$, where $\mathbf{A}$ is our matrix representation of α.

By using the previous terminology of groups, we can restate Euler's theorem in the following mathematical form:

Theorem 1.3. *Every matrix* $\mathbf{A} \in SO(3)$*, with* $\mathbf{A} \neq \mathbf{I_3}$*, has an eigenvalue* $+1$ [23].

This formulation, to which we arrived after some technicalities, is just a modern algebraic formulation of Euler's theorem. It is equivalent to our original formulation 1.1 once the appropriate physical interpretation is established. To establish this interpretation is to say that members of the group $SO(3)$ are linked to continuous motions in physical space to the extent that: ordinary space is three-dimensional Euclidean; the rotation of a rigid body around a point generates an orthogonal, real, proper transformation. On the other hand, observe that if we ignore these assumptions ("we remove the physics") we rest with a pure mathematical theorem, theorem 1.3, whose proof does not depend on any physical assumption.

[**Step** Σ] Steiner considers as explanatory a particular algebraic proof of this theorem, and it will be to this proof that I will refer to. However, before presenting the proof of theorem 1.3, let me consider some results which will be useful for the general proof.

If $\mathbf{A}$ is a real orthogonal 3×3 matrix, then

[23] [Sernesi, 1993, p. 305].

(a) the eigenvalues all have unit magnitude (i.e. the product of any eigenvalue with its complex conjugate is 1).

(b) it has at least one real eigenvalue.

Proof of (a). Result (a) follows from the orthogonality of $\mathbf{A}$. Although all the elements of $\mathbf{A}$ are real, we must consider the possibility that the characteristic equation has complex roots[24]. The magnitude of a complex vector is determined by the sum of the squares of the magnitudes of the components:

$$|X|^2 + |Y|^2 + |Z|^2 = \mathbf{R}^* \cdot \mathbf{R} = |\mathbf{R}|^2 \tag{1.4}$$

The orthogonality condition requires that the transformation leave the magnitude of the vector $\mathbf{R}$ unchanged. Therefore:

$$\mathbf{R}^{*\prime} \cdot \mathbf{R}' = \mathbf{R}^* \cdot \mathbf{R} \tag{1.5}$$

But if $\mathbf{R}$ is eigenvector it must be true that:

$$\mathbf{R}^{*\prime} \cdot \mathbf{R}' = \lambda^* \lambda \mathbf{R}^* \cdot \mathbf{R} \tag{1.6}$$

and hence we have our result:

$$\lambda^* \lambda = 1. \tag{1.7}$$

□

Proof of (b). Result (b) is a direct consequence of two theorems: the fundamental theorem of algebra (FTA) and the complex conjugate root theorem (CCRT). The first theorem states that any polynomial has at least a complex root. If we accept the factorization of a polynomial of degree n, it has exactly n complex roots. The CCRT says that if a polynomial in one variable with real coefficients, such as $|\mathbf{A} - \lambda \mathbf{1}|$, has a complex root λ, then the complex

[24]In this case the corresponding eigenvector is associated to a complex space, and it does not exist in the real space.

conjugate λ^* of λ is also a root of the polynomial. This means that, according to CCRT, if λ is a solution of the polynomial equation $|\mathbf{A} - \lambda\mathbf{1}| = 0$, also λ^* will be a solution of the same equation.

In general, the FTA says that $|\mathbf{A} - \lambda\mathbf{1}| = 0$ will always have complex solutions, but not necessarily real solutions. However, we are considering a real orthogonal 3×3 matrix, then the polynomial $|\mathbf{A} - \lambda\mathbf{1}|$ is a real polynomial (the entries of the matrix $\mathbf{A}$ are real, and so are the coefficients of the polynomial) of *odd* degree (degree 3). By FTA, then, the polynomial equation $|\mathbf{A} - \lambda\mathbf{1}| = 0$ has an odd number of solutions. Now, according to CCRT, complex solutions come in conjugate pairs, and then there are an even number of them in $|\mathbf{A} - \lambda\mathbf{1}| = 0$. Therefore the polynomial equation $|\mathbf{A} - \lambda\mathbf{1}| = 0$ has at least one real solution (i.e. at least one real eigenvalue). Moreover, by result (a) this real eigenvalue can only be $+1$ or -1. $\square$

It can also be proved that, if $\mathbf{A}$ is a real orthogonal 3×3 matrix, its determinant is the product of the three eigenvalues which are solutions to the characteristic equation $|\mathbf{A} - \lambda\mathbf{1}| = 0$: $\lambda_1\lambda_2\lambda_3 = |\mathbf{A}|$. Call this result (c)[25].

Observe that the previous results (a), (b) and (c) do not depend on the 3-dimensionality of the real space or on other physical assumptions. They only depend on some properties of matrices and odd polynomials with real coefficients. This observation is important because it makes clear that the use of these three results in the proof which follows does not require any physical assumption or dependence on the 3-dimensionality of the real space.

Finally, let's consider the proof of Euler's theorem 1.3. This proof, given in Goldstein's *Classical Mechanics* (1957), is the proof to which Steiner refers to[26].

Proof of Euler's theorem 1.3. The proof of the existence of such an eigenvalue is very short and could be stated via a speedy argument which requires

[25] See [Goldstein, 1957, p. 121-122] for a proof.

[26] The very same proof is common in textbooks of linear algebra and group theory. See, for instance, [Grove *et al.*, 1985].

no direct calculations but only results (a), (b), (c) and some background notions. We assume $\mathbf{A} \neq \mathbf{I_3}$ because $\mathbf{A} = \mathbf{I_3}$ is the uninteresting case of the identity transformation[27].

Consider a real matrix $\mathbf{A}$ member of $SO(3)$. For this matrix the determinant is equal to $+1$. Furthermore we know that, in general, the characteristic equation $|\mathbf{A} - \lambda \mathbf{1}| = 0$ has three roots which correspond to three eigenvalues. We want to prove that one of those eigenvalues is our $\lambda = +1$.

Suppose now λ_1, λ_2 and λ_3 are the eigenvalues of $\mathbf{A}$. They are the roots of a cubic polynomial with real coefficients (the entries of the matrix are real). Thus, according to (b), one of the eigenvalue (say, λ_1) is real. By CCRT, if λ_2 is not real, then its complex conjugate λ_2^* is also an eigenvalue ($\lambda_3 = \lambda_2^*$). Since from (c) we have that $|\mathbf{A}| = \lambda_1\lambda_2\lambda_3 = +1$, we have two possibilities for the eigenvalues:

(A) $\lambda_1 = +1$, $\lambda_2 = \lambda_3 = \pm 1$

(B) $\lambda_1 = +1$, $\lambda_2 = \lambda_3^* \notin \mathbf{R}$ (observe here the change in notation for λ_3^*)

In either case we have that $+1$ is eigenvalue[28] $\square$

How then are Steiner's criteria C_1 and C_2 supposed to operate in this case? Recall, first of all, that for Steiner a proof is explanatory only if it

[27] The case $\mathbf{A} = \mathbf{I_3}$ is uninteresting because it corresponds to a transformation which does not involve any change in the coordinate system. The situation corresponds to a rotation of zero degrees. It is easy to see that, for the particular case of $\mathbf{I_3}$, there exist three eigenvalues $\lambda = +1$; this is because the matrix is diagonal with all entries equal to 1, and therefore the characteristic equation is $(\lambda - 1)^3 = 0$. The eigenvectors of the identity matrix $\mathbf{I_3}$ are the unit vectors $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$; all corresponding to the eigenvalue $\lambda = +1$, of multiplicity 3.

[28] What interests us here is the *existence* of the eigenvalue. Observe, however, that the proof given also shows that there is only one eigenvalue $\lambda = +1$. The *unicity* of $\lambda = +1$ can also be proved as a corollary of the more general theorem of Cartan-Dieudonné, by showing that the dimension of the space $\mathrm{Fix}(\alpha)$ of the fixed points of the transformation is 1 [Grove, 2002, p. 49]. I sketch here the proof-idea. If we had that dim $\mathrm{Fix}(\alpha)$=2 (and not 1) we would have $\lambda = +1$ with multiplicity two ($\lambda_1 = \lambda_2 = +1$) and the third eigenvalue, say λ_3, would be -1 or +1. The first case $\lambda_3 = -1$ should be discarded because we know the determinant is +1, while the second case $\lambda_3 = +1$ should be rejected because three eigenvalues with value +1 would give us an eigenspace of dimension 3, which would correspond to the identity transformation (the uninteresting case).

makes evident that the conclusion depends on a property of some entity or structure mentioned in the theorem (criterion C_1). In other words, the first thing to do is to check if the proof does involve such a characterizing property.

Steiner's attention in his paper "Mathematics, explanation and scientific knowledge" [Steiner, 1978b], where the present example of Euler's theorem appears, is not explicitly directed at explanations in pure mathematics, but the focus is elsewhere[29]. This is, perhaps, what precludes him from offering a detailed description of how C_1 and C_2 are fulfilled in the proof above. However, in a personal communication he suggested me that we have to "pick the characterizing property of $SO(3)$ as having an odd dimension". This means that the entity (or structure) mentioned in the theorem is $SO(3)$, which belongs to the family $SO(n)$, while its characterizing property is 'having an odd dimension'. In what sense then does the previous proof *depend* on the odd dimensionality of $SO(3)$? If we concentrate, as Steiner suggests, on the proof strategy above, it is easy to notice that the existence of the eigenvalue $\lambda = +1$ *depends* on the fact that n is odd. For instance, we have seen how result (b) depends on the fact that the degree of the real polynomial $|\mathbf{A} - \lambda \mathbf{1}|$ is odd. And the degree of this polynomial is odd for the members of $SO(n)$ with n odd. Moreover, in the proof strategy we have used the fact that n is odd, together with some considerations on the product of signed numbers, to pick out two possible configurations (A) and (B) for the eigenvalues. A look to these configurations has showed that the eigenvalue $\lambda = +1$ is always a solution of the characteristic equation, which is the result we were looking for. Furthermore, as Steiner rightly observes, there is no *necessity* for any eigenvalue to be $+1$ [Steiner, 1978a, p. 18], because this does not hold for all the elements of the family $SO(n)$. In two or four dimensions, for instance,

[29] As we have seen, his attention goes to the existence of mathematical explanations in science and to the existential conclusions which follow once the existence of such explanations is accepted. I am particularly grateful to Mark Steiner for having discussed with me his account of MEPP, and for his clarifications about how his criteria C_1 and C_2 should be appreciated for the case of Euler's theorem.

the number of real eigenvalues is even and the theorem does not hold[30].

Second, recall that for Steiner 'generalizability' comes when we substitute the characterizing property of a related object and what we get is a related deformed theorem (criterion C_2). Concerning C_2, then, Steiner's idea is that the generalizability of the proof above is obtained by replacing the dimension 3 by some odd number: "There is an explanatory proof of this [existence of an eigenvalue +1] that extends to any Euclidean space of odd dimension" (personal communication). It is easy to see how we can use this strategy to get new theorems. By replacing 3 by some odd number we obtain our related theorem, i.e. a theorem which states the existence of the eigenvalue +1 for every real matrix $\mathbf{A} \in SO(n)$ with n odd (or, which is the same, a theorem which states the existence of an instantaneous axis of rotation for every rotation in a space having odd dimension). For instance, in his *The Classical Groups. Their Invariants and Representations* [Weyl, 1973], Hermann Weyl stated the general theorem as follows:

> A proper rotation in a space of odd dimension has an 'axis' through the origin whose points are fixed under the rotation. [Weyl, 1973, p. 58]

We finally have showed how Steiner's account of MEPP, and then criterion C_{MEPP}, works for the particular test case proposed by Steiner himself.

1.4 Some criticisms

Steiner's account of MEPP has not received much consideration among philosophers of science and of mathematics. To my knowledge, only a recent paper from Alan Baker [Baker, 2009] and Hafner and Mancosu's articles on explanation ([Mancosu, 2000], [Mancosu, 2001], [Hafner *et al.*, 2005], [Hafner *et al.*, 2008], [Mancosu, 2008b]) have devoted some attention to Steiner's

[30] For the case of dimension $n = 2$, real eigenvalues are roots of a quadratic equation and there is no vector in the space which is left unaltered by the rotation. The axis of rotation is perpendicular to the plane and therefore out of the space.

model for mathematical explanation in physics. However, these studies provide only a general discussion of the model and not a detailed analysis or even some testing of it. Perhaps, the only exception is [Baker, 2009], where Steiner's model of MEPP is checked against what Baker considers as a genuine case of MEPP. But, again, this assessment is only sketched and then it cannot be considered as a solid test of Steiner's model of MEPP. On the other hand, Steiner's model for explanation in mathematics has been tested and largely discussed by various authors ([Resnik *et al.*, 1987], [Butchart, 2001], [Hafner *et al.*, 2005] and [Cellucci, 2008]). Moreover, some efforts in improving it have been made [Weber *et al.*, 2002].

In what follows I will concentrate on some of the criticisms which have been leveled against Steiner's model of mathematical explanation in mathematics. This choice is motivated not only by the lack of criticisms about Steiner's model of MEPP. The usefulness of these criticisms in the context of MEPP is clear if we put forward the following observation: since Steiner's account of MEPP depends on his account of mathematical explanation in mathematics, to block or to criticize the latter inevitably means to block or to criticize his account of MEPP. My examination will skip, where possible, the technical details and I will focus more on the general remarks and on the moral of the criticisms. Since they are based on a more detailed testing of Steiner's model of explanation in mathematics, I will take into consideration Resnik and Kushner's [Resnik *et al.*, 1987] and Hafner and Mancosu's [Hafner *et al.*, 2005] criticisms. In the end, I will report Baker's considerations about the difficulties Steiner's model of MEPP has in accounting for a particular example of MEPP. For convenience sake, I will start each criticism with a corresponding label: **RK** for Resnik and Kushner's, **HM** for Hafner and Mancosu's and **Bk** for Baker's.

(**RK**) In their 1987 paper [Resnik *et al.*, 1987], Resnik and Kushner criticized Steiner on various levels (ontological and methodological). Since their analysis is essentially based on Steiner's 1983 paper "Mathematical Realism" [Steiner, 1983], I will begin by shortly illustrating some core-ideas from that paper. Although Steiner's 1983 article is not explicitly devoted to explanation, Steiner's account of mathematical explanation in mathematics is an essential ingredient of the discussion. Furthermore, the fact that this study is balanced towards ontological issues provides an example of how the debate concerning MEPP is linked to the ontological arena in philosophy of mathematics.

In his 1983 paper Steiner endorses Quine's argument that mathematical objects exist because quantifying over them is indispensible in contemporary science and he distinguishes mere existence (in Quine's sense of being the value of a variable) from two types of reality: ontic and epistemic. In what follows here I will consider only the epistemic reality of mathematical objects, since it is this notion which is discussed by Steiner in connection with explanation.

He starts his discussion by presenting his idea of "epistemic independence". For Steiner, an object is "real in the epistemic sense" if our epistemic access to it is conceptually independent of the theory in which the object has been initially postulated[31]. For instance, neutrinos were postulated in order to account for energy missing from certain quantum interaction, and only after they were detected via experiments. The experiments make sense independently of the specific laws that led to the postulation of the physical entities. Steiner goes on by claiming that an object is real in the epistemic sense if it has epistemically independent descriptions: "what is real is what is susceptible to independent description" [Steiner, 1983, p. 369].

By extending his discussion to the case of mathematical entities, he is faced with the following problem: to show that two mathematical descriptions refer to a same thing we can prove their equivalence, but what does it mean for two or more mathematical descriptions to be independent? The notion of independence for mathematical descriptions is thus defined by Steiner by appealing to his theory of mathematical explanation in mathematics. The

[31]Regarding Steiner's notion of ontic reality, it seems that Steiner regards an object as real in the ontic sense just in case reference to it is not eliminable by paraphrase from the language of science [Resnik *et al.*, 1987, p. 144].

definition amounts to the following: (I) two descriptions of a same mathematical object (for example π) are independent when we have a proof that they are coreferential, but there is no explanatory proof (explanatory in Steiner's sense) that they are. If there exist two descriptions of a mathematical entity and these descriptions are independent according to (I), then the mathematical entity described is epistemically real.

Now, recall the notion of characterizing property as defined above (section 1.2). What Steiner is suggesting here is that if we have a non-explanatory proof of the coreference (i.e. a proof which does not appeal to any characterizing property), then the separate descriptions are not connected through essences and should be considered as independent. Hence his notion of explanatory proof is regarded as suitable to distinguish between two kinds of proofs of coreference in mathematics: "those which merely demostrate the coreference and those which explain it" [Steiner, 1983, p. 376]. To have the former and *not* the latter does guarantee the independence of the descriptions (and thus the epistemic reality of some mathematical entity). As an example, he considers the case of the famous formula $e^{\pi i} + 1 = 0$.

Although π has a geometrical description as the ratio between the circumference of a circle to its diameter, we can obtain a separate analytical description of π, which is $arg(-1) = \pi$, through the "magic" identity $e^{\theta i} = cos\theta + sin\theta$ [32]. The coreference of the two descriptions (geometrical and analytical) is established by the following deductive proof:

Proof. We start from number e defined as

$$\lim_{n\to\infty}(1+\frac{1}{n})^n \tag{1.8}$$

We can now consider that, using calculus, is it possible to show that:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + ... \tag{1.9}$$

[32] Recall that $arg(x) = y$ if and only if $e^{yi} = x$.

By introducing complex exponentiations in the expression above (we substitute the imaginary $i\theta$ for x), we have:

$$e^{i\theta} = 1 + i\frac{\theta}{1!} - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} - ... \tag{1.10}$$

$$= (1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + ...) + i(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + ...) \tag{1.11}$$

Nevertheless, for real θ, the same techniques used for the expansion of e^x produce the following results:

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + ... \tag{1.12}$$

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + ... \tag{1.13}$$

Substituting the previous expressions we obtain the identity:

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{1.14}$$

Now, if $\theta = \pi$, we have the analytical result (or description of π):

$$arg(-1) = \pi \tag{1.15}$$

□

The proof we are faced with is a deductive proof which demonstrates the coreference, but which does not explain it. Why? Steiner observes that Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$ which appears in the deductive structure of the previous proof remains a mystery and "no explanatory proof of it has been given" [Steiner, 1983, p. 378][33]. This is why the previous proof must be

[33] Actually, Marcus Giaquinto has pointed out that "an explanation of the aptness" of the Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$ can be given [Giaquinto, 2005, p. 79]. In particular, Giaquinto's idea is that the "aptness" of the formula can be explained by presenting the geometric significance of it: "Consider the point on the unit circle at angle θ (anticlockwise from the unit vector on the x-axis). That point has coordinates $\langle\cos\theta, \sin\theta\rangle$. So it represents the complex number $\cos\theta + i\sin\theta$. Thinking of this as the vector from the origin to the point $\langle\cos\theta, \sin\theta\rangle$, Euler's formula tells us that $e^{i\theta}$ is that vector. If we expand

considered as non-explanatory. Thus the coreference of the two descriptions is established by a non-explanatory proof and, according to criterion (I), the descriptions are independent. Finally, from the independence of the two descriptions of π (the geometrical and the analytical), Steiner asserts that π is real in epistemic sense [Steiner, 1983, p. 376].

Resnik and Kusher's criticism attacks Steiner's notion of explanatory proof and its use as an instrument to establish the epistemic reality of mathematical entities. In order to support their criticism, they offer two counterexamples to Steiner's model of explanation in mathematics:

(α) A proof of the irrationality of $\sqrt{2}$ which meets Steiner's criteria of explanatoriness but which, according to Steiner, does not explain.

(β) A particular proof of the intermediate value theorem which, according to the authors, is explanatory but fails to meet Steiner's criteria of explanatoriness.

Concerning counterexample (α), recall that Steiner does not consider the Pythagorean proof of the irrationality of $\sqrt{2}$ as explanatory because it depends upon the crucial lemma that 'a^2 is divisible by 2 only if a is' (I called L_c such a lemma), and that lemma must be proved again in each case we use another number as substitute for 2. Resnik and Kusher consider again the Pythagorean proof and claim that, by adding a particular lemma, this impediment is removed and Steiner's criteria of explanatoriness are fulfilled. More precisely, they consider 'being the least integer x such that any integer that x divides is also divisible by 2' as characterizing property in the proof [Resnik *et al.*, 1987, p. 147]. This characterizing property allows us to pick out the number 2 among all the positive integers. Observe that the statement

(or contract) the x and y coordinates of that vector by real magnitude r to $r\cos\theta$ and $r\sin\theta$, it is clear that the corresponding vector must also expand or contract by a factor of r. This gives an immediate geometrical significance to the following trivial consequence of Euler's formula: $re^{i\theta} = r\cos\theta + r\sin\theta$. It tell us that $re^{i\theta}$ is the vector with length r at angle θ" [Giaquinto, 2005, p. 79-80]. Although providing a clear illustration of the significance of the Euler's Formula, however, this cannot be regarded as a proof of it.

'if $a^2 = 2b^2$ then 2 divides a' is true for any positive integer that we substitute for 2. Moreover, all but squares divide a^2 only if they divide a (here we use the crucial lemma L_c). By substituting any positive integer for 2, we have therefore a variation of the characterizing property and a deformation of the proof. Moreover, they claim that by deforming the proof in this way what we reach is exactly the same conclusion/generalization reached by Steiner by using the fundamental theorem of arithmetic: a^2 is equal to nb^2 only if the number to be substituted in the place of n is a square. However, note that the crucial lemma L_c is still used in the proof. Here is the key point: Resnik and Kusher point out that, contrary to what Steiner maintained, this lemma must not be proved again in each case, and this because it can be given in the more general form (L^*):

Lemma L^*. *If $a = kn + i$ where $0 < i < k$ (i.e. k does not divide a) and k is not a perfect square, then k does not divide a^2*

Proof of lemma L^*. We proceed by induction on a. We know that $a^2 = (kn)^2 + 2kni + i^2$. So k divides a^2 only if it divides i^2. But $k \neq i^2$, since k is not square; so, applying the inductive hypothesis to i, k divides i^2 only if it divides i. But the latter is impossible since $i < k$ [34]. □

Therefore, by using lemma L^* and the mentioned characterizing property, they offer a proof of the irrationality of $\sqrt{2}$ which meets Steiner's criteria of explanatoriness but which, according to Steiner, does not explain.

Let me now consider their counterexample (β), which will come out again in Hafner and Mancosu's criticism. Here Resnik and Kushner take into account the following version of the intermediate value theorem:

Theorem 1.4 (Intermediate value theorem). *If a real valued function f is continuous on the closed real interval $[a, b]$ and if $f(a) < c < f(b)$, then there is an x in $[a, b]$ for which $f(x) = c$.*

The proof they give, which according to the authors is a modification of a proof from [Rudin, 1953], runs as follows:

[34] The proof is given by Resnik and Kushner [Resnik *et al.*, 1987, p. 147].

Proof of theorem 1.4. Consider the set A of all the t in $[a, b]$ for which $f(t) < c$. This set contains a and is bounded from above by b; the real line is continuous, then the set A has a least upper bound x with $x \leq b$. Since $f(x)$ is continuous on $[a, b]$, $f(x)$ is defined. We will prove by contradiction that $f(x) = c$.
Suppose that $f(x) < c$. Because f is continuous, we can pick a point y to the right of x for which we have $f(y) < c$. But this contradicts the fact that we have isolated *all* such points to the left of x.
Suppose now that $f(x) > c$. Then for each point t in $[a, b]$ to the right of x we have $f(t) > c$. Again, by the continuity of f it will be possible to find a point y' to the left of x such that for all t in $[y', b]$, $f(t) > c$. But this contradicts the fact that x is the least upper bound of A with $x \leq b$. $\square$

Resnik and Kushner consider this proof as explanatory [Resnik *et al.*, 1987, 149]. However, they observe that nothing like Steiner's criteria of explanatoriness can be identified in the proof:

> [...] neither the theorem nor our proof is known to be 'deformable' to yield genuinely new results. In addition, as clear as the proof is, we find it hard to identify the characterizing properties on which it depends. According to Steiner, they should characterize something referred to in the theorem. To what does the theorem refer? Intervals, functions continuous on them and real numbers. The proof clearly depends upon properties of each – e.g., that the function be continuous on an interval-but none of these come close to characterizing any particular function, interval or real number. Perhaps the theorem is really about the class of real valued functions over the real numbers drawn from the family of all classes of functions. This analysis is favored by the use of the same proof-idea in intermediate value theorems about continuous functions on domains other than real intervals. Yet for the proof to count as explanatory in Steiner's sense it must make plain how the theorem changes to new theorems as we move from the class of continuous functions over real closed intervals to any other

> class of functions in the family. Nevertheless, the proof does not tell us anything about discontinuous functions [Resnik *et al.*, 1987, 149]

From their general discussion several considerations emerge. First of all, when we test Steiner's model of explanation on a proof we are faced with a major problem: since Steiner leaves quite indefinite his notion of characterizing property ("a property unique to a given entity or structure within a *family* or domain of such entities or structures" [Steiner, 1978b, p. 143]), it is not clear what we have to consider as "family of mathematical entities or structures" in our test case. To put it in a different way: where is the famous characterizing property in the proof? Furthermore, a second problem arises when we have to fix the limits of the proof deformation (recall the role of deformability in Steiner, and his generalizability criterion C_2). Let's call this latter problem the "deformability problem". These considerations are resumed by a more general criticism, which I will call the "matter of style criticism". Resnik and Kushner observe that:

> Whether or not something is evident from a proof is relative to subgroups of the mathematical community, at best. A proof that explains to a mathematical logician may be anything but evident to a topologist. Then there is the matter of the explicitness of the reference to a characterizing property. That is a matter of style. [Resnik *et al.*, 1987, p. 146]

The "matter of style criticism" suggests then that the two authors do attribute an active role to the context in the potential choice of a characterizing property. Hence, if this criticism is right, Steiner's attempt to "objectively" capture the notion of explanatoriness in proofs would not work, because to consider a proof as explanatory means to consider a proof as explanatory *in a context* (a mathematical community, a classroom, a mathematical tribe living in Amazonia). In other words: objectivity in explanation is lost! This introduces us to the point of view of the two authors. For Resnik and Kushner there is no objective distinction between explanatory and non-explanatory proof. In line with Van Fraassen's pragmatic approach to explanation, which I will present in the next chapter, their answer to the question "Are there explanatory proofs?" is negative and the distinction between explanatory and non-explanatory proofs can only be context-dependent [Resnik *et al.*, 1987, p. 153][35]. The context-dependence is due to the fact that one proof can be explanatory when given as an answer to a particular why-question raised by a questioner (presumably, a mathematician) in a given context. On the other hand, the very same proof can be discarded as non-explanatory by another questioner who is interested in a different why-question concerning the same result[36].

It is evident that, according to Resnik and Kushner's point that there is not explanation *simpliciter* (and, if there is, Steiner's model is not able to capture it), Steiner's account of explanation within mathematics should be rejected[37]. Consequently, his account of the epistemic reality of mathematical entities is in trouble, and the same holds for his model of mathematical explanation of physical phenomena.

In the final part of their paper, they add a comment on the fact that Steiner considers parts of mathematics (for instance, analysis and geometry in his example concerning π) as independent[38]. For Steiner this separability of

[35]It is very curious to observe, that, even if they repeat their resistance to believing in the existence of proofs which 'explain' ("We have doubts that any proof explain" [Resnik *et al.*, 1987, p. 146]), they claim that their examples (for instance, the intermediate value theorem) are examples of 'explanatory' proofs. This point will be considered again in Hafner and Mancosu's criticism.

[36]For instance, regarding the Pythagorean theorem, a questioner might be interested in why the Pythagorean theorem holds *only* for right triangles. And the proof provided by her friend, while correct, might not contain that kind of information.

[37]Resnik and Kushner doubt the existence of explanatory proofs in general [Resnik *et al.*, 1987, p. 146], denying an objective distinction between explanatory and non-explanatory proofs. Furthermore, they claim that "Mathematicians rarely describe themselves as explaining" [Resnik *et al.*, 1987, p. 151]. However, as observed by Hafner and Mancosu, "mathematicians often describe themselves and other mathematicians as explaining. And their judgments concerning explanatory vs non-explanatory proofs (and other varieties of explanation in mathematics as the case may be) has to figure as the basic evidence, however subjective or context dependent they may be. Claims to the effect that certain proofs are explanatory come from within mathematics not from philosophers of mathematics" [Hafner *et al.*, 2005, p. 223-224].

[38]After having given his example of epistemic reality of π, Steiner writes: "Assuming

mathematics into conceptually independent compartments is important because it is on this divisibility that his notions of independent descriptions and epistemic realism are grounded. However, as Resnik and Kushner observe, the history of mathematics shows us that boundaries between mathematical disciplines are *not* likely to exist:

> While we cannot establish conclusively that holism in mathematics is correct, it seems to us that one need only reflect on such disciplines as analytic geometry, algebraic topology, algebraic geometry or the grand foundations such as set theory or category theory to see how implausible it is that fixed conceptual boundaries between mathematical disciplines exist. (Recall, too, that analysis has its historical and conceptual roots in geometry and that its arithmetization gave it 'independent' foundations.) [Resnik *et al.*, 1987, p. 156]

Let me conclude this illustration of Resnik and Kushner's criticism by considering their last remark. Consider Benacerraf's paper "What Numbers Could Not Be" [Benacerraf, 1965]. In that study Benacerraf points out that the natural number sequence is isomorphic to any (recursive) mathematical progression. Furthermore, no mathematical argument (and thus no explanatory proof) can be used to decide if it is identical (or not) to one of these progressions. Hence, according to Resnik and Kushner, since those remarks seem to suggest us that the natural number sequence could be described independently from this or that progression, according to Steiner's criterion of independence it follows that the natural number sequence is epistemically real[39]. On the other hand, they observe, our notion of independence depends

that geometry and analysis are suitably 'independent', our criterion thus yields the reality of π" [Steiner, 1983, p. 376].

[39] Let me note, however, that Resnik and Kushner reading of Steiner's criterion of independent descriptions does not make justice to Steiner's original idea (at least in the context of Benacerraf's considerations). Recall that, for Steiner, two descriptions of a same mathematical object are independent when we have a proof that they are coreferential, but no explanatory proof that they are. In the case of the natural number sequence, and in the context of Benacerraf's 1965 paper, we do *not* have a proof that shows the coreferentiality of the different progressions. Therefore they cannot be considered independent *à la* Steiner. On the other hand, although I consider this particular point raised by the authors

upon a previous division of mathematics into separate theories (*of mathematics*), and thus the problem of holism occurs again:

> [...] the so-called independence of the descriptions in question is dependent upon a prior division of mathematics into separate theories. The claim that the natural number sequence is independently described by $T1$, $T2$, or by geometry or set theory presupposes that these theories are independent parts of mathematics. We are back to the holism issue again and to the problem that a mere reformulation of a branch of mathematics will shift the lines of division within mathematics. [Resnik *et al.*, 1987, p. 156-157]

(**HM**) Let's now consider a second criticism. In their [Hafner *et al.*, 2005], Hafner and Mancosu consider Resnik and Kushner's counterexamples as an insufficient challenge of Steiner's account of explanation in mathematics. The crucial point is that Resnik and Kushner do not offer a justification for regarding their examples as explanatory. They consider a particular proof of the intermediate value theorem as "explanatory" [Resnik *et al.*, 1987, p. 147, 149], but their claim is not grounded on any declaration coming from within the practice of mathematicians. On the other hand, Hafner and Mancosu's request is that:

> For counterexamples to Steiner's theory to carry real weight, they would have to be much more closely related to mathematical practice [Hafner *et al.*, 2005, p. 223]

In other words, Hafner and Mancosu's request is a demand for explanatory-evidence coming from the practice of mathematicians. This is why, to assess Steiner's account of explanations in mathematics, they choose as test-case a proof recognized as explanatory in mathematical practice. This proof comes

as lacking in power to face Steiner's view, their observation is particularly interesting because explores the question of the divisibility of mathematics into 'independent' domains (something Steiner seems to be very sympathetic to!), and I will refer to this question at the end of the next chapter.

from the work of Alfred Pringsheim in the theory of infinite series and concers Kummer's test for convergence[40].

Pringsheim considered as explanatory one particular proof of Kummer's test, then this proof should be considered a *bona fide* example in order to test Steiner's account of explanation in mathematics. Without entering in the technical details of Hafner and Mancosu's assessment of Steiner's account in this context, something which would take us too far from the present discussion, I report here some of their conclusions. Their general verdict is summed up by the following quotation:

> [Hafner and Mancosu] argue that the explanatoriness of the proof of the result in question cannot be accounted for in Steiner's model and, more importantly, this is instrumental in giving a careful and detailed scrutiny of various conceptual components of the model. [Mancosu, 2008b, p. 143-144]

The justification for such a claim comes from the following considerations. According to the authors, Steiner's definition of characterizing property is inapplicable in the context of the considered proof of Kummer's test, and then fails in considering this proof as explanatory (while in mathematical practice it is considered as such). The inapplicability of the notion of characterizing property is due to the fact that nothing like Steiner's characterizing property can be singled out in Pringsheim's proof:

> All 'entities' in Kummer's test are generic, no concrete objects are mentioned in it (apart from the number 0 of course, but the proof is clearly not based on any characterizing property of 0). This generality makes it hard to come up with a property that uniquely determines some entity within a family of them [Hafner *et al.*, 2005, p. 230]

Furthermore, they observe how the only potential candidate as characterizing property must be ruled out because, contrary to Steiner's desideratum,

[40] Kummer's test gives very powerful sufficient conditions for convergence or divergence of a positive series. For a presentation of the test see [Tong, 1994].

it does not pick out any particular entity or structure within a family or domain of entities.

Naturally, this 'lack of characterizing property' blocks also his second criterion, C_2 (generalizability), because the latter presupposes the former.

The impossibility of detecting a characterizing property is connected to a more general point raised by the authors. Recall Steiner's definition (the only one he offers) of characterizing property: a property unique to a given entity or structure within a family or domain of such entities or structures. Without having a definition of 'family', 'entity', 'structure', mentioned in the theorem, Steiner's theory is incomplete and vague, and leaves us with a deep sense of dissatisfaction. We are unable to exactly pick out the characterizing property that Steiner had in mind, and then it is hard to assess his theory of explanation, or even propose some refinement to it [Hafner *et al.*, 2005, p. 232]. Thus, the moral of the present criticism seems to be the following: our satisfaction in mathematics (Pringsheim's consideration about the explanatoriness of one proof of Kummer's test) leaves us with a deep dissatisfaction towards Steiner's model of explanation in mathematics.

To conclude, both criticisms above point to a major defect of Steiner's account of explanation in mathematics: without any constraint on what a 'family' (or 'domain') of mathematical entities is, the choice of characterizing property in a proof is left quite open and is subject to arbitrariness. Moreover, if we are interested in MEPP, this unsolved question also undermines Steiner's model of MEPP, since the latter is based on Steiner's account of explanations within mathematics.

(**Bk**) Finally, let me consider Baker's [Baker, 2009]. To my knowledge, this is the only study within the philosophical literature on explanation in which Steiner's model of MEPP is object of some sort of testing. This is why it is important to consider it here. Unfortunately, as I said at the beginning of this section, the assessment proposed by Baker is very general and it cannot be considered a solid testing of Steiner's model.

To test Steiner's model of MEPP Baker takes a case from evolutionary biology. This example, which is considered by Baker as a genuine case of MEPP and which is presented by him in another paper [Baker, 2005], will be discussed in the next chapter, in section 2.1. However, let me introduce here the general details of the test-case and the explanation regarded by Baker as a genuine mathematical explanation in science.

The specific biological phenomenon to explain concerns the life-cycle of an insect called *periodical cicada*. It has been noted that three species of periodical cicada share the same unusual life-cycles, 13 or 17 years (which are prime numbers). The question raised by biologists is: why are these life cycles prime? An explanation for why prime periods are evolutionary advantageous is given by biologists and is based on some ecological facts, some general biological laws, but also on a number theoretic result [Baker, 2005, p. 233]. More precisely, the number theoretic result which participates in the cicada explanation is a consequence of two lemmas. This amounts to saying that the appeal to mathematics, namely the two number theoretic lemmas, is essential to the overall explanation of the biological phenomenon. Moreover, biologists seem to welcome this explanation [Baker, 2009, p. 617]. This is why the cicada explanation is, according to Baker, a *genuine* case of mathematical explanation in science. There is, of course, more to add about the cicada case and the structure of what Baker considers as a genuine explanation. For the example is discussed by Baker in a very general way in the context of Steiner's model, however, the previous broad presentation is largely sufficient[41].

Let's now turn to Baker's test of Steiner's model of MEPP. Furthermore, keep in mind Steiner's account of MEPP discussed in the previous sections. Baker assumes that the cicada-explanation is a genuine case of MEPP. Therefore, if Steiner's model is correct, it should "recognize" the explanatory character of this test case. However, according to Baker, Steiner's theory faces

[41] Again, I will give a more detailed analysis of Baker's example in the next chapter. When [Baker, 2009] was published the general structure of this dissertation was already settled. This is why I do not offer here a detailed illustration of the cicada example but I remit it to the next chapter. I think, however, that the short presentation given here is sufficient to understand Baker's criticism of Steiner's model.

two difficulties to accomplish this task. A first difficulty is pointed out in the following passage:

> If we apply the "Steiner test" to the periodical cicada example then what is its verdict? We immediately run up against the first problem with the test, which is that it depends on a prior grasp of the notion of "internal" mathematical explanation (of mathematical truths), and this is something for which there is no widely accepted philosophical account [Baker, 2009, p. 623]

Observe that Baker refers to "internal mathematical explanation" as to indicate mathematical explanation within mathematics. Moreover, it is easy to recognize the criterion C_{MEPP} (given in section 1.3) behind what Baker calls "Steiner test". To perform a "Steiner's test" amounts to checking the applicability of C_{MEPP} to the cicada case.

In this passage Baker points out that that there is no general account for mathematical explanations within mathematics. This is something which is true, and which might be pointed out also in the context of models of MEPP. However, I do not see how this can be a problem for Steiner's account of MEPP, for as we have seen in the previous sections Steiner *does offer* his own model of mathematical explanation within mathematics. Rather then, if there is a difficulty for Steiner's model in accounting for the cicada case, this difficulty should be found by testing Steiner's account of 'internal mathematical explanation' (as Baker calls it). To perform a Steiner's test on the cicada example is to check the applicability of C_{MEPP}, but C_{MEPP} requires that the mathematical part of the explanation (in this case the two lemmas) be explanatory according to C_1 and C_2. To be more precise, we have to check if the proofs of the two lemmas which participate in the explanation are judged explanatory by C_1 and C_2. It seems that Baker is conscious of the weakness of his first observation against Steiner. This is why he reinforces his criticism by adding a second observation:

> If we rely here instead on intuitions then it would seem that the cicada explanation will probably not count as genuinely mathematical.

> The key mathematical results are the following two lemmas: [...] The proofs of these two lemmas, while relatively elementary, were not given in the paper; instead readers were referred to Edmund Landau's *Elementary Number Theory* [Landau, 1958]. My own feeling, on reviewing the proofs, is that neither is particularly explanatory. This may in part be because the results established by the lemmas in question are so basic: a few moments' reflection shows why they must be true,even without constructing a formal proof [Baker, 2009, p. 623]

Also in this case, however, Baker's remarks are too vague to block Steiner's account. Baker writes that his "feeling, on reviewing the proofs, is that neither is particularly explanatory". On what grounds is Baker rejecting the explanatory character of these proofs? More important for a testing of Steiner's account, I think, is the question: Are these proofs explanatory in Steiner's sense? What should be checked here is Steiner's account of mathematical explanation in mathematics. Nevertheless, in Baker's discussion there is nothing like a testing of Steiner's criteria C_1 and C_2.

After the two general observations above, Baker writes:

> I conclude that the Steiner test pronounces (albeit weakly) against the cicada explanation being a genuine mathematical explanation of a physical phenomenon. [...] Contra Steiner, I would argue that the evidence from scientific practice indicates that the internal explanatory basis of a piece of mathematics is largely irrelevant to its potential explanatory role in science. [Baker, 2009, p. 623]

Here Baker remarks that scientific practice does provide an evidence for the fact that the explanatory character of a piece of mathematics within mathematics does not affect the explanatory character of mathematical explanations in science. However, and I will mention this just as an aside here, there is no consensus on the fact that the intuitions from the practice of biologists do provide such "evidence" [Saatsi, 2011, p. 153].

Let me conclude by saying that, although relevant because it represents the unique attempt to test Steiner's model of MEPP, Baker's paper does not provide any detailed analysis or solid testing of Steiner's theory. Morever, I think that my remarks above show that Baker pronounces "albeit weakly" against Steiner's account. In Baker's discussion there appears no testing of Steiner's model. And such a testing would require: an identification of the physical (or biological, in this case) principles which links the mathematical ingredients of the cicada-explanation (the lemmas) with the biological part of the explanation; a check of criteria C_1 and C_2 on the proofs of the lemmas. Without such an investigation, it is hard to see how Steiner's model of MEPP might, or might not, account for this explanation.

Chapter 2

Is pragmatics enough?

The aim of this chapter is to give an example from which it emerges the difficulty for two classical theories of scientific explanation in accounting for MEPP. In order to do that, I will consider Baker's paper "Are there genuine mathematical explanations of physical phenomena?" [Baker, 2005] as a starting point for a discussion (advanced, but not elaborated further in Baker's paper) of the problem of extending the D-N model and Van Fraassen's theory of explanation (the 'pragmatic account') as to include MEPP.

In my discussion I will follow this order: after having introduced Baker's 2005 paper, together with his test case of MEPP and his considerations about the possibility of extending the two models, I will shortly consider the D-N model in the context of MEPP and then I will move to a deeper and more inclusive analysis of the pragmatic account. I will give a concise summary of Van Fraassen's theory of explanation, including some important criticisms (Salmon and Kitcher's criticism for Van Fraassen's account, and David Sandborg's criticism for an application of Van Fraassen's theory to mathematical explanations in mathematics), and a defense of it endorsed by Alan Richardson. Finally, I will return to Baker's case in order to analyze the difficulties of the pragmatic account as extended to cover mathematical explanation of physical phenomena (what I will indicate as PET: Pragmatic Extended Theory) and the possibilities to improve such an account in order to include mathematical explanations of physical phenomena. In particular, I will stress that to "save" Van Fraassen's account in the case of mathematical explanations (within mathematics or for physical phenomena) means to abandon some core ideas which stand behind the why-question strategy. These considerations will block Baker's positive idea about the possibility of using a pragmatic account in his analysis of MEPP.

The choice of this line of reasoning will have a twofold consequence: it will show how a classical theory of scientific explanation is in trouble when faced with a case of MEPP and, second, it will stress the importance of elaborating a more specific philosophical strategy as to account for cases of MEPP (such as Baker's) where traditional accounts of scientific explanation fail.

In presenting Sandborg's criticism, I will justify my choice to consider Van Fraassen's account among the WTA approaches to explanation.

2.1 Baker's test case

Baker's paper "Are there genuine mathematical explanations of physical phenomena?" [Baker, 2005] is structured along three interconnected challenges: show the existence of a case of "genuinely mathematical explanation of a physical phenomena" in order to give new insights into the Colyvan-Melia debate[1], offer an indispensability argument which does not rely on holism and

[1]The debate between Mark Colyvan and Joseph Melia, concerning the *right* indispensability of mathematics in science, took place in the review *Mind* along three papers ([Melia, 2000], [Melia, 2002] and [Colyvan, 2002]). More precisely, both the authors agree that the challenge for the realist is to show that there are convincing scientific examples in which positing mathematical abstracta "results in an increase in the same *kind of utility* as that provided by the postulation of theoretical entities" [Melia, 2002, p. 75]. For Melia and Colyvan, 'explanatory power' is an example of such a kind of utility, thus explanation has a role of key importance in this debate. Observe that, in his famous *Science Without Numbers* [Field, 1980], also Hartry Field points out that if the realist can find examples where mathematical posits are indispensable to explanations of physical phenomena, we should believe in the existence of mathematical posits via inference to the best explanation (IBE) [Field, 1980, p. 14-20]. An "Enhanced Indispensability Argument", which explicitly refers to the indispensable explanatory power of mathematics in scientific theories as an instrument to infer the existence of mathematical entities, was discussed during the workshop *Explanation, Indispensability of Mathematics, and Scientific Realism work-*

which shows how mathematics is indispensable to science *in the right way* and, third, support the platonist position with this argument. While the author does not develop a model for MEPP, and his attention is more focused on the implications that the existence of MEPP can have on the platonism-nominalist debate, his discussion of what counts as a "genuine mathematical explanation" is based on the assumption of some accounts of explanation.

Baker presents a case study from evolutionary biology. The specific biological phenomenon concerns the life-cycle of an insect called *periodical cicada*. Why is this phenomenon so interesting? The particularity lies in the fact that "three species of cicada of the genus *Magicicada* share the same unusual life-cycles", 13 or 17, which are prime numbers.

> [...] three species of cicadas of the genus *Magicicada* share the same unusual life-cycle. In each species the nymphal stage remains in the soil for a lengthy period, then the adult cicada emerges after 13 years or 17 years depending on the geographical area. Even more strikingly, this emergence is synchronized among the members of a cicada species in any given area. The adults all emerge within the same few days, they mate, die a few weeks later and then the cycle repeats itself. [Baker, 2005, p. 229]

Magicicada species emerge synchronously over a few weeks of the spring in each local population on scheduled years in eastern North America ([May, 1979], [Williams, 1995]). There are six species of periodical cicadas, three with a 17-year cycle and three with a 13-year cycle. The three species in each life-cycle group are distinctive in size, color, and song. The 17-year cicadas are generally northern, and the 13-year cicadas southern with considerable overlap in their distribution (both life-cycle types may occur in the same forest). While some features of the life-cycle could be explained by referring to ecological constraints (great duration of the cicada life-cycle and presence of two separate life-cycle durations within each cicada species in different regions)

shop (Leeds, January 2009). I will come back to these issues later in this dissertation, in sections 5.2.1 and 8.5.

and evolutionary biological laws (the periodic-syncronized emergence of adult cicadas), the prime-numbered-year cicada life-cycle is quite mysterious and needs a satisfactory explanation [Goles *et al.*, 2001, p. 33].

Why are life-periods prime? Two explanations based on the advantage of prime cycle periods have been offered: one based on avoiding predators [Goles *et al.*, 2001], and the other on the avoidance of hybridization with other species ([Cox *et al.*, 1998], [Yoshimura, 1997]). The former is based on the observation that the particular life-periods of cicadas avoid overlap with the life-periods of other periodical organisms and are therefore beneficial whether the other organisms are predators. The latter is based on the observation that these life-periods avoid overlap with different subspecies and are therefore beneficial since mating between subspecies would produce offspring that would not be coordinated with either subspecies. Both the explanation consider then that the particular life-periods 13 and 17 do minimize overlap with *nearby* (or *lower*) life-periods of other periodical organisms. In a particular ecosystem, the life-periods of cicadas are limited, and then the overlap with the life-period of other organisms must be minimized within a particular range (for instance, the 17-year cicadas are limited by biological constraints to periods from 14 to 18 years)[2].

Baker claims that both these explanations use a number theoretic theo-

[2]To see how a particular period length x minimizes the intersection with other period lengths in a specific range (fixed by biological constraints) is sufficient to find the least common multiple (LCM) for each pair (period length, period length) and observe that the LCM is always greater for the pairs which comprise x. In the context of cicadas, the LCM amounts to the number of years between successive intersections of the cicada life-cycle with the life-cycle of another organism (a predator or a subspecies of cicada). For instance, if we consider the periods 14,15,16,17 and 18, we will find that the LCM for the pairs one member of which comprises the number 17 will be always greater than the LCM for the pairs which do not comprise 17: LCM(16,17)=272, LCM(16,18)=144, LCM(14,17)=238, LCM(15,16)=240, LCM(14,18)=126, LCM(15,18)=90, LCM(15,17)=255, etc. Juha Saatsi has observed how this result can be found by playing with some sticks and assuming that each stick represents a period of time: “Take a bunch of sticks of 14, 15, 16, 17, and 18 cm. You’ll need fewer than 20 sticks of each kind. Lay down sticks of each kind one after another to find the least common multiple (LCM) for each pair (viz. the length at which the two lines comprising 14 cm and 15 cm sticks, say, coincide). You’ll soon find out that the least common multiple is almost always clearly longer for the pairs one member of which comprises 17 cm sticks” [Saatsi, 2011, p. 8].

rem ("prime periods minimize intersections compared to non-prime periods") as an essential element in giving the explanation for why is advantageous for the cicada to have such a prime life-cycle (17 or 13 years).

> The explanation makes use of specific ecological facts, general biological laws, and number theoretic result. My claim is that the purely mathematical component is both essential to the overall explanation and genuinely explanatory on its own right. In particular it explains why prime periods are evolutionary advantageous in this case [Baker, 2005, p. 233]

In particular, for Baker the structure of the explanation common to [Goles *et al.*, 2001] and [Cox *et al.*, 1998] is given by the following 5-steps argument:

(1) **Biological law**: Having a life cycle period which minimizes intersection with other (nearby/lower) periods is evolutionarily advantageous.

(2) **Number theoretic theorem**: prime periods minimize intersection (compared to non-prime periods)

---

(3) **Mixed biological / mathematical law**: Hence organisms with periodic life-cycles are likely to evolve periods that are prime.

Combining the previous result (3) with the following

(4) **Ecological constraint**: Cicadas in ecosystem type, E, are limited by biological constraints to periods from 14 to 18 years.

we obtain the prediction:

(5) Hence cicadas in ecosystem type, E, are likely to evolve 17-years periods[3].

Let me elucidate premise (2). Two numbers are said to be 'coprime' if their greatest common factor is 1 (for instance, the greatest common factor of 9 and 17 is 1, then they are coprime). Consider now the following two lemmas [Landau, 1958]:

[3] The same argument holds for cicadas having a life-cycle period of 13 years. As may be expected, in this case the ecological constraint (4) will be different.

L_1 The least common multiple of m and n is maximal if and only if m and n are coprime.

L_2 A number m is coprime with each number $n < 2m$, $n \neq m$ if and only if m is prime.

The number theoretic result which participates in the cicada explanation (or, better, which participates in the two explanations in terms of predators and hybridization) is a consequence of these two lemmas. It is very simple to see why. The first lemma says that the intersection frequency of two periods of length m and n is maximized when they are coprime. The second lemma says that two numbers m and n, where $n < 2m$ and $n \neq m$, are coprime only if m is prime. Consider now the explanation based on avoiding predators. Predators are assumed to have relatively low cycle periods (for instance, period m). If we use L_1 together with L_2, we have that for a given prime p, and for any pair of numbers n and m both less than p, the least common multiple of p and m is greater than the least common multiple of n and m. In other words, prime numbers maximize their least common multiple relative to all lower numbers. To read this in the context of cicada amounts to saying that a prime life-period (for instance, 17) maximizes the number of years between successive intersections with the life-cycle of predators with lower period lengths. From the same lemmas it follows also that prime numbers maximize their LCM relative to 'nearby' numbers, a situation which corresponds to the case in which a prime life-period maximizes the number of years between successive intersections with the life-cycle of subspecies with similar period lengths.

Thus the number theoretic result "prime periods minimize intersection (compared to non-prime periods)" is essential to the structure of the general explanation (which makes also use of specific ecological facts and general biological laws) and answers to the particular question: "Why are prime periods evolutionarily advantageous?".

Until now Baker has only claimed that "the application of mathematics yields explanatory power" [Baker, 2005, p. 233]. Furthermore, for him the fact that some biologists accept the explanation which relies on the concept of primeness (and then on the number theoretic result which appears in premise (2)) does provide an evidence for the fact that mathematics does play an explanatory role in the present case[4]. However, he has not showed *where* this explanatory power comes from, namely he has not provided or adopted an account of explanation. This issue is raised by Baker in part 3 of his paper, titled "Is the cicada example a genuinely mathematical explanation?".

Let's focus our attention on the term 'genuine'. What does Baker mean with this term? He presents three necessary conditions for a MEPP to be genuine: (A) the application be external to mathematics, (B) the phenomenon must be in need of an explanation, and (C) the phenomenon (primeness of the life-cycle of the cicadas) *must* have been identified independently of the putative explanation (explanation involving the mathematical theorem about primeness).

The last condition requires that the mathematical component involving primeness must not be used as to find a phenomenon that fits exactly for our case, because in this case we would have a prediction and not a genuine explanation. As both conditions (A) and (B), Baker affirms that also this condition is fulfilled in the present case of cicada, because the phenomenon of the 13 and 17 years life-cycles for the cicada was well-known before the developement of number theory as a independent branch of mathematics[5].

[4]This claim is made more explicit in [Baker, 2009]. He writes: "The way biologists talk and write about the cicada case suggests that they do take the mathematics to be explanatory, and this provides good grounds, at least *prima facie*, for adopting this same point of view" [Baker, 2009, p. 625].

[5]Let me note that, although periodical cicadas were discovered by colonists during the XVIIth century, number theoretical results are already presents in Euclid's *Elements*. In particular, in Book VII Euclid presents a method to find the greatest common factor or divisor of two numbers (the procedure is now called "Euclidean algorithm"). If the algorithm ends in 1, the original numbers are coprime, i.e. their greatest common factor is 1. While a modern proof appears in textbooks such as Landau's *Elementary Number Theory* [Landau, 1958], mentioned by Baker, the fact that the intersection frequency of two periods of length m and n is maximized when they are coprime is an implicit result of Euclid's algorithm. On the other hand, to show that *prime* numbers maximize their least common multiple relative to lower and nearby numbers and are then optimals for the present case requires a further step; the required result is contained in lemma L_2.

While satysfying the ‘genuineness’, the foregoing remarks do not offer a satisfactory answer to our complete question “Is the cicada example a genuinely mathematical explanation?”. What remains to be answered is: does the example constitue a genuinely *mathematical* explanation (for the physical phenomenon)? In order to check this, we need to assume an account of explanation. And this is what Baker does. In his words:

> What needs to be checked in the cicada example, therefore, is that the mathematical component of the explanation is explanatory in its own right, rather than functioning as a descriptive or calculational framework for the overall explanation. This is difficult without having in hand some substantive general account of explanation. The philosophical analysis of explanation is itself a thorny issue (and not one we shall attempt to settle here), but it may be useful to canvas the three leading contemporary philosophical accounts of explanation -the causal account, the deductive-nomological account, and the pragmatic account- to see if any of them can fruitfully be applied in the present context. [Baker, 2005, p. 234]

Surprisingly, in considering a case of mathematical explanation in science Baker does not test the application of a specific theory of MEPP such as Steiner's[6]. He considers three general models of scientific explanation: the causal model, the D-N model and Van Fraassen's pragmatic model.

The causal account is rejected as a possible account for MEPP because of its incompatibility with any genuine mathematical explanation [Baker, 2005, p. 234]. Of course, behind the use of the term “causal” there is the implicit adoption of a particular account of causal relation. However, limiting my considerations to standard theories of scientific explanation, here I consider that Baker is referring to the sense of causality as expressed by traditional views such as that put forward by Wesley Salmon in his *Scientific Explanation and*

[6] At the end of the previous chapter I have showed how Baker has recently proposed an assessment of Steiner's model on his cicada-case in his [Baker, 2009], although that assessment is largely inadequate to test and even criticize Steiner's account.

the Causal Structure of the World [Salmon, 1984a][7]. Without entering in the details of this account (see [Salmon, 1989] or [Salmon, 1984a] for a complete discussion), this analysis of causation is clearly inapplicable in cases of MEPP such as the cicada's example[8]. On the other hand, Baker claims that the deductive-nomological account and the pragmatic account "both support the claim that the cicada case study is an example of a genuinely explanatory application of mathematics to science" [Baker, 2005, p. 235]. Although he does not further substantiate this claim, I take it as starting point for a discussion of these models (and in particular the second) in the context of MEPP. In the next section I will focus on the deductive-nomological model.

2.2 Deductive-nomological account and MEPP

In the deductive-nomological model (D-N model) proposed by Carl Hempel and Paul Oppenheim in their famous 1948 paper "Studies in the Logic of Explanation" [Hempel *et al.*, 1948], the explanation of a phenomenon is given

[7]Salmon's idea of what a causal process consists in is based on Reichenbach's "mark criterion": a process is causal if it is capable of transmitting a "mark", i.e. a local modification in the process [Reichenbach, 1958, p. 136]. According to Salmon, causal processes are dinstinct from pseudo-processes (as shadows or moving spots of light on walls) which are not able to transmit marks (i.e. information) [Salmon, 1989, p. 107-111].

[8]A different analysis of causation (more precisely, a sketch of a counterfactual theory of causation), which does not use Salmon's link between causal processes, has been proposed, for instance, by Philip Kicther: "I suggest we can have causation without linking causal processes, and hence causal relations among events at which very peculiar interactions occur. What is critical to the causal claims seems to be the truth of the counterfactuals, not the existence of the processes and the interactions" [Kitcher, 1989, p. 472]. For a pioneer counterfactual theory of causation see [Lewis, 1973], while for recent developments see [Collins *et al.*, 2004]. The problem of causation was seriously addresed by David Hume in his *Enquiry concerning Human Undersanding* [Hume, 1999]. Section IV, part 1, and section VII, parts 1-2, contain his central thoughts on causation and are an attempt to answer the following question: "What is the foundation of all conclusions from experience?" [Hume, 1999, p. 113]. Hume concludes that it is only by repeatedly observing associated events that we can establish the existence of causal relations. According to him, we cannot find a physical connection between the cause and the effect; the connection does not exist in the physical world outside of our own minds, and the relation between cause and effect is custom and habit (in other words, belongs to the psychological domain and not to the physical world). For an abridged discussion of Hume on causation see [De Pierris, 2002], while for a survey of contemporary theories of causation see [Dowe, 2008].

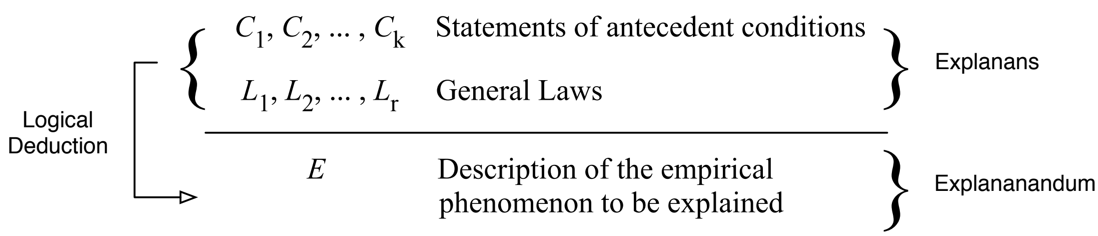

Figure 2.1: Schema summarizing the D-N model.

in terms of a logical relation between a class of sentences (explanans) and a singular sentence (explanandum). A phenomenon is explained when we are able to deduce a statement (the explanandum E), which describes the phenomenon, from some statements (the explanans) which include initial conditions C_k and law-like generalizations L_r (see Figure 2.1)[9]. Moreover, in order for the proposed explanation to be sound, its constituents have to satisfy certain conditions of adequacy, which may be divided into logical and empirical conditions [Hempel *et al.*, 1948, p. 137]:

- Logical conditions of adequacy:

  R1 The explanandum must be a logical consequence of the explanans.

  R2 The explanans must contain general laws, and these must actually be required for the derivation of the explanandum and use no accidental generalizations.

  R3 The explanans must have empirical content: that is, it must capable, at least in principle, of test by experiment and observation.

- Empirical condition of adequacy:

  R4 The sentences constituting the explanans must be true.

Two classical counterexamples undermine the claim that the D-N model provides *sufficient* conditions for successful scientific explanation. They are

[9]Feyerabend has called this the "principle of deducibility": explanation is achieved by deduction in the strict logical sense [Feyerabend, 1962, p. 30].

important for what follows, so I will shortly introduce them here.

The first problem concerns the so called *explanatory asymmetries* (see [Bromberger, 1963] and [Bromberger, 1966]). Explanatory asymmetries appear when we have pairs of deductively valid arguments which rely on the same law but which differ radically in explanatory potential. The classical example is that of the flagpole and the shadow. If we consider a flagpole and its shadow, from informations about the height of the flagpole, the angle θ it makes with the sun plus laws describing the rectilinear propagation of light, we can deduce the length of the shadow. This amounts to a reasonable explanation of the length of the shadow. Nevertheless, the deduction is perfectly legitimate, via the same laws and the same observation on the angle θ, the other way around. But here we have a problem. The problem with this second derivation, in the context of explanation, is that it seems nonsense to say that the length of the shadow *explains* the height of the flagpole (i.e. it is difficult to regard this as an explanation of why the flagpole has that particular height). Nevertheless, the D-N model considers as perfectly legitimate both directions of the explanation, thus lacking in resources to discriminate the good explanation.

A second problem derives from examples of *explanatory irrelevances*, i.e. cases where the logical derivation can satisfy the D-N criteria but it should be considered a faulty explanation because it contains irrelevancies other than those associated with the directional features of explanation. A well-known example is given by Wesley Salmon [Salmon, 1971, p. 34]: (L) All males who take birth control pills regularly fail to be pregnant; (C) Mario Rossi is a male who has been taking birth control pills regularly; (E) Mario Rossi fails to be pregnant. Despite the argument satysfies the requirements of the D-N model, it would be quite strange to consider (L) and (C) together as an explanation of (E).

The previous counterexamples focus on the role of causal considerations in explanation (the height of the flagpole causes the length of its shadow, and not the converse, while taking birth control pills does not cause Mario's failure to get pregnant). Moreover, they show that if we want to preserve the D-N model we have to add to nomic expectability, i.e. "expectability on the basis of lawful connections" [Salmon, 1989, p. 57], some other independent feature X in order to account for directional features of explanation and ensure the explanatory relevance that is missing in the birth control example[10].

A third and more radical counterexample to the D-N scheme is addressed to Hempel's thesis that there exists a logical symmetry between explanation and prediction [Hempel *et al.*, 1948, p. 139][11]. The classical example of "paresis and syphilis", proposed by Michael Scriven in his [Scriven, 1962], is intended to show that this is not always the case. What is more, the D-N model is not able to accommodate singular-causal explanations and it does not even state necessary conditions for acceptable explanation[12]. Consider the (true) statement: 'paresis is caused by untreated syphilis'. This statement does not contrast with the statement that 'syphilis is not often followed by paresis' (the chance that an individual syphilitic develops paresis is low, because only a small percentage – roughly 25% – of those who have untreated latent syphilis become paretic). Now, consider that Mario Rossi, a patient with untreated syphilis, has developed paresis. According to Scriven, we can explain Mario's paresis by saying that 'Mario's paresis was caused by his untreated latent syphilis'. The problem with the Hempelian model is that this explanation has not a D-N structure: the condition cited as explaining the paresis fails to be nomologically sufficient for paresis and it also fails to make paresis high probable. The latter consideration –the fact that the probabilistic cause of the phenomenon only gives it a low probability of occurring – also excludes the use of the Inductive-Statistical (I-S) model, which was proposed by Hempel in his [Hempel, 1965] as to cover the case of statistical

[10]The unification account, which will be discussed in the next chapter, might be regarded as an attempt to develop such an idea and add some feature X to logical conditions.

[11]Hempel advocates the "thesis of structural identity" [Ladyman, 2002, p. 205] according to which explanations and predictions have exactly the same structure. The only difference between them is that in the case of an explanation we already know that the conclusion of the argument is true, whereas in the case of a prediction the conclusion is unknown.

[12] For a discussion of Hempel's defense see [Dietl, 1966].

$$
\begin{array}{ll}
\left.\begin{array}{ll}
\Pr(B|A) = r & \text{Statistical law} \\
A_C & \text{Statement of particular fact}
\end{array}\right\} & \text{Explanans} \\
\hline\hline
\quad [r] & \\
\left.\begin{array}{ll}
B_C & \text{Individual event}
\end{array}\right\} & \text{Explananandum}
\end{array}
$$

Figure 2.2: Schema summarizing the I-S model. The double line indicates that the premises confer the conclusion the probability r, which is supposed to be close to one. The argument *explains* c's being B by showing that this is to be expected, with probability r, in view of the general statistical law and the statement of particular fact included in the explanans.

laws to explain things, and more precisely to include individual events under statistical laws (Figure 2.2). In this model, the relation between explanans and explanandum is inductive rather than deductive, in the sense that it lends more or less strong inductive support to the explanandum sentence. The I-S model could be seen as a natural generalization of the idea which stands behind the D-N model: while the D-N explanation shows that an explanandum was to be expected with certainty, the I-S explanation shows that it was to expected with high probability (to the extent that its explanans confers high probability on the explanandum outcome). Hence, the paresis example represents a case where an explanandum is explained but neither the D-N nor the I-S model are able to exhibit its nomic expectability. In other words, nomic expectability is not a necessary condition for explanation. In Scriven's words: "An event which cannot be predicted from a set of well-confirmed propositions can, if it occurs, be explained by appeal to them" [Scriven, 1959, p. 480].

After this short review of the D-N model, let me consider it in the context of Baker's paper. Baker refers to his 5 step procedure as a layout similar to the layout for the inference poposed by the D-N model. However, premise (2) of Baker's scheme ("prime periods minimize intersection") refers to a mathematical theorem, which does not have empirical content and does not represent a law of nature, thus violating conditions $R2$ and $R3$. Baker is conscious of this problem with the D-N requirements:

> But does the deductive-nomological model have the resources to dinstinguish explanatory from non-explanatory *components* of the explanations? One point in the platonist's favour is that the purely mathematical premise (2) of the cicada inference is in the form of a general law, in this case a theorem of number theory. A broadening of the category of laws of nature to include mathematical theorems and principles, which share commonly cited features such as universality and necessity, would count the mathematical theorem (2) as explanatory on the same grounds as the biological law (1) [Baker, 2005, p. 235].

Although Baker does not endorse any particular account of explanation (as himself has confirmed me in a private conversation), the previous quotation seems to suggest that, in order to distinguish an explanation of physical phenomena in which the mathematics plays an explanatory role from an explanation of physical phenomena in which it does not, the D-N model would need an extension (I will call this model 'D-N Extended'). And this extension should be based on the "broadening of the category of laws of nature to include mathematical theorems and principles". If we follow Baker's suggestion, condition $R3$ of the original D-N model would assume, in the D-N Extended, the following form:

R3* The explanans must have empirical content (i.,e. it must capable, at least in principle, of test by experiment and observation) or it must be a mathematical theorem (i.e. an analytical truth).

In passing, let me observe that it seems quite surprising that here Baker suggests to rehabilitate, for the case of MEPP, a model of scientific explanation which was addressed to explanations and predictions taking the form of logical derivations from observational statements. My aim here is just to jump on Baker's suggestion and show how the D-N Extended lacks the resources to correctly account for MEPP in its deductive structure.

A first observation concerning the D-N Extended has to do with the role of causal claims in scientific explanation. Advocates of the original D-N model referred to types of explanation where it is possible to trace a causal history from the sentences $C_1, C_2, ..., C_k$ to the singular event by empirical regularities $L_1, L_2, ..., L_r$[13].

> If E describes a particular event, then the antecedent circumstances described in the sentences $C_1, C_2, ..., C_k$ may be said jointly to "cause" that event, in the sense that there are certain empirical regularities, expressed by the laws $L_1, L_2, ..., L_r$, which imply that whenever conditions of the kind indicated by $C_1, C_2, ..., C_k$ occur, an event of the kind described in E will take place. Statements such as $L_1, L_2, ..., L_r$, which assert general and unexceptional connections between specified characteristics of events, are customarily called causal, or deterministic, laws. [Hempel *et al.*, 1948, p. 139]

Naturally, to adopt condition $R3^*$ is to abandon the idea that an explanation must provide such a causal linkage. Now, while in the Hempelian account explanations are arguments, by adopting $R3^*$ we are left with a schema which is nothing more than a purely logical deduction. Consequently, explaining an outcome E (in our case a phenomenon mathematically formulated) is just a matter of showing that it is nomically expectable, and then every phenomenon mathematically formulated is recognized by the D-N Extended as a case of MEPP. Although this would be at odds with the evidence coming from the scientific practice (where, as Baker's case shows, scientists do not consider that every application of mathematics yields explanatory power), it

[13]This made the D-N model vulnerable to the problem of adopting an account of cause or causal relation, without simply leaving those notions as primitive. For a discussion of this point and a survey of the major criticisms to the D-N model see [Woodward, 2003]. Although interesting, these criticisms will not affect my discussion in the following lines, so I will not comment on those.

might be thought that this observation cannot be used to block the use of the D-N Extended. In fact, once we adopt the Hempelian perspective that explanation *is* logical deduction, we do not need to resort to such 'evidences' from scientific practice. However, let me observe that the import of such 'evidences' is considered by some philosophers as of extreme importance for the philosophical investigation of MEPP [Mancosu, 2008a]. Therefore the intuitions coming from the scientific practice must be seriously taken into consideration. This is why, in the next lines, I am going to propose a case where the D-N Extended cannot mirror the intuitions coming from the scientific practice, and I will argue that this fact discloses a limitation of this model in the context of MEPP.

There are MEPP which are recognized as such in scientific practice and in which the mathematical component does not come in the form of a theorem, i.e. a statement (as required by $R3^*$). For instance, in their paper "The Explanatory Power of Phase Spaces" [Lyon *et al.*, 2008], Lyon and Colyvan have considered a particular test case and have focused on the explanatory role of phase-space theories [Lyon *et al.*, 2008]. In that case, the fact that the so called Hénon-Heiles systems, i.e. systems formed by a particle moving in a bidimensional potential $U(q_x, q_y) = \frac{1}{2}(q_x^2 + q_y^2) + q_x q_y^2 - \frac{1}{2}q_y^3$, exhibits regular or chaotic motion is deduced visually from a representation in the phase-space[14]. Thus the phase space, with its mathematical apparatus, is regarded to have an explanatory role:

> The explanatory power is in the structure of the phase space and the Poincaré map. So it seems that this is a case where using the phase space is essential to our understanding and ability to explain certain features of the world. [Lyon *et al.*, 2008, p. 14]

[14] I will come back to this example in the last part of this dissertation. Nevertheless, let me add here some details essential to understand how the regular or chaotic motion is deduced visually. We consider the total energy of the system E constant (and thus we lower the dimensionality of the space by one). Next we take a 2-dimensional cross section of this hypersurface in the phase space and then we map the intersections of the trajectories with the plane by using a function called Poincaré Map. Finally, we look at the "dots" made by the solutions (orbits) on the Poincaré section and we can visually grasp qualitative informations about the dynamics of the system at that particular energy.

Now, it is important to note that the mathematical procedure involving phase space is not the only alternative for the study of the system. In fact, it is possible to analyze the system via the Lagrangian formalism, although this route seems not to convey the sense of explanatoriness that we obtain from the use of the phase space theory in the Hamiltonian formalism:

> [...] although there is a Lagrangian formulation of the theory in question that does not employ phase spaces, the cost of adopting such an approach is a loss of explanatory power. [Lyon *et al.*, 2008, p. 2]

From this example two important points emerge: first, even if mathematics comes as an essential ingredient, it is not a particular theorem (i.e. a mathematical law) which participates in the explanation; second, although two mathematical procedures are acceptable as to study the physical phenomenon (regular or chaotic motion of the particle moving in the potential), only one of them contributes to the MEPP. Consequently, in the context of this example, the D-N Extended is confronted with the following problem: the model cannot deal with mathematical operations or procedures which do not come under the form of statements, and therefore it does not recognize the explanation as genuine. What is more, even if we would have such mathematical procedures under the form of theorems, the model would lack in resources to discriminate between the explanatory mathematical procedure and the non-explanatory one. In fact, these procedures are both formally correct, and therefore the D-N Extended would consider both equally explanatory (both are good ingredients of the logical deduction).

The moral of the previous lines is that this model does not accurately describe important aspects of scientific practice, and for this reason should be amended or abandoned for MEPP[15]. Of course, this claim is based on the assumption that the intuitions coming from the scientific practice provide

[15] I believe the situation is worse in cases of mathematical explanation within mathematics, where every legitimate deduction will assume the status of explanation when analyzed through a model which uses a deductive scheme similar to that of the D-N model. Every formal proof inevitably follows a logical deductive schema.

some guidelines in the study of MEPP, and that these intuitions should be mirrored by our philosophical accounts of explanation. Nevertheless, this is something that I regard as natural (and widely accepted) in the context of MEPP. I would be happy about the following situation: the D-N Extended model (or any other model) does identify a MEPP as genuine and that particular MEPP is recognized as genuine in the scientific practice as well. This situation would be, I think, an indicator of the fruitfulness of the philosophical investigation which lies behind that account of explanation. On the other hand, I regard neither reasonable nor philosophically fructuous that a philosophical model of explanation does impose a criterion of explanation on the scientists in a context where the scientists does not agree in considering a particular application of mathematics as explanatory. And this is precisely what happens if we endorse the D-N Extended account in the context of the Hénon-Heiles example. In fact, the model considers the Lagrangian path to the study of the system as explanatory, thus forcing the scientists to regard it as a genuine MEPP (something which scientists do not do in that context, because they consider as explanatory the Hamiltonian route).

Baker himself is conscious of the difficulties with the D-N model and seems to prefere a different approach: the pragmatic account developed by Bas C. Van Fraassen in his book *The Scientific Image* [Van Fraassen, 1980].

2.3 Van Fraassen pragmatic theory and MEPP

According to Baker, the pragmatic theory of explanation is the most apt in order to account for the explanatoriness of the mathematical component in the explanation of the cicada life-cycle.

> According to the pragmatic account, explaining a phenomenon involves providing an answer to a "why" question which shows how the phenomenon is more likely than its alternatives. This is the sketchiest of the three accounts, but perhaps also the most useful in the present context. It suggests that *genuinely explanatory applications of math-*

ematics ought to be reconfigurable as answers to questions about why a certain physical phenomenon occurred. [Baker, 2005, p. 235. My italics]

Therefore, we are confronted again with the idea of extending a model of scientific explanation as to include a mathematical subject and cover MEPP.

Before starting my discussion of the model, let me observe that there is an aspect of Van Fraassen's theory of explanation which is regarded by Baker in a way which is fundamentally different from Van Fraassen's original considerations. This aspect does not concern the technical details of the theory, which are not discussed by Baker, but the commitment to the ontological character of the theory itself[16]. While in Van Fraassen the pragmatic account is not committed to any form of realism[17], and the Inference to the Best Explanation (IBE) is rejected as a plausible criterion to establish the existence of entities postulated by a scientific theory [Van Fraassen, 1980, p. 19][18], in a footnote of his paper Baker points out that "there seems to be no reason why the pragmatic account cannot instead be combined with versions of realism" [Baker, 2005, p. 235][19]. As we will see further in this dissertation,

[16]In passing, let me not that in the previous quotation Baker gives the following characterization of the pragmatic account: "according to the pragmatic account, explaining a phenomenon involves providing an answer to a why question which shows how the phenomenon is *more likely* than its alternatives" (my emphasis). This characterization, however, seems to be more appropriate for a model of explanation such as the inductive statistical model (I-S), which considers explanation as involving the subsumption of individual events under statistical laws.

[17]As the term "pragmatic" suggests, in Van Fraassen the acceptance of a theory has only a pragmatic dimension. The anti-realist flavour of Van Fraassen's philosophy comes in various passages of his book *The Scientific Image*. His theory of explanation is developed as an alternative against "requests for explanation to which realists typically attach an objective validity which anti-realist cannot grant" [Van Fraassen, 1980, p. 13]. This point, as I will show further in this chapter, will be crucial to one of the major criticisms of Van Fraassen's theory of explanation, namely that advanced by Philip Kitcher and Wesley Salmon.

[18]For a criticism of IBE different from Van Fraassen's see, among others, [Friedman, 1983] and [Cartwright, 1983].

[19]Baker is not the only one who disentangles Van Fraassen's account from an anti-realist commitment. See, for example, a similar claim by Alan Richardson in the endnote 22 of his paper in defense of Van Fraassen [Richardson, 1995, p. 128], or Geoffrey Hellman in [Hellman, 1983, p. 232].

Baker admits IBE as a valid instrument for the realist and broadens the use of this principle in the context of the platonist-nominalist debate in philosophy of mathematics.

The previous observation is connected to a more general remark put forward by Baker, which I use as starter for this section. He suggests to consider the case of explanations in science, which involve concret theoretical posits, on par with mathematical explanations in science[20]. This is why Baker considers that the answer-question couple

- Why question: Why is the light from certain galaxies getting bent?
- Answer to the why question: Because there is a black hole between us and the distant galaxies.

could be paralleled to the couple of why-question and direct (*partial*, as Baker suggests) answer[21]:

- Why question: Why do periodical cicadas have prime periods?
- Answer to the why question: Because prime numbers minimize their frequency of intersection with other period lengths.

Let's keep in mind the example of cicadas and the fact that behind the answer which appears in the second couple why-question/answer there is a mathematical theorem.

I will refer to the pragmatic account for mathematical explanation in science (implicitly suggested in Baker but not developed by Van Fraassen or others) as the "Pragmatic-Extended Theory" (PET). While there have been

[20]Of course, it seems that this claim is in itself free from ontological import. However, differently from what we have seen with Steiner in the previous chapter, behind Baker's claim there is the convinction that concrete posits figuring in an explanation are real and so are the mathematical posits which appear in the same explanation. In fact, he applies IBE in the context of realism in mathematics.

[21]The terms "why question" and "direct answer" which appear here are borrowed from the language of Van Fraassen's theory. We will see them below in more detail, but for the present purpose it is sufficient to say that, in Van Fraassen's model, an explanation is an answer (better, a direct-answer) to a 'why-question'.

some efforts to assess and extend Van Fraassen's account in the context of mathematical explanation within mathematics [Sandborg, 1998], in Baker we have the first (implicit) attempt to discuss the pragmatic model for a case of mathematical explanation in science.

Let's now see what the original pragmatic account consists of. After that, I will come back to Baker in order to understand how the PET can be applied to his particular test case of cicadas. Therefore I will report some criticisms which have been leveled against the pragmatic model and, finally, I will make some general comments concerning the applicability of the pragmatic account in the context of MEPP.

2.3.1 Van Fraassen on explanation

Van Fraassen presented his pragmatic theory of scientific explanation in his book *The Scientific Image* [Van Fraassen, 1980][22]. Although sharing some similarities with other approaches to explanation, such as Peter Achinstein's illocutionary theory[23], Van Fraassen's account differentiates from those because it relies on the specific claim that explanations can be properly evaluated with respect to 'why-questions'.

Perhaps the best way to introduce his view on explanation is to quote two passages from Van Fraassen himself:

> There are no explanations in science. How did philosophers come

[22] Actually, the chapter of *The Scientific Image* in which his theory of explanation appears is based on Van Fraassen's previous paper "The Pragmatics of Explanation" [Van Fraassen, 1977].

[23] In his book *The Nature of Explanation* [Achinstein, 1983], Achinstein focused on the *act* of explanation, and not only on the product of such an act. Explanation may refer either to a process (a linguistic performance thorugh which someone explains something to someone) or a product (the content of the linguistic performance). For the same words can be used either to explain or to do not explain, the role of the intention (or illocutionary force) in characterizing the product must be central to an analysis of the product itself. The same sentence 'She drank too much last night' could be used by a physician in order to explain Mary's malaise this morning, and by Mary's husband to criticize his behaviour. Mary's husband speech act is not an explanation and, consequently, what he produced is not an explanation. For a short presentation of Achinstein's illocutionary theory of explanation see [Salmon, 1989, p. 146-150] or [Pitt, 1988, p. 199-222].

> to mislocate explanation among semantic rather than pragmatic relations? [Van Fraassen, 1977, p. 150]

> An explanation is not the same as a proposition, or an argument, or list of proposition. (Analogously, a son is not the same as a man, even if all sons are men, and every men is a son.) An explanation is an answer to a why question. So, a theory of explanation must be a theory of why-questions. [Van Fraassen, 1980, p. 134]

Thus, according to Van Fraassen, there is no scientific explanation *simpliciter*, but explanations are relative to the context dependent why-questions they answer[24]. What is requested in order to respond to the question 'Why is it the case that P?' differs from context to context, and the question arises in a context with a certain body of accepted theory plus information. As answers to context-dependent questions, explanations (better: explanatory evaluations) are themselves context-dependent. The same thesis about the context-dependence of explanatory evaluation has been expressed by other authors as well. Resnik and Kusher, assessing the difficulty of Steiner's model in their paper "Explanation, Independence and Realism in Mathematics" [Resnik *et al.*, 1987], point out that

> Without committing ourselves to the details of Van Fraassen's analysis nor its *prima facie* ontic commitment to propositions we shall adopt its moral that nothing is an explanation *simpliciter* but only relative to the context dependent why-question(s) that it answers. [Resnik *et al.*, 1987, p. 153]

The context dependency of scientific explanation comes, according to Van Fraassen, from the fact that scientic explanations are not pure science but an

[24]Nevertheless, I am considering Van Fraassen's account among the WTA conception of explanation, i.e. the conception of explanation in which the model is designed in order to capture explanation through a single model (and then, we might say, a general notion of explanation). How do I justify the choice to consider Van Fraassen among the WTA models? I will provide a motivation for this choice further in this section, when I will present David Sandborg's criticism of Van Fraassen's model.

application of science, i.e. the use of science to satisfy certain of our desires in a specific context [Van Fraassen, 1980, p. 156]; our desires are always desires for descriptive information and are specific depending on the context, then our evaluation of the information provided differs from context to context. This idea is evident from the way in which Van Fraassen formulated his theory of why-questions, which I am going to summarize here.

Van Fraassen's approach to the general logic of questions was inspired by Belnap and Steel's book *The Logic of Questions and Answers* [Belnap *et al.*, 1976], with some refinements in order to fit the theory of telling answer with other studies on scientific explanation. For Van Fraassen, a necessary prerequisite for an explanation is that there is a why-question. But what exactly a why question is? In the pragmatic model a why-question is a triple $Q = \langle P_k, X, R\rangle$ consisting of:

- a topic P_k
- a constrast class $X = \{P_1, ..., P_k, ...\}$
- a relevance relation R

When we ask "Why P_k?" we refer to a proposition P_k called the *topic* of our question (P_k expresses the fact to be explained, i.e. the explanandum). The *contrast-class* of the question is a set of alternatives, that is, a class X of propositions $\{P_1, ..., P_k, ...\}$ which includes the topic P_k. The propositions (or alternatives) P_i belonging to X are propositions expressing possibilities the questioner is willing to consider, including P_k. Finally, a *relevance relation* R is the "respect-in-which a reason is requested". The relevance relation is used to constrain admissible answers, by specifying what factors will count as explanatorily relevant and thus by distinguishing between different senses of the question. A proposition A is called *relevant* to a why question Q if A bears relation R to the couple $\langle P_k, X\rangle$[25]. Answers to such a question Q

[25]The only formal constraint on the relevance relation is that it obtains between proposed answers and topic/contrast-class pairs. Van Fraassen does not offer any other relevance requirement on R in the formal characterization. I will return to this point, central to Kitcher and Salmon criticism of Van Fraassen model [Kitcher *et al.*, 1987, p. 318], below.

differ from non-answers because they have the following form of words: "P_k *in constrast to* (the rest of) X *because* A", where the word "because" indicates that A is a *reason*. More precisely, the word 'because' guarantees that A is relevant, in this context, to the question, i.e. that it bears relation R to $\langle P_k, X \rangle$. Van Fraassen observes that to consider the claim of relevance as implicitly contained in such an answer (behind the linguistic signal "because") is just a matter of regimentation, in order to avoid the building of the claim of relevance into the question as an explicit conjunct. Thus a definition of the notion of *direct answer*, i.e. what counts as an answer to a why-question, is the following:

> B is a *direct answer* to question $Q = \langle P_k, X, R \rangle$ exactly if there is some proposition A such that A bears relation R to $\langle P_k, X \rangle$ and B is the proposition which is true exactly if (P_k; and for all $i \neq k$, not P_i; and A) is true, where $X = \{P_1, ..., P_k, ...\}$. [Van Fraassen, 1980, p. 144].

For simplicity, call A the *core* of answer B (so that the answer can be abbreviated to "Because A") and the proposition (P_k; and for all $i \neq k$, not P_i; and A) the *central presupposition* of question Q. The above definition of direct answer determines the *presupposition*, namely what a why-question exactly presupposes:

(a) that its topic is true (P_k is true).

(b) that in its contrast class, only the topic is true (each P_i in X is false if $i \neq k$).

(c) that at least one of the proposition that bears its relevance relation to its topic and constrast-class, is also true (there is at least one A true which bears R to $\langle P_k, X \rangle$).

Van Fraassen claims that the presupposition of a why-question, plus some additional requirements, allows us to solve one of the major puzzle theories of scientific explanation are confronted to, that is, the kind of problem introduced by Scriven paresis' example in section 2.2. Following Van Fraassen's terminology, let me call this problem the *problem of the rejection of explanation requests*.

As we have seen when we introduced Hempel's D-N model, one question which we were not able to answer, and which caused tension with that picture of explanation, was: "why did Mario Rossi contract paresis rather than some of his luckier fellow syphilitics"?[26]. Van Fraassen points to Scriven paresis' example as to show that there are cases, in a theory's domain, where the request for explanation is nevertheless rejected because at *that stage*, i.e. during that particular period, that particular request is considered as intrinsically illegittimate. And this is the case of the paresis example, where the medical science still does not have an answer for why Mario Rossi was the only syphilitic who developed paresis[27]. For instance, the Aristotelians asked the Galileans the (illegitimate) question: "Why does a body free of impressed forces retain its velocity?". Newton's theory of gravitation did not offer an explanation of gravitational phenomena, but only a description. It is thus clear that the problem is to have a criterion of what to consider as a *legitimate* explanation request in a particular period, because, to put the problem in Van Fraassen's language, "not everything in a theory's domain is a legitimate topic for a why-question, and that what is, is not determinable *a priori*" [Van Fraassen, 1980, p. 112].

A first solution (but not complete, as we will see in a moment) to the problem of the rejection of explanation request is given, according to Van Fraassen, by the constraints $(a), (b), (c)$ expressed by his presupposition. For instance, if the why question has a false presupposition (ex: the topic P_k is false), then the best response in not a direct answer, for no direct answer could be true, but a corrective answer (a denial of one or more parts of the presupposition, such as 'P_k is not true!'). Nevertheless, something more, as

[26]The expression "rather than some of his luckier fellow syphilitics" comes from Philip Kitcher [Kitcher, 1985b, p. 635].

[27]Of course, the same medical science hopes to find such an answer someday soon.

the example of the paresis shows, must be required if we want an answer *telling* for the topic[28]. In the paresis example, where the why-question has topic 'Mario Rossi contracted paresis, rather than some of his luckier fellow syphilitics', the presupposition is true because: the topic P_k is true; the answer satisfies the relevance relation, i.e. the answer 'because Mario had untreated syphilis' gives the sort of information the questioner has in mind; in the contrast class only the topic P_k is true. However, no telling answer could be given (there is nothing that favours Mario's developing paresis among the contrast class). As Van Fraassen himself remarks: "However, as we shall see, if all three of these presuppositions are true, the question may still not have a *telling* answer" [Van Fraassen, 1980, p. 145]. David Sandborg puts the point in this way:

> Van Fraassen needs an evaluative component for answers beyound the relevance relation, because even if an answer gives the sort of information the questioner has in mind (i.e. satisfies the relevance relation), and all of that information is true, it may still not have any bearing on the topic with respects to the contrast class. [Sandborg, 1998, p. 606]

The strategy in order to decide if an answer is a good answer or an answer better than other answers that might have been given is called 'the process of *evaluation*'. The evaluation of answers to why-questions is made with reference to background information and to the part of science accepted as background theory in the particular context in which the question is posed. We call 'K' this body of accepted background theory and factual information, assumed as the same for the questioner and the audience[29].

[28] While Van Fraassen does not offer a definition of "telling answer" during his discussion of his theory of why-questions, for him "telling" means 'revelant for the topic with respect to its alternatives'. This idea is explicitly stated in his article "The Pragmatics of Explanation" [Van Fraassen, 1977, p. 150]. The paresis' example is intended to illustrate this point and introduce Van Fraassen's theory of telling answers.

[29] Kitcher and Salmon have observed that the body of knowledge of the person S_q who poses the question could not be the same body of knowledge of the respondent S_r, and then two different contexts are involved [Kitcher *et al.*, 1987, p. 318].

Van Fraassen stresses the role of K in the emergence of questions in a precise context. He suggests the expression "the question Q arises in the context" to mean that K implies the central presupposition (P_k; and for all $i \neq k$, not P_i; and A) and K does not imply the denial of any presupposition (a),(b),(c), that is, it is altogether appropriate to raise Q even if we do not know whether there is a direct answer or not, provided the central presupposition is fulfilled.

Assume we are in a context with background K of accepted theory plus information. A question Q with topic P_k and contrast class $X = \{P_1, ..., P_k, ..., P_n\}$ arises. How could the answer *Because A* be evaluated as *telling*, good or better? Van Fraassen suggests that the evaluation of "how much an answer is telling" relies on three different criteria:

1. The fact that A itself is more probable (than other reasons) in light of our knowledge K.

2. The probability that A, and thus the answer, favors the topic P_k against the other members of the contrast class relative to background knowledge (favoring criterion).

3. The fact that the answer is made wholly or partially irrelevant by other answers that could be given.

While the first criterion is straightforward (if the question arises in a context, we ask what probability K bestows on A and later we compare this with the probability which K bestows on the cores of other possible answers), the favoring criterion is less intuitive. To score well, the answer should increase the distance between the probability of the topic and the probabilities of the other members of the contrast class. The idea here is that, in order to avoid trivializations, the evaluation of the answer '*Because A*' to question Q should be made with reference only to a part $K(Q)$ of our knowledge K[30]; $K(Q)$

[30]If it would not be so (and we would consider all knowledge K), the favoring criterion would express exactly the information that the topic is true and the alternatives in X are not (the topic would be already part of our background knowledge), thus making possible trivialization.

constitutes the general theory about these phenomena together with other auxiliary facts which are known but which do not imply the fact to be explained. The way in which the subset $K(Q)$ of K is selected is not specified, and this problem is common to other theories of scientific explanation, as acknowledged by Van Fraassen himself [Van Fraassen, 1980, p. 147][31]. The last criterion concerns the availability of superior answers. It indicates that the answer A loses marks if it has a rival that fares better, perhaps because the rival receives higher probability in light of K, perhaps because the rival favors the topic more than A does[32], perhaps because the rival screens off A from the topic.

Now, with Van Fraassen's theory of *telling answers* in our hands, let us focus again on the why-question "Why is the case that Mario Rossi contracted paresis in contrast to other syphilitics?". For this why-question, as Van Fraassen puts it, there is no answer (and then no explanation) because in the present context we have no information that favours P_k in contrast to other members of X. At the present time, we do not have something (a medical theory, for instance) which makes possible to favour the topic 'Mario Rossi contracted paresis' in contrast to other members of the contrast class which, exactly as Mario himself, have an history of untreated syphilis and could have contracted paresis. Thus the answer 'because Mario had syphilis' is not explanatory in this context. On the other hand, if we consider the different why-question 'Why did Mario get paresis, rather than his colleagues?', the answer 'Because he had syphilis' will represent an explanation because the property of Mario that his colleagues did not have makes his probability of getting paresis higher than theirs (thus excluding other possible answers).

[31] As observed by Wesley Salmon, the subset $K(Q)$ of background knowledge we have to choose in order to evaluate a direct answer reminds Hempel's problem of maximal specificity, i.e. a traditional problem in scientific explanation [Salmon, 1989, p. 145]. On the problem of maximal specificity see [Massey, 1968].

[32] Salmon has pointed out that: "In Van Fraassen's erotetic version of the epistemic conception the 'rather than' condition is preserved through the requirement that an adequate explanation favor the topic of the why-question. This requirement is not tantamount to the high-probability requirement of the inferential version [Hempel's model]" [Salmon, 1984b, p. 301].

To put it simply, Van Fraassen's claim in the context of the paresis' example amounts to the following: by taking into account the contrast-class, a step which is made explicit in the criteria of evaluation 1-3, we have an evaluation of the telling character of our answer and therefore we solve the traditional problem of rejection of requests for explanation. Finally, the context has done the work for us.

The previous lines, together with the paresis' example, were aimed at giving the basic structure of Van Fraassen's theory of why-questions. Nevertheless, now we want to see how this theory might work in the context of MEPP.

Van Fraassen's theory addressed general scientific explanation and did not take into account mathematical why-questions. However, according to Van Fraassen, scientific explanations do not distinguish themselves from ordinary explanations by the form or by the sort of information adduced, but only by the fact that such type of explanations "draw on science to get this information" and that "the criteria of evaluation of how good an explanation it is, are being applied using a scientific theory". Therefore, as observed by David Sandborg:

> An adequate why-question oriented theory of explanations should address mathematical, as well as empirical, why-questions. [...] Mathematical explanations should differ from other types only in their subject-matter; not in any fundamental way. Since why-questions can be and are asked and answered about mathematical facts (for instance, after having been informed that $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + ...$ converges to $\frac{\pi}{4}$, it is certainly reasonable to ask why this is so) as well as empirical ones, the why-question approach should be adaptable to mathematical explanations. [Sandborg, 1998, p. 604]

Let's then address again the cicada case.

2.3.2 Pragmatic account and cicadas

As we have seen, Baker claims that in the case of cicada we have the following couple why-question-(direct) answer:

(*) Why question Q: Why do periodical cicadas have prime periods?

(**) (partial) Answer: Because prime numbers minimize their intersection with other period lengths.

Here the topic P_k is 'periodical cicadas have prime periods', while the proposition A is 'prime numbers minimize their intersection with other period lengths'. Moreover, let me observe that Baker considers (**) as a "partial" answer. The reason for this is that, as we have seen in section 2.1, the number theoretical result 'prime numbers minimize their intersection with other period lengths' represents only an ingredient of the cicada-explanation. The additional ingredients of the explanation come from biology and concern some ecological constraints and evolutionary laws. However, I am going to show that these ingredients can be included in the why-question (*), thus leaving the answer (**) unchanged and complete.

First of all, it is important to remark that there could be different 'readings' of the topic P_k, and every reading has a different contrast class X. For example, if in the topic we stress the expression 'periodical cicadas', our reading of the why-question (*) will be: Why do *periodical cicadas* have prime periods? That is, why periodical cicadas (and not humans, cats or Chinese pandas)? Depending on the interest of the questioner, the same form of words in Q can pose different *contrastive why-questions*, each one with a different "constrastive focus"[33]. To fix a particular reading amounts to indicating what sort of explanatory information is wanted, and therefore it yields the exclusion of events which do not belong to the relevant range of events

[33]Contrastive why-questions have been studied for the first time by [Dretske, 1972]. He calls *constrastive statemens* that statements which embody a particular focus, called "contrastive focus". To have a constrastive focuse implies a "featured exclusion of certain possibilities" [Dretske, 1972, p. 412].

(observe that sometimes Q contains an explicit "rather then..." which facilitates the individuation of the focus). In Baker's example the constrastive focus is located on 'prime periods', and not on 'periodical cicadas', and thus the explanation requests concern the primeness of the life-cycle of periodical cicadas. To put it differently, in the biological context the why-question (*) has the following reading: 'Why do periodical cicadas have *prime periods*?'. More precisely, we do not consider all prime periods but we want an explanation of why periods are particular primes (13 and 17). Therefore the why-question must include some further information about the considered periods, and (*) should be reformulated in a more correct form as follows:

(***) Why cicadas in ecosystem type E have a period of n years, where n is the only prime between p and q, with p and q limits stated by the biological constraints and the evolutionary laws?

Observe that I have included the biological constraints and the evolutionary laws in the why-question (***). Consequently, such constraints do not appear in the answer (**), which is left unchanged and must be regarded as a complete answer.

Now, we want to see if Van Fraassen's theory regards the answer (**) as a correct explanation. What kind of propositions do we find in the contrast class X? Consider, as I did in the illustration of the cicada-case, 17-years periodical cicadas (but the same reasoning holds for 13-years cicadas). Therefore (***) will take the form: 'Why cicadas in ecosystem type E have a period of 17 years, where 17 is the only prime between 14 and 18, with 14 and 18 limits stated by the biological constraints and the evolutionary laws?'. The topic P_k is 'cicadas in ecosystem type E have a period of 17 years, where 17 is the only prime between 14 and 18, with 14 and 18 limits stated by the biological constraints and the evolutionary laws'. By assuming that the period of life-cycle (in years) of cicadas is an integer, in X we find all the alternatives including the topic: {'periodical cicadas have 14-years period', 'periodical cicadas have 15-years period', 'periodical cicadas have 16-years period', 'periodical cicadas have 17-years period', 'periodical cicadas have 18-years period'}. Now, let's see if the central presupposition is fulfilled:

1. P_k is true
2. In X only P_k is true
3. A bears relation R to $\langle P_k, X\rangle$.

At a first look it seems that point 1 and 2 are not problematic. We have a quantity of biological observations that support the evidence of the topic P_k. Moreover, the alternatives in the contrast class are excluded by such observations. However, from a deeper examination we notice that there are two problems with presuppositions 1 and 2. First, the topic include in itself the expression "prime number", which is a mathematical expression. This turns the topic in a very different topic from a topic such as "bodies fall to the Earth". A similar observation is made by Sorin Bangu in the context of the debate about the indispensability arguments, a debate which took place around Baker's paper [Bangu, 2008][34]. Second, the word 'period' refers to a precise time evaluation, which presupposes some chosen unit for time evaluation. But such a unit could be conventional (How do we value the duration of a year? When does a year finish?). Are these real difficulties? The latter problem could be dismissed by considering that years are intrinsically, and not arbitrarily, rooted in the physical characteristics of the example[35].

[34] Bangu points out that, in Baker's example, the mathematical language is essential to the formulation of the question to be answered ("Why cicadas in ecosystem type E have a period of n years, where n is the only *prime* between p and q, with p and q limits stated by the biological constraints?"). Thus the truth of the explanandum presupposes or depends on the truth of a mathematical statement (*primeness* is a mathematical property which applies to numbers). By taking the explanandum, which is a mixed statement (i.e. it contains mathematical plus non-mathematical facts), as being true to comply with the requirements of the IBE strategy (IBE strategy assumes that both the explanans and the explanandum are true statements), Baker assumes mathematical realism *before* he argues for it. This undermines Baker's argument for realism and begs the question against the nominalist. In passing, let me note that Bangu's argument is very similar to Steiner's negative argument about the use of MEPP as to infer the existence of mathematical entities (we have seen Steiner's argument in the previous chapter).

[35] This is the line of defense adopted by Baker [Baker, 2009]. He observes that the choice of year as time unit is not arbitrary because it depends upon the revolution of the

The former problem, which concerns more the indispensability argument's context, can be rejected simply adopting Van Fraassen's point of view: the topic P_k is just a true proposition.

However, the third presupposition is problematic in our case of MEPP, and I am going to show why. Before doing that, and discussing the process of evaluation for PET, let me consider some criticisms of Van Fraassen's theory of explanation advanced by Philip Kitcher, Wesley Salmon and David Sandborg. I will parallel these criticisms with a defense of Van Fraassen endorsed by Alan Richardson. These comments will be very useful to evaluate a possible extension of the pragmatic account in the context of MEPP.

2.3.3 Criticisms

The criticism advanced by Kitcher and Salmon in their paper "Van Fraassen on Explanation" [Kitcher *et al.*, 1987] can be summed up in three mutually dependent points: the impossibility for Van Fraassen's theory of explanation of solving the traditional problem of asymmetry; the fact that his account is subject to trivialization; the fact that, in order to "save" his account by introducing a formal constraint on the relevance relation, Van Fraassen would

Earth around the Sun. There is a direct relationship between the internal rhythms of an organism and the external rhythms of the environment. However, observe that chronobiology, i.e. the systematic scientific study of living timing processes in plants and animals, distinguishes between endogenous (driven internally) and exogenous (driven externally) rhythms (a *rhythm* is a change that is repeated with a similar pattern, for instance the emergence of the periodical ciacada) [Koukkari *et al.*, 2006, p. 90]. The length of time required to repeat a rhythmic cycle is called the *period*, a characteristic that has been used to categorize rhythms into three major groups: circadian (20-28 h), ultradian (<20 h), and infradian (>28 h). For instance, the 27-34 days cycle period of the women's menstrual cycle is considered as an example of infradian rhytms [Presser, 1974]. For particular cases such as the emergence of the periodical cicada, chronobiologists speak of "multiyear cycles". Those kinds of long-term cycles are under study and are considered as *far* from being fully understood: "Multiyear cycles, such as the emergence of the periodical cicada (*Magicicada* spp.) every 13 years in the south and midwestern USA or every 17 years in the northeastern USA, the 8- to 10-year population cycles of the ruffed grouse (*Bonasa umbellus*) in Minnesota, and the 15- to 120-year cycles in flowering of various species of bamboo, are among the spectacular rhythmic events found in nature *that are little understood*" [Koukkari *et al.*, 2006, p. 8. My emphasis]. *Contra* Baker, this claim seems to suggest that there is no consensus within the biologists on the fact that the peculiar life-period of cicadas has been sufficiently explained.

commit himself to the sort of realism he wants to avoid. Let's consider them one by one (I will indicate them with KS_1, KS_2 and KS_3).

(KS_1) Van Fraassen's solution to the classical problem of explanatory asymmetries in explanation comes, as in the case of rejection of explanation requests, from a contextual factor, namely the "contextual relevance". The classical example of the flagpole and the shadow, seen in the previous section during the discussion of the D-N model, is paraphrased by Van Fraassen using the story of the tower and the shadow. Here is the story[36]. Van Fraassen and his friend, the "Chevalier", are enjoying a cup of tea together while sitting on the Chevalier's terrace. There is a tower in the garden, and this tower casts a long shadow on the ground. When asked why the shadow of the tower has such a long length ('Why must that tower have such a long shadow?'), Van Fraassen's host says that it is cast by the tower, of a certain height, and he adds that the tower was built to that height on that particular spot for certain historical reasons. The tower had been erected in honour of Queen Marie Antoniette, in 1930, and at that time the Queen would have been one hundred and seventy-five years old, so the tower had been built exactly that many feet high. This, plus laws of trigonometry, the fact that light travels in straight lines and the sun is not alterable in its course, concludes the Chevalier's explanation for the length of the shadow. That is *his* explanation. However, the Chevalier's housemaid has a different version of the facts: "That tower marks the spot where he [the Chevalier] killed the maid with whom he had been in love to the point of madness". Moreover, she adduces that the Chevalier has built the tower of that precise height in order the shadow would cover the terrace where he first proclaimed his love, every setting sun. According to Van Fraassen, the explanation (the answer to the question "Why is the shadow so long?") given by the Chevalier, based on laws of trigonometry and the straight path of light rays, is exactly on a par with the explanation (the answer to the question "Why is the tower so high?") given by the housemaid and based on the love-story between the

[36]See [Van Fraassen, 1980, p. 132-134] for the whole anecdote.

Chevalier and the maid. The moral is very clear: in this case, it is appropriate to explain the tower's height in terms of the length of the shadow it casts. The particular context will fix a particular relevance relation (for instance, the relevance relation is given by the intentions of the Chevalier in building the tower), and this relevance relation will specify the direction of the explanation. To put it in other words, the asymmetries are accounted for simply by taking into consideration the contextual character of the relation of relevance for a why-question. What determines the relevance is specifically the interest of the questioner and that of the audience (the interlocutors), i.e. the context. Therefore asymmetries are reversible through a change in context, as the story of the tower and the shadow shows.

In *Posterior Analytics* (I.13) Aristotle gives examples of asymmetries. In Van Fraassen's reading of Aristotle, the four causes represent the relevance relations (or explanatory factors) for a why-question; thus the same why question, if formulated in different contexts, is a request for different types of relevance relations[37]. Van Fraassen refers to the following passage from Aristotle (*Posterior Analytics* II.11) as a way of taking into account the importance of the context in requesting explanatory factors for a particular why-question:

> It is possible for the same thing to be the case both with some aim and from necessity -e.g. the light through the lantern; for the finer body passes through the larger pores both from necessity (if light comes about by passing through), and with some aim (in order that we shan't stumble) [Aristotle, CWA 1984, p. 156]

In their analysis Kitcher and Salmon claim that, contrary to what Van Fraassen pointed out, his theory of explanation does not solve the traditional problem of asymmetries. Their point is that the change in context can reverse

[37]However, differently from Van Fraassen, Aristotle's approach to explanation treats explanation as objective, such that x explains y just in case (i) x and y are states of affairs in the world, and (ii) states of affairs of the x-type cause states of affairs of the y-type. See [Shields, 2007, p. 40] for a discussion of Aristotle's conception of explanation in natural philosophy.

the direction of explanation, but every asymmetry that could be generated in Hempel's model is reproducible in Van Fraassen's. Therefore pragmatic constraints cannot really eliminate spurious explanations generated by the asymmetries[38].

For instance, take the case of the tower and the shadow. By considering the why-question Q "Why is the height of the tower h?", we can construe the relevance relation R to be that of intentional relevance (exactly as Van Fraassen does); in this way, the proposition "the tower was built on the spot where he [the Chevalier] killed the maid and it was build of that precise height in order the shadow would cover the terrace where he first proclaimed his love" (the core A of the answer) will be relevant to the topic P_k of Q. The relation R holding between A and $\langle P_k, X\rangle$, where X is a collection of propositions ascribing different heights, can be seen as a 'censored' D-N derivation, i.e. a relation that holds between A and $\langle P_k, X\rangle$ just in case there is a D-N argument that derives the height of the tower (topic P_k as explanandum) from A together with additional premises in $K(Q)$ (the explanans)[39]. Now, if we want to admit the proposition "the height of the tower is h" as the core of the answer to the why-question 'Why is the length of the shadow l?", $K(Q)$ must also contain the proposition ascribing the elevation of the sun and the laws of propagation of light. Thus, according to the authors, "Van Fraassen's theory allows explanations which correspond to those D-N explanations which intuitively run the wrong way", and this should be considered as a mistake [Kitcher *et al.*, 1987, p. 328]. As the D-N model, in fact, Van Fraassen's approach does not privilege the direction of the explanation which gains some sort of explanatory potential from the causal objective relevance relation, and thus it does not distinguish the explanatory merits of the two derivations. This is why they consider Van Fraassen's solution of the problem as a mere presupposition of the solution to the problem, and regard Van

[38]However, let me note that Van Fraassen would say that there is nothing spurious to eliminate, because the "wrong direction" of the explanation should be considerate as legitimate too.

[39]$K(Q)$ is a subset of our knowledge K.

Fraassen's claim to have solved the problem of asymmetry as incorrect.

(KS_2) The second and core criticism to Van Fraassen's theory raised by Kitcher and Salmon is strictly linked to the remarks of the previous lines and points to the fact that the pragmatic approach is subject to trivialization. They illustrate this by showing formally that "any true proposition A can be an indispensable part of an explanation of any topic P_k (with respect to a constrast class X that contains P_k and any assortment of false propositions), and, indeed, that it gets highest marks as an explanation of P_k" [Salmon, 1989, p. 143][40]. This means that, for *every* context and for any pair of true propositions, the first proposition explains the second (i.e. the first proposition is the core of the only explanation of the second). Let's call this claim

(TRV1) For any pair of true propositions A and B, there is a context in which A explains B

They offer the following astrological example which, they claim, has the status of a legitimate and acceptable explanation in Van Fraassen theory. The why-question is given by:

- P_k = 'JFK died on 11/22/1963'
- X = { 'JFK died on 1/1/63'; 'JFK died on 1/2/63'; ..., 'JFK died on 12/31/1963'; 'JFK survived 1963' }
- R = a relation of astral influence defined as a relation between propositions describing the positions of the stars and planets at the time of a person's birth and propositions about that person's fate.

The answer to the why question might have as core A a true description of the positions of the stars and planets at the time of JFK's birth. This, joined with an appropriate astrological theory, would make possible to infer that JFK died on that precise day. Such a piece of information will answer

[40]See [Kitcher *et al.*, 1987, p. 319-322] for the formal argument.

the original why-question and will be maximally telling. Nevertheless, even if Van Fraassen theory allows an expanation of JFK's death in astrological terms, it is quite reasonable to reject (at least in the context of modern science) this answer because astral influence has nohing to do with JFK's death.

Salmon and Kitcher's diagnosis is that, in order to escape trivialization (TRV1), Van Fraassen "needs to supplement his theory of explanation with an account of relevance relations" [Kitcher *et al.*, 1987, p. 323].

(KS_3) However, and here is their third criticism, if Van Fraassen adopts some objective relevance relation among propositions (for instance in order to avoid examples such as the astrological example), then he will be committed precisely to the kind of realism he would like dismiss in his *The Scientific Image*, i.e. the scientific realism that claims for objective relevance relations among propositions a scientific theory must capture to be explanatory ("[the] request for explanation to which realists typically attach an objective validity which anti-realist cannot grant" [Van Fraassen, 1980, p. 13]). Therefore, according to Salmon and Kitcher:

> Van Fraassen would have to revise his account of what it is to accept a scientific theory by adding the idea that acceptance involves believing that the theory has explanatory power as well as believing that it saves the phenomena (or, perhaps, believing that the theory offers the best tradeoff between saving the phenomena and having explanatory power). [Kitcher *et al.*, 1987, p. 330]

An attempt to defend Van Fraassen's solution to the problem of asymmetries from the attacks of Salmon and Kitcher has been made by Alan Richardson [Richardson, 1995]. Without entering in the full details of his paper, let me observe that his defense of Van Fraassen is a reply to the three main points raised by Kitcher and Salmon. Behind the first criticism of those authors, as we have seen in the foregoing lines, there is the assumption that the difference in explanatory potential in the asymmetry problem is due to the fact that one direction of the explanation gest the objectively asymmetric causal order right (the length of the shadow is causally conditioned by the height of the tower), i.e. the objective relevance relation which provides explanatory power is a relation of causal order. By leaving his relevance relation R without any formal constraint, Van Fraassen is accused of being not able to account for this or other kinds of objective relevance relations, and this leaves the original problem of asymmetry without a 'real' solution. Moreover, if we agree with Salmon and Kitcher, this lack of characterization for R leaves Van Fraassen's account open to trivialization. However, Alan Richardson points out that, in his story of the tower and the shadow, Van Fraassen is not trying to provide any kind of objective relevance relation among propositions which, for Kitcher and Salmon, a scientific theory must capture (in order to have explanatory power and support that scientific realism the two authors seem to argue for). What he is offering is a totally new (pragmatic) approach to the problem. In Van Fraassen the theories are used in the form of certain propositions to provide an answer to a particular question raised in a particular occasion, thus the asymmetries do not come out from some objective (causal, for instance) structure. What is more, Van Fraassen does not consider explanations as arguments, then the solution to the asymmetry problem as illustrated in the tower and the shadow story could not be transposed in the Hempelian language (as Kitcher and Salmon do), but should be looked through the lens of Van Fraassen's pragmatic machinery. In the tower and the shadow story, we are interested in judging the explanatory value of the statements:

(8) The height of the tower explains the length of the shadow

(9) The length of the shadow explains the height of the tower

In the opinion of Richardson, the crucial point is that, *pace* Kitcher and Salmon,

> [...] for Van Fraassen, (8) is not an objective fact about the causal structure of the world that our scientific theories seek to capture and which, if captured, provides additional reason to believe these theories; nor (9) is objectively false. Neither (8) is a theoretical judgement

> made from within some particular scientific theory. Rather, judgements like (8) and (9) are themselves radically context dependent judgements about the relation of information provided by scientific theories to types of information requested by people in particular contexts. [Richardson, 1995, p. 113]

The fact that the asymmetries in explanation are, at least in certain cases, reversible comes from the fact that when the context is fixed, the relevance relation is contextually fixed also (because we use information that is relevant to a particular question on a particular occasion)[41]; coming back to the example of the tower and the shadow, we have that in the first context proposition (8) is assertable and (9) is not, while in the second context (9) is assertable and (8) is not[42]. For Richardson (and presumably Van Fraassen) this is enough to understand the asymmetric structures that claims like (8) and (9) have.

The second point, the defense of Van Fraassen's model from Kitcher and Salmon's trivialization (TRV1), is based, again, on the contextual-dependent

[41] Van Fraassen claims that there exist cases of asymmetries which cannot be reversed [Van Fraassen, 1980, p. 132]. A typical example is that of red shift and galactic motion. It is generally believed by cosmologists that the distant galaxies are receding from us at high velocities. The main evidence for this hypothesis is the fact that the light from these galaxies is shifted toward the red end of the spectrum (this phenomenon is called 'red-shift'). Now, given classical physics, the velocity of the galaxy explains the red shift. But here the asymmetry cannot be reversed because there is a context in which the velocity of the galaxy can be used to explain the red shift, while on the contrary there is no context in which the red shift explains the velocity of the galaxies. In fact, the fact that the light from these galaxies is shifted toward the red end of the spectrum does not explain why the galaxies are traveling away from us (no relevence relation can be adduced in this context). The recession of the galaxies is explained on the basis of the big bang, and not by the red shift.

[42] The first context is that in which the Chevalier understands the question of the protagonist as a request of physical details of why the shadow has that length. Consequently, the relevance relation is fixed so as the answer to the question "Why is the shadow so long?" requires the causal details. It is quite evident that if we ask "Why is the tower so high?" in this context, with the same relevance relation, to cite the length of the shadow is irrelevant. Hence (9) is not assertable. On the other direction, the maid answers a question about the height of the tower in a context which fixes the relevance relation as psychological details about why the Chevalier has built a tower so high. She answers the question by adducing the Chevalier's desire that the shadow be long enough to cover the terrace. The interlocutors have fixed the relevance relation and, as a result, in this context (9) is assertable and (8) is not.

character of the pragmatic theory. With respect to (TRV1), Richardson observes that this claim is perfectly consonant to Van Fraassen's theory, and that it does not represent a trivialization. Van Fraassen would accept (TRV1) because from the point of view of his pragmatic theory this is perfectly sound. However, Richardson observes, by offering their astrological example Salmon and Kitcher seems to be making a claim stronger that (TRV1), and precisely they are claiming that:

(TRV2) Within a *given* context, for any pair of true propositions A and B, A explains B because it is always possible to find an adequate relevance relation (i.e. a relevance relation which makes the answer 'Because A' maximally telling)[43]

Now, if Kitcher and Salmon are attacking Van Fraassen by claiming that the pragmatic theory implies (TRV2), Richardson stresses how, in Van Fraassen, the context does not guarantee the availability of *every* extensionally definable relations among propositions as relevance relations. It is the context in which the interrogative occurs that cointain certain theories and

> [...] these theories *tell us* what is and isn't scientifically relevant to the occurrence of P_k. It is from among these scientifically relevant factors that the relevant relation of the question expressed by that interrogative must come. [Richardson, 1995, p. 121]

Therefore he dismisses Kitcher and Salmon's trivialization (TRV2): the why-question, in their counterexample (JFK), does not arise because the astral influence relation is scientifically irrelevant to the occurrence of the topic[44]:

[43] Differently from (TRV1), the claim (TRV2) is not explictly made by Salmon and Kitcher. However, this is what they seem to suggest with their astrological example, and what Alan Richardson takes as their 'strong' claim against Van Fraassen. In think that Richardson's reading of Kitcher and Salmon makes perfectly sense and makes fully explicit their criticism.

[44] Recall that a question Q arises in context K just in case the presuppositions of Q are consistent with K: K implies the central presupposition (P_k; and for all $i \neq k$, not P_i; and A) and K does not imply the denial of any presupposition (a), (b), (c): (a) P_k is true; (b) each P_i in X is false if $i \neq k$; (c) there is at least one A true which bears R to $\langle P_k, X \rangle$.

> No astral influence relation between propositions describing positions of planets and propositions about personal fate is in the models of modern science. [...] No relation containing $\langle A, \langle P_k, X \rangle \rangle$ exists and, hence, no such question arises. [Richardson, 1995, p. 122]

Richardson's point is that Kitcher and Salmon's relation of astral influence does not exist in the present context of modern science, and then no astrological answer will be relevant and maximally telling to the question "Why did JFK die on 11/22/1963?". However, he observes, there exist other contexts as well. In a different context (not scientific, I suppose!) in which the conversants share the belief in the observational adequacy of some astrological theory, it is perfectly reasonable to accept and use a relation of astrological influence as to answer the same why-question. It should be observed that Van Fraassen's constructive empiricism has, at its heart, the notion of *empirical adequacy*, taken to be the aim of science and characterized in model-theoretic terms. A theory is said to be empirically adequate if it "saves the phenomena" by representing that phenomena in terms of *appearances* which are effectively embedded in the theory[45]. To say that a particular (astrological) theory explains a fact (JFK's death) does not entail that that the theory be true or empirically adequate. This emerges very clearly from Van Fraassen's methodological remarks:

> So I conclude that (a) the assertion that theory T explains, or provides an explanation for, fact E does not presuppose or imply that T is true or even empirically adequate, and (b) the assertion that we have an explanation is most simply constructed as meaning that we "have on the books" an acceptable theory which explains. I shall henceforth adopt this construal. [Van Fraassen, 1980, p. 100]

[45]The notion of *embedding* used here is a mathematical one in the sense that there is an isomorphism (a mapping that is one-to-one and onto) between the appearances and sub-structures of the theory, known as the "empirical substructures". Roughly, a theory is said to be *empirically adequate* if all of its claims about observables are true. To use Van Fraassen's words: "A theory is empirically adequate exactly if what it says about the observable things and events in this world, is true – exactly if it saves the phenomena" [Van Fraassen, 1980, p. 12].

In the specific example of JFK's death, when we say that the astrological theory explains JFK's death on that precise day, this does not mean that the astrological-circle-explanation provides some sort of correspondence between the astrological theory and JFK's death (the observable phenomenon) in astrological terms. To put it in a more 'pragmatic' form: the members of the astrological-circle use the astrological theory to explain JFK's death (as written on their astrological books), thus providing an explanation, even if there is no correspondence between the astrological theory and JFK's death *en tant que* fact (i.e. the theory is not empirically adequate). This would save Van Fraassen from Kitcher and Salmon attacks.

Let's move to David Sandborg's criticism of Van Fraassen's theory. His analysis concerns the possibility for Van Fraassen's pragmatic account of dealing with mathematical explanations within mathematics, where the latter come in the form of proofs. Thus the relation R, for the specific case considered by Sandborg, is a relation between a mathematical statement P_k and a class of mathematical statements X. Keep in mind that the criteria of evaluation in the original pragmatic account were expressed by

1. The fact that A itself is more probable (than other reasons) in light of our knowledge K.
2. The probability that the answers favor the topic P_k against the other members of the contrast class relative to background knowledge (favoring criterion).
3. The fact that the answers are made wholly or partially irrelevant by other answers that could be given.

In order to test Van Fraassen's theory on a case of mathematical explanations within mathematics, Sandborg presents an example of proof from Polya's book *Patters of Plausible Inference* [Polya, 1968, p. 147]. The theorem to be proved is

Theorem 2.1. *If the terms of the sequence $a_1, a_2, a_3, ...$ are non-negative real numbers, not all equal to 0, then* $\sum_1^\infty (a_1 a_2 a_3 \dots a_n)^{\frac{1}{n}} < e \sum_1^\infty a_n$.

To prove the theorem we need to introduce an auxiliary sequence $\{c_i\}$ by the formula $c_1\, c_2\, c_3 \dots c_n = (n+1)^n$ for $n = 1, 2, 3 \dots$. If the sequence is introduced in the first step of the proof we are able to go on and, after a series of inequalities and the observation that the sequence defining e – whose general term is $(\frac{k+1}{k})^k$ – is increasing, prove the theorem. The proof runs as follows:

Proof. We introduce the auxiliary sequence $\{c_i\}$ in $\sum_1^\infty (a_1 a_2 a_3 \dots a_n)^{\frac{1}{n}}$ and we obtain the equality:

$$\sum_1^\infty (a_1 a_2 a_3 \dots a_n)^{\frac{1}{n}} = \sum_1^\infty \frac{(a_1 c_1 a_2 c_2 a_3 c_3 \dots a_n c_n)^{\frac{1}{n}}}{n+1} \tag{2.1}$$

By using the following inequality between the arithmetic and the geometric means

$$(a_1 a_2 a_3 \dots a_n)^{\frac{1}{n}} \leq \frac{(a_1 + a_2 + a_3 + \ \dots \ + a_n)}{n} \tag{2.2}$$

we obtain:

$$\begin{aligned}
\sum_1^\infty \frac{(a_1 c_1 a_2 c_2 a_3 c_3 \dots a_n c_n)^{\frac{1}{n}}}{n+1} \quad &\leq \quad \sum_1^\infty \frac{(a_1 c_1 + a_2 c_2 + a_3 c_3 \dots + a_n c_n)}{n(n+1)} \\
&= \sum_{k=1}^\infty a_k c_k \sum_{n=k}^\infty \frac{1}{n(n+1)} \\
&= \sum_{k=1}^\infty a_k c_k \sum_{n=k}^\infty (\frac{1}{n} - \frac{1}{n+1}) \\
&= \sum_{k=1}^\infty a_k \frac{(k+1)^k}{k^{k-1}} \frac{1}{k} \\
&< e \sum_{k=1}^\infty a_k
\end{aligned}$$

□

Even if the proof is rigorous and leads to the desired result, it leaves us with a deep sense of dissatisfaction. We ask ourselves: where does the auxiliary sequence come from? The step of introducing the sequence $\{c_i\}$ is a *deus ex machina* step, as Polya writes, because it is visibly important but its connection with the aim is not visible at all. The sequence "appears as a rubbit pulled out of a hat" [Polya, 1968, p. 147] and we have the impression that it comes out from nowhere. This is why Polya tried to reconstruct the proof in a different way, as to make the introduction of $\{c_i\}$, and thus the derivation, more understandable. He thus provided a mathematical explanation of the proof by following successive steps. Since an understanding of his procedure is important for what follows, I will shortly resume it[46].

First of all, Polya observed that theorem 2.1 could be seen as a lemma of the theorem "if the series with positive terms $\sum_{n=1}^{\infty} a_n$ is convergent, then the series $\sum_{n=1}^{\infty}(a_1a_2a_3 \dots a_n)^{\frac{1}{n}}$ also converges". By focusing on the term $\sum_1^{\infty}(a_1a_2a_3 \dots a_n)^{\frac{1}{n}}$, we can apply the inequality between the geometric and the arithmetic means (equation 2.2), thus obtaining:

$$\sum_{n=1}^{\infty}(a_1a_2a_3 \dots a_n)^{\frac{1}{n}} \leq \sum_{n=1}^{\infty} \frac{(a_1 + a_2 + a_3 \dots + a_n)}{n} \tag{2.3}$$

$$= \sum_{k=1}^{\infty} a_k \sum_{n=k}^{\infty} \frac{1}{n} \tag{2.4}$$

This result nevertheless is of no help for what we want to prove, because the series $\sum_{n=k}^{\infty} \frac{1}{n}$ diverges. Polya's next step, which he calls the 'learning from failure step', consists in individuating what was wrong with the foregoing reasoning. The series $a_1, a_2, \dots, a_n$ converges, then a_n is small when n is large. The mistake in the previous procedure is quite evident if we observe that the two sides of the inequality 2.3 will tend to somewhat not equal. Thus, to apply the inequality 2.2 to those too unequal quantities was not a good choice. What we need is to balance the two sides (i.e. making the terms in the inequality more equal) by introducing some increasing compensating factors. We thus multiply a_i by some increasing factor, and by doing so we have

[46] For the complete argument see [Polya, 1968, p. 148-152].

modified our approach to the problem (step three). Now, what quantities do we use as increasing factors? Polya proposes to consider the very general quantities $1^\lambda a_1, 2^\lambda a_2, 3^\lambda a_3, ..., n^\lambda a_n$, where the most advantageous value of λ is to be found. By using those quantities as increasing factors we obtain:

$$\sum_{n=1}^{\infty}(a_1\, a_2\, a_3\, ...\, a_n)^{\frac{1}{n}} = \sum_{n=1}^{\infty}\frac{a_1 1^\lambda \cdot a_2 2^\lambda \cdot\ ...\ \cdot a_n n^\lambda}{(1\cdot 2\cdot ...\cdot n)^{\frac{\lambda}{n}}} \tag{2.5}$$

$$\leq \sum_{n=1}^{\infty}\frac{a_1 1^\lambda + a_2 2^\lambda + ... + a_n n^\lambda}{n(n!)^{\frac{\lambda}{n}}} \tag{2.6}$$

$$= \sum_{k=1}^{\infty} a_k k^k \sum_{n=k}^{\infty}\frac{1}{n(n!)^{\frac{\lambda}{n}}} \tag{2.7}$$

The problem, at this point, is that we cannot calculate the last sum. If we proceed by approximation, from the fact that $n(n!)^{\frac{1}{n}} \approx ne^{-1}$ we have

$$\begin{aligned}\sum_{n=k}^{\infty}\frac{1}{n(n!)^{\frac{\lambda}{n}}} &\approx e^\lambda \sum_{n=k}^{\infty} n^{-1-\lambda}\\ &\approx e^\lambda \int_k^\infty x^{-1-\lambda}dx\\ &= e^\lambda \lambda^{-1} k^{-\lambda}\end{aligned}$$

Introducing this approximation into 2.7, we arrive very close to the desired result:

$$\sum_{n=1}^{\infty}(a_1\, a_2\, a_3\, ...\, a_n)^{\frac{1}{n}} \leq C\sum_{k=1}^{\infty} a_k \tag{2.8}$$

where C is a constant $(e^\lambda\lambda^{-1})$. But we are now in a position to choose a value for λ, and precisely we try with that value that makes $e^\lambda\lambda^{-1}$ a minimum. By differential calculus we have $\lambda = 1$. We can look back at the original question about the choice of increasing factors (call this step four or 'looking back step'), and assume that the choice of a compensating factor multiplying a_n as n, or some quantity similar to n when n is large, could be a good choice in order to reach the value $C = e$ in equation 2.8. However, as we have seen, to fix $\lambda = 1$ leads to something that we are not able to calculate. One of the ideas in the foregoing passages was to to leave λ indeterminate. So, Polya asks, why not introduce more flexibility in our strategy? This is why in the final step of his explanation ('more flexibility step') Polya suggests to leave the compensating factor that multiplies a_n indeterminate (call it c_n), thus obtaining

$$\begin{aligned}
\sum_{n=1}^{\infty}(a_1\, a_2\, a_3\, ...\, a_n)^{\frac{1}{n}} &= \sum_{n=1}^{\infty} \frac{a_1c_1 \cdot a_2c_2\, ...\, \cdot a_nc_n}{(c_1 \cdot c_2 \cdot ... \cdot c_n)^{\frac{1}{n}}} && (2.9)\\
&\leq \sum_{n=1}^{\infty} \frac{a_1c_1 + a_2c_2 +\ ...\ + a_nc_n}{n(c_1c_2\, ...\, c_n)^{\frac{1}{n}}} && (2.10)\\
&= \sum_{k=1}^{\infty} a_kc_k \sum_{n=k}^{\infty} \frac{1}{n(c_1c_2\, ...\, c_n)^{\frac{1}{n}}} && (2.11)
\end{aligned}$$

Nevertheless, now we want to find such a sequence c_i. In order to use the advantageous consideration on λ and escape the problem with divergent series, we can use a sequence which is asymptotically equivalent to $1, 2, 3, ..., n$, i.e. a sequence c_i which is close to $1, 2, 3, ..., n$ in the limit. Moreover, we need an evaluation or a simplification of the summation $\sum_{n=k}^{\infty} \frac{1}{n(c_1c_2\, ...\, c_n)^{\frac{1}{n}}}$ which appears in 2.11. At this point we use our previous knowledge about series and we observe that

$$\sum \frac{1}{n(n+1)} = \sum(\frac{1}{n} - \frac{1}{n+1}) \tag{2.12}$$

and

$$\sum_{n=k}^{\infty}(\frac{1}{n} - \frac{1}{n+1}) = \frac{1}{n} \tag{2.13}$$

Finally, by choosing the sequence $\{c_i\}$ as $c_1\, c_2\, c_3\, ...\, c_n = (n+1)^n$, the sum $\sum_{n=k}^{\infty} \frac{1}{n(c_1c_2\, ...\, c_n)^{\frac{1}{n}}}$ takes the form 2.13 and we can simplify and proceed in the proof. Moreover, we observe that $n + 1 \approx n$ for a large n, and $c_n = \frac{(n+1)^n}{n^{n-1}} = (1 + \frac{1}{n})^n n \sim en$ (c_n is asymptotically proportional to n and the number e arises). This concludes the explanation of why we have introduced the sequence c_i, then we can now move to the formal proof with more confidence. As Polya affirms: "now, we may understand how it was humanly possible to discover that definition of c_i which appeared as a *deus ex machina*. The derivation became also more understandable" [Polya, 1968, p. 152].

Is Van Fraassen's theory able to account for mathematical explanations such as Polya's? As we have seen, Polya's explanatory reasoning aims to justify the introduction of the auxiliary sequence $\{c_i\}$ into the heuristic of the proof. Thus, it is quite natural to formulate our demand for explanation under the form of a why-question, and precisely: "Why is it appropriate to introduce the $\{c_i\}$ sequence in the proof?". This why-question has, in Van Fraassen's sense, two possible readings: 1) "Why should *a sequence* be introduced?", and 2) "Why exactly *this sequence*?". Prima facie, as Sandborg observes, it might seem that two possible satisfactory answers to 1) and 2) would be: (1*) An auxiliary sequence was used in order to replace a divergent series by a convergent one; (2*) the particular sequence was chosen because it had a favourable growth behaviour and it allowed to simplify a crucial term in the derivation [Sandborg, 1998, p. 612]. However, as Sandborg points out:

> though this why-question analysis is useful in pointing out how Polya's explanation performs two distinct functions, it does not correctly account for *what makes Polya's explanation good*. [Sandborg, 1998, p. 612. My emphasis]

Since what is important for our discussion about MEPP are the conclusions which follow from Sandborg's analysis, I will not fully address his treatment of Polya test-case here. However, these conclusions will justify the previous quotation about the difficulties Van Fraassen's theory has in accounting for mathematical explanations such as Polya's.

Sandborg distinguishes two kinds of difficulties (α and β below) with Van Fraassen's theory of explanatory evaluations as extended to the case of mathematical explanations. Moreover, a third and more general problem (γ) is referred to any approach to explanation in terms why questions:

(α) Van Fraassen's theory can't account for mathematical answers to a common type of why questions, in which the members of the contrast class are mutually exclusive (*exclusive contrast-questions* have the form: "*Why X rather than Y?*", where X and Y are mathematical statements. Ex: Why does $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7}$ converge to $\frac{\pi}{4}$ rather than some other real number? Other sequences *could* also converge to the same number)[47].

(β) It cannot account for explanatory proofs, because mathematical explanations in the form of proofs are not recognized as answers by Van Fraassen's theory[48].

(γ) A why-question approach misses an important aspect of the context in which mathematical explanations are given, and precisely the conceptual resources the questioner has available to analyze the situation.

I will consider these points one by one. However, as in the case of Kitcher and Sandborg's criticism, the points raised by Sandborg are not independent and deal with difficulties which are strictly interconnected. Let's start from (α).

(α) Keep in mind Van Fraassen's three criteria of evaluation, and focus on a why-question in which the members of the contrast class are mutually exclusive. In this case, following Van Fraassen, any formally correct proof of the (mathematical) topic will be considered as telling with respect to such a question; moreover, every proof which follows from accepted mathematical propositions automatically will attain the maximal score (i.e. it will be considered as maximally probable). The reasoning is quite straightforward.

[47] It should be observed here that the set of exclusive contrast-questions is an important subset of the why-questions. However, in general, a constrast class might consist of members which are not mutually exclusive. For instance, this is the case of Polya's question "Why does this sequence serve to complete the proof?", which is not an exclusive contrast question since other sequences might be useful as well as $\{c_i\}$.

[48] Again, this criticism is not explicitly addressed to Polya's case, where the explanation does not come under the form of proof. However, proofs are commonly offered as explanations in mathematics (for instance, see [Steiner, 1978a], [Weber *et al.*, 2002], [Mancosu, 1999] and [Hafner *et al.*, 2008]). Observe that proof is neither necessary (Polya's formal proof is not explanatory) nor sufficient (Polya's explanation is not a proof) for explanation.

Consider, as in Polyas's case, a (correct) formal proof A which establishes the truth of the topic (for instance, the initial proof we gave before Polya's explanation). This proof follows from accepted mathematical propositions plus some axioms. If we consider the set formed by those axioms plus the accepted mathematical propositions as a subset of our mathematical background knowledge, and we call it $K(Q)$, the proof will be judged maximally probable with $P(A|K(Q)) = 1$. The topic, having being proved, will have probability 1, while the other members of the contrast class will have null probability[49]. Therefore there is no possibility of having other more probable answers, or answers which favour the topic better or screen it off. Here we are faced with a trivialization: from the point of view of Van Fraassen's theory of explanatory evaluation, every formally correct proof of the topic is considered explanatory when we are faced with exclusive-contrast mathematical questions. However, we know from the mathematical practice that some proofs explain better than others, and this suggests that an evaluation of mathematical explanation in mathematics should not rely on probability calculus, but on some other criteria. Although the why-question in Polya's case was not an exclusive contrast question, Polya's example is quite representative of the necessity to account for the explanation by going *behind* the mere formal aspects of the proof and choosing between a range of possibilities (the sequence c_i was not the unique possibility). A possible solution to this problem could be to fix some constraint on the relevance relation, in order to distinguish explanatory from non-explanatory answers to exclusive-contrast why questions. But we have seen in the previous pages and in Van Fraassen's treatment of asymmetric explanation that Van Fraassen does not introduce such a constraint and seems to favour an unrestricted relevance relation. Consequently, the trivialization which Sandborg underlines has something in

[49]Let me comment here that this is not the case if we accept that there exist visual proofs. The particular status of visual proofs would suggest an approach (and a criterion of evaluation) different from that used for formal proofs. Nevertheless, the problem of evaluation arises again: what are the criteria which permit us to say that a visual proof is better than another visual (or formal) proof? We are back to our original question: when does a mathematical proof explain?

common with the sort of trivialization which was pointed out by Salmon and Kitcher (trivialization ($TRV2$)) and which concerned Van Fraassen's theory of telling answers. In fact, recall that in their analysis the trivialization $TRV2$ pointed to the fact that within a fixed context we can construct every answer to every question –were the answer and the question are both true propositions–, by fixing a suitable relevance relation in order for the answer to be maximally telling to that why-question[50]. Besides the similarity with that criticism, by considering mathematical explanations the difficulty is even bigger and a defense such as that of Alan Richardson cannot save Van Fraassen's theory of telling answers from trivialization. This is because, even if we may dismiss the contrived example of JFK by saying that astrology is "bad science" (i.e. the astral influence relation is scientifically irrelevant to the occurrence of the topic where we consider as context a scientific community), we cannot dismiss a proof as "bad-maths". The moral here is, again, that in order to avoid his theory being trivialized, Van Fraassen needs some sort of restriction on his relevance relation in order to have a *genuine* relevance relation (not gerrymandered). However, also this possibility (of fixing some constraint on the relevance relation) seems to not do the trick in the case of mathematical explanations by proofs. As Sandborg observes:

> Also, the reason we do not consider some proofs to be explanatory don't seem to have to do with *relevance* at all, or at least not relevance that the questioner can specify in advance. An early unintellegible proof of a result (such as one using 'brute force' calculational techniques) may be obviously relevant to that result, even if considered non-explanatory, while an explanatory proof employing more abstract mathematical resources may not be so obviously relevant; it may not be immediately clear that the resources used are appropriate to the problem. Indeed, discovering these resources may be a large

[50]In their JFK's example, by constructing an appropriate relevance relation, the answer together with a selected subset of background knowledge will imply the topic of the question and the negations of all members of the constrast class. Therefore the answer will be considered as maximally telling.

> step towards solving the problem in the first place. No relevance relation specifiable at the outset would be able to distinguish between explanatory and non-explanatory proofs, because what is relevant to the explanandum is not prior to the explanation itself. [Sandborg, 1998, p. 615]

Let's now switch to (β). In point (α) we have considered a proof as a (valid) answer to a why-question and we have put in evidence as Van Fraassen's theory consider all proofs as explanatory under its criteria of evaluation. The second problem (β) highlighted by Sandborg is connected to (α) but addresses a more delicate aspect of Van Fraassen's treatment of explanation. More precisely, it concerns the *descriptive* form an explanation (an answer) must assume, in Van Fraassen's theory, in order to support the topic.

As Van Fraassen himself points out:

> If you ask a scientist to explain something to you, the information he gives you is not different in kind (and does not sound or look different) from the information he gives you when you ask for a description. [...] To call an explanation scientific, it to say nothing about its form or the sort of information adduced, but only to say that the explanation draws on science to get this information (at least to some extent) and, more importantly, that the criteria of how good an explanation it is, are being applied using a scientific theory [Van Fraassen, 1980, p. 155-156]

To evaluate an answer, in Van Fraassen's pragmatic account, is to consider the contribution that it makes to favouring the topic when it is used to supplement some subset $K(Q)$ of our background knowledge. If we pass to mathematical explanations, it is then natural to ask whether Van Fraassen's theory can account for answers coming under the form of proofs.

In such cases it is quite natural to consider our background knowledge K as either the set of axioms of a theory, or their deductive closure. The topic "theorem" will then be either implied or contained in K. Nevertheless, if we restrict, as in the empirical case, K to some part $K(Q)$ consisting in a smaller set of axioms or its deductive closure, then an answer to the mathematical why-question will be some proposition which, added to $K(Q)$, will imply the topic "theorem". So considered, this answer represents that part of logical information missing for the theorem to be true: a missing axiom, or some statement that imply the theorem[51]. Therefore, Sandborg observes, "Van Fraassen's theory suggests that mathematical explanations would have to be analyses of the preconditions required for a theorem to be true, such as showing that a theorem depended on the axiom of choice" [Sandborg, 1998, p. 616]. Consequently, through the lens of Van Fraassen's theory, mathematical explanation reduces to a form of analysis of axioms required for theorems, a method of deducing consequences from previously given propositions. There is no room for new information, because the explanatory activity admitted assumes the form of a display of consequences of what we have already accepted as given[52].

However, to deduce consequences from previously given propositions is far from being considered an explanatory activity in mathematics[53]. The explanatory activity in mathematics often assumes a very different form. For instance, in his explanation Polya does not pick out any proposition that, in conjunction with part of his background knowledge, leads to the result. This suggests that the explanatory activity in mathematics does not assume that form Van Fraassen's pragmatic theory seems to require, and consequently this theory cannot account for explanatory mathematical proofs. As Sandborg

[51]In passing, let me note that a theory of why-questions focused on explanations as providing additional logical information is presented by Hintikka and Halonen [Hintikka *et al.*, 1995].

[52]Furthermore, Sandborg observes that a "proof does not fill in any missing information, but instead draws out consequences from previously given propositions" [Sandborg, 1998, p. 616]. And this seems to constrast with Van Fraassen's picture of explanation, which considers that an answer must provide some extra-information which is not present in the subset $K(Q)$ of our background knowledge.

[53]Observe that this is exactly the problem which emerges if we consider the D-N account in the context of mathematical explanations within mathematics: by admitting mathematical explanations in the D-N account (with the adequate changes), every mathematical proof turns to be an explanation because it fits perfectly within the deductive schema.

observes, "any theory [of explanation] that takes explanations to increase our propositional knowledge will tend not to regard any proofs as explanatory, since they do not do this, but at best display the consequences of what we have already accepted as given" [Sandborg, 1998, p. 616].

A third and more general remark (γ) raised by Sandborg follows directly from points (α) and (β) and concerns the impossibility for a why-question approach (a general why-question approach, not only Van Fraassen's) of accounting for the conceptual resources available to the questioner in the analysis of the situation. Let me illustrate the situation with the aid of Polya's test case.

Consider again the why-question answered by Polya: "Why is it appropriate to introduce the c_i sequence in the proof?". And remember the two possible readings of it: 1) "Why should *a sequence* be introduced?"; 2) "Why exactly *this sequence*?". Polya's explanation clearly addressed both readings, then it should be considered as formed by the two answers. His answers are, respectively: (1*) An auxiliary sequence was used in order to replace a divergent series by a convergent one; (2*) the particular sequence was chosen because it had a favourable growth behaviour and it allowed to simplify a crucial term in the derivation. However, in completing his explanation by answering 2), Polya showed *how* the growth rate of c_i was favourable for completing the proof, but it did not show that this sequence was the *only* appropriate to use (other sequences with the same growth rate were good candidates as well). The sequence was chosen simply because it allowed to simplify the term $\sum_{n=k}^{\infty} \frac{1}{n(c_1 c_2 \dots c_n)^{\frac{1}{n}}}$. Hence, though it provided a solution to our initial puzzle, Polya showed that this sequence worked for our purpose, but he did not show *why* it worked. Here is the crucial point: Polya's explanation does not distinguish between this sequence and other sequences, and this is why it cannot be considered a telling answer to the why-question "Why is it appropriate to introduce the c_i sequence in the proof?". Therefore, if we accept Polya's explanation as a good explanation, we have to conclude that it does not provide a good answer to its motivating why-question. This seems to suggest that the why-question analysis lacks some facets of Polya's explanation. In particular, as we are going to see in the following lines, the why-question approach does not account for the conceptual resources introduced by Polya in his explanation.

Polya's explanation cannot be considered a telling answer to the why question "Why is it appropriate to introduce the c_i sequence in the proof?". Consequently, the why-question approach does not account for this particular explanation. Why? Sandborg's diagnosis is that the difficulty for the why-question approach comes from what implicitly a why-question presupposes, namely that the way it regards its topic being fixed. In line with Belnap and Steel's dictum "to understand a question is to know what would count as an answer to it" [Belnap *et al.*, 1976, p. 35], this roughly means that we know *a priori* what an acceptable answer to our why-question would be. There is then no room for conceptual resources we introduce into the analysis of the situation. On the other hand, our initial state of puzzlement may be due to not even knowing how best to regard the topic. In fact, an explanation can gain most of its virtue by showing us an effective way to understand the subject-matter, rather than through any particular why-questions it happens to answer[54]. For example, in Polya's case, the mathematician shows the importance of *the* growth rate, and not the importance of *that* growth rate rather than another. Thus, if we want to account for this situation in terms of why-questions, the why-question "should allow an answer in terms of growth rates to be judged explanatory by a questioner who couldn't even talk about growth rates *until* the explanation had been given" [Sandborg, 1998, p. 621. My emphasis]. In so far as asking a why-question fixes a way of looking at the explanandum and demands an explanation in those terms, the why-question approach will not be able to account for a variety of mathematical explanations. More importantly, this point is relevant for our discussion

[54]This observation is connected to Sandborg's quotation in point (α) about the impossibility (for the questioner) of specifying *a priori* a relevance relation in cases where: we discover new resources during the solving procedure and those resources are not manifest at the outset.

because the problem is not restricted to mathematical explanation within mathematics but it is shared also by explanations of physical phenomena analyzed through the lens of why-questions techniques. For instance, consider Newton's explanation for the motion of planets in terms of action at a distance. His explanation clearly did not answer the question the Cartesians had in mind, but it introduced a new way of looking at the topic "the planet P moves around the sun following path x". This has been possible through the introduction of concepts not available before[55]. Therefore "it would have been impossible to specify a Newtonian answer as an appropriate answer to a question posed before the *Principia*; the pertinent concepts couldn't yet be given to indicate that *kind* of answer was appropriate" [Sandborg, 1998, p. 622].

Sandborg proposes to correct the problems highlighted in (α), (β) and (γ) by offering an alternative to the why-question approach, namely a picture of explanation evaluation which takes into account conceptual resources not previously available: "an explanation may be significant because it deploys relevant conceptual resources not previously available" [Sandborg, 1998, p. 621]. Even if the idea of conceptual resources is left undeveloped in his study, such approach to the problem of explanation would preserve contextual factors (what was not previously available depends upon the context), including the aspects of the context which the why-question approach was not able to capture (difficulty γ). Moreover, it would avoid the problems α (problem of explanatory evaluation) and β (problem of recognition of explanations under the form of proof) we have seen Van Fraassen's model has when faced with

[55] It is very interesting to note that Van Fraassen points to the same example (Newton's explanation of phenomena in terms of gravitation) as to show that there are cases, in a theory's domain, where the request for explanation is rejected because at *that stage*, i.e. during that particular period, that particular request is considered as intrinsically illegittimate. This was, in fact, the case of the paresis example (see section 2.3.1, and precisely Van Fraassen's treatment of the problem of the rejection of explanation requests). According to Van Fraassen, Newton did not provide an explanation of gravitational phenomena, but only a description [Van Fraassen, 1980, p. 112]. Differently from Van Fraassen, Sandborg considers that Newton did provide an explanation, even if in a totally new form with respect to the explanations which came before [Sandborg, 1998, p. 622].

cases of mathematical explanations[56].

Observe that up to now I have considered Van Fraassen's among the WTA approaches to explanation. However, Van Fraassen considers that there no exists explanation *simpliciter* and he proposes a model in which the relevance relation is open (it depends on the context we are in). Therefore his account seems not to fit into the WTA conception of explanation presented in this part of the dissertation, but it looks to be much more related with a pluralist perspective on MEPP. Why then I have considered Van Fraassen's account among the WTA models? It is time to justify this choice, and Sandborg's criticism (especially remark γ) provides the right 'context' to do that.

According to Van Fraassen, explanations *always* come under the form of answers to why-questions. In regarding explanations as coming under this form, Van Fraassen does impose a general schema (or form) on explanation, and proposes a single model. This idea does not fit within a pluralist view on explanation, which considers that there are different kinds of explanation and these explanations cannot be captured by a single account.

Of course, it might be noted that Van Fraassen leaves open his relevance relation exactly because he wants his model to mirror this variety of explanations, namely the different kinds of relevant answers which can be given depending from the contexts. However, Sandborg's criticism shows that explanations do not always come under the form of an answer to a why questions. In particular, the why-question schema cannot be applied when we do not know *a priori* how best to regard the topic, and our explanation is given in terms of procedures or considerations which were not known or

[56] According to Sandborg's sketch, when a mathematical explanation "does not look at the explanandum in a new way" (i.e. it does not deploy new conceptual resources), it could not be considered explanatory. This would make possible to bypass the trivialization seen in (α), concerning the insufficiency of a probabilistic criterion of explanatory evaluation in mathematical explanations. As regards (β), this approach would have the advantage to explain why we consider a proof of a result as an explanation of it: the explanatory character of the proof would be not evaluated according to its capacity of providing new propositional information, but to the fact that it invokes new conceptual resources not previously available. Sandborg's picture of explanation is not developed further in his paper.

taken into account at the outset (as Sandborg shows in Polya's case)[57]. As put forward by Matti Sintonen:

> It also becomes obvious that not all explanations are answers to why-questions. Depending on the type of inquiry at hand they could be how-questions, how possible-questions, what-questions, or the like [Sintonen, 1999, p. 134].

Therefore, by considering that *every* explanation in science comes under the form of a why-question, Van Fraassen commits himself to a WTA view on explanation. This is why I consider his model as a WTA approach.

Before coming back to Baker's MEPP and PET, it seems to me important to add a comment on Richardson's and Sandborg's remarks about the possibility (or not) of dismissing an answer (for instance, the answer about JFK's death in terms of positions of the planets on that particular day) just by looking at the particular context (for instance, modern science or the astrological-club). To this comment I will add some considerations about the applicability of the pragmatic account to mathematical explanations. This is relevant to MEPP because has to do with the process of evaluation of a mathematical answer to a why-question.

Recall Richardson's defense of Van Fraassen: the relation of astral influence, i.e. a relation between propositions describing positions of planets and propositions about personal fate, does not exist in the present context of modern science, and then no astrological-answer is relevant and can be used as telling answer to the why-question "Why did JFK die on 11/22/1963?". Let me comment on this.

In the context of modern science there is no relation of astral influence. How can we say that? If we want to throw away astral influence as bad science (and thus dismiss Kitcher and Salmon's criticism), it seems that Van

[57]I will provide a similar example of such explanations, i.e. explanations which are not given in terms of why-questions but which are regarded as such in the scientific practice, during my analysis of a particular test-case, in chapter 7. This example will concern MEPP, and not mathematical explanation within mathematics (as in Sandborg's case).

Fraassen and Richardson need a definition for what is “scientifically acceptable”. A possible solution could be to consider the astral theory as ‘scientific’ if it is empirically adequate. But, as we have seen, Van Fraassen’s idea is that empirical adequacy does not affect explanation:

> So I conclude that (a) the assertion that theory T explains, or provides an explanation for, fact E does not presuppose or imply that T is true or even empirically adequate [Van Fraassen, 1980, p. 100]

Therefore, how can Richardson say that such a astral theory does not belong to the present context of modern science *without* giving a criterion to judge what is scientific and what is not? It seems that only something very similar to an objective relevance relation (for instance, a relation expressed as a law of that theory) should do the trick. This means that the problem to have some objective (or not) constraint elsewhere comes out again.

Let’s make a step further by making a parallel between scientific explanation and mathematical explanation. Even if we assume that astral influence is not science (according to a criterion of ‘scientificity’ which Van Fraassen does not offer), in the case of mathematical proof we cannot exclude a (formally correct) proof as “bad maths”. Therefore no criteria for “bad proofs” (proofs to be considered as not relevant to that context) are established for formally correct proof. A possible criterion to evaluate the ‘goodness’ of a proof could be given, precisely, by its explanatory power (or its aesthetic virtues, or other virtues). But the explanatoriness is exactly what we are looking for. Now, Richardson claims that the assertion ‘For any two true propositions, A and B, A explains B’ is true only if, given any two propositions (A and B), there is some context in which there is a why-question with B as topic which A answers. Nevertheless, from the analysis of David Sandborg clearly emerges that this context is hard to find in the case of mathematical explanations in mathematics, and in particular in the case of explanations under the form of proofs. What is more, let me note that while astrological circles and scientific communities have different rules and utilize different theories, the mathematical community uses a common instrument called mathematics, which is made by extremely interconnected domains that are very difficult to 'cut' into separate contexts. Thus the mere propositional-context alone (the mathematical knowledge which is employed in the resolution of the proof), which is considered by theories as Van Fraassen's, is insufficient as to discriminate an explanatory from a non-explanatory proof. Sandborg's suggestion of focusing on the role of "conceptual resources not previously available" points exactly in this direction: we need to include an extra ingredient into our model of explanation, at least if we want our model to mirror cases of mathematical explanations in mathematics which occur under the form of proofs[58].

Therefore it seems that, for the case of mathematical explanations in mathematics, the "austerely beautiful landscape of our empiricist philosophical homeland" [Richardson, 1995, p. 126] is far from being reached. And we have to choose between the Scylla of a new theoretical approach (as proposed by Sandborg) and the Charybdis of an objective relation as a virtue of theory that provides us explanatoriness [Kitcher *et al.*, 1987, p. 330]. However, the third way proposed by Van Fraassen is still open, and it might save the pragmatic account in the case of mathematical explanations. What is the price to pay? I will shortly discuss this third way below, in the context of MEPP and our example of cicadas.

2.4 Is PET a good candidate for MEPP?

Let us now return to Baker's test-case and the possibility for the pragmatic account of covering MEPP. As we have seen, Van Fraassen's third presupposition (A bears relation R to $\langle P_k, X\rangle$) involves the relation R, which is left unrestricted by the author. The criticisms presented in the previous section suggest that this is *the* crucial aspect of Van Fraassen's model. It is therefore important to address this aspect.

In subsection 2.3.2 I have claimed that the third presupposition 'A bears

[58]Recall that, while "proofs are often vehicles for mathematical explanation" [Sandborg, 1998, p. 616], not all explanations come under the form of proofs, as Polya's example suggests.

relation R to $\langle P_k, X \rangle$' is problematic in our cicada-case of MEPP. However, I have not provided a justification for that claim. Remember now that A is the proposition 'prime periods minimize intersections compared to non-prime periods' (the mathematical statement). We want to know how R can be found (or how R *is* found) in this context.

Of course, in the case of cicadas we cannot assume that the reason A bears a *causal* relation R to $\langle P_k, X \rangle$. Therefore, what kind of relation R does a number theoretic theorem (A) bear to $\langle P_k, X \rangle$? It seems that, if we want to extend the pragmatic account and save it for this particular case, we have two possibilities concerning R:

1. to agree with Salmon and Kitcher, and try to find an *objective* relevance relation.

2. to leave the relevance relation indefinite (without any constraint, objective or not), as Van Fraassen does, and justify this choice somehow depending on the context we are considering.

Leaving apart the question whether Van Fraassen's theory is committed to realist or anti-realistic positions, if we choose the first solution we come back to our starting point. To find an objective relevance relation R between a mathematical fact and a physical phenomenon might be regarded exactly as the problem of what makes a mathematical description of that phenomenon explanatory. This is the challenge of accounts like Steiner's, where Van Fraassen's machinery is not needed. Steiner's solution in terms of "separability" of the mathematical part from the physical one and his definition of characterizing property could be read as an attempt to define an objective relevance relation, which does not require the why-questions approach[59]. The difficulty in characterizing the relevance relation in the case

[59]Let me clarify this point. Recall what we saw in the previous chapter. For Steiner, a mathematical proof is explanatory if: it depends on a characterizing property mentioned in the theorem (C_1); it is possible to deform that proof "substituting the characterizing property of a related entity" and getting "a related theorem" (C_2). We have a MEPP when, after the separation of the mathematical part from the physical part of the explanation

of cicadas is well expressed by Baker when he affirms (in private communication): "If I were to try to say more about how the pragmatic model might work here, I think I would analyze the relevance relation as some sort of counterfactual. One problem, of course, is that many mathematical counterfactuals are also counterpossibles, e.g. 'If 17 were not prime then ...'. So one would need an account of how to make sense of these, so that they do not all come out as vacuously true. Or perhaps it is better to focus on the 'mixed' mathematical-physical claims and evaluate the counterfactuals with respect to these".

Alternatively, if we choose the second strategy in the case of mathematical answers, we can consider those answers as relevant depending on the interests of the questioner. The interests of the questioner would fix a criterion of relevance, thus providing a relevance relation. However, in order the questioner can fix a criterion of relevance, he needs alternatives answers. In the context of mathematical explanation within mathematics and MEPP, to have alternative answers amounts to have more than one mathematical theory of the same mathematical result or physical phenomenon. This is the possible alternative and third way, between Sandborg's Scylla and Kitcher and Salmon's Charybdis. The idea has been expressed by Van Fraassen to me in a private communication[60]: " [...] it might still be possible to point first of

(we isolate the bridge-principles), we obtain an explanatory proof in the sense of C_1 and C_2. I have called C_{MEPP} such a criterion. Therefore, a possible way of fix Van Fraassen's relevance relation R, according to Steiner's criterion of explanatoriness, might be to model C_{MEPP} on R. The relevance relation R_{st} would be the relation that the mathematical explanans A bears to $\langle P_k, X\rangle$. For instance, the relation that the mathematical statement 'prime periods minimize intersections compared to non-prime periods' bears to the couple $\langle P_k, X\rangle$ in the case of cicadas. Of course, the relation R_{st} would make the answer 'Because A' favour the topic with respect to the other alternatives, thus picking out only the topic within the constrast class. We would obtain a super-telling answer to our why-question! Is this circular in any sense (because we are already assuming an account of MEPP)? Would Van Fraassen agree on this reading? Moreover, let me observe that this relation would be objective because based on the notion of characterizing property. I will come back to the possibility to 'read' Steiner's account of MEPP in terms of a relevance relation in chapter 4.

[60]Note that I call this approach the 'third way' just to emphasize that in this case the pragmatic approach is specifically used in the context of mathematical answers to why-questions. The idea expressed by Van Fraassen in the private communication is the

all to the interests of the questioner which might determine to some extent the relevance of the answer. For example, some problems are in principle treated equally well in set theory and in category theory, but the questioner might not be wanting the first rather than the second. And while "more likely" than its alternatives seems inapplicable in the mathematical case if "likely" is taken in an objective sense, an answer to a mathematical question might eliminate more or fewer of the relevant alternatives (while not offering a complete proof)".

Nevertheless, to adopt an account of explanation with this pragmatic function seems largely insufficient for our purposes. By accepting this strategy we would be forced to reduce our analysis only to cases where mathematical or physical results are explained *via* different mathematical theories, a possibility which is not so common. If we concentrate on the case of physical phenomena treated mathematically, it is not common to have different mathematical theories which describe the same natural phenomenon. For instance, this is not the case of the cicada example, where there is (as least to our knowledge) only one mathematical answer. Beside this, let me note that there is also a problem with what we should consider as a mathematical theory (can call this problem the "boundary" problem). Even if we agree that set theory and category theory are considered distinct theories by a questioner, there are cases of MEPP where two or more distinct mathematical *procedures* can lead to the desired result, but those procedures do not belong to different theories; even worse, those procedures might be closely interconnected because they use shared mathematical tools. As an example, take the case of Euler's theorem we saw during the discussion of Steiner's account of MEPP. We can have prove the theorem, and then the existence of the instantaneous axis of rotation, via a geometric procedure (this is what Euler did in his [Euler, 1750], or what is made in [Whittaker, 1904, p. 2] and [Targ, 1987, p. 221-222]) or via an abstract algebraic procedure (considering matrices, vector space, etc..). Although different, these mathematical

same idea which stands behind the pragmatic approach. However, Van Fraassen does not consider the case of *mathematical* answers to why-questions in his [Van Fraassen, 1980].

procedures do not belong to different mathematical *theories*. Modern linear algebra embeds geometrical concepts (for instance, the distance between points) as a consequence of an historical process, and then the procedures cannot be considered as independent one from the other.

Therefore, at least when we cannot find different theories which describe the same physical phenomenon, it seems that this third pragmatic way to explanation must be rejected. When we have at disposition only *one* mathematical theory (or description) of the physical phenomenon, or one mathematical theory of the mathematical result, the questioner has only one (mathematical) answer at disposition, and the theory of why-question is of not interest for an analysis of the explanatoriness of the argument. This is the case, for instance, of the cicada example. Thus there seems to be little chance to save PET for mathematical explanations in mathematics and in science by considering the interests of the questioner as a criterion of relevance.

Finally, also if we assume PET as way to eliminate relevant alternatives, a further problem emerges: the problem of evaluation. This is precisely the trivialization (point α) raised by David Sandborg in his discussion of Van Fraassen's theory as applied to the case of mathematical explanations in mathematics. It takes us back to the problem of finding explanatory factors in order to evaluate the mathematical component on a part $K(Q)$ of our knowledge K. If we consider mathematical explanations within mathematics, and in particular the case of proofs, no probabilistic tools are suitable to evaluate our explanations because all proofs would be explanatory with maximal score. If we consider MEPP, we do not have a common framework which links mathematics to our background knowledge (or to a part $K(Q)$ of it), thus how do we evaluate A in light of it? As an example, consider that we are able to find a different proof of the number theorem which Baker refers to in his paper (or, more luckily, a different theorem which offers a useful connection between primeness and biological laws). Does the PET tell us (as Steiner's account does) if the new proof (or theorem or mathematical procedure) plays a more or less explanatory role than the previous one? How other answers can favour the topic better or screen it off? How the interests of the questioner might be taken into account in the process of evaluation of the telling character of such an answer? Again, it seems that the problem of evaluation strictly depends on the lack in an evaluative component coming from R, i.e. a relevance relation as a sufficient criterion in order to evaluate the two differents proofs (or theorems) in light of our knowledge or part of this knowledge.

2.5 Conclusions

In order to consider MEPP in Van Fraassen's model we can choose between two general strategies. If we agree with Paolo Mancosu that a theory of mathematical explanations of scientific phenomena is not completely independent of a theory mathematical explanation of mathematical facts [Mancosu, 2008b], we can consider the pragmatic account from this perspective. Nevertheless, as emerged from my analysis, Van Fraassen's theory is not suitable to deal with mathematical explanation within mathematics, thus the pragmatic model should be rejected also as a model of MEPP. The alternative strategy consists in taking into account an extension of the pragmatic theory (I called PET this extension), assume some sort of methodological continuity between the mathematical world and the physical one and focus directly on MEPP. However, as we have seen in the previous section, also this extension is problematic because it could be made only via two approaches: define an objective relevance relation, but this is exactly the problem of defining what counts as an objective criterion for explanatoriness; or assume Van Fraassen's theory as a theory through which evaluate the relevance that different answers have for the questioner. The latter case, as we have seen, is insufficient for our purposes. The cicada case shows well how these alternative answers are often impossible to find in the context of MEPP. Moreover, in both cases we have the problem of the evaluation of the mathematical argument, which makes Van Fraassen's model of explanation problematic to extend to cases of MEPP. It seems that the PET model is not able to tell us *why* a mathematical procedure is regarded as providing explanatoriness (with respect to another mathematical procedure).

To conclude, it seems that, to save Van Fraassen's account in the case of mathematical explanations (within mathematics or for physical phenomena), we need some extra criterion (on the relevance relation and on the evaluation step). This extra criterion is not offered by Van Fraassen's theory, and it is not clear how to offer it in the context of MEPP. Furthermore, my previous considerations block (or, at least, make less plausible) Baker's positive idea about the possibility of using the pragmatic account in the cicada-case. Finally, at least in the context of MEPP, I have to agree with Kitcher and Salmon in concluding that Van Fraassen's theory of explanation seems to be a very useful theory *of the* pragmatics of explanation, but not *a* pragmatic theory of explanation [Kitcher *et al.*, 1987, p. 315].

Chapter 3

Unification as a way to explanation: a uniform, global approach

> The aim of scientific explanation throughout the ages has been unification, i.e., the comprehending of a maximum of facts and regularities in terms of a minimum of theoretical concepts and assumptions. The remarkables success achieved, especially in the theories of physics, chemistry, and to some extent recent biology, has encouraged pursuit of a unitary system of explanatory premises. Whether this aim is attainable depends, of course, both on the nature of the world and on the ingenuity of the scientists. I think this is what Einstein had in mind in his famous sayings: "God is subtle but He is not malicious"; "The only thing that is incomprehensible, is that the world is comprehensible"
>
> Herbert Feigl, *The "Orthodox" View of Theories*, p. 12.

The concept of unity in physics has a long history. As Klein and Lachieze-Rey have well illustrated in their book *The Quest for Unity: The Adventure of Physics* [Klein *et al.*, 1999], this search for unification begins with Greek conceptions of unity and arrives until our day. The effort to reconcile opposites (fire and water, being and not being, finite and infinite, abstract and concrete, and so on) and subsume regularities under a same framework, is as ancient as philosophy. Passing through differents ways of thinking, contexts and authors such as Thales, Heraclitus, Nicholas of Cusa, Kepler, Descartes, the concept of unity has showed its force through the history of human thought. With respect to physics, it is well-known how important has been, and continues to be, the role of mathematics in the process of inclusion of separate theories and phenomena into one single framework. This is the case, for example, of Maxwell's famous unification of electromagnetism and optics through the Lagrangian formalism[1]. Now, although it is largely accepted that "appearances of homogeneity in physics are due to a generalized use of mathematics" [Klein *et al.*, 1999, p. 103], we are left with the following problem: how does mathematics play this unifying role? And, for what interests us, does this unification have something to do with explanation? If mathematics is to serve as a unifying tool, how should we characterize a theory of explanation in terms of unification?[2].

In this chapter I will present the unification model for explanation, firstly proposed by Michael Friedman in his [Friedman, 1974] and successively modified and extended by Philip Kitcher in his [Kitcher, 1981] and [Kitcher, 1989]. I will focus on Kitcher's version of unification, and this choice is motivated by the fact that Kitcher's model can be regarded as a refined version of Friedman's (I will substantiate this claim during my discussion). However, the reader might be surprised at my choice. The original construction of Kitcher's model was addressed to general scientific explanation, and in particular to explanation under the form of laws. Why then do I discuss it in the context of MEPP? I can adduce three general motivations for this choice. First of all, observe that the fact that such a theoretical account could *potentially* cover mathematical explanations of physical phenomena as well as mathematical explanations within mathematics is now a shared opinion by a number of em-

[1] In his [Maxwell, 1865], Maxwell deduced analytical mechanics, electromagnetism, and wave mechanics from a variational principle. See [Morrison, 1992] and [Morrison, 2000], especially chapter 3 of the latter, for a detailed history of the development of Maxwell's electrodynamics.

[2] As we are going to see, the unification theory of explanation proposed by Philip Kitcher will be more general and it will consider "patterns", and not mathematics, as a unifying tool in science.

inent philosophers (for instance, see [Hafner *et al.*, 2005], [Mancosu, 2008b] and [Tappenden, 2005]). As pointed out by Jamie Tappenden:

> However, mindful of the fact that some explanations in physics and mathematics *do* seem to be governed by the same principles, I'll count it as an advantage of an account that it supports a uniform treatment of some mathematical and some physical explanations. A promising candidate to support a uniform treatment of some pure mathematical cases and some non-mathematical ones is the treatment of explanation as unification as proposed in the seventies by Michael Friedman and Philip Kitcher. [Tappenden, 2005, p. 158-159]

Moreover, as we are going to see, also Kitcher considers that his unification model is well suited to account for cases of mathematical explanations. Therefore, since this model has been regarded as a promising candidate as to cover MEPP, it is important to include it in this study. Second, let me observe that, despite the great interest in the linkage scientific explanation-MEPP, an extensive discussion of this model in the context of MEPP has not been offered and work is just beginning [Mancosu, 2008c]. This is why in the final part of the dissertation I will come back to Kitcher's model and I will discuss it in the context of MEPP. In chapter 7 I will assess Kitcher's model on a case of MEPP coming from the scientific practice, while in the final chapter I will discuss Kitcher's example of Newtonian unification in the context of my perspective on MEPP and of a generalization of my approach. Moreover, I will provide a sketch of how my perspective on MEPP can be potentially extended to scientific explanation, and Kitcher's model will be very important to state this point. Consequently, the space devoted to Kitcher's unification and the detailed presentation of such an account are worth for the discussion I will deal with in the last part. Finally, there is a third reason (less substantial and more 'pragmatic') which has led me to include this model in my analysis. Kitcher's model represents a good example of WTA model of explanation, and therefore I will use it to constrast WTA with pluralism. Moreover, its exhaustive structure has permitted me to pick out from that model some specific aspects, and I have used these aspects to fix some general characterizations concerning the WTA models (I will offer such characterizations in chapter 4).

The chapter is structured as follows. In the first section, I will present Friedman's original idea concerning explanation as unification. Second, I will skip to Kitcher's more elaborated formulation of explanatory unification. This will require some technicalities, which will be introduced gradually in my presentation. Next I will report some criticisms against the unification approach to explanation and, more particularly, against the fact that there exist some relationships between explanation and unification. Finally, I will dedicate the conclusive section to the 'methodological holism' on which Kitcher's model is based, and I will stress the relevance of this chapter for the topic and the general strategy of this dissertation.

3.1 Friedman's unification

Friedman's discussion in his [Friedman, 1974] starts with two important observations: for the most part, what is explained in science is a general regularity or pattern of behaviour (explanations of particular events are relatively rare)[3]; second, the explanation of a phenomenon often involves another phenomenon by a relation known as *reduction*. For example, we explain the fact that water turns to steam when heated by relating this phenomenon with the behaviour of the molecules of water[4]. The initial phenomenon to be explained (the behaviour of water) is then reducted to an explaining phenomenon (the behaviour of molecules). This situation is very common in scientific explanation. Thus, concerning the main desiderata a theory of scientific explanation should have, Friedman points out that

> [...] the central problem for the theory of scientific explanation comes

[3]We remind here that the D-N model had been drawn to take into account particular events. I will return to this point during the presentation of Friedman's model.

[4]When water is heated the energy of the molecules increases. If their energy is sufficient to overcome intermolecolar forces, the molecules fly away and escape in the atmosphere.

> down to this: what is the relation between phenomena in virtue of which one phenomenon can constitute an explanation of another, and what is it about this relation that gives us understanding of the explained phenomenon? [Friedman, 1974, p. 6]

The previous quotation underlines two crucial points of Friedman's discussion: the reductionistic character a theory of explanation should have (and we will see how Friedman's model tries to capture this feature) and the focus on the linkage explanation-understanding such a theory should offer. Referring to the latter, Friedman's idea is that the notion of explanation and that of understanding should not be studied separately. We cannot start by defining what "scientific understanding" is and then requiring that a theory of explanation captures this sense of understanding. As Friedman observes: "We can find out what scientific understanding consists in only by finding out what scientific explanation is and vice versa" [Friedman, 1974, p. 6].

In order to define a new account doing better than the preexistent accounts of scientific explanation, in his paper Friedman follows a three-step method: (A) analyse the problems which the notion of understanding has in traditional accounts of scientific explanation; (B) extract from the analysis the properties a good notion of understanding should have; (C) sketch a new account of scientific explanation in which these good properties are included, and which offers the linkage explanation-understanding we are searching for. Here I will consider steps (A) and (B), and then I will concentrate on his account of explanation.

(A) According to Friedman, if we focus on the notion of understanding, classical approaches to explanation could be divided in two categories: those like Hempel's, which have a precise definition of the nature of the explanation relation but nothing to say about the linkage between this relation and the notion of understanding; and those like Toulmin's, Scriven's or Hanson's which offer a clear definition of what understanding is but they do not have a story of how this understanding is produced[5].

[5]These approaches to explanation are given, respectively, in [Hempel *et al.*, 1948], [Toulmin, 1963], [Scriven, 1962] and [Hanson, 1963].

If we consider the D-N model, Hempel argued against a notion of scientific understanding because of its non-logical, pragmatic aspect in explanation. Here, with the term "pragmatic", Hempel means psycological, i.e. having to do with individual beliefs or attitudes of persons. However, as Friedman observes, the term pragmatic can also mean subjective as opposed to objective [Friedman, 1974, p. 8]. In the latter case, the term pragmatic must be regarded as a relative notion (a notion varying from individual to individual), while in the first sense it might be not. Hence, while agreeing with Hempel that a good theory of explanation should offer an objective notion of explanation, Friedman maintains that the notion of understanding can be pragmatic in the first sense (and then psychological) but also objective (and then constant for a group of persons, for instance a community sharing a system of beliefs):

> Similarly, although the notion of understanding, like knowledge and belief but unlike truth, just is a psychological notion, I don't see why there can't be an objective or rational sense of 'scientific understanding', a sense on which what is scientifically comprehensible is constant for a relatively large class of people [Friedman, 1974, p. 8]

This is why, for Friedman, we have to take into account such a notion of understanding (i.e. a notion depending on psychological factors but having an objective value for a group of individuals) in building our theories of scientific explanation.

Despite rejecting understanding as a relevant feature of his account, Hempel suggests that in the D-N model the understanding is achieved by rational expectation. In other words, the passage from the ontic to the epistemic is made by rational expectation. However, as Friedman observes, "to have grounds for rationally expecting a phenomenon is not the same thing as to understand it" [Friedman, 1974, p. 8]. A typical situation of this is provided by the so called 'indicator laws', as the law which connects the indication on a barometer and the arrival of a storm. Those laws offer predictions, i.e. rational expectations, on the basis of initial conditions and laws, but do not offer the understanding of *why* the phenomena does occur (we look at the barometer and we know there will be a storm, but we do not know *why* this occurs). Those counterexamples concern explanation of particular events, which are concerned by the D-N model. Friedman observes that the situation is even worse if the phenomenon to explain is a general regularity or a pattern of behaviour, since they do not occur at a definite time and then there is no question of expecting them. Thus we have only rational grounds for believing the phenomenon does occur, which is part of the understanding of the phenomenon but it is not a complete and sufficient story for such an understanding.

The "familiarity" approach advanced by P. D. Bridgman in his book *The Logic of Modern Physics* [Bridgman, 1927] maintains that we have understanding of the world through scientific explanation when we reduce unfamiliar phenomena to familiar ones[6]. Nevertheless, as stressed by Friedman, this account is not acceptable because most of the explanations offered by modern physics are given in terms of phenomena stranger and less familiar than the original phenomena they explain. The same kind of criticism has been expressed by Hempel himself:

> The free fall of a physical body may well be said to be a more familiar phenomenon than the law of gravitation, by means of which it can be explained; and surely the basic ideas of the theory of relativity will appear to many to be far less familiar than the phenomena for which the theory accounts. [Hempel *et al.*, 1948, p. 145]

A view similar to that expressed by the familiarity account is that of Michael Scriven in his "Explanation, Prediction and Laws" [Scriven, 1962]. Scriven suggests that explanation has a logical function. This logical function consists in relating some phenomena (not understood) with the set of phenomena understood by a person at a particular time, a "realm of understanding" [Scriven, 1962, p. 202]. In this account, the phenomenon being

[6] The term "familiarity view" has been conied by Friedman [Friedman, 1974].

explained must be related to a phenomenon that is already understood[7]. Nevertheless, according to Friedman, Scriven's account is not immune to the problem the familiarity view on explanation has: in science we very often explain *via* a phenomenon which is less understood, in the relevant sense, than the phenomenon to explain.

Let me add a comment on Friedman's observation. I think that Friedman's criticism should be reinforced by the following observation: Scriven does not offer any precise way of discriminating between different degrees of understanding of a phenomenon, then his account suffers from the (even more urgent!) problem of giving such a comparison. For instance, in what sense is gravitation better understood than the fall of bodies? The problem is to define a condition for which a phenomenon belongs to the realm of understanding. What is involved is, again, the notion of understanding we are considering. Observe that Scriven makes reference to a "proper context" in which the explanation is given, but does not fix any constraint on what he calls "levels of understanding":

> Hence the notion of the proper context for giving or requesting an explanation, which presupposes the existence of a certain level of knowledge and understanding on the part of the audience or inquirer, *automatically* entails the possibility of a complete explanation being given. And it indicates exactly what can be meant by the phrase "*the* (complete) explanation". For levels of understanding and interest define areas of lack of understanding and interest, and the required explanation is the one which relates to these areas and not to those other areas related to the subject of the explanation but perfectly well understood or of no interest (these would be explanations which could be correct and adequate but inappropriate) [Scriven, 1962, p. 202]

Even if we accept that explanation has the logical function of relating a phenomenon (not understood) with a set of phenomena which are under-

[7]Note the difference with the familiarity account: the familiarity partisans maintain that the phenomenon being explained is related to a *familiar* phenomenon, whereas for Scriven the phenomenon being explained must be related to a phenomenon which is already *understood*.

stood (realm of understanding), Scriven does not offer any characterization of the elements of such a set (for instance, by introducing a comparison of "levels of understanding"). Consequently, his account should be considered as incomplete.

Finally, these two positions (the familiarity view and Scriven's) assign to the explanans a necessary special epistemological feature. According to Friedman, this special epistemological status is not necessary, and these views suffer from lack of generality because they do not account for all cases of explanation in science.

Since what interests us are MEPP, observe here that in order to extend such accounts to the case of mathematical explanations in physics we would need a criterion to consider the explanans as more (or less) "familiar"/"understood" than the phenomenon which is explained. It is true that, in mathematical explanations of physical phenomena, a phenomenon is often explained by recurring to a mathematical fact which is considered as more familiar than the phenomenon itself. This is the case, for instance, of the explanation of the life-cycle of cicadas through the specific property of prime numbers, in Alan Baker's example seen in the previous chapter[8]. However, there are also cases where a less familiar or less understood "mathematical argument" is used to explain a more familiar or more understood phenomenon. For instance, the delta Dirac function was justified mathematically late after his introduction, in distribution theory. Nevertheless its use in mathematical physics came before its mathematical justification[9].

The third approach to explanation considered is that called by Friedman the "intellectual fashion view". In this group we find the positions endorsed by Stephen Toulmin [Toulmin, 1963] and, in a very different flavour, that of N. R. Hanson [Hanson, 1963]. In common with the familiarity account and Scriven's account, this view assigns to the explanantia special epistemolog-

[8]Naturally, the notion of "familiar" is not defined here. However, it is reasonable to say that the particular property of primes -known- is more familiar than the life-cycle of cicadas -unknown- (at least until the first is used in the explanation of the second).

[9]I will come back to the notion of understanding (in relation with MEPP) in the conclusive part of this work.

ical status, but contrary to the previous theories this special status is not static and varies with history. The first position, which is represented by Toulmin, states that during a particular historical period and within a particular historical tradition there exist phenomena which need no explanation (they are "ideals of intelligibility" or, as Toulmin calls them, "ideals of natural orders"). The business of explanation is to relate those self-explanatory phenomena to other unexplained phenomena (unexplained at *that* particular time). The second position, endorsed by Hanson, is different in the fact that the choice of an ideal of intelligibility is due to its qualities, for example its predictive power. According to Hanson scientific theories pass through three stages, and they finally become standards of intelligibility ("glass boxes") in the third stage, when they reach the ability to connect previously disconnected areas of research via predictions. In those stages of glass boxes the phenomena described by the theory are taken as paradigms of naturalness and comprehensibility.

The intellectual fashion approach offers then a notion of "scientific understanding" which varies through history, because ideals of intelligibility do. In spite of having a lot of historical support, for what counts as explanatory during a specific historical period is regarded as non-explanatory during another[10], Friedman observes that this view does not offer a common, objective and rational sense of explanation in which scientific theories explain; it does not offer a sense of explanation in science which is *constant* throughout the history of science.

(B) After his analysis of the problems that the notion of understanding has in some traditional accounts of scientific explanation, Friedman moves to step two of his project by listing the properties that a theory of explanation should have:

1. It should connect explanation and understanding: it should tell us what kind of understanding scientific explanations provide and how they provide it.

[10] See [Mischel, 1966] for various examples.

2. It should be sufficiently general: when tested on (most of) the theories that we consider to be explanatory it should mirror the explanatoriness of these theories.

3. It should be objective: it should capture the objective and rational sense in which scientific theories explain (if there is any!), i.e. it should not consider an explanation as depending on non-rational factors like "ideals of natural order" or "changing tastes of scientists and historical periods".

As we have seen, according to Friedman's analysis none of the views on explanation seen before satisfy both those conditions. This is why Friedman proposes an account of explanation which, according to him, satisfies these three desiderata.

3.1.1 Friedman's model

The idea that a scientific theory explains an empirical law by showing how this is an aspect of more comprehensive regularities, i.e. the idea that such a theory provides a "systematically unified account of many empirical laws", was suggested by Hempel himself [Hempel, 1966, p. 83]. However, the idea of unification in the sense of reduction of the number of independent phenomena appears for the first time in William Kneale's *Probability and Induction* [Kneale, 1949, p. 91]. This is the starting point of Friedman's construction of his model of explanation, whose 'unificatory' flavour emerges from the following passage:

> From the fact that *all* bodies obey the laws of mechanics it follows that the planets behave as they do, falling bodies behave as they do, and gases behave as they do. Once again, we have reduced a multiplicity of unexplained, independent phenomena to one. I claim that this is the crucial property of scientific theories we are looking for; this is the essence of scientific explanation science increases our understanding of the world by reducing the total number of independent phenomena

> that we have to accept as ultimate or given. A world with fewer independent phenomena is, other things equal, more comprehensible than one with more. [Friedman, 1974, p. 15]

But what does it mean "to reduce the total number of acceptable independent phenomena"? Friedman assumes that phenomena, i.e. general uniformities or patterns of beheviour, are representable by lawlike sentences. Thus, instead of speaking of independent phenomena, we can speak of logically independent lawlike sentences. The notion of 'acceptance' is defined by supposing that at any given time there is a set K of accepted lawlike sentences (i.e. phenomena), where accepted means "accepted by the scientific community". This set K is deductively closed in the sense that if S is a sentence and $K \vdash S$, then S is a member of the set K. So defined, K contains all lawlike consequences of members of K. The reduction is made on the total number of phenomena (lawlike sentences) we have to accept, thus the next step is to define what do we mean by "reduction of independent phenomena". In other words, we want to know when a given lawlike sentence S_i permits a reduction of the number of independent sentences of K. Nevertheless, here we are confronted with a famous problem (call this the *problem of conjunction*), which was acknowledged by Hempel and Oppenheim in footnote 28 of their 1948 paper on explanation. Here is the famous footnote 28:

> The precise rational reconstruction of explanation as applied to general regularities presents peculiar problems for which we can offer no solution.The core of the difficulty can be indicated briefly by reference to an example: Kepler's laws, K, may be conjoined with Boyle's law, B, to a stronger law $K.B$; but derivation of K from the latter would not be considered as an explanation of the regularities stated in Kepler's laws; rather it would be viewed as representing, in effect, a pointless "explanation" of Kepler's laws by themselves. The derivation of Kepler's laws from Newton's laws of motion and of gravitation, on the other hand, would be recognized as a genuine explanation in terms of more comprehensive regularities, or so-called higher-level laws [Hempel *et al.*, 1948, p. 159]

Thus the conjunction does not offer any interesting sense of reduction (of different laws to one), since the conjunction (one law expressed as the conjunction of n-laws) is equivalent to the set of independent laws (n-conjuncts). In order to surmount this problem, and have a useful sense of reduction, Friedman introduces the notion of 'independent acceptability'. Intuitively, the notion of independent acceptability captures the idea that if we have two sentences, S_1 and S_2, of which one (say, S_1) is acceptable independently (AI) of the other, we have grounds for accepting S_1 while the same grounds are not sufficient for accepting S_2. This would provide a possible way to deal with the problem of conjunction: although every sentence is equivalent to a set of n independent sentences, it is not the case that every sentence is equivalent to a set of n *independently acceptable* sentences.

Assuming that sufficient grounds for accepting a sentence S are also sufficient for accepting any consequence of S, if $AI(X, Y)$ means "X is acceptable independently of Y" we have:

(1) If $S \vdash Q$ then $\neg AI(S, Q)$ (where $\neg AI(S, Q)$ means that grounds for accepting S are also grounds for accepting Q)

(2) If $AI(S, P)$ and $Q \vdash P$, then $AI(S, Q)$.

Concerning $\neg AI(S, Q)$, observe that, since S could be a stronger statement than Q, grounds for accepting Q may be insufficient grounds for accepting S.

We can reformulate (1) by using the law of contraposition:

$$\text{If } AI(S, Q) \text{ then } \neg(S \vdash Q)$$

While the rationale for (1) is easy to see, we can reformulate (2) in the following more readable form:

$$\text{(2') Given } Q \vdash P, \text{ if } \neg AI(S, Q) \text{ then } \neg AI(S, P)$$

If Q entails P, by condition (1) we have that grounds for accepting Q are also sufficient for accepting P (i.e. they are not AI). Thus we conclude that grounds for accepting S are also sufficient grounds for accepting P, i.e. $\neg AI(S, P)$.

The next notion to define is that of "reduction" (of independent sentences). In order to obtain this definition Friedman introduces some terminology.

A *partition* of a sentence S is a set of sentences Γ such that

- $\Gamma \Longleftrightarrow S$ (logical equivalence)
- $AI(S', S)$ for every $S' \in \Gamma$.

Note that the members of the partition do not have to be mutually exclusive.

A K-*atomic* sentence is a sentence S which has no partition (in the set K of accepted lawlike sentences). This means that there is no set of sentences $\{S_1, S_2\}$ in K such that S_1 and S_2 are AI of S and $S_1 \wedge S_2$ is logically equivalent to S. For example, if S_1 is Boyle-Charles law, S_2 is Graham's law of diffusion and S_3 is Galileo's law of free fall, then the conjunction $S = S_1 \wedge S_2 \wedge S_3$ is *not* K-atomic, because there exists a partition Γ: $AI(S_1, S)$, $AI(S_2, S)$, $AI(S_3, S)$ and $\Gamma = \{S_1, S_2, S_3\}$ is logically equivalent to S[11]. The scheme in Figure 3.1 illustrates the example.

Given the notion of K-atomic sentence, let the K-*partition* of a set of sentences Δ be the set Γ^* of K-atomic sentences which is logically equivalent to Δ. In general, for a set Δ there may exist more than one K-partition. Then the K-*cardinality* of the set Δ is the number of the members of the smallest K-partition of Δ, i. e. the greatest lower bound of the cardinality of the K-partitions of Δ (Friedman assumes here the existence of the K-partition for every set of sentences Δ):

$$K\text{-card}(\Delta) = \inf\{\text{card}(\Gamma^*)\colon\ \Gamma^* \text{ a } K\text{-partition of } \Delta\}$$

[11] Boyle-Charles law and Graham's law of diffusion are both derived from the kinetic theory of gases. For instance, in deriving the Boyle-Charles law we make use of Newton's laws together with the assumption that gases are composed of molecules interacting only in collisions. Furthermore, the kinetic theory permits to integrate the behavior of gases with that of falling bodies near the earth. The behaviors of gases and falling bodies are both derived by using the laws of mechanics.

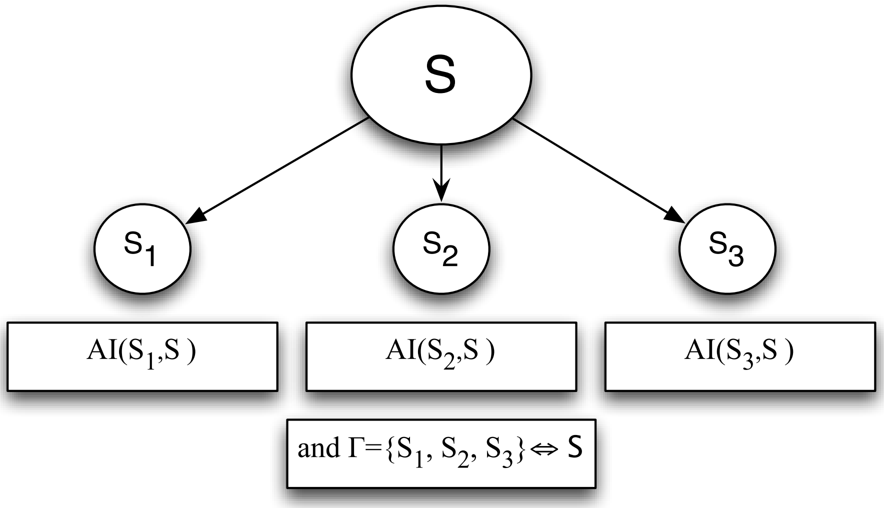

Figure 3.1: The conjunction $S = S_1 \wedge S_2 \wedge S_3$ is *not* K-atomic (S admits a partition Γ).

Note that the number of K-atomic sentences contained in any K-partition Γ^* of Δ is greater or, at least, equal to the number of elements of the set Δ. Thus, if $\Delta = BC \wedge Gr \wedge Gal \wedge Kep$, the K-card(Δ) is at least 4. The diagram in Figure 3.2 illustrates this situation in the case we have more than one K-partition of a set of n sentences $\Delta = \{\alpha_1, ..., \alpha_n\}$, and K-card(Δ)=inf$\{card(\Gamma_i^*)\} = p$.

We thus come to define what "a sentence reduces a set of sentences" stands for. A sentence S *reduces* the set of sentences Δ if the K-cardinality of the union of $\{S\}$ with Δ is smaller than the K-cardinality of Δ. More formally:

$$[reduction]\ S \text{ reduces } \Delta \text{ iff } K\text{-card}(\Delta \cup \{S\}) < K\text{-card}(\Delta).$$

What does this exactly mean? In our example of the conjunction of four independent laws (Boyle-Charles' law, Graham's law, Galileo's and Kepler's laws) we have the following situation: we want to know if $S = BC \wedge Gr \wedge Gal \wedge Kep$ reduces the set of its conjuncts $\Gamma = \{BC, Gr, Gal, Kep\}$. By applying the definition of reduction we have that S does *not* reduce Γ (and

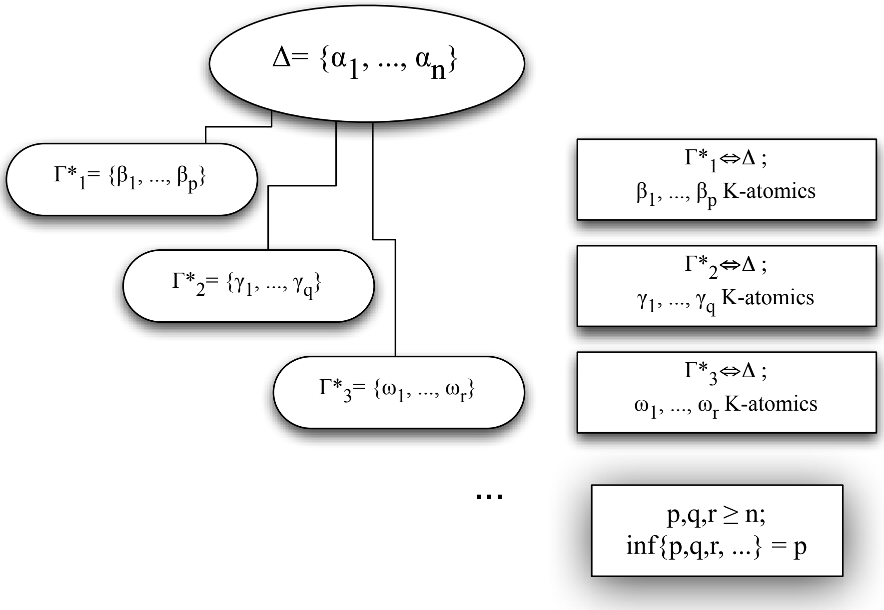

Figure 3.2: More than one K-partition of a set of n sentences $\Delta = \{\alpha_1, ..., \alpha_n\}$.

thus we have a solution to the problem of conjunction for the particular case):

- K-card($\Gamma \cup \{S\}$) $= K$-card($\{BC, Gr, Gal, Kep, BC \wedge Gr \wedge Gal \wedge Kep\}$) $=$ 6, because $\{BC, Gr, Gal, Kep_1, Kep_2, Kep_3\}$ is the smallest K-partition of the union[12].
- K-card(Γ)$=$ 6.
- Thus K-card($\Gamma \cup \{S\}$) $\nless K$-card(Γ) and we have that S does *not* reduce the set of its conjuncts Γ, as desired.

As a further step, Friedman introduces the idea of the *set of independently acceptable consequences* of a sentence S, namely $\text{con}_k(S)$: a sentence $S_1 \in \text{con}_k(S)$ iff $S \vdash S_1$ and $AI(S_1, S)$. The basic idea which he wants to capture

[12]Observe that $Kep = Kep_1 \wedge Kep_2 \wedge Kep_3$, but this conjunction is not K-atomic because there exists a partition Γ of it: $AI(Kep_1, Kep)$, $AI(Kep_2, Kep)$, $AI(Kep_3, Kep)$ and $\Gamma = \{Kep_1, Kep_2, Kep_3\}$ is logically equivalent to $Kep = Kep_1 \wedge Kep_2 \wedge Kep_3$.

by introducing the set $\mathrm{con}_k(S)$ is that a sentence S *explains* when it is able to reduce its set of independently acceptable consequences.

Friedman's definition of "explanation between laws" is given in terms of the previous terminology:

> How can we define *explanation* in terms of these ideas? If S is a candidate for explaining some S' in K, we want to know whether S permits a reduction in the number of independent sentences. I think that the relevant set we want S to reduce is the set of *independently acceptable* consequences of S ($\mathrm{con}_k(S)$). For instance, Newton's laws are a good candidate for explaining Boyle's law, say, because Newton's laws reduce the set of their independently acceptable consequences – the set containing Boyle's law, Graham's law, etc. On the other hand, the *conjunction* of Boyle's law and Graham's law is not a good candidate, since it does not reduce the set of its independently acceptable consequences. This suggests the following definition of explanation between laws:
>
> (D1) S_1 explains S_2 iff $S_2 \in \mathrm{con}_k(\mathrm{S}_1)$ and S_1 reduces $\mathrm{con}_k(\mathrm{S}_1)$

The previous definition of explanation (D1), however, is soon weakened by Friedman because it rules out the case of the conjunction of a sentence which explains with an irrelevant law:

> Actually this definition seems to me to be too strong; for if S_1 explains S_2 and S_3 is some independently acceptable law, then $S_1 \wedge S_3$ will not explain S_2 – since $S_1 \wedge S_3$ will not reduce $\mathrm{con}_k(\mathrm{S}_1 \wedge S_3)$. This seems undesirable – why should the conjunction of a completely irrelevant law to a good explanation destroy its explanatory power? So I will weaken (D1) to
>
> (D2) S_1 explains S_2 iff there exists a partition Γ of S_1 and an $S_i \in \Gamma$ such that $S_2 \in \mathrm{con}_k(\mathrm{S}_i)$ and S_i reduces $\mathrm{con}_k(\mathrm{S}_i)$
>
> Thus, if S_1 explains S_2, then so does $S_1 \wedge S_3$; for $\{S_1, S_3\}$ is a partition of $S_1 \wedge S_3$, and S_1 reduces $\mathrm{con}_k(\mathrm{S}_1)$ by hypothesis. [Friedman, 1974, p. 17-18]

Let's consider an example. The set $\{BC, Gr, Gal, Kep\}$ be the set of independently acceptable consequences of Newton's laws N_1,N_2,N_3. Does the sentence $N = N_1 \wedge N_2 \wedge N_3$ reduce $\Delta = \{BC, Gr, Gal, Kep\}$?

$$\{N\} = \{N_1 \wedge N_2 \wedge N_3\}$$

$$\Delta = \{BC, Gr, Gal, Kep\}$$

thus

$$K\text{-card}(\Delta \cup \{N\})=K\text{-card}(\{BC, Gr, Gal, Kep, N_1 \wedge N_2 \wedge N_3\})$$

The K-cardinality of the union set $\Delta \cup \{N\}$ is the number of the members of its smallest K-partition, which is $\{N_1, N_2, N_3\}$ (the set $\{N_1, N_2, N_3\}$ is logically equivalent to $\{BC, Gr, Gal, Kep, N_1 \wedge N_2 \wedge N_3\}$ and N_1, N_2, N_3 are K-atomics). K-card($\Delta \cup \{N\}$)=3 $<$ K-card(Δ)=6 and thus the sentence N containing the conjunction of Newton's laws reduces the set Δ of their independently acceptable consequences (as desired). As we have seen from a previous example, the conjunction of those laws $S = BC \wedge Gr \wedge Gal \wedge Kep$ does not reduce the set of its conjuncts $\Gamma = \{BC, Gr, Gal, Kep\}$ and then S it is not a good candidate for the explanation of one of the conjuncts BC, Gr, Gal, Kep, while Newton's laws (N) are. Hence, for example, Kepler laws $Kep \in con_k(N)$ and N reduces con$_k(N)$, thus Kep is explained by Newton's laws N (as stated by D1)[13].

However, as Friedman points out, definition (D1) used in the previous example seems to be too strong because in the case of the conjunction of a good explanation (for example N explains Kep) with an independently acceptable law (say, a law of quantum mechanics QML –Quantum Mechanical Law–), definition (D1) tells us that the conjunction $N \wedge QML$ does *not* explain Kep. This could be seen by observing that, according to (D1), we have that $N \wedge QML$ explains Kep iff

- $Kep \in$ con$_k$(N $\wedge QML$)

[13] The sentence $Kep \in con_k(N)$ because $N \vdash Kep$ and $AI(Kep, N)$.

- $(N \wedge QML)$ reduces con$_k$(N $\wedge QML$)

While the first condition poses no problem[14], the second is problematic because K-card(con$_k$(N $\wedge QML) \cup \{N \wedge QML\}$) will be greater or equal than K-card(con$_k$(N $\wedge QML$)). Thus, the sentence $N \wedge QML$ does not reduce the set of its independently acceptable consequences con$_k$(N $\wedge QML$) and therefore, by (D1), the conjunction $N \wedge QML$ does not explain Kepler laws Kep. More generally, $S_1 \wedge S_2$ does not reduce con$_k(S_1 \wedge S_2)$ if $AI(S_1, S_2)$ and $AI(S_2, S_1)$.

In order to solve this problem, the new definition of explanation proposed by Friedman states that:

> (D2) S_1 explains S_2 iff there exists a partition Z of S_1 and an $S_i \in Z$ such that $S_2 \in$ con$_k$(S$_i$) and S_i reduces con$_k$(S$_i$)

Let's now examine the same example of Newton's laws N and QML according to (D2). $(N \wedge QML)$ explains Kep iff

- there exists a partition Z of $(N \wedge QML)$
- there exists an $S_i \in Z$ such that $Kep \in$ con$_k$(S$_i$) and S_i reduces con$_k$(S$_i$)

The partition Z of $(N \wedge QML)$ exists and is $Z = \{N, QML\}$, where $N = N_1 \wedge N_2 \wedge N_3$ [15]. If we take N in this set Z, we see that $Kep \in$ con$_k(N)$ and N reduces, as we saw in the previous example, the set of its independently acceptable consequences con$_k(N)$. Thus the conjunction $(N \wedge QML)$ explains Kep, and the new definition (D2) permits that, if a lawlike sentence S_1 explains S_2, then so does the conjunction $S_1 \wedge S_3$, as desired. To put it roughly, Friedman solves the difficulty with (D1) by permitting an higher step in the projection of laws and integrating (D1) in the lower step. This is showed in Figure 3.3.

[14] The sentence $Kep \in con_k(N \wedge QML)$ because $(N \wedge QML) \vdash Kep$ and $AI(Kep, N \wedge QML)$.

[15] The partition exists because: the set Z is logically equivalent to the sentence $(N \wedge QML)$; furthermore, $AI(N, N \wedge QML)$ and $AI(QML, N \wedge QML)$.

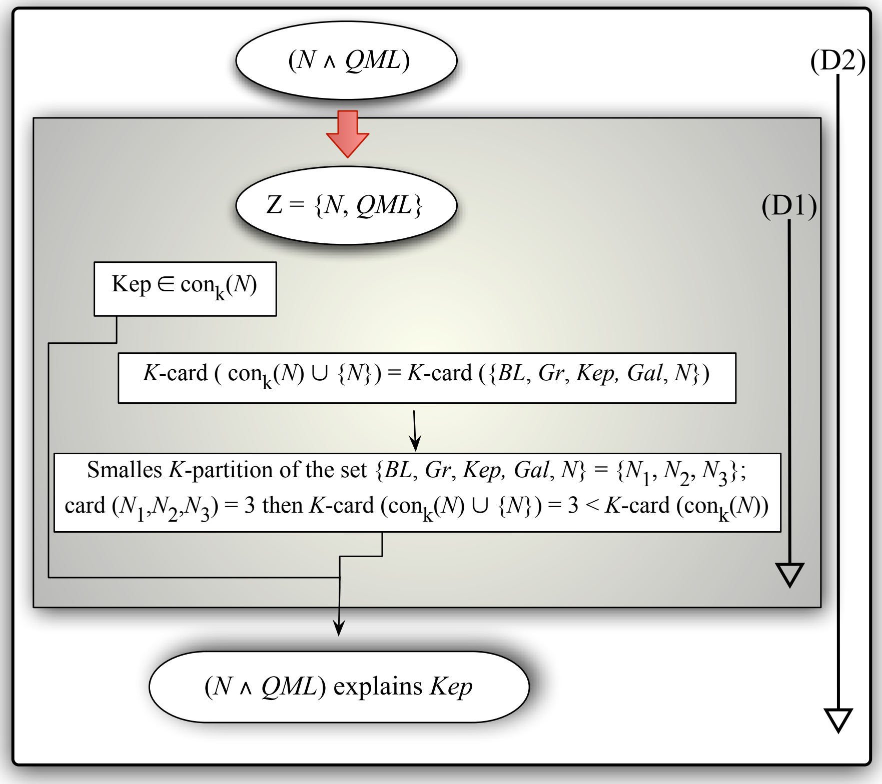

Figure 3.3: The conjunction $(N \wedge QML)$ explains Kep according to (D2).

According to Friedman, his definition (D2) permits to solve the difficulty with the D-N model, where the conjuction of two independent laws always entails each of the conjuncts while it does not necessarily explains them. In addition, he considers that his approach provides an account of when we increase our understanding, and thus a linkage between an objective definition of explanation and scientific understanding.

By concentrating on what he calls the "local aspects of explanation" (the explanandum, the explanans, the deductive relation between the two), traditional theories of scientific explanation have given a special epistemological status to the explanans. Nevertheless, according to Friedman, this adoption of non-rational factors (familiarity, etc..) should be abandoned, as is clear from the previous discussion. Friedman's idea is that in order to connect explanation and understanding, and avoid the problem of giving a special epistemological status to the explanans, the focus should be put on the "global aspects of explanation" and on the "global nature of scientific understanding".

In Friedman's view, thus, we increase our (global) understanding when we can simplify our total picture of nature by a reduction in the number of independent phenomena we accept as ultimate (where the notion of reduction is definited as above). Therefore the sense in which we replace a phenomenon with another, in order to increase our understanding, does not amount to a simple replacement of a puzzling phenomenon with another one (recall that Friedman's assuption is that phenomena are always representable by law-like sentences). The laws of Newtonian mechanics allow us to derive the fact that planets obey Kepler's laws, the fact that terrestrial bodies obey Galileo's laws of free fall, the behaviour of gases, etc. If we agree with Friedman, we should be able to trace the path of the scientific understanding enterprise by focusing on scientific laws, at a global level.

> [...] Scientific understanding is a global affair. We don't simply replace one phenomenon with another. We replace one phenomenon with a *more comprehensive* phenomenon, and thereby effect a reduction in the total number of accepted phenomena. We thus genuinely increase our understanding of the world. [Friedman, 1974, p. 19]

Thus, while trying to find an answer to the general question "What is the *nature* of the understanding scientific explanations are supposes to convey?" [Salmon, 1989, p. 127], Friedman addressed the different question "*How* do we increase our understanding?". As observed by Salmon, Friedman's answer to this question in terms of unification can also be read in terms of information theory [Salmon, 1989, p. 131], for what is important for explanation is the *way* in which our descriptive knowledge is organized and not some particular kind of explanatory knowledge (as, for instance, causal knowledge).

3.1.2 Kitcher's criticism of Friedman's model

Friedman's model was put under discussion by Kitcher in his [Kitcher, 1976]. According to Kitcher, Friedman's definition of explanation (D1) admits only K-atomic sentences as explananda, while (D2) rules out K-atomic sentences as explananda. From these observations there follow further objections, which are summarized by two types of counterexamples. Before seeing these counterexamples, let me illustrate Kitcher's arguments for the two claims:

Cl_1 According to (D1) only K-atomic sentences can explain

Cl_2 According to (D2) no K-atomic sentences can explain

The latter claim Cl_2 comes directly from Friedman's definition (D2): "S_1 explains S_2 iff there exists a partition Z of S_1 and ...". If S is K-atomic, it has no partition (from the definition of K-atomic sentence)[16] and then it does not explain.

Kitcher's claim Cl_1 amounts to saying that: if S is not k-atomic, then it does not reduce the set of its independently acceptable consequences $\text{con}_k(S)$. He reaches this conclusion by the following argument: we assume that S is not k-atomic; then there exists a partition Γ such that for each sentence $A_i \in \Gamma$ it holds $AI(A_i, S)$ and Γ is logically equivalent to S; but then, since for each $A \in \Gamma$ it holds $S \vdash A$, it follows that $\Gamma \subseteq \text{con}_k(S)$; if $S \vdash C$ we have that $\Gamma \vdash C$ (since S and Γ are logically equivalent) and thus that $\text{con}_k(S) \vdash C$ (since $\Gamma \subseteq \text{con}_k(S)$); conversely, if now we consider $\text{con}_k(S) \vdash C$, then only a finite subset of $\text{con}_k(S)$ is used in the deduction; if B is the conjunction of the sentences in this finite subset, then $B \vdash C$ and $S \vdash B$; hence $S \vdash C$; we have showed that S is logically equivalent to its set of independently acceptable consequence $\text{con}_k(S)$; now, if S is logically equivalent to $\text{con}_k(S)$, then $\text{con}_k(S)$ is logically equivalent to $\text{con}_k(S) \cup \{S\}$; thus any K-partition of $\text{con}_k(S)$ is also a K-partition of $\text{con}_k(S) \cup \{S\}$, and conversely. Therefore K-card$(\text{con}_k(S))$ is equal to K-card$(\text{con}_k(S) \cup \{S\})$, and consequently

[16] If A is K-atomic, $\Gamma = \{A\}$ is logically equivalent to A but A is not independently acceptable from A.

(from Friedman's definition of reduction) S does not reduce the set of its independently acceptable consequences K-card(con$_k(S)$). Finally, only K-atomic sentences can explain according to (D1).

Kitcher shows that the problem with (D1) is that, by allowing the fact that only K-atomic sentences can explain, Friedman's account rules out trivial conjunctions (and then solves the problem of conjunction) but it also rules out a number of *bona fide* scientific explanations. Kinds of genuine scientific explanations which are ruled ou by (D1) are illustrated by Kitcher by proposing two sorts of counterexamples. Furthermore, Kitcher points out that these counterexamples, to which Friedman's theory is vulnerable, are not solved by (D2). The first type of counterexamples (CE1) occurs when

> we have independently acceptable laws which belong to the same theory and which can be put together in genuine explanations. The explananda that result are not K-atomic and hence fail to meet the necessary condition derived from Friedman's theory [Kitcher, 1976, p. 209]

For instance, the law of adiabatic expansion of an ideal gas could be derived from the conjunction of two laws (Boyle-Charles law and the first law of thermodynamics) acceptable on the basis of quite independent tests. Therefore the conjunction is not K-atomic (as required by D1), but the derivation of the law of adiabatic expansion from this conjunction seems to be a genuine explanation. The second type of counterexample (CE2) arises when we use laws coming from different theories in order to explain a complex phenomenon. In this case, as Kitcher observes, the theories are often independently acceptable and the laws drawn from them are also independently acceptable. For example, the explanation of why lighting flashes are followed by thunderclaps utilizes laws of electricity, thermodynamics and acoustics, which are independently acceptable. Or

> a complete explanation of why human eyes are sensitive to a particular range of light frequencies (an explanation that would involve indepen-

dently acceptable laws drawn from evolutionary biology, geophysics, and optics) [*Ibid.*, p. 210]

Does Friedman's definition (D2) avoid the previous counterexamples? Kitcher's opinion is that it does not. On the basis of claims Cl_1 and Cl_2, Kitcher reinforces Friedman's definition (D2) by using the right-hand side of (D1):

> I suspect that Friedman's intentions would be better captured by an equivalence whose right-hand side consisted of the disjunction of the right-hand sides of (D1) and (D2) [*Ibid.*, p. 211]

More explicitly, the new definition D3 is the following:

(D3) S_1 explains S_2 iff $[(S_2 \in \text{con}_k(\text{S}_1)$ and S_1 reduces $\text{con}_k(S_1)) \vee$ (there exists a partition Γ of S_1 and an $S_i \in \Gamma$ such that $S_2 \in \text{con}_k(\text{S}_i)$ and S_i reduces $\text{con}_k(\text{S}_i))]$

Observe that, since the right-hand side of D3 contains the right-hand sides of (D1) and (D2) in a disjunction, to provide a counterexample for (D3) amounts to providing a counterexample for (D1) and also for (D2). We have already seen Kicther's counterexamples (CE1) and (CE2) for (D1). Kitcher's next step is to show that (D3) –and then (D2)– does not avoid these counterexamples. For instance, he focuses on counterexample (CE1), i.e. the explanation of the law of adiabatic expansion for an ideal gas.

As we have seen above, (D1) does not capture the explanatory character of the conjunction of the two laws (Boyle-Charles law and the first law of thermodynamics – acceptable on the basis of quite independent tests) from which the law of adiabatic expansion for an ideal gas can be derived. What about (D2)? Let T be the first law of thermodynamics, B the Boyle-Charles law, and A the law of adiabatic expansion for an ideal gas. According to (D2), A can be derived from $T \wedge B$ (as a genuine explanation) only if the following two conditions hold:

- There exists a partition Γ of $T \wedge B$
- There is an $S_i \in \Gamma$ such that $A \in \mathrm{con}_k(S_i)$ and S_i reduces $\mathrm{con}_k(S_i)$

By what we said above (Cl_1 and Cl_2), the sentence S_i must be K-atomic. The problem is then to find Γ and S_i meeting this condition.

Assume that there were such a Γ and S_i. Consider the set $\{(S_i \vee T), (S_i \vee B)\}$, which is equivalent to the sentence $S_i \vee (T \wedge B)$. Since Γ is a partition of $T \wedge B$, we have that $(T \wedge B) \vdash S_i$. But then the set $\{(S_i \vee T), (S_i \vee B)\}$ is equivalent to S_i. Since S_i is K-atomic, it has no partition. Therefore one of $(S_i \vee T)$ and $(S_i \vee B)$ is not acceptable independently of S_i. As a consequence, if we want to find a partition Γ and a sentence S_i that will allow the explanation of the law of adiabatic expansion for an ideal gas to stand as a genuine explanation, we have to find a S_i such that either T or B is not acceptable independently of S_i. However, when we consider our grounds for accepting B and our grounds for accepting T, there is *no* law L that meets the two conditions:

- that all our sufficient grounds for accepting B (or all our sufficient grounds for accepting T) be sufficient grounds for accepting L
- $L \vdash A$

Therefore it seems that Γ and S_i are not available and that his (D2) does not save his theory from the types of counterexamples raised by Kitcher. However, as Kitcher himself observes, these counterexamples are not avoided by Friedman's definitions exactly because they do not correspond to Friedman's sense of explanation, i.e. they do not work by reducing the number of independent phenomena:

> Moreover, it would be contrary to the spirit of Friedman's whole enterprise if (D2) (or the disjunctive condition I have suggested as a more exact representation of his views) *did* allow that my counterexamples are genuine explanations because of the existence of some unobvious

> way of partitioning the explanantia. For Friedman's central insight is that we attain understanding when we see how to reduce the number of independent phenomena. The explanations cited above give genuine understanding in their own right; we do not obtain this understanding by supposing that some partition Γ containing an appropriate S_i lurks behind them and does the real work of reducing the number of independent phenomena. [Kitcher, 1976, p. 211-212]

Finally, we come to what Kitcher considers the moral of his criticism. While he agrees with Friedman that a theory of explanation should connect explanation and unification, his point is that Friedman was wrong in considering *laws* as the key aspect of unification. This is why Friedman's model is not able to account for some cases of explanations which seem to be genuine.

> Finally, I think that it is not hard to see why Friedman's theory goes wrong. Although he rightly insists on the connection between explanation and unification, Friedman is incorrect in counting phenomena according to the number of independent laws. [...] What is much more striking than the relation between these numbers is the fact that Newton's laws of motion are used again and again and that they are always supplemented by laws of the same types, to wit, laws specifying force distributions, mass distributions, initial velocity distributions, etc. Hence the unification achieved by Newtonian theory seems to consist not in the replacement of a large number of independent laws by a smaller number, but in the repeated use of a small number of types of law which relate a large class of apparently diverse phenomena to a few fundamental magnitudes and properties. Each explanation embodies a similar pattern: from the laws governing the fundamental magnitudes and properties together with laws that specify those magnitudes and properties for a class of systems, we derive the laws that apply to systems of that class. [Kitcher, 1976, p. 212]

Thus, for Kitcher, not simply laws but something else should be central to explanatory unification: patterns.

3.2 Kitcher's unification

In this section I am going to present Kitcher's unification model, which has been presented by Kitcher's in two different papers: [Kitcher, 1981] and [Kitcher, 1989][17]. In addition to those two papers, an important key in order to understand Kitcher's model is his book *The Nature of Mathematical Knowledge* [Kitcher, 1984], which specifically concerns the rational growth of mathematical knowledge. By focusing specifically on mathematical practice, this book represented a decisive step in the integration of history with philosophy of mathematics, against that form of "mathematical apriorism" which denied any active role of history in the study of mathematical knowledge[18].

As Friedman, Kitcher does not make distinction between mathematical explanations and scientific explanations. Furthermore, as Steiner, Kitcher considers mathematics and science on par from a methodological point of view:

> [...] given my own views on the nature of mathematics, mathematical knowledge is similar to other parts of scientific knowledge, and there is no basis for a methodological division between mathematics and natural sciences [Kitcher, 1989, p. 423]

In addition to this methodological continuity, Kitcher explicitly claims that his model is able to cover mathematical explanations as well:

> [...] even in areas of investigation where causal concepts do not apply – such as mathematics – we can make sense of the view that there are patterns of derivation that can be applied again and again to generate a variety of conclusions. [...] The fact that the unification approach provides an account of explanation, and explanatory symmetries, in mathematics stands to its credit. [Kitcher, 1989, p. 437]

Let's now see his model in detail.

[17]Now systematized in his [Kitcher, 1993].

[18]For Kitcher's objections against the "apriorist epistemology of mathematics" see his [Kitcher, 1980], and the recapitulation he gives in his [Kitcher, 1988].

3.2.1 Kitcher's model

Kitcher's starting points are Friedman's 1974 paper and the idea that behind the D-N model there was an "unofficial model" which considered explanation as unification and which seemed to be more promising than the official one[19]. However, differently from Hempel, Kitcher regards explanation as an "activity". More precisely, following Peter Achinstein [Achinstein, 1983, p. 84-85], he regards an explanation as an ordered pair consisting of a proposition p and an act type *(Mr X's act of) explaining* q: (p, explaining q). For Achinstein, what makes the ordered pair an explanation is the appropriate relation the sentence expressing p bears with a particular argument[20]. Kitcher's idea is that there exist acts of explanation which draw on scientific arguments. Hence the basic problem is to determine the features such scientific arguments should have in order to serve as the basis for acts of explanation, i.e. the conditions which must be met if a scientific argument whose conclusion is S is used in answering the question 'Why is it the case that S?'.

Thus, although adopting the notion of argument to characterize that of explanation, Kitcher abandons the Hempelian (ontological) thesis that ex-

[19] I have already mentioned Hempel's passage in subsection 3.1.1. Here is the entire passage: "What scientific explanation, especially theoretical explanation, aims at is not [an] intuitive and highly subjective kind of understanding, but an objective kind of insight that is achieved by a systematic unification, by exhibiting the phenomena as manifestations o common, underlying structures and processes that conform to specific, testable, basic principles" [Hempel, 1966, p. 83].

[20] More formally, for Achinstein E is an explanation of q if and only if (i) Q is a content-question (in direct form; the indirect form of Q is q), (ii) E is an ordered pair whose first member is a complete content-giving proposition (p) with respect to Q and whose second member is the act-type 'explaining q'. Question Q is a content-question if and only if there is a complete content-giving proposition with respect to Q. A complete content-giving proposition with respect to Q is a proposition that entails all of Q's presuppositions. For instance, if Q is 'Why did Peter get a stomach ache?', the propositions 'The reason that Peter got a stomach ache is that he ate spoiled meat' and 'Bill got a stomach ache because he ate spoiled meat' are complete content-giving propositions, while the proposition 'Bill ate spoiled meat' is not. Furthermore, Achinstein shows how an explanation acceptable within the D-N model can be viewed in terms of ordered pairs: E is an explanation of q if and only if (i) Q is a question of the form 'Why is it the case that p?', (ii) E is a D-N argument, whose conclusion is p, which contains lawlike sentences and satisfies Hempel conditions (i.e. logical and empirical conditions of adequacy) [Achinstein, 1983, p. 93].

planations *are* arguments. Furthermore, his position is very different from Van Fraassen's, for Kitcher claims that *there are* context-*in*dependent features of arguments which make them explanatory, and scientific theories can be investigated under the light of their capacity to provide us with with such arguments[21]. He regards Newtonian theory and Darwin's evolutionay theory as examples of such theories.

Nevertheless, if we want to consider particular arguments with some context independent feature, we are confronted with two problems: where do those arguments come from? What are their particular features? Kitcher answers the first question in a way very similar Friedman answered the question about the existence of a set of lawlike sentences:

> The set of arguments which science supplies for adaptation in acts of explanation will change with our changing beliefs. Therefore the appropriate analysandum is the notion of the store of arguments relative to a set of accepted sentences. Suppose that, at the point in the history of inquiry which interests us, the set of accepted sentences is K. (I shall assume, for simplicity's sake, that K is consistent. Should our beliefs be inconsistent then it is more appropriate to regard K as some tidied version of our beliefs.) The general problem I have set is that of specifying $E(K)$, the explanatory store over K, which is the set of arguments acceptable as the basis for acts of explanation by those whose beliefs are exactly the members of K. (For the purposes of this paper I shall assume that, for each K there is exactly one $E(K)$.) [Kitcher, 1981, p. 512]

While the previous quotation states that for each consistent and deductively closed set K of beliefs (endorsed by a scientific community at a particular time) there is exactly one set $E(K)$ of arguments acceptable as the basis for acts of explaining (i.e. the *explanatory store over* K), we do not

[21]As we have seen in chapter 2, for Van Fraassen there are no context-independent features (beyond those of simplicity and empirical adequacy) which distinguish arguments for use in explanations.

still have an answer to our second question about the features of the arguments which characterize this set. A very general and indirect answer to this question is that $E(K)$ is the set of arguments which *best unifies* K, or, in Kitcher technical language (that we will see below), that $E(K)$ is the *best systematization of* K.

As I have already observed, the notion of *pattern*, and precisely that of *pattern of arguments*, will be central to the definition of Kitcher's unification. In particular, for him, the unifying power of a theory (for instance Newton's theory) consists in the fact that, by using the same pattern of derivation again and again, the theory shows us how to derive a large number of sentences which we accept. This is the way in which science advances our understanding of nature [Kitcher, 1989, p. 432]. The linkage between $E(K)$ and the notion of pattern emerges from the following passage:

> So the criterion of unification I shall try to articulate will be based on the idea that $E(K)$ is a set of derivations that makes the best tradeoff between minimizing the number of patterns of derivation employed and maximizing the number of conclusions generated. [Kitcher, 1989, p. 432]

Thus, by considering arguments as derivations, Kitcher considers his proposal different from that of Hempel, for the latter looked at arguments as premise-conclusion pairs. In Kitcher's discussion an argument is a derivation, i.e. "a sequence of statements whose status (as a premise or as following from a previous members in accordance with some specified rule) is clearly specified" [Kitcher, 1989, p. 431]. For a derivation to count as an acceptable explanation, it must belong to the explanatory store over K. In order to clarify the previous notions, let me introduce the formal structure of Kitcher's model.

We have already introduced the basic notion of set of accepted sentences K. A given set of beliefs K may have different ways of deriving some of its sentences from others, and each of these ways -consisting in a set of derivations- is a systematization for K. Thus a systematization Σ of K is a set of arguments which derives some members of K from other members of K. The explanatory store over K, $E(K)$, is the best systematization of K, i.e. the systematization which provides the highest degree of unification. Here Kitcher makes the idealization that $E(K)$ is unique for every K. He also requires that all the arguments in $E(K)$ be *acceptable relative* to K. To say that a set of derivation is acceptable relative to K is to say that each step in each derivation in the set of derivations is deductively valid and each premise of each derivation belongs to K. Thus, according to Kitcher, in finding systematizations of K we restrict our attention to sets of arguments (or derivations) which are acceptable relative to K. But now we are left with the following question: what are the criteria which permit to say that one systematization is better than another (and that $E(K)$ is the *best* systematization?).

The evaluation of a systematization Σ is made by using three different notions:

- *argument pattern*
- *generating set* for Σ
- *conclusion set* for Σ

The concept of pattern discussed here by Kitcher, i.e. that argument pattern offered by theories which can be used to derive a large number of accepted sentences, is not the same as that familiar from formal logic (say, a purely logical pattern). However, due to the particular way in which the non-logical vocabulary (the vocabulary made by non-logical terms as "force", "mass", etc) occurs in the instantiation of a pattern by an argument, Kitcher observes that logic (and logicians) helps us to isolate and study the notion of argument pattern. A *schematic sentence* is an expression obtained by replacing some or all of the non-logical expressions occurring in a sentence with dummy letters, while a set of *filling instructions* tells us how to replace the dummy letters in a schematic sentence (the "direction" for replacing dummy letters). A *schematic argument* is a sequence of schematic sentences, and a *classification* for it is a set of sentences which describe the inferential characteristics of the schematic argument and the role each schematic sentence has in it.

Finally, a *general argument pattern* $\langle s, f, c \rangle$ is a triple consisting of a schematic argument s, a set f of filling instructions and a classification c for s. Kitcher's favorite example of theories which offer unification in his sense are Newtonian's theory and Darwin's evolutionary theory. Concerning the latter, Kitcher observes that the strategy in Darwin's *On the origin of species* consists precisely in showing that "certain kinds of modifications of beliefs make possible an increase in unification of a large number of biological phenomena" [Kitcher, 1989, p. 491][22]. Here I will use the Newtonian case to give a concrete idea of the notion of argument pattern[23].

The following schematic sentences (1)-(5) form a schematic argument s_N:

1. The force on α is β
2. The acceleration of α is γ
3. Force = mass $\cdot$ acceleration
4. (Mass of α)$\cdot$(γ) = β
5. $\delta = \theta$

The set of filling instructions f_N contains the directions for replacing the dummy letters α, β, γ, δ, θ in every schematic sentence. The members of f_N are: "all occurrences of α are to be replaced by an expression referring to the body under investigation"; "occurrences of β are to be replaced by an algebraic expression referring to a function of the variable coordinates and of time"; "γ is to be replaced by an expression which gives the acceleration of the body as a function of its coordinates and their time-derivatives"; "δ is

[22] For a discussion of Darwin's example see [Kitcher, 1981], [Kitcher, 1985a] and [Kitcher, 1989].

[23] For an example of argument pattern in the case of mathematics see [Hafner *et al.*, 2008, p. 213-214]. I will come back to this in the last section.

to be replaced by an expression referring to the variable coordinates of the body, and θ is to be replaced by an explicit function of time". The sentences contained in the classification set c_N for the schematic argument s_N give us the inferential informations about the schematic argument: "(1)-(3) have the status of premises"; "(4) is obtained from (1)-(3) by substituting identicals"; "(5) follows from (4) using algebraic manipulations and the techniques of the calculus". Thus we have that a particular derivation in Newtonian mechanics, i.e. a sequence of sentences and formulas which accord Newton's laws, instantiates the general argument pattern $\langle s_N, f_N, c_N \rangle$ just in case: (*i*) the derivation has the same number of terms as the schematic argument s_N, (*ii*) each sentence or formula in the derivation can be obtained from the corresponding schematic sentence in accordance with the filling instructions f_N, (*iii*) the terms of the derivation have the properties assigned by the classification c_N to members of the schematic argument s_N.

As an illustration, consider the following example: a projectile P of mass m_p is fired horizontally from a tower of height y_0. Take the top of the tower as the origin of an xy Cartesian coordinate system (the projectile starts its flight parallel to the x axis, i.e. parallel to the ground), and make the idealization that there is no air-resistance. The initial speed of the projectile is v_0. Of course, also this condition on the speed is an idealized one: we assume that when the projectile is fired, at $t = 0$, it has velocity v_0; however, at time $t = 0$ the projectile is at rest and then it is subject to an instantaneous acceleration in some small interval of time $t + \varepsilon$. The horizontal and vertical components of the velocity vector $\mathbf{V_0}$ are, respectively, $\dot{x} = v_{x0} = v_0 \cos\alpha$ and $\dot{y} = v_{y0} = v_0 \sin\alpha$, where α is the angle formed by the velocity vector and the x-axis. Velocity v_0 also corresponds to the value of the initial horizontal component of the velocity ($\dot{x} = v_{x0} = v_0 \cos\alpha = v_0$, with $\alpha = 0$), which in absence of air-resistance remains constant throughout the motion. On the other hand, the initial vertical component of the velocity is 0 ($\dot{y} = v_{y0} = v_0 \sin\alpha = 0$, with $\alpha = 0$), and it increases in magnitude during the motion (at a costant rate g, as we are going to see in a moment). The only force the projectile is subject to is the gravitational force F_G, which acts downward.

The motion of the projectile in the plane xy is described by the vector equation:

$$m_p \frac{d^2\mathbf{r}}{dt^2} = -m_p \mathbf{g}\hat{j} \tag{3.1}$$

or, in component form,

$$m_p \frac{d^2x}{dt^2} = 0 \quad \text{and} \quad m_p \frac{d^2y}{dt^2} = -m_p g \tag{3.2}$$

Using the initial conditions given above and the equation of the motion, we can now consider the first and second time integrals for the two components (horizontal and vertical):

$$\begin{array}{rl} \text{Horizontal} & \text{Vertical} \\ \dot{x} = v_{x0} = \text{constant} & \dot{y} = v_{y0} - gt = -gt \quad (3.3) \\ x = v_{x0}t & y = -\frac{1}{2}gt^2 \quad (3.4) \end{array}$$

Finally, the two parametric equations $x = v_{x0}t$ and $y = -\frac{1}{2}gt^2$ could be written as a single vector equation (an expression in which the variable coordinates of the body are espressed in function of time)[24]:

$$\mathbf{r} = \mathbf{V_0}t + \frac{1}{2}\mathbf{g}t^2 \tag{3.5}$$

Let's now see how the previous derivation in Newtonian mechanics instantiates the general Newtonian argument pattern $\langle s_N, f_N, c_N \rangle$. The set of filling instructions f_N permits to substitute the dummy letters α, β, γ, δ, θ in every schematic sentence: α is replaced by 'projectile'; β by '$-m\mathbf{g}$'; γ by '$\frac{d^2\mathbf{r}}{dt^2}$'; δ by '$\mathbf{r}(x, y)$' and θ by '$\mathbf{V_0}t + \frac{1}{2}\mathbf{g}t^2$'. Therefore we have, for the particular case of the projectile seen above, the following schematic argument:

[24]Note the sign plus in the equation. The vector $\mathbf{g}$ is directed downward.

1* The force on the projectile is $-m\mathbf{g}$

2* The acceleration of the projectile is $\dfrac{d^2\mathbf{r}}{dt^2}$

3* Force = mass $\cdot$ acceleration

4* $-m\mathbf{g} = m_p \cdot \dfrac{d^2\mathbf{r}}{dt^2}$

5* $\mathbf{r}(x, y) = \mathbf{V_0}t + \dfrac{1}{2}\mathbf{g}t^2$

The classification set c_N for the schematic argument s_N gives us the inferential informations about the schematic argument. It tells us that: "(1*)-(3*) have the status of premises"; "(4*) is obtained from (1*)-(3*) by substituting identicals"; "(5*) is deduced from (4*) using algebraic manipulations and the techniques of the calculus –integration above". Finally, (*i*) the derivation has the same number of terms as the schematic argument s_N (observe that velocity can be seen as the first integral of acceleration, while the coordinates of the body as the second integral of acceleration); (*ii*) each sentence or formula in the derivation can be obtained from the corresponding schematic sentence in accordance with the filling instructions f_N; (*iii*) the terms of the derivation have the properties assigned by the classification c_N to members of the schematic argument s_N.

By emphasizing the differences with a purely logical formulation, Kitcher stresses the fact that in his schema mathematical assumptions do not occur as terms of the schematic argument. This could be seen very easily from the Newtonian example, where, for example, β must be replaced with an algebraic expression. The focus here is not on the importance of the particular algebraic expression used, but on the importance of having a pattern of derivation of this type. The attention must be turned on the general theoretical framework rather than on some particular feature of the argument or of the phenomenon under study, as it was the case for Steiner's account.

> Whereas logicians are concerned to display all the schematic premises which are employed and to specify exactly which rules of inference are used, our example allows for the use of premises (mathematical

> assumptions) which do not occur as terms of the schematic argument and it does not give a complete description of the way in which the route from (4) to (5) is to go. Moreover, our pattern does not replace all nonlogical expressions by dummy letters. Because some nonlogical expressions remain, the pattern imposes special demands on arguments which instantiate it. In a different way, restrictions are set by the instructions for replacing dummy letters. [Kitcher, 1981, p. 517-518]

The explanatory activity of scientists results from a compromise in demaning two kinds of similarity for arguments: similarity in the terms of logical structure and similarity in terms of the non-logical vocabulary they employ at corresponding places. Kitcher's idea is that arguments similar in either of these ways share a common pattern, hence he tries to capture those desiderata by introducing the notion of *stringency* of a pattern. Without offering a detailed analysis of this notion, he proposes that the stringency of a pattern is subject to two different constraints: 1) the conditions on the logical structure, imposed by the classification (strictness in the characterization of the inferential principles); 2) the conditions on the substitution of dummy letters, jointly imposed by the presence of non-logical expression in the pattern (nature of the schematic sentences) and by the filling instructions (strictness in the instantiation conditions). Consider now our Newtonian example: although not having the same logical structure, arguments instantiating the Newtonian pattern share some similarities in the logical structure (imposed by the classification c_N), while they make use of non-logical vocabulary at same places.

A second notion we need in order to test a systematization is the notion of *generating set* for a set of derivations Σ. Within each systematization there are derivations which instantiate different argument patterns. A generating set for a given systematization Σ is a set of argument patterns Π such that every acceptable derivation in Σ instantiates an argument pattern in Π. A generating set will be said to be *complete with respect to K* if and only if every derivation which is acceptable relative to K and which instantiates a pattern in Π also belongs to Σ[25].

Now, as we have seen before, $E(K)$ is that set of derivations which makes the best tradeoff between minimizing the number of patterns of derivation employed and maximizing the number of conclusions generated. Kitcher suggests the following steps for the determination of $E(K)$:

(1) Choose between the different systematizations of K only those systematizations (set of arguments) which are acceptable relative to K, i.e. the systematizations for which each step in each derivation is deductively valid and each premise of each derivation belongs to K.

(2) Consider, for each systematization, the various generating sets Π which are complete with respect to K.

(3) For each (acceptable) systematization select among the collection of generating sets Π that set with the best unifying power. This should be done according to the criterion of minimizing the number of patterns of derivation employed (criterion of *paucity of patterns*) and that of maximizing the stringency of a pattern. Call this set the *basis* B for that systematization. A generating set which contains few patterns scores better than a generating set which contains a bigger number of patterns.

(4) The explanatory store over K is that systematization whose basis does the best in terms of unifying power: $E(K)$ is the best systematization of K.

While in step (3) the criterion of minimizing the number of patters of derivation permits us to choose a basis for each systematization, in order to

[25] To put it in a different way, Π is said to be complete if and only if there is no acceptable derivation in K which instantiates a pattern from Π and is not included in the systematization generated by Π (incomplete generating sets should be forbidden since they use patterns selectively by not recognizing all possible acceptable instantiations of a pattern).

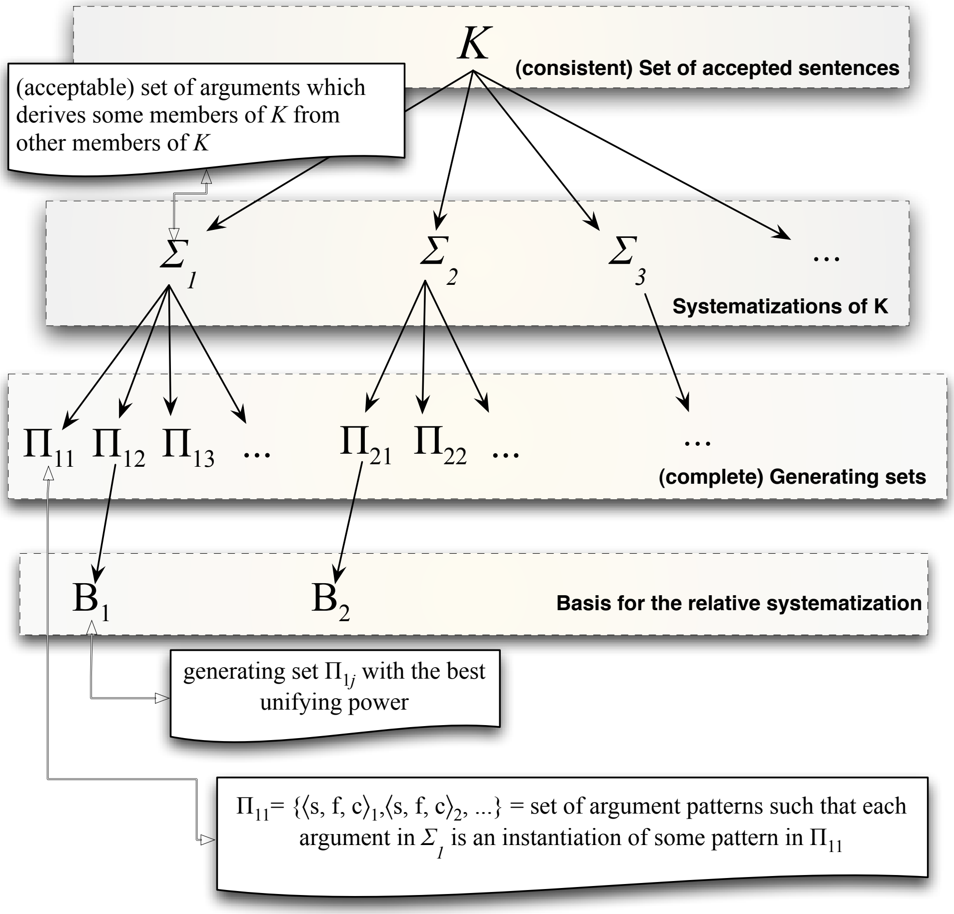

Figure 3.4: Kitcher's unification scheme.

rank the basis (and then the systematization, for each basis refer to one systematization) in step (4) we need something more. Kitcher's intuitive idea is that we have unifying power for a set of argument patterns Π "by generating a large number of accepted sentences as the conclusions of acceptable arguments which instantiate a few, stringent patterns" [Kitcher, 1981, p. 520]. Thus the criterion of evaluation must say something about the number of conclusions generated by a set of argument patterns. We introduce here the third key-notion, i.e. that of conclusion set for Σ: the *conclusion set* $C(\Sigma)$ of a set of arguments Σ is the set of sentences which occur as conclusions of some argument in Σ.

Thus the unifying power (or degree of unification) of a systematization Σ_i, which corresponds to the unifying power of a basis B_i, is directly proportional to the size of the conclusion set $C(\Sigma_i)$, directly proportional to the stringency of patterns which belong to the basis B_i, and inversely proportional to the number of argument patterns of B_i. I resume the basic notions in Figure 3.4[26].

Finally, we can say that the explanatory store over K, $E(K)$, contains arguments which are explanations of members of K. An important observation is required here. As pointed out by Kitcher, to say that the explanatory store over K is a function of K does not imply that the acceptance of K must be temporally prior to the adoption of $E(K)$. This is because the promises of explanatory power of a theory (for example Darwin's promises in the *Origins*) enter in the modification of our beliefs, or more precisely in the beliefs of a scientific community, thus modifying K [Kitcher, 1981, p. 519]. Now, if we come back to Kitcher's original question "When (under what conditions) does a derivation (an argument) explain why its conclusion is true?", we have that in order for a derivation (an argument) to be an acceptable explanation of its conclusion (relative to a belief corpus K), it must meet some specific conditions: a particular argument counts as an explanation only if it belongs to the set $E(K)$, i.e. the best systematization of K according to the conditions on unifying power. It is important to observe that Kitcher leaves open the possibility that, for some possible K, there could be indeterminacy in deciding how to weigh stringency, paucity of patterns and the number of conclusions, with consequent indeterminacy about $E(K)$ [Kitcher, 1989, p. 435].

[26]The diagram is basically that presented by Kitcher in his [Kitcher, 1981, p. 520].

3.2.2 Asymmetries, irrelevance and spurious unification

As Friedman before him, Kitcher also claims his model can solve three traditional difficulties the covering law model could not: the problem of asymmetry, that of irrelevance and the "problem of conjunction". Without giving all the details of his solutions here, I am going to illustrate the general strategy[27].

To exclude the wrong explanation both in the case of the "bad direction" in the asymmetry problem and in the irrelevance problem we need to show that the arguments we want to exclude are not in the explanatory store over K. According to Kitcher's model, when we admit the intuitively non-explanatory arguments what we are left with is a set of arguments whose basis has less unifying power than that of the basis for the set of arguments we normally accept as genuinely explanatory. Thus we rule out the wrong explanations which appear in one direction in the asymmetry problem and in the irrelevance problem. The criteria used for weighing the unifying power are those proposed by Kitcher and based on paucity of patterns used, size of conclusion set, stringency of patterns:

> Intuitively, the line of solution that we want toa dopt consists in showing that those who accept the wrong derivations are committed to accepting more patterns than they need, or to accepting less stringent patterns than they should, or to generating a more restricted set of consequences [Kitcher, 1989, p. 480-481]

Let's consider Kitcher's strategy for the case of the asymmetry problem. As we saw in the previous chapter, the asymmetry problem arises in the following situation: if A can be used to explain A', it is generally not the case that A' can be used to explain A. Van Fraassen's case of the tower and the shadow was precisely an illustration of this. Kitcher's challenge is to show

[27] The full technical story is too long and of no interest here. For a discussion of his solution see [Kitcher, 1981, p. 178-181] and [Kitcher, 1989, p. 482-488].

that his theory of explanation respects this asymmetry, i.e. that his account considers the shadow-based derivation of the tower height as not-explanatory.

Consider now e as some explanandum of type "X has length y". Let D_1 and D_2 two derivations of e, valid with respect to K and whose premises are in K. While D_1 is considered intuitively explanatory, D_2 violates asymmetry and we do not consider it as a genuine explanation (as in the case of the explanation of the height of the tower from that of its shadow length). Now, let P_1 and P_2 be, respectively, the patterns instantiated by D_1 and D_2. In the example of the tower and the shadow, call the pattern instantiated by the derivation P_1 *origin and development pattern of length explanation* (OD pattern), while the pattern P_2 instantiated by the derivation of the tower's height from the length of the shadow is called *shadow-based explanation pattern* (SBE pattern). Kitcher observes that the two patterns are not structurally identical. The OD pattern is more general and it derives all kinds of explananda of the type "Object O has length L"[28]. In Kitcher's words: the OD pattern is

> a general pattern of tracing the present dimensions to the conditions in which the object originated and the modifications that it has since undergone [Kitcher, 1989, p. 485]

For instance, in the case of our tower, the standard explanation of its length might refer to facts about the intentions of its architect, the process by which it was constructed and the tower's resulting propensity, once built, to remain rigid over time.

In order to account for the asymmetry, Kitcher points out that P_1 fares better than P_2 in terms of unifying power relative to K. By extending the analysis to the class of explananda E of the same type of e (and thus $e \in E$)[29],

[28]Kitcher's claim is that the class of all statements in K of the form "Object O has length L" (the class of explananda of object length) are derivable from statements in K describing facts about the origin and development of the object whose length is to be explained. All these derivations instantiate the unique origin and development pattern of length explanation (OD).

[29]Barnes has defined this way of proceeding from e to the class E "Kitcher's widening strategy" [Barnes, 1992, p. 562].

Kitcher observes that P_1 fares better than P_2 in generating derivations (acceptable relative to K) of the various members of E. In the case of the tower and the shadow, this amounts to saying that the OD pattern fares better than SBE pattern in generating derivations, acceptable relative to K, of object length. Why? The crucial observation is that the SBE pattern only works to generate conclusions about the length of a relatively small number of objects (the ones that actually possess shadows, i.e. the illuminated objects). Thus, if we incorporate the SBE pattern in the explanatory story $E(K)$, we have to include in $E(K)$ also the traditional OD pattern to successfully generate derivations of the length of *all* objects (including the objects which possess no shadow[30]). However, since one additional pattern has been included in $E(K)$ with *no* corresponding increase in the number of things which may be explained, the incorporation of both patterns would reduce the unifying power of $E(K)$ from what it was when it contained only the OD pattern[31]. For Kitcher this argument shows why, on his account, the shadow-based derivation of tower height fails to explain the tower height and then it must be rejected as non-explanatory.

In the previous section we saw how Michael Friedman tried to solve the problem avanced by Hempel and Oppenheim in their footnote 28. We called that problem the "problem of conjunction". Kitcher addressed the same problem, and he called the phenomenon of conjunction of laws which is at the origin of the difficulty the phenomenon of "spurious unification". Kitcher claimed that this phenomenon could be avoided in his account. The possibility of ruling spurious unification out is offered by the requirement that argument patterns should be stringent (we saw the notion of stringency before). A pattern of argument $\dfrac{\alpha \wedge B}{\alpha}$ used by a scientific theory B, where

[30]The length of some unilluminated, shadowless object O, for example, cannot be derived by an instantiation of the SBE pattern simply because O possesses no shadow from which its length may be derived

[31]This according to Kitcher's idea that the unifying power of $E(K)$ is inversely proportional to the number of patterns of derivation employed (criterion of paucity of patterns) and directly proportional to the number of the conclusions generated (size of the conclusion set).

α can be replaced by any sentence we accept, will satisfy our criterion of using few patterns of arguments as to generate many beliefs, but it will be unable to meet one of the two constraints set by the criterion of stringency of pattern: although arguments instantiating this pattern will have a similar logical form, they will not have similar non-logical vocabulary at similar places, since any kind of vocabulary can appear in the place of α. Therefore patterns as $\dfrac{\alpha \wedge B}{\alpha}$ or α should be excluded because as non-stringent[32].

3.2.3 Unification and scientific change

Is the previous discussion the end of the story? Unfortunately, no. We are faced with a very basic problem which has been put off until now. We assumed that there is a set K. However, as history goes on, our beliefs and our science change. Therefore the set K changes as a direct consequence of our change in beliefs. Furthermore, it is also reasonable to expect a change in the explanatory store $E(K)$. How Kitcher's model can account for these transitions?

Kitcher considers two contexts in which a methodological principle directing us to unify our beliefs can be expected to operate: a simple context, in which the corpus of beliefs and its language of formulation are fixed and the challenge is to see how the principle permits to select the best set of derivations that unify the corpus; and a more complex context, in which the corpus of beliefs and even the language in which our beliefs are formulated can change. I will present Kitcher's analysis for the two separate contexts. This is important for two different reasons. First of all, until now we have *not* considered situations in which the derivations of two (or more) systematizations are instantiated by two (or more) set of patterns (each of one is the basis for the considered systematization) with different virtues and for which the criteria of unifying power pull in different directions, neither we have

[32]For a complete discussion of the problem of spurious unification, including how to to rule out the artificial introduction of restrictions on the patterns in order to make them more stringent, see [Kitcher, 1981, p. 181-184].

explicitly discussed the notion of 'stringency of a pattern' by comparing two different patterns[33]. Second, a robust model of explanation must account for explanatory changes in scientific changes, thus permitting to justify the fact that our theories evolve and permit to explain new phenomena (or already known phenomena in a different way). In the following two subsections I am going to propose Kitcher's treatment of those points.

Fixed corpus, fixed language

First of all, let's focus on the first context, that in which we have to assess the merits of two different systematization S and S' of the same body of beliefs K (at some stage in the development of science). The systematizations are formulated in the same language L and all members of each set are acceptable relative to K. The principle adopted by Kitcher is the following:

(U) S should be chosen over S' as the explanatory store over K, $E(K)$, just in case S has greater unifying power with respect to K than S'.

How do we compare S and S'? Remember that Kitcher proposes that the assessment of unifying power is made by evaluating the paucity of patterns used, size of conclusion set, and stringency of patterns. In particular, for him the unifying power of a systematization Σ_i is inversely proportional to the number of argument patterns of B_i, directly proportional to the stringency of patterns which belong to the basis B_i and directly proportional to the size of the conclusion set $C(\Sigma_i)$. However, it is reasonable to expect the three criteria to pull in different directions. In this case, even if it is difficult to propose a comparison and make a tradeoff between two sets of rival patterns with different virtues (for instance, U and U', which instantiates the derivations in S and S'), Kitcher proposes the idea that we can always find an acceptable way to combine the virtues of those patterns in a new set of patterns (for

[33] In showing Kitcher's solution to the problem of Asymmetry, for instance, we have considered only the notion of consequence set and that of paucity of patterns. The notion of stringency of pattern was left quite free until now, and it has not been subject to any quantitative analysis.

instance, the set of patterns U^*, corresponding to a systematization S^* that combines the merits of the two systematizations S and S'). To capture this idea, he proposes a criterion (**O**), which tells us that U^* does at least as well as its rivals by the criteria of stringency and paucity of patterns, and that it does at least as well as generating consequences over K:

(**O**) Let U, U' be sets of patterns. Then there is a set of patterns U^* such that:

- O.a there is a one-one mapping from U^* to U, f, and a one-one mapping from U^* to U, f', such that for each pattern p in U^*, p is at least as stringent as $f(p)$ and at least as stringent as $f'(p)$; (one or both of f, f' may be injections rather than surjections)
- O.b let S, S', S^* be the sets of derivations that are the complete instantiations of U, U', and U^* with respect to K; then the consequence sets $C(S)$, $C(S')$, $C(S^*)$ are such that $C(S)$ and $C(S')$ are both subsets (not necessarily proper) of $C(S^*)$.

However, as Kitcher himself admits, this idea is quite optimistic[34]. For instance, consider the particular situation in which $f : U^* \to U$ is a surjection and $f' : U^* \to U'$ is an injection, with sets of patterns $U = \{\langle p\rangle_1, \langle p\rangle_2\}$, $U' = \{\langle p'\rangle_1, \langle p'\rangle_2, \langle p'\rangle_3, \langle p'\rangle_4\}$ and $U^* = \{\langle p^*\rangle_1, \langle p^*\rangle_2, \langle p^*\rangle_3\}$ (where p stands for 'pattern'). Figure 3.5 is an illustration of this. The new systematization will cover the derivations covered by S and S' by incorporating the virtues of both S and S'. To state that a systematization S^* is *always* possible to find is to say that it is always possible to find a new theory which instantiates a

[34]In his paper "Problems for Kitcher's Account of Explanation" [Sabatés, 1994], M. H. Sabatés has argued that the optimistic move proposed by (**O**) (i.e. for some pairs of S's there might be a combination $S*$ which synthesizes their virtues) seems implausible. He points to three major difficulties: "First, it will be hard (to say the least) to find a general criterion which gives us exactly those systematizations that in fact can be combined. Second, it will be even harder to find a rationale for such a general criterion (why must those systematizations which cannot be combined be ruled out? why can't one of them be the best systematization?). Third, once we have combined the virtues of different pairs of systematizations, how shall we find the best among the combinations (S, S', S''', ...) themselves. Should we maintain our optimism about combination to the very end? In sum, restricting the combination-candidates doesn't seem to lead us too far" [Sabatés, 1994, p. 281-282].

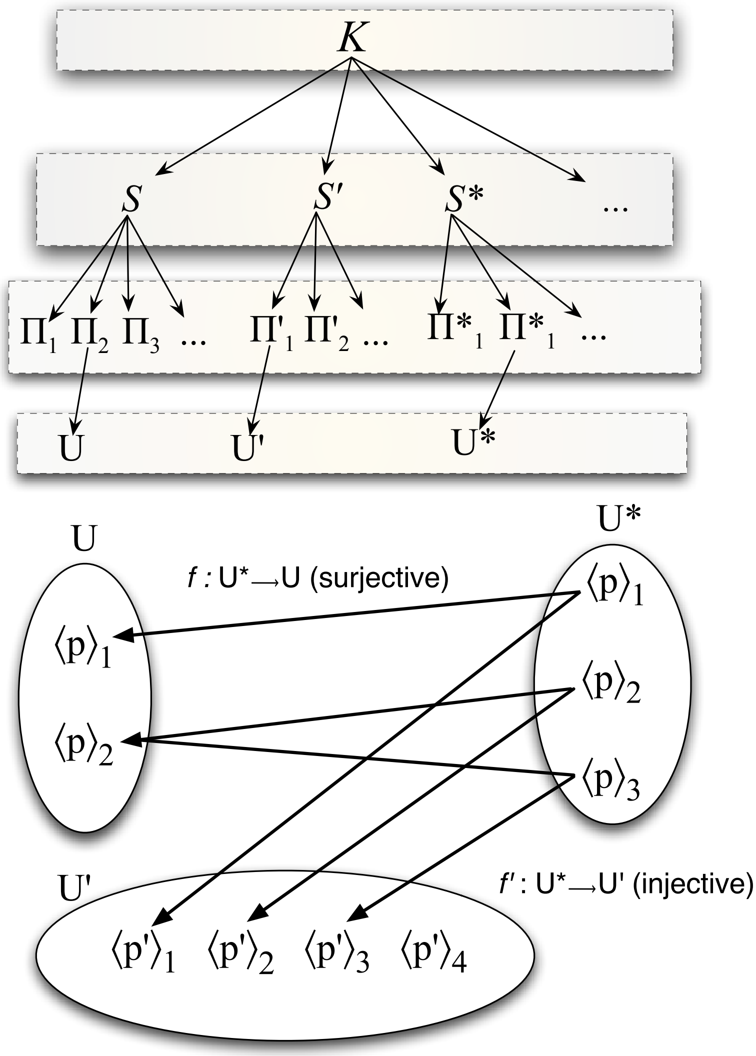

Figure 3.5: Optimism: The systematization S^*, whose basis is U^*, combines the merits of the two systematizations S and S'. Each pattern p^* in U^* is at least as stringent as $f(p^*)$ and at least as stringent as $f'(p^*)$, while the consequence sets $C(S)$ and $C(S')$ of S and S' are both subsets (not necessarily proper) of the consequence set of S^*, namely $C(S^*)$.

set of patterns which incorporates patterns included in U or U'. For instance, in the situation considered in Figure 3.5, the derivation intantiated in S' by pattern $\langle p' \rangle_4$ will be intantiated in S^* by some new pattern $\langle p' \rangle$. Nevertheless to have a new pattern $\langle p' \rangle$ capable to instantiate derivations which were instantiated by old patterns (for instance, $\langle p \rangle_1$ and $\langle p' \rangle_4$) amounts to saying that there exists a new theory from which the two patterns can be obtained. By presupposing that it is *always* possible to find such a new encompassing theory, Kitcher makes a quite optimistic asssertion.

Since the precedent condition (**O**) seems extremely optimistic, Kitcher's second task is to propose an explicit condition to compare the unifying power of sets of patterns (U and U'):

(**C**) Let U, U' be sets of patterns and S, S' their complete instantiations with respect to K. Then U has greater unifying power than U' if one (or both) of the following conditions is met:

C1 $C(S')$ is a subset of $C(S)$, possibly though not necessarily proper, and there is a one-one mapping f from S to S' such that for each pattern p in S, p is at least as stringent $f(p)$, and such that either f is an injection or f is a surjection and there is at least one pattern p in S such that p is more stringent than $f(p)$.

C2 $C(S')$ is a proper subset of $C(S)$ and there is a one-one mapping f from S to S' (either an injection or a surjection) such that for each p in S, p is at least as stringent as $f(p)$.

In particular, the conditions C1 and C2 apply in the following situations:

- (Cl) applies if S uses fewer or more stringent patterns to generate the same conclusions as S'.
- (C2) holds if S does equally well as S' by criteria of stringency and paucity of patterns and is able to generate a broader class of consequences.

Observe that the comparative relation introduced by (**C**) is both asymmetric and transitive. Thus it defines an order relation which permits to order sets of patterns with respect to unifying power. Furthermore, the comparison suggested by (**C**) is made on patterns which are assumed to have a similar structure. The latter requirement is introduced in order to escape further complications in the evaluation process.

Third, Kitcher finally comes to the notion of stringency of patterns, which is used freely both in (**O**) and (**C**). Again, as in condition (**C**), for simplicity Kitcher considers here the task of comparing pattern with a similar structure.

The relative stringency of two different patterns could be compared by using two different criteria. If we consider patterns which have a common classification c, we can say that a pattern is more stringent than the other if the schemata which correspond to the first pattern are subject to more rigourous demands on instantiation than the schemata in the second pattern. Call this criterion **T**:

(**T**) Let $\langle s, i\rangle$ be a pair whose first member is a schematic sentence and whose second member is a complete filling instruction for that sentence, and let $\langle s', i'\rangle$ be another such pair. Suppose that s and s' have a common logical form. Let g be the mapping that takes each nonlogical expression (or schematic letter) in s to the nonlogical expression (or schematic letter) in the corresponding place in s'. For any schematic letter t occurring in s, $\langle s, i\rangle$ is *tighter* than $\langle s', i'\rangle$ with respect to t' just in case the set of substitution instances that i allows for t is a proper subset of the set of substitution instances that i' allows for $g(t)$; $\langle s, i\rangle$ is at least as tight as $\langle s', i'\rangle$ with respect to t just in case the set of substitution instances that i allows for t is a subset of the set of substitution instances that i' allows for $g(t)$. $\langle s, i\rangle$ is tighter than $\langle s', i'\rangle$ just in case, (i) for every schematic letter occurring in$s$, $\langle s, i\rangle$ is at least as tight as $\langle s', i'\rangle$ with respect to that schematic letter, (ii) there is at least one schematic letter occurring in s with respect to which $\langle s, i\rangle$ is tighter than $\langle s', i'\rangle$ or there is a nonlogical expression e occurring in s such that $g(e)$ is a schematic letter, and (iii) for every schematic letter t occurring in s, $g(t)$ is a schematic letter. If only conditions (i) and (iii) are satisfied, then $\langle s, i\rangle$ is at least as tight as $\langle s', i'\rangle$.

Let p, p' be general argument patterns sharing the same classification. Let

$\langle p_1, p_n \rangle$ and $\langle p'_1, ..., p'_n \rangle$ be the sequence of schematic sentences and filling instructions belonging to p and p' respectively. Therefore p is more stringent than p' if for each $j(1 < j < ri)$ p_j is at least as tight as p'_j and there is a k such that p_k is tighter than p'_k.

However, we can have a second criterion of comparison (i.e. another way in which a pattern can be more stringent than another). This criterion operates when the inferential transition indicated by the classification of the first pattern refers to certain kinds of principle while the inferential transition indicated by the classification of the second patter consists of a precise linking of schematic premises. Thus, if the classification of the second pattern precludes possible instantiations which are left open by the classification of the first pattern, then the second pattern should be regarded as more stringent. The idea is captured in criterion **R**, which Kitcher formulates in the following way:

(**R**) Let p, p' be general argument patterns such that the sequence of schematic sentences and filling instructions of p is $\langle p_1, ..., p_n \rangle$ and the sequence of schematic sentences and filling instructions of p' is $\langle p_1, ..., p_r, q_1, ..., q_s, p_{r+1}, ..., p_r \rangle$. Suppose that the classifications differ only in that for p one or more of the p_{r+j} is to be obtained from previous members of the sequence by derivations involving some further principles of a general kind G, while for p' that (or those) p_{r+j} are to be obtained from the same earlier members of the sequence and from some of the q_k by specified inferential transitions. Suppose further that in each case of difference the set of subderivations allowed by p' is a subset of the set of subderivations allowed by p, and that in at least one case the relation is that of proper inclusion. Then p' is more stringent than p.

As Kitcher observes, the previous comparison based on demands of instantiantion is what is required when we are faced with a case of what he calls "explanatory extension", i.e. a situation in which one or more generalizations employed in a derivation are enriched by being treated by a different formal perspective. Kitcher gives an example of this kind of situation by considering a particular pattern of derivation in early nineteenth-century chemistry, and then showing how this pattern has been extended by subsequent work until the mid-twentieth century [Kitcher, 1989, p. 446]. The pattern considered, named "Dalton", answered why-questions of the form: "Why does one of the compounds between X and Y always contain X and Y in the weight ratio $m : nT$?". The Dalton-pattern is so constructed:

- (1) There is a compound Z between X and Y that has the atomic formula $X_p Y_q$.
- (2) The atomic weight of X is x; the atomic weight of Y is y.
- (3) The weight ratio of X to Y in Z is $px : qy$ $(= m : n)$.
- *Filling Instructions*: X,Y,Z are replaced by names of chemical substances; p,q are replaced by natural numerals; x,y are replaced by names of real numbers.
- *Classification*: (1) and (2) are premises, (3) is derived from (1) and (2).

With the successive development of post-Daltonian chemistry and atomic theory, the introduction of new concepts extended Dalton: the concept of valence and rules for assigning valences began to be used (although, at this first stage, the attributions of valence are unexplained and there is no understanding of why the constraints hold); with the successive introduction of the shell model of the atom chemists were given the possibility of explaining the results about valences. In particular, by appealling to the shell model of the atom and to principles about ionic and covalent bonding, scientists were able to derive instances of (1) and (2) from premises that characterize the composition of atoms in terms of protons, neutrons, and electrons. Finally, the derivations given at the stage in which the shell model was introduced have been showed to be embedded within quantum mechanical descriptions of the atoms. Furthermore, the shell structures and possibilities of bond-formation have been found to be consequences of the stability of quantum mechanical systems[35]. In this situation, the pattern which extends the original Dalton-pattern should be considered as more stringent according to **R**. This is because the classification of the extended pattern indicates a sequence of inferential transition which links in a more precise way the schematic premises and which precludes possible instantiations which were left open by the Dalton-pattern.

Naturally, it may happen that, by comparing two patterns, the first pattern turns out to be more stringent than the second pattern according to **T**, while the second pattern results as more stringent than the first when analyzed by **R**. In this case, again, Kitcher leaves open the (optimistic) possibility of combining the merits of two patterns in one new pattern.

To conclude, in the case in which the corpus of beliefs and the language are both fixed, by combining the criteria introduced (**C**, **T**, **R**), we are now able compare the merits of two different systematizations (say, S and S') of K: we use criterion **C** to choose the best generating set between the set of patterns chosen as generating sets for S and S' (where, in using criterion **C**, the comparison in stringency is given by **T** and **R**); to the best generating set there will correspond the preferred systematization between S and S'; according to **U**, to the preferred systematization there corresponds the set of derivations that best systematizes K in terms of unification – the explanatory store over K.

New corpus, new language

The second context the unification view should be confronted with is the context in which we have the possibility of scientific change. This situation

[35]Observe that, by using the idea of argument pattern (together with an extension of it), the notion of explanatory extension does not necessarily require reduction (derivability of laws of the reduced theory from laws of the reducing theory). This is because it might be possible that some of the concepts of the extended theory cannot be formulated in terms of the concepts of the extending theory. Therefore explanatory extension is considered by Kitcher as primary to the notion of reduction in the analysis of the relations among successive theories and in the development of a philosophical account whose aim is to capture the idea of accumulation of knowledge [Kitcher, 1989, p. 448].

appears more realistic and also more interesting, especially if we observe that we often justify transitions in science by appealling to the fact that a new corpus of beliefs (and, eventually, a new language) has greater explanatory power than an older one.

Consider science as a sequence of practices, each one distinguished by a language L, a corpus of belief K formulated in that language and a store of explanatory derivations $E(K)$. Therefore we can indicate a scientific practice with the triple $< L, K, E(K) >$. The challenge for the unification account is the following: to account for the gain in explanatory power in the transition from a scientific practice $< L, K, E(K) >$ to a new scientific practice $< L', K', E(K') >$ in terms of greater unification in our beliefs. As natural, Kitcher's request is that such an account should allow transitions in science without damaging his solution to the problem of asymmetry and irrelevance. Here we are confronted with two different and separate tasks:

(α) to specify the conditions under which a systematization Σ' of K' provides a better unification of K' than Σ of K does of K.

(β) given the fact that the best systematization $E'(K')$ of K' provides a better unification of K' than the best systematization $E(K)$ of K does of K, to say when this accounts for the transition from $< L, K, E(K) >$ to $< L', K', E(K') >$.

Here Kitcher does not offer a general account, but only partial conditions which we can identify in the history of science and which seem to underlie some transitions in history of science. On the other hand, he suggests that the search for unification of beliefs is subject on some kinds of principles that govern the modification of language and that rule on the acceptability of the proposed beliefs.

By considering the first task (α), Kitcher restricts his analysis to the case in which the shift from $< L, K, E(K) >$ to $< L', K', E(K') >$ involves *no* explanatory loss. Assume the validity of criteria **U**, **C T**, **R** seen above about the relative stringency between patterns. By using these criteria we can find the unique and best systematizations of K and of K', which are respectively $E(K)$ and $E(K')$. The condition of no explanatory loss is stated as follows:

> [No Explanatory Loss Condition] For any statement that occurs as a conclusion of a derivation in $E(K)$ there is an extensionally isomorphic statement that occurs as a conclusion of a derivation in $E(K')$. Two statements are extensionally isomorphic just in case they have the same logical form and the nonlogical expressions at corresponding places refer to the same entities (objects in the case of names and sets in the case of predicates). [Kitcher, 1989, p. 489]

This condition (which presupposes realism: objects exist independently from theories[36]) is introduced in order to permit, in a shift from L to L', the kinds of change ("refixing of reference") which, according to Kitcher, underlie the phenomenon of incommesurability championed by authors such as Kuhn and Feyerabend. According to the doctrine (conceptual relativism) of those authors, for two languages used in a same scientific field at different times (times separated by large-scale changes or so called "revolutions") there are expressions in each language whose referents are not specifiable and translatable in the other language (we can't formulate – translate – past theories in the language of a modern theory)[37]. Thus the impossibility of comparing concepts belonging to those theories (concepts, and then theories, are "incommensurable")[38]. Naturally, if we agree with this version of the

[36]The following quote is quite indicative of Kitcher's realism about objects as expressed in the previous condition of no explanatory loss: "Trivially, there are just the entities there are. When we succeed in talking about anything at all, these entities are the things we talk about, even though our ways of talking about them may be radically different. However variable the connections we draw among its constituents, the world supplies a common content for our references" [Kitcher, 1978, p. 547]. However, note that in another passage Kitcher claims that his account is not committed to any metaphysical view: "I have been trying to show that we can make sense of scientific explanation and our view of the causal structure of nature without indulging in the metaphysics" [Kitcher, 1989, p. 500].

[37]For a defense of conceptual relativism in terms of referential change see [Feyerabend, 1965, especially p. 270-274], [Feyerabend, 1981], and [Kuhn, 1970].

[38]As observed by Kitcher, referential change is neither necessary nor a sufficient condition for conceptual relativism, for conceptual relativism can occur if the languages involved contain completely different expressions (and then no referential change is involved) [Kitcher, 1978, p. 521].

incommesurability thesis, then there would be no way to study any rational transition (in Kitcher's sense), for it might be that the referents of their terms cannot be compared. On the contrary, Kitcher claims that we can allow for conceptual revision without conceptual discontinuity [Kitcher, 1978, p. 544]. The changes he calls "refixings of reference" allow for a redescription of a phenomenon (explained in a old and abandoned theory, as Priestley's phlogiston theory) in a new language, hence discarding Kuhnian incommensurability. The notion of extensional isomorphism captures this continuity[39].

If the condition of no explanatory loss is satisfied, then we can justify the explanatory gain in the transition from $< L, K, E(K) >$ to $< L', K', E(K') >$ in terms of explanatory unification. This could be made by looking at the intuitive fact that we gain in explanatory power when the same number of equally stringent patterns generate more consequences, or fewer or more stringent patterns generate the same consequences. Thus, under the assumption of no explanatory loss, the idea that $E(K')$ unifies K' better than $E(K)$ unifies K is formulated by Kitcher in the following way:

> **C'** Suppose there is a one-one mapping f from the basis of $E(K')$ to the basis of $E(K)$ such that for each p in the basis of $E(K')$ p is at least as stringent as $f(p)$; and (i) f is an injection, or (ii) there is some p in the basis of $E(K')$ such that p is more stringent than $f(p)$, or (iii) there is some statement in the consequence set of $E(K')$ that is not extensionally isomorphic to any statement in the concequence set of $E(K)$. Then $E(K')$ unifies K' better than $E(K)$ unifies K [Kitcher, 1989, p. 490]

Observe that **C'** can be seen as a sort of generalization of criterion **C** seen above: while criterion **C** was used to compare the virtues of two basis in a same set of beliefs K, **C'** is designed by Kitcher to compare the merits

[39]For a detailed discussion of Kitcher's account of conceptual change in science and some examples where later theories *refine* concepts of earlier theories see [Kitcher, 1978], [Kitcher, 1982] and [Kitcher, 1984, especially p. 152 and p. 165-170]. What is of interest to our discussion is that this change is possible and is captured through Kitcher's notion of extensional isomorphism.

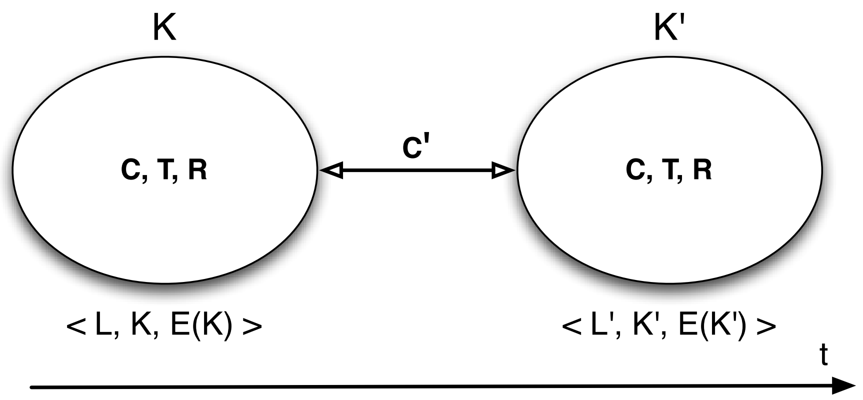

Figure 3.6: Criteria **C**, **T** and **R** operate within a set of beliefs, while criterion **C'** permits to compare sets of patterns belonging to distinct sets of beliefs (K and K').

of two basis belonging to two distinct sets of beliefs K and K' (Figure 3.6).

Nevertheless, in order to use **C'**, we need a sense of the notion of relative stringency to compare patterns which are formulated in different languages[40]. According to Kitcher this could be obtained by observing that the old patterns of explanation (i.e. the patterns of explanation in $E(K)$) are isomorphic to subpatterns of the new patterns of $E(K')$. Unification is achieved because the new patterns can be partially instantiated in different ways to generate patterns that were previously viewed as belonging to different fields.

As an illustration of a situation in which the comparison of patterns accords with the idea sketched above (previous paragraph), Kitcher refers to the case of Maxwellian electromagnetic theory and Darwinian theory. In both cases the theory came with a new language and new theoretical claims, which introduced explanatory advantages in science. For instance, in the first case

[40]The strategy followed here by Kitcher is the same as in the case of fixed context and fixed language: introduce a criterion (**C**, and **C'** for the present contex) to compare set of patterns, and finally propose some criteria for comparing the relative stringency of patterns. The only difference is that in the present case we are considering the comparison between sets of patters (and patterns) which belong to two different set of beliefs. Moreover, each set of beliefs is formulated in a specific language. The assumption of no explanatory loss -which involves the notion of extensional isomorphism- is what permits the present comparison.

(Maxwellian electromagnetism), the new theory provided a variety of patterns which were able to generate explanatory derivations within theories which were considered as distinct (and independent) before Maxwell. The patterns which intantiated these derivations in Maxwell electromagnetism are, according to Kitcher, isomorphic to the old (and distinct) patterns of explanation coming from the various theories:

> Consider Maxwellian electromagnetic theory. This supplies a variety of patterns for generating explanatory derivations within geometrical optics, the theory of diffraction, electrostatic interactions, and so forth. The old patterns of explanation are isomorphic to subpatterns of the new patterns, and unification is achieved because the same underlying pattern can be partially instantiated in different ways to generate patterns that were previously viewed as belonging to different fields [Kitcher, 1989, p. 490]

This is why Kitcher proposes to regulate the notion of stringency in **C'** (in a comparison of patterns which are formulated in different languages!) through the notion of extensional isomorphism between patterns (as his example of Maxwell electromagnetism inuitively suggests)[41]:

> **P'** I shall therefore propose that **C'** can be satisfied by meeting condition (i) if there is one (or more) pattern of p of $E(K')$ such that there are at least two patterns of $E(K)$ that are extensionally isomorphic to subpatterns of p and if all other patterns of $E(K')$ –that is patterns that are not partially instantiable by isomorphs of patterns of $E(K)$– are themselves extensionally isomorphic to patterns of $E(K)$. It seems to me that many episodes from the history of science will require comparisons of unyfing power that are based on more subtle conceptions of relative stringency, but I shall not try to pursue this difficult issue further here. [Kitcher, 1989, p. 490]

[41] Although Kitcher does not propose the following criterion under the form of definition, to simplify the discussion I present the criterion as **P'**.

Let's now focus on task (β), i.e. to see when appeals to explanatory unification (the fact that $E'(K')$ provides a better unification of K' than $E(K)$ does of K) justify the transition from $< L, K, E(K) >$ to $< L', K', E(K') >$.

According to Kitcher, in order for the transition to be guaranteed by explanatory unification, the condition is that the shifts from K to K' and from L to L' be defensible. For instance, in the shift from beliefs K to beliefs K', this might mean that there are no strong arguments from the perspective of $< L, K, E(K) >$ against the addition to K of some statements (K plus this set will result in K'). If we have this kind of modification we say that K is *neutral* towards the changes. More precisely, K is neutral towards the changes when

> a modification of K to K' may involve the additions of statements for which there was previously no positive evidence but which were not precluded by strong arguments from well established principles of K (or conversely, such modification may involve abandoning statements in such a way that the prior view that there was evidence in favor of such statements is explained as illusory) [Kitcher, 1989, p. 491].

We say that K is *negative* towards the changes when there are arguments using premises that are common both to K and to K' either against statements that would be added or in favour of statements that would be dropped. In the case K is neutral towards the changes, the fact that the new corpus would permit greater unification of belief justifies the transition, while in the case K is negative towards the changes the condition for the appeal to greater unification is not satisfied.

The theory of evolution by natural selection is an example of the former case, while the reception of the theory of continental drift in the 1920s and 1930s is an example of the latter[42]. In the second case, the theory proposed by A. L. Wegener offered an interesting perspective to "unify" geological, biogeographical, and paleometeorological beliefs. Unfortunately, the geologists

[42]See [Kitcher, 1989, p. 491-492] for a discussion of the two examples.

had strong arguments against the fact that continents could move. Hence, the shift from beliefs K to beliefs K' was *not* defensible and to appeal to explanatory unification (the fact that $E'(K')$ provides a better unification of K' than $E(K)$ does of K) did not justify the transition (the theory was not accepted in the 1920s and 1930s). To come to the first example, Darwin's arguments in the *Origins* were similar to the theory of continental drifts in the promise of increased unification. However, the strategy through which Darwin proposed his arguments was intended to eliminate the accepted forms of reasoning which would have been considered incompatible with his theory. In other words, he provided a sufficient demostration that the shift from K to K' was defensible, and then the proviso for the appeal to greater unification was satisfied.

Now, if we consider the second shift, i.e. that of language (from L to L'), an analogous proviso could be given in order for the modification of language to be defensible: L should be neutral towards the projectability of predicates, i.e. in the transition some predicate already viewed as projectable from the standpoint of L is taken to cover phenomena which have been regarded as separate because previous scientists had seen nothing in common between them. An example is given by Maxwell's electromagnetic theory. With the birth of this theory, phenomena in the propagation of light and phenomena involving electromagnetic effects, which had been considered unconnected, were subsumed under (projectible) predicates such as "transverse wave propagated with velocity c".

After this illustration of Kitcher's theory, I am going to present some criticisms which have been leveled against this model.

3.3 Is unification enough?

Criticisms against Kitcher's unification approach to explanation are various and attack the model on different levels (metaphysical, methodological, epistemological)[43]. With regard to the general intuition which stays behind the unification project, one famous objection is that of David Lewis[44]. I will start this section by illustrating this criticism and Kitcher's defense. As it will be clear in a moment, Lewis' objection will introduce some metaphysical considerations. In line with what I have done in the first chapter, I will identify each criticism by a label: for Lewis' criticism I will use **DL**, for Margaret Morrison's **MM**, for Paul Humphreys' **PH**, for Eric Barnes' **EB**, for Jamie Tappenden's **JT**, for Johannes Hafner and Paolo Mancosu's **MH**.

(**DL**) In illustrating his account, Kitcher proposes some criteria to say when a derivation should be considered an *acceptable* explanation of its conclusion (relative to a set of beliefs). However, the partisans of the causal approach to explanation (such as Lewis) clearly distinguish between an acceptable explanation and a *correct* (or *true*) explanation. For them, an acceptable explanation is an explanation which identifies what would be rational (for a person who has a particular set of beliefs) to take as the causal structure underlying the explanandum phenomenon. On the other hand, a correct explanation is an explanation that identifies the causal structure underlying the phenomenon[45]. Now, although Kitcher does recognize the importance of causal explanations, he does not focus extensively on the distinction between a correct (or true) and an acceptable explanation. This creates a tension with the partisans of the causal view on explanation. Furthermore, this tension sparks criticisms such as that of David Lewis.

Lewis' argument can be summed up in the following way: unification's partisans consider as explanatory relevant those factors that figure in a unifed treatment of phenomena; however, if the world is messy and confused (why not?), then factors which are causally relevant to phenomena may be a mix-

[43]From now on I will refer to the "unification account" in Kitcher's sense. Whether I will discuss Friedman's model, I will explicitly refer to that account.

[44]Lewis' criticism is reported by Kitcher in his [Kitcher, 1989, p. 494-497], while a reply to the objection is given by the author in [Kitcher, 1989, p. 498-499].

[45]Naturally, behind this distinction there is a particular metaphysical view. Observe how the causal view on explanation comes with a strong form of realism: world has a structure and a true explanation identifies the causal structure which lies under the phenomena.

ture of disconnected facts; hence the unification model of explanation (which seems to assumes a priori that our nature is not such a messy place) would not reveal ('trace') the causal structure of phenomena and would even reject causal explanations because not acceptable (in other words: unification would not be crucial to explanation). If the world is not unified, then the unification account is attributing to nature an *a priori* structure that nature *may not* have. A similar observation about the possibility of attributing the world a unified structure it might not possess has been proposed by Peter Railton: "If unification provides a criterion of explanation, and if explanation is evidence of truth, then unification is evidence of truth. Yet how does the realist know a priori that the world we inhabit is a unified one?" [Railton, 1989, p. 228]. Observe that both Lewis' and Railton's arguments assume that the unification account aims to provide a *true* explanation. Although Kitcher is more focused on acceptable explanations, the tension here is acute because Kitcher accepts the existence of "true causal statements" [Kitcher, 1989, p. 494]. How then does Kitcher reject this argument, and more particularly Lewis'? He tries to offer a characterization of correct explanations by taking as metaphysical assumption that causal truths are *not* independent of our search for order in the phenomena [Kitcher, 1989, p. 497]. This marks an evident dividing line with causal partisans who, on the contrary, believe in this kind of independence[46]. Furthermore, to endorse this metaphysical view does not commit Kitcher to give a justification (from the point of view of unification) or an identification of the causal structure of the world which is independent of our search for order in the phenomena (this is what Lewis seems to presuppose in his argument against Kitcher, i.e. the fact central to explanation is the identification of the causal structure of the world). As Kitcher puts it:

> [...] on the version of the epistemic conception developed here (the

[46]For instance, by claiming that there exists a number of fundamental causal mechanisms in the world. According to Salmon, "to explain a particular occurrence is to show how it fits into the causal network of the world" [Salmon, 1984a, p. 276]. The causal network of the world exist independently of our way to identify them in explanations.

unification approach) unification is *constitutive* of explanation, while on Salmon's version of the causal approach, unification is at best a *contingent concomitant* of the tracing of causal structure. The heart of the unification approach is that we cannot make sense of the notion of a basic mechanism apart from the idea of a systematization of the world in which as many consequences as possible are traced to the action of as small a number of basic mechanisms as possible. In short, on the unification approach, the basic mechanisms must be those picked out in the best unifying systematization of our best beliefs, for if they were not so picked out then they would not be basic. [Kitcher, 1989, p. 497]

We have already seen as Kitcher proposes to conceive science as a sequence of practices, where each practice is distinguished by a language, a body of belief, and a store of explanatory derivations. His strategy is to define what counts as a correct explanation in terms of modification of practices in science. More precisely, he considers that correctness is obtained in the ideal long run when the principles of rational modification (a *static* principle of unification which operates in a body of belief and permits to obtain the explanatory store for that a body of belief K; plus a *dynamic* principle, subject to the condition seen in the previous section, that directs us to modify practice so as to achieve advances in unification) are followed:

> [...] true statements are those that belong to the belief corpus of scientific practice in the limit of its development under the principles of rational transition. Finally, and most important for present purposes, *correct* explanations are those derivations that appear in the explanatory store in the limit of the rational development of scientific practice. [Kitcher, 1989, p. 497-498]

With the previous considerations in his hands Kitcher observes that he is now able to face Lewis' criticism. In particular, retaining the connection between explanatory relevance and causal relevance (i.e. "If F is causally relevant to P then F is explanatorily relevant to P"), his defense from Lewis' objection consists in providing grounds for rejecting the following claim:

> It is possible (and may, for all we know, be true) that there is a factor F that is causally relevant to some phenomenon P such that no derivation occurring in the explanatory store in the limit of scientific practice derives a description of P from premises that make reference to F.

The grounds adduced to reject this claim are that the notion of causal relevance depends on the notion of explanatory relevance, but what is explanatorily relevant is what figures in the systematization of belief in the limit of scientific inquiry (as guided by the search for unification). In other words: in the limit of our attempts to systematize our beliefs (and to achieve a unified view of the world), what is explanatorily relevant will emerge and therefore also all basic causal explanatory mechanics will be captured in the same limit[47].

> The growth of science is driven in part by the desire for explanation, and to explain is to fit the phenomena into a unified picture insofar as we can. What emerges in the limit of this process is nothing less than the causal structure of the world [Kitcher, 1989, p. 500]

(**MM**) A second general line of criticism is put forward by Margaret Morrison. Although she agrees on the primary role of mathematics in attaining unity ([Morrison, 2002, p. 247] and [Morrison, 2000]), she claims that unification has little if anything to do with explanation. Explanation and unification are different (and sometimes conflicting) businesses, they pull in different directions. The point is stated many times throughout Morrison's book *Unifying Scientific Theories*:

> Rather than analysing unification as a special case of explanatory power, as in commonly done in the literature, I claim that they frequently have little to do with each other and in many cases are actually at odds [Morrison, 2000, p. 2]

[47]By observing that "there is no a priori guarantee of how successful we shall be in achieving unification", Kitcher does not propose any particular degree of unification that this limit must reach [Kitcher, 1989, p. 499].

> [...] explanation and unification may not be as closely related as has typically been thought; unity is possible without a satisfactory level of explanatory power [Morrison, 2000, p. 4]

> Unity and explanatory power are different and often conflicting goals [Morrison, 2000, p. 34]

In her discussion of Kitcher's example of Darwinian theory, Morrison points out that the unification permitted by natural selection is not able to account for its explanatory power ("I do not want to deny that natural selection can have explanatory power; rather, my claim is that its explanatory power cannot be understood in terms of its unifying power" [Morrison, 2000, p. 201]). Thus, contrary to what is claimed by Kitcher, we can have cases of unification without explanatoriness. In addition to this, Morrison underlines that unification and explanatory power could be *conflicting* goals. This claim is justified by observing that unification is facilitated by generality and abstraction, but in this process the details which provide us with explanatoriness, i.e. the details which should give us the explanation of the physical dynamics of the unified theory, are sacrified [Morrison, 2000, p. 5]. For instance, Morrison shows how, in the case of Maxwell's electrodynamics, the initial encompassment of electromagnetism and optics was obtained by Maxwell by demonstrating that the velocity of electromagnetic waves travelling through a material medium (aether) was equal to the velocity of light v. However, later versions of the theory did not rely on the aethereal medium and derived the velocity v from the field equations formulated in the Lagrangian formalism. Hence, at this mature stage (the stage of the field-theoretic description), no explanation is given of *how* electromagnetic waves are propagated through space. Therefore, while the unification is obtained through abstraction and generality by mathematizating the phenomena using the mathematical Lagrangian apparatus, the explanatory power is lost in this process. A similar example is given by Morrison for the case of the unification of terrestrial and celestial phenomena made by Newton's *Principia*. In this case, it is a well-known fact that with the *Principia* there was a move away from explanations of planetary motions in terms of mechanical causes[48]. Thus Morrison observes that:

> Of course, the inverse-square law of gravitational attraction explains why the planets move in the way they do, but there is no explanation of how this gravitational force acts on bodies (how it is transported), nor is there any account of its causal properties. [Morrison, 2000, p. 4]

However, it seems to me that Morrison's criticism lacks of power here. The kind of (acceptable) explanations which are linked to the unification process in Kitcher does not necessarily have to be causal (i.e. mechanical). In Kitcher's sense, what is important is that the shift from the mechanical to the new mathematical explanation (inverse-square law) produces "good" unification in terms of patterns. In the new framework of the Newtonian theory, it could make no sense to ask *how* the gravitational force is transported. The right question (in Kitcher's sense) is: does the Newtonian general pattern provide us with the correct unification? What is important is that the derivation satisfies certain unification criteria, which are stated in Kitcher's model. Morrison's objection seems to be based on the view that an acceptable explanation must provide the causal history, while Kitcher explicitly denies that causal explanations are primary to theoretical explanation and that we obtain explicit knowledge of causal dependency in his account:

> Thus the picture advanced by the unification approach shows the concept of causal dependence as derivative from that of explanatory dependence, but it does not promote the dubious idea that each of us gain explicit knowledge of causal dependencies through recognition of the structure of the explanatory store [Kitcher, 1989, p. 436]

(**PH**) By offering one example concerning elementary logic, Paul Humphreys has tried to attack one core idea of the unification account and show that

[48] In the Introduction I have remarked how Yves Gingras considers this explanatory shift as an 'epistemic effect' of the process of mathematization started with Newton [Gingras, 2001].

the connection between increased understanding and a process of abstraction (a process of abstraction such as the process of unification, i.e., the process of seeing common patterns underlying different phenomena) is more complicated and could not be completely rendered by Kitcher's account in terms of patterns of unification [Humphreys, 1993].

Take two different axiomatizations of classical propositional logic L and L' which differ only in their axiom schemata. According to Kitcher's rules of comparing unifying power of patterns, a pattern instantiated by L is supposed to provide greater unifying power than that instantiated by L'. Moreover, in Kitcher's general framework, this means that L is also supposed to provide better *understanding* than that given by L'. Nevertheless, Humphreys remarks, to the eyes of logicians the understanding of the formal system provided by the two axiomatizations is exactly the same, and therefore "it is obvious that the single argument pattern of L does not give us this [understanding]" [Humphreys, 1993, p. 187]. Here the criticism put forward by Humphreys points to a more general level:

> Understanding is an epistemological concept. Unless we can have some, perhaps imperfect, epistemic access to the ideal agent's criteria for comparative understanding, we have no grounds for assessing the correctness of the ordering relations imposed on understanding by appeal to the ideal agent. And in the present case, it is scarcely credible that by extrapolation from your present epistemic state, your understanding of propositional logic through L will surpass (and not just equal) your understanding of it through L'. [Humphreys, 1993, p. 187]

In passing, let me report here that the linkage unification-understanding has been criticized by various authors[49]. By denying the fact that explanatory

[49]For instance, Jaegwon Kim: "But what does the unification approach tell us about explanation and understanding? Surprisingly little, I think. [...] is it so clear that the concept of unification, whether taken pre-analytically or in either of the senses explicated by Friedman and Kitcher, is any closer to understanding?" [Kim, 1994, p. 65].

power follows from unification, also Margaret Morrison rejects that unification is a key to understanding on the basis that the mathematical part of a MEPP confers the explanation a high contingent character, thus reducing the explanatory power and our potential of understanding of the phenomena involved[50]:

> Often an identification of a phenomenon with a particular mathematical characterization is highly contingent, and the generality of such frameworks is such that they provide no unique or detailed understanding of the physical systems that they represent. That is to say, we can predict the motions of the phenomena from dynamical principles, but we have no understanding of the causes of motion. [Morrison, 2000, p. 30]

(**EB**) While Humphreys focused on the linkage understanding-unification, Eric Barnes has proposed examples where Kitcher's model is unable to account for the asymmetric structure of the explanation [Barnes, 1992][51]. In particular, Barnes offers examples of maximally unifying but nonexplanatory argument patterns, i.e. the sort of patterns that Kitcher denies to exist. In one of his examples Barnes considers a closed Newtonian system S whose laws are temporally symmetric. This example is important because the case of Newtonian mechanics is exactly the case considered by Kitcher in his example of argument pattern.

Barnes considers the class of explananda E as the set of all statements in K of the form "Object O in system S has position P and velocity v at time t". We are then in the following situation: we have at least one complete description D of the system at some time t in K; K is deductively closed; Newton's laws L are in K. With the foregoing stipulations, the class of explananda E will contain a complete set of descriptions of the position and

[50] Again, it seems to me that Morrison is wrong in trying to criticize Kitcher's approach by requiring as good understanding that understanding which emerges from causal factors.

[51] Remember that the ability for Kitcher's model to solve the asymmetry problem and overcome this obstacle was considered by Kitcher as one of the fundamental advantages of the unification account [Kitcher, 1989, p. 487-488].

velocity of all objects in S for each moment (past, present and future) in the history of the system. Those derivations will be the result of combining D together with Newton's laws L. The classical kind of Newtonian pattern (the same considered by Kitcher) from which we deduce a statements e in E is called by Barnes the 'Newtonian predictive pattern'. Nevertheless, as Barnes correctly observes, in our system S the fact that Newton's laws are temporally symmetric permits the introduction of another symmetric pattern, called the 'Newtonian retrodictive pattern', which permits to retrodict the same explanandum e (in both the two patterns we derive the particular position and velocity of an object O, namely our explanandum, starting from a state of the system –initial conditions– plus Newton's laws). The only structural difference between the predictive and the retrodictive pattern will be in the temporal reference which appears in the filling instructions. For instance, if we consider a moving particle, by using Newton's dynamic equations plus *later* states of the system (say, at time t'') we can derive conclusions about any earlier states (say, for example, at time t', where $t' < t''$). In this case, in which we consider as explanandum an event at time t', in the set of filling instructions the dummy letter θ would be replaced by an explicit function of time t'. To give a concrete illustration of retrodictive pattern, let me take the example of moving projectile I considered in subsection 3.2.1, where I offered a concrete example of the Newtonian pattern.

In that case, we started with initial conditions t_0 and velocity v_0 and we showed how from the Newtonian pattern we are able to predict a future state of the system (explanandum e), at time $t' > t_0$. However, suppose we would have known the state (position and velocity) of the system at a time $t'' > t'$ (on the temporal axis: $t_0 < t' < t''$). In that case, using the same derivation, we could have used the same pattern to infer that the projectile had a particular state at time t' *before* t''. The explanandum, i.e. the position of the projectile at time t', is the same. The only difference in the derivation would have been in the filling instruction set of the retrodictive pattern, where θ would have been replaced by '$\mathbf{V_0}t' + \frac{1}{2}\mathbf{g}t'^2$'. We could then have used the retrodictive pattern to derive the projectile position at time t' by using a later configuration at time t''.

Now, it is then easy to see what's wrong with this latter derivation. The problem comes from the fact that, while the Newtonian predictive pattern is intuitively explanatory of the members of E (such as e), the Newtonian retrodictive pattern is not. And this is because to explain some explanandum e by using facts that occurred subsequent to it is, intuitively, a not genuine explanation. However, the patterns are identical with respect to their unifying power (to make a substitution in the filling set does not affect unifying power, at least according to the criteria proposed by Kitcher). This is exactly the problem of asymmetry. Thus, if the two patterns have equal unifying power under Kitcher's model but one of them is (intuitively) nonexplanatory, it seems that we have to give up or modify the linkage proposed between unification and explanation[52].

> Clearly, if two argument patterns [the Newtonian predictive and the Newtonian retrodictive patterns] may be of equal unifying power but are such that one pattern is explanatory and the other not explanatory, then we cannot in general settle the question of a pattern's explanatory force on the basis of considerations about its unifying power. [Barnes, 1992, p. 565-566]

In general, according to Barnes, Kitcher's 'widening strategy' (presented in subsection 3.2.2) will not able to "save the wrong explanation" (and thus

[52]In his paper "How the Unification Theory of Explanation Escapes Asymmetry Problems" [Jones, 1995], T. Jones offers a potential way to escape Barnes' criticism and permit Kitcher's model to avoid the wrong retrodictive explanation. In particular, Jones adduces a very simple argument to save Kitcher from Barnes' objection. He points out that although the two patterns (Newtonian predictive and retrodictive patterns) can generate the same conclusion, when we use the retrodictive pattern we are adding a new type of pattern to our systematization. Nevertheless this pattern would give us the very same conclusion that we get from a commonly used family of patterns, i.e. the family of patterns that are used to explain something by describing its forms/condition at the time of origin and how subsequent forces have altered that condition to produce a present or future state (among them figures the Newtonian predictive pattern). As a consequence, if the unifying power is inversely proportional to the number of patterns used (criterion of paucity of patterns), to add this new pattern would decrease the unifying power of our systematization and this is why the retrodictive 'backward' pattern should be rejected as not explanatory.

the problem of asymmetry), because the character of temporal symmetry of Newton's laws L will make the wider classes E_1, E_2, E_3, ... always coverable by both the predictive and the retrodictive pattern. In other words, the widening strategy will not allow to eliminate the "wrong" explanation as nonexplanatory, and thus account for the intuitive asymmetry.

Barnes' criticism is twofold. It also attacks Kitcher on the epistemological level. Kitcher considers "the concept of causal dependence as derivative from that of explanatory dependence" [Kitcher, 1989, p. 436]. Thus, the fact that predictive and retrodictive patterns have, in the unification model, the same explanatory force, entails in Kitcher's antirealist account of causation that the symmetry is transmitted also to the causal relation: there is *no* causal asymmetry between earlier and later states of a Newtonian system. In other words, the causal relation between earlier and later states of the system is symmetric, i.e., state 1 causes state 2 and state 2 causes state 1, which is an obvious counterintuitive claim. This is why, against Kitcher's claim that the causal structure of the world derives from explanatory (unificatory) stories, for Barnes the causal dimension must be considered primary to scientific explanation. However, while Barnes claims for the essential role of causation in *any* account of explanation, Kitcher's remarks about the existence of non-causal explanations in such domains as formal syntax or mathematics [Kitcher, 1989, p. 422-428] underline the fact that is very likely that causation could be only *part* of the story about explanation.

While the previous citicisms concern general aspects of Kitcher's account (impossibility of having a unified structure for the world, disjunction between explanation and unification, impossibility for Kitcher of solving the traditional problem of asymmetry, non-linkage unification-understanding), the last two criticisms presented here, i.e. that of Jamie Tappenden in his [Tappenden, 2005] and that of Johannes Hafner and Paolo Mancosu in their [Hafner *et al.*, 2008], point to some technical difficulties of the unification approach. Both criticisms deal with real cases of mathematical explanation within mathematics, and both give primary importance to the practice of mathematics in their discussions. The latter is the reason why those authors claim for a bottom-up methodology, i.e. they consider that the philosophical investigation of the notion of explanation should start by considering real cases from the mathematical practice. Moreover, as we will see below, they agree on demanding Kitcher's theory to introduce some qualitative reinforcement in its apparatus.

(**JT**) Jamie Tappenden gives a general discussion of the unification approach for cases of mathematical explanation within mathematics [Tappenden, 2005]. He points out that existing accounts of unification are more balanced on quantitative restrictions (for instance, the quantity of patterns in Kitcher's formulation) and need to be supplemented with qualitive reinforcements. His idea is that only such qualitative reinforcements would permit the unification model to reflect the actual mathematical practice. This idea occurs many times in his paper. Just to quote a significant passage:

> Though unification accounts have a grain of truth (since a phenomenon (or cluster of phenomena) called "unification" is in fact important in many cases) we are far from an analysis of what "unification" is. In particular, the degree of unification cannot be usefully taken to turn upon simple syntactic criteria such as counting axioms or argument patterns. I'll argue that existing unification - based accounts need to be supplemented by an account of qualitative distinctions between homogeneous and heterogeneous theories, between "natural" and "artificial" predicates. I'll argue further that in both mathematical and broader scientific practice, rational distinctions between more and less natural properties are made systematically. [Tappenden, 2005, p. 147-148]

Thus, according to Tappenden, we need to study the qualitative features of a theory, i.e. "what makes a framework, and the categories in it, natural and homogeneous or whatever" [Tappenden, 2005, p. 167]. As an example, he takes the framework of geometric algebra as offered in Artin's textbook *Geometric Algebra* [Artin, 1957]. In Artin's framework the role of visualization is central and contributes to the theoretical fecundity of his theory. In particular, as Tappenden suggests, visualization is the qualitative feature that contributes to the assessment of the framework as "natural" [Tappenden, 2005, p. 180, 182-83].

> To find a place for unification as a scientific and mathematical success, as it is treated in practice, we need to clarify certain qualitative features of theories and the properties they deal with. Which classes and theories are homogeneous and which are heterogeneous? Which classifications and properties are natural and which artificial? We need to be clear about what sorts of considerations are brought to bear, in deciding what formulations are the right ones to use. The conclusion suggested here, especially as exemplified in the case of *Geometric Algebra*, is that these distinctions are, in practice, made out in a way that is rationally justifiable, but also that they appeal to details of mathematical and scientific practice that are more involved and case-specific than philosophical accounts of explanation as unification have appreciated. This suggests that we reorient our conception of the methodology of mathematics in a "bottom up" direction. [Tappenden, 2005, p. 187]

I will return to the distinction between "bottom-up" and "top-down" methodologies in the next chapter. Nevertheless, it is worth observing here that for Tappenden qualitative injections would permit Kitcher to have a unified treatment of explanations in mathematics and in natural sciences [Tappenden, 2005, p. 174].

(**MH**) Finally, Johannes Hafner and Paolo Mancosu's criticism is an attempt to test in detail Kitcher's theory of explanation by comparing it with a case taken from real algebraic geometry [Hafner *et al.*, 2008].

As we have seen in the subsection 3.2.3, Kitcher discussed the case of a rational transition from $< L, K, E(K) >$ to $< L', K', E(K') >$. In particular, his attention was focused on the conditions under which a systematization Σ' of K' provides a *better* unification of K' than systematization Σ of K does of K. Thus Kitcher's model permits to compare only different systematizations of K belonging to K itself, with premises and conclusions in K (and the same holds for K'). Now, there are cases in mathematics where the kind of systematization which appears in the rational shift is called by Kitcher "systematization by conceptualization"[53]. Kitcher's discussion of Lagrange's analysis of resolvent equations and permutation of equations, which permitted to say when the solution of some particular equations can be reduced to the solution of equations of lower degree, is one of them [Kitcher, 1984, p. 221].

> Generally, systematization by conceptualization consists in modifying the language to enable statements, questions, and reasonings which were formerly treated separately to be brought together under a common formulation. The new language enables us to perceive the common thread which runs through our old problem solutions, thereby encreasing our insight into why those solutions worked. This is especially apparent in the case of Lagrange, where, antecedently, there seems to be neither rhyme nor reason to the choice of substitutions and thus a genuine explanatory problem. [Kitcher, 1984, p. 221]

Consider now an alternative axiomatization to Euclid (say, Euclid*), which uses different axioms from Euclid but which is constructed from the same sentences used by Euclid in K. Kitcher's criteria would permit the comparison among the axiomatizations. However, as Hafner and Mancosu observe, it should be possible that a new systematization would use axioms formulated in a richer language (L^*) than the sentences of K. In that case, the new axiomatization would not use sentences coming from K, as in the case of Euclid and Euclid*, although these sentences would turn out to be equally accepted by the members of K (this would represent, according to Kitcher, a case of systematization by conceptualization). Now, as we said above,

[53] The other kinds of systematization suggested by Kitcher is "systematization by axiomatization". See [Kitcher, 1984], chapter 9, for his discussion of the pattern of axiomatization.

Kitcher's model permits a comparison only between systematizations of K, where derivations have premises and conclusions belonging to K. What then if we would like to make a comparison between the new axiomatization and one which uses sentences from K? We would like to say when the set K^* of new sentences formulated in the new language L^* offers a better explanation of K. To be more precise, we would like to decide when the new systematization, which appeals to a class of sentences K^* richer than K, gives a better explanation of a sentence (a mathematical theorem, in the present case) already contained in K, and thus a better unification of K. Also Kitcher seems to be aware of this point, when he writes:

> For we ought to allow for the possibility that a why-question might be answered by producing a derivation among whose premises is some proposition (or propositions) that is not expressed by any statement in K but which would be rationally accepted by those who believe the members of K. It is even possible that why-questions should be answered by derivations instantiating patterns that are not in the basis of $E(K)$ but that would be included in the basis of $E(K^*)$ where K^* would be rationally accepted by those who accept K and who recognize the validity of the derivations in question. [Kitcher, 1989, p. 435]

Hence Hafner and Mancosu propose a modification to Kitcher's theory, by allowing that in the evaluation of systematizations of K it should be possible to appeal to a class of sentences K^* richer than K. This is made by introducing the following amendment: a systematization of K is any set of arguments which derive sentences in K from other sentences in K^*, where K^* is a consistent superset of K and where K^* can be rationally accepted by those who accept K [Hafner *et al.*, 2008, p. 219][54]. We have then a modification in what counts as an *acceptable* set of derivations: a set of derivations is acceptable relative to K just in case the conclusion belongs to K, every step in the derivation is deductively valid and each premise and each derivation

[54]Observe that this extension still leaves open the possibility that K^* be identical with K.

belong to K^*. This modification, they suggest, allows Kitcher's model to analyze a variety of situations which are very common in mathematics and in science [Hafner *et al.*, 2008, p. 219].

The test case chosen by the authors is taken from the domain of real algebraic geometry and is a particular theorem about real closed fields (RCF)[55]:

Theorem 3.1. *A polynomial $f(x_1, ..., x_n)$ assumes a maximum value on any bounded closed semi-algebraic set $S \subset R^n$*

For this theorem, in his book *Partially Ordered Rings and Semi-Algebraic Geometry* [Brumfiel, 1979] the mathematician Gregory W. Brumfiel has offered three different methods of proofs, which correspond in Kitcher's terminology to three different systematization of RCF.

A first proof strategy is based on the so called Tarski-Seidenberg (T-S) decision procedure, which is considered by Brumfiel as a 'useful tool' for proving theorems in RCF but not a proof technique which offers an explanatory proof[56]. A second proof strategy appeals to transcendental methods (and more particularly on a consequence of the Tarski-Seidenberg decision procedure), while a third proof draws on purely algebraic means and is considered by Brumfiel as a proof which also 'explains' the result. The modification to Kitcher's model permits Hafner and Mancosu to evaluate the different systematizations of RCF, which in Kitcher's sense is the set K to be systematized (in fact, RCF is the consistent and deductively closed set of elementary sentences true in any real closed field), and this by going *beyond* the language and the sentences of RCF. In this case this modification is essential because the T-S decision alghoritm does not belong to RCF as a theorem but it is a

[55]Informally, a real closed field is a field which admits a unique ordering, such that every positive element has a square root and every polynomial of odd degree has a root. One example of RCF is the set of real numbers R. The theory of RCF is the deductive closure of its axioms: Axioms for field; Order axioms; $\forall x \exists y(x = y^2 \vee -x = y^2)$; For each natural number n, $\forall x_0 \forall x_1 ... \forall x_{2n} \exists y(x_0 + x_1 \cdot y + x_2 \cdot y^2 + ... + x_{2n} \cdot y^{2n} + y^{2n+1} = 0$.

[56]Given any first order sentence φ in the language of RCF (an "elementary sentence"), the Tarski-Seidenberg decision alghoritm outputs 1 if RCF proves φ, 0 if RCF proves $\neg\varphi$. The procedure is described in [Tarski, 1951] and [Seidenberg, 1954]. See also [Van den Dries, 1988] for a concise history of how the procedure was introduced.

statement in its metatheory, and thus requires the introduction of a consistent superset (of RCF) as the set from which sentences of RCF may legitimately get premises and derivations. The core of their criticism is that by evaluating the different systematizations, Kitcher's model considers as the best systematization of K that provided by the T-S decision procedure[57], which is explicitly rejected by Brumfiel as non-explanatory [Brumfiel, 1979, p. 166]. Moreover, the model is not able to compare two rival systematizations (the proof strategy which appeals to transcendental methods and that which uses algebraic resources) and discriminate between them. Therefore, by remaining silent in choosing between two rival proof techniques and in failing in accounting for Brumfiel's consideration about the non-explanatoriness of the proof coming from the T-S decision procedure, Kitcher's model conflicts with the mathematical practice.

> Hence Kitcher's model of explanation would declare the set of all instantiations of this single argument pattern [the argument pattern coming from the instantiation of the T-D procedure] as the explanatory store over K, i.e. the set of – explanatory – arguments which best unifies K. This result clearly conflicts with mathematical practice since Kitcher's model ends up positing as the best systematization one which in practice does not enjoy the properties of explanatoriness that Kitcher's model would seem to bestow upon it. Even worse, not only do arguments in this "explanatory store" in general fail to be considered as paradigm explanations, they are hardly ever used at all by working mathematicians because of the limited feasibility of the decision algorithm. [Hafner *et al.*, 2008, p. 228]

Finally, as Tappenden, also Hafner and Mancosu agree on the fact that Kitcher's model should be supplemented with some *qualitative* reinforcement in order to account for the intuitions coming from the practice of the mathematicians:

[57] This is because the systematization corresponding to the T-S procedure permits to generate all of K with only one pattern, thus scoring best in terms of unification power.

> Yet, despite its focus on unification Kitcher's account of explanation apparently does not have the resources to provide insight into the controversy over the "right" proof methods or at least enhance our understanding of Brumfiel's motivations. One of the reasons for Kitcher's failure may lie in the fact that his account, although much more sophisticated than Friedman's model, still shares the latter's basic intuition, namely that unifying and explanatory power can be accounted for on the basis of quantitative comparisons alone. However, in the controversy over the use of transcendental methods in real algebraic geometry the point at issue concerns *qualitative* differences in the proof methods. [Hafner *et al.*, 2008, p. 233; my italics]

3.4 Where are MEPP?

What interests us are mathematical explanations in science. Thus it is very natural to ask: where are MEPP? In what sense is Kitcher's model relevant to the topic of this dissertation?

In the Introduction to this dissertation I quoted a passage from [Mancosu, 2008b] and I observed that MEPP are explanations of physical phenomena where mathematics plays an essential role in the explanation provided. Now, an account of MEPP should address such kinds of explanations. However, there is also the possibility that an account of explanation addresses MEPP and other kinds of explanation as well, and that this choice derives from a particular theoretical standpoint. With respect to Kitcher's example of the Newtonian pattern, we have seen that the mathematical ingredients did not explicitly occur as terms of the schematic argument. They were replaced in the schematic sentences according to the filling instructions. For instance, the dummy letter β in the schematic argument was replaced by an algebraic expression according to a specific filling instruction. On the other hand, in Kitcher's example of the Dalton pattern that I reported in subsection 3.2.3, the filling instructions contained the directions to replace the dummy letters by names of chemical substances (and not mathematical expressions). What is then the common ingredient to these patterns? I think that Kitcher would answer: the fact that these patterns are instantiated by some particular derivation. The fact that these patterns make (or not make) use of mathematical ingredients does not make difference for the unification model. The focus of Kitcher's unification is not on the mathematical or on the empirical ingredients of the pattern (or on the mathematical or empirical ingredients of the derivations which instantiate these patterns), but more on the fact that the pattern is instantiated by some derivation. In fact, as I have already put forward at the beginning of section 3.2, Kitcher considers his model of unification as a *unique* model covering both explanations in natural science and mathematics. Therefore, differently from Steiner, he does not provide a specific answer to the question "When do we have a *mathematical* explanation for a physical phenomenon?" According to the unification way of thinking, the criterion in order to have an explanation (both in science and mathematics) is given by the fact that we have an argument belonging to the explanatory store $E(K)$ over K, and this explanatory store contains particular derivations which minimize the number of patterns of derivation employed and maximize the number of conclusions generated. This is the essence of Kitcher's "theoretical" unification. In this sense, Kitcher's model must be regarded *also* as an account of MEPP. Moreover, behind this theoretical unification there is a methodological holism, because the theory proposed by Kitcher considers the system "science-mathematics" as a whole and it is not focused at the sublevel of structures of natural science or mathematics. As Jaegwon Kim observes:

> To put it somewhat crudely, explanation is a matter of the shape and organization of one's belief system, not of its content. Both Kitcher's and Friedman's accounts make explanation a holistic affair: whether or not a given derivation is an explanation cannot be determined locally, just by looking at the derivation; it depends on facts about the whole belief system. [Kim, 1994, p. 64]

These considerations give me the occasion to trace a dividing line between Kitcher's conception of MEPP and Steiner's. Steiner's model, as we have seen in chapter 1, is parasitic on his account of mathematical explanation within mathematics. Steiner thinks that an account of MEPP must be dependent on a separate account of explanation in mathematics. On the other hand, Kitcher regards his model as able to cover MEPP as well as other kinds of explanations in mathematical and in empirical science. Therefore he does not think that we do necessarily need to construct a theory of MEPP which relies on some particular account of mathematical explanations within mathematics. This is in contrast with Mancosu's suggestion that:

> [...] it is conceivable that whatever account we will end up giving of mathematical explanations of scientific phenomena, it won't be completely independent of mathematical explanation of mathematical facts (indeed for Steiner the former is explicated in terms of the latter). [Mancosu, 2008b, p. 192-193]

Now, if Kitcher's model has been regarded as an overarching model for explanation in mathematics and science (and it has been considered by the author himself as such), it is therefore important to present it and discuss its applicability in the context of MEPP. This is why in the final part of the dissertation (chapter 7) I will come back to this account and I will assess it on a case of MEPP coming from the scientific practice. Now, the reader might be surprised for this choice, and may ask: 'Why don't you give that assessment *here*?' I have a very natural reason for my choice. To test Kitcher's model later in the dissertation, and not in this chapter, is perfectly consonant with the general strategy that I have adopted for my investigation. In chapter 7 I will point to the difficulties that this and other accounts have in capturing a MEPP which is recognized as such in scientific practice. And from that analysis I will propose a diagnosis which will lead me to a change of perspective toward MEPP (chapter 8).

Unfortunately, the complex structure of Kitcher's unification model of explanation makes very difficult to illustrate the model without giving a detailed presentation of it. And I have to admit that I have not succeeded in such a task of reduction. This is why this chapter has filled more space than the previous ones, thus testing the patience of the reader. On the other hand, I hope to have persuaded the reader that the space devoted to unification and the detailed presentation of Kitcher's model are needed for the general topic of the dissertation. And that the content of this long chapter is not only relevant to MEPP, but it is also necessary to develop my original ideas in the final part of this study.

In the next chapter I am going to individuate some general features of the WTA approaches. In this context, Kitcher's theory (together with its detailed structure) will be useful because it will help me to individuate these features.

Chapter 4

Some features of the winner-take-all approaches

In this first part I have presented three WTA accounts of explanation which have been proposed as candidates as to cover MEPP. While the pragmatic and the unification account have been originally developed as general accounts of scientific explanation, Mark Steiner's model has been proposed by the author in order to explicitly cover MEPP (even if, as we have seen, Steiner considers that MEPP depend on mathematical explanations within mathematics).

The fact that these positions are very different among themselves indicates well how the topic of MEPP could be considered from very different perspectives. Furthermore, as we have seen, every account is inevitably interconnected with other general problematics in philosophy of science and with the position that those authors endorse toward these problematics. Of course, this should not come as a surprise. Behind any philosophical account there are men, and behind those men there are ideas, conceptions or misconceptions about the world. Nevertheless, as I am going to suggest below, an evaluation of such a close linkage between explanation and general topics in philosophy of mathematics and of science is useful to provide some broad characterization of theories of explanation. Moreover, I consider that this characterization is not only relevant to the general debate on MEPP, for instance because it would permit to discriminate one account of MEPP from another (in some specific sense), but it is also worth for the second and the third part of this dissertation, where I will make use of it. Finally, giving such a characterization is necessary in order to give a more uniform view of this part I. For expository's sake, in giving it I will take three themes as general guidelines:

- Explanatoriness as a global or local feature
- Tension ontic-epistemic (expressed in the interplay explanation-understanding)
- Relevance relation

4.1 Explanatoriness: global or local feature?

I am not interested in commenting here the difference between a causal theory of explanation (as Salmon's, Railton's or Lewis') and Kitcher's unification approach[1]. On the other hand, I am going to begin this section by illustrating one of Kitcher's motivations for developing his account, and to do that I will need to follow Kitcher in constrasting causal with non-causal explanations. This short discussion will substantiate what I have said in the previous chapter, namely that Kitcher's account has been proposed by the author to cover also mathematical explanations. Moreover, it will permit me to introduce the characterization of the accounts in terms of the global or local character they attribute to explanatoriness.

For Kitcher, to take the concept of causality as the central concept for a theory of explanation is not the most fruitfull approach to a good theory of explanation. This remark is supported by the observation that in domains such as formal syntax or mathematics we have explanations which are not

[1]The potential fruitfulness of the unification approach, and the fact that theoretical explanation should be regarded as primary with respect to causal explanation, is largely discussed by Kitcher in his [Kitcher, 1985b] and [Kitcher, 1989].

causal ([Kitcher, 1985b, p. 637], [Kitcher, 1989, p. 422-428]).

The examples from mathematics reported by Kitcher include (A) the case of Bolzano's proof of the intermediate value theorem, (B) the proof of a property of finite groups by means of one specific axiomatization of the theory of finite groups, and (C) Galois' theory as to explain *why*, for a specific class of equations in one variable (linear, quadratic, cubic and quartic), it is possible to express roots as rational functions of the coefficients. Furthermore, Kitcher's claim about the limited scope of a causal theory of explanation is reinforced by the observation that there exist also physical explanations which are not causal. Consider, for instance, the following situation: someone has knotted a telephone cord around a pair of scissors. We are asked to free the scissors. The scissors can easily be removed if we make a right twist at the start. If not, we will not be able solve the trick but we will only obtain a more complicated configuration. Suppose we do not the right twist at the start. How do we explain our failure in solving the trick and the tangled situation we obtained? According to Kitcher, the causal history (i.e. the sequence of actions actually performed leading to the tangled configuration) is not sufficient enough for the explanation of the failure [Kitcher, 1989, p. 426]. We need something more, namely, to know the topological features of the situation that lie behind the causal history and that permit us to say why our attempt was doomed to failure. More precisely, we can say that we did not solve the trick because our sequence of actions did not satisfy the topological condition necessary to solve it[2]. The moral of the latter example is that even when we have in our hands the causal history of the phenomenon, we may not have at disposition what the explanation requires.

These examples of non causal explanations (in mathematics and in physics) provide, according to Kitcher, sufficient reasons to switch the focus from causal explanation to "theoretical" explanation, and consider the latter as primary to the former. As he suggests, while "at each stage [of science] the

[2] A very similar example, based on geometrical rather than topological considerations, has been proposed by Peter Lipton [Lipton, 2004, p. 9-10]. I reported this example in the Introduction to the dissertation.

explanatory store supplies an ordering for the phenomena and serves as a basis for the introduction of causal concepts" [Kitcher, 1989, p. 477], the converse is not true because causal notions are inapplicable in domains such as mathematics nor even physics[3].

> For even in areas of investigation where causal concepts do no apply –such as mathematics– we can make sense of the view that there are patterns of derivation that can be applied again and again to generate a variety of conclusions. Moreover, the unification criterion seems to fit very well with the examples in which explanatory asymmetries occur in mathematics. [Kitcher, 1989, p. 437]

During his discussion of the mathematical examples (A) (B) (C), Kitcher suggests that to cases of explanatory irrelevances and asymmetries in empirical science there correspond analog cases of explanatory irrelevances and asymmetries in mathematics [Kitcher, 1989, p. 424-425][4]. Concerning example (A), Kitcher observes that Bolzano considers the geometrical proofs of the intermediate value theorem as not explanatory. In particular, Bolzano regards these proofs as not explanatory because geometrical facts are foreign to the analytical domain under investigation[5]. Therefore, Kitcher observes,

[3]Even Salmon is conscious of the difficulties faced by his causal-account, for instance in the case of explanations in quantum mechanics: " [...] there are fundamental difficulties in principle in attempting to provide causal explanations in terms of spatiotemporally continuous causal processes and localized interactions in the quantum domain. I am not inclined to dispute this claim. Rather, I should say, it appears that causal explanations of the sort discussed above are adequate and appropriate in many domains of science, but that other mechanisms –possibly of a radically noncausal sort– operate in the quantum domain. If that is true, then we need to learn what we can about those mechanisms, so that we can arrive at a satisfactory characterization of quantum mechanical explanation. It may turn out that the causal conception of scientific explanation has limited applicability" [Salmon, 1984b, p. 298].

[4]I introduced the problem of explanatory irrelevances and the problem of asymmetry in my presentation of the D-N model, in section 2.2.

[5]For a discussion of Bolzano's distinction between explanatory and non-explanatory proofs see [Kitcher, 1975], [Mancosu, 1996], [Mancosu, 1999] and [Mancosu, 2000]. Bolzano explicitly adopted the Aristotelian distinction between proofs of the fact (*hoti* proofs) and proofs of the reasoned fact (*dioti* proofs). While hoti proofs shows *that* something is, only the latters (dioti proofs) give the reason *why*, the cause (see Aristotle's *Posterior Analytics* I.13).

Bolzano's remark about the irrelevance of geometrical considerations to the proof of the intermediate value theorem can be paralleled with the same considerations that have been raised in the context of the D-N model for the problem of explanatory irrelevances. In this sense, Bolzano's case represents the mathematical analogue to the irrelevance problem for explanation in the natural sciences. Concerning example (B), Kitcher observes how one particular axiomatization containing the existence of the inverse and idempotent elements is preferred by the mathematicians in order to explain why finite groups satisfy the division property. On the other hand, the reverse derivation, i.e. the derivation of the existence of an idempotent element and of inverses from the division property, is regarded as a less natural and non-explanatory derivation (although formally valid). Example (B) can therefore be paralleled with the asymmetry problem in empirical science, where only one direction is considered as explanatory[6].

The previous considerations can be summarized as follows. First, for Kitcher there exist explanations which are not causal. Second, the problems of asymmetry and irrelevance are not a privilege exclusive of the causal debate on explanation. Explanatory asymmetries and irrelevances arise also in the domain of mathematics. And Kitcher's methodological attitude towards those kinds of problems in mathematics is the same as in empirical science, namely, to see how they can be solved through the lens of the unification account.

Contrary to Salmon's causal approach, then, Kitcher's theoretical account is built to capture also mathematical explanations (and causal explanations as well). This, according to Kitcher, should be considered a remarkable advantage[7]. Observe, however, that in considering theoretical explanation as

[6]To my knowledge, no authors have proposed this analogy, i.e. the analogy between the problem of asymmetries and irrelevance in science and a same kind of problem in mathematics, before Kitcher in his [Kitcher, 1989].

[7]In passing, let me observe that the "greater applicability" of the unification model (with respect to a causal approach such as Salmon's) is evident if we consider the different domains in which the account has been tested. For instance, Maki and Marchionni recently tested Kitcher's account on a case of "explanatory unification in the social science", in the domain of geographical economics [Maki *et al.*, 2009].

primary to causal explanation, Kitcher is adopting a perspective on explanation which is based on a completely different paradigm:

> Salmon's approach to explanation is "bottom up". Explanation consists in identifying causal relations. Causal relations primarily relate individual events; so the explanation of particular occurrences is fundamental. [...] Hempel's approach to explanation was "top down". Explanatory concepts were conceived as prior to causal concepts. But the D-N model foundered on its liberality. The asymmetries of explanation invited philosophers to make explicit appeal to causal notions. I claim that a more radical "top down" approach is possible. Begin from the idea that explanation is directed at an ideal of scientific understanding. We achieve that ideal by giving a unified, deductive systematization of our beliefs. *Our views about genuine properties and explanatory dependence emerge from the project of unifying the regularities we discover in nature.* On this approach, theoretical explanation is primary. Causal concepts are derivative from explanatory concepts. In explaining particular events we answer as many questions as we can, drawing on our view of the order of natural phenomena. In some cases, our ideal of understanding may not be completely realizable [Kitcher, 1985b, p. 638-639. My italics]

In Salmon's account, explanation consists in identifying causal relations. Therefore we can say that "explanatoriness" is a (local) property possessed by the local relation between the explanandum and the explanans. On the other hand, for the unification-believer, "explanatoriness" must be considered a *global* property of a theory or framework (of scientific beliefs). This is in line with Kitcher's holistic picture of explanation. For instance, we have seen how Kitcher considered the Newtonian derivation explanatory because it belongs to $E(K)$, which is the best systematization of our beliefs. According to the unification view, explanations depend on global constraints (they belong to $E(K)$). On this basis, Jaegwon Kim has defined Kitcher and Friedman as "explanatory internalists":

> It is useful to view Friedman and Kitcher as explanatory internalists [...] what makes these derivations explanatory, on Friedman's and Kitcher's accounts, is their relationship to other items within our epistemic system, not some objective facts about external events or phenomena. On Kitcher's account, for example, what makes a given D-N argument explanatory is the fact that it is a member of a class of arguments (a "systematization", as Kitcher calls it, of our belief system) which best unifies our belief system. And the measure of degrees of unification depends solely on factors internal to the epistemic system, such as the number of argument patterns required to generate the given class of arguments, the "stringency" of the patterns, etc., not on any objective relations holding for events or phenomena involved in the putative explanations. [Kim, 1994, p. 63-64][8]

In Mark Steiner we are faced with an opposite situation. In his model, in fact, the "explanatoriness" property is a *local* feature, and precisely a local feature of the proof. In other words: while in Kitcher the fact that mathematics be explanatory in the description of a physical phenomenon depends from the fact that the explanation (the derivation) be part "in a certain way" of a whole theory or system of beliefs, Steiner's idea is that in every particular mathematical explanation of a physical fact we have a local property which provides explanatoriness (in such an explanation we can separate the mathematical part from the physical one and we identify a specific local characteristic, called characterizing property, which provides us with explanatoriness). As put forward by Mancosu:

> this [the contraposition between local and global] captures well the difference between the two major accounts of mathematical explanation available at the moment, those of Steiner and Kitcher. [Mancosu, 2008b, p. 195]

[8] Kim observes that "any unification approach to explanation will be holistic, although this isn't true of all internalist theories (compare, e.g., Hempel's covering-law theory)" [Kim, 1994, p. 64, footnote 22]. This is because any explanation considered under a unification perspective (along the lines of Kitcher or Friedman's views) would be characterized not locally, but in terms of facts about our whole belief system.

Finally, as it emerges from the detailed discussion we made for the pragmatic account, Van Fraassen's theory does not characterize "explanatoriness" in any (local or global) sense. This simply because in his view explanatoriness can be a local property, but it can be a global property as well. For instance, if we consider the relevance relation as that of causation (something that Van Fraassen accepts[9]), explanatoriness will be a local property of a state of affairs. On the other hand, if we fix the relevance relance as that of intentional relevance, as it is the case in his example of the tower and the shadow, we must consider that explanatoriness is a global property of a whole system of belief, since it is this system of beliefs which fixes the relevance relation. We can thus say that this theory is *neutral* towards the view of explanatoriness as a global or local feature. Nevertheless, following Van Fraassen in his "pragmatic" characterization of explanation (context-dependence, etc.), it is easy to understand why he claims that explanation (and *not* explanatoriness) is not a global affair but a local one [Van Fraassen, 1980, p. 109-110].

4.2 Ontic versus epistemic

To every account on explanation there corresponds a tension towards the 'epistemic' or the 'ontic' côté[10]. I use here the terms epistemic and ontic with the following semantic values: I call *ontic* an account of explanation which is based on a relation that is characterized independently from the categories linked with the subject who knowns (the understanding is an example of such categories); I call *epistemic* an account which uses a relation which is defined through categories, such as the understanding, which are linked to the subject[11].

[9] Although, of course, he regards a caual relation as no more than 'what because must denote' [Van Fraassen, 1980, p. 155].

[10] The importance of this tension as a guideline to 'read' the recent history of philosophical models of explanation has been pointed out by Paolo Mancosu in his lecture "Understanding, Explanation and Unification", during the workshop *Mathematical Understanding* (Paris, June 2008).

[11] The semantic values of those terms change with respect to the authors who use them. In his [Salmon, 1984b], Wesley Salmon refers to the "epistemic conception of scientific

While authors as Toulmin, Scriven or Hanson have made an explicict claim to the notion of understanding as a desideratum of their model, other authors as Hempel did not consider understanding as central or relevant to the discussion about explanation and offered only a precise (formal) definition of the nature of the explanation relation. Restricting my analysis to the authors considered, we have already seen how Friedman tried to solve this tension by linking an objective definition of explanation with that of understanding (in particular by offering a theory of explanation that would able to incorporate a notion of understanding in science). While Kitcher followed the same way[12], in Van Fraassen we have that the focus is much more balanced on the epistemic side. According to Van Fraassen, there is no specific characteristic of an explanation which gives us explanatoriness[13], but explanatoriness is given within a specific context by the interests of a questioner. Explanation are answers to why questions, and the theory addresses the problem of how these answers are formulated (theory of why questions) and when they are legitimate (evaluation). The fact that the answer be legitimate depends on the relevance relation, and the latter is (roughly speaking) fixed by the questioner and his interests.

Mark Steiner is more focused on the ontic perspective and he does not refer to the notion of understanding or to any epistemic value in his explanatory theory. In his theory of MEPP a particular objective feature ("characteriz-

explanation" as that view which regards explanations as arguments, while the "ontic conception" is considered as the conception which "sees explanations as exhibitions of the ways in which what is to be explained fits into natural patterns or regularities" [Salmon, 1984b, p. 293]. In this classification, Hempel and Van Fraassen both belong to the epistemic conception of explanation, while the second is advocated by authors as Michael Scriven and Salmon himself.

[12]For instance, he writes: "Understanding the phenomena is not simply a matter of reducing the 'fundamental incomprehensibilitie' but of seeing connections, common patterns, in what initially appeared to be different situations. Here the switch in conception from premise-conclusion pairs to derivations proves vital. Science advances our understanding of nature by showing us how to derive descriptions of many phenomena, using the same patterns of derivation again and again, and, in demonstrating this, it teaches us how to reduce the number of types of facts we have to accept as ultimate (or brute)" [Kitcher, 1989, p. 432].

[13]This point is stressed by Salmon in the passage I quoted at the end of subsection 3.1.1 [Salmon, 1989, p. 131].

ing property") is the kind of (objective and epistemic-neutral) relation that provides us with explanatoriness. Therefore we can regard Steiner's account as an ontic account of explanation.

It should not come as a surprise that, as pointed out by Peter Railton, different views in explanations correspond to different metaphysical views:

> Except in a few polemical places, theories of explanation were described by their formal features – "covering law", "why-questions", "speech-act", "statistical relevance" – and did not come prefixed with such metaphysical codes as "empiricist", "pragmatist", or "realist". Yet at the table sat empiricists, pragmatists, and realists. [Railton, 1989, p. 220]

Thus, for instance, Salmon's causal account corresponds to a strong version of scientific realism. Concerning the authors discussed in this first part of the dissertation, Mark Steiner's ontic perpective on explanation is connected to his realist position in philosophy of mathematics, while Van Fraassen's pragmatic model mirrors his anti-realist conception in science[14]. Baker, as we have seen, defends his platonist position in philosophy of mathematics by appealing to the role that MEPP have in the enhanced indispensability argument. Although Kitcher seems to accept a kind of methodological realism, without endorsing any particular metaphysical position[15], his and Friedman's unification proposals are congenial to scientific realism. In particular, the fact that Friedman and Kitcher's positions are more inclined towards scientific realism has been observed by Peter Railton in his [Railton, 1989, p. 227, 230]. The point is that the unification approach gives a role to the postulation of structures and mechanisms (different phenomena are traced to a common structural basis), and therefore it involves commitment to the existence of the reducing entities or properties. For instance, suppose we have a physical model of the atom which unifies diverse phenomena (emission and absorption

[14] Nevertheless, as we have observed in chapter 2, his model is not incompatible with realism.

[15] See [Kitcher, 1989, p. 500].

spectra, conductivity of metals, etc.) into a common structural (or causal) basis (the model of the atom). This basis is characterized theoretically, then if we were to eliminate the commitment to the reducing entities (atoms, electrons, etc.), we would lose the physical model by means of which the unification is achieved. In Friedman's book *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science* [Friedman, 1983], this commitment is explicitly stated. Here the author claims that the principle of unifying power provides a defense for a realist interpretation of theoretical structures. In particular, we should attribute a realist interpretation only to all the theoretical elements which have uniyfing power (we shall call those elements "good" theoretical elements, while the theoretical elements without unifying power, such as Newton's absolute space, are called "bad" theoretical elements), because in such cases the theory has a higher degree of confirmation under a realist interpretation than under a non realist one[16]. To quote a long but very informative passage of Friedman's text:

> More generally, we can put the matter as follows. Physical theories postulate a structure $\mathcal{A} = \langle A, R_1, ..., R_n \rangle$ [where A is an open region of space-time and $R_1, ..., R_n$ are observational properties] that is intended to be taken literally, this is supposed to have physical reality. Physical theories also typically invoke various representative elements – pieces of mathematical structure that are not intended to be taken literally. In this latter category we find such things as choices of coordinates, units, and so on; in our present context, we are always dealing with a representation $\phi : \mathcal{A} \to R^4$. Our problem is to find a rationale for this practice. On what basis do we assign some pieces of structure to the "world" of physical reality $\mathcal{A}$ and other pieces to the "world"

[16]See [Hiskes, 1986] for a detailed review of Friedman's book. She underlines one important aspect of Friedman's discussion: "unification between theories limits one's choice between alternative theoretical models of a given observational structure, thus providing a defense against the 'underdetermination' of all theory. For example, there may exist quite different models of space-time geometry, both of which embed all possible data, but only one of these can be unified with a given theory of kinematics. As Friedman correctly points out, taken to its extreme, this problem is just the problem of alternative total theories of the world, and it extends to observational descriptions as well" [Hiskes, 1986, p. 123].

> of mathematical representation R^4? My answer is that this practice is based on the unifying power of theoretical structure. A particular piece of structure postulated by an initial theory of the form
>
> $$\exists \phi \in \Phi : \mathcal{A} \to R^4$$
>
> (where Φ is a class of mappings) has unifying power in the context of a second theory of the form
>
> $$\exists \psi \in \Psi : \mathcal{A} \to R^4$$
>
> (where Ψ is a second class of mappings) just in case it facilitates the inference to
>
> $$\exists \chi \in \Phi \cap \Psi : \mathcal{A} \to R^4.$$
>
> If this inference goes through even without the structure in question, as in the example of absolute rest and gravitational theory, it has no unifying power and can be safely dropped from $\mathcal{A}$. If, on the other hand, the structure in question plays a necessary role in many such inferences, we have no choice but to take it literally, to assign it a rightful place in the "world" of physical reality $\mathcal{A}$. For otherwise our total theory of $\mathcal{A}$ is much less well-confirmed. [Friedman, 1983, p. 250]

Friedman's "inference to the unified explanation" (i.e. the use of unifying power as a criterion for a realist interpretation of a theoretical structure) has been strongly criticized by Margaret Morrison in her [Morrison, 1990][17]. Among the main objections to Friedman, Morrison raised the folowing point: while Friedman's unification is constructed as a context-dependent process (here context is used in a wider way than in Van Fraassen; it refers to Kitcher's set K of accepted belief at a *particular* time), it cannot motivate realism because this context-dependence provides no ontological stability [Morrison, 1990, p. 327-328].

[17] Morrison's paper, which is basically a previous version of chapter 3 of [Morrison, 2000], contains a short and clear illustration of Friedman's characterization of the relationship between observational and theoretical structure as that of submodel to model.

The foregoing remarks show very well how the problematics of explanation could not be disentangled from the metaphysical arena and might play a discriminant role in general disputes in philosophy of science and mathematics. In 1989, more than ten years before Mark Colyvan's book *The Indispensability of Mathematics* [Colyvan, 2001] and the 'advent' of the enhanced indispensability argument, in his study "Explanation and Metaphysical Controversy" [Railton, 1989] Peter Railton observed::

> Interestingly, within that dispute [the dispute between realists and irrealists] *the concept of explanation proved to be indespensable*: one side often claimed that a realist interpretation of scientific theory was justified by inference to the best explanation; the other side often responded that the realist's posits could do no explanatory work, since they yielded no empirical predictions beyond those already afforded by the observational reduction of the theory. [Railton, 1989, p. 221; my italics]

This is evident if we consider the key role that is given to the notion of explanation in books which deal with specific subjects of philosophy of mathematics, for instance Colyvan's book. Moreover, in the previous chapters I have shown how the role of explanation was central in Baker's and Steiner's realist claims, although in a different flavour[18]. In the context of these debates between the realists and the anti-realists (in general philosophy of science or in philosophy of mathematics), the characterization in terms of the epistemic/ontic distinction as that introduced here might be useful to disentangle the various positions and might contribute to make the ontological debate more transparent[19].

[18] Alan Baker justifies the existence of numbers thorugh MEPP and IBE's criterion (in the enhanced indispensability argument). On the other hand, although endorsing realism towards mathematical entities, Mark Steiner denies the possibility to use MEPP in inferring the existence of numbers.

[19] I leave this task here as an aside, and in this dissertation I will not use the ontic/epistemic characterization of the accounts in the context of debates between realists and anti-realists.

4.3 Relevance relation

The tension ontic-epistemic is linked to a very particular feature that we can individuate in the WTA view: a relevance relation. In this section I want to suggest that the three WTA models can be 'read' in terms of this relevance relation, and this relevance relation can be objective (as for Kitcher's and Steiner's model) or not (as in the case of Van Fraassen's model).

Let me reconsider here shortly Van Fraassen's definition of relevance relation. For Van Fraassen a *relevance relation* R is the "respect-in-which a reason is requested". We saw that in his theory of why-questions an explanation takes the form "Because A", where the proposition A bears relation R to the couple $\langle P_k, X \rangle$. Now, even if Van Fraassen did not offer any constraint or formal characterization of relevance requirement on R, I think that we can read Steiner's and Kitcher's model in terms of Van Fraassen's *objective* relevance relation[20].

In Steiner's account of explanation in mathematics the characterizing property is exactly what permits us to fix an objective relevance relation between a mathematical entity (mentioned in a proof) and a theorem (or a class of theorem –take in mind deformability!)[21]. To switch to Steiner's account of MEPP, let's consider A as a proposition which has mathematical content (for instance, Euler's theorem), while the topic P_k as a proposition expressing the phenomenon to be explained (for instance, the fact that when a body is rotated there is an axis which passes through the center of the body and which does not move). As I have suggested in a footnote of section 2.4, we might consider that Steiner's criterion C_{MEPP} can be used to model a particular relevance relation R_{st} that A bears to the couple $\langle P_k, X \rangle$. This relation would be objective because dependent on Steiner's objective charac-

[20]We can also read Salmon in Van Fraassen's terms, as Kitcher does: R is a relation of particular causal relevance. Kitcher calls "causal why-questions" those particular why-questions [Kitcher, 1989, p. 420-421]. In what follows I am not going to offer a "paraphrase" of Steiner's and Kitcher's ideas in the form of relevance relation. Rather, I consider that such a task might be undertaken.

[21]In passing, let me observe that Steiner's account of explanation in mathematics has been formulated in terms of why-questions by Weber [Weber *et al.*, 2002, p. 300].

terizing property. Moreover, once R is adopted, the answer 'Because A' will be (of course!) relevant to the topic P_k with respect to its alternatives and will favor the topic with respect to the other members of the contrast class X. In Kitcher's case, it is the author himself who suggests a sketch of how his account can be read in terms of Van Fraassen's pragmatic account and a relevance relation[22]:

> An ideal why-question acceptable relative to K is a triple $\langle P, X, R \rangle$ where P is expressed by some member of K, X is an admissible contrast-class, and R obtains between a sequence of propositions A and $\langle P, X \rangle$ just in case A is expressed by a derivation in $E(K)$ whose conclusion expresses the conjunction of P and the negations of the remaining members of X. An actual why-question acceptable relative to K is a triple $\langle P, X, R \rangle$ where P, X must satisfy the same conditions as before and R holds between A and $\langle P, X \rangle$ just in case A is a subsequence of a sequence of propositions expressed by a derivation in $E(K)$ whose conclusion expresses the conjunction of P and the negations of the remaining members of X [Kitcher, 1989, p. 435]

All the three accounts illustrated in this part share then a same feature: they all define explanation in terms of a particular kind of logical relevance relation between the explanans and the explanandum (at least under the reading of this section). In Kitcher and Steiner the relation is objective and it is provided, while in Van Fraassen this relation remains free and Van Fraassen excludes the possibility for such a relation being objective.

4.4 Conclusion

As we have seen through the three guidelines proposed above (explanatoriness as a global or local feature, tension ontic-epistemic, relevance relation),

[22]See [Kitcher, 1985b, p. 633-634] and [Kitcher, 1989, p. 435-436] on the possibility of integrating the unification account with a pragmatic approach. Of course, Van Fraassen would refuse such integration on the basis that, for him, there no exists a criterion for constraining the relevance relaiton.

the three accounts I have examined share some general features but differ with respect to other aspects.

These authors adopted different methodologies of research to build their models, and these methodological choices are inevitably connected with some metaphysical convinction. This makes extremely difficult to analyze the models under an uniform framework (as, for instance, this dissertation). Therefore, by identifying three general guidelines, I have tried to build an instrument to constrast a particular account of explanation with another (at least with respect to some aspects) and disentangle these models (together with the respective positions of the authors).

There are, of course, other aspects which could have been picked out from the models and which might have been considered. For instance, although I have based my choice on the exigence of the strategy of this dissertation, another aspect concerns the role that the notion of 'generality' play in such accounts. Let me shortly consider how this aspect can be appreciated in Kitcher's model, and how it is connected to the global characterization I offered in section 4.1[23].

In his book *The Nature of Mathematical Knowledge* [Kitcher, 1984], Kitcher considers the development of mathematics as rational and offers an account of the growth of mathematics through the rational steps which characterize its practice. The question is: if we assume that mathematics develops according to some rational criteria, what are the patterns of change which are typical of mathematics? This question, or better an answer to this question, might have strong consequences for any model of explanation which refers to the realm of mathematics. In his discussion, Kitcher identifies the role of generalization as of primary importance in mathematical explanation. In particular, what

[23]The choice to focus on the role of generality in Kitcher is motivated by the fact that the role of 'generality' (and generalizability) in Steiner's account has been already discussed in the relative chapter. Moreover, in Steiner the role of generality concerns his account of mathematical explanation within mathematics. In Van Fraassen the notion of 'generality' seems to play no role (as far as I see), and no such analysis has been proposed for his account (at least to my knowledge). An analysis of the role of generality in connection with explanation in Kitcher's account as applied to explanations in mathematics is offered in [Mancosu, 2008b].

he calls "significant generalizations" are *a* kind of pattern of mathematical change which provide explanatory power (other patterns of change, and thus good candidates for offering explanation in mathematics, are "rigorization" and "systematization"). Referring to his mathematical examples, Kitcher remarks that "the preferred derivation can be generalized to achieve more wide-ranging results [...] the explanatory derivation is similar to derivations we could provide for a more general result; the nonexplanatory derivation cannot be generalized, it applies only to the local case" [Kitcher, 1989, p. 425]. It is easy to see here how the role of generality in Kitcher is strictly connected to what I said in section 4.1, namely that in Kitcher's model explanatoriness is a global feature of a framework. Recall that the intuitive idea behind unification is that $E(K)$ is the set of derivations that makes the best trade-off between minimizing the number of patterns of derivation employed and maximizing the number of conclusions generated (the 'unification criterion'). Now, we can say that explanatoriness is a (global) property of such a framework $E(K)$. Therefore, to see how the notion of explanatoriness (as a global feature of the framework) and generalizability are connected in Kitcher's model, it is sufficient to observe that in $E(K)$ we find those derivations which maximize the number of conclusions, namely the (explanatory) derivations which can be generalized. On this account, a derivation which cannot be generalized does not belong to $E(K)$ and is therefore not explanatory, i.e. it does not contributes to the explanatoriness of $E(K)$[24].

In giving the three guidelines, I emphasized the fact that the debate around scientific explanation has been a bridge to the arena realism-antirealism

[24]Kitcher's claim that generality provides explanatory power is criticized by Margaret Morrison. Although, as we have seen, Morrison does not agree with Kitcher on the linkage explanation-unification, she looks at generality as the "basis of the unifying power" [Morrison, 2002, p. 348]. However, for her this generality reduces explanatory power (while in Kitcher's view generality provides explanatory power). In the particular case of Maxwell unification, the generality of the framework is the basis which permits the unification through the Lagrangian approach [Morrison, 2002]. But, as we have seen in the previous chapter, Morrison points out that in this process "explanatory" details are lost: "[...] the generality provided by such mathematical structures [the mathematical structures through which the unification is achieved] can actually detract from rather than enhance the theory's overall explanatory power" [Morrison, 2000, p. 31].

in science, and that the philosophy of mathematics has inherited the same mechanism. The fact that MEPP are now seen to play a key role in the ontological debate between nominalists and platonists is something that I have already mentioned in chapter 2, where in footnote I gave a sketch of the Melya-Colyvan debate. On the other hand, we have seen in this chapter that it is not only the notion of explanation that plays a role in our metaphysical convictions, but the converse is also true (i.e. that our conceptions of the world do play a role in our conception of explanation). As Peter Railton has pointed out:

> To say that one's background picture of the world is involved in one's *conception* of explanation is to suggest that one's intuitions about particular kinds or instances of purported explanation may not constitute a body of neutral data for testing theories of explanation. [...] Equally, it is to suggest that there may not be a unitary, substantial *concept* of explanation to analyze, or, more accurately, that the concept of explanation is rather thin, too slight, perhaps, to be asked to resolve deep philosophical disputes [Railton, 1989, p. 224].

Now, as Peter Railton suggests, there may not be a unitary, substantial *concept* of explanation to analyze. This is perfectly in line with the idea that stands behind the pragmatic account. Nevertheless, let me add to Railton's remark that there may not be a single model to capture this not unitary concept of explanation. Now, bearing these two remarks in mind, namely that there may not be a single model and explanations may be heterogeneous, we can ask if there is an approach to MEPP which goes in this direction. An answer for this is 'Yes, there is such a perspective on MEPP', and it is to this perspective that I am going to turn my attention in part II.

In the next part, I am going to present a view on MEPP which is very different from the WTA view. The next authors consider there are different kinds of MEPP but they do not think that there is a *single* model to which the notion of explanation can be reduced. These two ingredients characterize their 'pluralism' toward MEPP. Now, I think that it is necessary to add here a remark concerning the contrast WTA/pluralism. More particularly, I want to make plain that there is an essential difference between the pluralist view I am going to present and the WTA view adopted by the approaches that I have analyzed in this first part.

Consider Kitcher's and Steiner's models. Kitcher proposes an holistic picture of explanation in which explanations (among them MEPP) are derivations belonging to the explanatory store $E(K)$. Steiner gives a single model for MEPP and looks at MEPP as those explanations which are in tune with his criterion C_{MEPP}. It is easy to see that these authors are not pluralist toward MEPP. With Van Fraassen things are more subtle. In fact, as we have seen, Van Fraassen accept that there are different kinds of explanation (and then different kinds of MEPP, if we consider the extension of the model to MEPP). This is clear if we consider that he leaves open his relevance relation R. On the other hand, and here is the aspect which makes his model *not* pluralist, Van Fraassen regards an explanation as an answer to a why-question and therefore he does offer a single model to capture the notion of explanation[25].

Finally, according to the fresh perspective adopted by the authors of the next part, it is more favourable a "bottom-up" methodology in the study of MEPP, i.e. a methodology according which the philosophical investigation of MEPP should start from the observation of scientific and mathematical practice[26]. As we are going to see, for these philosophers Hempel's dictum

[25]Moreover, the model, as emerged from Sandborg's criticism (subsection 2.3.3), does not capture some explanations (in that case mathematical explanations in mathematics) which do not come under the form of an answer to a why-question. In the final part of the dissertation, in chapter 7, I will offer an example which shows that there are also MEPP which do not come under the form of answers to why-questions. Here, again, I must pray the patience of the reader.

[26]Paolo Mancosu has stressed the importance of such a type of methodological approach in the context of mathematical explanation: "Previous theories of mathematical explanation proceeded top-down, that is by first providing a general model without much concern for describing the phenomenology from mathematical practice that the theory should account for. Recent work has shown that it might be more fruitful to proceed bottom-up, that is by first providing a good sample of case studies before proposing a single encompassing model of mathematical explanation" [Mancosu, 2008c]. The same idea, namely to focus on the mathematical practice by get our hands dirty with the details of mathematics

"rules as limits" disappears into the jungle of explanation.

and then proceede in a "bottom up" direction, is shared by Tappenden [Tappenden, 2005, p. 187-188]. Mancosu and Tappenden's considerations about the importance of such a methodological approach are reflected, as we have seen in the previous chapter, in their criticisms to Kitcher's model. Moreover, let me observe that the way in which I am using here "bottom-up approach to explanation" and "top-down approach to explanation" is different from the way Kitcher used the same expression in his passage from [Kitcher, 1985b] (I have quoted the passage in section 4.1). I use bottom-up/top-down with the (methodological) connotation proposed by Mancosu.

Part II

The pluralist way to MEPP

> The prospect of any one of these models being developed to cover all good scientific explanations (let alone good explanations in general) are dim. Perhaps we sould opt for pluralism. Perhaps there are several types of explanation, each with its appropriate model.
>
> W. H. Newton-Smith, *A Companion to the Philosophy of Science*, p. 132.

In this part II, I will illustrate two positions which diverge from the three WTA views presented in the first part and which offer a very different perspective on MEPP. The defenders of these views do not propose a solution to the whole problem of MEPP, i.e. they do not propose a single encompassing model of MEPP as Steiner, Kitcher and Van Fraassen did, but focus on specific kinds of explanations and try to tell only *part* of the story about MEPP. The views I am going to illustrate are that of Christopher Pincock, who attempts to accommodate the explanatory role of mathematics in his mapping account of the application of mathematics, and that of Robert Batterman, who focuses on a particular form of scientific reasoning ("asymptotic reasoning") which he considers as central to a particular form of MEPP. To discuss these positions will highlight the relation mathematical modelling and idealization have with explanation, a connection which has received much interest among various philosophers of science.

The main aim of this part is thus to present some positions on MEPP which explore only fragments of the problem of explanation, thus contributing to the idea that does not exist a WTA approach to MEPP but such explanations are heterogeneous and cannot be captured by a global model. Putting to one side the embarrassment for the philosophy of science (why are we unable to find a general model? Is our investigation for a global model such a visionary search?), to endorse the idea of 'pluralism in explanation' (there are several kinds of explanation, and these explanations cannot be captured by a single model) has strong consequences in terms of what I said in the previous chapter[27]. First, if we are pluralists, there is nothing like

[27]I use the term 'pluralism in explanation' in reference to the passage above, quoted from Newton-Smith [Newton-Smith, 2000, p. 132]. Henceforth I will refer to it simply as 'pluralism', and 'pluralists' will be the authors who endorse such a view.

a *unique* objective relevance relation which permits to capture a sense of explanation efficient enough as to cover any instance of MEPP. This point seems to confirm a well-known criticism of theories of scientific explanation, put forward by Paul Feyerabend and concerning the impossibility of having a general objective account of explanation:

> This criterion [the criterion involving "subjective" elements] would seem to be somewhat arbitrary. It is easily seen, however, that it cannot be replaced by a less arbitrary and more "objective" criterion. What would such an objective criterion be? It would be a criterion which is either based upon behavior that is not connected with any theoretical element –and this is impossible– or it would be behavior that is tied with an irrefutable and firmly established theory –which is equally impossible. We have to conclude, therefore, that a formal and "objective" account of explanation cannot be given[28] [Feyerabend, 1962, p. 95]

Second, the following problem arises: if we accept that does not exist a general theory of MEPP, how can we characterize an explanation with respect to another explanation, and more precisely how do we compare explanation E_1 of P_1 from explanation E_2 of P_2, or even explanation E_1 of P_1 from explanation E_2 of the same physical fact P_1? Let me call the latter problem the *incommensurability-problem of explanation*. Is there a way to compare two MEPP (for instance, in terms of their explanatory power) if we have at our disposal such a jungle of perspectives on MEPP?

A further and general remark is important for what follows. Robert Batterman has rightly emphasized as most (if not all) of the recent investigation

[28]Observe that Feyerabend's rejection of objectivity does not entail pluralism. A pluralist does not exclude that there might be different accounts of MEPP based on a "objective criterion". For instance, objectivity expressed by account A captures explanation E_A, while account B is based on a different objective relation which captures explanation E_B, and so on. However, a pluralist would not accept a *unique* objective criterion for explanation. This is the point I want to stress here: to accept pluralism is to reject the existence of a *unique* objective account of MEPP. A conclusion which coincides with that of Feyerabend.

of the notion of MEPP originated in the Indispensability Argument debate [Batterman, 2010]. I have sketched the general lines of this debate in my discussion of Baker's test case (see especially footnote 1 of section 2.1). In order to show that mathematical entities play (or do not play) an indispensable explanatory role in our best science, the interest of much philosophers was directed to cases of MEPP which involved properties of mathematical entities (this was the case, for instance, of Steiner and Baker[29]). The negative effect of that was that other cases of MEPP which do *not* make reference to mathematical entities were left outside. For instance, as we will see, Pincock is interested in "explanation that appeal primarily to formal relational features of a physical system" [Pincock, 2007a, p. 257]. On the other hand, Batterman is concerned with "explanations which involve mathematical operations", rather than mathematical entities or properties of entities [Batterman, 2010, p. 5].

Finally, it should be noted that Pincock's and Batterman's thesis are very recent and they have not been subjected to an intensive criticism yet. Even if we do not dispose of this criticism, however, I will try to put on the table some important points of the discussion which took place around these positions on MEPP. There is no doubt that the perspectives contained in this second part represent fresh and stimulating advancements in the study of MEPP.

Before presenting the accounts, let me add two general observations. The first concerns the methodology adopted by pluralists. The second concerns the consequences that the adoption of pluralism has (or has not) at the level of the ontic/epistemic character of the resulting model. To put the latter point differently: once we adopt pluralism, are we necessarily committed to an ontic (or epistemic) model of explanation?

[29]Baker focused on the property of primeness, while Steiner focused on the property (the 'characterizing property') possessed by particular mathematical objects.

Pluralism and methodology

If we accept that pluralism *is* the way, we are well-suited to fix a methodological question. In chapter 4 I made a distinction between top-down and bottom up methodological approaches to MEPP. A top-down strategy amounts to providing a general encompassing model and then test the model on specific case studies from scientific practice. On the contrary, to proceed bottom-up is to begin the analysis by taking into consideration some cases of MEPP coming from scientific practice and then try to propose a model or some philosophical considerations. We have already seen how Mancosu and Tappenden, in their analysis of Kitcher's unification model, have suggested that the bottom-up procedure might constitute a more fruitful strategy ([Tappenden, 2005, p. 187], [Mancosu, 2008c]). Let me now provide a linkage betwen this methodological point and the pluralist view on MEPP.

If we follow a bottom-up strategy, we do not necessarily have to endorse pluralism. For instance, we might start from a specific set of cases of MEPP as recognized in scientific practice and then search for a single model. On the other hand, if we are pluralist, we are necessarily committed to a bottom-up methodology. This is quite natural, and it is evident if we consider the methodology of the pluralist authors I am going to present in the following two chapters. They always start their analysis from a singular case-study coming from scientific practice, and next they propose their idea of what 'species' of MEPP we are confronted with by looking at the specific test-case itself.

Pluralism and ontic/epistemic commitment

In chapter 4 I introduced a characterization of the accounts in terms of the ontic/epistemic distinction. I defined an account of explanation as ontic when it is based on a relation that is characterized independently from the categories linked with the subject who knows (the understanding is an example of such categories). On the other hand, I defined as epistemic an account which uses a relation which is defined through categories, such as the understanding, which are linked to the subject. We have seen how the adoption of a WTA view on explanation was neutral on the ontic/epistemic character that such a theory of explanation should have, in the sense that a WTA approach might be ontic or epistemic as well. While Van Fraassen's model was much more balanced on the epistemic side, and Steiner on the ontic one, Kitcher's model (as Friedman's) stood between the ontic and the epistemic[30]. It is then interesting to see whether to adopt a pluralist view on explanation might have consequences on the level of ontic/epistemic character that the resulting theory of explanation has. In order to investigate this point, it is sufficient to anticipate here some basic features of the models which will be illustrated in detail in the following two chapters.

Batteman's account is based on the idea that a particular kind of 'reasoning' (asymptotic reasoning) does characterize a particular kind of explanation (asymptotic explanation). As a consequence, his account must be seen as epistemic, because the particular way of reasoning he refers to is dependent on the subject who is doing the explanation. On the other hand, the account proposed by Christopher Pincock is based on the explicit assumption that there exists a (partial or total) mapping with certain structural properties between the world and the mathematical domain. The existence of this mapping (which for Pincock is essential to the MEPP) is independent of us, and therefore it introduces an ontic component into the model[31]. However, as we will see, for Pincock (as for other mapping-account supporters such

[30]Kitcher and Friedman incorporated in their models a linkage between understanding and explanation, where explanation was objectively characterized, i.e. characterized independently from the subject doing the explanation, but understanding was not.

[31]It would be more easy to delineate this ontic component by adopting some terminology that will be introduced in subsection 5.2. The kinds of statements in which we find the occurrence of mathematical together with non-mathematical terms are called "mixed statements". The mapping accounts of the application of mathematics are *external-relation* accounts, i.e. they claim that mixed statements require, for their truth, the existence of an external relation (the mapping relation between the mathematical domain and the world). The picture of MEPP proposed by Pincock is based on such external-relation view, and therefore it incorporates the idea that a mapping between the mathematical domain and the world does exists independently from us. This represents the ontic component of his model.

as Octavio Bueno and Mark Colyvan) mathematical structures play the role of explanatory devices depending upon the context and our pragmatic motivations. This is to say that what counts as explanation also depends on the intentions of the scientists who are doing the explanation in a specific scientific context. The effect of these latter (pragmatic) considerations is to create a tension with the ontic component and shift the characterization of the account towards the epistemic side as well. This is why we can say that, in this case, the resulting account of MEPP has an ontic as well as an epistemic component[32].

The moral of the previous paragraph is that pluralism is neutral towards the ontic/epistemic commitment a theory of MEPP should adopt. Moreover, it is easy to see that the adoption of an epistemic (or ontic) account of explanation is compatible with the pluralist view. Let me make this claim more explicit. Suppose we propose an epistemic account of explanation. This account will use a relation, say R_{ep}, which has been defined by taking into consideration the epistemic access of the subject to a particular state of affairs (for instance, by taking into consideration the capacity the subject has to visualize a state of affairs). In that case, the acceptance of pluralism does not exclude the possibility of having other epistemic or ontic accounts, based on different relations R_{ep2}, R_{ont1}, ... On the other hand, suppose now we propose an ontic account of explanation. This account will be based on a relation which is characterized independently from the subject, for instance it will be based on an objective relation R_{ont} such as that offered by Steiner[33]. Also in this case we can be pluralists simply by accepting the idea that our account does capture one particular objective relation R_{ont}, but others objec-

[32]Observe that to investigate the ontic/epistemic characterization of the accounts does not mean to investigate the realist/anti-realist ontological commitment of the authors towards mathematical entities or structures (the latter point will be discussed in subsection 5.2.1). The ontic/epistemic characterization that I am discussing here is something which concerns the account itself. However we find that, in cases such as that of Van Fraassen's or Batterman's, to the epistemic account proposed there corresponds an anti-realist position of the author.

[33]In the previous chapter I have showed how Steiner's account can be read in terms of an objective relevance relation.

tive or epistemic relations (for instance R_{ont2}, R_{ep1} and so on) are captured by other accounts of explanation as well. The pluralist option is thus compatible with ontic as well as with epistemic accounts of MEPP.

Chapter 5

Christopher Pincock: mapping accounts and MEPP

We use mathematics to represent and study our world. For instance, we use a mathematical model to represent a piece of actual world (a target system) and then we study it mathematically, possibly improving the mathematical model by making further considerations about the actual system under study. However, the effectiveness of this practice leaves us with a question: How does mathematics apply to the world? This is the famous "unreasonable effectiveness of mathematics in the natural sciences" to which Eugene Wigner referred to in his [Wigner, 1960]. One possible line of answering this question is to say that we can account for the applicability of mathematics to the world by demonstrating the existence of the right kind of mapping from a mathematical structure to some appropriate physical structure. The philosophical accounts that try to describe how this mapping works (i.e. how the structural correspondences between some structural aspects the target system and the correspective objects in the mathematical structure are established) have been called "mapping accounts" [Pincock, 2004a].

The motivation which stands behind these studies could be traced back to some worries which were not adequately faced by the semantic view on scientific representation. As observed by Christopher Pincock:

> The traditional advocates of the semantic view seemed to downplay the distinctively mathematical nature of the descriptions that were used to pick out their models. Here they followed much of the foundational tradition in the philosophy of mathematics in assuming that first-order logic and set theory were sufficient to characterize scientific models and their representational relationship to the world [Pincock, 2011c, p. 328-329]

Right now, two accounts are representative of the mapping view: Christopher Pincock's approach to the application of mathematics ([Pincock, 2004b] and [Pincock, 2007a]), and Octavio Bueno and Mark Colyvan's "inferential conception of the application of mathematics" [Bueno *et al.*, 2011][1]. Their attempt to account for how the mapping from an actual target system to the mathematical structure works is particularly interesting for my study. This is because of the emphasis they put on the explanatory role that mathematics can have in the process. By not endorsing the idea that to speak of explanation is to speak of causal aspects of the phenomena, these authors leave room for MEPP.

A fine-grained look at the representational problem (the problem of representing the world mathematically) permits to single out at least two distinct questions: the semantic question of whether or not a mathematical representation is about a target system, and the question of the respects in which the mathematical representation is an accurate representation of the target system. Naturally, the latter question has a subquestion concerning the accuracy of the representation in terms of the details of the target system: how many details do we have to mirror in our mathematical model to improve the accuracy? If there exists a dependence between explanatoriness and representativeness, as some partisans of the mapping accounts seem to suggest,

[1]Let me add that the mapping accounts can be traced back to the 'projectivist' family in theory of representation, and in particular to one of its origins in the works on representational measurement theory by Suppes and collaborators ([Krantz *et al.*, 1971], [Suppes *et al.*, 1989], [Lute *et al.*, 1990]), which focuses on structural similarities between empirical and mathematical systems. The approach here, though, is quite more general for the similarity is not restricted to empirical comparison systems as it is the case in measurement theories.

the latter question is relevant to our investigation.

In the first section of this chapter, I am going to present a famous problem, the Königsberg bridges problem, whose solution was found by Euler in 1735. In particular, I will show how a solution to this problem can be given in the modern language of graph theory. Next, I will present Pincock's account of abstract explanations in the context of his mapping account view, as offered in his papers [Pincock, 2004b] and [Pincock, 2007a]. As we are going to see, in order to illustrate his idea of abstract explanations, Pincock discusses the modern solution to the problem originally solved by Euler. Section 5.3 will contain a discussion of how mapping accounts accommmodate the role of idealizations and how the idea that idealizations can play explanatory roles can be defended. Finally, in the last section of the chapter (section 5.4), I will address the following question: Is representation a necessary condition for explanation?

Before starting with Euler and the Königsberg bridges let me observe that, although Bueno and Colyvan's inferential conception of the application of mathematics is seen by the authors as an attempt to improve Pincock's model of application of mathematics [Bueno *et al.*, 2011, p. 371 endnote 10], I will not consider their position in my analysis. This choice is motivated by the fact that Bueno and Colyvan do not offer a precise characterization of what they consider as "explanation", but only a general and intuitive idea of how MEPP can be accommodated within their model of application [Bueno *et al.*, 2011, p. 353]. On the other hand, Pincock's treatment of explanation is explicitly aimed to capture a particular instance of MEPP (abstract explanation) and thus it offers a more precise and workable notion.

5.1 A walk across the seven bridges of Königsberg. Does mathematics help?

Let's begin with a famous recreational problem concerning the seven bridges of Königsberg. In the 18th century the city of Königsberg consisted

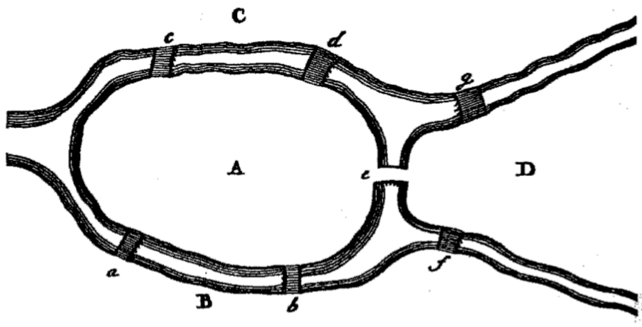

Figure 5.1: The seven bridges' diagram as appeared in Euler's 1736 paper.

of four land areas separated by branches of the river Pregel over which there were seven bridges[2]. The problem was the following: is it possible for somebody to make a complete tour of the town crossing each of the seven bridges exactly once and returning to the starting point?

In 1727, Leonhard Euler began working at the Academy of Sciences in Saint Petersburg. He presented a paper to his colleagues on 26 August 1735 on the solution of "a problem relating to the geometry of position". The problem to which Euler referred to was the Königsberg bridges problem[3]. He proved that it was *not* possible to plan a route that would cross each of the seven bridges of Königsberg exactly once, whether or not you ended up in the same place as you began[4].

[2]Königsberg, along with the rest of northern East Prussia, became part of the Soviet Union (now Russia) at the end of World War II and was renamed 'Kaliningrad'. The river Pregel was renamed 'Pregolya' and only two original bridges still survive, while a third bridge is a version rebuilt by Germans in 1935. See [Taylor, 2000] for the story of the seven bridges.

[3]The solution to the problem appears in 1736, in Euler's paper in the *Commentarii Academiae Scientiarum Imperialis Petropolitanae* [Euler, 1736]. Although dated 1736, Euler's paper was not actually published until 1741.

[4]Euler reformulated the problem as one of trying to find a sequence of letters of length eight, containing only letters A, B, C, D (the land areas in Figure 5.1), such that the pairs AB and AC are adiacent twice (corresponding to the two bridges between A and B and A and C), and the pairs AD, BD, and CD are adiacent just once. He showed by a counting argument that no such sequence exists, thus proving that there is no solution to the Königsberg problem, i.e. it is not possible to make such a tour. More precisely, Euler showed that if there are more than two areas to which an odd number of bridges

Imagine that Kant sets out to prove Euler's point by walking back and forth across the seven bridges. The reason why he, as anyone else, was not able to cross the bridges once and return to the starting point, is provided by Euler's mathematical result. This reason is, of course, far from being a "causal" reason. Now, Christopher Pincock focuses on the modern solution, in terms of graph theory, to the seven bridges problem and considers as explanandum the impossibility of walking the desired route which crosses each of the seven bridges exactly once:

> As my example I take an explanation of why it was impossible to walk a certain kind of path across the bridges of Königsberg [Pincock, 2007a, p. 257]

Before seeing what is considered an explanation by Pincock, let me introduce some notions of graph theory[5].

A *simple graph* G consists of a non-empty finite set $V(G)$ of elements called *vertices* (or *nodes*, or *points*) of the simple graph, and a finite set $E(G)$ of distinct unordered pairs of distinct elements of $V(G)$ called *edges* (or *lines*). The set $V(G)$ is called the *vertex set* of G, while $E(G)$ is the *edge set* of G. An edge $\{v, w\}$ is said to *join* the vertices v and w, and is usually abbreviated to vw.

In any simple graph there is at most one edge joining each pair of vertices.

lead, then the 'Eulerian journey' is impossible [Hopkins et al., 2004, p. 418]. In his solution, Euler made no mention of graphs. The connection between the Königsberg bridges problem and diagram-tracing puzzles was not recognized until the end of the 19th century. It was pointed out by the British mathematician (and amateur magician) W. W. Rouse Ball in his book *Mathematical Recreations and Problems of Past and Present Times* [Rouse Ball, 1892]. Rouse Ball seems to have been the first to use a diagram to solve the problem (the diagram he used is that of Figure 5.4). For an historical discussion of Euler's solution and the development of the present-day solution see [Hopkins et al., 2004], [Sachs *et al.*, 1988] and [Wilson, 1986]. In what follows I do not consider Euler's original solution but only the modern one in terms of graph theory. This choice is motivated by the fact that Pincock's discussion (and his account of abstract explanations) is illustrated in terms of the modern solution.

[5] In presenting the basic notions of graph theory I follow Robin J. Wilson's *Introduction to Graph Theory* [Wilson, 1996]. For an introduction to graph theory see also [Diestel, 2005], while for a general handbook on graph theory (including some historical aspects and curiosities) see [Gross *et al.*, 2004].

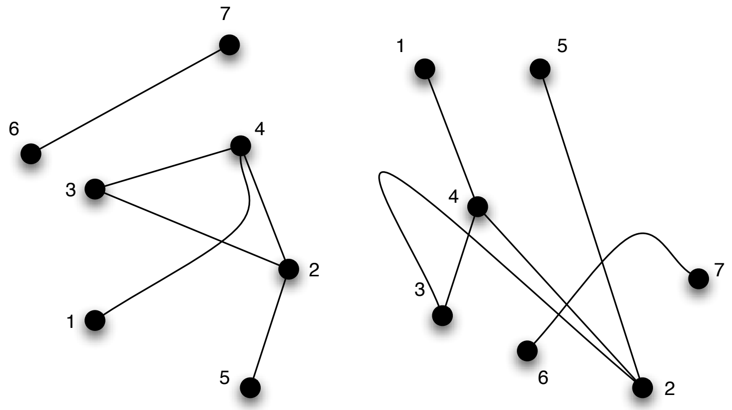

Figure 5.2: Two representations of the same graph G^* of order $|G^*| = 7$ and number of edges $||G^*|| = 6$.

However, many results that hold for simple graphs can be extended to more general objects in which two vertices may have several edges joining them (we speak of *multiple edges*). In addition, we may remove the restriction that an edge joins two distinct vertices, and allow edges joining a vertex to itself (*loops*). The resulting object, in which loops and multiple edges are allowed, is called a *general graph* or, simply, *a graph*. Thus every simple graph is a graph, but not every graph is a simple graph[6]. Thus, a *graph* G consists of a non-empty finite set $V(G)$ of elements called vertices, and a finite family $E(G)$ of unordered pairs of (not necessarily distinct) elements of $V(G)$ called edges. Here the word 'family' means a collection of elements, some of which may occur several times (for example, $\{a, b, c\}$ is a set, but (a, a, c, b, a, c) is a family). Note that the use of the word 'family' permits the existence of

[6]Observe that the language of graph theory is not standard and many authors adopt their own terminology. For instance, some authors use the term 'graph' for what I defined as a simple graph, and then speak of *multigraph* for a graph which has more than one edges between two same vertices (this terminology is adopted in [Diestel, 2005]). Here I follow Wilson's terminology in his [Wilson, 1996] and therefore I include the condition of multiple edges between two vertices (a condition which corresponds to our seven bridges scenario) into the definition of general graph (or graph). This notation seems to be much more clear, and it makes things easier for us.

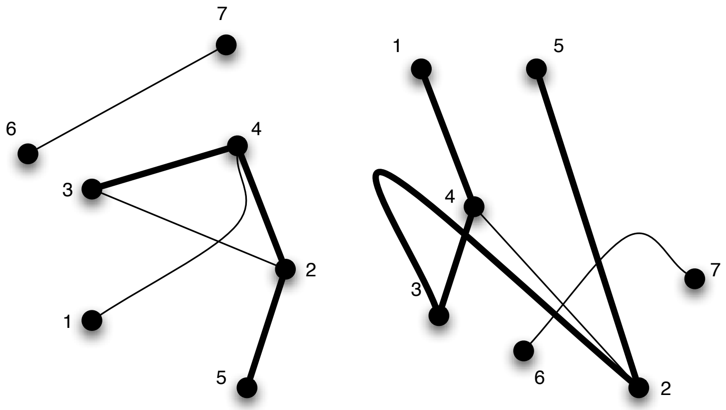

Figure 5.3: Two walks on G^*: 3425 and 14325.

multiple edges[7]. We call $V(G)$ the vertex set and $E(G)$ the *edge family* of G. An edge $\{v, w\}$ is said to join the vertices v and w, and is again abbreviated to vw. The number of vertices of a graph G is its *order*, written as $|G|$, while its number of edges is denoted by $||G||$. A graph whose $E(G)$ is empty is a *null graph*, while a graph in which both the edge family and the vertex set are empty is called the *empty graph*.

Usually, a graph is pictured by drawing a dot for each vertex and joining two of these dots and lines if the corresponding vertices form an edge (if there is more than one edge between two vertices, we draw two lines). The way in which these dots and lines are drawn is irrelevant. What is important is the information of which pairs of vertices form edges and which no. For instance, the two diagrams in Figure 5.2 represent the same graph G^* with $V^* = \{1, ..., 7\}$ and edge family $E^* = (14, 23, 24, 25, 34, 67)$.

The *degree* (or *valence*) of a vertex v of G is the number of edges incident

[7]Here (again) I follow Wilson's textbook [Wilson, 1996] and I adopt his notation in what follows. Nevertheless, let me remark that the multiedge condition (more than one edges between two same vertices) can be given in a more formal (and more orthodox) way. For instance, if we define a graph as a pair $G = (V, E)$ of sets such that $E \subseteq \left[V\right]^2$ (where the elements of the set V are vertices and the elements of the set E are edges), we can say that a multigraph is a pair (V, E) of disjoint sets (of vertices and edges) together with a map $E \longrightarrow V \cup [V]^2$ assigning to every edge either one or two vertices.

with v, and is written $deg(v)$ (in the example in Figure 5.2, $deg(4) = 3$)[8]. We say that two distinct edges are *adjacent* if they have a vertex in common (in G^*, the edges 52 and 24 are adjacent because they have vertex 2 in common). Given a graph G, a *walk* in G is a finite sequence of edges of the form $v_0v_1, v_1v_2, ..., v_{m-1}v_m$, in which any two consecutive edges are adjacent or identical. Such a walk determines a sequence of vertices $v_0, v_1, ..., v_m$, and we speak of *a walk from* v_0 *to* v_m with *initial vertex* v_0 and *final vertex* v_m. We often refer to a walk by the natural sequence of its vertices. For instance, when we write $v_0v_1v_2...v_m$, this is a walk from v_0 to v_m. The number of edges in a walk is its *length*. In Figure 5.3 I traced two possible walks for the graph G^*: 3425 of length 3 (on the left) and 14325 of length 4 (on the right). A walk in which all the edges are distinct is a *trail*. If, in addition, the vertices $v_0, v_1, ..., v_m$ are distinct (except, possibly, $v_0 = v_m$), then the trail is a *path*. In other words, a path is a walk in which all the edges and all the vertices are distinct, with the only possible exception $v_0 = v_m$ (for instance, the two walks 3425 and 14325 on G^* are two paths). If $v_0 = v_m$ we say that the path or trail is *closed*. A graph is said to be *connected* if and only if there is a path between each pair of vertices. For instance, the graph G^* is not connected because there is no path between vertex 3 and vertex 6.

A connected graph G is *Eulerian* if there exists a closed trail containing *every* edge of G. Such a trail is called an *Eulerian trail*. Observe that this definition requires each edge to be traversed once and once only.

Now, with the aid of the previous notions, let's come back to our promenade across the Königsberg bridges. If we treat the islands and the banks (A, B, C, D in Euler's original diagram in Figure 5.1) as objects labeled as $1, 2, 3, 4$, and we use the bridges to form edges a, b, c, d, e, f, g, the physical configuration of the seven bridges is represented by the graph G_b with $V_b(G_b) = \{1, 2, 3, 4\}$ and $E(G_b) = (12, 12, 42, 42, 13, 23, 43)$. This graph, which is pictured in Figure 5.4, is connected because there is path between each pair of vertices. Therefore the seven bridges problem can be rephrased

[8]In calculating the degree of a vertex v, we usually make the convention that a loop at v contributes 2 (rather than 1) to the degree of v.

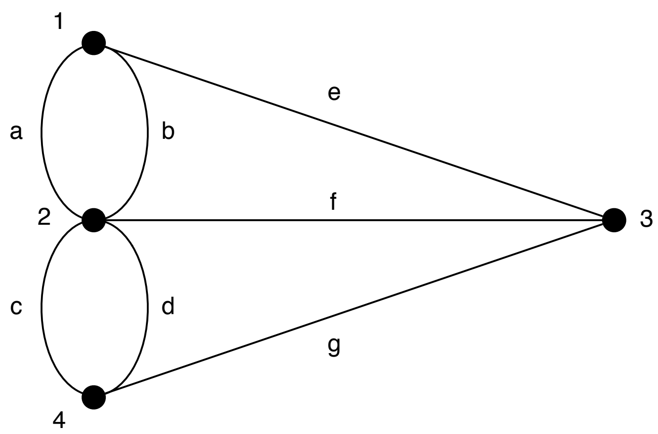

Figure 5.4: Picture representing the graph G_b. This graph formalizes the seven bridges problem.

as follows: Is the connected graph G_b Eulerian? (Or, which is the same: Does the connected graph G_b admit an Eulerian trail?).

Although the solution to the seven bridges problem was given by Euler in his 1736 paper (not in terms of graph theory!), in every textbook of graph theory the solution to our puzzle is given in terms of the following theorem:

Theorem 5.1. *A connected graph G is Eulerian if and only if the degree of each vertex is even* [9]

The impossibility of crossing all the bridges exactly once and return to the starting point is then stated by previous theorem, because at least one of the vertices of our graph G_b has an odd degree (more precisely, they *all* have odd valence!), and then the graph is *non-Eulerian*. Therefore, we may conclude (as did Euler) that the attempts by the residents of Königsberg were in vain. Conversely, a graph such that pictured in Figure 5.5 is Eulerian. This means that, if the Königsberg scenario had been that corresponding to Figure 5.5, it would have been possible to cross all the bridges exactly once and return to the starting point.

[9] For a proof see [Wilson, 1996, p. 32].

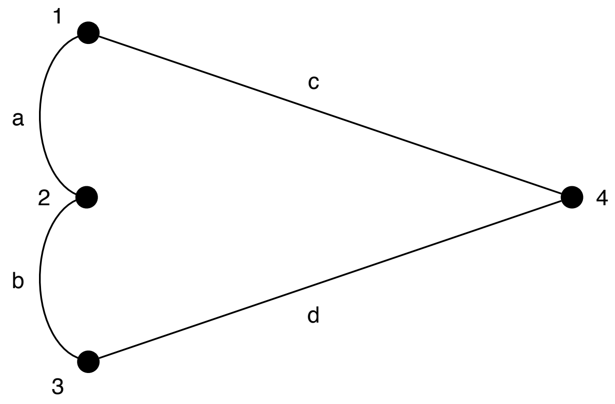

Figure 5.5: Representation of an Eulerian graph, where the Euler tour is possible.

5.2 Pincock's abstract explanations

In his paper "A role for mathematics in the physical science" [Pincock, 2007a], Pincock considers the seven bridges problem as an illustration of a particular kind of explanation: *abstract explanation*.

As we have seen in the previous section, in the seven bridges problem the explanandum is not a physical phenomenon in the classical sense (for instance, the particular behavior of a physical system such as a gas or the motion of a planet around the Sun), but a particular actualizable situation which depends on an actual physical system (the impossibility of making the Euler tour across the seven bridges). This is what Pincock takes as the explanandum X: the fact that it is impossible to walk a certain kind of path across the bridges of Königsberg.

But what exactly is the 'explanation' for the case considered? Pincock regards the property of vertices of having odd valence as providing an explanation of X:

> If I was asked to explain why it is impossible to make such a crossing, then I would appeal to the fact that one of the vertices has an odd valence. [Pincock, 2007a, p. 259]

> I claim that it is a fact about the bridges of Königsberg that they are non-Eulerian and that an explanation for this is that at least one vertex has an odd valence. Whenever such a physical system has at least one bank or island with an odd number of bridges from it, there will be no path that crosses every bridge exactly once and that returns to the starting point. If the situation were slightly different, as it is in K' [Eulerian graph], and the valence of the vertices were to be all even, then there would be a path of the desired kind. [Pincock, 2007a, p. 259]

Naturally, to know that at least one vertex has an odd valence would be meaningless for our explanation without the specific knowledge of theorem 5.1[10]. However, the fact to consider that it is a property of vertices, and not the particular theorem, which provides the essential ingredient in the explanation, makes a considerable difference. To appeal to the fact that the number of vertices/banks is even is to appeal to a structural property of the actual physical system (the system bridge-banks). This kind of explanation has, according to Pincock, "different features from the explanations involving coordinate systems" [Pincock, 2007a, p. 257]. In fact, in the case of the seven bridges problem, without using any coordinate system or unit of measurement, we are offering an explanation which draws only on a formal relational feature of the system bridges-banks: the parity of the number of vertices, where vertices/banks are defined in relation with edges/bridges. Now, an essential ingredient in order to have our mathematical explanation of the path restrictions is the procedure of mapping *some* of the structural (actual) relations between parts of the bridge-system into the graph structure. For instance, while being a structural feature, the fact that one bridge is longer than the others is irrelevant to the mapping relation and does not figure among the elements of our representational mathematical graph. Although the actual bridge-system is *not* a graph, Pincock maintains that there

[10] For instance, Pincock writes: "Knowing Euler's theorem, I could now explain why we could cross all the bridges exactly once and end up at the starting point by appealing to the fact that all the vertices have an even valence" [Pincock, 2007a, p. 259].

is a structural similarity between the bridge-system and the graph: "[...] the bridge system has the structure of a graph, in the sense that the relations among its parts allow us to map those parts directly onto a particular graph" [Pincock, 2007a, p. 260]. This is why, for him, to explain why an Euler tour *is* possible would amount to saying that all the vertices have an even valence (situation of Figure 5.5). This explanation is a particular kind of explanation, which he calls *abstract explanation*:

> By an abstract explanation I mean an explanation that appeals primarily to the formal relational features of a physical system. Some abstract explanations that employ mathematics seem to qualify as intrinsic explanations. This is because even though they can be thought of as involving mappings between a physical system and a mathematical domain, these mappings do not turn on any arbitrary choice of units, but concern only the intrinsic features of the systems represented [Pincock, 2007a, p. 257]

As Pincock observes, this kind of explanation seems superior to other kinds of explanations that might be given in terms of the microscopic configuration of the system. For instance, consider that we have an explanation of the impossibility to make our Euler tour across the bridges in terms of the microscopic configuration of the bridge-system. Can we use the same explanation if we turn the bridges into silver? The answer is no because the microscopic configuration will be altered. On the other hand, our explanation in terms of the parity of the vertices would still work. Therefore it seems that "abstract explanation seems superior because it gets at the root cause of why walking a certain path is impossible by focusing on the abstract structure of system" [Pincock, 2007a, p. 260].

In passing, let me note that Pincock considers abstract explanations as a particular species of *structural explanations*, where the latter are defined by Ernan McMullin:

> When the properties or behavior of a complex entity are explained by alluding to the structure of that entity, the resultant explanation may

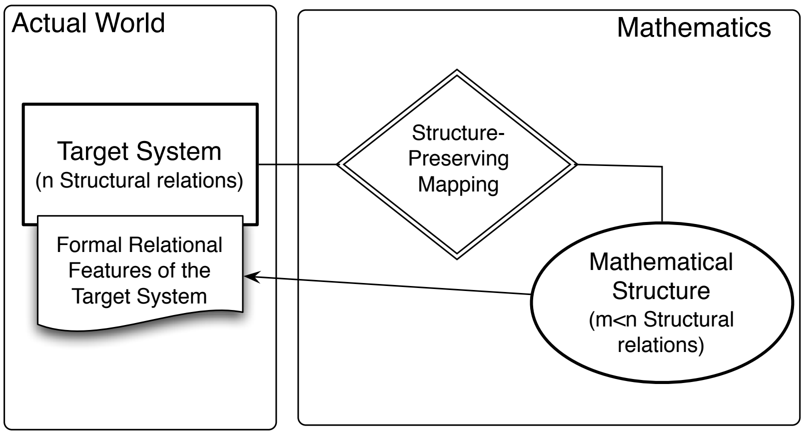

Figure 5.6: Pincock's abstract explanations are explanations which appeal to formal relational features of the physical system.

> be called a *structural* one. The term 'structure' here refers to a set of constituent entities or processes and the relationships between them. [McMullin, 1978, p. 139]

We can say that Pincock's idea is that the representational capacity of the graph (which, keep in mind, is a mathematical entity!) is a source of explanatory power due to the ability of the graph to pick out structural relational features of the actual system (as my diagram in Figure 5.6 suggests). Note how this idea is very close to Margaret Morrison's: mathematical models are explanatory because they "exhibit certain kinds of structural dependences" [Morrison, 1999, p. 63]. Hovewer, there are some differences between Pincock's and Morrison's view. Perhaps the most evident, and also the most relevant for our discussion, is that while for Pincock the representational capacity is defined in terms of mapping relation of structures ('mapping criterion'), and explanatory power is the *result* of this representational capacity, for Morrison the explanatory power of a model is a *function* of its representational feature where the nature of the representation cannot be uniquely characterized. In Morrison there is nothing like a mapping criterion (models are seen as "autonomous" agents which mediate between theories and phenomena). Again, for Pincock the 'picking out' of structural relational features of the actual system is possible because there is some structure-preserving mapping between the world and the mathematical structure in question, and then the mapping is essential to the explanatory role of the representation[11]. Observe here that the mapping is itself a mathematical object, therefore to say (as Pincock suggests within his account) that the structure-preserving mapping is essential to his abstract explanations amounts to saying that we are considering a case of MEPP.

There is an additional story about Pincock's structuralist position, which concerns the interpretation of statements arising in applied mathematics and which is supposed to give a possible ground to his approach to the applicability of mathematics. This story is relevant to my discussion for two reasons: it complements his view on the applicability of mathematics, thus providing a more distinct picture of the framework in which his abstract explanations are discussed; it permits me to present the ontological commitment which follows from his structuralism, and to show how abstract explanations help Pincock to defend his ontological position (once more, we will see how the

[11] Observe that in his [Pincock, 2007a] Pincock only considers isomorphisms and homomorphisms as the mapping to be used in his structuralist proposal. An *homomorphism* is a mapping from one structure A (with a domain D, and a family of relations R among the elements of D) to another structure B (with codomain D', and a family of relations R' among the elements of D') that respects the relations of A by assigning each element of R to a corresponding element of R'. That is, an homomorphism maps not only the objects of one domain to another; intuitively, it does so in such a way that preserves certain aspects (although typically not all) of the structures involved. An *isomorphism* is a bijective homomorphism, i.e. an homomorphism which is injective and surjective (it is one-to-one and every member of the codomain is the image of at least one member of the domain). As pointed out by Bueno and Colyvan, Pincock's silence on the potential use of different types of mapping comes presumably from his conviction that the kind of mapping changes depending on the application in question [Bueno *et al.*, 2011, p. 347]. This observation fits well with what Pincock himself says in another article: "Each kind of application, of course, needs a different kind of a mapping, but it does offer a unified account of mixed statements." [Pincock, 2004b, p. 150]. What about the status of these mappings? Pincock considers mappings involved in the applications of mathematics as relations, with the requirement that relations bear an external relation to their relata and these relations exist independently of whathever entities they happen to relate (the same relation exists regardless of which objects it has as relata). He points out how this demand is fulfilled by adopting an account of relations (and a criterion of identity for intensional entities) such as that developed by George Bealer in model theoretical terms [Bealer, 1982].

topic of MEPP comes as crucial ingredient in the ontological debate in philosophy of mathematics).

In his [Pincock, 2004b], Pincock focuses on the content and the meaning of statements such as 'The mass of the satellite is 1500Kg'. Those kinds of statements are called *mixed statements* because in them we find the occurrence of mathematical terms together with non-mathematical terms. Naturally, to think about the meaning of mixed statements is to think about some relationship between the actual object (the satellite), the actual standard gram, the mathematical term (the real number 100). Pincock individuates three approaches which target the meaning of mixed statements and which claim to solve the problem of applicability:

α *No-relation* account: this view denies the existence of any relation which connects the mathematical and the physical domain. For instance, fictionalists such as Hartry Field, considers true a claim like 'The mass of the satellite is 1500Kg' but deny that there exists a relation between the satellite, the gram and the real number 1500. According to Field, the truth of statements like that depends only on properties of the satellite and its physical relations to other physical objects[12].

β *Internal-relation* account: according to this view mixed statements require (for their truth) that an internal relation obtain between the mathematical domain and the physical world. An internal relation for an object a is a relation that a must stand in if it is to be that particular object: a stands in an internal relation R to b just in case aRb's obtaining is involved in a's criteria of identity. According to Pincock, an example of internal relation is that between a set and its members[13], while an example of internal-relation account is offered by Frege's definition of natural numbers in terms of their paradigm application to counting [Frege, 1980][14].

[12]Remember that Field denies the existence of mathematical objects and argues for the view that mathematics is dispensable from science [Field, 1980].

[13]For instance, if b and c are members of the set S, the set S bears an internal relation to b and c. A set which has different members will be different from S.

[14]Pincock also considers Steiner's approach to application expressed in his book *The Applicability of Mathematics as a Philosophical Problem* [Steiner, 1998] as an internal-

γ *External-relation* account: in this picture, mixed statements require (for their truth) the existence of an external relation between the mathematical domain and the physical situation, i.e. a relation that does not involve the criteria of identity of the mathematical objects.

Both the internal-relation and the external-relation accounts accept that mixed statements depend on a relation between mathematical and physical domains for their truth. However, they differ in what is considered the *right* kind of relation (internal or external). The mapping-structuralist approach proposed by Pincock in order to account for mixed statements and for the application of mathematics is an external-relation account. It is structuralist because "the truth of mixed statements depends on the existence of a mapping with certain structural properties, and these properties can be picked out in structural terms" [Pincock, 2004b, p. 146][15]. For instance, the mixed statement 'there are five apples on the table' comes true just in case there is an isomorphism, i.e. a mapping preserving cardinality and structure, from apples to the initial segment of the natural numbers ending with 5.

There is a feature that external-relation accounts have with respect to no-relation and internal-relation accounts, and concerns what Pincock calls the "Dummett's dilemma" [Pincock, 2004b, p. 141]. The dilemma consists in the choice between the two desiderata[16]:

relation account, although observing that some of Steiner's remarks point to a non-internal or perhaps to an hybrid solution [Pincock, 2004b, p. 142-143].

[15]Pincock notes that, while being compatible with, its structuralism is independent of a structuralist account of pure mathematics, i.e. the view that the subject matter of mathematics is abstract structures. This structuralist view is maintained by Stewart Shapiro [Shapiro, 1997] and Michael Resnik [Resnik, 1997]. In passing, let me observe that when structuralism in pure mathematics is demanded to account for the applicability of mathematics to reality, the two structuralisms (Pincock's and Shapiro and Resnik's) are more than compatible and seem to converge. For instance, Shapiro writes: "According to the structuralist, the application of mathematics to science occurs, in part, by discovering or postulating that certain structures are exemplified in the material world. Mathematics is to material reality as pattern is to patterned" [Shapiro, 2005, p. 21]. In the latter quotation from Shapiro, the process of "exemplification" presupposes a mapping.

[16]In his paper "What is mathematics about?" [Dummett, 1993], Michael Dummett has observed how the two desiderata are often in conflict. Here is Dummett's original passage quoted by Pincock: "The difficulty about mathematical objects thus arises because we want our mathematical theories to be pure in the sense of not depending for the existence

(1) we want an account of mathematics that guarantees the applicability of mathematics to the world

(2) we want to preserve the necessity of mathematical truths

The difficulty for no-relation and internal-relation accounts in satisfying both of the two desiderata is noticeable. By providing a direct relation which links mathematical objects to the world, an empiricist philosophy of mathematics will guaranteee (1) at the cost of sacrifying (2). On the other hand, an account of mathematics that would preserve the necessity of mathematical truths might be in danger of offering only a weak relation between mathematics and the physical world. According to Pincock, an external relation account "promises to resolve Dummett's dilemma" [Pincock, 2004b, p. 145]. What is then the feature that an external-relation account has with respect to these positions (and which would permit to solve the dilemma)?

Recall that, in an external-relation account, mixed statements require (for their truth) the existence of an external relation between the mathematical domain and the physical situation (a relation which does not involve the criteria of identity of the mathematical objects). In Pincock's structuralism, mixed statements require (for their truth) the existence of a mapping with some structural properties. Now, according to Pincock, the crucial feature of external-relation accounts lies in the fact that, if based on a genuine external relation, such accounts potentially offer a connection between the physical world and mathematics which does *not* affect the necessity of mathematical truths (thus satisfying both the desiderata above). Nevertheless, to solve the dilemma amounts to showing that the external-relation account is able to provide enough external relations between the physical and mathematical objects to ensure the applicability of mathematics to science. Concerning this point, Pincock's claim is that his structuralist approach, if generalized further, would provide such relations (in terms of mappings), thus preserving

of their objects on empirical reality, but yet to satisfy axioms guaranteeing sufficiently many objects for any application that we may have occasion to make" [Dummett, 1993, p. 437-438].

the applicability and the independence of mathematics[17].

5.2.1 Pincocok's structuralism and ontological commitment

Let me spend some further words on what Pincock considers the epistemic utility of applied mathematics and the ontological commitment which follows from his structuralist position. In this discussion, as we are going to see, his idea of abstract explanation plays a very important role. Moreover, the following lines are intended to show that (and *how*) the notion of MEPP is deeply involved in the ontological debate between platonists and nominalists.

Keep in mind that in the seven bridges problem there is nothing like a causal-history of the system, and causal relationships (whatever they are!) and unit of measurement do not have any influence in elaborating the mathematical solution to the problem. The mathematics of graph theory offers an explanation of X (which is an actual situation, and then a sort of physical phenomenon) although it is no possible to trace a causal history of X or individuate a causal relation relevant to our problem (for instance, as we have seen, the interactions between the molecules of the bridges are largely irrelevant). Therefore mathematics does play an essential role in the explanation provided. More precisely, to put it in Pincock's structuralist language, this abstract explanation appeals to mappings that do not turn on any arbitrary choice of units but concern only the intrinsic features of the systems represented.

Now, according to Pincock, his structuralist approach to the application of mathematics defends the theoretical indispensability of mathematics from attacks such as that of Hartry Field. As it is well-known, Field argued not

[17]The independence of mathematics is preserved because, differently from what an empiricist philosophy of mathematics would require, the relation used in evaluating the truthful of mixed statements in Pincock's account (as in every external-relation account) is independent from what happens in the world (roughly, the world does not affect mathematical objects), and therefore mathematical truths are preserved [Pincock, 2004b, p. 153]

only for the metaphysically dispensability of mathematics (i.e. the view that mathematical entities have no relevant causal role in the actual world), but also for its theoretically dispensability. In fact, he tried to show that mathematics is theoretically dispensable by offering non-mathematical versions of some parts of classical mechanics [Field, 1980]. His program is based on a conception of the application of mathematics according to which the ontology includes points and regions of space time (which are, in Field's opinion, concrete entities[18]), and every non-mathematical conservative theory to use in his representational theorems will invoke only such space-time regions and points (while magnitude will be defined as properties of such entities).

Against Field's theoretical dispensability, Pincock points out how, in the case of the graph-solution to the bridge problem, Field's strategy would be impossible to apply because the treatment of the system in terms of graph theory does not make reference to points or regions of space-time. A similar line of criticism against Field's strategy was advanced by David Malament in his review of Field's *Science Without Numbers* [Field, 1980]. More precisely, Malament pointed out how scientific theories formulated in terms of phase-spaces cannot be handled by Field, and this exactly because these theories represent abstract objects and *not* space-time points or regions [Malament, 1982, p. 533][19]. A recent revival of this criticism, advanced in connection with MEPP, has been proposed by Aidan Lyon and Mark Colyvan [Lyon *et al.*, 2008]. They picture a possible nominalist response to Malament: phase-space theories do not add any new physics to the picture, and so any physical law that can be stated in terms of points in phase space has an equivalent nominalist counterpart (which we do not have at the present day!). However, even if we grant this response, they add that there is an im-

[18] Although considering points in space time as *concrete* entities, Field does not trace a clear distinction between abstract and concrete objects [Malament, 1982, p. 532].

[19] Points in phase spaces are abstract objects that represent other abstract objects (possible dynamical states), whereas the points in manifolds are abstract objects that represent other concrete objects (points in space-time). So, applying the kind of representation theorems that Field employs to deal with Newtonian space-time, clearly will not do the trick for the case of phase-spaces theories.

portant role that phase-space theories play in science apart from their ability to provide a neat expression of the relevant laws of physics: phase-space theories are considered to play an explanatory role [Lyon *et al.*, 2008, p. 7]. The point is of crucial importance if we note that Field accepts the principle of inference to the best explanation [Field, 1989, p. 15-16]. So, according to Lyon and Colyvan, any nominalist reformulation of our scientific theories that he provides must have at least the same explanatory resources as the platonist counterparts. But this, again, seems to be impossible in cases such as that of phase-space theories, which represent mathematical explanations to which there is non-mathematical counterpart [Lyon *et al.*, 2008, p. 3].

In situations like these (the graph-solution to the bridges problem or the use of phase-space), Pincock observes, mathematics appears to be theoretically *in*dispensable while metaphysically *di*spensable. The metaphysically dispensability comes from the fact that, even if mathematics has an essential and epistemic role in the formulation of theories which are confirmed by the evidence, the process of abstraction involved in a situation as that of the seven bridges problem does not attribute to mathematical entities any relevant causal role in the world. This traces a dividing line between Lyon and Colyvan's position and Pincock's. In fact, while Lyon and Colyvan attack Field and maintain that mathematics is theoretically *and* metaphysically indispensable, Pincock considers that cases such as the seven bridges problem show that mathematics is theoretically indispensable although being metaphysically *di*spensable. There is, for Pincock, an epistemic utility in using mathematics in cases such as the seven bridges problem. However, this essential epistemic role of mathematics does not entail its metaphysical indispensability. Let me conclude this section by disentangling better the difference between Pincock's ontological commitment and Colyvan's.

For Pincock, the structural approach to the application of mathematics requires the truth of applied mathematical assertions, which are confirmed within mathematics itself, but it does not require that mathematical entities exist. He points out that mathematics plays its role of being theoretical indispensable when it is confirmed *prior* to its application to science. Therefore he remarks that its point in favor of a reconciliation of theoretical indispensability with metaphysical dispensability must not be read as an attempt to argue for platonism on the basis of indispensability arguments [Pincock, 2007a, p. 255][20]. On this point the difference with Mark Colyvan is evident. While they both share Malament's criticism against Field, contrary to Pincock, Colyvan is favourable in considering the explanatory power of mathematics in scientific theories as an instrument to infer the existence of mathematical entities and to 'embrace platonism' [Colyvan, 2002, p. 69][21]. In giving a MEPP this essential role for supporting his realist view, he is adopting the so called "enhanced indispensability argument":

1 We ought rationally to believe in the existence of any entity that plays an indispensable explanatory role in our best scientific theories.

2 Mathematical objects play an indispensable explanatory role in science.

---

3 Hence, we ought rationally to believe in the existence of mathematical objects.

I will come back to the enhanced indispensability argument in the final part of the dissertation, where I will maintain that, contrary to what Colyvan thinks, the inference obtained through this argument is not viable. For the present discussion, let me add that, for Pincock, an abstract explanation such as that given in terms of graph theory in the case of the seven bridges cannot be used to infer the existence of some mathematical entity.

[20]But what about confirmation of mathematics *prior* to its application? Pincock does not offer a precise definition of how this confirmation should be made: "Ideally my account of the role of mathematics in physical theories would be supplemented with a story of exactly how mathematics is confirmed by mathematicians. I do not have such a story ready to hand, and so must fall back on the naturalistic premise that if mathematicians accept a given body of mathematical theory, they must have taken appropriate steps to confirm it" [Pincock, 2007a, p. 264].

[21]In section 2.1, when in footnote I sketched the general lines of the Melia-Colyvan debate, I have already pointed to the importance that Mark Colyvan gives to MEPP in the context of the enhanced indispensability argument.

In the next section I am going to consider the following question: if the mapping accounts are representative, how do they account for the use of idealizations in science? This question is important for my investigation, especially if we note that mathematical idealizations are often considered to play an explanatory role in science.

5.3 Mapping accounts and idealizations

There is an important aspect of mapping accounts which is relevant to MEPP and which has not yet been addressed. Mapping accounts are representative, but in science we usually represent the actual system by introducing a deliberate falsification (for instance, in order for the system to be mathematically tractable). Call 'idealization' this deliberate distortion of some features of the actual system under study [Cartwright, 1989, p. 187]. In these situations, the resulting mathematical representation comes from nothing as a structure-preserving mapping and is known to be false of the actual world (since idealizations are necessarily false of the physical world, there can be no physical structure to be mapped onto an appropriate mathematical structure). How then do the partisans of mapping accounts, who claim to solve the problem of application of mathematics in terms of structure-preserving mapping between the empirical world and mathematics, accommodate such representations resulting from idealization? This is a serious problem for the mapping accounts, and Pincock, Bueno and Colyvan are well aware of that ([Pincock, 2004b, p. 137] and [Pincock, 2007a, p. 271], [Bueno *et al.*, 2011, p. 351]). A second question comes as a direct consequence: if we accept that idealizations do contribute to explanation, that is something well accepted ([Cartwright, 1983], [Portides, 2008], [Morrison, 1999])[22], how do mapping accounts accommodate their explanatory role and classify idealizations as

[22]While not explictly, even Pincock and Bueno and Colyvan seem to agree on this. Batterman makes the explicit claim that they accept the idea that idealizations can be explanatory: "My argument depends upon accepting the idea that idealizations can indeed be explanatory (This is something both Pincock and Bueno and Colyvan accept)" [Batterman, 2010, p. 23].

more or less explanatory? The (recognized) explanatory role of idealizations might then be 'read' as a motivation for accomodating idealizations in mapping accounts. For instance, Robert Batterman observes:

> Many explanatory models appear to involve idealizations. We speak of frictionless planes when there are no such things, and we idealize fluids to be continua when, in fact, they are composed of discrete finite collections of molecules. If we accept that idealizations can and do play important roles (perhaps even explanatory roles), then that raises a deep problem for mapping accounts of the applicability of mathematics. The problem is simple: Nothing in the physical world actually corresponds to the idealization. So in what sense can we have a mapping from a mathematical structure to an existing physical structure? Mapping accounts are representative and good representations reflect the truth about the world. Idealizations, however, are false. [Batterman, 2010, p. 11-12]

The connection between the two questions is obvious if we state the point in the following way. Consider Pincock's view that explanatory power is a consequence of representational capacity (as we have seen in his abstract explanations). Now, consider we have idealized assumptions in play. In that case, the representational capacity of the model will not say anything about the explanatory power of the model, and this simply because the represented model is, strictly speaking, false. How then do we 'weigh' the explanatory power of the idealized model in structuralistic terms? There is clearly something missing.

To sum it up, we have so far individuated two general and dependent questions:

α How do mapping accounts accommodate idealizations?

β How do mapping accounts accommodate the explanatory role played by idealizations?

Naturally, to answer β by providing some criteria to evaluate the explanatory power of idealizations would amount to having a potential way to rank idealizations in terms of explanatory power.

In the following subsection I will present Pincock's solutions to the first question, which is a necessary bridge to the second one[23]. After this, in the final section, I will propose a recent criticism put forward by Robert Batterman [Batterman, 2010]. In particular, he has observed how question β has not yet received a sufficient answer from the partisans of mapping accounts. According to Batterman, a possible solution for them would be to adopt a "Galilean" conception of idealizations (discussed below), and this would provide them with some kind of method to rank the degree of explanatory power of idealizations. However, as Batterman observes, to adopt this solution does not eliminate a further difficulty: there are idealizations which play explanatory roles *without representing* the system under study. This points to a more general limitation of mapping account of applied mathematics, which are representative, and leads to the conclusion that representation is *not* a necessary condition for explanation. I will discuss the latter point in the final part of the chapter.

5.3.1 Ranking idealizations

The missing part of Pincock's structuralist account, i.e. the story about idealizations, is given in his paper "Mathematical Idealizations" [Pincock, 2007b]. Pincock defines idealization in a way equivalent to that of Nancy Cartwright [Cartwright, 1989, p. 187]:

[23] Bueno and Colyvan propose to accommodate idealizations in their inferential conception by appealling to partial mappings between actual world and mathematical structures [Bueno *et al.*, 2011, p. 358]. The formal background of the partial structures approach is given in [Bueno *et al.*, 2002] and [Da Costa *et al.*, 2003]. In what follows I will concentrate only on Pincock's proposal, which I have considered up to now. There is another motivation behind this choice, and amounts to observing that Bueno and Colyvan's answer to question α in terms of partial representation is to some extent analogous to Pincock's answer to the same question in terms of equation and matching model [Batterman, 2010, p. 16].

> I will say that a representation results from *idealization* when the steps leading up to the representation involve deliberate falsification, that is, assumptions are invoked that the agents constructing the representation believe to be false [Jones, 1998]. And an idealization will be *mathematical* just in case these assumptions, or the resulting representation, involve mathematics in some crucial way. [Pincock, 2007b, p. 957]

How then can those false assumptions, which are essential to our science, contribute to good representations? Pincock answers this question by proposing a ranking of idealizations in terms of their representational capacity and contextual factors (beliefs and intentions of scientists doing the representation).

The starting point for this ranking is given by defining two kinds of models [Pincock, 2007b, p. 961]. The physical situation under study is perfectly reflected (in all the physical features) in what Pincock calls the "matching model", i.e. a model which ideally mirrors (mathematically) all the physical magnitudes of the target system. The mathematical model (for instance, an equation) resulting from the idealization is called the "equation model". Thus every physical parameter has a counterpart in the matching model, that is, there is an isomorphism between the target system and the matching model, while the equation model represents a class of models and comes from a mathematical idealization. At this point, the question α for Pincock's mapping account can be stated in the following way: How the (false) equation model can be representative? Pincock's idea is that the equation model represents a physical situation when:

(A) there is an isomorphism between the matching model and the physical situation

(B) there is an acceptable mathematical transformation between the equation model and the matching model[24]

[24]Observe that Pincock leaves open the possibility that, in cases where we have an

The notion of 'acceptability' of mathematical transformation is central to Pincock's proposal. To consider the mathematical transformation as acceptable draws on contextual factors:

> A mathematical transformation will be acceptable when it is consistent with the goals of the scientists in terms of scale and accuracy. [Pincock, 2007b, p. 963]

I will consider his example in a moment. Before passing to that, let me add an observation. Recall that, for Pincock, in line with the semantic view, a representation is a mathematical model or set of mathematical models [Pincock, 2007b, p. 959]. However, as it emerges from the previous quotation, his position differs from the semantic view with respect to a major point: Pincock rejects something like a "naturalistic account of scientific representation" [Suárez, 2003] and proposes an account in which the beliefs and the intentions of the scientists doing the representation are taken into consideration[25]. Bueno and Colyvan's claim that Pincock's proposal is "purely" structural [Bueno *et al.*, 2011, p. 352] should then be rejected because it is not true. What is more, the contextual and pragmatic extension proposed in their inferential conception of the application of mathematics has much in common with Pincock's appeal to "goals of scientists"[26].

In order to illustrate how his proposal works, Pincock proposes a simple example of representation which results from an idealization. Consider the amount of heat per unit of time passing from a warmer plate 2 to a cooler plate 1. This quantity is given by the following discrete equation ("Newton's

idealized representation, such transformation (relating the equation and the matching model) is not found by scientists, or even it does not exist [Pincock, 2007b, p. 963-964].

[25] There is also a second point of difference between the semantic view and Pincock's: while for the classical semantic theorist the models involved in scientific representation are models as we find in model theory, for him the models which represent a physical system via a structure preserving mapping are "wholly mathematical" (the entities in the domain are mathematical entities like real numbers and pure sets) [Pincock, 2007b, p. 960].

[26] This similarity is stressed by Batterman [Batterman, 2010, p. 14].

law of cooling states"):

$$\frac{\Delta Q}{\Delta t} = \frac{kA}{d}|T_2 - T_1|, \tag{5.1}$$

where ΔQ is a discrete quantity of heat, k is the termal conductivity of the material, T_1 and T_2 are the respective temperatures of the plates, A is their area and d is the distance from one another.

Equation 5.1, which is formulated in terms of finite differences of heat over finite periods of time, can be replaced by the 'more idealized' one dimensional heat equation

$$\frac{\partial}{\partial t}u(x,t) = \frac{k}{\rho s}\frac{\partial^2}{\partial x^2}u(x,t), \tag{5.2}$$

where $u(x,t)$ is a function describing temperature at a point x at time t, ρ is the material density and s its specific heat[27]. The heat equation 5.2 is then a partial differential equation in which discrete quantities are considered as continua. Nevertheless, to see these quantities as continua (in the passage from 5.1 to 5.2) presupposes the assumption that the material being investigated is continuous, something which is evidently false. For instance, if we use the heat equation to study a bar made of iron, the bar will contain atoms of iron and other impurities. Moreover, another iron bar will have a different microscopical structure. The question is: How do we account for the representativeness of equation 5.2?

In Pincock's terminology, equation 5.2 is what stands for the equation model, that is, what cuts down the "complete class of models reflecting all logically possible combinations of position, time and temperature to those that the equation will permit" [Pincock, 2007b, p. 962]. The equation model is the idealized mathematical model. In the case of the iron bar, the matching model is a model in which not only the physical magnitudes which appear in the continuous heat equation are mapped (via an isomorphism), but also other physical features like the color of the bar, the positions of the iron

[27]For the details of the derivation of the heat equation from equation 5.1 see [Cannon, 1984].

molecules over time, and so on. How do we pass from the matching model to the equation model? More precisely, how do we recognize that there exists a mathematical transformation between the matching model and the equation model? Pincock's idea is that we use equation 5.1 as starting point, where we have good reasons to think that this equation accurately describes *some* features of the matching model [Pincock, 2007b, p. 963]. Among these features it does not figure the color of the bar, and this because we are not interested in this feature but in the representation of the scale temperature dynamics on the iron bar. By working mathematically on equation 5.1, and introducing some constraints, we extract the equation model 5.2. According to Pincock, to accept that there exists a transformation between the matching model and the equation model amounts to proving that the model equation and the original equation from which we started will agree on certain magnitudes within certain constraints [Pincock, 2007b, p. 963]. In the case of the iron bar, this agreement (on temperature) is found by considering a specific (medium) range of time and distances. The equation model is then "good" because the mathematical relation captures this relevant feature of the matching model. On the other hand, if we consider shorter and longer timescales, this agreement will not be found (this is because in the long term the heat loss to the environment will become a dominating factor, while in the short time the particle-particle interaction will become more significant to the temperature dynamics). Finally, the fact that the equation model is representational (although idealized) depends from some contextual consideration, and precisely from the fact that scientists can accept a transformation from the equation to the matching model:

> My proposal is to go contextual.We bring in the goals that the scientists have in mind for their representation. In the heat equation case, the goal is most likely to be to represent the medium scale temperature dynamics of the iron bar for a short period of time. This provides for a certain threshold of error. So, in such a case, if there is a mathematical transformation from the equation model to the matching model that

> falls within this threshold, then we have a good or adequate idealized representation. If, despite the beliefs and intentions of the scientists, there is no such mathematical transformation, then the idealized representation is bad or inadequate. The upshot of this proposal is that we must look at the goals that the scientists have for the representation if we are to evaluate its goodness. What is an adequate idealized representation for some purposes may be inadequate for other purposes. Obviously, the heat equation is not going to be adequate to represent the color of the iron bar as the associated equation model contains nothing relevant to color. But even though it does have features tied to temperatures, there is also not going to be an acceptable mathematical transformation that gets the temperature dynamics right on the microscale. [Pincock, 2007b, p. 962]

These contextual factors permit Pincock to rank idealizations according to their adequacy for the specific situation under study. He writes:

> In my paper on idealization, I distinguish between good and bad idealizations. This distinction draws not only on the mapping account of content, but also appeals to other factors like the beliefs and goals of the scientists which deploy the model [...] it is important to be clear that I did not intend to offer such a ranking of idealizations only in terms of their goodness. There is a second respect in which idealizations are ranked. This is in terms of our knowledge of the goodness of these idealizations [Pincock, 2011a, p. 213]

More recently, Pincock has shifted from the two-models picture presented in the previous lines to a simplified version of it [Pincock, 2011d], while retaining the basic idea of his original approach to idealizations. More precisely, he simplified the above two-models picture by eliminating the idea of matching model, thus allowing to consider the direct (mathematical) relationship between the target system and the equation model. According to this refinement, in the previous example we focus on the transformation which links the target system to the equation model (alway starting from the discrete equation, which picks out some features of the target system). At this point, Pincock's new claim is that the false assumptions which come into play (for instance, those which permit to go from the discrete equation to the continuous one) should be considered only as a mean to obtain the representation. This amounts to saying that the false assumptions are no longer physically interpreted in the equation idealized model:

> false assumptions take us from an interpreted part of a scientific representation to an idealized representation where this particular part is no longer interpreted [Pincock, 2011d, p. 11]

For instance, in the case of the heat equation, this would mean that the hypothesis about the structural 'continuity' of the iron bar does not correspond to any interpretation (concerning the continuity of the iron bar) in the equation model, but it is used to fix some genuine representational content like temperature. Again, Pincock stresses the point that there exists an acceptable scale on which the target system and the idealized model equation "agree".

5.4 Is representation a necessary condition for explanation?

Pincock, as Bueno and Colyvan, welcomes the idea that idealizations can be explanatory. However, as we have seen, he provides a possible answer to question α without addressing question β, i.e. how do we account for the explanatory role idealizations play in applied contexts? Let's now see to what would amount a potential answer to the latter question.

In his [Batterman, 2010], Robert Batterman observes that there is a unsolved tension in Pincock's account of idealizations. The tension comes from the fact that Pincock wants to offer a ranking of idealizations, but he denies that the goodness of a model can be weighed by introducing a *global* metric or distance measure between the matching model and the (idealized) equation model [Pincock, 2007b, p. 964-965]. In fact, Pincock is convinced that a *local* measure of goodness, in terms of scale and accuracy acceptable to scientists (as in the heat equation case), will do the work[28].

Now, Pincock does propose only a local measure for ranking idealizations in terms of their representational capacity and contextual factors, but without an absolute measure it is not clear how different idealized models can be ranked in term of their of goodness. For instance, according to some global measure of the representational capacity of an idealization, it might be possible to say that an idealization is more representative than another, and this would provide also a possible criterion to account for the explanatory role of idealizations (for example, by considering that a less idealized model is more explanatory because is more representative than another one). This, according to Batterman, would provide Pincock with an answer to question β[29].

Batterman has observed how even Bueno and Colyvan's answer to α in terms of partial representation is to some extent analogous to Pincock's answer to the same question in terms of equation and matching model [Batterman, 2010, p. 16]. Furthermore, Batterman observes, although their partial representation approach suggests a possible ranking of idealizations in terms of representativeness (the less idealized the model is, the more representational capacity it has), Bueno and Colyvan remain silent about the degree of explanatori-

[28] Observe how, on this point, Pincock's view seems to converges on Margaret Morrison's and Mary Morgan's. According to Morrison and Morgan we will not be able to rank the representational capacity of models because models can represent the target system in different ways ([Morrison, 1999], [Morgan *et al.*, 1999b]).

[29] In his discussion note of Batterman's paper "On the Explanatory Role of Mathematics in Empirical Science", Pincock says that, differently from what Batterman thinks [Batterman, 2010, p. 15], he does not want to offer a ranking of idealizations only in terms of their representational goodness [Pincock, 2011a, p. 213]. Unfortunately, this discussion note is very recent and the structure of this dissertation was already settled when I had access to it. This is why I cannot provide here a full discussion of the various points raised by Pincock. On the other hand, this does not affect the point I want to make in this section, namely, that representation is not necessary to explanation (a point on which, I think, Pincock would perfectly agree). Moreover, although it is not given in terms of representational capacity alone (as Batterman claims), Pincock's measure of the goodness of an idealization is a local measure, and in this sense Batterman's suggestion is relevant to my discussion here.

ness of idealizations. A possible way to get out of this silence and provide an answer to β would be for them to endorse the following view: partial mappings, which are required because of idealizations, can play explanatory roles because their partialness can be potentially eliminated thus obtaining a complete representational mapping (a full isomorphism between between the world and the mathematical model)[30]. A global measure of the explanatoriness of the idealization would then be provided by a measure which points to the more or less idealized character of the model. In this framework, by paying attention to the details that are ignored by the idealized models (what is left out by the partial mapping), we would be able to say that an idealized model plays an explanatory role because we can (at least in principle) complete the story. Hence, the more the idealized model approaches the full story, less de-idealization it will require, the more explanatory power it will have. However this idea, which would provide an answer to question β in terms of ranking of idealizations, is not endorsed by Bueno and Colyvan[31]. Furthermore, even Pincock does not seem to agree on this solution. In fact, for him the explanatory power of abstract explanations comes exactly from the detraction (rather than the addition) of certain structural details of the target system:

> In the Königsberg bridges case, the explanatory power is tied to the simple way in which the model abstracts from the irrelevant details of the target system. It throws out what is irrelevant and highlights what is relevant. Crucially, what is relevant is the mathematical structure found in the target system itself. [Pincock, 2011a, p. 213]

Finally, the same tension that Batterman attributes to Pincock's account of idealization, i.e. that his ranking of idealizations does not allow a comparison of their explanatory power, is present for Bueno and Colyvan as well.

[30]As I will show in a moment, there are situations where this complete representation is not possible to obtain. An example of such a situation, concerning the use of models in social sciences, is given by Marco Panza in his [Panza, 2001].

[31]In a private communication Colyvan says to Batterman that, for him, the less degree of idealization does not necessarily correspond to a more degree of explanatory power [Batterman, 2010, p. 18].

In other words, question β has not yet received an answer from Bueno and Colyvan's inferential conception and *this* answer (the answer in terms of de-idealization) from Pincock's structuralism. Pincock, in fact, has provided a partial answer to question β. For him, at least in some cases, an idealization can play an explanatory role because it helps to indicate what is irrelevant and relevant to the phenomenon being explained. To illustrate his point he considers the following example [Pincock, 2011a, p. 214]. Drop a rock in a calm ocean. The rock will produce an irregular disturbance in the surface. Now, as the disturbance propagates outwards, it becomes more regular as the waves in the original superposition with a longer wavelength move more quickly than the waves with a shorter wavelength. This is an instance of wave dispersion. In order to explain this phenomenon, we first represent the situation through the Navier-Stokes equations for fluids. This is our model A. Next, we consider the limit where the ratio of the depth of the ocean to the wavelength goes to infinity, thus obtaining a second model B. In model B, which is our idealized model, we derive the following equation:

$$c = \sqrt{\frac{g\lambda}{2\pi}} \tag{5.3}$$

The equation expresses the velocity c of a wave with wavelength λ (g is the gravitational acceleration). From this equation we see that the speed of a wave will increase as its wavelength increases, and this is an essential component to the explanation of our phenomenon. Moreover, the mathematical passage from A to B shows that the specific depth of the ocean is irrelevant as long as it exceeds a certain threshold. In fact, we can classify waves as deep-water waves when the depth is greater than 0.28 times the wavelength λ. Therefore equation 5.3 can explain the wave dispersion for all such waves. In this case, the mapping step in which we find our model A is not sufficient to explain the phenomenon, but according to Pincock it is the mathematical passage from A to B (the limiting operation) which highlights what is relevant or not to explain our phenomenon. This mathematical operation shows that the depth of the ocean is irrelevant if it surpasses a certain threshold, while the wavelength and the gravitational acceleration g are relevant to the speed of the wave. In terms of what we have seen at the end of the previous section, here false assumptions take us from an interpreted part of a scientific representation (for instance, the depth of the ocean interpreted in the model A) to an idealized representation in which this part is no longer interpreted (in the idealized model B the depth of the ocean does not figure). It is the mathematical passage from model A to model B which removes the physical interpretation, picking out what is relevant and can be used to explain the phenomenon. Consequently, at least in this case, Pincock's mapping account can accommodate the explanatory role played by an idealization. More precisely, this example shows that a mapping account can contribute to an account of the explanatory role played by idealizations (thus providing an answer to question β). But this is only a partial answer, as Pincock observes: "But, contrary to what Batterman suggests, I do not claim that my account of idealization offers enough to make sense of the explanatory power of scientific models, including those pertaining to phase transitions and supernumerary bows. An idealization's goodness, either in the sense that it accomplishes the purposes of scientists or in the sense that we know that it does so, may have little to do with its explanatory power" [Pincock, 2011a, p. 213].

Now, let me shortly resume what Batterman considers a possible solution to the *impasse* (or partial *impasse*) mentioned above, i.e. the fact that without an absolute measure for ranking idealizations it is not clear how different idealized models can be ranked in term of their of goodness. In order to provide such a ranking and answer question β, the mapping partisans might adopt a particular view on idealizations, namely, Ernan McMullin's "Galilean" understanding of idealizations [McMullin, 1985][32]. According to McMullin, idealizations are compatible with science to the extent that they can be eliminated through further work that fills in the details ignored or distorted in the idealized model. On this account, the idealized model can

[32]See also [Suárez, 1999] for a presentation of McMullin's view.

be de-idealized by adding the appropriate corrections, and the latter operation is possible because we have 'tested' our initial idealized model on the actual system. In this way we would account for the initial idealized assumptions and simplifications, thus obtaining a more accurate representation of the actual system. Discussing his "formal idealizations", i.e. a particular type of idealization in which we simplify factors for mathematical-conceptual tractability (even if those factors are known to be relevant to the situation), McMullin writes:

> [...] models can be made more specific by eliminating simplifying assumptions and 'de-idealizing', as it were. The model then serves as the basis for a continuing research program. This technique will work only if the original model idealizes the real structure of the object. To the extent that it does, one would expect the technique to work. If simplifications have been made in the course of formulating the original model, once the operations of this model have been explored and tested against experimental data, the model can be improved by gradually adding back the complexities. [McMullin, 1985, p. 261]

By adopting this Galilean view on idealizations, then, the mapping accounts partisans might classify idealizations on the various degrees under which these idealizations can be de-idealized, and to a de-idealization there would correspond a more representative picture of the system. However, as we have seen above, neither Pincock nor Bueno and Colyvan adopt this conception[33].

[33]It might be observed that to adopt McMullin's Galilean view on idealization would commit those authors to a scientific realistic view, and their accounts would loose their neutrality about realist and anti-realist issues (at least for what concerns the truth or falsity of scientific theories). However, if we consider McMullin's approach as a 'tool', a technique of application [Suárez, 1999, p. 179], realism can be left apart and his conception can be adopted intrumentally, without any precise realist restriction. The committment of McMullin's Galilean view of idealizations to scientific realism (and anti-realism) is extensively discussed in [Suárez, 1999]. Note also that, if the fact that the model improves by making the customization suggested by the theory is used by McMullin in favour of the truth of the theory, a similar move (towards the theory) is made by Lakatos by arguing that such a customization must be seen under the light of the progressiveness of the research

Unfortunately, Batterman observes, there is a further difficulty with idealizations. In fact, there are also cases of idealizations which are not Galilean, i.e. which cannot be de-idealized by observing how the model can be potentially improved (Batterman call these idealizations 'non-traditional idealizations' [Batterman, 2010, p. 17]). A very similar observation about the impossibility of (always) improving the representation capacity of a model by adding the "rest of the story" is made by Mauricio Suárez [Suárez, 1999] and by Margaret Morrison and Mary Morgan [Morgan *et al.*, 1999b]. While Suárez observes that a "final representation" of a system might be impossible to be found, Morrison and Morgan remark how sometimes (as in the case of the various models of the nucleus) the addition of corrections to the original model results in a new model that describes the original system in a way which is inconsistent with the model of departure[34]:

> Often, models are partial renderings and in such cases, we cannot always add corrections to a stable structure to increase the accuracy of the representation. For example, models of the nucleus are able to represent only a small part of its behaviour and sometimes represent nuclear structures in ways that we know are not accurate (e.g. by ignoring certain quantum mechanical properties). In this case, the addition of parameters results in a new model that presents a radically

programme in which the theory is embedded [Cartwright, 1999, p. 250-251] (see the previous quotation from McMullin, where he writes that a model can serve as a potential basis for a 'continuing research program'). In contrast to this tendency which considers theories as having a primary importance, in her *How the Laws of Physics Lie* [Cartwright, 1983] Nancy Cartwright claimed that this sort of corrections take you away from theory and closer to the truth (the corrections to the model are *ad verum* corrections).

[34]In nuclear physics exist different models of nuclear structure, and each of them describes the nucleus in ways which are mutually incompatibles. Each model incorporates significant and different features of the nucleus, depending on what property or behavior of the nucleus is under investigation. For instance, the liquid drop model is useful in the explanation of the nuclear fission and ignores quantum statistical by treating the nucleus classically. The optical model serves as a basis for high energy scattering experiments. The shell model, on the other hand, treats nucleons in nuclei as moving independently in a central potential and takes into account the quantum behaviour, which is inexplicable using the liquid drop model. Note how contextual factors play a decisive role in the choice of the right model to be used. Nuclear models function as "epistemic resources for dealing with specific kinds of nuclear phenomena" [Morrison, 1999, p. 61].

> different account of the nucleus and its behaviour. Hence in describing nuclear processe, we are left with a number of models that are inconsistent with each other. [...] In some cases abstract representations simply cannot be improved upon; but this in no way detracts from their value. [Morgan *et al.*, 1999b, p. 28][35]

Batterman offers an example in which the mathematical idealization results from a limiting operation that relates one model to another (we will see his example in the next chapter). This operation is supposed to play an explanatory role, but it does not appeal to the static mirroring of the empirical structure by mathematics. These kinds of non-traditional idealizations play explanatory roles by involving operations and mathematical processes rather than the representation of the system. As we will see later, with his example Batterman wants to show that mapping accounts are unable to account for the role of non-traditional (limiting) idealizations in mathematical explanation:

> Surely something is right about the mapping account. In particular, when it comes to representing physical structures, mathematical structures often provide useful models that abstract (as Pincock stresses) from various explanatorily irrelevant physical details. My disagreement, as will become evident, concerns the necessity of representation for explanation. Instead, what is often explanatorily essential is the mediating limiting relationship between the representative models. To put this slightly differently, mapping accounts focus on "static" relationships between mathematical models and the world. My view is that this misses, in many cases, what is explanatorily relevant about idealizations; namely, that they often involve processes or limiting operations [Batterman, 2010, p. 10]

Hence, if Batterman is right, we have to conclude that representation is *not* necessary for explanation and more is needed. On this point Pincock agrees:

[35] See also [Morrison, 1998] and [Morrison, 1999] for similar claims.

> More generally, we should not expect an account of how mathematics describes a target system to be able to provide a complete account of how mathematics can be used to explain features of the target system. This is because *explanation usually requires more than merely accurate description* [Pincock, 2011a, p. 212-213. My italics]

This conclusion accords well with the observation that there are cases of MEPP where the mathematical model provides an explanation of the behaviour of the system but the model functions as a 'representative', i.e. its elements do not denote any element of the actual system (the model is, roughly speaking, an 'icon' which we use to study the actual world), rather than a 'representation' of the physical system under study [Morgan *et al.*, 1999b, p. 33]. For instance, R. I. G. Hughes provides a clear example of such a situation [Hughes, 1999]. In that case the Ising model, which is employed in the study of phenomena associated with diverse group of physical systems, provides a significative understanding of critical point phenomena. However, the elements in the model do not 'map' any element of the physical system under study (for instance, a magnet or a fluid studied at its critical temperature)[36].

[36] In his [Hughes, 1999], Hughes provides a short but very clear illustration of the Ising model, together with an historical background and some useful references for the study of critical point physics. I will not pause here on the details, for what interests us is the fact that it is recognized that the abstract Ising model provides a MEPP without representing the system studied. In passing, let me note that, in order to accommodate the multiple roles played by the model and its simulation, Hughes proposes a general account of representation (the *DDI* account) which is extremely similar to Bueno and Colyvan's inferential conception (the diagram of the *DDI* account is an exact copy of Bueno and Colyvan's illustration of the inferential conception, although with different labels [Hughes, 1999, p. 125]). The account takes the theoretical representation as having three components: *Denotation*, *Demonstration* and *Interpretation*; Elements of the subject of the model (a physical system showing a particular behavious) are *denoted* by elements of the model; the model possesses an internal dynamic that allows us to *demonstrate* theoretical conclusions (answers to specific questions); these conclusions can then be *interpreted* in terms of the subject of the model. Bueno and Colyvan see their account as an extension of Hughes' [Bueno *et al.*, 2011, p. 372 endnote 18]. The difference with Bueno and Colyvan's three steps approach (Immersion, Derivation, Interpretation) resides in the fact that Hughes explicitly note that "representation does not involve a similarity or resemblance between the representation and its subject", thus excluding the existence of some kind of mapping between the actual phenomena and the model, and therefore emphasizing the independence of the model from the target system [Hughes, 1999, p. 126]. I will not pursue this comparison here. The *DDI* account is discussed in detail in [Hughes, 1997].

Thus the model provides a MEPP without the representation of the system (as we will see in the next chapter, the fact that some models do not need any representation of the physical details of the systems studied is generally seen as a consequence of the universal behaviour that those systems exhibit at or near their critical temperature). In the next chapter we are going to see how Batterman proposes a possible new approach to MEPP which is independent of mapping accounts.

Finally, let me conclude this chapter with three general remarks. The first concerns the pluralist position that the mapping account partisans adopt.

As we have seen, Pincock does not exclude the possibility of having other types of explanations. Furthermore, he does not provide a single model of MEPP but he focuses on some particular kinds of MEPP, abstract explanations. The same pluralist view seems to be shared by Bueno and Colyvan when they affirm that the inferential conception "provides *one* way to understand applications of mathematics" [Bueno *et al.*, 2011, p. 370]. If other ways of applying mathematics to the world are permitted, it seems that they welcome the idea that there are different kinds of explanations and different accounts of MEPP can be given as well (remember that Bueno and Colyvan accept the idea that mathematics plays an explanatory role in science). Now, the present discussion has showed how the incommensurability problem of explanation[37] is not solved *within* the particular kind of approach the mapping accounts partisans offer. And this because, by not proposing some kind of global metric or criterion for ranking idealizations in terms of their representational capacity, the mapping accounts are not able to compare and to evaluate the explanatoriness of two *structural* explanations which involve idealizations. The incommensurability problem of explanation recurs then at

[37] I stated the problem in the introduction to this second part. If we accept that does not exist a general theory of MEPP, how can we characterize an explanation with respect to another explanation, and more precisely how do we compare explanation E_1 of phenomenon P_1 from explanation E_2 of phenomenon P_2, or even explanation E_1 of P_1 from explanation E_2 of the same physical fact P_1? Is there a way to compare two MEPP, for instance in terms of their explanatory power?

a lower level, between the very *same* kinds of explanations[38].

Second, observe how the interest of the mapping account partisans was focused on mathematical entities (for instance, a graph in Christopher Pincock's example) and properties of mathematical entities (Eulerian or not-Eulerian structure of the graph). This can be seen as a natural consequence of the interest of these authors in the ontological debate, and more particularly in the indispensability debate. My previous discussions about the linkage between Pincock's account and the ontological commitment which results from his position was intended to stress this point.

Third, in the final part of the previous chapter I raised the following question: Does a more detailed model represent better the physical phenomenon? Naturally, there are different answers to this question and nothing like a shared idea. However, what is interesting for us is to observe that, in formulating their accounts, some mapping account partisans start from the idea that the more details of the phenomenon the model (mathematically) mirrors, the better the model permits to understand that phenomenon. Batterman calls this traditional approach to modelling the 'details are better' approach [Batterman, 2002b, p. 22]. The goal is to have the best fit (a kind of convergence) between the mathematical representation and the physical phenomenon, and then more details will be welcome in this sense[39]. However, there is also another perspective on the table, and in particular a conception of mathematical modelling according to which the fine details reduce (rather than improve) our understanding of the phenomenon under study. The point

[38]As I have already noted, for Pincock this is *not* a problem, and precisely because he does not want to offer a ranking of idealizations only in terms of representational capacity.

[39]Batterman considers Pincock as a partisan of the 'details are better' approach. However, again, this is not true. In my discussion of Pincock's mapping account in the context of idealizations, I have showed how he gives a particular importance to contextual factors in the evaluation of the goodness of a representation. Moreover, in his recent discussion note Pincock explicitly rejects Batterman's reading of his position [Pincock, 2011a]. In his account the degree of goodness of a representation is given also by contextual factors, and in several cases there are details of the target system which are irrelevant to the representation itself. Pincock's point is that representations are evaluated *not only* according to their capacity to mirror the system, but also considering factors like the beliefs and goals of the scientists who deploy the model [Pincock, 2011a].

is relevant to our discussion because there are cases in which this operation of "throwing away the details" can be regarded as providing explanatory power. A study of these situations requires, according to Batterman, a totally new approach with respect to that of mapping accounts.

Chapter 6

Batterman's asymptotic explanations: painting, lack of details and mathematical operations

In the previous chapter I have presented one influential mapping account of the application of mathematics, Pincock's account, together with a particular kind of MEPP which is called by Pincock abstract explanation. Furthermore, I showed that the partisans of the mapping account view seem to endorse a pluralist view on MEPP, and that there are still open problems for their mapping accounts in the context of MEPP. In particular, in the conclusion of the chapter, I suggested that representation is not necessary for explanation. This observation, which has been pointed out by Robert Batterman in his [Batterman, 2010], does not conflict with a pluralist view but leaves us with the necessity to say something more on cases of MEPP which are recognized as such in scientific pratice but which mapping accounts are not able to deal with. In particular, if there are cases of MEPP where we have explanation without representation (one example will be provided in this chapter), we want to account for the explanatory role played by mathematics in these situations (or at least a subset of these). Batterman's view on explanation offers a possibility of this sort. It provides a way to account for the explanatory role of some non-traditional idealizations (idealizations for which, as we have seen in the previous chapter, it is not possible to tell any potential de-idealizing story) which involve a limiting operation. In particular, Batterman proposes the idea that, in these situations, the explanatory power comes from the "systematic throwing away of various causal and physical details". His view is expressed in his book *The Devil in the Details* [Batterman, 2002a] and in his papers [Batterman, 2002b], [Batterman, 2005a] and [Batterman, 2010].

There are several topics which come into play during the presentation of Batterman's view on MEPP. I introduced some of them in the previous two chapters –for instance, idealization and mathematical modelling–, while others –repeatability, universality– will require a short presentation in what follows. My purpose here is to present his idea of asymptotic explanation quite gradually, first by proposing a concrete example as an illustration of one of Batterman's core ideas and then passing to a more technical analysis of his view. This is why, in the first section, I will concentrate on two artistic artifacts (respectively, a painting and a reflection-fracturing surface of hexagonal tiles). These œuvres have been explictly constructed in order to produce a comprehensive idea of the work, or a particular effect, *only* when the observer stands not too close to the artifact and does not concentrate on its details. This short discussion (which, evidently, will be extremely rude and uninteresting from an artistic point of view) will introduce an idea that is central to Batterman's approach to MEPP, which I will discuss in section 6.2. Finally, in the third and conclusive section of this chapter, I will concentrate on Batterman's idea that his approach in terms of asymptotic reasoning is well-suited to account for the explanatory role of non traditional idealizations. In this section I will also point out some questions which are (or are not) answered by Batterman's account.

6.1 More details you have, less comprehension will result

In order to introduce Pincock's abstract explanations, we took a walk across the bridges of Königsberg. This section starts with a visit to the Metropolitan Museum of Art in New York. Consider two works which are exposed in that museum: a study of the famous painting *Un dimanche aprés-midi à l'Ille de la Grande Jatte* (henceforth *La Grande Jatte*), from the French painter and initiator of the Neo-Impressionist movement Georges Seurat (1859-1891), and the installation from the contemporary indian artist Anish Kapoor, *As Yet Untitled*[1].

La Grande Jatte is a composition made by tiny, detached strokes of pure colour too small to be distinguished when looking at the entire work but making the painting shimmer with brilliance (Figure 6.1). The pointillist technique is well-known: the dots of color become well blended when the viewer looks at them from a suitable distance. Kapoor's work is a big reflection-fracturing surface of hexagonal tiles (Figure 6.2). At a certain distance, the observer can distinguish his image without difficulty, but when too close to the surface the image becomes very confused and it becomes harder for him to appreciate the reflection of his body. This is because the numerous hexagonal mirrors will distort his image.

Although very different under various aspects (technique, material, aesthetic interests of the author, etc..), the two artifacts share something which will be central to this chapter and to Batterman's view on explanation: in order to appreciate, or discern, the scene of *La Grande Jatte* and the image in Kapoor's installation, the observer which is placed in front of both artifacts must perform the same operation, i.e. he must stay at some intermediate

[1]Seurat's painting *Un dimanche aprés-midi à l'Ille de la Grande Jatte* is exposed at the Art Institute of Chicago. Seurat's study for that painting, exposed at the Metropolitan Museum of Art in New York, is obtained through the same technique, which is what I take as revelant to my discussion here. This is why, in referring to *La Grande Jatte*, I will not make a distinction between Seurat's study and the final version of the painting (in Figure 6.1 I have used the final version of the painting).

Figure 6.1: *Un dimanche aprés-midi à l'Ille de la Grande Jatte* (Art Institute of Chicago).

distance from the artifact and not focus on the details of it. Naturally, if we try to look at *La Grande Jatte* from one kilometer of distance, we will not be able to discern anything. As for every painting or artistic work, in fact, there is an optimal distance or a range of distances which allow the viewer to better discern and understand the work. However, in the case of *La Grande Jatte* and *As Yet Untitled*, the artist has precisely introduced a sort of constraint (the pointillist technique and the fracturing surface of hexagonal tiles) to obtain an explicit aesthetic effect on the observer. To take Seurat's case, the artist intention was to make the artifact discernible and appreciable in all its aspects (for instance, relief and separation of forms, luminosity and chromatic phenomena by using the effect of "lustre"[2]), only when looked at without concentrating on the fine details[3]. Naturally, it could be noted that

[2]The effect of "lustre" is given by the perception of a partial fusion of colors. It holds when the viewer is moving back from the canvas and he has not reached the distance in which the colors are completely blended. The German physicist Heinrich-Wilhelm Dove offered an explanation of the phenomenon in terms of the effect of two masses of light which simultaneousely act on the eyes [Dove, 1853]. Probably Seurat knew Dove's writings on color [Homer, 1964, p. 143].

[3]On Seurat's method of painting see the book *Seurat and the Science of Painting* by William Innes Homer [Homer, 1964]. Seurat's techniques of paintings were strongly influenced by Charles Blanc's *Grammaire des Arts du Dessin* (Paris, 1867). In discussing the laws governing the vibration of colors, Blanc suggested a method of mixing colors

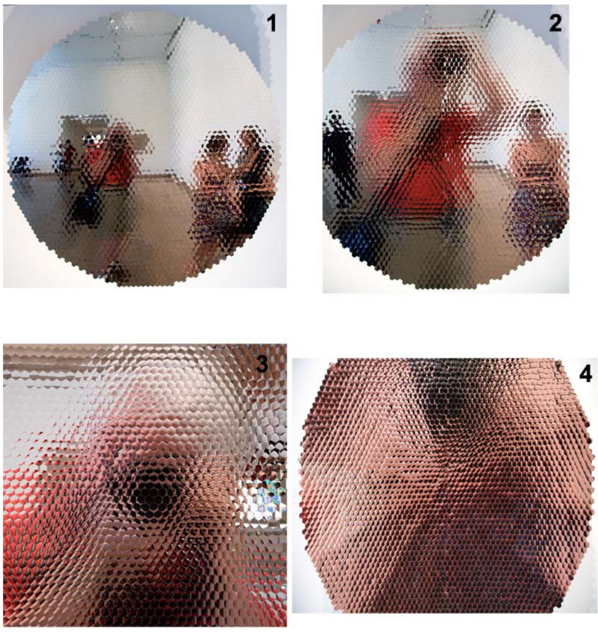

Figure 6.2: Anish Kapoor, *As Yet Untitled* (Metropolitan Museum of Art). Photos courtesy of Bill Holmes.

the observer might be interested in the details of the painting or in details of the artifact. Thus it seems that here we are confronted with different senses, or degrees, of "appreciability". Obviously, it is very natural to feel free when appreciating a canvas or a modern installation. Nevertheless, for the present discussion I am assuming that the kind of understanding the observer is immediately confronted with is that which would permit him to appreciate the full scene offered by the painting, in the case of *La Grande Jatte*, and the

---

optically through small spots or stars of pigment, and this had a direct repercussion on Seurat's pointillist technique [Homer, 1964, p. 32]. It is also very interesting to note the impact that the book *Student's Texbook of Color* [Rood, 1881], written by the physicist Odgen N. Rood and in which an explanation of physical phenomena relevant to the problem of painting was given, had on Seurat and other painters. In the chapter "On the Mixture of Colors", Rood took into consideration the rotation of Maxwell's discs at high speed as to show how the mixture of light could be effected (see [Homer, 1964, p. 37] for an illustration of the experiment). In discussing other techniques of mixing colored light, Rood also illustrated one technique contrived by Dr. Jean Mile: small dots of color placed next to each other, when viewed at a distance, yield the same effect as mixtures obtained by rotating discs. As is evident, the latter observation was of paramount importance for the Neo-Impressionist theory.

full and clear reflection of his image in the case of Kapoor's work. This, I think, also fits well with the intentions of the two artists to produce a particular effect at a particular distance. In order to not get lost in the details, the viewer must not place himself too close to the artifact. Moreover, by doing this, the observer is performing some kind of operation. For instance, the acts the viewer is requied to perform when in front of Seurat's work are well summarized by Homer in his book *Seurat and the Science of Painting* [Homer, 1964]. He writes:

> The effects of lustre described by Rood are immediately evident in *La Grande Jatte*, where the degree of optical mixture of the various colored elements depends on the viewer's distance from the canvas. *If one stands one foot from the painting*, for example, one can see all of the individual constituent colors in any given small area; *but upon moving back gradually*, these hues begin to fuse and coalesce *until one reaches a point about twenty feet from the canvas*, where fusion is complete and the individual colored strokes are no longer discernible. But before this point is reached, as Dove pointed out, effects of partial fusion or "lustre" become evident, and it is just this quality that Fénéon observed. [Homer, 1964, p. 143. My emphasis]

The motivation for this artistic discussion will be more explicit in the next section, which introduces Batterman's view on MEPP.

6.2 Asymptotic explanation: from art to science

What is the moral of the previous section? Basically, we have learned that there are cases in which the details of an artifact are mostly irrelevant to our interests, and in particular for our understanding of the artifact itself[4].

[4]Where 'understanding' must be intended in the broad sense I proposed in the previous section: the appreciation of the full scene in *La Grande Jatte* and the appreciation of the reflection of our body in the case of *As Yet Untitled*. I think that this understanding

Moreover, to have this understanding we have to perform a specific operation (we have to move away from the artifact). Now, if we identify the artifact with a specific physical phenomenon, and we substitute the word understanding with "scientific understanding", we find one of Batterman's core ideas:

> The idea that scientific understanding often requires methods which eliminate detail and, in some sense, precision, is a theme that runs throughtout this book. Suppose we are interested in explaining some physical phenomenon governed by a particular physical theory. That theory may say a lot about the nature of the phenomenon: the nature of its evolution, and what sorts of details–for example initial and boundary conditions– are required to "solve" the governing equations, and so on. One might think that the theory will therefore enable us to account for the phenomenon through straightforward derivation from the appropriate initial data, given the governing equation(s). However, I will show that, with respect to other critically important why-questions, *many theories are explanatory deficient*. [...] The kind of explanatory questions for which the detailed accounts simply provide explanatory "noise" and for which asymptotic methods fill in the explanatory lacunae are questions about the existence of patterns noted in nature. [Batterman, 2002a, p. 3-4]

From the previous quotation, various considerations emerge. First of all, Batterman focuses on particular explanatory questions which concern the existence of patterns noted in nature. He addresses, as we are going to see, only a particular kind of explanation, and accepts the existence of other kinds of explanations as well. Therefore Batterman shares a pluralist view on explanation with the supporters of mapping accounts seen in the previous chapter, i.e. he focuses only on a specific kind of explanation and he does

---

offers us a general understanding of the artifacts. In the case of *La Grande Jatte* this is straightforward. When we discern the full scene we can appreciate various aspects of the canvas and of Seurat's technique. In the case of *As Yet Untitled*, the appreciation of our image gives us a sort of 'coordinate system' which is essential to appreciate Kapoor's work and his technique.

not propose any encompassing model[5]. Second, Batterman observes that there are some methods which permit us to offer an explanatory answer to specific questions about the existence of patterns noted in nature. These methods, which "eliminate detail and, in some sense, precision", are called by Batterman *asymptotic methods*. Let me anticipate that, for Batterman, these methods involve a specific form of reasoning, which Batterman calls *asymptotic reasoning* [Batterman, 2002b, p. 3].

In the previous section I stressed the importance of the operation involved in our understanding of the artifact. More precisely, the observer had to move away from the canvas in order to reach an understanding of the artifact. That operation had the precise effect of "throwing away the details". To parallel our artistic example with Batterman's ideas, we find that it is exactly through asymptotic methods that the operation of "eliminate detail and, in some sense, precision" is performed. In these methods we use a particular form of reasoning, namely, asymptotic reasoning. As we will see shortly, Batterman considers mathematical limiting operations as paradigm instances of asymptotic reasoning [Batterman, 2002a, p. 16].

But what about explanation? Batterman calls *asymptotic explanation* that kind of explanation which utilizes such specific kind of reasoning:

> [...] asymptotic explanations gain their explanatory power by the systematic throwing away of various causal and physical details [Batterman, 2010, p. 3]

I will disentangle the notions of asymptotic method, asymptotic reasoning and asymptotic explanation in a moment, by exploring one of Batterman's examples. Before this step, let me introduce some terminology which is essential for what follows.

In certain cases, physical systems which have different molecular constitution display the same type of behavior. The patterns of behavior expressed by such physical systems are called *universal*. *Universality* is a feature of

[5] Batterman has confirmed to me his pluralism in a private conversation.

those patterns. For instance, in thermodynamics, fluids and magnets display an identical behavior at their respective critical points (this is an experimental fact), thus we speak of universality of critical phenomena[6]. In this case, as we are going to see, the same dimensionless number called ‘critical exponent’ characterizes the behavior of the system at criticality. Different systems (for instance, different fluids) have the same critical exponent.

To take another example of universality, consider many pendula with bobs of different masses, rods of different lengths and composed of different materials. In this case, the microstructural differences in the bobs or in the rods do not affect the general result that (for small oscillations) the period is directly proportional to the square root of the rod from which the bob is hanging. In other words, even if the micro-details of the pendula are extremely different, the following (mathematical) relation holds:

$$\theta = 2\pi\sqrt{\frac{l}{g}} \tag{6.1}$$

Universal phenomena have a very important feature: under perturbation of the microscopic details, their mathematical representation (the model) remains stable[7]. This is a mark that those phenomena (for instance, critical phenomena) are repeatable or reproducible. We say then that a phenomenon, or a pattern of behavior, has the feature of *repeatability* if it shows at various times and places the same macro-level phenomenology even if there are differences in the various microscopic models of it (as in the case of the pendula).

To return to our artistic example, it makes no difference if one point of color of *La Grande Jatte* is 2 or 2.03 millimeters far from another point. When the viewer reaches a distance of about twenty feet from the canvas, where the fusion of colors is complete, the micro-disposition of points is irrelevant

[6]The mathematical physicist Michael Berry has pointed out that “The most familiar example [of universality] from physics involves thermodynamics near critical points (of, say, fluids and magnets)” [Berry, 1987, p. 185].

[7]The stability under perturbation of the microscopic details is called *structural stability*. See [Rueger, 2000] for a discussion of structural stability.

to the visual result[8]. Now, keep in mind that Batterman is not interested in particular events but in repeatable patterns of behavior, and these patterns of behavior are what he considers ‘phenomena’ [Batterman, 2002b, p. 26-27].

But how do our mathematical representations of the system remain stable under changes of the microscopic details? Batterman observes that this is possible through the taking of limits:

> One important way mathematics allows us to do this [to permit structural stability] is through the taking of limits. Limits are a means by which various details can be thrown away. (For instance, in taking the thermodynamic limit in the context of explaining fluid behavior, we eliminate the need to keep track of individual molecules and we remove details about the boundaries of the container in which the fluid finds itself, etc.) [Batterman, 2010, p. 23]

As we are going to see, the discussion about limits and limiting operations will bring the debate about idealization and modelling back. In particular, concerning mathematical modelling, at the end of the previous chapter I traced a distinction between a representative conception of modelling, the “details are better view” [Batterman, 2002b], and a second view which considers that the fact to mirror too much details into a model *detracts* from an understanding of the phenomenon under scrutiny. The partisans of the second view maintain that the good model is a *minimal* model, i.e. a model which “most economically caricatures the essential physics” [Goldenfeld, 1992, p. 33]. Minimal models are then good candidates for studying the universality

[8]Naturally, my observation implicitly assumes that there is some threshold in the microscopic disposition of points above which the global effect of color perceived by the viewer is affected. Furthermore, it is the human eye which perceives the global fusion, as in the case of the pendula it is the scientist who is investigating a particular phenomenon. What is relevant here is that the global effect of color at *that distance* is exactly what the artist wanted to produce, as in the case of pendula the range of small oscillation is the exact ‘distance’ which permits the scientist to appreciate some particular behavior of the physical system. Here I will not push further the analogy. The basic idea is that at different scales (or distances) *we* appreciate different phenomena. Michael Berry has recently drawn attention to this point during his talk “Emergence and asymptotics in physics: how one theory can live inside another” (Conference on *Mathematical and Geometrical Explanations in Physics*, Bristol, December 2009).

of patterns of behavior belonging to different physical systems. As evident, Batterman's considerations will fit well in this latter direction.

We have seen that Batterman considers repeatable patterns of behavior as explananda. Then what is the 'fundamental explanatory question' he wants to answer?

> We need to understand why we have these regularities and invariances. We need, that is, to ask for an explanation of those very regularities and invariances. This is the fundamental explanatory question. The other accounts don't ask that question, in that they typically treat those regularities and invariances as given. The answer to this fundamental question necessarily will involve a demonstration of the stability of the phenomenon or pattern under changes in various details. [Batterman, 2010, p. 24]

Let's now see the previous terminology in action in one of Batterman's examples.

6.2.1 Asymptotic explanation of the universality of critical phenomena

The most exploited (and perhaps the most elucidative) of Batterman's examples concerns the explanation given in consensed matter physics for the universality of critical phenomena[9].

[*Universality of critical phenomena*] Consider the temperature-pressure diagram for a typical fluid, in Figure 6.3, where the bold lines represent thermodynamical states where two phases of the fluid can coexist (coexistence regions). The lines also represent states in which the system is subject to a first order phase transition[10]. For instance, along the line AC we find that

[9]The example is illustrated again and again in [Batterman, 2002a], [Batterman, 2002b] and [Batterman, 2010].

[10]Phase transitions are usually divided into two classes according to the behavior of derivatives of the Gibbs free energy. Phase transitions which exhibit a discontinuity in the first derivative of the free energy with respect to a thermodynamic variable are called *dis-*

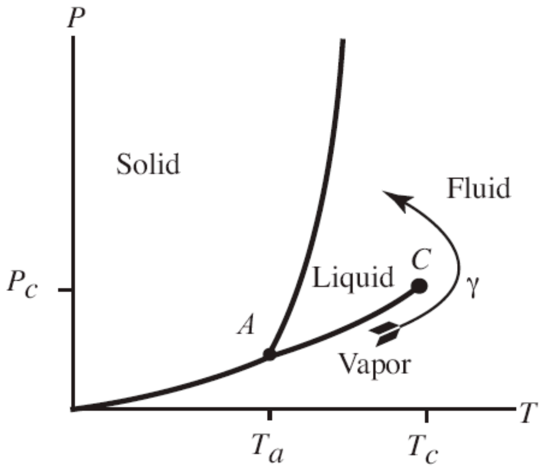

Figure 6.3: Temperature-pressure diagram for a fluid [Batterman, 2010, p. 6].

the fluid exists in both its vapour and liquid phase (it's the water boiling in a container!). In a coexistence region the temperature, pressure, and chemical potentials of each type of particle must be equal; hence, as we move along the coexistence curve, the changes in the chemical potentials and temperature of the two phases must be equal. *Triple point* A marks conditions at which three different phases can coexist. For example, the water phase diagram has a triple point corresponding to the single temperature and pressure at which solid, liquid, and vapor phases can coexist in a stable equilibrium. Nevertheless at point C, called *critical point*, the system has a totally different behavior. A *critical temperature* T_c corresponds to that point. For example, in the case of water, the critical point occurs at around 647 °K (374 °C or

continuous or *phase transitions of first order*. The various solid/liquid/gas transitions are classified as first-order transitions because they involve a discontinuous change in density, which is the first derivative of the free energy with respect to the chemical potential. First-order phase transitions are also connected with an entropy discontinuity. On the other hand, *second-order phase transitions* are accompanied by a continuous change of state, i.e. the first derivative of the free energy is continuous, but they exhibit discontinuity in the second (or higher-order) derivative of the free energy. In higher-order transitions the entropy S is continuous. See [Reichl, 1998], especially chapter 3, for an introduction to thermodynamics of phase transitions.

705 °F), and at this temperature:

> the distinction between water and steam disappears, and the whole boiling phenomenon vanishes. [...] one finds bubbles of steam and drops of water intermixed at all size scales from macroscopic, visible sizes down to atomic scales [Wilson, 1982]

Between T_a and T_c every passage from the vapor phase to the liquid phase (or the other way around) crosses the line AC and then the system goes into a state where the vapor and liquid phases coexist. But above the critical temperature ($T > T_c$) the system can go from its vapor phase to its liquid phase *without* going into this coexistence regime (by following the path γ in Figure 6.3)[11]. In other words, there is a qualitative change in the behavior of the system, and this qualitative distinction is 'mathematically' represented by a singularity in the function (free energy) which characterizes the state of the system. For $T = T_c$ the liquid-vapor transition (as well as the paramagnetic-ferromagnetic phase transition in the case of magnets) is of second order, and then continuous.

At critical point C, the same behavior of systems with *different* microscopic structure (for instance, different kinds of fluids or magnets) is described by a particular dimensioneless number β called 'critical exponent'. Surprisingly, the critical exponent has the very same value for all those different systems (this is an experimental result). This means that, even if the value of T_c changes depending on the system considered, the identical and universal behavior of those systems at criticality is described by the same β.

Now, consider what Batterman takes as explanandum: the similar behavior of fluids of different molecular constitution when at their respective critical points (universality of critical phenomena). How is condensed matter physics able to account for this remarkable fact? The key strategy to "explain" this fact is provided by the so called "renormalization group" (RG). I will introduce the RG strategy after a short presentation of how the critical

[11] To obtain the path γ: increase temperature T beyond T_c, then increase pressure P beyond P_c, and finally decrease T below T_c.

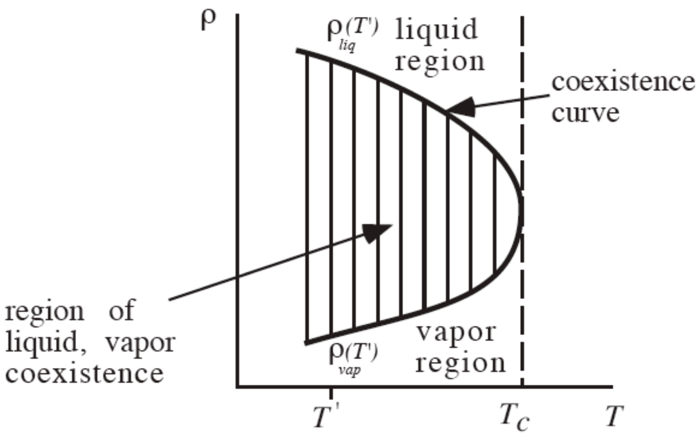

Figure 6.4: Diagram density-temperature and coexistence curve for a fluid [Batterman, 2010, p. 7].

exponent β is found[12].

In order to analyze the rearrangements of the structure of the fluid at its critical point we introduce a quantity Ψ called "order parameter". The order parameter was firstly proposed by Lev Landau in his 1937 influential study [Landau, 1937]. For a fluid, Ψ is the difference between the liquid and vapor densities when in the coexistent phases (along the line AC in Figure 6.3)[13]:

$$\Psi = |\rho_{liq} - \rho_{vap}| \tag{6.2}$$

The order parameter thus represents the main qualitative difference between the various phases. Above T_c, outside the region of liquid-vapor coexistence (region of vertical lines in Figure 6.4), Ψ vanishes for the liquid-gas phase transition, since a distinction between both phases is no longer possible. However, if we observe that the order parameter is small in the vicinity

[12]The present discussion, far from being an exhaustive presentation of the thermodynamics of systems at critical points, is intended to illustrate Batterman's example of asymptotic explanations. This is why I will not report all the technical details. For a comprehensive treatment of the critical phenomena topic see [Pfeuty *et al.*, 1977] or [Amit, 1978].

[13]For a magnet the order parameter is the net magnetization M, or magnetic moment, which measures the cooperative alignment of the atomic or molecular dipole moments.

of the critical point, it can serve as expansion parameter in the description of critical phenomena. How does the order parameter vanish? In other words, what is the shape of the coexistence curve near T_c? We can introduce a distance, the "reduced temperature" t, which allows us to say how far the system is from criticality:

$$t = |\frac{T - T_c}{T_c}| \tag{6.3}$$

We now make the assumption (experimentally supported) that, as T approaches T_c from below, the order parameter vanishes as some power β of t[14]:

$$\Psi = |\rho_{liq} - \rho_{vap}| \propto |t|^{\beta} \tag{6.4}$$

Then the critical exponent β characterizes the shape of the coexistence curve near T_c (β is also called the *degree of the coexistence curve*). Experimentally, β is found to be a number close to 0.33 [15]. All distinct fluids (and, again, even magnets with net magnetization as order parameter) exhibit an *extremely similar* shape in their coexistence curve near T_c.

Now, the critical exponent can be computed starting from mean field theories such as the Van Der Waals theory or the Ginzburg-Landau theory. The common feature of these theories is that they can be derived assuming that the particles move in a mean field due to all other particles. However, the mean field theories do not give a correct result for the critical exponent,

[14] In the case of magnets, where the order parameter is the net magnetization M, the magnetization is positive below temperature T_c (which, for magnets, is called the *Curie Temperature*) and zero above T_c. At T_c we have the transition between the paramagnetic and the ferromagnetic states of magnetic materials. As for fluids, approaching T_c the net magnetization vanishes as some power β of t: $M \propto |t|^{\beta}$.

[15] Note that in textbooks of thermodynamics and statistical mechanics the experimental value of β for fluids and ferromagnets is reported to be "in the neighborhood of 0.3 and 0.4" [Callen, 1985, p. 267], depending of the system under examination. For instance, in the case of gasses, β of He^3 is found to be 0.361, while β of Ar is 0.362 and that of CO^2 0.34 [Greiner *et al.*, 1995, p. 425]. I will return to this point in the final part of the chapter. Furthermore, it seems to me important to report a quite obvious, but extremely important, observation: "the experimental determination of critical indices is very difficult, and may contain large errors" [Greiner *et al.*, 1995, p. 425].

i.e. they are quantitatively incorrect. For instance, by using the Van Der Waals equation for a gas we obtain that the degree of the coexistence curve is $\beta = 1/2$ (see [Reichl, 1998, p. 139-141] for the exact calculation). So, in order to provide an answer to the question "Why does Ψ of different fluids F', F'', ... scale as a specific power law $|t|^{\beta}$?", and thus account for the universality claim which stands behind this question, condensed matter physics makes use of the Renormalization Group Theory (RGT). RGT is a powerful analytical theory formulated in general terms in 1971 by the high-energy theorist Kenneth Wilson [Wilson, 1971][16]. The essential contribution of RGT to the study of the universality of critical phenomena is well expressed by Greiner, Neise and Stöcker in their book *Thermodynamics and Statistical Mechanics*:

> [...] phase transitions of second order show an approximately universal behavior which does not depend on the details of the interaction, but only on a few global properties of the system, like dimension, number of components, and range of the interaction. Only after renormalization group theory was developed, which also gained large importance in quantum field theory, was it possible to establish this universality hypothesis from the theoretical point of view. [Greiner *et al.*, 1995, p. 428]

Batterman claims that the RG methods are able to provide an *explanation* (and then a MEPP) of the universability for the particular scaling of order parameters of different fluids ([Batterman, 2002a, p. 39] and [Batterman, 2010, p. 8])[17]. Observe that the same Kenneth Wilson, in his Nobel lecture *The Renormalization Group and Critical Phenomena* (8 December 1982), gives particular emphasis to the important role played by RGT in the *explanation*

[16] Even before the formal and mathematically controlled formulation of the renormalization group techniques became available, the theoretical physicist Leo Kadanoff provided a conceptual basis for the scaling behavior. His approach, which fixed the standard language of critical phenomena, gave rise to the ideas of renormalization [Wilson, 1971].

[17] Observe that RGT is a mathematical theory. Consequently, to accept that it can explain a particular pattern of behavior of a physical system (a phenomenon) amounts to saying that we considering a case of MEPP.

of the universal behavior of different systems. Illustrating the RG approach to critical phenomena, he says:

> As L [the correlation length] becomes large the free energy F_L approaches a fixed point of the transformation, and thereby becomes independent of details of the system at the atomic level. *This leads to an explanation of the universality of critical behavior for different kinds of systems at the atomic level.* Liquid-gas transitions, magnetic transitions, alloy transitions, etc. all show the same critical exponents experimentally; theoretically this can be understood from the hypothesis that the same "fixed point" interaction describes all these systems. [Wilson, 1982, My emphasis]

Even if the technical details of RGT lie beyond the scope of this chapter, let now see what Renormalization Group (RG) strategy amounts to and how it can play this explanatory role[18]. After this, we will see how Batterman accounts for the explanatory role played by the RG strategy.

[*RG strategy*] We can represent every (distinct) system by a function called Hamiltonian. The Hamiltonian characterizes the interactions between the system's components and the effect of an external field. The correlations between the different components of a system (the particles of our fluid, for instance) are usually short-ranged far from the critical point, i.e. in these regions the interaction is significant only between nearby components. As we approach the critical point, however, the system anticipates its new behavior by making "adjustments" on a microscopic scale and increasing the "length" of the correlations[19]. More precisely, at critical points, the length of the correlations diverges to infinity, thus making the mathematical problem intractable. Fortunately, RG analysis provides a method to skip this problem and compute the desired result for β.

[18] An extensive formal discussion of the technical details of the RG strategy, together with some concrete examples of RG analysis, are provided in chapter 8 of Reichl's book *A Modern Course in Statistical Physics* [Reichl, 1998].

[19] Despite the fact that the interactions between particles remain local.

The main idea is to switch from the intractable problem (analytically intractable Hamiltonian) to a tractable problem (analytically tractable Hamiltonian) preserving the functional form of the initial Hamiltonian and the fact that it describes a system with that particular behavior. To assure that the transformed Hamiltonian describes a system with the same behavior, thermodynamics parameters are properly adjusted (renormalized). More precisely, the strategy of the RG analysis is based on a systematic *rescaling* of the effective Hamiltonian which describes the system near the critical point. As the correlation length increases in proximity of a critical point, one repeatedly integrates out the effect of shorter-ranged correlations and requires that the Hamiltonian retain the same functional form. During the process the number of coupled components (the degrees of freedom) is reduced within the correlation length. This leads to nonlinear recursion relations between the effective coupling constants on different length scales. The critical point (for which the correlation range goes to infinite) corresponds to a fixed point of these recursion relations (as the number of transformation goes to infinity). This sequence of transformations, say T, is what is called the *renormalization group*[20]. At this point, the transformation matrix which yields the recursion relation can be linearized about the fixed points and the eigenvalues of the matrix can be expressed in terms of the critical exponents. Therefore, if we can find the eigenvalues, the problem is solved. Through the RG procedure *different* Hamiltonians flow to the *same* fixed point, which is quantitatively identified and accords with experimental data. Very importantly, it is in the thermodynamic limit (limit in which the number of particle of the system approaches infinity) that the fixed point of the recursion relation converges to the exact critical temperature [Lewis, 1977]. To take the thermodynamical limit amounts to taking the volume and number of particles of a bulk system to infinity, while keeping the density finite. Finally, RGT is able to tell us that the critical behavior of different systems is characterized by the *same* critical exponents.

[20]In passing, let me note that transformation T only has properties of a semigroup, and not of group.

Note that the RG procedure considers the microscopic details (degrees of freedom) of the different systems irrelevant :

> This onset of long-range correlated behavior is the key to the statistical mechanical (or "renormalization group") solution to the problem. Because large regions are so closely correlated, the details of the particular atomic structure of the specific material become of secondary importance! The atomic structure is so masked by the long-range correlation that large families of materials behave similarly–a phenomenon known as "universality," to which we shall return subsequently. [Callen, 1985, p. 265]

The microscopic details, irrelevant for the macroscopic phenomenology, are thrown away by the RG analysis. Therefore, RG analysis shows what are the details irrelevant for the system's behavior at criticality. Moreover, it allows for the determination of the physical details which *are* relevant for this behavior. In fact, in RG analysis, the critical exponent shows a dependence (only) upon the spatial dimension of the system and on the symmetrical properties of the order parameter[21]:

> Renormalization group theory demonstrates that the numerical values of the exponents of large classes of materials are identical; the values are determined primarily by the dimensionality of the system and by the dimensionality of the order parameter [Callen, 1985, p. 274]

> [...] a universality, in the sense of independence of any microscopic details, has survived the scrutiny of recent investigations, both experimental and calculational. As we shall see, this type of universality can be understood within the theory, once fluctuations are treated properly. On the other hand, the exponents show a very marked dependence on symmetry and on the number of space dimensions. What

[21]For instance, the computation of critical exponents for a triangular planar lattice will give a value of β different from that of three-dimensional systems such as fluids.

> is meant here by symmetry is the internal spin symmetry. Thus, an Ising model has a discrete symmetry of two elements – changing the sign of all spins [Amit, 1978, p. 8]

RG analysis is then able to tell us what (and *why*) various properties are (or are not) relevant for the critical behavior. Therefore it is able to tell us why the same exponent describes the critical behavior of different fluids (fluids which have a different miscroscopic composition). This is how it provides an explanation for the critical behavior of different fluids.

Now, let me put in order all these 'details'. First, RG analysis investigates the asymptotic regime (where the correlation length diverge and many molecules find themselves correlated), and it is exactly by going in this limiting regime that it makes possible to account for the universality of critical behavior. In doing that, RG analysis systematically throws away the microscopical details of the system. Therefore, since for Batterman asymptotic methods are methods which "eliminate detail and, in some sense, precision", RG analysis is our asymptotic method. Second, remember that there is an essential ingredient which permits to explore the limiting regime and find the exact critical temperature, and this ingredient is the thermodynamic limit. In fact, it is in the thermodynamic limit that the fixed point of the recursion relation converges to the exact critical temperature. Batterman defines asymptotic reasoning as "the taking of limits as a means to simplify, and the study of the nature of these limits" [Batterman, 2002a, p. 132]. Moreover, he writes:

> Limits are a means by which various details can be thrown away. (For instance, in taking the thermodynamic limit in the context of explaining fluid behavior, we eliminate the need to keep track of individual molecules and we remove details about the boundaries of the container in which the fluid finds itself.) [...] in taking such limits we are often led to focus on mathematical singularities that can emerge in those limits. The divergence of the correlation length in the renormalization group explanation of the universality of critical phenomena is one such

emergent singularity [Batterman, 2010, p. 20]

In our example, then, the asymptotic reasoning involved is the taking of the thermodynamical limit as a means to throw away the details[22].

Finally, the explanation of the critical phenomena given in terms of the RG analysis is for Batterman a case of asymptotic explanation. This explanation, according to him, gains its explanatory power by the systematic throwing away of various causal and physical details, namely, by the use of an aymptotic method (RG analysis in our example).

6.3 Managing (explanatory) idealizations and some criticisms

We have seen how Batterman proposes to capture the explanatory strategy (asymptotic explanation) for cases where the phenomena are patterns of behavior. Asymptotic explanation is, as Batterman points out, a "form of explanation largerly missed by current philosophical conceptions" [Batterman, 2002a, p. 37].

According to him, for this particular *form* of explanation it is the mathematical operation (the mathematical operation which permits the passage to the limiting regime, as in the example of critical phenomena), rather than a mathematical entity or a property of a mathematical entity, which is essential to the explanation. Furthermore, I have already mentioned Batterman's claim that his approach in terms of asymptotic reasoning is well-suited to account for the explanatory role of non traditional idealizations. In this final

[22]Let me note that, although Batterman uses the expression "asymptotic reasoning" throughout his [Batterman, 2002a], in his recent [Batterman, 2010] the same expression does not appear. And this although he discusses the very same example (universality of critical phenomena). Nevertheless, in his [Batterman, 2002a] he explicitly writes that asymptotic reasoning is the reasoning involved in asymptotic methods: "I call these methods asymptotic method and the type(s) of reasoning they involve asymptotic reasoning" [Batterman, 2002a, p. 13]. Therefore, in my analysis of his account, I retain this as his definition of asymptotic reasoning.

section, I will show how Batterman justifies this latter claim and I will add some general considerations to what has been presented in this chapter.

6.3.1 Idealizations, operations, singularities and minimal models

Basically, we have already seen how Batterman accounts for the explanatory role played by non traditional idealizations. Now it is just a matter of making his claim more explicit, by regarding the 'details' of his account and his example of critical behavior.

In the example given in the previous section, the renormalization group invokes the thermodynamic limit. This limit is, evidently, false and therefore must be considered as an idealization (better, a limiting idealization) [Liu, 2001, p. S332][23]. Batterman points out that, while this idealization is essential to the explanation (without the thermodynamic hypotesis RG analysis does not work for critical phenomena [Liu, 2001]), it is not of the kind of

[23]More precisely, the thermodynamic limit is the limit in which the number of particles of the system $N \longrightarrow \infty$, the volume $V \longrightarrow \infty$ with the constraint that the density $\frac{N}{V} \longrightarrow constant$. To take this limit usually amounts to saying that statistical mechanics *reduces* to thermodynamics. I say "usually" because for the present example of critical phenomena we take the thermodynamic limit but we do not pass from statistical mechanics to thermodynamics -for instance, at critical point various continuum thermodynamic quantities such as the compressibility k of the fluid diverge and become infinite. Singularity equals breakdown of the continuum limit. No *reduction* holds at critical points. See [Batterman, 2005a] for the technical argument. This conception of reduction, used by physicists and different from the sense of reduction used by philosophers (for instance, reduction used "à la Nagel" [Nagel, 1961]: T' reduces to T if the laws of T' are derivable from the laws of T), can be represented as follows: $\lim_{\epsilon\to 0} T_f = T_c$, where T_f is statistical mechanics, T_c is thermodynamics, and $\epsilon = \frac{1}{N}$. For a justification of the claim that asymptotic reasoning is also important in the investigation of intra-theoretic-relations and theory-reduction see [Batterman, 2002a]. Just to sketch the general moral of his discussion: while traditional philosophical accounts of reduction and physicists' sense of reduction are not able to account for emergent properties we find in the asymptotic limit (such as critical phenomena), asymptotic reasoning plays an interpretative role for such situations and provides a better instrument of investigation. The impossibility for such senses of reduction of accounting for those phenomena comes from the fact that neither T_f nor T_c are able to account for the specific emergent phenomenology within their theoretical framework. On the distinction between the physicists' and the philosophers' sense of reduction, see [Nickles, 1973].

the Galileian idealization for which it is possible to tell a sort of de-idealizing story. More precisely, this idealization plays an essential role in the explanation involving operations and processes without involving any kind of representation of the system. The mathematical idealization results from a limiting operation that relates one model (the finite statistical mechanical model) to another (the continuum thermodynamical model) *without* appealling to the static mirroring of the empirical structure. This is how Batterman's view can account for the explanatory role of non-traditional idealizations. Moreover, the focus here is on operations (mathematical limiting operations) rather than on mathematical entities. This would potentially put an end to the dependence of the explanation-debate from the indispensability-arena, or better give a fresh and innovative twist to it:

> But the main point here is that if I am right, and taking the thermodynamic limit is an explanatory essential mathematical operation, then this is a case in which, while we have a genuine mathematical explanation of physical phenomenon, there is no appeal to the existence of mathematical entities or their properties. Instead, the appeal is to a mathematical idealization resulting from a limit operation that relates one model (the finite statistical mechanical model) to another (the continuum thermodynamic model) [Batterman, 2010, p. 8]

> In contrast to explanations that appeal to mathematical entities (or properties of such entities) to explain physical phenomena, there are explanations which, while mathematical, do not make reference to such objects. Rather, these explanations appeal to (or better "involve") mathematical operations [Batterman, 2010, p. 5]

These considerations permit me to point to three aspects of Batterman's account. First, Batterman draws attention to the importance asymptotic reasoning has in underlining the crucial role of singularities in those particular forms of MEPP. In the previous example asymptotic reasoning, i.e. the taking of limits, leads to the study of a fixed point (corresponding to the behavior of the system at critical point). Singularities are often associated in physics with the breakdown of a particular pattern of regularity. In the case of the RG theory applied to critical phenomena analysis, the study of singular points (fixed point of the transformation) is used to study the regular behavior of the system as it approaches criticality (where the regularity is broken).

> [...] in taking such limits we are often led to focus on mathematical singularities that can emerge in those limits. The divergence of the correlation length in the renormalization group explanation of the universality of critical phenomena is one such emergent singularity. [Batterman, 2010, p. 23]

Second, observe how the "details-are-better" approach of mathematical modelling is not useful in the study of the behavior of the system at criticality. The model for critical behavior is a minimal model, i.e. an idealized caricature of the system under study[24]. It's only through this caricature that the universal (and repeatable) pattern of behavior is accounted for. Therefore Batterman endorses this "minimal" conception of modelling, at least for those sorts of repeatible patterns of behavior we encounter in science:

> Minimal models play crucial computational and *explanatory* roles. [...] Let us return to the question raised at the beginning: is there a kind of methodological priority among the different approaches to modelling in science? The answer, I believe, is 'yes'. Given that our primary interest is to understand repeatable phenomenological behavior, it seems we ought first to search for minimal, exactly solved models of that behavior. The asymptotic methods involved in justifying the use of such models to explain universality themselves provide the understanding of this type of repeatable phenomena. [Batterman, 2002b, p. 37]

[24]Let me observe that Batterman does not trace a clear dinstinction between the word "model" as used in "thermodynamic/statistical model" and the same word used in "model for critical behavior". However, it is clear that they refer to different things. Intuitively, Batterman uses the word "model" in the expression "model for critical behavior" to refer to the mathematical form of the transformed Hamiltonian, which models the physical situation without mirroring the details of the system and thus represents a "minimal model" (see, for instance, [Batterman, 2002b, p. 25]).

Finally, it should be noted that there are interesting similarities between Batterman's view and Pincock's. For instance, both consider that the mathematical limiting operation and the lack of detail play a positive role in their accounts. Consider the following passages from Pincock's papers:

> [...] mathematics allows us to make claims about higher-order or large-scale features of physical systems while remaining neutral about the basic or micro-scale features of such systems. [Pincock, 2007a, p. 255]

> All that I have done is described the physical system at a higher level of abstraction by ignoring the microphysical properties of the bridges, the banks and the islands. [Pincock, 2007a, p. 259]

With respect to the role that limiting operations play in idealizations, we have see in the previous chapter how Pincock agrees in considering that such limiting operations can have an explanatory role. And he considers that the limiting operation has exactly the function of highlighting what is relevant and what is not relevant to the explanation. On this point it seems that Batterman and Pincock converge. Moreover, Pincock suggests that his mapping account can contribute to an account of the explanatory power of a case involving such a limiting operation (as in his example of wave dispersion reported in section 5.4)[25]. This is why he maintains that a focus on asymptotic reasoning (the taking of a limit) provides new opportunities to combine a mapping account with a positive account of explanatory power [Pincock, 2011a, p. 216].

Nevertheless, abstract explanations and asymptotic explanations are considered (by both Pincock and Batterman) as *distinct* kinds of MEPP:

> I disagree with Pincock's claim that asymptotic explanations are a subspecies of abstract explanations. At least I believe that, by and

[25] In the case of wave dispersion considered by Pincock the limit is regular, while Batterman focuses on examples in which the limits are singular. However, Pincock considers that also in cases where the limit is singular "it is possible to reconstruct these sorts of explanations [the explanations considered by Batterman, in which the taking of a limit is the taking of a singular limit] with the aid of the mapping account of content" [Pincock, 2011a, p. 216]. This proposal is not developed further in his [Pincock, 2011a].

> large, asymptotic explanations often do not proceed by focusing on an abstract structure realized by the physical system. [Batterman, 2010, p. 3]
>
> But I explicitly noted in my discussion of the bridges case that this sort of 'abstract explanation' is different from Batterman's cases: 'some abstract explanations are not asymptotic explanations [...] [and] abstract explanations generally require philosophical examination' [Pincock, 2011a, p. 213]

In particular, as Batterman points out, asymptotic explanations are explanations that do not necessitate the 'representativeness' that appears in the case of abstract explanations. This is the case, for instance, of the explanation of critical phenomena:

> [...] there are no structures (properties of entities) that are involved in the limiting mathematical operations. That is, limiting mathematical operations typically do not yield anything like the abstract non-Eulerian structure of the bridge system in Pincock's example [Batterman, 2010, p. 21]

6.3.2 Open questions and some criticisms

Naturally, there are still open questions with respect to Batterman's account, and a lack of criticism of his view among the contemporary literature on MEPP does not mean that his position has not been criticized.

For instance, some philosophers have expressed their skepticism toward Batterman's approach to explanation during the Conference *Mathematical and Geometrical Explanations in Physics*, at the University of Bristol (11th-12th December 2009). Of particular interest was the lecture given by Steven French, "Disentangling Mathematical and Physical Explanation", in which he proposed the idea that symmetry (or symmetry principles) might play a crucial role in mathematical explanations [French, 2010]. Very roughly, the point is that symmetries are a fundamental feature of physical structure, and they can play an explanatory role. Now, concerning Batterman's example of critical phenomena, we have seen that the critical exponents show a dependence only upon the spatial dimension of the system and on the symmetrical properties of the order parameter. It is then reasonable, I think, to ask whether in this case symmetrical properties are the essential explanatory ingredient. Batterman does not give any essential role to symmetric considerations in illustrating his account of asymptotic explanations for the example of critical phenomena, but it seems that scientific practice attributes to symmetry an important role for the specific case under study. Even if a study of the role of symmetry principles in MEPP has not been proposed, it would be interesting to further investigate French's considerations on the explanatory role of symmetries for the particular case considered by Batterman[26]. This would also potentially provide a useful lever on Batterman's claims. However, here I will concentrate on what I consider two general weak points of Batterman's account: (α) it does not offer a ranking of idealizations in terms of their explanatory power (this difficulty relies on a more general problem: the lack of precision in what is seen by Batterman as the key explanatory factor in asymptotic reasoning); (β) it is not able to capture cases of asymptotic reasoning where the throwing away of the details comes *without* appealling to a limiting operation. Concerning point (β), I will report one position (that of Margaret Morrison) which is in conflict with Batterman's idea that explanatory power comes with the procedure of "throwing away the details" (as in Batterman's asymptotic reasoning). Finally, I will conclude this section by considering Batterman's answer to a question which has not been addressed yet: (γ) What about the applicability of others models of explanation to the example of the universality of critical phenomena?

(α) This point concerns what I have called the incommensurability problem of explanation[27]. Note that Batterman is silent on the ranking of ide-

[26]In particular, in a private conversation French observes how it may be that the role of symmetries can be captured within Dorato and Fellini's account of 'structural explanation' [Dorato *et al.*, 2011]. Nevertheless, this is a direction which has not been explored yet.

[27]I stated the problem in the introduction to this second part of the dissertation.

alizations according to their explanatory power. Let me put the problem in the following way: if we have two asymptotic explanations of the same phenomenon which appeal to a limiting operation, and this limiting operation is clearly an idealization as in the case of the thermodynamic limit, do we have to consider both as equally explanatory? Two cases are possible. If we answer positively, and we consider them as equally explanatory, it is reasonable that we need to characterize a possible set of limiting operations which are considered as explanatory (for instance, limiting operation σ involved in explanation E_1 and limiting operation τ involved in explanation E_2 both belong to the set of limiting operations X, which is seen as the set containing the -explanatory- limiting operations). Nevertheless, in order to characterize a set of explanatory limiting operations, we need an entrance requirement, i.e. a condition that σ and τ must satisfy in order to belong to this (same) set, and this condition is not provided by Batterman. On the other hand, if our answer is "No! This or that explanation is recognized in scientific practice as more (or less) explanatory than the other", we are using some criterion to compare their explanatory power. Hence, in both cases, it would be reasonable for Batterman to say something more on this point. More exactly, Batterman needs a more precise characterization of what is intended as "mathematical limiting operation". The problem of not disposing of a metric for weighing the explanatory power of idealizations relies, I think, on a general difficulty that Batterman's account has. The difficulty comes from the fact that, from Batterman's discussion and more precisely from the example I have reported above, it is not clear what is the (essential) explanatory factor to be considered. Compare the following statements:

(a) " [...] these singularities are essential for genuine explanation" [Batterman, 2010, p. 22]

(b) " [...] taking the thermodynamic limit is an explanatory essential mathematical operation" [Batterman, 2010, p. 8]

(c) "Minimal models play crucial computational and explanatory roles" [Batterman, 2002b, p. 37]

Recall that the transformed tractable Hamiltonian in the RG procedure represents a minimal model. The renormalization group invokes the so called thermodynamic limit and the singularity, i.e. the fixed point corresponding to the behavior of the system at the critical point, is what emerges in this limiting operation. Therefore, while in (a) (b) and (c) the essential explanatory factor is seen as something different (in *a* the fixed point, in *b* the thermodynamic limit, in *c* the Hamiltonian itself), if we follow Batterman it seems that it should be more correct to say that it is the investigation of the limiting regime (and then the operation involved in *b*) which makes the Hamiltonian and the singularity explanatory. This is because the thermodynamic limit is essential to the RG procedure and therefore to the finding of the fixed point. The focus is then on the passage to the limit (the same emphasis on the thermodynamic limit as the limiting operation is given by Batterman in his [Batterman, 2010]).

Now, another example of asymptotic reasoning is given by Batterman in chapter 6 of his book *The Devil in the Details* [Batterman, 2002a]. In it he considers as explanandum a particular (observable) pattern of behavior exhibited by rainbows. This universal feature of rainbows emerges in the asymptotic domain as the theory of light approaches the ray theory (or geometrical optics), i.e. the universal feature emerges in the limit in which the wavelength of light approaches zero ($\lambda \longrightarrow 0$). However, neither ray theory nor wave theory can account for the phenomena (patterns of behavior) inhabiting this asymptotic domain. This is because in the zero wavelength limit we find mathematical singularities and therefore the smooth limit in which we would find ray optics simply does not exist. The existence of this 'no man's land' between two theories is important enough to let some physicists speak of a *new* theory capable of characterizing this asymptotic domain. This new theory has been called *catastrophe optics* [Berry *et al.*, 1980]. Catastrophe optics incorporates elements of wave theory as well as ray theory[28]. Furthermore, under the light of this new theory, it is asymptotic reasoning which, as in the case of critical phenomena, permits to account for the stability of patterns under perturbation of microscopic details, thus leading to an explanation of their universality. The point I want to make here is the following. This example, in which the limiting operation is the passage $\lambda \longrightarrow 0$, seems to confirm that in the example illustrated above (critical phenomena) it is *the* thermodynamic limit ($N \longrightarrow \infty$) *the* limiting operation central to asymptotic reasoning (and, consequently, to asymptotic explanation). However, consider again what we said earlier in the chapter: asymptotic methods are methods which "eliminate detail and, in some sense, precision"; the reasoning involved in them is asymptotic reasoning; and asymptotic explanation is exactly that kind of explanation which utilizes such specific kind of reasoning. We also said that the role of the limiting operation involved had the precise effect of "throwing away the details", and this "throwing away the details" is what grants asymptotic explanations their explanatory power [Batterman, 2010, p. 3]. Now, if we focus on the thermodynamic limit as the limiting operation, there is nothing like an asymptotic reasoning in taking it. This is because to take the thermodynamic limit does not amount to any operation of "throwing away the details" *per se*. It is the RG procedure which, through a series of transformations, permits the throwing away of details. To take

[28] An important observation should be added here. What Batterman wants to illustrate with his example is that there exist cases of fundamental physical theories (FT) which are explanatorily inadequate and whose conceptual apparatus must be supplemented by concepts of a less fundamental theory (LFT) in order for these theories to be suitable to explain some phenomena. In the example of the rainbow, the wave theory is explanatorily inadequate to account for the phenomenology of interest, and the explanation of the phenomena requires the introduction of concepts from geometrical optics. A criticism of this argument is found in [Belot, 2005]. In a nutshell, Gordon Belot's objection is that in Batterman's example the mathematics of the LFT can be defined in terms of the mathematics of the FT, and only the latter need be given a physical interpretation, i.e. the explanation draws only upon resources which are internal to the more fundamental theory. Similarly, in his discussion note of *The Devil in the Details*, Michael Redhead accuses Batterman of improperly reifying the mathematical structures of the superseded emeritus ray theory when claiming that such structures are required for genuine physical understanding [Redhead, 2004]. See [Batterman, 2005b] for Batterman's response to Belot and Redhead.

the thermodynamical limit into account is essential to this procedure, but it cannot be considered as *the* limiting operation by which various details are thrown away[29]. Therefore the limiting operation would be given by the RG procedure, but this assertion is controversial because the RG procedure is a mathematical theory and not a mathematical operation. An operation could be performed within a theory, but then something more must be said to characterize such an operation. Furthermore, Batterman seems to be more inclined to consider the thermodynamic limit as the mathematical limiting operation which permits to reason asymptotically. What is then the *key* to (asymptotic) explanation? "Limits are a means by which various details can be thrown away", but what kind of *limits*? Again, we find the problem of characterizing the mathematical limiting operation as a subproblem to the more general problem of characterizing the key explanatory factor in asymptotic reasoning.

(β) Let me pass to what I consider a second weak point of Batterman's account. Even if we accept that asymptotic reasoning is essential to our scientific practice, it seems that there are cases where the procedure of throwing away the details comes *without* appealling to limiting operation. For instance, consider Margaret Morrison's analysis of the Lagrangian formalism as applied to the study of mechanical systems[30]. She stresses how the lack of details (concerning the nature of the system studied and its motion) was an essential feature of the Lagrangian approach to mechanics. What is more, she points out that the fact that the Lagrangian approach is a strategy which does not consider details is exactly what confers to it its unifying power. Con-

[29]Recall a previous quotation from Batterman: "Limits are a means by which various details can be thrown away. (For instance, in taking the thermodynamic limit in the context of explaining fluid behavior, we eliminate the need to keep track of individual molecules and we remove details about the boundaries of the container in which the fluid finds itself, etc.)" [Batterman, 2010, p. 23].

[30]The analysis of the Lagrangian formalism as an essential ingredient to Maxwell's unification of electromagnetism and optics appears in [Morrison, 1992] and is reworked in chapter 3 of her [Morrison, 2000]. The key role played by the abstract structure of the Lagrangian formalism in the development of Maxwell's electrodynamics is well summarized in her review article of Etienne Klein and Marc Lachieze-Rey's book *The Quest for Unity: The Adventure of Physics* [Morrison, 2002].

sider the following (long) quotation from Morrison's book *Unifying Scientific Theories*:

> The unifying power of the Lagrangian approach to mechanics lay in the fact that it ignored the nature of the system and the details of its motion; one did not start with a set of variables that had immediate physical meanings. Indeed, insofar as the formal structure of the theory is concerned, analytical mechanics, electromagnetism, and wave mechanics can all be deduced from a variational principle; the result being that each theory has a uniform Lagrangian appearance. The velocities, momenta, and forces related to the coordinates in the equations of motion need not be thought of as collective variables that refer to some microscopic order. Instead, their physical significance is simply a measure of their practical value in solving Lagrange's equations. Hence, we can have a quantitative determination of the field without knowing the actual motion, location, and nature of the system. This degree of generality allows us to apply the Lagrangian formalism to a variety of phenomena regardless of their specific nature; something that is particularly useful when the details of the system are unknown or when it is assumed to lack a mechanism. One can see, then, how this generality is also the basis of its unifying power. Because very little information is provided about the physical system, it becomes easier to bring together diverse phenomena. Only their very general features need be accounted for, yielding a unification that to some extent is simply a formal analogy between different kinds of phenomena. Because the Lagrangian approach did not require that Maxwell provide details about the propagation of electromagnetic waves due to internal structural features of the field, a unified theory was significantly easier to produce. [Morrison, 2002, p. 348]

For this case, mathematical limit operations do not constitute paradigmatic instances of asymptotic reasoning. And this simply because in the case of the Lagrangian approach there is no appeal to limiting operations. For instance, consider the case of the different pendula seen above, where the Lagrangian approach can be used to study the systems. The Lagrangian puts emphasis on the energetical properties of a system, rather than its internal structure. Hence the situation is the following: "throwing away the details" and study the behavior of the system(s) via the Lagrangian formalism gives rise to the form of MEPP considerd by Batterman (it involves asymptotic reasoning), but this procedure is performed *without* appealling to any limiting operation and therefore escapes one of Batterman's core intuitions[31]. Therefore it is not clear if the role of limiting operations in asymptotic explanations must be considered as an essential ingrendient or not (Morrison, for instance, does not even refer to limiting operations in considering the Lagrangian formalism). Furthermore, as we have seen in chapter 3, for Morrison unification and explanation are often constrasting goals. In the case of Lagrangian mechanics she points out that the fact that Lagrangian approach lacks of details is exactly what confers to it its unifying power. For her, an increase in unification here means that we lose in explanatory power. If we translate the situation into Batterman's terminology (and we retain Morrison's conclusion): asymptotic reasoning leads to unification but not to explanatory power. Nevertheless this openly conflicts with Batterman's core idea, i.e. that "asymptotic explanations gain their explanatory power by the systematic throwing away of various causal and physical details" (previous quotation). I think a similar remark could be made by considering the unifying power of the RGT when applied to critical phenomena. RGT is able to account for a pattern of behavior showed by systems with different mi-

[31]Batterman could reply that the feature of universality of pendula, i.e. the fact that $\theta = 2\pi\sqrt{\frac{l}{g}}$ holds for small oscillations, involves a limiting operation in the sense that the small angle hypothesis ($\theta \ll 1$) allows the use of a Taylor expansion of the function $\sin\theta$ about the equilibrium point which leads to the approximation of $\sin\theta$ by θ. Furthermore, the original function is simply the limit of Taylor polynomials ($f(x) = \sum_{n=0}^{\infty} \frac{f^n(x_0)}{n!}(x - x_0)^n$), hence a mathematical limiting operation holds. Observe, however, that here we 'cut' the Taylor expansion (we omit terms of some higher order). A similar defense could be used against the use of the Lagrangian formalism plus a variation principle in the study of a variety of phenomena. In this case, Batterman could say that a variational principle is an integral principle, and then a limiting operation is necessarily involved (the use of infinitesimals quantities/displacements). But do we have to consider the use of infinitesimals in physics as 'limiting operations'?

croscopic composition, thus providing a sort of unifying structure capable of accounting for the universal behavior. In that case, I think, Margaret Morrison would claim: "That's right! But are you sure we gain in explanatory power *via* the general unified mathematical description?".

(γ) Finally, let me conclude this chapter with a (partial) answer to the following question: 'What about the applicability of others models of explanation to universal phenomena?'. Batterman rejects the applicability of the mapping accounts in the case such as that of critical phenomena. But what about other accounts of explanation? Batterman addresses this question in chapter 3 of his *The Devil in the Details*. The conclusion he reaches is that neither the causal-mechanical models nor the unification models manage to do a very good job [Batterman, 2002a, p. 35]. The motivation for rejecting the causal-mechanical models is that the explanation of universal behaviour does not seem to involve causal considerations, at least according to the particular characterization of causal mechanism or process given by authors as Salmon (observe that it is precisely the insensitivity to detailed causal mechanisms that characterizes universality)[32]. Concerning Kitcher's unification, Batterman's claim is that in order to account for universal behavior Kitcher has to include into the explanatory store $E(K)$ some kind of scheme reflecting asymptotic methods[33]. However this cannot easily be done, because asymptotic methods come in very different forms (associated with different types of limiting operations):

> Asymptotic analyses show that many, if not most, of the details required for characterizing the types of systems in question are irrelevant. But in each case that reasoning may be very different and the form of the limiting "laws" may also be quite different. I do not think that the purely internalist account of explanatory unification can easily respect these differences. That is to say, I don't think the unification theorist can simply add to $E(K)$ some schema that says

[32] I sketched Salmon's conception of 'causal process' in chapter 2.

[33] Remember that the explanatory store $E(K)$ over K is the set of arguments which best unifies K, where K is the set of accepted sentences at a particular point in time.

"use asymptotic analysis" [Batterman, 2002a, p. 33]

Another claim against the possibility to use a unification account in situations such as that concerning the universality of critical phenomena has been made by Michael Redhead in his discussion note of Batterman's *The Devil in the Details*. For Redhead, accounts *à la* Kitcher and *à la* Friedman are not appropriate for covering explanations of critical phenomena because "far from tying together apparently *diverse* phenomena in unified patterns of explanation, universality concentrates attention on a range of *similar* phenomena instantiated by diverse causal mechanisms" [Redhead, 2004, p. 529].

But what about the classical Hempelian D-N model? Batterman distinguishes betwen two different sorts of why-questions [Batterman, 2002a, p. 23]:

(i) Why questions which ask for the explanation of why a given instance of a pattern obtained.

(ii) Why questions which ask for the explanation of why patterns of a given type can be expected to obtain.

Why-questions (ii) are questions about the existence of universal behavior, i.e. the kinds of why-questions Batterman claims asymptotic explanation is able to cover. On the other hand, according to Batterman, the Hempelian model for explanation is able to answer why-questions of type (i), but it remains silent on why-questions of the second species. This is a first objection raised by Batterman against the applicability of the D-N model in the case of universal phenomena. For instance, in his reply to Gordon Belot, who claims that asymptotic explanations can be given under the form of D-N explanations/derivations [Belot, 2005, p. 144-145], Batterman writes:

> The question now concerns how to understand Hempel's idea that we may "speak derivatively of a theoretical explanation of solar eclipses or rainbows in general." The D-N model provides an account whereby we can explain the particular occurrence of a solar eclipse or of a rainbow. It is unclear how such an account (or even a collection of such

> accounts) can allow us to speak 'derivatively' of a general account that would answer the corresponding type (ii) question. The point here, as I have emphasized in the book, is that each such D-N account of a particular occurrence will be remarkably different from all of the others. As Belot himself notes, each account will involve different initial and boundary conditions –different shapes of the raindrops, for example. Had we an explanation that answers the type (ii) question, then it seems that, yes, we may very well speak 'derivatively' of explaining any given instance. We will have an account that tells us why many/most of those individual details can be ignored. That is just what asymptotic explanation, as I have presented it, provides. [Batterman, 2005b, p. 156-157]

A second argument given by Batterman against the possibility of using the D-N schema in explaining universal patterns of behavior concerns the use of idealization in the explanation involved. In chapter 2 I presented the classical Hempelian model of explanation. Remember that, according to Hempel, an adequate D-N explanation must meet the 'empirical condition of adequacy' (R4), i.e. the sentences constituting the explanans must be true. Hence any explanation appealling to idealized structures, which are false, is automatically excluded as good candidate for the D-N logical machinery. In Batterman's words:

> Asymptotic explanation, in that it typically involves idealizations, fails (at least in some instances) to meet one of Hempel's conditions on an adequate explanation –namely, truth. [Batterman, 2005b, p. 162, footnote]

Asymptotic explanation is, if Batterman is right, just another form of explanation. It is a different kind of MEPP. However, as we have seen in the previous pages, there are still open problems for this account.

Part III

A new approach to MEPP in terms of intellectual tools and conceptual resources

Don't treat your commonsense like an umbrella. When you come into a room to philosophize, don't leave it outside but bring it with you

Ludwig Wittgenstein's slogan, in [Wittgenstein, 1975, p. 68]

In the previous parts I have analyzed two general philosophical trends towards MEPP, together with their most significant positions. In part I it has been shown how some authors propose a general solution to the problem of MEPP, namely, they intend to capture the nature of MEPP by proposing a single encompassing model. Their models, however, are subject to criticisms and counterexamples and they do not seem to offer a robust notion of MEPP. On the other hand, in part II we have seen how other authors endorse a pluralist view on explanation and are then committed only to a partial solution of the problem of what constitutes a MEPP (and how do we capture such a notion). For these authors there are different kinds of MEPP, i.e. they consider that what makes something a good explanation can vary from case to case. Moreover, they offer a characterization of a specific type of explanation but they do not provide any single encompassing model of MEPP. When I presented these pluralist positions I have shown how also the accounts proposed by these authors are open to criticism, and the philosophical discussion of these positions is still at an early stage. In this final part, which stands as a conclusion of my work, I will resume and extend some general results of the previous discussions in order to support two thesis:

α The three WTA accounts analyzed in the first part do not capture some cases recognized as genuine MEPP in scientific practice. This lack of success might be due to the fact that, contrary to what the scientific practice itself suggests, those models do not consider qualitative factors as essential to MEPP. As a consequence, the models should be refined, in order to capture the test cases proposed, or rejected in favour of a different approach. This argument will be defended in chapter 7.

β However interesting and successful might be the consequences which would derive from the refinement of each one of the WTA approaches considered, the burden of the proof is shifted towards the partisans of the WTA approach and I will not concentrate on such a task. By adopting a pluralist view and focusing on qualitative factors, I will suggest a new approach to MEPP which is supposed to account for our scientific practice and offer new directions of analysis. Even if we cannot characterize MEPP *simpliciter*, we can say *why* a MEPP is regarded as explanatory in scientific practice. This could be done by appealling to two categories: conceptual resources and intellectual tools. Through these categories it is possible to account for the different species of MEPP we are confronted with in scientific practice. This thesis, together with the presentation of the relative notions, will be defended in chapter 8.

I will consider the two thesis in more detail in the next two chapters. However, let me spend some words on thesis β.

To accept a pluralist view on MEPP entails admitting that it is impossible to provide an encompassing model of MEPP (together with a general notion of explanatory power). In claiming that I am adopting a pluralist view and I am taking into consideration qualitative factors, I am not denying that it is possible to provide some objective-based (*à la* Steiner) or formal-based (*à la* Kitcher) notion of MEPP which works well in specific cases (this attitude reflects exactly the pluralist position, which I totally endorse). The moral of my claim in β is much more fine-grained: I think that a promising task in the philosophical analysis of MEPP is not to focus on *what* a MEPP is (this is the task of the scientist who recognizes a MEPP as genuine in his scientific practice, not the task of philosophers!), but on *why* a MEPP is recognized as providing explanatoriness in scientific practice. The latter task is very different from the former (maybe the first could be seen as a natural subtask of the second). The pluralist principle supports this claim in the sense that it suggests that the investigation of a general universal notion of MEPP would result as meaningless, and then such a direction should be discarded because it is not fruitful. By accepting pluralism, I will not try to develop a new and distinct model of MEPP, neither will I refine an existent account of MEPP. I will try to offer a possible schema to capture MEPP, under the form of a general framework, which will be compatible with this pluralist principle. Such a schema will provide a representation of MEPP in terms of two categories (conceptual resources and intellectual tools). Although offered in terms of two categories, this schema will permit us to represent within its structure different kinds of MEPP.

Let me anticipate how this schema is supposed to work. Consider the following example from the game of soccer. To simplify, we can assume that every tactic of a soccer team can be represented in terms of the number of players in every unit of the field (defense, midfield, attack)[34]. For instance, in a 4-3-3 tactical system we will have four players in defense, three midfielders, and three attackers. We use the same field-schema made of defense, midfield and attack zone to represent teams' tactics which are very different among them, for instance that of Manchester United and that of Barcelona. Every distinct tactic can be described, and even classified, according to the schema provided by the field schema. In other words, the field-schema does offer a way to capture the differences among the various tactics employed by distinct teams. The classification is provided by the possible combinations (4-3-3, 4-4-2, 5-3-2 and so on). Moreover, every kind of tactical system informs us of the defensive or attacking character of the team (a team which adopts a 5-3-2 will have a more defensive character). In the same way, my schema will accept within its structure different species of MEPP (different teams' strategies or tactics, in the soccer-analogy), and these MEPP will be described in terms of conceptual resources and intellectual tools (the three field units in the analogy). As in the case of the team's tactics, this will point to a potential way to classify MEPP.

The compatibility of my approach with the pluralist hypothesis therefore comes from the fact that I will try to identify a trait common to *different*

[34]Here I am not considering the goalkeeper, whose tactical position is supposed to be the same in every soccer team.

species of MEPP[35]. This common trait concerns the *use* of mathematical concepts, rather than the properties of mathematical entities involved in the explanation or some objective relevance relation, and the possibility of using our abilities to reason through these concepts. In terms of the distinction ontic/epistemic introduced in section 4.2, I will propose then an epistemic approach to MEPP.

My approach will have various favourable consequences. In particular, as I will show, the payoff of my analysis in terms of conceptual resources and intellectual tools will be to:

- Offer a potential solution to the classical asymmetry problem discussed in connection with scientific explanation.

- Offer a new insight into the ontological debate between platonists and nominalists in philosophy of mathematics. More specifically, if endorsed, my framework offers a possible ground to reject the validity of the Enhanced Indispensability Argument proposed by platonists as to show the existence of mathematical objects.

- Offer a possible linkage with the notion of scientific understanding.

- Integrate an historical approach to MEPP (which is, for the most part, banned by contemporary models of MEPP).

Far from giving a general characterization of MEPP, I will suggest a different perspective of study. The general moral will be: even if we cannot say *what* a MEPP is, perhaps we can try to answer some more simple and pragmatic questions: Why do scientist regard a MEPP as such? Is there a common trait which is shared by the different species of MEPP and which can permit a classification of MEPP? Naturally, even if I will suggest a way to answer these questions, my approach is far from giving a whole story about MEPP. This is why I will point to some problems which are still open and

[35] In the analogy of soccer, the common trait shared by every team's strategy is the fact that a strategy is built on defenders, midfields and attackers.

are not solved or addressed within my framework. As the painter Nikolaus Hipp said: "It is better to show restraints that include too much!".

Finally, I want to add something concerning the motivations for proposing such an original approach. Basically, four general ideas led me to develop this analysis:

- The demand to have some qualitative reinforcement in the study of MEPP (as we have seen in the analysis of the three WTA models, this demand is urgent and is required by various philosophers); quantitative restrictions are not enough to capture MEPP!

- The important role that some authors give to contextual factors in the analysis of explanatory power.

- The idea that the acceptance of the pluralist principle makes the explanation-scholar better suited to capture the intuitions coming from the scientific practice (implicit in my first claim above, in thesis α).

- The idea that the driven-force in the study of MEPP must come from scientific practice itself (for instance, from physics itself). To fix a philosophical model and to fit a number of scientific examples into its framework does not do justice to the rich scientific methodology and this should be regarded as an *a-priori* operation. We must focus on scientific practice itself, and then go on philosophically. In other words, our methodology must follow a bottom-up direction.

If we follow Wittgenstein, we will enter into our room to philosophize (about MEPP) and we will bring our umbrella with us. Our umbrella will be the observation of our scientific practice.

Chapter 7

WTA models and qualitative reinforcements

In this short chapter I will defend the following thesis:

α The three WTA accounts analyzed in the first part do not capture some cases recognized as genuine MEPP in scientific practice. This lack of success might be identified with the fact that, contrary to what the scientific practice itself suggests, those models do not consider qualitative factors as essential to MEPP. As a consequence, the models should be refined, in order to capture the test cases proposed, or rejected in favour of a different approach.

Thesis α contains two subthesis:

α_1 The WTA models seen in the first part do not account for some cases recognized as genuine MEPP in scientific practice.

α_2 Scientists do attribute an essential role to some qualitative factors which operate in the explanation provided. Nevertheless, those qualitative factors are not captured by the three WTA models.

As the numerous criticisms proposed in the first part have shown, there is an unquestionable link between explanation and qualitative (rather than purely quantitative) factors in the practice of scientists who explain phenomena in science. In this chapter, as an example, I will show how the three major winner-take-all models of MEPP fail in considering as genuine a MEPP recognized by scientists as such. This lack of success might be due to the fact that the three models were designed to capture the notion of explanation according to a formal schema or through the identification of some objective property, and specific qualitative factors coming from the scientific practice were not sufficiently taken into account in them.

As a first step, I am going to present a case of MEPP recognized as such in scientific practice: the behaviour of Hénon-Heiles systems explained via the phase space formalism. This example has been sketched in section 2.2, but it will be discussed here in a more comprehensive manner. Note that to take the behaviour of Hénon-Heiles systems as a test case of MEPP does not represent a novelty. The example of Hénon-Heiles system has been used by Lyon and Colyvan in their [Lyon *et al.*, 2008]. However, these authors discussed it in the context of the nominalist-platonist debate in philosophy of mathematics. My analysis here points to aspects which are quite far from Lyon and Colyvan's ontological considerations. I will not concentrate on the role that MEPP like this are supposed to play in the ontological dispute, but rather on why this case is regarded as a genuine explanation by scientists. This marks an essential difference in the analysis and in the directions of investigations which follow. In the second place, in section 7.2 I will show that the three WTA accounts fail in considering this as a genuine MEPP, thus contradicting the intuitions of scientists. In my assessment I will not evaluate the two pluralist accounts introduced in the second part, and this simply because the MEPP considered is neither a case of abstract explanation nor a case of asymptotic explanation. The accounts proposed by Pincock and Batterman do not aim at accounting for cases of MEPP such as that of the Hénon-Heiles system. Finally, in section 7.3, my diagnosis for this failure will confirm the moral of some criticisms which have been addressed to the three WTA models in part I: qualitative ingredients do play an essential role in

MEPP[1]. This consideration will establish a bridge between this chapter and the final chapter, in which I am going to propose an approach to MEPP in which these qualitative ingredients are taken into account.

7.1 Hénon-Heiles systems

Four decades ago, Michel Hénon and Carl Heiles were investigating the motion of stars about the galactic center. Rather than studying the problem with the actual potential of the galaxy (something which would have been quite difficult to achieve!), they restricted the motion to the xy plane, as in the Kepler problem, and studied a relatively simple analytic potential $U(q_x, q_y)$ [Hénon *et al.*, 1964]. This potential, called 'Hénon-Heiles potential', exhibits two cubic perturbation terms which couple together two standard harmonic oscillators:

$$U(q_x, q_y) = \frac{1}{2}(q_x^2 + q_y^2) + q_y q_x^2 - \frac{1}{3}q_y^3\,, \tag{7.1}$$

where the coordinates q_x and q_y are called 'generalized coordinates' (I will come back to these notions in the next paragraphs). Accordingly, we call "Hénon-Heiles systems" those systems formed by a particle moving in such a bidimensional potential.

Consider now the physical phenomenon under study, i.e. the motion of one particle moving in the Hénon-Heiles bidimensional potential $U(q_x, q_y)$. And take the motion of the system as our explanandum. More precisely, we want to explain the behaviour (regular or not) of the system for different energies.

There are two mathematical ways to study the system. We can study the system through the Lagrangian analysis. Or, alternatively, we can study it through the Hamiltonian formulation which comes with a particular mathematical structure called *phase space*. The Lagrangian formulation is obtained

[1] I consider quantitative factors as those factors which can be captured through a formal scheme or analysis. On the other hand, qualitative factors are pragmatic factors, which cannot be captured through such a formal scheme or analysis.

introducing the Lagrangian function $L = T - U$, where T is the kinetic energy of the system, and successively obtaining the equations of the motion from the so called *Lagrange's equations*:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \tag{7.2}$$

In the Lagrangian formulation (nonrelativistic), a system with n degrees of freedom possesses n (second-order) differential equations of motion of the form (7.2)[2]. The state of the system is represented by a point in an n-dimensional *configuration space* whose coordinates are the n generalized coordinates q_i (q_x and q_y for the present bidimensional example). The motion of the system (as a function of time) can be interpreted as the path traced by this point as it traverses the configuration space. In the Lagrangian formulation, all the n coordinates must be independent. However, there is another formulation of the problem. This formulation is called the Hamiltonian formulation, and it is "based on a fundamentally different picture" [Goldstein *et al.*, 2001, p. 335]. In the Hamiltonian formulation, we want to describe the motion in terms of first-order equations of motion. In order to do that, we double our set of independent quantities (thus obtaining $2n$ independent variables) by adding to our generalized coordinates q_i the new variables *conjugate* (or generalized) *momenta* p_i, which are defined as follows:

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \tag{7.3}$$

The quantities (q, p) are known as *canonical variables*.

From a mathematical point of view, the transition from the Lagrangian to the Hamiltonian formulation corresponds to changing the variables in our mechanical functions from $(q, \dot{q}, t)$ to (q, p, t), where p is related to q and $\dot{q}$ by equation (7.3). The procedure for switching variables in this manner, which also gives the so called *Hamiltonian function* associated, is provided by a

[2]Consequently, for a system of N particles in physical space, the system of equations above is a system of $3N$ equations.

Legendre transformation:

$$H(q,p,t) = \sum_{k=1}^{n} p_i \dot{q}_i - L(q,\dot{q},t) \tag{7.4}$$

Now, if we consider the differentials of the Lagrangian $L(q,\dot{q},t)$ and of the Hamiltonian (7.4), we will obtain the $2n+1$ relations

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \tag{7.5}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{7.6}$$

$$\frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t} \tag{7.7}$$

The first two equations above are known as *Hamilton's canonical equations of motion*. They are the desired set of first-order equations of motion which replace the n second-order Lagrange equations.

The space of the q and p coordinates is known as the *phase space*[3]. Hence, the $2n$ canonical equations of the motion describe the behavior of the system point in the *phase space*, which has $2n$-dimensions and whose coordinates are the $2n$ independent variables q_i, p_i. In other words, in the Hamiltonian formulation of mechanics the dynamics of the system is defined by the evolution of points ('trajectories') in this phase space.

For the case of our system, i.e. a particle moving in the bidimensional potential (7.1), the Hamiltonian function will be:

$$H = T + U = \frac{1}{2}(p_x^2 + p_y^2) + \frac{1}{2}(q_x^2 + q_y^2) + q_y q_x^2 - \frac{1}{3}q_y^3 \tag{7.8}$$

And the respective (nonlinear) equations of motion:

[3]Since in doubling our independent quantities we have chosen half of them to be the n generalized coordinates q_i, the configuration space can be thought of as the half of the phase space that contains the position coordinates q_i.

$$\frac{d^2 q_y}{dt^2} = \frac{dp_y}{dt} = \frac{\partial H}{\partial q_y} = -q_y - q_x^2 + q_y^2 \tag{7.9}$$

$$\frac{d^2 q_x}{dt^2} = \frac{dp_x}{dt} = \frac{\partial H}{\partial q_x} = -q_x - 2q_x q_y \tag{7.10}$$

These equations may be obtained from either Lagrange's equations or Hamilton's equations. Although both the routes are admissible, however, scientists agree that the mathematical procedure involving the Lagrangian formalism does not convey the sense of explanatoriness that we obtain from the use of Hamiltonian formalism involving the phase space. This is because by using the Hamiltonian formalism we can visually deduce, from a representation in the phase space, if the system has a regular or chaotic motion[4]. To be more precise, in order to do that we start by considering the total energy of the system E constant, and thus we lower the dimensionality of the phase space by one:

$$E = \frac{1}{2}(p_x^2 + p_y^2) + \frac{1}{2}(q_x^2 + q_y^2) + q_y q_x^2 - \frac{1}{3}q_y^3 \tag{7.11}$$

We then take a 2-dimensional cross section (*surface of section* or *Poincaré section*) of this hypersurface in the phase space and we map the intersections of the trajectories with the plane by using a function called *Poincaré map* (Figure 7.1). Finally, we look at the "dots" made by the solutions (orbits) on the surface of section and we can visually grasp qualitative informations about the dynamics of the system. We do that by following the order in which the dots appear. Solutions that never pass through the same arbitrarily small neighborhood of a point twice are chaotic (instead of following a regular curve, they are scattered and jump around in a more or less random fashion from one part of the Poincaré section to another). On the other hand, a dynamic

[4]Chaos is a motion which is, simultaneously: (a) irregular in time (it is not simply the superposition of periodic motions, it is really aperiodic); (b) unpredictable in the long term and sensitive to initial conditions; (c) complex, but ordered, in the phase space (it is associated with a fractal structure) [Tamás and Márton, 2006, p. 22].

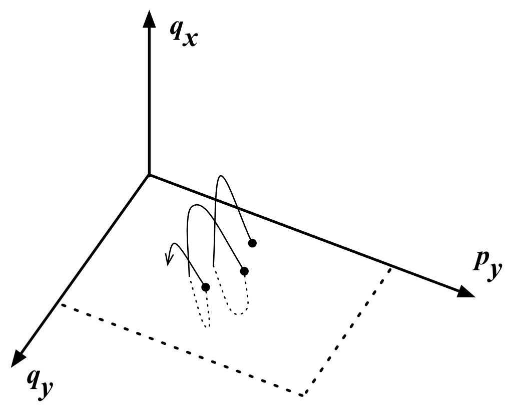

Figure 7.1: The $q_x = 0$ plane is a 2-dimensional cross section of the phase space and is called surface of section or Poincaré section. Poincaré map $p : S \longrightarrow S$ associates successive intersections of a trajectory with the $q_x = 0$ plane in the direction of increasing q_x.

state that gives rise to regular motion will have the property that nearby dynamic states will stay close to it as they get mapped around the plane. We use the mathematical 'resource' of Poincaré map because in the Hamiltonian formulation of our problem the phase space is 4-dimensional (points in the phase space are represented by quadruples of the form (q_x, q_y, p_x, p_y)) and therefore in this space the trajectories (which define the dynamics of the system) are not directly visualizable on a diagram.

The Poincaré sections for various energies summarize the dynamics of the system at that energy. Studying the diagrams for different energies we can observe how at low energy the section is dominated by regular orbits (and then the associated motion is regular), at intermediate energy the section is divided more or less equally into regular and chaotic regions, while as we increase the total energy of the system the orbits become chaotic and the section is dominated by a single chaotic zone (Figure 7.2)[5].

The importance of explaining the behavior of the system through this

[5]Such transitions from regular to chaotic behavior are quite common; similar phenomena occur in widely different systems, though the details naturally depend on the system under study.

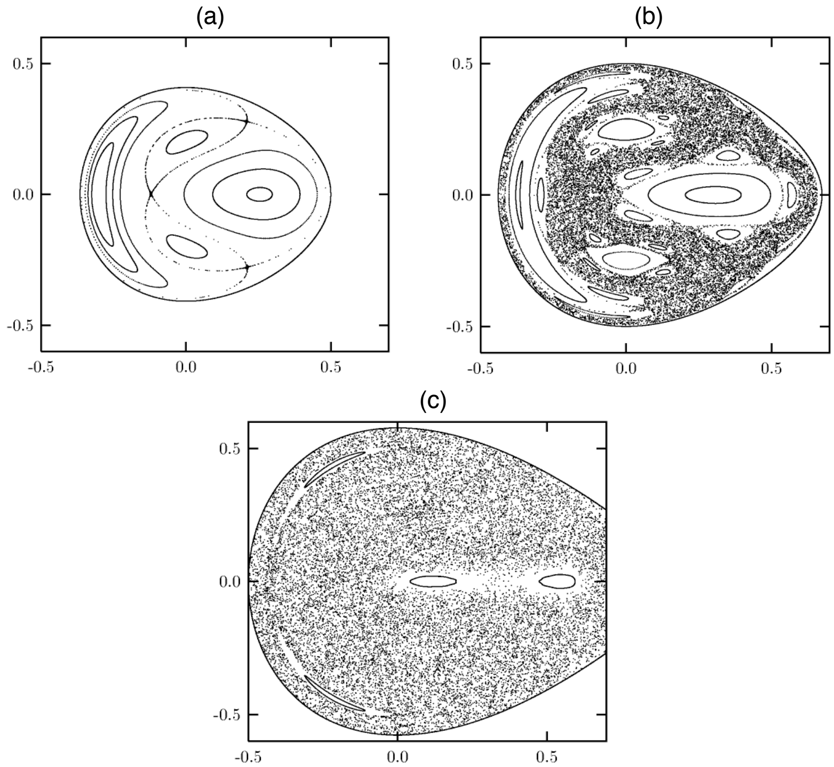

Figure 7.2: Poincaré sections of the Hénon-Heiles system in the $q_y p_y$ plane ($q_x = 0$), showing several Henon-Heiles orbits. For $E = \frac{1}{12}$ orbits are regular (a); for $E = \frac{1}{8}$ we observe regions of regular motion and regions of chaos (b); finally, for energy $E = \frac{1}{6}$ chaos is dominant (c). Based on [Hénon *et al.*, 1964].

procedure is well recognized by scientists. For instance, just to cite a popular textbook on mechanics and dynamical systems:

> It is in such maps [Poincaré maps] that the fractal structure of chaotic dynamics becomes plausible. Only in special cases (like those of the magnetic pendulum, the mirroring spheres and advection) can fractal structures be observed in real space. Therefore the use of phase space is inevitable as a means of understanding the structure accompanying chaos [Tamás and Márton, 2006, p. 22]

Thus the phase space, together with the Poincaré map, is recognized to have an explanatory role[6]. It permits to say *why* at that particular energy the system has that particular behavior. We do not explain the behavior of the system by tracing causal processes and interactions, but the explanation is carried out by essential appeal to mathematics (the phase space and the Poincaré map).

It is important to note, again, that the mathematical procedure involving phase space is not the only alternative for the study of the system. It is possible to analyse it via the Lagrangian formalism, although this route does not seem to carry the sense of explanatoriness that we obtain from the use of phase space theory.

> [...] although there is a Lagrangian formulation of the theory in question that does not employ phase spaces, the cost of adopting such an approach is a loss of explanatory power [Lyon *et al.*, 2008, p. 2]

To sum it up, two mathematical formalisms are acceptable to study the physical phenomenon (regular or chaotic motion of the particle moving in the bidimensional potential 7.1), but only one of them –the Hamiltonian formalism including Poincaré map– contributes to the MEPP because it permits us to visually grasp the behaviour of the system.

7.2 Testing the accounts

It is now time to ask ourselves the following question: Are the three WTA accounts presented above able to account for the mathematical explanation of the behavior of the Hénon-Heiles system?

[6] Observe that, with respect to my claim here and to the previous quotation, I am considering that to have an understanding of a phenomenon amounts to having a genuine explanation of the phenomenon. This is why I consider that phase space, together with Poincaré map, have an explanatory role. I will address the linkage understanding-explanation, thus providing a justification for the present claim, in subsection 8.6.1.

Matching Steiner's model

Let's begin with Steiner's model of MEPP illustrated in chapter 1. Recall that Steiner's account of MEPP depends on his account of mathematical explanations in mathematics, which has been proposed by the author to cover explanations coming under the form of proofs. In particular, for Steiner we have a MEPP when the following criterion is satisfied:

C_{MEPP} If we remove the physics (physical assumptions or bridge principles) we rest with a mathematical proof which satisfies criteria C_1 and C_2

where C_1 and C_2 are his criteria for explanatory proofs:

C_1 The proof depends on a characterizing property mentioned in the theorem (dependence criterion)

C_2 It is possible to deform the proof "substituting the characterizing property of a related entity" and getting "a related theorem" (generalizability criterion)

In order to be applied, Steiner's C_{MEPP} requires a mathematical theorem and a relative proof. Nevertheless the mathematical explanation of why Hénon-Heiles systems exhibit regular or chaotic motion does not come under the form of a mathematical theorem (together with the relative proof). Therefore Steiner's account for MEPP should be considered to have no application for cases as this one.

However, it might be observed that for Steiner having a proof is not a necessary condition in order to have a mathematical explanation within mathematics, i.e. that he accepts the existence of explanations in mathematics which do not come under the form of proof. In fact, in a passage of his paper "Mathematics, explanation and scientific knowledge" [Steiner, 1978b], Steiner claims that we have a mathematical explanation in physics when removing the physics (physical assumptions or bridge principles) we rest with a "mathematical explanation of a mathematical truth" [Steiner, 1978b, p. 18]. In this passage, the "mathematical truth" to be explained is not identified with the result of a proof, then it might be thought that Steiner is referring here to explanations in mathematics which are not proof-explanations. This reading would provide a potential way to defend Steiner from what I said in the previous paragraph[7]. In order to test his model, this possibility must be taken into account.

Let's then assume that Steiner's account can cover MEPP which do not directly involve mathematical proofs. I will show that, even with this assumption, Steiner's model must be considered to have no application for the case presented in the previous section. As it will emerge from my discussion below, a general difficulty for the account comes from the fact that we are not faced with a theorem and therefore the choice of the mathematical truth that the mathematical formalism is supposed to explain turns out to be extremely subject to arbitrariness.

To "remove the physics" in the case of the Hénon-Heiles system is to consider that are left with the first-order differential equations (7.5) and (7.6). As we have seen above, the mathematical study of the solutions of this system is made by fixing the value of E, and introducing a 2-dimensional cross section in phase space. Finally, we map the intersections of the orbits with the plane by using the Poincaré function. This is *only* mathematics, and no physical assumptions are required to proceed in this way. According to this 'rescue'-reading of Steiner, a MEPP requires that we remain with a mathematical explanation of some mathematical truth. Now, by considering that the dots on the Poincaré section are not (physical) dynamical states of the system but intersections of the solutions of our system of differential equations with a section of the plane in phase space, we are left with a particular succession of dots on a plane. Does this represent a mathematical explanation of a mathematical truth? Is the proposition 'the succession of the dots

[7]Note, however, that Steiner's examples of explanation in mathematics always concern proofs (see his [Steiner, 1978a], where the examples are given). This suggests that for him a mathematical explanation of a mathematical truth is an explanatory mathematical proof. Furthermore, it is natural to identify a 'mathematical truth' with the result of a proof. Otherwise, what would "mathematical truth" stand for? Steiner does not offer any definition.

is so and so...' the mathematical truth we are explaining here? Certainly not. However, we might reason in the following way.

The differential equations (7.5) and (7.6) describe the (continuous) time evolution of our system. A Poincaré map can be interpreted as a discrete dynamical system with a state space that is one dimension smaller than the original continuous dynamical system. It preserves many properties of periodic and quasiperiodic orbits of the original system (differential equations 7.5) and has a lower dimensional state space, and this is why we use it for analyzing the original system (even if, in practice, this is not always possible as there is no general method to construct a Poincaré map). The stability of a periodic orbit of the original system is closely related to the stability of the fixed point of the corresponding Poincaré map. A periodic orbit of our original continuous dynamical system is stable if and only if the fixed point of the Poincaré map is stable[8]. This is a mathematical theorem [Arnold, 1992, p. 258]. Nevertheless the stability of a periodic orbit is a mark of regular behavior, which is something we want to explain for our system at a particular energy E fixed. Hence, in our case, the mathematical truth to be considered here might be the proposition 'the fixed point of the corresponding Poincaré map is stable or not', which is a pure mathematical statement because stability can be defined analytically. However, observe that this result (stability of periodic orbits) is obtained by computation (by computing the orbits through the numerical integration of the equations of motion, for instance using the Runge Kutta integration scheme) and not via analytical methods[9]. The same holds for the investigation of quasi-periodic orbits (the

[8]The stability (or *Lyapunov* stability) of an orbit of a dynamical system characterizes whether nearby orbits will remain in a neighborhood of that orbit without being repelled away from it. The same idea characterizes the stability of fixed points of a mapping: a fixed point $\mathbf{x_0}$ of the mapping A is called *Lyapunov stable* (or *stable*) if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $|\mathbf{x} - \mathbf{x_0}| < \delta$ implies $|A^n\mathbf{x} - A^n\mathbf{x_0}| < \varepsilon$ for all n, $0 < n < \infty$ simultaneously.

[9]For fixed $E = 1/12$, Hénon and Heiles found that the stable fixed points of the Poincaré map, corresponding to stable periodic orbits, are located near the middle of the nested, closed curves (Figure 7.2a). To obtain this result they calculated the four regions with oval-shaped orbits for smaller and smaller circumferences, and they observed that they shrank to four fixed points [Hénon *et al.*, 1964, p. 75]. The point I want to stress here

continuous lines in Figure 7.2, which do represent regular behavior) and of the irregular aperiodic orbits (the "dots", representing chaotic behavior), which is performed by numerical analysis, for a number of values of E. Therefore, due to the mere calculatory aspect of these steps, it would be difficult to say that we are faced with a mathematical *explanation* of some mathematical truths. In fact, there is evidence that the explanatory activity in mathematics is driven by factors other than justificatory aims such as establishing the truth of a mathematical fact [Hafner *et al.*, 2005][10]. This conclusion is reinforced if we consider the explanation of the behaviour of the Hénon-Heiles system as given in the previous section, where the regular or chaotic behavior is deduced from qualitative considerations and analytical considerations about the stability or the periodicity of the orbits are not necessary. Consequently, to take into consideration such analytical aspects as to make Steiner's model fit with the test-case would represent an *ad hoc* move, which does not mirror the practice of scientists.

Let me add a remark. The very same test-case concerning the behavior of Hénon-Heiles systems has been used in section 2.2 to show that the D-N Extended (deductive-nomological extended) model is not apt to capture MEPP. In particular, I suggested that, when faced with this case of MEPP, the D-N Extended is confronted with two major problems. First, it cannot deal with mathematical operations or procedures (which do not come under the form of statements). Second, even if we would have these procedures or operations under the form of theorems, the model would lack in resources to discriminate between the explanatory mathematical procedure and the non-explanatory one. And this because, since the two mathematical procedures (the Lagrangian and the Hamiltonian) are both formally correct, the

is that we cannot find the difference equations of the Poincaré map for our Hénon-Heiles system. The Poincaré map is found only by solving numerically the differential equations of the system, finding the successive intersections of various trajectories with the surface of section. Is this a mathematical explanation?

[10]For instance, when confronted with a theorem, the mathematician prefers a particular proof-strategy or procedure because that provides more than a mere justification of the mathematical truth ([Kitcher, 1984],[Sandborg, 1998], [Tappenden, 2005]).

model would consider both as explanatory and would regard them as good ingredients for its deductive schema. Therefore it seems that the fallibility of Steiner's model and that of the D-N Extended derive from a very similar cause. The D-N Extended fails to account for the Hénon-Heiles MEPP because this explanation appeals to a factor (the possibility of visualizing a particular state of affairs) which cannot be captured by a purely formal deductive structure. On the other hand, Steiner's model fails in capturing the MEPP because there is nothing like a formal proof from which it would be possible to derive the explanandum. But every formal proof inevitably follows a deductive schema, therefore the same 'purely formal deductive ingredient' required by the D-N Extended model is required by Steiner's model as well. Hence, the common insufficiency seems to come from the fact that both the two accounts consider deductive (mathematical or law-based, in the case of the D-N model) arguments as central to explanation.

Matching Kitcher's unification model

In chapter 3 I illustrated the unification model proposed by Kitcher and I presented several criticisms. The case used by Hafner and Mancosu to test Kitcher's model in mathematics concerned Brumfiel's work on real closed fields. In that particular case, Brumfiel explicitly indicates that he adopts a certain method of proof even though another method that would provide a more unified method of proof is available as well (here 'unified' is used in Kitcher's sense). This would undermine Kitcher's proposal that unification is the key to explanation not only in science, but also in mathematics.

Another observation against the possibility of using Kitcher's model to account for cases recognized as explanatory in mathematical practice has been proposed by Jamie Tappenden. By giving a short analysis of the unification approach for cases of mathematical explanation within mathematics, Tappenden observed that existing accounts of unification are more balanced on quantitative restrictions (for instance, the quantity of patterns in Kitcher's formulation) and need to be supplemented with qualitative reinforcements.

These authors (Mancosu, Hafner, Tappenden) agree on demanding Kitcher's theory to introduce some qualitative reinforcement in its apparatus in order to reflect the actual mathematical practice. Furthermore, as we we have seen in chapter 6, Redhead has observed how, in case such as Batterman's example of critical phenomena, the unification approach proposed by Friedman and Kitcher is useless because that account has been created to fit together *different* phenomena into unified patterns of explanations, when on the other hand universality concerns *similar* phenomena which come from different causal stories [Redhead, 2004, p. 529]. Moreover, also Batterman considered that these unification models are not well suited to account for critical phenomena. In particular, Batterman remarks that unification models are not able to answer the specific why-questions which concern these phenomena. These why-questions ask for the explanation of why patterns of a given type can be expected to obtain[11].

In this subsection I will concentrate on the case of Hénon-Heiles system presented above, and I will show that Kitcher's account is faced with evident difficulties when it is demanded to account for such a MEPP. By paralleling what Kitcher considered as a paradigmatic case of explanation by unification (that given Newtonian theory) and the example considered above (Hénon-Heiles system), I am going to defend the idea that Kitcher's unification approach lacks resources to account for the latter MEPP. To anticipate my argument, I will point to the fact that the regular or chaotic behavior of the system is explained by appealling to the possibility of visualizing the trajectories on the surface of section. According to scientists, this is an essential ingredient in the explanation provided. However, such an inferential step (the inferential step in which we infere the regular or chaotic behaviour of the system by visualizing the trajectories) cannot be modelled by Kitcher's idea of argument pattern, which is designed according to a formal deductive schema.

[11] Nevertheless, as I remarked in the Introduction to this dissertation, Batterman does not provide any 'robust' assessment of Kitcher's model but only a sketch (see [Batterman, 2002a, p. 32-33]).

In chapter 3 we have seen that the explananda considered in Kitcher's unification model are members of K. Informally, we can think of K as the set of statements endorsed by an ideal scientific community at a specific moment in time. A statement in K, for instance that of the form "Object O_1 has position P_1 and velocity v_1 at time t_1", is derived in Kitcher's model by using a particular argument pattern (for example, the Newtonian pattern). Furthermore, the same pattern is used to derive statements which do represent different phenomena. For instance, the Newtonian pattern is also used to derive the statement "Object O_2 has position P_2 and velocity v_2 at time t_2", which refers to a physical phenomenon different from that considered by the previous statement having the same form. The unification model is then able (if we agree with Kitcher) to tell that statements which represent different phenomena are derived from arguments that instantiate a *common* argument pattern[12].

Now, in order to apply Kitcher's account to the present case of MEPP, two essential requirements should be fulfilled. First, the pattern of derivation used in the MEPP of the Hénon-Heiles system must represent an instance of a pattern used in deriving statements which concern the behavior of other physical phenomena as well. Otherwise, there would not be the unification idea that Kitcher assumes for explanation, and my testing would result trivial. Second, the statement concerning the behaviour of Hénon-Heiles system must belong to the set K of statements accepted by a scientific community at a particular time. For simplicity, let me indicate with the expression 'behaviour-statement' a statement concerning the behaviour of a physical system[13]. In our case, the behaviour-statement to consider is the following:

[12]Consider, for instance, that O_1 is a ball and O_2 is a satellite. The orbiting of the satellite around the Earth and the falling of the ball from a tower are different phenomena. However, according to Kitcher, they are covered (and unified) by the same Newtonian pattern, i.e. the arguments from which we derive the two statements "Object O_1 has position P_1 and velocity v_1 at time t_1" and "Object O_2 has position P_2 and velocity v_2 at time t_2" do instantiate the same (Newtonian) argument pattern.

[13]Intuitively, statements such as "Mark played his new guitar during the concert" are not among the behaviour-statements I am considering here.

S_1 'The particle P has regular (or chaotic) behaviour at energy E'

Fortunately, both of the requisites are met: the behaviour-statement S_1 belongs to K; second, there exist behaviour-statements concerning physical phenomena different from that related to S_1, and these statements are deduced via the same procedure used to derive S_1. With respect to the former, the behaviour-statement "The particle P has regular (or chaotic) behaviour at energy E" belongs to K because it expresses a belief shared by scientists (it comes from our scientific practice, and it expresses a belief which is shared by the scientific community at a particular time). Furthermore, it is easy to see how the second requisite is satisfied as well. The very same procedure involving the Hamiltonian formalism, the phase space and the Poincaré map, can be used to infer statements concerning the behaviour of physical phenomena different from the motion of a particle in a potential. For instance, the same procedure can be used in the case of the double pendulum, which is a different physical phenomenon[14].

If distinct behaviour-statements about the different phenomena are inferred through the very same procedure, according to Kitcher there should be an argument pattern which is used to derive these behavior-statements, thus providing unification. This suggests that the notion to be checked here is that of Kitcher's argument pattern.

To test Kitcher's notion of argument pattern for the present case amounts to answering the crucial question: is there a pattern (in Kitcher's sense) which is able to instantiate the Hénon-Heiles derivation and the derivation of the double pendulum? If yes, by considering the case of Hénon-Heiles system, an instantiation of this pattern would account for the particular derivation of the statement "The particle P has regular (or chaotic) behaviour at energy E". In my testing below I am going to show that Kitcher's idea of argument pattern does not capture this particular derivation. More generally, I will

[14] In the same way of the Hénon-Heiles system, the regular or chaotic behaviour of this system can be established by using the Hamiltonian formalism and then looking at the trajectories made on the surface of section. See chapter 5 of [Korsch *et al.*, 2008] for a study of the double pendulum.

point to the fact that Kitcher's argument pattern does not admit such types of derivations within its structure, thus constrasting with the intuitions of the scientists who consider this derivation as a genuine MEPP.

For Kitcher, a *general argument pattern* $\langle s, f, c\rangle$ is a triple consisting of a *schematic argument* s, a set f of filling instructions and a classification c for s. In the Newtonian case, the following schematic sentences (1)-(5) form a schematic argument s_N:

1. The force on α is β
2. The acceleration of α is γ
3. Force = mass $\cdot$ acceleration
4. (Mass of α)$\cdot$(γ) = β
5. $\delta = \theta$

The members of the set of filling instructions f_N are: "all occurrences of α are to be replaced by an expression referring to the body under investigation"; "occurrences of β are to be replaced by an algebraic expression referring to a function of the variable coordinates and of time"; "γ is to be replaced by an expression which gives the acceleration of the body as a function of its coordinates and their time-derivatives"; "δ is to be replaced by an expression referring to the variable coordinates of the body, and θ is to be replaced by an explicit function of time". The set of filling instructions f_N contains the directions for replacing the dummy letters α, β, γ, δ, θ in every schematic sentence. The sentences contained in the classification set c_N for the schematic argument s_N give us the inferential information about the schematic argument: "(1)-(3) have the status of premises"; "(4) is obtained from (1)-(3) by substituting identicals"; "(5) follows from (4) using algebraic manipulations and the techniques of the calculus". Thus we have that a particular derivation in Newtonian mechanics, i.e. a sequence of sentences and formulas which accord Newton's laws, instantiates the general argument pattern $\langle s_N, f_N, c_N\rangle$ just in case: (*i*) the derivation has the same number of terms as the schematic argument s_N, (*ii*) each sentence or formula in the derivation can be obtained from the corresponding schematic sentence in accordance with the filling instructions f_N, (iii) the terms of the derivation have the properties assigned by the classification c_N to members of the schematic argument s_N. The unifying power of Newton's theory consists in the fact that, by using the Newtonian pattern $\langle s_N, f_N, c_N \rangle$ again and again, the theory shows us how to derive a large number of statements accepted by the scientific community.

Let's now consider the laws of mechanics and the theory of differential equations as belonging to the corpus K of our beliefs. We want to construct an argument pattern (of the same kind of Kitcher's Newtonian pattern) which does instantiate the particular derivations which lead to the following behaviour-statements:

S_1 'The particle P has regular (or chaotic) behaviour at energy E'

S_2 'The double pendulum S has regular (or chaotic) behaviour at energy E'

Every statement above (S_1, S_2) is accepted in K, and each statement concerns a different physical phenomenon (respectively, the motion of a particle in a potential and the motion of a double pendulum). The behaviour-statement S_1 and S_2 are both obtained by finding the equations of motion, constructing the Poincaré section with the relative map, and finally grasping visually the behaviour of the system for fixed energies. Observe that there are plenty of behaviour-statements (concerning different physical phenomena) that can be obtained through the same procedure. For instance, statements about the regular or chaotic behaviour of the voltage in a triode circuit (modelled by the Van der Pol oscillator), about the behaviour of a spring pendulum whose spring's stiffness does not exactly obey Hooke's law (modelled by the Duffing oscillator), or even about the behaviour of the simple pendulum or the undamped spring-mass system (the latters modelled by a simple harmonic oscillator)[15]. The list of statements above contained

[15]In the latter unidimensional cases (simple pendulum and undamped spring-mass system) the surface of section coincides with the whole phase space. To study the trajectories

only two of them. Now, if the steps in the derivation involving the equations of motion and the construction of the Poincaré section can be mirrored by an argument pattern *à la* Kitcher, the inferential step which appeals to the possibility of visualizing the trajectories cannot. This is manifest if we look at the Newtonian pattern presented above. How would such a (visual) inferential step appear in the schematic argument s? And what kind of filling instructions and classification would be able to capture it? The derivation which is performed in the Newtonian case can be mirrored by a formal deductive schema of the kind Kitcher proposes, but the derivation used in the MEPP of Hénon-Heiles system does appeal to an ingredient which cannot be mirrored by the idea of argument pattern. Moreover, this ingredient is recognized as essential by scientists and therefore it is reasonable to include it in the pattern-structure. If there exits a common pattern for behaviour-statements like S_1 and S_2, then, it seems that this pattern has a structure essentially different from that of the Newtonian pattern $\langle s_N, f_N, c_N \rangle$ proposed by Kitcher.

To sum up, an argument pattern *à la* Kitcher does not admit inside its structure a particular inferential step recognized by scientists as essential to the explanation provided. It is then inappropriate to use it for derivations of behaviour-statements such as that concerning the Hénon-Heiles system, or for similar derivations of behaviour-statements concerning different phenomena. Hence Kitcher's unification model, at least in its original form, is not able to account for the MEPP concerning the behaviour of the Hénon-Heiles system[16].

on the surface of section is the same as to study the trajectories in the phase space.

[16]Observe that the fact that Kitcher's argument pattern does not reflect the kind of inferences made in the Hénon-Heiles example is sufficient to show the inapplicability of Kitcher's account to the present case. Further considerations about the number of conclusions generated by the pattern are not necessary once the basic idea of argument pattern comes as inapplicable to our case.

Matching the pragmatic model

In section 2.4 I showed how the PET account, i.e. Van Fraassen's account extended to the case of MEPP, had problems in accounting for Baker's example of MEPP. Furthermore, from that discussion I advanced an argument against the possibility of using PET as model of MEPP. The motivation for discarding the traditional pragmatic account, as well as PET, as potential accounts of MEPP came from an observation I made in section 2.5. First, in line with Mancosu's intuition [Mancosu, 2008b, p. 140], I assumed that a theory of mathematical explanations of scientific phenomena might be not completely independent of a theory of mathematical explanation of mathematical facts. Therefore I considered the traditional pragmatic account from this perspective (by considering it as a model for explanation within mathematics). However, as seen from my considerations, Van Fraassen's theory is not suitable to deal with mathematical explanation within mathematics, thus the pragmatic model should also be rejected as a model of MEPP. The alternative strategy was to consider the extension of the pragmatic theory (I called PET this extension), assume some sort of methodological continuity between the mathematical word and the physical one, and focus directly on MEPP. But also this extension came out as problematic because it could be made only via two approaches: define an objective relevance relation, but this is exactly the problem of determining what counts as an objective criterion of explanatoriness (something Van Fraassen does not want and firmly excludes from his account!); or assume Van Fraassen's theory as a theory through which one evaluates the relevance of answers for the questioner (in the context of MEPP). The latter case, however, also turned out to be insufficient for our purposes. In fact, when we have at disposition only *one* mathematical theory of the physical phenomenon, the questioner has only one mathematical answer at her disposition and the PET does not tell us anything interesting about MEPP. In both cases we are faced with the problem of evaluating the mathematical argument, and this is something that creates problems to Van Fraassen's model as extended to MEPP. In this subsection I will take into consideration Van Fraassen's original model and I will show that it is not able to account for the Hénon-Heiles' MEPP for a reason different from those reported in section 2.4. My argument below will reinforce my considerations about the difficulty a why-question approach has when faced with MEPP.

In the case of Van Fraassen's model, the model looks at explanations as answers to why-questions and it needs a relevance relation in order to account for them. Consider our why-question 'Why does Hénon-Heiles system have chaotic behaviour at energy E?', where the topic P_k is 'Hénon-Heiles system has chaotic behaviour at energy E'. In the contrast class X we will have, together with the topic, the alternative proposition 'Hénon-Heiles system has regular behaviour at energy E'. According to what we saw in the previous section, the answer B to the why question is given by the following proposition: 'Because solutions –the dots– are scattered on Poincaré section. They never pass through the same arbitrarily small neighborhood of a point twice'.

Now, consider our answer B in the form 'Because A', where A is the proposition 'Solutions –the dots– are scattered on Poincaré section. They never pass through the same arbitrarily small neighborhood of a point twice'. According to Van Fraassen, B is a *direct answer* (an explanation) to our why-question $Q = \langle P_k, X, R \rangle$ if the following conditions are both satisfied:

(i) proposition A bears a relation R to $\langle P_k, X \rangle$

(ii) B is the proposition which is true exactly if: the topic P_k is true; only the topic is true in the constrast class X (formed only by the two propositions 'Hénon-Heiles system has chaotic behaviour at energy E' and 'Hénon-Heiles system has regular behaviour at energy E') ; and A is true.

There are at least two problems with this approach as applied to our test-case, namely: a difficulty in determining the relevance relation R and a problem in what a why-question presupposes. Both point to the fact that the explanation considered cannot be captured by a why-question analysis. Here I will only address the second problem (this will be sufficient, however, to reject Van Fraassen's model as suitable to cover the present case).

Consider again the second part of Van Fraassen's definition of direct answer: B is a *direct answer* to Q if (ii) B is the proposition which is true exactly if: P_k is true; only P_k is true in X; and A is true. Moreover, keep in mind that we want Van Fraassen's account to agree with scientific practice in considering the procedure which involves phase space and the Poincaré map as a genuine explanation. The difficulty for Van Fraassen's approach is noticeable. To say that only P_k is true in X (and then to regard the proposition "Hénon-Heiles system has regular behaviour at energy E" as false) is grounded exactly in our possibility of grasping qualitatively the behaviour of the system via the Poincaré map and claim that proposition A is true. In other words, we know that Hénon-Heiles system has chaotic behaviour at energy E because we use the Poincaré map and we infer visually that solutions never pass through the same arbitrarily small neighborhood of a point twice. Nevertheless this means that, according to a why-question analysis, the explanatory activity admitted assumes the form of a display of consequences (the topic P_k) of what we have already accepted as given (the proposition A). Very roughly, this would amount to saying: "this is an explanation because we have already accepted that we have an explanation". This was, in fact, the general moral of Sandborg's criticism reported in subsection 2.3.3. To use his words:

> The key point is that a why question is taken to implicitly fix the way an answer must regard its topic [Sandborg, 1998, p. 621]

Therefore either the why-question approach proposed by Van Fraassen is not able to account for our MEPP[17].

[17] The failure of the *why*-question approach for cases of MEPP such as the case of the Hénon-Heiles I propose here can be attributed to the fact that those cases do not come under the form of *why*-questions. As I showed above, to subsume our explanandum in a *why* question amounts to fixing 'a priori' the way the answer regards the topic. Perhaps, in cases

7.3 The moral: the importance of qualitative factors

A general conclusion which can be drawn from my testing above is that the three WTA models seems to be insufficient in capturing one MEPP recognized as such in scientific practice, and consequently should be abandoned or refined. Moreover, we have seen in section 7.1 that in such a MEPP it is possible to appeal to our ability to reason visually on a diagram. This qualitative component is essential to the explanation: by looking at the diagram in the right way we can explain why the system has a particular behaviour at fixed energies. This suggests that there is a link between explanation and specific qualitative (rather than purely quantitative) factors in the practice of scientists who explain phenomena in science. Despite having an essential role in the explanation provided by the scientists, however, the qualitative ingredient which is involved in the explanation of the behavior of the Hénon-Heiles system is not captured by the three WTA models, as my assessment shows.

Observe that my evalution of the models, and in particular my considerations about the inadequacy of the three WTA models to account for qualitative factors essential to the explanation, fit well with some of the major criticisms of the WTA accounts presented in the respective sections (section 1.4 for Steiner, section 3.3 for Kitcher and subsection 2.3.3 for Van Fraassen). Concerning Steiner, Michael Resnik and Johannes Hafner drew attention to the fact that context-dependence does affect explanation, but it is not possible to capture this contextual factor via some kind of purely quantitative factor or through the kind of 'objective property' Steiner proposes. Fur-

such as that of the Hénon-Heiles system, a more promising direction of analysis would be to adopt a *What*-question approach. For instance, re-formulating the explanatory question as "*What* is the behaviour of the system at Energy E?". A motivation for such a change of perspective comes from the observation that, just as it occurs for explanations given in commonly spoken language, explanations in science are associated with informative answers which are responses not only to why-questions, but also to *what* or *how*-questions ([Faye, 1999], [Sintonen, 1999]).

thermore, Johannes Hafner and Paolo Mancosu pointed out that Steiner's model of explanation does not account for the intuitions of the scientists in their practice. With respect to Kitcher's unification account, the same authors (Hafner and Mancosu) and Jamie Tappenden observed how the account must be emended in order to also consider qualitative (rather than purely quantitative) factors within its structure. Finally, David Sandborg advocated the introduction of qualitative factors ("conceptual resources not previously available") into Van Fraassen's model.

It can be noted that MEPP do not always involve visual reasoning (and our ability to reason visually) and it is recognized that there are MEPP that involve other qualitative ingredients which come under the form of peculiar forms of reasoning as well. An example is given by asymptotic reasoning. In chapter 6 we have seen how Robert Batterman has argued that particular species of MEPP, called "asymptotic explanations", gain their explanatory power by the systematic throwing away of various causal and physical details. By using asymptotic methods, i.e. methods which eliminate detail and precision, the scientist is able to obtain a particular mathematical explanation for a phenomenon (where the phenomenon considered was a pattern of behavior). The reasoning which is involved in asymptotic methods is called by Batterman asymptotic reasoning, and asymptotic explanation is exactly the kind of explanation which involves such specific kind of reasoning. As we have seen, the main example discussed by Batterman concerns the explanation offered in condensed matter physics for the so called universality of critical phenomena. In particular, for that case the mathematical technique of renormalization group theory is what permits us to reason asymptotically and obtain an explanation for the universality of critical phenomena. Note that Batterman's intuitions about the explanatory role played by such particular kind of reasoning is strongly supported by examples from scientific practice. As I remarked in chapter 6, the same high-energy theorist Kenneth Wilson, who formulated the renormalization group theory in general terms in 1971, in his Nobel lecture gave particular emphasis to the crucial role played by this theory (and then the asymptotic method) in the explanation of the universal behaviour of different systems [Wilson, 1971]. Moreover, this opinion is shared by other scientists as well. For instance, the mathematical physicist Michael Berry has recently drawn attention to the importance of asymptotic methods in physics during his talk "Emergence and asymptotics in physics: how one theory can live inside another" [Berry, 2009].

Another form of reasoning which is recognized to come as essential ingredient in our explanatory scientific practices is analogical reasoning. Analogical reasoning is the process of reasoning by analogy, i.e. reason and learn about a new situation (the 'target' analog) by relating it to a more familiar situation (the 'source' analog) that can be viewed as structurally parallel [Holyoak *et al.*, 1997]. While this particular kind of reasoning is used extensively in our everyday-life, some philosophers welcome the idea that the use of analogical reasoning in science does provide an essential contribute to scientific explanation [Hesse, 1966].

Far from giving a bestiary of the kinds of reasoning we find associated to MEPP in scientific practice, the purpose of this chapter is to suggest that MEPP involve specific qualitative ingredients which are recognized as essential by the scientific community and which come under the form of particular kinds of reasoning. More precisely, these kinds of reasoning come with an ability to reason. For instance, visual reasoning come with the ability to reason visually: when we reason visually on a diagram we use our ability to reason visually on that diagram. In Batterman's case, when we reason asymptotically we are using our ability to reason asymptotically (I will say more on this point in the next chapter). These qualitative factors are not captured by the three WTA models presented in the first part of this dissertation.

Now, these forms of reasoning which we find in MEPP are essentially distinct (for instance, analogical reasoning is distinct from visual reasoning) and it might be thought that they characterize different 'species' of MEPP. To accept this new perspective has an immediate consequence on our methodology and on the possibility to study MEPP by proposing a WTA model. In order to account for this variety, or species, of explanations we have to accept that pluralism is the best attitude to adopt (at least if we take the intuitions of scientists seriously, which is what I assume as a basic premise of my investigation). It is very hard, in fact, to see how this picture of MEPP can fit within the traditional WTA view. In order to account for this variety, or species, of explanations we have to accept that the pluralist principle provides a more promising way for the study of MEPP.

Observe that to defend the 'advantage' of a pluralist approach is not only to say that it is preferable because the WTA conception of explanation (as expressed by the three models analyzed here) is in trouble when faced with MEPP which appeal to such qualitative ingredients. Perhaps, another point which supports the advantage of the pluralist way comes from the fact that by adopting this perspective we can account for a greater number of cases of MEPP (cases of MEPP recognized as such in scientific practice) which otherwise would be excluded from our philosophical investigation. This constitutes, I think, a quite important step ahead in the study of MEPP.

Naturally, all the previous considerations do not exclude that a WTA model might be proposed (or that a WTA model is able to capture the very specific case of MEPP for which it was designed –something which, if true, would corroborate the need to adopt a pluralist principle). My modest claim in section 7.2 was that such a single encompassing model is not within the WTA models I have analyzed during this work. But if my testing above is correct, and if scientific practice suggests that specific forms of reasoning are a 'mark' of MEPP, how could the WTA partisan argue against the need of adopting a pluralist perspective? There are, I think, three possible ways to reject pluralism as the better attitude to adopt toward MEPP. The first strategy would consist in denying that such qualitative ingredients play an essential role in MEPP or that examples such as that of Hénon-Heiles system do represent genuine cases of MEPP. Nevertheless, this option is not the option our contemporary science seems to suggest us. If we want to take the work of scientists seriously, I think, we have to accept the test-case considered as a genuine case of MEPP and follow the scientists in looking at the form of reasoning which is involved in MEPP as a crucial explanatory ingredient. Naturally, a second move would be to propose a new encompassing WTA model, i.e. a model able to account for all the varieties of MEPP (among them the case of Hénon Heiles systems). However, at the best of my knowledge, we do not dispose of such a model. Finally, it might be thought that one of the WTA models considered up to now, or possibly all, can be refined in order to capture the test case proposed and reflect the intuitions of the scientists. But this obviously shifts the burden of the proof to the partisans of the WTA approach. However interesting and successful might be the latter two options, in the next chapter I will concentrate on a completely different strategy.

To sum up, I did not provide any *a priori* account of MEPP but I focused on a MEPP recognized as such in scientific practice. I evaluated the three WTA accounts seen in the first part on this test case and I showed that they have difficulties in accounting for the explanatory character of this MEPP. Moreover, I pointed out that my evaluation and independent considerations coming from the literature reveal a different picture of MEPP which is in need to be explored. In this picture specific forms of reasoning, and the abilities to reason which come with these forms of reasoning, are crucial to MEPP. In the next (and final) chapter I am going to turn my attention to this picture of MEPP.

Chapter 8

A new approach to MEPP in terms of intellectual tools and conceptual resources

In the previous chapter I showed that the three WTA accounts presented in the first part of my dissertation are not able to capture some cases which have been recognized as genuine in scientific practice. Furthermore, by focusing on qualitative factors which act in MEPP, I put forward the idea that a pluralist position provides a more promising way for the study of MEPP. The motivation for this claim lies in the following two remarks: a) the qualitative ingredients which are not captured by the WTA models analyzed can be identified with specific abilities to reason (for instance, the ability to reason visually in the Hénon-Heiles case) which are employed in MEPP and which are seen by scientists as essential to the explanation itself; b) there are other examples of MEPP in which different abilities to reason are recognized to play an essential role as well (for instance, the ability to reason asymptotically for the case of Batterman's asymptotic explanations).

In this last chapter, I will take into account these considerations and I will propose my original approach to MEPP. In particular, I am not going to propose a model of explanation (I will consider explanation as a 'practice', or better as the result of a practice in science), but a way to account for the explanatoriness scientists attribute to a specific MEPP in their scientific practice. In order to do that, I will introduce two categories: *intellectual tools* and *conceptual resources*. I will suggest the idea that, through these categories, we can capture the qualitative reinforcements which have been demanded by some authors to the WTA conception of explanation.

Even if I will come back to the notions of intellectual tools and conceptual resources, let me anticipate here a clarification on my notion of conceptual resources. In general, I will propose the idea that the intellectual tools are our abilities to reason when used in the practice of explaining, while the conceptual resources are the concepts which permit the use of our intellectual tools in a particular situation. I will consider that the conceptual resources are particular concepts which permit us to reconceptualize a state of affairs in a fruitful way. The fruitfulness of the reconceptualization comes from the possibility of applying a particular ability to reason. To take a very simple example, the concept of 'integral' might act as a conceptual resource because it allows us to redescribe a particular function (for instance, the function velocity $v(t)$) as an area (the area under the function $a(t)$). To see it as a conceptual resource is to consider that it makes possible such reconceptualization and the application of our abilities to reason (for instance, our ability to reason visually on a time-acceleration diagram).

I will show how my framework in terms of intellectual tools and conceptual resources is supposed to subsume under an umbrella the different species of MEPP which are assumed to exist according to the pluralist principle[1]. Moreover, this framework does provide a way to discriminate between different kinds of MEPP and it can be generalized.

The outline of this chapter will be the following. In the next section I will illustrate Henk De Regt and Dennis Dieks' recent attempt to capture the

[1]Observe that, in proposing such an approach, I will not come up with a new WTA model. My framework will be perfectly compatible with the pluralist hypothesis. I will come back to this point in the final part of the chapter, in section 8.4, where I will offer a possible way to generalize my approach.

notion of scientific understanding. A short examination of their perspective on scientific understanding is important because, in presenting my notion of intellectual tools, I will make a parallel with De Regt and Dieks' notion of conceptual tools. Soon after I will move to my own approach to MEPP. I will introduce the two basic notions (intellectual tools and conceptual resources) and I will illustrate, *in concreto*, how this framework can be used in the example of MEPP introduced in the previous chapter (the case of the Hénon-Heiles system). Next, I will show how this approach can be applied to other species of MEPP which have been considered in the previous chapters, thus suggesting a way to generalize it. Finally, I will point to the potential payoff of adopting my framework, to some questions which have not been sufficiently answered by it, and I will report my conclusions.

8.1 De Regt and Dieks on scientific understanding: conceptual tools

In their 2005 paper on scientific understanding [De Regt *et al.*, 2005], Henk De Regt and Dennis Dieks suggest to consider causality, visualization and unification as examples of "intelligibility standards" that vary through history and depend on the specific meso-level scientific context in which they are employed (where the meso-level scientific context is the context of scientific communities in a specific historical period). They propose the idea that intelligibility standards function as context-dependent "conceptual tools" for achieving understanding in science [De Regt *et al.*, 2005, p. 165]. More precisely, De Regt and Dieks' idea is that those tools permit the intelligibility of a scientific theory (in context C) by making possible the circumvention of a calculatory stage and the direct jump to a conclusion which concerns qualitative characteristic consequences of the theory itself. This is their criterion for the intelligibility of theories (CIT: criterion for the intelligibility of theories). Their criterion for the understanding of phenomena (CUP) is parasitic on CIT and is formulated in the following way: A phenomenon P

can be understood if a theory T of P exists that is intelligible (and meets the usual logical, methodological and empirical requirements).

In the unification account of explanation, understanding is obtained through unfying power [Kitcher, 1989, p. 432], while in the classical version of the causal account it is the knowledge of causal relations that is decisive for scientific understanding [Salmon, 1984a, p. 260]. In these theories of explanation the standards of intelligibility (causality and unifying power) have a universal status and the context does not influence them. By considering standards of intelligibility as *contextual-dependent* conceptual tools, De Regt and Dieks' approach to understanding is an attempt to reconcile these different views of explanatory understanding: scientific understanding is provided by scientific explanations of diverse types, but the conceptual tools which participate in every type of explanation are different (and their availability depends on the scientific community in which they are employed).

To illustrate how their criteria CIT and CUP work, they consider the way Ludwig Boltzmann presented the kinetic theory of gases in his 1964's *Lectures on Gas Theory* [Boltzmann, 1964]. In the introductory section Boltzmann considered that a gas can be pictured as a collection of freely moving molecules in a container. To picture the gas in this way allows us to obtain qualitative insight on the behaviour of gases. According to this representation, by identifying heat with molecular motion, it follows that to an increase of temperature there corresponds an increase in the average kinetic energy of molecules. Using the same picture of the gas, it is easy to consider that the collision of every gas molecule with the wall of the container results in a little push, and the sum of all the molecules-pushing produces the pressure. Hence "the picture immediately gives us a qualitative explanation of the fact that a gas exerts pressure on the walls of its container" [De Regt *et al.*, 2005, p. 152]. By using the same picture, again, we do not need to use calculations to see that if we decrease the volume the pressure will increase. This is because a decrease of volume means an increase in the number of molecules per unit of volume, i.e. an increase in the number of impacts per unit of time on the walls of the container, and then an increase in pressure. Finally, De Regt and Dieks write:

> In this way we obtain qualitative understanding of the relations between temperature, pressure and volume of a gas. If one adds heat to a gas in a container of constant volume, the average kinetic energy of the moving molecules –and thereby the temperature– will increase. The velocities of the molecules therefore increase and they will hit the walls of the container more often and with greater force. The pressure of the gas will increase. In a similar manner, we can infer that, if temperature remains constant, a decrease of volume results in an increase of pressure. Together these conclusions lead to a qualitative expression of Boyle's ideal gas law. It is important to note that the above reasoning does not involve any calculations. It is based on general characteristics of the theoretical description of the gas. Its purpose is to give us understanding of the phenomena, before we embark on detailed calculations. Such calculations are subsequently motivated, and given direction, through the understanding we already possess [De Regt *et al.*, 2005, p. 152-153]

According to De Regt and Dieks, the ability to develop qualitative insight into the consequences of the kinetic theory requires a conceptual tool, which for the present example is causal reasoning [Eigner, 2009, p. 277]. Gaseous phenomena, i.e. the behaviour of macroscopic properties of gases, can be understood in the kinetic theory because the conceptual tool of causal reasoning permits the intelligibility of the theory (causal reasoning makes possible to recognize qualitative consequences of the theory without performing exact calculations). We are able to use the conceptual tool of causal reasoning because we possess the skill of making causal inferences and Boltzmann's theory of gases has the virtue of providing a causal mechanical picture. The contextual dependence is given by the fact that a scientific theory T could be intelligible for a scientist who operates in context C (a scientific community, like that to which Boltzmann belonged) while it might be unintelligible for a scientist belonging to another context C'. Hence, again, conceptual contingent tools depend on the skills of the scientific community during a precise historical period or in a very specific methodological context[2].

This rapid presentation of De Regt and Dieks' view will be useful in the illustration of my notion of intellectual tools, and in my discussion of the linkage understanding-MEPP which I will address later in the chapter. Before closing the present section and moving to my own approach to explanation, let me add a clarification concerning De Regt and Dieks' conception of understanding.

As we have seen above, these authors consider that scientific understanding is obtained from an appropriate combination of "intelligibility-enhancing theoretical virtues" of a theory [De Regt *et al.*, 2005, p. 142], i.e. virtues of the theory which contribute to the intelligibility of that theory, and specific skills possessed by the scientists in a determinate scientific community. Nevertheless this means that, contrary to what the words "intelligibility" and "contextual" might suggest to the reader, their idea is that scientific understanding is neither subjective nor individual. They accept, of course, that there is a subjective ingredient which participates in scientific understanding. This ingredient is given by our capacity to intelligibly grasp a scientific theory. Nevertheless, for De Regt and Dieks this is not a sufficient condition for understanding. Theoretical virtues of theories, as the acquisition of some skills in a particular context, contribute to understanding as well and make it possible to fulfill their criteria CIT and CUP[3]. This point will be useful for

[2] In a recent paper, Steffen Ducheyne has proposed a test case to show that De Regt and Dieks' criterion for the intelligibility of theories CIT is too strong and must be modified or extended [Ducheyne, 2009]. He argues that, although endorsing different (and incompatible) standards of intelligibility, Newton and Huygens could understand their respective gravitational theories (both qualitatively and quantitatively). According to Ducheyne, "if De Regt and Dieks' criterion was adequate, there would be a deep incommensurability between Newton's and Huygens's gravitational theories: since Newton and Huygens did not share the relevant intelligibility-enhancing theoretical virtues, both could not understand one another's theory quantitatively. Moreover, since De Regt and Dieks furthermore assume that further quantitative understanding requires that a theory should be seen as intelligible in the first place, Newton and Huygens could not understand each others' theories quantitatively" [Ducheyne, 2009, p. 256]

[3] De Regt and Dieks point out this hybrid, not purely subjective and invididual, nature of scientific understanding, in a passage of their 2005 paper [De Regt *et al.*, 2005, p. 143].

what I am going to present in the next section, where I will draw attention to the difference that there is between my notion of intellectual tools and De Regt and Dieks' notion of conceptual tools.

8.2 Intellectual tools and conceptual resources

Even if we agree that pluralism is the principle which governs the study of MEPP, and that our methodology must necessarily go in the bottom-up direction, a fundamental problem has not been addressed yet. It seems that to accept that MEPP come in different types suggests the idea that we get lost in this jungle of possibilities. However, we would like to know if something more can be said about some common characteristics shared by the different species of MEPP which can be captured by adopting the pluralist approach. To put it in other words, we want to see if there exist a language (i.e. if there exist some categories) through which the different species of MEPP can be described and distinguished.

Recall now the moral from the testing of the three WTA accounts I proposed in the previous chapter and the demand for qualitative reinforcements to be used in models of MEPP expressed by various philosophers. Moreover, consider the following quote from Jamie Tappenden:

> The theoretical virtues that led to the choice of a mathematical framework (and that consequently inform the ideas of understanding and explanation that the framework induces) influence the explanation of physical events as well [Tappenden, 2005, p. 177]

What are these theoretical "virtues" which are responsible for the choice of a mathematical framework and which inform the idea of explanation that the framework induces? I will try to offer a possible answer in what follows, by proposing the idea that the virtues in question regard the capacity that some mathematical concepts have to make the explanation more 'conceptually accessible'. This conceptual access is permitted to us by the use of our abilities to reason. Those modes of reasoning are exactly the qualitative reinforcements that have been demanded to the WTA models.

The view I am going to propose is based on a paradigm very different from that which stands behind the majority of the models of MEPP which I have analyzed in the previous chapters. For the most part, in fact, these models were based on the idea that the feature which contributes to the explanatory power of a MEPP is an objective feature, i.e. a feature which does not depend upon the observer performing the explanation. According to this view, the task of a theory of explanation is to identify this particular feature (a particular quality of the mathematical formalism, a particular state of affairs, a fact, a relation which holds in the world or in mathematics), because it is this feature which 'explains'. For instance, in Steiner's model this particular feature was given by a property of a mathematical object or structure involved in the explanation ('characterizing property'). In Kitcher's model, the search for the best unifying systematization of our best beliefs is dependent on the fact that there exists such unification 'trend' in the world (and the notion of pattern is intended to capture this particular feature of the world). In Pincock's case, we have seen how the representational capacity of the graph was a source of explanatory power due to the ability of the graph to pick out structural relational features of bridge-system. In this case, the explanatory power was given by the fact that a mathematical structure has the quality to pick out some feature of the real world. This quality, together with the existence of the structural relational features of the actual system, is something independent from us. On the other hand, my approach shifts the attention to the particular forms of reasoning which are employed in MEPP. To put it in a very simple way, while these theories of MEPP have given importance to *what* explains, which is independent from us and which according to these models does characterize a genuine explanation, my approach is to consider that it is the *way* in which we identify a particular state of affairs that does contribute to the genuineness of a MEPP. This marks, I think, an essential difference with respect to some pictures of explanation that we have seen up to now. In my view, a genuine explanation does not result from the identification of a state of affairs or a property of the world or mathematics, but rather from the fact that we can look at that property or state of affairs in a specific way. Very roughly, I consider that the explanatory power is not given by any particular feature, but by how a feature or a phenomenon are explained. In this sense, my approach is similar to Batterman's and Van Fraassen's. In fact, Batterman considered that it is a particular kind of reasoning (asymptotic reasoning) which is essential for explaining and which characterizes a specific kind of explanation (asymptotic explanation). In Van Fraassen we have seen how an answer to a why-question (an explanation) gives "the sort of information the questioner has in mind", and therefore also in this case the *way* in which we explain is crucial. However, in the details, my approach considerably departs from Batterman's and Van Fraassen's. In particular, I am going to argue that when we can reconceptualize a particular state of affairs (through conceptual resources), and this reconceptualization does permit to use our abilities to reason, we do have a genuine explanation. Of course, this is not to say that explanation is purely subjective. Rather, I will argue that there is a subjective factor (ability to reason) which comes into play. But other ingredients are involved as well. As I will show in what follows, the reconceptualization permits us to apply an ability to reason. This reconceptualization is permitted only by particular mathematical concepts (conceptual resources). Therefore my approach is not subjective because these mathematical concepts have some virtues (they permit a reconceptualization), and this is something which is completely independent from us.

8.2.1 Intellectual tools

In order to account for a MEPP I introduce two categories: *intellectual tools* and *conceptual resources*. The intellectual tools are our abilities to reason when used in the practice of explaining. For instance, it is generally recognized that we acquire the ability to reason analogically[4]. This means that we are able to apply analogical reasoning in different situations, where analogical reasoning is the process of reasoning by analogy, i.e. to reason and learn about a new situation –the *target* analog– by relating it to a more familiar situation –the *source* analog– that can be viewed as structurally parallel [Holyoak *et al.*, 1997]. This particular kind of reasoning is used extensively in our everyday-life. It is used in linguistic, when we make a textual comparison between two words (or sets of words) and we highlight some form of semantic similarity between them; in speech, when we use metaphors (a metaphor is a figure of speech that constructs an analogy between two things or ideas); even in anatomy when we consider two anatomical structures to serve similar functions but are not evolutionarily related. It can also be used in morality: if it is wrong to do something in a situation A, and situation B is analogous to A in all relevant features, then it is also wrong to perform that action in situation B. Nevertheless, the ability to reason analogically is recognized to be extensively used in science as well. To take two very straighforward cases, consider the analogy between the atom and the solar system, which is given in order to stress the relative similarity of behaviour of the systems, or that between billiard balls and gas molecules[5]. To come to my view, I consider that an ability to reason (such as the ability to reason analogically) does function as an intellectual tool when it is used in the practice of explaining[6].

[4]Although it is not relevant here to focus on the process of acquisition, it is important to note that the ability to reason analogically is not considered as a primitive cognitive ability but it is seen as emerging and developing under the guide of certain basic constraints ([Gentner *et al.*, 2001], [Holyoak *et al.*, 1997]).

[5]By explicitly comparing the atom with a (tiny) solar system the learner is provided with a useful starting point for learning about atomic structure and the forces acting in an atom. Observe, however, that there are clearly limitations in the use of analogies in explaining scientific ideas. For instance, some aspects of the target (the atom) which we wish to emphasize or explain might not have a counterpart in the source system (the solar system), thus making inadequate the use of analogical reasoning [Taber, 2001].

[6]Concerning the case of analogical reasoning, observe that there is a number of philosophers who welcome the idea that the use of this kind of reasoning in science does contribute to scientific explanation ([Hesse, 1966], [Cartwright, 1983]). Nervetheless, in what follows I will concentrate on two other abilities to reason: the ability to reason visually and that to reason asymptotically, with a particular accent on the former.

Intellectual tools are not the unique ingredient in my approach to MEPP. To consider an ability to reason as an intellectual tool presupposes, in my view, that there are particular concepts which permit the use of that ability. The other important notion to be introduced, which concerns those particular concepts, is that of conceptual resources. Before passing to this, however, let me add an observation concerning my notion of intellectual tools and De Regt and Dieks' notion of conceptual tools.

In section 8.1 I have introduced De Regt and Dieks' account of scientific understanding, together with their notion of conceptual tools. Even if I borrow the general idea of intellectual tools from their account, and more precisely the idea that there exist some 'tools' which permit a qualitative insight in scientific practice, there is a fundamental difference between my notion and De Regt and Dieks'[7]. I totally agree with De Regt and Dieks on the importance of the particular skills of the scientists in the process of understanding and explaining a phenomenon, and on the fact that such skills can vary from context to context:

> [...] possessing a theory is not enough: in addition one should be able to use the theory to derive predictions or descriptions of the phenomenon. And this implies that not only knowledge of laws and theories (and background conditions) but also particular skills of the user of this knowledge are involved in achieving the epistemic aim of science. This introduces a pragmatic element that, we will argue, is part of an epistemically relevant notion of understanding [De Regt *et al.*, 2005, p. 142]

However, they consider that it is a property of some theories to have special pragmatic virtues such as visualizability (or simplicity), and these virtues

[7] Observe that De Regt and Dieks welcome the idea that the notion of conceptual tools also has a potential application in the realm of mathematical understanding [De Regt *et al.*, 2005, p. 163-164], although account does not explicitly refer to mathematical explanation and their examples of conceptual tools do not concern purely mathematical cases. Moreover, in a private conversation Henk De Regt has confirmed to me that he sees as a potential promising step the use of the notion of conceptual tool also in connection with mathematical explanation.

influence our epistemic access to the theories themselves (our intelligibility of them):

> Not only skills of scientists but also properties of theories play a role in this dimension: whether scientists are able to apply a theory to a particular phenomenon depends both on their skills and on the pragmatic virtues of the theory, e.g., visualizability or simplicity. These virtues may contribute to the intelligibility of the theory, thereby facilitating the use of the theory in the construction and application of models, and accordingly they contribute to the achievement of the epistemic aims of science. The appropriate combination of scientists' skills and intelligibility-enhancing theoretical virtues is a condition for scientific understanding [De Regt *et al.*, 2005, p. 142]

In other words, the fact that a theory is intelligible depends on contextual factors (capacities, background knowledge and background beliefs of the scientists in a particular context), but also from the fact that the particular theory under study has the property (or virtue) to be, for instance, visualizable. In De Regt and Dieks' example concerning the kinetic theory of gases, we are able to use the conceptual tool of causal reasoning because we do possess the skill of making causal inferences and Boltzmann's theory of gases has the virtue of providing a causal mechanical picture. My idea of intellectual tools diverges from their notion of conceptual tool on the latter point. I do not consider that a theory (or a mathematical framework under study) has virtues such as visualizability, or the virtue of providing a causal-mechanical picture. It is hard to see how a classification of theories according to their 'capacity to be visualizable' or 'capacity to provide a causal mechanical picture' might be obtained. Furthermore, it should be noted that physical phenomena are rarely explained or understood by appealing to a single theory. For instance, in Batterman's example discussed in section 6.2 there are at least two physical theories involved (the kinetic theory and the theory of condensed matter), plus a mathematical theory (renormalization group theory). What is *the* theory which contributes to the understanding of the phenomenon of universality at critical temperature?[8] Very differently, I consider that we do possess abilities to reason and these abilities are employed as tools (this is why I call them 'intellectual tools') in a particular situation and convey the sense of explanatoriness that we recognize in scientific practice. For instance, to take Boltzmann's example as an illustration of my point, I consider that the kinetic theory of gases does not have the virtue to give us a causal mechanical picture. Rather, *we* do have a particular ability to reason (the ability to reason causally) and this ability is used within the kinetic theory. This ability to reason is extra-theoretical, i.e. it is independent from the theory (or from the mathematical framework in question), and it belongs to what De Regt and Dieks call the "skills" of the scientists doing the explanation. In this sense, an ability to reason is subjective but not individual because the very same ability is used by the members of a scientific community[9].

This is not the end of the story. The possibility of applying such abilities to reason is permitted by the use of particular concepts, which I call "conceptual resources". These concepts have a particular virtue: they do permit a reconceptualization of a particular state of affairs[10]. Here I will concentrate on the case in which the conceptual resources come from mathematics.

[8]Let me note as an aside that my observation here points to a pending aspect of De Regt and Dieks' proposal. They consider virtues of theories. However, in order for their account of scientific understanding to be tested or even refined, they need a precise characterization of what they regard as a 'theory'. Moreover, they should discuss the possibility that two (or more) theories having a particular virtue could participate in the understanding of the same phenomenon.

[9]For instance, our ability to reason analogically is subjective because it depends on our intellect, but its use is shared within a scientific context and therefore this ability is not individual.

[10]I will also consider the possibility that conceptual resources permit a conceptualization, i.e. the making of a new concept from a previously known state of affairs. I will give an example of such a situation in section 8.4, when using my approach to account for Kitcher's example.

8.2.2 Conceptual resources

The conceptual resources are the concepts which permit the reconceptualization of a particular state of affairs and allow the use of our abilities to reason in a particular MEPP.

First of all, let me illustrate what I take to be the notion of reconceptualization[11]. We can describe the set of the real numbers R as the power set of the set of the natural numbers N (the power set of the naturals can be put in a one-to-one correspondence with the set of the real numbers). Thus the notion of 'power set' allows us to redescribe, or reconceptualize, the set of the reals as the power set of the naturals. The notion of 'vector' is another example of a concept which permits a reconceptualization. It provides an interpretation of any quantity X that has both a magnitude and direction as a point in a vector space. The result of the interpretation is a point in a vector space, of course, but this point can be thought as a redescription of the original quantity X.

Observe, however, that to recognize that the notion of vector (or that of power-set) permits a reconceptualization is not enough to give it the status of conceptual resource. A further requirement is that this reconceptualization must be fuitful, i.e. it must permit the use of an ability to reason. The moral is that the virtue possessed by a concept to permit a reconceptualization is not sufficient to consider that concept a conceptual resource. Perhaps, a quick and "explicative" analogy is to consider that the conceptual resources act as the ribosomes during the translation process in the synthesis of proteins (Figure 8.1). Ribosomes (conceptual resources) bind to the messenger RNA (our mathematical formalism), thus permitting the reading and the matching of the transfer-RNA (our ability to reason –intellectual tool) which will finally lead to the synthesis of the protein (which corresponds to our genuine MEPP). The virtue ribosomes have to bind to the messenger RNA

[11] Unfortunately, I can only provide here an intuitive connotation of it. However, I think that the idea is sufficient for what I claim below. I leave the task of offering a rigorous definition for a further study.

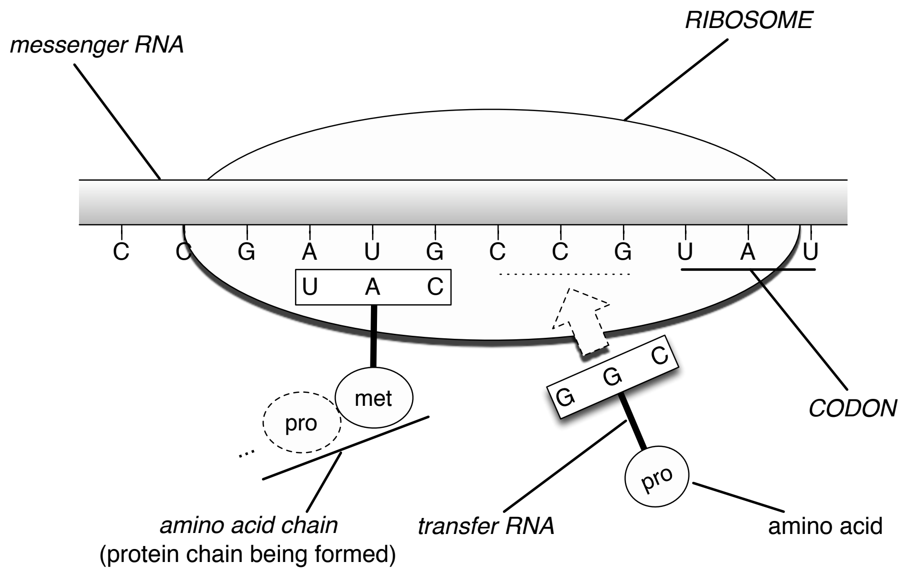

Figure 8.1: Diagram representing the translation process in the synthesis of proteins.

is not sufficient for the matching of the transfer-RNA and the final synthesis of the protein. In order for the synthesis of the protein to be successful, ribosomes must permit the reading and the matching of the transfer-RNA. In order for the explanation be 'successful', conceptual resources must permit the reconceptualization and the use of an ability to reason.

The analogy between my schema and the translation process in the synthesis of proteins can be pushed a little further[12]. Consider our abilities to reason as belonging to a continuous strand, which intuitively can be identified with our intellect (Figure 8.2). Every ability in the strand has a specific 'form', i.e. it is intrinsically distinct, as in the continuous strand of transfer-RNA there are three distinct base regions called "anticodons'[13].

[12]Evidently the translation process in the synthesis of proteins is much more complicated than the illustration I follow here. However, this simplification is advantageous to illustrate my ideas, and this is why I adopt it.

[13]An anticodon is a unit made up of three nucleotides. It corresponds to the three bases (three nucleotides) of the codon in the messenger-RNA. For instance, the codon for the amino acid lysine is AAA (where A stands for 'Adenine'); the anticodon of a lysine tRNA might be UUU (where U stands for 'Uracil'). In Figure 8.1, the codon for proline is CCG, and its complementary anticodon is GGC.

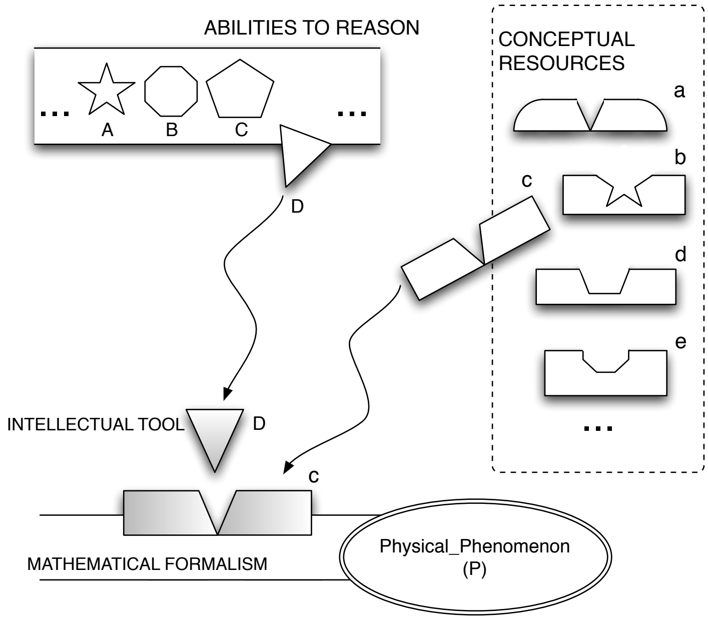

Figure 8.2: A possible representation of my framework.

Each anticodon on the transfer-RNA molecule can base pair to a corresponding three base (codon) region on the messenger-RNA (as showed in Figure 8.1). This particular three base region in the messenger-RNA strand does correspond, in my analogy, to a particular part of the mathematical formalism where the reconceptualization and the application of an ability is made possible. Finally, the binding of our abilities to reason to the mathematical formalism is permitted by the conceptual resources, which act as the ribosomes in the protein synthesis[14].

[14]It might be observed here that my analogy between conceptual resources and ribosomes can be thought as not fully adequate. This is because in my diagram (Figure 8.2) I represent conceptual resources as having a different form, thus suggesting the idea that they have a different structure. On the other hand, even if archaeal, eubacterial and eukaryotic ribosomes differ in their size and composition, ribosomes of a particular organism are considered to have a very similar structure and size. For example, eukaryotic ribosomes are between 25 and 30 nanometers and share an extremely similar structure. This is why, from this perspective, the analogy is not fully adequate. However, note that with this analogy I want to stress another point. The *function* of ribosomes in a particular cell-type (for instance, in eukaryotic cells) is the same, and namely that of permitting the matching of the transfer-RNA with the messenger-RNA and the consequent protein

In Figure 8.2 I represented the case of a particular ability to reason (triangle D) whose use is permitted by a conceptual resource (object c). The ability to reason D is therefore an intellectual tool. The attentive observer might have noted that the conceptual resources a and c permit the binding of the very same ability to reason (the triangle D), even if they have a different form. This state of affairs expresses the idea that the use of an ability to reason (for instance, the triangle D) can be permitted by different conceptual resources (a and c in the illustration). To put it in a more intuitive way, one conceptual resource can be available in a scientific context and can permit the use of an ability to reason, even if other conceptual resources might permit the use of the same ability as well (in the very same context or in another context, to explain the same phenomenon or different phenomena). In the same scientific context shared by James and Dan (for instance, James and Dan both work on dynamical systems and they agree on the mathematics they use in their research work), scientist James might prefer a to use D, while scientist Dan might opt for c to use the same ability D. In a situation when James and Dan do not share the same context (for instance, James lived in 17th century and Dan is a 20th century scientist), it is reasonable to think that they will adopt different conceptual resources, even if they will have at their disposal the same ability to reason D. If we think at our abilities to reason as a set X, and at the conceptual resources as a set Y, we can model this intuition by saying that the function $f : X \longrightarrow Y$ is *not* one-to-one (injective).

There is, I think, much more to say and to investigate about the possible combinations of our abilities to reason and conceptual resources in MEPP. Perhaps, when looking at Figure 8.2, it is very natural to ask if the following two cases are possible:

synthesis. Moreover, there is no structural but considerably functional homology between prokaryotic and eukaryotic ribosomes. This is something which is well mirrored in my analog, because every conceptual resource has exactly the same function: it permits a reconceptualization and the application of our abilities to reason. It is this functional aspect that I would like to emphasize in my analogy between ribosomes and conceptual resources.

(1) two or even more conceptual resources can be used together to permit the use of an ability to reason.

(2) one conceptual resource can permit the use of more than one ability to reason at the same time.

I think that both the cases are plausible. The first case can be well illustrated with a straightforward hypothetical example. Consider that mathematical concept A makes possible a reconceptualization $O_1 \xrightarrow{\text{A}} O_2$, where O_1 and O_2 are two distinct state of affairs, while another concept B makes possible a reconceptualization $O_2 \xrightarrow{\text{B}} O_3$. Therefore it will be possible to say that O_1 is redescribed by O_3 (reconceptualization, in this sense, can be considered as a transitive relation). If the final reconceptualization $O_1 \xrightarrow{\text{AB}} O_3$ permits to use an ability to reason, for instance the ability to reason visually, we will be exactly in the situation expressed by (1): two conceptual resources can be used together to permit the use of an ability to reason. Regarding the situation expressed by (2), things are a little bit more complicated. An illustration of this situation would require a hypothetical case of MEPP in which two (or more) abilities to reason are employed at the same time. Perhaps an interesting and very simple case would be that in which the ability to reason analogically is used in combination with the ability to reason visually. For instance, by considering that two systems are analogous we might trace a diagram and reason visually on that diagram, and these operations can be performed at the same time. Our ability to reason visually can help us to better grasp the analogy, or it can provide new insight into the similarities shared (or not) by the two systems. This is what happens, I think, in the present situation, where I am trying to 'explain' to the reader how my framework works by using my ability to reason analogically (through the analogy between the translation process in the synthesis of proteins and the functioning of my framework) and my ability to reason visually (by using shapes and trying to connect the elements on the diagram 8.2 in the right way). I am using both these abilities, at the same time. Undoubtedly, in this case mathematics plays no role and it would be meaningless to speak of mathematical explanation. Furthermore, it might be thought that in this situation the ability to reason visually is not a distinct ability but it is embedded in the ability to reason analogically. This is because the analog components are visualizable on a diagram. Although at this stage I do not have any argument against this possibility, I think that the ability to reason analogically and that to reason visually should be considered as distinct abilities. And this despite the fact that some analogies make the analogs visualizable on a diagram or a schema. For instance, in their paper "Analogy Theory for a Systems Approach to Physical and Technical Systems" [Hezemans *et al.*, 1991], Hezemans and van Geffen specifically focus on the conditions which permit the use of analogies in science and write: "an analogy can be critically applied, such that it can be seen visually" [Hezemans *et al.*, 1991, p. 170]. This may suggest that the visualization of such analogy, and more importantly the reasoning on the analogy itself, does require a further step, i.e. the application of a distinct ability to reason.

Let me provide a more concrete example of situation (2). This example requires, in addition to the use of the two abilities, the fact that the use of this combination of abilities has been made possible by a particular mathematical concept (a concept that has permitted the reconceptualization of a specific state of affairs and the application of the two abilities). Consider, for instance, the following case. In his *The Science of Mechanics* [Mach, 1893], in order to illustrate the virtual work principle and explain the difference between stable and unstable equilibria of a mechanical system, Ernst Mach proposes to regard a system of bodies acted by various forces as a machine for doing work. A hanging weight is attached to the system and this weight carries a pencil pressing against a sheet of paper, carried past it horizontally. When the system is allowed to move, "the depth of the hanging weight below its original position will be an indicator of the work done by the system in reaching any other configuration. The pencil will record this depth in a curve, as in Figure" [Mach, 1893, p. 64]. I have reported Mach's diagram in Figure 8.3. The curve expresses the evolution of the system, and its maxima

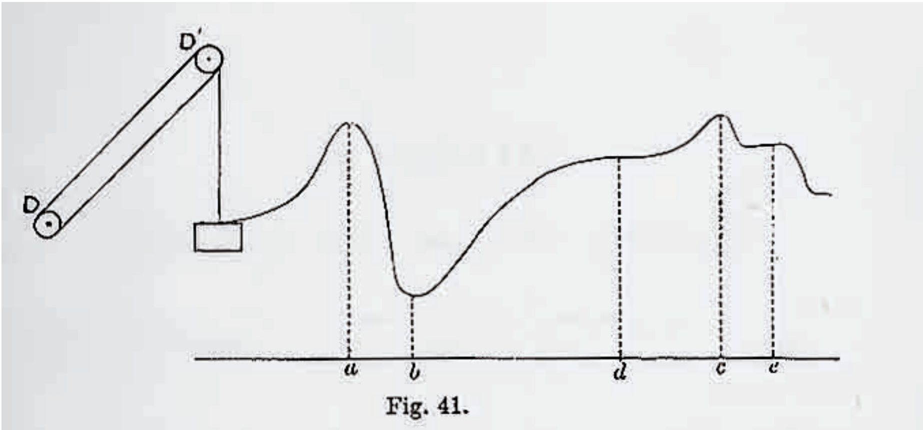

Figure 8.3: Mach's original diagram [Mach, 1893, p. 65].

and minima are, of course, the equilibrium points of the system[15]. When the system arrives at a position of equilibrium, the work done is in general a maximum or a minimum. The weight is at a turning point of the curve. The maxima are unstable because, if the weight is disturbed (its position is slightly changed), it will descend due to the effect of gravity. This point corresponds to the minimum amount of work that can be done by the system. On the other hand, the minima are stable because, when its position is slightly changed, the weight will return to its position because of the force of gravity. In this case, the system can only do work by returning to the position of equilibrium if disturbed. Stable equilibrium therefore corresponds to a maximum of work done by the system, unstable equilibrium to a minimum. Finally, when the curve remains horizontal (as at d, e), the system is in a state of "neutral" equilibrium, "as when a sphere rests on a horizontal plane" [Mach, 1893, p. 65].

In this case we are using two abilities to reason, namely that to reason analogically (by using the analogy of the falling weight) and that to reason visually (by visualizing the behavior of the system through the weight-curve diagram). Furthermore, it is the concept of the global extrema of a function which permits us to reconceptualize the problem and apply these two abili-

[15]See [Panza, 1995] for a discussion of Mach's example.

ties[16].

I assume that there is a mutual interaction between intellectual tools and conceptual resources: we acquire conceptual resources through the use of intellectual tools and vice versa (for instance, through education and practice in science). Intellectual tools, like conceptual resources, are therefore not unchangeable but can vary over time (I will come back to this point below, in subsection 8.6.3).

After this short presentation of the notions, it is now time to pass to a more concrete setting. In the next section I will illustrate how this framework functions by applying it to the example of MEPP considered in the previous chapter.

8.3 Intellectual tools and conceptual resources at work

Consider again the case of the Hénon-Heiles system, where, when faced with two mathematical paths to study the same phenomenon (the behaviour of the particle in the considered bidimensional potential), scientists do regard as explanatory the formalism involving the Poincaré map. Recall that, studying on the diagram the Poincaré section for various energies, we obtain information about the dynamic of the system at that energy (Figure 7.2). The scientist's ability to reason visually on the diagram is considered then as essential to explain why the system has that particular behavior at the fixed energy. However, as we have seen, the possibility of performing such a qualitative analysis is permitted by the use of the Poincaré map, a particular function which "dots" the solutions (orbits) on the respective section. On the view I am proposing, the intellectual tool which is used is our ability to reason visually, and its use is made possible through the particular function

[16]Naturally, there is another situation which can be imagined as well: various conceptual resources which permit the application of more than one ability to reason. Observe, however, that this case is somewhat analogous to case (1) and similar considerations hold for it.

Poincaré map, which acts as a conceptual resource.

In the case of the Hénon-Heiles system we fix the Energy to reduce the dimensionality of the space by one, and then we choose the $q_y p_y$ plane ($q_x = 0$) as 2-dimensional cross section (S) of the hypersurface in the phase space. By doing that, we obtain a discrete dynamical system of the continuous 3-dimensional Hamiltonian flow (the flow of trajectories in phase-space), with a state space that is one dimension smaller than the original dynamical system. This reduced system inherits many properties, e.g. periodicity or aperiodicity of the original system, and can be interpreted as a discrete function $p : S \longrightarrow S$ which associates consecutive intersections of a trajectory of the 3-dimensional flow with the surface of section S. Function p is exactly the Poincaré map.

Now, there is a fundamental difference in introducing the Poincaré map in our reading of the formalism. As Aubin and Dalmedico point out:

> This method [the use of Poincaré map] made it natural to think about the states of the system considered as points in phase space [Aubin *et al.*, 2002, p. 286].

As we have seen in the previous chapter, when I illustrated the Hénon-Heiles example, the consecutive disposition of the points trace a (regular or scattered) sequence on the surface of section (Figure 7.2). By identifying the points with the states of the system, we associate a particular disposition of points with a particular behaviour of the system. To reconceptualize the states of the system as points is therefore essential to grasp, by observing the successive disposition of the points on the surface of section, qualitative conclusions about the behaviour of the system. For instance, this convinces us that a point which is consecutively mapped into points very close to each other will represent a dynamical state which gives rise to regular motion. Furthermore, in grasping such qualitative information from the diagram we are reasoning in a very particular way. More precisely, we are using our ability to reason visually. The Poincaré map is thus the conceptual resource which permits the use of our ability to reason visually on a diagram (at some particular step in the reasoning). Without the introduction of the Poincaré map it would have been impossible to make use of this particular ability to reason.

Obviously, before the step which concerns the qualitative (visual) analysis of Poincaré map in the diagram, we *do* exact calculations through the Hamiltonian formalism. However, I think that this should not be considered as a real difficulty with respect to what I said above. Before arriving to the step which involves the visualizable situation, some calculations are necessary. What is really central to my discussion is that at a particular stage in the mathematical procedure we grasp the behavior of the phenomenon by visualizing its dynamic through the Poincaré map, and the role of mathematics is essential in performing such a move.

This evaluation reflects the intuitions of the scientists who consider one mathematical path (that involving the phase-space and the Poincaré map) as more powerful and explanatory than another. To choose this route as the explanatory one, again, derives from the fact that a conceptual resource permits the application of an ability to reason. The Poincaré map has the virtue to permit a reconceptualization, and such a reconceptualization is what permits us to reason visually and grasp the behaviour of the system. Of course, other mathematical concepts might have permitted such a reconceptualization as well. And the use of such theoretical resources clearly depends upon our scientific education and the scientific context we are working in.

In the next section I am going to propose the idea that my language of intellectual tools and conceptual resources can be used to describe the cases of MEPP proposed by Batterman, Pincock, Kitcher and Steiner. Moreover, my approach will offer a way to differentiate them and it will support the idea that they do represent genuine examples of MEPP, thus preserving the general intuitions of the authors who investigated these cases through their respective accounts.

8.4 Generalization

A detailed application of my framework to the cases discussed by Batterman, Pincock, Kitcher and Steiner would require an in-depth analysis which cannot be proposed here. However, let me indicate very shortly how my schema is supposed to account for such MEPP.

8.4.1 Batterman

As I said in the section 8.2.1, our ability to reason visually in not the only intellectual tool we use in our explanatory practices and others abilities to reason are at our disposal as well. An example is provided by the ability to reason which acts in Batterman's example of asymptotic explanation (the explanation offered in condensed matter physics for the universality of critical phenomena).

Asymptotic reasoning is the kind of reasoning which is generated from asymptotic techniques. Asymptotic techniques permits the elimination of causal and physical details which are not essential to the explanation of the phenomenon, thus highlighting relevant factors for the phenomenon which is explained. Observe, however, that Batterman considers that the throwing away of the causal and physical details which are not relevant to the physical phenomenon under study is given in these techniques by the passage to a limit (a limiting operation). For instance, in the example of the renormalization group explanation of the universality of critical phenomena, it is in the thermodynamic limit (limit in which the number of particle of the system approaches infinity) that the fixed point of the recursion relation converges to the exact critical temperature. As I showed in chapter 6, the main idea in the renormalization group procedure investigated by Batterman is to switch from an intractable problem (analytically intractable Hamiltonian) to a tractable problem (analytically tractable Hamiltonian), preserving the functional form of the initial Hamiltonian. To assure that the transformed Hamiltonian describes a system with the same behavior of the original systems, thermodynamic parameters are properly adjusted (renormalized). More precisely, the strategy of the RG analysis is based on a systematic rescaling of the effective Hamiltonian which describes the system near the critical point. In this context, the taking of the thermodynamical limit is essential to use the renormalization group and therefore to eliminate the particular causal and physical details which are not essential to understand the phenomenon of universality. As batterman writes:

> Limits are a means by which various details can be thrown away (For instance, in taking the thermodynamic limit in the context of explaining fluid behavior, we eliminate the need to keep track of individual molecules and we remove details about the boundaries of the container in which the fluid finds itself.) [Batterman, 2010, p. 20]

> In the thermodynamic limit, as it is employed in the explanation of the universal scaling of the order parameter for critical systems, there emerges a host of divergences and singularities. Crucial among these is the divergence of the so-called correlation length. This implies a loss of any sort of characteristic length scale and allows the various systems to be compared with one another asymptotically. Such a loss of scale is required to demonstrate the genuine qualitative change in the states of matter that occur at criticality [Batterman, 2010, p. 18]

What characterizes asymptotic reasoning, therefore, is not the simple thought process of eliminating causal and physical details of the system. Otherwise, asymptotic reasoning would correspond to the Aristotelian procedure of abstraction, through which we 'remove' specific properties from an object to study a particular phenomenon. Rather, the thought process which characterizes asymptotic reasoning consists in the study of the nature of an asymptotic regime by considering only these features which have been identified as relevant by asymptotic methods. This marks an essential difference with the Aristotelian procedure of abstraction. Asymptotic explanation is an explanation which is essentially characterized by this kind of reasoning. But how can my perspective on MEPP accommodate such type of (asymptotic) explanation?

Where we appeal to the thermodynamical limit we use the concept of limit to redescribe a finite system of n particles as a system of infinite particles. Observe, then, that the concept of limit permits us to reconceptualize the system in a different way (as made of infinite particles). The mathematical concept of limit acts, according to my terminology, as a conceptual resource. Now, the identification of the physical quantities which does not affect the particular phenomenology in the limiting regime is not given by asymptotic techniques but requires something more. What the asymptotic method (through the renormalization group procedure) tells us is that at critical temperature there are some features which are irrelevant to the estimation of the critical exponents. Nevertheless, it does not suggest what are the structural features of the system which are relevant or irrelevant to the asymptotic behavior. For instance, in his paper "Infinite Systems in SM Explanations: Thermodynamic Limit, Renormalization (Semi-) Groups, and Irreversibility" [Chuang, 2001], Liu Chuang observes:

> If one is looking for a causal explanation for the critical phenomena, i.e., why we have the same asymptotic behavior (cf. Balashov 1997) near criticalities, the renormalization group does not deliver it. It does not tell us which compositional or structural features are relevant or irrelevant for the asymptotic behavior. What it does tell us is, at a critical point, which features are relevant or irrelevant to the estimation of the critical exponents. We should note that the space of Hamiltonians is not a configuration space, and the flows on Hamiltonian space are not physical processes. Therefore, the relevancy of the flows to a fixed point determined by the renormalization group should not be confused with the relevancy to the asymptotic processes of systems in the same universal class approaching their respective critical points. The fact that the latter may provide an explanation for the former shows why the former alone is not yet an explanation for the universal asymptoticity of critical phenomena. [Chuang, 2001, p. S337]

To pass from the features which are irrelevant to the estimation of the critical exponents to the structural features which are irrelevant for the asymptotic behavior requires a further step. To put it roughly, the renormalization group alone does not suggest what actual features of the physical system do not affect the value of critical exponent. For instance, the fact that correlation lengths diverge 'on the road to' criticalities does not say anything about the actual system, but it is informative about how microscopic variables at different positions are correlated. However, we associate this divergence with the fact that certain kinds of micro-degrees of freedom become asymptotically frozen and irrelevant to the behaviour of the real system, i.e. that the microscopic composition of the system is irrelevant to the behaviour of the system at criticality. To perform such a step is to employ a specific ability to reason, which I call 'ability to reason asymptotically'. This specific ability is the ability to recognize, in an asymptotic/limiting regime, what are the physical features which are not essential to the description of the phenomenon. This is crucial to explanations such as that of the universality of critical phenomena. It is this ability to reason, in fact, which permits us to look at the particular phenomenology in the limiting regime as representative of a class of physical systems. For instance, by recognizing that, in the limiting regime, the microscopic composition of the fluids are not essential to obtain a particular value for the critical exponent, we can consider that the particular value of that critical exponent is shared by other fluids as well. Note, however, that it is only through the reconceptualization permitted by the concept of limit that it is possible to use this particular ability to reason. The reconceptualization of the system (to see the system as made by an infinite number of particles) is essential to have access, through the renormalization group, to the limiting regime, where this ability to reason is applied. Finally, in line with Batterman, I welcome the idea that this should be considered as a particular species of MEPP (asymptotic explanation), and precisely as an explanation which uses the ability to reason asymptotically as an essential ingredient.

8.4.2 Pincock

Recall now Pincock's account of abstract explanations, presented in chapter 5. For Pincock an abstract explanation is an explanation that appeals primarily to the formal relational features of a physical system. As an illustration of such species of explanation, Pincock focuses on the explanation (in terms of graph theory) of the inability to make a particular path across the Königsberg bridges. In that case, the impossibility of walking across all of the bridges exactly once and returning to one's starting point is explained by appealling to the fact that the system bridges-paths exhibits the structure of a non-Eulerian graph[17]. In his words:

> In the Königsberg bridges case, the explanatory power is tied to the simple way in which the model abstracts from the irrelevant details of the target system. It throws out what is irrelevant and highlights what is relevant. Crucially, what is relevant is the mathematical structure found in the target system itself. [Pincock, 2011a, p. 3]

Like Batterman's asymptotic explanations, Pincock's abstract explanations gain their explanatory power by the systematic throwing away of various causal and physical details of the physical system under study. Moreover, according to Pincock, abstract explanations proceed by focusing on an abstract structure realized by such a system. The Königsberg bridges example shows that abstract explanations are not asymptotic explanations. This is because the kind of asymptotic explanations studied by Batterman involves mathematical equations that result from taking one or more quantities in a fundamental mathematical law to a limit (such as 0 or infinity)[18], but in the example of the Königsberg bridges there is no reference to any limiting

[17] As we have seen in chapter 5, there is a theorem in graph theory which says that a connected graph is Eulerian if and only if every vertex has an even degree [Wilson, 1996, p. 32].

[18] This was the case of the transformation performed through the renormalization group procedure in Batterman's example. In that case the renormalization group invokes the so called thermodynamic limit and the singularity, i.e. the fixed point (in the space of Hamiltonians) corresponding to the behavior of the system at critical point, is what emerges in this limiting operation.

operation in this sense.

To come to my approach, I consider that the ability to reason which is employed in Pincock's example of the Königsberg bridges is different from the ability to reason asymptotically which we have seen in Batterman's example of asymptotic explanation. But what about this ability to reason? And what about the conceptual resources which do permit its application?

In the case of the bridges, the mathematical conceptual resources of graph theory (vertex of a graph, edge of a graph) permit us a reconceptualization of some relevant structural relational features of the actual system (the bridge-system). For instance, bridges are seen as edges, or islands and banks as vertices. Of course, the mathematical entities coming out from this reconceptualization are elements of a graph. Otherwise, such operation would be useless[19]. Now, when this reconceptualization is put in practice, it is possible to obtain (for instance through a theorem) new knowledge about our graph. For example, we obtain the property a connected graph with all vertices of even degree has to be Eulerian. From this new (mathematical) property we will go back to our actual system (the bridge-system), and this step is made by identifying some particular entities of the graph (the entities which have been reconceptualized), together with their relational features, with some actual aspect of the physical system (together with its relational properties). This is to say that, at this step, we are using our ability to reason analogically: we are learning about a new situation, in this case a structural property of the actual system, by relating it to a more familiar situation, in this case a property of the graph (surprisingly, the graph is more familiar than the actual system). Observe, however, that to see the actual system and the graph as structurally parallel is permitted by the previous reconceptualization, i.e. the conceptual resources of graph theory allow to apply our ability to reason analogically. Furthermore, there is another ability to reason which is permit-

[19]The same holds, I think, when we reconceptualize a pencil as a line belonging to a plane. This line should be seen as an element of the euclidean plane, and after the reconceptualization it will be subject to the rules of euclidean geometry. In the same way the lines and the dots in the case of the Königsberg bridges are seen as edges and vertices of a graph, and therefore they are subject to the rules of graph theory.

ted by the reconceptualization and which is used as well in Pincock example.

In their paper "Graph theory representations of engineering systems and their embedded knowledge" [Shai et al., 1999], Shai and Preiss consider the usefulness of a mathematical representation which is isomorphic to the elements of an engineering system. In particular they focus on the benefits which derive from the use of graph theory in representing an engineering system. The study of graphs provides interesting insights concerning the knowledge of the engineering systems. This knowledge is called by them "embedded knowledge":

> The properties of the mathematical elements of those graphs and the relations between them are then equivalent to knowledge about the engineering system, and are hence termed "embedded knowledge". The use of this embedded knowledge is illustrated by several examples: a structural truss, a gear wheel system, a mass-spring-dashpot system and a mechanism. Using various graph representations and the theorems and algorithms embedded within them, provides a fruitful source of representations which can form a basis upon which to extend formal theories of reformulation. [Shai et al., 1999, p. 273]

Moreover, they remark how the embedded knowledge that results from the use of a graph-representation is conveyed to the students by the "ability" to "switch seamlessly from one representation to its analogy":

> Other results of this project have included successful use of these reasoning methods in high school classes, where students have assimilated the experience of using several representations to solve, or reason about, an engineering system. In the last decade over 300 high school students have successfully attained a much better-than-usual grasp of both mathematics and physics by using a variety of representations. We postulate that this success is partially owing to the use of multiple approaches and the ability to switch seamlessly from one representation to its analogy. [Shai et al., 1999, p. 284]

It seems then that this ability should be regarded as a distinct ability to reason which participates, together with the ability to reason analogically, within Pincock's abstract explanation of the impossibility of making an Euler tour across the seven bridges of Königsberg. Once the reconceptualization is put in place, we can use our ability to reason analogically and our ability to switch from the graph structure to the actual structure (and the other way around). It is by using these abilities that we are convinced that it is impossible to perform such a walk across the bridges, i.e. that the relational features of the bridge-system do make impossible such a route.

It might be observed that in Pincock's case I consider that mathematical conceptual resources can permit a reconceptualization of a physical object as a mathematical object, while up to now I have considered reconceptualization of mathematical objects as mathematical objects alone. Reconceptualization in this sense, from a physical object to a mathematical object via mathematics, is something which I consider perfectly admissible within my framework. For instance, I will consider the same sort of reconceptualization in section 8.5, when focusing on the asymmetry problem. Moreover, as I suggest below, this sort of reconceptualization (to interpret a physical object or an actual state of affairs as a mathematical entity) might be seen as associated with the possibility of having a mapping between the actual world and a mathematical structure, i.e. reconceptualization in this sense is associated with representation.

My discussion of Pincock's example in terms of conceptual resources and intellectual tools manifests a further aspect of my account. To say that it is possible to reconceptualize a piece of the world through the mathematics of graph theory amounts to saying, in Pincock's structuralist terms, that it is possible to identify a mapping between the world and some abstract mathematical model[20]. This is something which I welcome and which is per-

[20]To be more precise, I suppose that Pincock would say that it does exist a mapping from an actual target system to the mathematical structure.

fectly compatible with my idea of conceptual resources[21]. Observe, however, that reconceptualization alone is not sufficient, according to my framework, for genuine explanation. A genuine MEPP results from the fact that the conceptual resources make possible a *fruitful* reconceptualization, i.e. the conceptual resources must permit to apply a particular ability to reason (intellectual tool). Therefore, according to my framework, representation alone (associated with the reconceptualization as in the example proposed by Pincock) is not enough for a genuine explanation. Furthermore, conceptual resources are not always linked to this representational aspect (as the conceptual resource acting in Batterman's example shows). Finally, behind my framework lies the idea that representation is neither necessary nor sufficient for explanation. And this accords with what we have seen in section 5.4, where from the discussion it emerged that representation is not necessary for explanation.

8.4.3 Kitcher

Let me pass to consider Kitcher's example of unification through the Newtonian pattern. In the previous chapter I showed how Kitcher's unification model, in its original form, is not able to account for the MEPP concerning the behaviour of the Hénon-Heiles system. The weak point of the account stood in the inadequacy of Kitcher's argument pattern to capture a particular inferential step inside its structure. The inferential step in question concerned the grasping of the particular behaviour of the system obtained though our ability to reason visually on a diagram.

What I want to suggest in this subsection is that Kitcher's example of the Newtonian case can be rewritten in terms of my categories, while preserving the idea that the Newtonian pattern does provide unification. Furthermore,

[21]Perhaps a further step in the investigation of the relation between mapping and conceptual resources would be to explore the possibility that conceptual resources do permit such a mapping, i.e. they permit the passage from the actual situation to the graph. For my present focus is elsewhere, I will not address this issue and I leave this question to the structuralist explorer.

the idea suggested by Kitcher, i.e. that different phenomena which do obey Newton's laws can be *explained* by using the very same pattern, is preserved in my analysis too.

The Newtonian pattern is used to derive sentences which do represent different phenomena. In order to use my framework of conceptual resources and intellectual tools, and preserve the idea that through the Newtonian pattern we do have genuine MEPP, it must be shown that to the use of this pattern there corresponds the use of a particular ability to reason through some conceptual resources.

As we have seen in chapter 3, in the Newtonian pattern $\langle s, f, c\rangle$ the set of filling instructions f_N contains the directions for replacing the dummy letters α, β, γ, δ, θ in every schematic sentence. For instance, consider the schematic sentence "The force on α is β". As stated by the two filling instructions contained in f_N, the letters α and β are to be replaced, respectively, by an expression referring to the body under investigation and by an algebraic expression referring to a function of the variable coordinates and of time. The same operation of replacement can be made with the dummy letters which appear in the remaining schematic sentences. Furthermore, the classification set c_N for the schematic argument s_N gives us the inferential information about the schematic argument (how we move, in the schematic argument, from the premises to the conclusion). Finally, this inferential schema is used to derive sentences which represent different phenomena. This derivation, once made according to the classification set c_N, has the form of a simple derivation performed in first order logic[22]. Here is a simplified schema of the derivation:

(1) P_1

(2) P_2

(3) P_3

[22]The classification set indicates that: P_1, P_1 and P_3 are premises; that P_4 is obtained by P_1-P_3 by substituting identicals; that P_5 is deduced from P_4

(4) P_4 (from P_1-P_3 by substituting identicals)

---

(5) C (from P_4)

The conceptual resources of Newtonian mechanics (force, acceleration, trajectory, ...) permit us to describe some actual state of affairs as configurations of magnitudes which are related among them through differential equations. Next, the resources of first order logic (with schematic letters) permit to fit these informations (the relations among the various differential equations) into an inferential schema[23]. On this schema the scientist reasons deductively, in terms of a first order predicate language and in terms of the non-logical instructions given by the classification set. In fact, the deduction of C is not purely logical. For instance, C follows from P_4 by using algebraic manipulations and the techniques of the calculus (as stated by one member of the classification set), but this information does not appear in the inferential schema above (the schematic argument). As Kitcher points out:

> Whereas logicians are concerned to display all the schematic premises which are employed and to specify exactly which rules of inference are used, our example allows for the use of premises (mathematical assumptions) which do not occur as terms of the schematic argument and it does not give a complete description of the way in which the route from (4) to (5) is to go. Moreover, our pattern does not replace all nonlogical expressions by dummy letters. Because some nonlogical expressions remain, the pattern imposes special demands on arguments which instantiate it. In a different way, restrictions are set by the instructions for replacing dummy letters. [Kitcher, 1981, p. 517-518]

Finally, through the pattern the scientist is able to derive a large number of accepted statements (accepted by the scientific community to which

[23]As Kitcher observed, the tools of logics are used to isolate the notion of pattern and costruct this inferential schema: "the logician's approach can help us to isolate the notion of argument pattern which we require" [Kitcher, 1981, p. 516].

the scientist doing the explanation does belong). For instance, two statements concerning the behaviour of two physical phenomena are derived from arguments that instantiate the common argument pattern. This is how phenomena which do obey Newton's laws can be *explained* by using the very same pattern.

Now, let me switch to my approach to explanation. Differently from Kitcher, I would propose the following reading of what it means to explain a phenomenon through the Newtonian pattern: to explain a phenomenon through a Newtonian pattern is to be able to reason deductively in terms of this pattern. Let me elucidate this intuition with an example. Suppose that we want to explain why a projectile fired horizontally covers a certain distance x after a certain time t (for some initial conditions). The statement 'the projectile fired horizontally covers a certain distance x after a certain time t' belongs to the set of our accepted beliefs, i.e. it belongs to what Kitcher called the set of accepted sentences K. In order to explain why the projectile fired horizontally covers a certain distance x after a certain time t, the first step is to use the conceptual resources coming from the Newtonian theory (concepts of force, acceleration, trajectory) to describe the actual state of affairs as configurations of magnitudes which are related through differential equations. Once this conceptualization is made, we use the Newtonian pattern as to derive the statement 'the projectile fired horizontally covers a certain distance x after a certain time t'. The crucial point is that the use of the Newtonian pattern presupposes the ability to reason in terms of this pattern, i.e. the ability to follow the schematic argument by taking into account the sets of filling instructions and the classification set, plus the particular initial conditions. To rephrase this in terms of my approach: to explain that 'the projectile fired horizontally covers a certain distance x after a certain time t' we use the ability to reason (deductively) in terms of the Newtonian pattern. The use of this ability is permitted by the conceptual resources of Newtonian mechanics (something which sounds very natural!). Of course, when we are confronted with a different phenomenon, there is also the possibility that these conceptual resources do not permit such a conceptualization and the application of this ability. For instance, in the case of a chemical reaction, if we want to explain why we obtain the product C from two particular reagents A and B, the Newtonian conceptual resources will not permit any fruitfull conceptualization (and therefore we will not able to use the ability to reason in terms of the Newtonian pattern). On the other hand, although this particular ability to reason is used for a class of phenomena which can be studied and conceptualized through the Newtonian theory, it is possible that a physical phenomenon be explained by using the ability to reason in terms of a different pattern having the structure of the Newtonian pattern. This can happen, for instance, when some theoretical concepts of a theory do permit a conceptualization and on this conceptualization it is possible to reason deductively in terms of a certain pattern (similar in structure to the Newtonian pattern but with different sets s ,f, c)[24].

Therefore, some conceptual resources coming from Newtonian theory permits to apply the ability to reason in term of the Newtonian pattern. Is this circular in some sense? I think it is not. The general idea which stands behind the previous lines is that the Newtonian pattern is a tool that we acquire. This instrument indicates a particular scheme of inference, and to use this instrument amounts to being able to reason accordingly to the deductive schema which is stated by it[25]. Of course, this pattern is found before we use it to explain a phenomenon. And the ability to reason in terms of this pattern is something that we learn in our scientific education and practice. However, the ability to reason in term of this pattern must be keep separated from the conceptualization which is permitted by the Newtonian concepts. We might be able to find some conceptualization of an actual state of affairs by means of Newtonian concepts, but we might not be able to reason in terms of the Newtonian pattern. This is what happen, I think, in the following situation:

[24]For instance, Kitcher shows how such a pattern can be found in the Darwinian theory of evolution [Kitcher, 1981].

[25]Let me add that to be able to reason deductively in terms of the pattern requires the ability to reason schematically in terms of a predicate language.

a student is asked to explain some phenomenon; he conceptualizes the situation by using the Newtonian concepts; he is not able to reason deductively in terms of the pattern and then he is not able to *see* how the explanandum is derived according to the pattern. In this case, he will not be able to explain the phenomenon.

Observe that, in Kitcher's example, conceptual resources do not have the function of permitting a *re*conceptualization from A to B, namely the redescription of A as B where A and B are known states of affairs. Every time that we use such Newtonian conceptual resources we conceive *ex novo* an actual state of affairs as configurations of magnitudes which are related among themselves through differential equations. This is different from the case of the Hénon Heiles system. In that case the conceptual resource Poincaré map did permit the passage from A to B, where A and B were known before the reconceptualization, but here the conceptual resources of Newtonian mechanics permit a "conceptualization" (they permit us to conceive B *ex novo*). This is why I used the term 'conceptualization' instead of 'reconceptualization'. Therefore, together with reconceptualization, here is a second function of conceptual resources: they permit the conceptualization of a particular state of affairs. This conceptualization must allow us to apply an ability to reason, which in this case is our ability to reason in terms of the Newtonian pattern.

Finally observe that, although my approach does not consider that it is the unification function of the Newtonian pattern which confers an explanatory character to the deduction, the idea that the pattern does provide unification (in Kitcher's sense) is left untouched. On the other hand, according to my approach, the unification function of the pattern is related to explanation only in the following sense: the pattern 'defines' (broadly speaking) an ability to reason, and to the use of this ability to reason there corresponds a particular form of explanation which is exploited for a class of phenomena (the Newtonian phenomena).

8.4.4 Steiner

Let's come back to the existence of an instantaneous axis of rotation in the kinematics of rigid body motion. This result, as obtained through a particular theorem (and the relative proof), has been taken by Mark Steiner as an example of genuine MEPP[26].

As we have seen in our discussion of Steiner's model, the mathematical theorem that states the existence of such an instantaneous axis is called "Euler's theorem". In order to see how the example taken by Steiner can be described in terms of my categories, it is useful to come back to Euler's original formulation of the theorem.

Euler gave a geometrical proof of the existence of the instantaneous axis, for the first time, in his E177 [Euler, 1750][27]. In E177, after a geometrical-analytical argument, Euler adds a purely geometrical proof of the existence of the instantaneous axis of rotation, discussing the infinitesimal motion of a spherical surface with a fixed point. This proof, which is given in the framework of Euclidean geometry and which persuades Euler of the existence of the instantaneous axis of rotation, is performed by reasoning visually on a diagram[28]. As Euler affirms:

> Sans entrer dans le détail du calcul, que je viens de déveloper, on peut aussi prouver la même vérité par la seule Géométrie. Qu'on considère dans le corps une couche sphérique, dont le centre soit dans le centre du gravité du corps, car il est évident, qu'ayant connu le mouvement de cette superficie sphérique, le mouvement du corps tout entier sera déterminé. [Euler, 1750, p. 96]

Euler considers the geometrical proof as an explanation of the physical phenomenon (take in mind the explanandum: the fact that for every rotation of a rigid body with a fixed point there exists an instantaneous axis around

[26] See section 1.3 for a presentation of the theorem's proof and Steiner's account.

[27] In 1910 and 1913, the swedish mathematician Gustav Eneström completed the first comprehensive survey of Euler's works. Each work is classified by using the letter E and a number (the works are referred to as "Eneström number").

[28] For a reconstruction of Euler's geometrical proof see [Koetsier, 2007, p. 184-185].

which the rotation is made), and he explicitly affirms that such an explanation is given without recurring to exact calculations ("Sans entrer dans le détail du calcul, que je viens de déveloper, on peut aussi prouver la même *vérité* par la seule Géométrie")[29]. The instantaneous axis exists because we can easily construct it geometrically. In such a construction Euler makes use of the ability to reason visually on a diagram. The conceptual resources which permit Euler to apply such an ability come from the Euclidean geometrical framework (for instance, the concept of euclidean distance makes possible to reconceptualize a fixed distance between two points of the actual body).

Now, Euler's geometrical argument is perfectly valid from a formal point of view[30]. However, in modern textbooks of classical mechanics the explanation of *why* there exists such an axis is given by recurring to the formalism of linear algebra and group theory[31]. The interesting observation for us is that scientists regard the modern proof as genuinely explanatory [Goldstein, 1957, p. 156]. This is what happens in Steiner's case, where the theorem is proved by using the tools of linear algebra. It would be interesting, then, to investigate the transition from Euler's formulation to the modern one in terms of linear algebra. This sort of investigation might reveal that there is some common feature which is supposed to preserve the 'explanatoriness' of the old formulation in the modern formulation as well. Another option would be that, in this transition, standards of explanations have changed. Thus the modern formulation in terms of linear algebra would be regarded as explanatory in our context, but it would not possess any explanatory virtue which is shared with the original formulation. In this subsection, I suggest the idea that the effect of this transition has been to generate a new ability

[29] Note that in considering Euler's geometrical proof as a genuine explanation I am not endorsing any particular account of mathematical explanation in physics, such as Steiner's. The fact that this is an explanation of the phenomenon is recognized by Euler himself.

[30] The possibility of proving Euler's theorem via a geometrical procedure is not lost, i.e. euclidean geometry still works (why should it not?). For instance, although different from Euler's, a purely geometrical proof of the theorem is given by Whittaker [Whittaker, 1904, p. 2] and Targ [Targ, 1987, p. 221].

[31] I have presented the theorem and the relative proof in section 1.3.

to reason which is now employed as intellectual tool in the example analyzed by Steiner. This ability to reason embeds our ability to reason visually, but should be considered as a different ability to reason. Its use in MEPP is made possible by conceptual resources. However, these conceptual resources are different from those which permitted Euler to use his ability to reason visually.

As a consequence of a historical process, concepts like determinant, matrix, linear systems, orthogonality or point-to-point transformation are today included in the mathematical apparatus of linear algebra and we can profit from their interplay without exiting from this framework. Those concepts are linked together in the solid framework of linear algebra and such a linkage provides, potentially, conceptual resources which were not available before. The interplay of concepts in this network of concepts is so complex that it does not permit any easy separate analysis of the mathematical elements which are found in the proof structure of a theorem such as Euler's. This difficulty is well remarked by Israel Kleiner in his *History of Linear Algebra*:

> Among the elementary concepts of linear algebra are linear equations, matrices, determinants, linear transformations, linear independence, dimension, bilinear forms, quadratic forms, and vector spaces. Since these concepts are closely interconnected, several usually appear in a given context (e.g. linear equations and matrices) and it is often impossible to disengage them [Kleiner, 2007, p. 79]

To come to our case, in the shift of the formulation various different geometrical and analytical concepts have been included in the framework of linear algebra. What is more important, in the modern proof, the geometrical part comes as already 'included' in the algebraic formalism and we do not need a purely geometrical argument to state the result (the existence of the axis). It is interesting to report here what Peano himself, in his *Analisi della Teoria dei Vettori* [Peano, 1898], observed:

> Thus the theory of vectors appears to be developed without presupposing any previous geometrical study. And since, by means of this

> theory, all of geometry can be treated, there results thereby the theoretical possibility of *substituting* the theory of vectors for elementary geometry itself [Peano, 1898, p. 513]

Thus vector space theory substitutes geometry, and embeds geometrical concepts within its structure. Consequently, since vector theory is a subfield of linear algebra, linear algebra will embed geometrical concepts. Euler could not profit from this 'substitution' and his proof structure had to turn to a purely geometrical system of concepts.

In the case of the modern proof we prove the result (eigenvalue +1 with eigenspace of dimension 1) by using concepts and results which belong to linear algebra[32]. Observe, moreover, that the eigenvector corresponding to the eigenvalue +1 is associated to a vector in a 3-dimensional vector space. This step is essential in the reasoning because we are convinced that the (actual) axis exists only if we assume that it denotes such a vector. In this case, the association of a particular vector with an actual object (the axis), makes it possible to have a visual picture of the situation (the eigenvector associated to an eigenvalue is visualized as a vector in a 3-dimensional vector space[33]). However, this kind of visualization seems to be somewhat different from that which appears in Euler's original geometrical proof. Where in Euler's case the visualization was performed on a diagram, here we have a sort of mental picture of the situation. The ability to reason which is employed is different from the ability to reason visually. Plausibly, the ability which operates when we 'see' the instantaneous axis of rotation behind an eigenvector embeds our ability to reason visually, but it employs an abstract ingredient as well. We have a mental image of the situation, and we do not necessitate to reason on a diagram (although such a diagram can be constructed). For simplicity, let

[32] Keep in mind that the proof I am referring to is that which follows Goldstein's book [Goldstein, 1957]. I reported the proof in section 1.3.

[33] Here I am not claiming that the geometrical interpretation of eigenvectors is intrinsic in their definition. I am assuming that under a particular 'reading' (in our case Euler's theorem in kinematics of rigid body motion), a subset of vectors of the vector space considered (the subset containing the instantaneous axis) has a geometrical representation in a diagram at time t. This representation can be obtained, for instance, by a computer graphic simulation.

me call 'ability to reason abstractly' this particular ability to reason[34].

According to my framework, in a genuine MEPP the use of this ability to reason is permitted by some conceptual resources. These conceptual resources are provided by linear algebra. Moreover, they embed some geometrical concepts which were present in Euler's geometrical proof. This is why, although different, these conceptual resources permit a reconceptualization which is very similar to that permitted by the conceptual resources used by Euler. For instance, in the modern algebraic proof the condition of invariance of distances is not mentioned because it is already included in the invariance of the euclidean scalar product which is defined in the euclidean 3-dimensional space[35]. The modern formulation includes the concept of topological space (an euclidean 3-dimensional space is a topological space) and as a consequence we do not need to work (for the particular case of our proof) with *distances* between points. What is important for a topological space is the 'form' of the space, its topological properties, which are given by the metrics induced by the norm (the norm defined in our vector space). To underline the shift from Euler's concept of distance (euclidean distance) to the modern one based on the norm is important because it allows us to look at the change in the structure of the explanation through the lens of conceptual resources. Euclidean distances are *already* included in the mathematical concept of norm, which stands as a new conceptual resource available in the framework of linear algebra. This conceptual resource, as that given by the concept of vector, permits a reconceptualization of the actual situation and makes it possible to use our ability to reason abstractly.

To observe that it is possible to reason qualitatively through an ability to reason (to have a mental picture of the axis) does not mean that the mathematical formulation does not include any (quantitative) calculatory step, or that such calculations have no role in MEPP. Of course, calculations are necessary in the process of explaining a scientific phenomenon. For instance,

[34]The ability to reason abstractly considere here should not be confused with abstract explanation, which is the particular kind of explanation considered by Pincock.

[35]Recall that orthogonal transformations preserve scalar product.

in this case algebraic calculations are necessary to state the existence of an eigenvalue +1 (which corresponds to an eigenvector). However, in my view, these calculatory steps are not sufficient to have a genuine explanation and they do not carry any explanatory power. A genuine explanation results from the use, at some step in the calculations, of one or more abilities to reason. It is the use of these abilities to reason that must be considered as an essential ingredient in a genuine MEPP.

To sum up, in the example given by Steiner there is a particular ability to reason which acts, our ability to reason abstractly, and the use of this ability to reason is permitted by some conceptual resources (by the concept of vector and by the concepts provided by the framework of linear algebra, for instance the concept of norm). Our ability to reason abstractly embeds our ability to reason visually, but the latter is different from the former. There is probably much more to say on the development of this ability, which is used as a natural instrument in linear algebra (for instance, when we sum two vectors we can have a mental image of the algebraic operation we are performing). And it is reasonable to think that the progressive embedding of geometrical concepts into the abstract concepts of linear algebra, passing through the theory of vector spaces, has contributed to the development of this ability and its acceptance as a natural tool in our mathematical practice.

Finally, it should be noted that seeing the existence of a physical phenomena *behind* the modern proof (i.e. when we consider an eigenvector we have a mental picture of an instantaneous axis of rotation) it is possible only via an assumed isomorphism between physical space and the euclidean 3-dimensional space (the vector space) we are working with. The isomorphism in question is, to use Steiner's words, a necessary 'bridge principle' in order to analyse the physical system in terms of vectors. Also in the case of Euler, where the physical space is considered isomorphic to the geometrical space, there is a bridge principle which implictly operates. However, as stressed by Lyon and Colyvan [Lyon *et al.*, 2008, p. 16], the fact that some necessary bridge principle should be adopted in MEPP does not reduce the importance of the mathematical part of the explanation in question (the bridge principle is itself a piece of mathematics, because it is defined in terms of isomorphism). As I suggested in Pincock's case discussed above, the mapping is correlated to some extent to the conceptual resources which act in the MEPP.

8.4.5 Generalization: strategy

In the previous four subsections I suggested that the examples of MEPP seen throughout this thesis, and proposed by the different authors analyzed in parts I and II, can be rewritten according to my categories of conceptual resources and intelectual tools. This would preserve the intuition that these cases are genuine cases of MEPP.

The interesting point is that the use of one or more intellectual tools might correspond to a particular species of explanation, and this idea is perfectly in line with the pluralist hypothesis. In Batterman's example, the ability to reason asymptotically is what characterizes asymptotic explanations, while in Pincock's example I have suggested that two abilities to reason are employed (the ability to reason analogically and the ability to switch from the graph to the actual system). For Kitcher's example and Steiner's I put forward the idea that in those cases two other abilities to reason are used as well, the ability to reason in terms of the Newtonian pattern and the ability to reason abstractly. Table 8.1 provides a summary of the abilities to reason employed in every explanation. Again, this kind of investigation would require a more comprehensive analysis. I have only offered a very general intuition about how such an analysis, if performed, would confirm the applicability of my framework.

What then about a possible generalization of my schema? A potential way to generalize it would consist in identifying the abilities to reason which are used in scientific practice in a particular scientific context (a task which might be well delegated to cognitive science), and then consider how such abilities are used through mathematical concepts in the practice of explain-

Example	discussed by	Intellectual tool(s)
Universality of critical phenomena	Batterman	ability to reason asymptotically
Königsberg bridges	Pincock	ability to reason analogically, ability to switch from the analog to the actual system
Newtonian phenomena	Kitcher	ability to reason in terms of the Newtonian pattern
Existence of an instantaneous axis of rotation	Steiner	ability to reason abstractly

Table 8.1: Intellectual tools used in the respective examples.

ing. This would result in the identification of a variety of explanations, where each explanation makes use of one or more particular abilities to reason[36]. If my intuition is correct, the use of these abilities to reason is permitted by one or more conceptual resource. As natural, this procedure might result to be interesting for the study of new species of MEPP (explanations which use abilities to reason different from these considered here), but it might even reveal the fallacy of my approach.

8.5 Payoff, directions of analysis

In this section I will suggest how my framework is supposed to provide insights into two distinct topics which appear in the contemporary philosophical debate on explanation.

First of all, I will reconsider the classical asymmetry problem of explanation (presented in section 2.2 and discussed in relation with the accounts in various sections of part I). This problem is not discussed in the contemporary

[36]As I remarked in section 8.2, and as we have seen in my discussion of Pincock's example in subsection 8.4.3, I do not exclude that two or more intellectual tools might be used together and then combined in an explanation. For instance, the ability to reason visually and that to reason asymptotically might both be used in an explanation of a physical phenomenon, and this would result in a species of explanation different from that involving only one of such abilities.

philosophical literature in relation to MEPP, and this is because of the intervention of causal claims in what is regarded as the correct strategy to solve it. I will show how my framework can provide a possible way to account for what is generally seen as a solution to the asymmetry problem, and do this without losing the reference to the causality claims. In other words, I will show that what is considered as the right solution to the asymmetry problem can be written in terms of my categories. This will suggest the idea that my approach is not limited to MEPP but can be extended to cover scientific explanation as well.

Secondly, I will show how my framework, if adopted, would have a thought-provoking repercussion for the ontological debate between platonists and nominalist in the philosophy of mathematics. In particular, it might provide a useful lever on realist's claims about the existence of mathematical objects as they appear in the Enhanced Indispensability Argument (EIA).

Even if I will not report a comprehensive analysis here, I think that the lines which follow contain sufficient detail to show that the framework I propose can provide new directions of investigation.

8.5.1 Asymmetry problem revisited

For the sake of clarity, let me shortly reconsider the asymmetry problem of scientific explanation. The problem of asymmetry arises when we have pairs of deductively valid arguments which rely on the same law(s) and which differ radically in explanatory potential. The classical example is that of the flagpole and the shadow, which has been paraphrased by Van Fraassen by using the example of the tower and the shadow[37]. If we consider a flagpole and its shadow, by using the laws of optics together with the laws of trigonometry and some physical assumptions we can explain why the shadow has that particular length by considering the length of the flagpole. Now, while the same kind of deduction is perfectly legitimate via the same laws

[37]See subsection 2.3.3 for Van Fraassen's example and his story of the tower and the shadow.

the other way around (deduce the length of the flagpole from that of the shadow), it seems nonsense to say that the length of the shadow *explains* the height of the flagpole. Observe that trigonometry laws alone do not permit the choose that one direction is preferable, and therefore these laws do not act as discriminant in the choice of the preferred explanation. Thus, in order to solve the problem of asymmetry, namely, in order to pick out what is generally considered as the genuine explanatory direction, we need something more.

Recall now Wittgenstein's motto: bring you commonsense with you and don't leave it outside, when you enter the room to philosophize. In this case, it is reasonable to consider that a specific ability, our ability to reason causally, has a discriminant role in the choice of the 'good direction' of the explanation (that in which we explain the length of the shadow through the length of the flagpole by using geometrical considerations and physical assumptions). If this intuition is right, the difference in explanatory potential in the asymmetry problem is due to the fact that one direction of the explanation gets the asymmetric causal order *right* (the length of the shadow is causally conditioned by the height of the flagpole), while the other does not. Intuitively, we know that the shadow is "caused" by the flagpole because we assume that light rays travel in straight paths and the dark area (the shadow) appears *when* there is an object (the flagpole) between a source of rays (the sun) and a surface (the ground). This information is in our background knowledge under the form of physical concepts, which are not timeless but are context-dependent and could vary over time[38]. For simplicity, let's call (α) and (β) the two situations:

(α) the length of the shadow is explained through the height of the flagpole

[38] Try to think, for instance, to a solution of the asymmetry problem in a world *à la* Lewis Carroll in which shadows do cause flagpoles by means of some strange mechanism. In this case, if we assume that the Pythagorean theorem is right and trigonometry laws have not changed, we will accept as good explanation that which uses the shadow's length to explain the flagpole height. This exactly because Carroll's world, as scientists have discovered during their investigations, affects our commonsense in that precise way (call the relation between shadows and flagpoles in that strange world 'acausation').

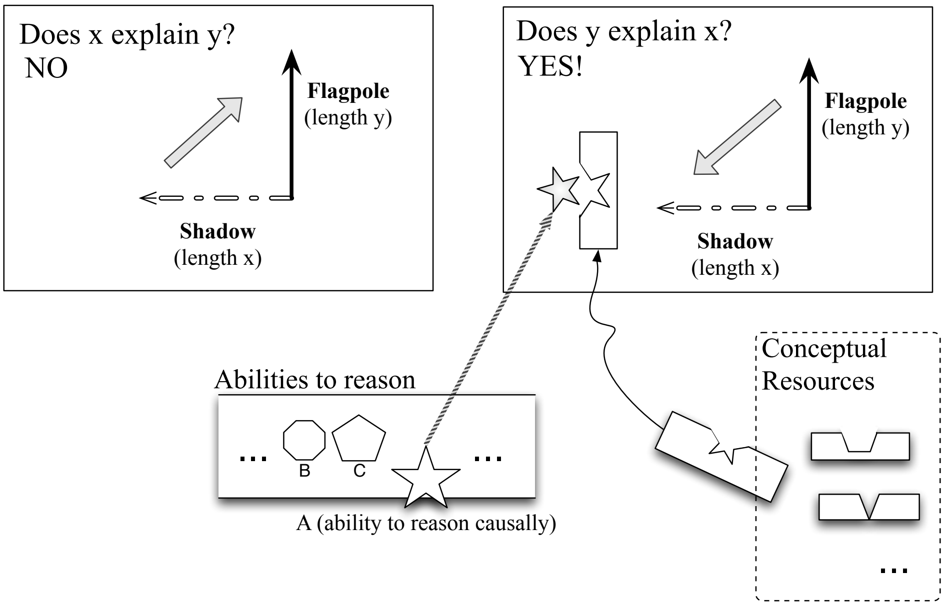

Figure 8.4: My framework for the case of the asymmetry problem of explanation. Our ability to reason causally is an intellectual tool, used through conceptual resources.

(β) the height of the flagpole is explained through the length of the shadow

To anticipate my point, what I want to show is that we can solve the asymmetry problem because in (α) we dispose (with respect to the explanation the other way around) of some extra utility which comes into play and permits us to discriminate between the two putative explanations. This extra utility is our ability to reason causally, i.e. our ability to recognize that there is a functional dependency between two or more determining elements. Our ability to reason causally corresponds, in my view, to an intellectual tool which is used in the explanation and which contributes to the acceptance of one particular direction of the deduction as explanatory. However, to say that we dispose of such an intellectual tool is not enough. According to my framework, the use of this tool is possible because we dispose of some conceptual resources which permit a reconceptualization and the application of our ability to reason causally (as shown in Figure 8.4). In the following

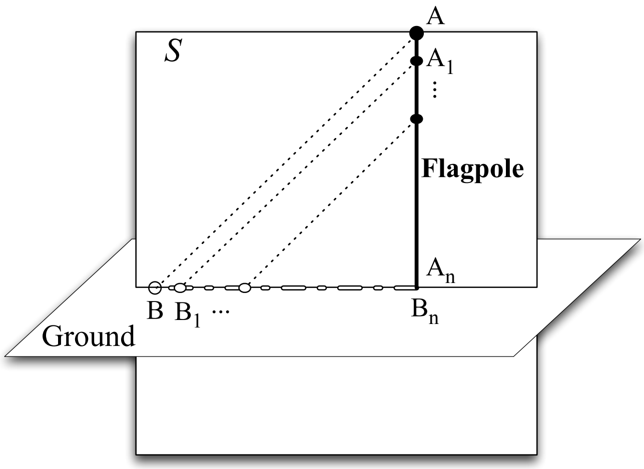

Figure 8.5: The situation reconceptualized through euclidean geometry.

paragraph I provide a justification for this claim.

In this case geometrical concepts (distance, angle, etc...) permit us to reconceptualize the physical objects 'flagpole' and 'shadow'. Through the lens of euclidean geometry we see the flagpole and the shadow as segments on a plane (the two-dimensional surface S of Figure 8.5). Morever, from the theory of light we assume that light rays travel in straight paths, and then the same euclidean concepts permit us to also see the path of one light ray coming from the sun as a segment of the same plane (more precisely, we focus on a part of that segment, i.e. that going from the flagpole to the ground). We have, finally, three segments which can be seen as sides of a triangle. To reconceptualize the problem as such allows us to make derivations about the length of the shadow from the length of the flagpole, but also the other way around (through trigonometrical laws). However, observe that even if the reconceptualization permits deductions in both senses, we dispose of an extra ingredient which can be added to our story. We know from our experience that to an occurrence of an event at point A on the flagpole (A is the edge of the flagpole/segment) there corresponds another event at point B on the ground. The event in A, i.e. the interaction of the light ray with the flagpole and its blocking, is distinct from the event in B, i.e. the emergence of a shadow (or shadow-point). Moreover, our (accepted) physical theory says that the event at B is a consequence of the event at A (event at B is posterior to A and it depends on the occurrence of A). The same holds for all the other points $A_1 \dots A_n$ and $B_1 \dots B_n$, belonging, respectively, to the flagpole-segment and the shadow-segment. In other words, we know that the shadow-segment is the sum of each causally connected event $AB, A_1B_1, \dots, A_nB_n$. By reasoning in this way, we are using an ability to reason, and precisely our ability to reason causally. This ability can be used, in this particular situation, because some conceptual resources (coming from euclidean geometry) make possible a reconceptualization and the application of our particular ability. Therefore our ability to reason causally is an intellectual tool. It injects an extra ingredient into our purely deductive considerations. This ingredient provides us with a sort of natural 'persuasion' and permits us to pick out one direction as genuinely explanatory. In situation (β) the very same reconceptualization is not sufficient to provide the same sense of explanatoriness, and this is exactly because our ability to reason causally suggests the other direction as natural.

My discussion above provides, I think, a possible way to account for a correct solution to the asymmetry problem without recurring to any specific model of explanation. We can discriminate *the* genuine explanation, but we do not endorse any specific *account* of MEPP in performing such an evaluation[39]. However, and this would agree with my idea that the different species

[39]Remember that, as we have seen in section 2.3.3, Van Fraassen gave the same status of genuine explanation to both the directions in the asymmetry problem. For him the explanation (the answer to the question "Why is the shadow so long?") given by the Chevalier, and based on laws of trigonometry and the straight path of light rays, was exactly on a par with the explanation (the answer to the question "Why is the tower so high?") given by the housemaid and based on the love-story between the Chevalier and the maid. Van Fraassen claimed that context makes appropriate to explain the tower's height in terms of the length of the shadow it casts, and the resulting explanation is a genuine explanation. What then about his approach to the asymmetry problem and my perspective? I think that in the case proposed by Van Fraassen the explanation given by the housemaid and based on the love-story between the Chevalier and the maid must be excluded as genuine. This is because, in my view, conceptual resources must be *scientific* concepts used in scientific practice, and the direction of the explanation making use of the love-story between the Chevalier and the maid does not make use of scientific concepts.

of explanation can be discriminate by making reference to the intellectual tool which employ, the use of our ability to reason causally as an intellectual tool denotes a particular form of explanation (causal explanation).

Observe that the observation that causation constitues an *extra* requirement to be added when we are faced with particular scientific explanation is something which is not new and which has been stressed by various authors. For instance, the 'unificationist' Schurz points out that:

> it seems to follow that in the area of explanations of singular events, we have to add the causality requirement as an *extra* requirement, which goes beyond the idea of unification in a merely inferential (or information-theoretic) sense [Schurz, 1999, p. 100]

In my argument above I did not necessarily have to endorse some account of causality and causal relation (such as Salmon's [Salmon, 1984a], Dowe's [Dowe, 2008] or Woodward's [Woodward, 2003]). The fact that the flagpole *causes* the shadow perfectly accords with the conceptual schemes developed, used, and tested in scientific practice, no matter what causality means. This 'practice-driven' attitude towards causality does not undermine the useful role that our ability to reason causally has in science, neither does it make the philosophical analysis of it meaningless[40]. To attribute to causality such an heuristic value, without defining a sense of causal relation or adopting a theory of causality (and then without offering a potential solution to Hume's

---

On the other hand, the explanation in terms of trigonometrical and physical considerations makes use of scientific concepts and permits a fruitful reconceptualization. As I proposed above, these concepts do permit to apply an ability to reason. This marks an essential difference between my approach to explanation and Van Fraassen's. More generally, I think that the task to provide some explanation in science is a quite different affair from that of providing other sorts of explanation in our everyday life. For instance, it would be very hard to imagine a physicist who, when asked for a genuine explanation of some natural phenomenon, points to a love story as an essential ingredient of *why* that phenomenon occurs.

[40]The term "practice-driven" was used in opposition to "methaphysically-driven" accounts of causality during the *Barcelona Conference on Causality and Explanation in Physics, Biology and Economics* (Barcelona, February 2010) [Sus et al., 2010]. In particular, a practice-account of causation involves no metaphysics and does not provide answers to what causation *is*.

problem), is an attitute which is shared by a contemporary trend in philosophy of science. For instance, in his paper "Causation as Folk Science" [Norton, 2007], John Norton endorses this anti-"causal-fundamentalist" attitude[41].

Let me spend some additional words on the ability to reason causally. Note that I am not claiming that our ability to reason causally is a primitive cognitive faculty. We do not *possess* this ability. On the other hand, I think that it is reasonably to say that we do possess the faculty to cognitively discriminate events. For instance, we can discriminate the event 'emission of light from the sun' from the event 'appearance of the flagpole's shadow', and the latter from the event 'blocking of the light rays by the flagpole'. The same happens when we observe two balls moving on a billiard table and hitting the rail in two distinct points. In that case we are able to say that the event 'collision of the first ball with the rail' is distinct from the event 'collision of the second ball with the rail'. Call *event discrimination* the cognitive faculty which permits to discriminate between these events. Event discrimination, therefore, can be reasonably thought as a primitive cognitive faculty. Although necessary to reason causally, however, this faculty is not sufficient to it. Event discrimination permits to recognize separate events, but it does not permit to individuate a causal connection (if there is any!) between them. Unfortunately, I do not have very much to say on the relations between event discrimination and our ability to reason causally. For my focus here is on the latter, and my claim is that such ability is acquired and is not a cognitively primitive faculty, perhaps the best way to clarify my intuition is to consider that our ability to reason causally results from the use of our faculty

[41]Where *causal-fundamentalism* is defined as the doctrine according to which "Nature is governed by cause and effect; and the burden of individual sciences is to find the particular expressions of the general notion in the realm of their specialized subject matter" [Norton, 2007, p. 13]. Recall how Bertrand Russell claimed against a concept of causal law defined in terms of relation betwen events and, even in a stronger way, against "universal determinism" (i.e. the idea that every event has a cause). Here is the famous quote: "The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving like the monarchy, only because it is erroneously supposed to do no harm" [Russell, 1913, p. 1].

of event discrimination together with the use of our scientific theories and empirical considerations. Roughly, the use of event discrimination and that of our best science provides us with an ability to reason which is now seen as natural in a specific context and which can be used in scientific pratice. This is what happens, for instance, when we say that the flagpole 'causes' the shadow. There are two separate events, for instance in A and B, and we have a scientific theory which suggests that a causal linkage exists between these events. Our ability to reason causally comes then from empirical considerations, from a scientific theory plus event discrimination as an essential ingredient. Again, the use of this ability does not require some account of causality or causal relation. Causality has an heuristic value, and to reason causally in a particular situation (as in the asymmetry problem) does not require a knowledge of how the causal connections are supposed to operate. Paul Klee once said: "do not reproduce; but make visible". In the context of the asymmetry problem, I would say: do not reproduce the causal relation (if there is any), but make visible how that relation is used.

Finally, observe that in the asymmetry problem I considered that conceptual resources do permit a reconceptualization of a physical object as a mathematical object, thus allowing us to apply our ability to reason causally. In doing that, I implicitly expressed the idea that my framework can be applied to causal explanations as well. Unfortunately, given the topic of this dissertation, it is not feasible to reinforce this intuition by providing an analysis of other cases of causal explanation in terms of my framework.

8.5.2 My approach and the Enhanced Indispensability Argument

In section 2.1 and subsection 5.2.1 we have seen as some philosophers committed to the existential attitude (notably Mark Colyvan [Colyvan, 2002] and Alan Baker [Baker, 2005]) attempt to show the existence of mathematical objects by the so-called "Enhanced Indispensability Argument". Those philosophers refer to the indispensable explanatory power of mathematics in scientific theories as an instrument to support the claim that some mathematical objects exist[42]. More precisely, they use the following deductive schema to infer the existence of mathematical objects:

Enhanced Indispensability Argument (EIA)

(1) We ought rationally to believe in the existence of any entity that plays an indispensable explanatory role in our best scientific theories.

(2) Mathematical objects play an indispensable explanatory role in science.

(3) Hence, we ought rationally to believe in the existence of mathematical objects.

However, this argument is extremely controversial and various criticisms have been leveled against it. Some authors have pointed out that the explanatory power of mathematics cannot be used in such existential inference and the explanatory utility of a mathematical model does not depend upon the actual existence of the mathematical objects posited by the model ([Leng, 2005], [Bangu, 2008])[43]. Furthermore, it has been observed that mathematics does not play any explanatory role in science, but only a representational one [Saatsi, 2011]. Very generally, the strategy adopted by the nominalists and by the opponents of the EIA in order to reject the realist's ontological claim is to focus on premise (2) and consider that it is not true. Additionally, it should be noted that the expression 'indispensable explanatory role' in the EIA is used without the necessary clarifications

[42]See [Baker, 2009] and [Saatsi, 2011] for a survey of the debate concerning the enhanced indispensability argument. Baker considers this argument as an improved ('enhanced') version of the Quine-Putnam indispensability argument, whose first explicit formulation appears in [Putnam, 1971, p. 65]. The improvement is given by the fact that the enhanced argument refers to the *explanatory* role of mathematics in science, while in the Quine-Putnam indispensability argument there is no such reference and the argument focuses on indispensability *simpliciter*. The enhanced argument appeals to the role of inference to the best explanation (IBE) in the defense of scientific realism. See [Colyvan, 2001] and chapter 6 of [Panza *et al.*, 2010] for a survey of the Quine-Putnam indispensability argument.

[43]For instance, Mary Leng has rejected the argument on the grounds that mathematical explanations need not have *true* explanans, and thus the objects posited by such explanations do not necessarily exist [Leng, 2005].

[Panza *et al.*, 2010]. What does it mean for a mathematical object to play an 'explanatory role' in science? Further clarifications on the use of the notion of explanatoriness (of mathematics or mathematical objects in science) are required, and the same holds for the notion of 'indispensability' (of a mathematical object playing an explanatory role). There is no general consensus on how these notions are used in the EIA. What is indispensable in the EIA? The explanatory role played by some mathematical objects, or the explanatory role played by mathematical objects in general? It is not clear if the indispensable explanatory role is that of mathematics *in toto*, of mathematical theories, of properties of mathematical objects, of mathematical objects or of some particular mathematical objects. As Baker rightly observes, these are distinct levels, and an attack on EIA should take into account all these different levels [Baker, 2009, p. 615].

I will not give here the a full presentation of the debate about EIA, neither I will address the numerous questions which belong to that debate and which are resulting in a very rich discussion in philosophy of mathematics. Rather, for my discussion here, I will assume the following reading of premise (2): there are some mathematical objects that play an essential role in every genuine explanation of some scientific fact. These mathematical objects are then indispensable to explain a scientific fact because without a reference to them the genuine explanation could not be given. This reading of premise (2) does consider those and only those mathematical objects that are explanatorily indispensable in the scientific practice. The resulting argument is referred to as 'strong EIA' [Baker, 2009, p. 616]. In addition to this first EIA, one may adopt a weaker reading of premise (2) and consider that every adequate explanation of a scientific fact involves reference to some mathematical objects. This reading does correspond to a weaker version of the EIA. Both these readings of premise (2) lead to an argument for realism concerning mathematical entities, but while the first leads to an argument for the existence of some mathematical entities (the strong EIA), the latter leads to an argument for the existence of mathematical entities in general (the weak EIA). Henceforth I will concentrate on the strong EIA and for simplicity I will refer to it simply as EIA[44].

Does my approach have something interesting to say on the EIA? I think it does. In particular, my approach to MEPP supports the idea that the inference obtained through EIA is not viable, but for reasons different from those of Bangu, Leng or Saatsi. In this subsection I am going to suggest how such an argument can be constructed.

There are two possible strategies to see how my framework can be used to undermine the correctness of the EIA, and both focus on the expression 'indispensable explanatory role' which appears in the two premises. According to my framework, in MEPP the mathematical conceptual resources permit a reconceptualization and the application of an ability to reason (intellectual tool). Hence, if there is some mathematics which plays an indispensable explanatory role, the attention should be addressed to the particular pieces of mathematics I call conceptual resources.

A very natural move to start would be to consider that conceptual resources should be seen as 'explanatorily indispensable' because they permit us to apply a particular ability to reason. At face value, this idea accords well with my intuitions about the role of conceptual resources and intellectual tools. Without a conceptual resource it is not possible to apply any ability to reason, therefore a conceptual resource is indispensable for explanation. However, my conceptual resources are concepts, not objects, and premise (2) is about mathematical objects. A first strategy to attack EIA is then straightforward and it consists in observing that, from the perspective of my approach, there are no mathematical objects which do play an 'indispensable

[44]Observe that the distinction between a strong and a weak indispensability argument has been inherited from the discussion in the context of the original Quine-Putnam indispensability argument: "One should distinguish two Quinean indispensability arguments. In the next section I will propose a weak indispensability argument: one can have a warrant for positing mathematical entities on the basis of their indispensability in all normal scientific theories. This is a real argument for the existence of mathematical objects, and is hard to overcome. In the third section I will propose a strong indispensability argument: the ontology consists of all the mathematical entities that are really indispensable and only these" [Decock, 2002, p. 232].

explanatory role' in science. Therefore the EIA is not correct simply because, once my approach is endorsed, premise (2) must be regarded as false. There is, however, a second strategy to attack EIA, and it is to this strategy that I will concentrate in what follows below. It consists in assuming that there are some mathematical objects which are purported to fall under conceptual resources. With this assumption and my framework in mind, EIA can be rewritten in a new form without losing the intuition which stands behind the original argument. However, this new formulation of the argument must be considered as not correct because based on a false premise.

Assume that every mathematical conceptual resource has one or more mathematical objects purported to fall under it (the mathematical concept which acts as a conceptual resource is a sortal concept). Then, if some mathematical object purports to fall under a concept (a conceptual resource) that plays an indispensable explanatory role, we can focus on this role of mathematical concepts without losing the reference to mathematical objects. This would preserve the general idea of the original argument, namely that *some* mathematical objects are supposed to exist because there is some piece of mathematics which plays an indispensable explanatory role in our science, although slightly modifying it by introducing the mention of concepts in premises (1) and (2). Call this argument EIA*:

EIA*

(1*) We ought rationally to believe in the existence of any entity that is purported to fall under concepts that plays an indispensable explanatory role in our best scientific theories.

(2*) Mathematical objects are purported to fall under concepts that play an indispensable explanatory role in science.

(3*) Hence, we ought rationally to believe in the existence of mathematical objects (purported to fall under these concepts).

With respect to the original EIA, the argument above shifts the attention from the indispensable explanatory role of mathematical entities to that of concepts. The explanatory role is thus played by concepts, and not by mathematical objects. The motivation for this change of perspective is not only based on the demand to adapt the original argument to my perspective on MEPP. There is a further reason to consider EIA* rather than the original argument, and precisely the fact that EIA is subject to a phenomenon that Crispin Wright and Martin Davies have called 'transmission failure' ([Wright, 2000], [Wright, 2002]). This phenomenon concerns the failure, on the part of some deductively valid inferences, to transmit one's justification for believing the premises. It is quite simple to see how this occurs in the EIA. Premise (2) states that mathematical objects do play an explanatory role in science. To justify this claim we have to suppose that they exist. However, the existence of some mathematical objects is exactly the conclusion (3) we want to reach. EIA, therefore, results in a logically valid argument but it exhibits circularity. To consider that mathematical objects play such an explanatory role introduces this epistemic feature into the deductively valid inference, and premise (2) must be seen as begging the question. This is why I find more plausible to focus on the idea that the explanatory role is played by concepts, and mathematical objects get involved in such a type of argument only because they are purported to fall under these concepts.

Let's now return to EIA*. What I want to show is that the argument is not correct. More precisely, I want to show that the choice of the mathematical concepts playing an explanatory role in science is arbitrary because sensitive to contextual factors. To establish such arbitrariness would amount to showing that these concepts are not explanatorily *indispensable*. As Baker writes in the context of his example of indispensable mathematical explanation of a physical phenomenon:

> One way to attack the claim that the mathematics involved in the explanation of the cicada period lengths is indispensable is to show that somehow the choice of mathematical apparatus here is arbitrary. The thought is that if it can be shown that the choice of mathematical apparatus is just one of many equally good alternatives then the

> particular mathematical objects involved cannot be indispensable to the overall explanation. [Baker, 2009, p. 614-615]

Remember that, according to my approach, the use of conceptual resources depends on contextual factors. For instance, in one particular scientific context a particular conceptual resource might be employed because it allows to apply a particular ability to reason. In another context, to explain the very same phenomenon, a different conceptual resource might be preferred because it permits us to apply another ability to reason which is considered to be more natural or preferable in that particular context. This suggests the idea that a particular conceptual resource can be *dispensable*, and that is because its explanatory role is extremely sensitive to the context in which the explanation is made. Furthermore, let me mention that in a specific context the use of an ability to reason through a conceptual resource might be regarded by scientists as providing zero explanatory power. This can happen, for instance, in a context where that particular ability to reason has not been developed or is not accepted by the scientific community. In that context, the particular conceptual resource which permits the use of that ability will not carry any explanatory power. Conceptual resources are necessary, in general, for explanation. Nevertheless, specific conceptual resources are used in a specific context, i.e. their use is context-dependent. These considerations suggest, again, that explanation is a contextual affair. More importantly, they suggest that if the explanatory role of a conceptual resource is dependent on the specific context in which the explanation is made, the fact that it appears or not in the EIA* must also be regarded as a contextual affair. To sum it up, the explanatory role of mathematical concepts is not a function which has universal validity and premise (2*) in the EIA* must be rewritten with an explicit mention of the context, namely, by including the following premise:

(2**) Mathematical objects are purported to fall under concepts that play an explanatory role *in context C*.

Observe that, with respect to premise (2*), premise (2**) does not mention the *indispensable* explanatory role of concepts (conceptual resources). And this because, if the explanatory role of conceptual resources depends on contextual factors, the choice of conceptual resources is arbitrary because based on the preferences of the scientists belonging to a particular context. To accept this contextual sensitivity in EIA* would amount to accept the (bizarre) idea that the existence of some mathematical objects is sensitive to contextual and pragmatic factors. Premise (2**) is therefore in conflict with premise (2*) appearing in EIA*, and the latter must be considered as not true (there are no such *indispensable* concepts playing an explanatory role in *every* context in science).

Nevertheless, it might be observed that some mathematical concepts do play an explanatory role in context C, and then these concepts are explanatorily indispensable in that context. Roughly, in that context the explanation of a physical phenomenon can be given only by making reference to some particular conceptual resources. In this case, premise (2**) should be rewritten with an explicit mention of the indispensable role played by these concepts: 'Mathematical objects are purported to fall under concepts that play an indispensable explanatory role in context C'. Call this premise (2***). This would not block the existential inference in the EIA*, because to say that mathematical concepts have an indispensable explanatory role depending on the context does not undermine the following argument (EIA* with context-dependence, or EIA***):

EIA* context-dependent (EIA*)**

(1***) We ought rationally to believe in the existence of any entity that is purported to fall under concepts that plays an indispensable explanatory role in science in context C.

(2***) Mathematical objects are purported to fall under concepts that play an indispensable explanatory role in context C.

(3***) Hence, we ought rationally to believe in the existence of mathematical objects (purported to fall under these concepts).

In fact, if some mathematical concepts do play an indispensable explanatory role in context C, the EIA* with context dependence (EIA***) can be used to infer the existence of the mathematical objects purported to fall under these concepts. However, according to my approach, the explanatory role of conceptual resources, and then of mathematical concepts, is not indispensable within a particular context. This is because, in order to explain a physical phenomenon, a conceptual resource can be used in context C (for instance because it permits to use a particular ability to reason), but in the same context C another conceptual resource might be used as well. In this case, there is no concept which would play an *indispensable* explanatory role in context C, but its role would be dispensable. Therefore premise (2***) would not be compatible with my approach and the inference obtained through EIA*** cannot be accepted.

Finally, my discussion above discredits the conclusion that the existence of some mathematical objects can be inferred from considerations about the explanatory role of mathematical concepts. The revised version of the EIA (EIA*), seen from the perspective of my approach, must be considered as not correct because premise (2*) is false[45].

A possible moral of this subsection, at least if we accept my approach to explanation and the argument I proposed against EIA*, is that the notion of

[45]Remember that the strong EIA concerns the existence of some mathematical objects, while the weak EIA that of mathematical objects. Of course, I have focused on the strong EIA and my argument does not undermine the existential claim resulting from the weak EIA. In fact, according to my framework, mathematics is necessary to genuine MEPP, and then every genuine explanation of a scientific fact will inevitably involve reference to some mathematical objects. To introduce a mention of the context would not undermine the weak EIA simply because in every context the genuine explanation involves reference to mathematical concepts (conceptual resources) and to the mathematical objects purported to fall under these concepts. Reference to mathematics is indispensable for MEPP. Unfortunately, I do not have any argument against the weak EIA. I suppose, however, that a possible strategy to undermine the weak EIA and preserve my approach would be to propose a nominalist paraphrase of the mathematics which is involved in the explanation, thus showing that mathematics is dispensable (although being explanatorily relevant).

explanation does not carry any ontological value. Furthermore, a general remark concerning EIA would be to note that the transmission failure in that argument reveals a confusion between ontological and epistemological levels. Explanation, in line with my approach in terms of conceptual resources and intellectual tools, seems to be an epistemic affair. And the notion of explanation need not track the ontological status of mathematical entities. If influenced by subjective factors, as my approach suggests, the notion of explanation cannot be used in inferences such as EIA whose conclusion has ontological nature[46] .

8.6 Three big questions for my approach

There are, of course, various difficulties with my framework which have not been solved or even approached in this final chapter. My ideas are still at a preliminary stage, and the fact that the notion of explanation appears in numerous and distinct topics in the philosophy of science makes things even more difficult.

There are at least three big questions which are still unanswered (at least explicitly) and which are extremely relevant for my approach to explanation:

α What about the notion of understanding and its linkage with explanation?

β It is reasonable to maintain that in science we *always* use an ability to reason, even if we welcome the idea that these abilities come under different species and are employed in particular contexts. Hence, according to my notion of intellectual tools which act in genuine MEPP, every practice in science must be regarded as genuinely explanatory and the notion of MEPP would result as meaningless. How can we avoid this trivialization?

[46]Panza and Sereni raise a similar point about the necessity to make clear how the notion of explanation, which might be influenced by subjective factors that depen upon our (limited) cognitive capacities, can be used in an argument such as EIA which leads to a conclusion of ontological nature [Panza *et al.*, 2010, p. 202].

γ It has been claimed that there is a mutual interaction between intellectual tools and conceptual resources: we acquire conceptual resources through the use of tools and we acquire intellectual tools through the use of conceptual resources. Intellectual tools, as conceptual resources, are therefore not unchangeable but can vary over time. How is this interplay between the conceptual resources and intellectual tools supposed to work?

8.6.1 Understanding and explanation (α)

Let me begin by concentrating on question α. It is often assumed that explanation and scientific understanding bear a direct and intimate connection [De Regt, 2009]. Moreover, the wish of having a relationship explanation-understanding as a desideratum of a theory of scientific explanation has been expressed by different authors. Just to list some of them: [Friedman, 1974, p. 6], [Tuomela, 1980, p. 212], [Railton, 1981, p. 243-244], [Kitcher, 1981, p. 508], [Achinstein, 1983, p. 16], [Salmon, 1989, p. 134-135], [Lewis, 1993, p. 185], [Weber, 1996, p. 1], [Schurz, 1999, p. 98], [Tappenden, 2005, p. 166]. For instance, in section 3.1 we have seen how Michael Friedman built his theory of explanation starting from the assumption that every theory of explanation should connect explanation and understanding: "We can find out what scientific understanding consists in only by finding out what scientific explanation is and vice versa" [Friedman, 1974, p. 6]. Friedman regarded the notion of scientific understanding as depending on psychological factors, but having an objective value for a group of individuals and being epistemically relevant for the philosophical analysis of the notion of explanation. In his unification account, scientific understanding increases as we decrease the number of independent assumptions that are required to explain what goes on in the world.

On the other hand, some philosophers have refused the idea that a theory of explanation should inform on such a linkage and they have adopted what Henk De Regt has called the "objectivist view" of the relation between explanation and understanding [De Regt, 2009]. According to this view, whose major advocate is Carl Hempel, understanding should be banned from the philosophical discourse concerning explanation. The objectivist view looks at the nature of explanation as objective, while the notion of understanding is seen as having a pragmatic and subjective value, i.e. it has to do with the individual beliefs or attitudes of the scientists involved in the process of explaining [Hempel, 1965]. The philosophers of science, whose aim is to give an objectivist account of explanation (at least on this Hempelian view), must then regard understanding as philosophically irrelevant for explanation. For instance, J.D. Trout, a contemporary defender of the objectivist view, cosiders the feeling of understanding as a subjective experience that may be induced by explanations but which should not be regarded as epistemically relevant to these [Trout, 2002].

Therefore, there is no consensus on the fact that understanding does play an epistemic role in explanation and that our theories of explanation should be built in order to incorporate such a notion (whatever this notion may be). There is no doubt, however, that a well formulated picture of explanation should, in principle, have something to say on the linkage explanation-understanding in science (or on the fact that there is no such linkage)[47]. The aim of this subsection is to propose a characterization of such a linkage from the perspective of my approach. In doing that, I will use as a guide the following remark put forward by Wesley Salmon:

> Scientific understanding is, after all, a complex matter; there is every reason to suppose that it has various different facets [Salmon, 1989, p. 183]

Observe that, in line with Van Fraassen's approach to explanation, I consider 'explanation' as the result of a practice in science (as in our everyday life). This practice is performed by a scientist or a scientific community

[47]For instance, Friedman observed: "I don't see how the philosopher of science can afford to ignore such concepts as 'understanding' and 'intelligibility' when giving a theory of the explanation relation" [Friedman, 1974, p. 8].

in a precise scientific context (a community where some corpus of beliefs is shared). The same holds for MEPP, where the practice of explaining draws on mathematical facts and causal considerations are supposed to play no role. However, very differently from Van Fraassen, I claim that when an explanatory practice is performed, scientists do express their preference for a particular way to explain the phenomenon (even if they dispose of an alternative formulation of it), *and* this preference can be accounted for in terms of intellectual tools and conceptual resources. To put it in other words, scientists do attribute explanatoriness to a particular account, and to this genuine MEPP there corresponds the use made by scientists of intellectual tools through mathematical conceptual resources. The crucial point, which condenses (and anticipates) my intuitions about the linkage explanation-understanding, is that to this genuine MEPP there corresponds a genuine sense of understanding (an explanatory understanding). In this sense, I will make no sharp distinction between a mathematical explanation of a phenomenon X and the understanding of the same phenomenon, for having an understanding of X amounts to having a genuine explanation of X. However, I will claim that there is a 'pragmatic' understanding which is involved in such a genuine explanation and which is necessary to it. Before I substantiate this claim, let me turn to De Regt's 2009 paper "The Epistemic Value of Understanding" [De Regt, 2009] as to make clear to what 'facets' of understanding I am referring to.

De Regt distinguishes three different types of understanding which can be used in connection with scientific explanation [De Regt, 2009, p. 588]:

(FU) Feeling of understanding = the subjective psychological experiences accompanying an explanation.

(UT) Understanding a theory = being able to use the theory (pragmatic understanding).

(UP) Understanding a phenomenon = having an appropriate explanation of the phenomenon.

The feeling of understanding (FU) is neither necessary nor sufficient for (UP), as Hempel and Trout observed. On the other hand, contra the objectivist view, De Regt claims that (UT) is a necessary condition for (UP), where (UT) amounts to "the ability to use relevant theories to construct explanations". He writes:

> I will argue that actual scientific explanation involves a kind of understanding that is pragmatic and hence not purely objective. This type of understanding is based on skills and judgments of scientists and cannot be captured in objective algorithmic procedures. It is therefore incompatible with the objectivist conception of explanation and understanding favored by Hempel and Trout. The pragmatic kind of understanding that I claim is crucial to scientific explanation is not a product of explanation [De Regt, 2009, p. 587]

Now, observe that the pragmatic understanding (UT) is pragmatic in the sense that it pertains to the scientists involved in the process of explanation. More precisely, it has to do with the attitudes and abilities of the scientists involved in this process. However, it is not arbitrary because within a scientific context a group of individuals might possess the same abilities or skills. According to De Regt, a skill is "the ability to construct deductive arguments from the available knowledge"[48]. To consider that particular skills of scientists are crucial, in a specific context, for constructing explanations and for achieving understanding (UP) in that context entails that understanding has a pragmatic dimension that is relevant to the epistemic aim (UP)[49].

[48]For instance, when asked to explain why jets fly, the scientist should be able to use the Bernoulli's principle together with the background conditions in order to derive the explanandum. This fits the phenomenon into a broader theoretical framework. To merely know the Bernoulli's principle and all the background conditions is not enough to explain the phenomenon. We have to be able to use these informations in the right way. The "extra ingredient needed to construct the explanation is a skill: the ability to construct deductive arguments from the available knowledge" [De Regt, 2009, p. 588].

[49]Moreover, as De Regt observes: "although it is possible and useful to distinguish analytically between the epistemic and the pragmatic, the two are inextricably intertwined in scientific practice: epistemic activities and evaluations (production and assessment of knowledge claims) are possible only if particular pragmatic conditions are fulfilled" [De Regt, 2009, p. 590].

Furthermore, it seems that by proposing (UP), De Regt is conflating understanding (of a phenomenon) and explanation (of the same phenomenon), where he regards an 'explanation' as an "argument that fit a phenomenon into a broader theoretical framework" [De Regt, 2009, p. 593-594][50]. This is, I think, a crucial point. And the sense of 'understanding a phenomenon' which appears in (UP) can be made clear further. Very generally, we can individuate two uses of 'understanding': when understanding is claimed for some object, such as some subject matter, and when it involves understanding that something is the case. As Jonathan L. Kvanvig has pointed out, these two uses of understanding can elucidate other uses of understanding, for instance understanding why, understanding when or understanding where [Kvanvig, 2003]. Now, I suppose that the use of understanding in (UP) is the following: 'understanding why something is the case'. For instance, understanding why the particle has passed through the wall, why jets fly or why the particle has that particular behaviour in the Hénon-Heiles potential. Nevertheless, understanding why something is the case can be accounted for in terms of understanding that something is the case. As Kvanvig observes:

> Understanding why, when, where, and what are explicable in terms of understanding that something is the case. In each such case there is some truth that explains the special kind of understanding in question, and the person's relationship to that truth can be explicated in terms of understanding that something is the case. For example, understanding why something is the case requires understanding that a certain explanation is correct, and understanding what happened requires understanding that such-and-such happened [Kvanvig, 2003, p. 189-190]

This suggests that a "correct" explanation can be used to understand why something is the case. And Kvanvig's claim is very similar to De Regt's

[50]There is, of course, more on this point and De Regt's idea of the connection between explanation and scientific understanding. However, here I will limit my discussion to De Regt's notions that I consider as relevant to elucidate the linkage understanding-explanation in my approach.

claim that to have understanding of a phenomenon (UP) amounts to having an "appropriate" explanation of it. Both the authors require that the 'understanding of why something is the case' be based on an explanation which has been previously evaluated as "correct" or "appropriate". However, differently from Kvanvig, De Regt considers that the appropriateness of an explanation depends on the various virtues of the theory in question (visualizability, causality, unifying power, simplicity) and on the capacity (the skill) scientists have to use these virtues in constructing their explanations:

> In this pragmatic dimension two elements play a crucial role: whether scientists are able to use a theory for explaining a phenomenon depends both on their skills and on the virtues of the theory. More precisely, it depends on whether the right combination of scientists' skills and theoretical virtues is realized. Particular virtues of theories, e.g., visualizability or simplicity, may be valued by scientists because they facilitate the use of the theory in constructing models and predicting or explaining phenomena; in this sense they are pragmatic virtues. Nevertheless not all scientists value the same qualities: their preferences are related to their skills, acquired by training and experience, and to other contextual factors such as their background knowledge, metaphysical commitments, and the properties of already entrenched theories. [De Regt, 2004, p. 105].

Where does my approach to MEPP locate with respect to De Regt's and Kvanvig's claims? A clarification of my position in this context requires a step-by-step strategy. I will begin by considering how De Regt's sense of understanding (FU), (UT) and (UP) can be used in the context of my approach. Next, I will reconsider the Hénon-Heiles example to make clear how a particular sense of understanding is involved in this MEPP (and in MEPP in general). Finally, I will point to the general linkage explanation-understanding from the perspective of my approach and I will maintain that understanding why something is the case is, at least in the context of MEPP, just a gloss of "we have a genuine MEPP". On the other hand, a pragmatic sense of understanding does operate in MEPP at the level of intellectual tools.

I agree with De Regt (and Trout and Hempel) on the fact that (FU) is not relevant for the scientific understanding of a phenomenon and for scientific explanation (it is neither necessary nor sufficient for (UP)). And I assume that it is neither relevant for MEPP. I also agree with De Regt in considering that there is a pragmatic sense of understanding (UT) which has to do with some skills possessed by scientists and which operates in explanation. However, in the context of MEPP (and of my approach), I consider that this pragmatic understanding concerns the ability to reason qualitatively on a mathematical state of affairs (and not on a scientific theory) and connect the qualitative information obtained to an accepted body of background knowledge. Concerning the ability to 'connect the qualitative information to an accepted body of background knowledge' employed in pragmatic understanding, observe that the fact that the matter understood should be put in a wider context and related to a body of previously accepted beliefs is often regarded as a precondition for understanding. For instance, Kvanvig writes:

> Understanding requires the grasping of explanatory and other coherence-making relationships in a large and comprehensive body of information. One can know many unrelated pieces of information, but understanding is achieved only when informational items are pieced together by the subject in question. [Kvanvig, 2003, p. 192]

Hence, instead of (UT), henceforth I will adopt the following sense for pragmatic understanding:

(UM) Understanding a mathematical state of affairs = being able to reason qualitatively on a mathematical state of affairs by connecting this qualitative information to an accepted body of background knowledge (pragmatic understanding)

Finally, about the third sense of understanding which is used in explanation, namely (UP), I agree with De Regt in considering that the understanding of a phenomenon amounts to having an appropriate explanation of the phenomenon. However, I consider that an "appropriate" explanation of a phenomenon is an explanation in which: i) an ability to reason is used through a conceptual resource, and ii) this ability to reason permits to reason qualitatively on a mathematical state of affairs and connect this qualitative information to an accepted body of background knowledge (thus obtaining a pragmatic understanding (UM) of such a state of affairs). Let me elucidate these intuitions in the context of the Hénon-Heiles example[51].

In the context of the Hénon-Heiles example, I claimed that we are able to perceive the relations between the energy and the behaviour of the system through our ability to reason visually (intellectual tool). Of course, visual reasoning is necessary but not sufficient for explanation. To reason visually on the surface of section does permits to make qualitative considerations and a number of connections and inferences, but these must be considered as adequate from a scientific point of view (they must conform to our scientific background knowledge). The crucial step consists then in recognizing that these connections, for instance the relationship between the energy and the trajectories on the Poincaré section, are 'good' because they are conformal to our background scientific knowledge and our beliefs (in a specific context C). To this step there corresponds a pragmatic understanding (UM), which is necessary to the overall MEPP. Pragmatic understanding (UM) is therefore given, in line with De Regt's idea, by the use of a particular ability possessed by the scientist providing the explanation in a specific context[52]. This partic-

[51] I presented the example in section 7.1. Here I will use only some general considerations about the example, those which I consider as necessary to state my point about the linkage explanation-understanding.

[52] In passing, let me note that although there are various convergences between De Regt's approach and mine on this point, there are also some substancial differences. Perhaps the most evident concerns the fact that De Regt considers that pragmatic understanding concerns skills of scientists, where a skill is defined by him as the ability to construct deductive arguments from the available knowledge [De Regt, 2009, p. 588], while in my approach pragmatic understanding has to do with an ability to reason (which is not necessarily the ability to construct deductive arguments from the available knowledge). Furthermore, remember that in subsection 8.2.1 I have presented De Regt and Dieks' notion of 'intelligibility' of a theory. According to De Regt, the notion of intelligibility rephrase the notion of pragmatic understanding of a theory: If scientists have a pragmatic understanding of a theory, i.e. they are able to use that theory, the theory is intelligible to them [De Regt, 2009,

ular ability to reason cannot be captured objectively, but it is neither purely subjective nor relative because it is shared (and used) by a group of scientists in a specific context. Finally, to have an understanding of the phenomenon (UP) is, to use De Regt's terminology, to have an appropriate explanation. This explanation is appropriate (in a specific context) because it involves the use of an ability to reason and this ability incorporates the pragmatic understanding expressed by (UM).

Now, let me move to some more general considerations about the linkage explanation-understanding and my approach to MEPP. Remember that my approach is neither objective, because explanation is dependent on the subject, nor entirely subjective, because the use of intellectual tools is permitted by a particular virtue of some mathematical concepts, which is something external to the scientist performing the explanation. To say that we have a genuine MEPP when we have a pragmatic understanding (UM) is a very natural idea within my framework. Explanation has to do with our abilities to reason in science, and these abilities to reason are developed according to our background knowledge. It is then reasonable to expect that a phenomenon will be considered as genuinely explained in a specific context when the scientist will be able to use an ability to reason that he considers as natural in that context. The use of these abilities makes the scientist confident in the explanation. And to the use of these abilities, as I claimed in the previous paragraph, there corresponds a pragmatic understanding (UM) of the mathematics involved in the explanation. This pragmatic understanding is not given by a sequence of calculations, and there is 'something extra' in it which goes behind mathematical knowledge and which is necessary to explanation.

p. 593]. Intelligibility (and then pragmatic understanding) is obtained through a particular combination of skills and theoretical virtues. The theoretical virtues used in combination with the skills function as "conceptual tools" for achieving understanding and explanation. However, in subsection 8.2.1 I have pointed to the fact that there is a difference between De Regt and Dieks' conceptual tools and my intellectual tools. To put this difference in the present context, in my approach the pragmatic understanding is obtained through an intellectual tool, and not through a conceptual tool.

Of course, the negative condition on calculus is not the full story. Calculations could be necessary to obtain, at a precise step, a pragmatic understanding (UM). Furthermore, there are other conditions of adequacy that are required in order to apply an ability to reason and have a understanding (UM) and (UP). These conditions regard the relevant aspects of the physical phenomenon which are mirrored by the mathematical model. For instance, in the case of Hénon-Heiles, the energy of the particle is a relevant aspect of the phenomenon which has a corresponding term in the Hamiltonian formulation. When we apply our ability to reason visually on the Poincaré section, we are considering that such a mapping is established and that the mathematical formalism maps this (and others) relevant feature(s).

Finally, according to my intuitions, having a genuine MEPP amounts to an understanding of why a phenomenon occurs. And in this genuine MEPP a pragmatic understanding (UM) operates and it is conveyed through an intellectual tool. On the other hand, of course, I do not exclude that in science there are other kinds or senses of 'understanding why' which might be obtained without recurring to genuine MEPP[53]. Unfortunately, I do not have anything interesting to say on these other facets of understanding and on the difference between these and the two senses of understanding (UM) and (UP) that I consider as associated to MEPP.

Very curiously, by performing such a move in my analysis, namely by identifying a genuine MEPP with understanding, it seems that I am falling back to the origins of the debate on scientific explanation. In fact, at *prima facie*, my considerations well accord with a general observation that has been proposed in the context of scientific explanation and that appears in the opening lines of the famous Volume XIII of the Minnesota Studies in the Philosophy

[53]For instance, Peter Lipton considers various kinds of knowledge (causation, necessity, possibility and unification) as cognitive benefits given by an explanation, and focuses on those rather than on the explanation itself [Lipton, 2009]. He identifies each of those benefits with understanding and claims that each of them might be acquired "by routes that do *not* pass through explanation" [Lipton, 2009, p. 44]. This is why, for him, we can have understanding without explanation. While I consider Lipton's argument very thought-provoking, I will not explore this issue here.

of Science devoted to the topic of scientific explanation:

> The search for scientific knowledge extends far back into antiquity. At some point in that quest, at least by the time of Aristotle, philosophers recognized that a fundamental distinction should be drawn between two kinds of scientific knowledge –roughly, knowledge that and knowledge why. It is one thing to know that each planet periodically reverses the direction of its motion with respect to the background of fixed stars; it is quite a different matter to know why. Knowledge of the former type is descriptive; knowledge of the latter type is explanatory. It is explanatory knowledge that provides scientific understanding of our world. [Salmon, 1989, p. 3]

I accept that explanatory knowledge provides scientific understanding of our world. However, there is a profound difference between the conception of explanatory knowledge and scientific understanding that the author of the previous passage had in mind and the conception which results from my approach to MEPP. In my view, explanation has not a purely objective value, i.e. it is not independent of the subject doing the explanation, and the same holds for understanding (UP). To say that we have a genuine MEPP in a context X does not entail that in a different scientific context Y this explanation will be regarded as genuine. And this has a direct repercussion on the character of understanding (UP). To have understanding of why a phenomenon occurs is to have understanding in context X, and not to have understanding *simpliciter* of why that phenomenon occurs[54].

Although still vague at this stage, I suppose that the considerations put forward in this subsection can be accommodated further within the view on scientific understanding proposed by Henk De Regt and Dennis Dieks ([De Regt *et al.*, 2005] and [De Regt, 2009]). For instance, I regard their criteria CUP (criterion for understanding phenomena) and CIT (criterion

[54]This means, of course, that we can understand a phenomenon in a context X where our scientific beliefs are *not* correct (for instance, they are not empirically adequate). In that context an explanation will draw on such a corpus of beliefs, and the resulting understanding will be considered as genuine (in that context!).

for the intelligibility of theories) as good candidates to model, if properly adjusted in the context of MEPP, my considerations about the role of intellectual tools in the explanation and the understanding of a physical phenomenon. Furthermore, I think that a potential direction for future analysis would be to explore the use of intellectual tools in mathematics alone and see how my intuitions are supposed to work in the context of mathematical explanation within mathematics. However, I have not explored these routes here and I have opted for a more general discussion which is, again, only a sketch of the nature of understanding and of the role that such a concept can play in MEPP once my approach is endorsed.

8.6.2 Abilities to reason (β)

Question β concerns our abilities to reason and a possible trivialization of my approach. In particular, if in science we always use one o more abilities to reason, therefore according to my approach every scientific practice should be regarded as genuinely explanatory. In this subsection I am going to propose some speculations about a possible way to avoid this trivialization.

A first argument again this trivialization would be: the kind of abilities to reason that I am considering are employed in genuine MEPP *only* when there are some conceptual resources which permit their use. For instance, someone might claim that the ability to perform a very complicated calculus should be considered as an ability to reason. However, this ability does not correspond to an intellectual tool because it can be used without recurring to any conceptual resource. Examples from scientific practice would support the latter claim by showing that such an ability is not used through a conceptual resource. On the other hand, this defense seems to be too weak and even *ad hoc*. The problem is that I do not have an argument to support the idea that an ability to reason used without a conceptual resource does not carry explanatory power. Perhaps, a justification for this argument might be found by analysing further cases from scientific practice. Nevertheless, I do not have such a robust analysis at this step.

There is a second, but still uncomfortable, way to reject the accusation of trivialization. Consider the ability to reason visually, the ability to reason asymptotically and the ability to reason causally. Are these abilities cognitively primitive? Are they based on a cognitively primitive faculty of our mind? Perhaps my observations about the ability to reason causally, put forward at the end of subsection 8.5.1, can be applied to our ability to reason visually as well. Although I consider that our ability to reason visually is an ability we acquire, and therefore it is not cognitively primitive, our capacity to grasp a kind of knowledge that has a non-propositional complement by visualizing a particular state of affairs might be reasonably considered as a primitive faculty (here the parallel is with the faculty I called 'event discrimination', which I regarded as cognitively primitive and necessary to reason causally). For instance, it is reasonable to expect that a child will grasp some kind of knowledge from a pictorial proof of the Pythagorean theorem (for example, just by playing with the squares and learning that there exist some particular proportions between those squares), even if he will not be able to fully grasp the Pythagorean theorem. On the other hand, it is reasonable to expect that, when using his ability to reason visually on the pictorial proof of the Pythagorean theorem, a trained mathematician will use the same faculty used by the little child. As in the case of the ability to reason causally, then, there might be a primitive cognitive faculty which is involved and which is necessary (but not sufficient) to reason in a particular way. Concerning our ability to reason analogically, it can be thought that this ability results from our primitive faculty to recognize structural similarities between two states of affairs. Nevertheless, again, although there is some faculty which can be thought as primitive from a cognitive point of view, and which is essential to develop an ability to reason, our ability to reason analogically can be regarded as based on that particular faculty but requiring something more[55].

[55]A child might be able to recognize very simple structural similarities, but it might be unable to recognize similarities which have a more complicated character (for instance, functional similarities). As I have remarked in a footnote at the beginning of subsection 8.2.1, cognitive studies suggest that our ability to reason analogically is not a primitive cognitive ability but should be considered as emerging and developing under the guide of

To come to my point, although I favour the idea that our abilities to reason are acquired (through practice or education, for instance), I do not have good arguments to exclude that some particular abilities to reason can be cognitively primitive (and therefore not acquired), or that there are abilities to reason which are not based on some cognitively primitive faculty. Consequently, the questions "Are our abilities to reason used in science always acquired?" and "Are our abilities to reason used in science always based on some cognitively primitive faculty?" remain open questions. A possible characterization of our abilities to reason (and then of intellectual tools) in terms of the cognitive primitiveness of the faculties which are employed in them, or in terms of the cognitive primitiveness of the abilities themselves, would require an extra analysis which I am not able to propose here. However, I suppose that such an investigation, which should be carried out under the guidance of cognitive studies, might provide a possible answer to question β (and therefore a defense from the accusation of trivialization). For instance, once such a characterization would be available, it might result that only the abilities to reason which are based on some cognitively primitive faculty do act in genuine MEPP, and this would screen off these abilities from other abilities to reason used in science as well. Furthermore, it might be found that only the abilities to reason which are cognitively primitive do act in genuine MEPP, and therefore the abilities to reason which do act in science but which are not cognitively primitive could not act as intellectual tools[56]. In both cases the trivialization advanced in β would be avoided because it would not be true that *all* the abilities to reason used in science do contribute to genuine MEPP.

A final strategy of defense, which is the option I prefere and I consider to be more natural, is to recognize that there is a particular set of abilities to reason which acts in MEPP, and to a specific MEPP there correspond the use

certain basic constraints.

[56]Of course, negative examples might be proposed. For instance, examples in which two abilities to reason A and B are regarded as being cognitively primitive, or based on a faculty which is considered as cognitively primitive, but respectively are acting and not acting in genuine MEPP coming from the scientific practice.

of one or more of these abilities. Our task as explanation-scholars is to find the elements of such a set, and identify cases where these abilities are employed. Of course, these abilities are necessary but not sufficient for MEPP. As I have pointed out in the previous subsection, their use in genuine MEPP is extremely sensitive to contextual and pragmatic factors. For instance, a scientist will consider as genuine explanatory only the MEPP in which are used the abilities to reason belonging to his educational background, and these abilities permit to reach a result which is consistent with a previously accepted corpus of scientific knowledge. By focusing on some cases of MEPP (MEPP recognized as such in scientific practice), I have proposed the idea that these abilities act in MEPP when their use is made possible by one or more conceptual resource. I defined these abilities in action as intellectual tools. Although we do possess a number of abilities to reason, then, the abilities which do operate in MEPP would be a subset of our abilities to reason and would be preferred for pragmatic factors which may be attributed to the scientific context in which the scientist doing the explanation is operating. This would provide an answer to β and a reply to the accusation of trivialization.

8.6.3 Mutual interactions between conceptual resources and intellectual tools (γ)

At the end of section 8.2 I suggested that there is a mutual interaction between intellectual tools and conceptual resources: we acquire conceptual resources through the use of intellectual tools and, conversely, we acquire intellectual tools through the use of conceptual resources. As a consequence, intellectual tools, as conceptual resources, are not unchangeable but can vary over time[57]. Now I want to suggest how this claim can be justified, thus an-

[57]Let me note that this does not undermine what I have suggested in the final lines of the previous subsection, namely that the abilities which do operate in MEPP would be a subset of our abilities to reason. In fact, there might be the possibility that, in the temporal transition, our abilities to reason in science will change but the abilities which are used in MEPP will continue to be a subset of these abilities. In this subset there might be

swering question γ.

We use our abilities to reason and our mathematical concepts to explain phenomena, but as science changes the use of those concepts and abilities becomes more (or less) preferable depending on the context. In claiming that intellectual tools are not immutable over time I want to say that as our scientific knowledge increases we develop new theoretical concepts, and the repeated use of these concepts can affect our modes of reasoning in science, thus providing us with new abilities to reason. For instance, an example could be provided by the development of the asymptotic techniques analyzed by Batterman (such as that involving the renormalization group theory)[58]. In that case, there will be some initial stage in which the mathematical concepts that come with these techniques will permit the scientist to reason in a particular way, but this particular way of reasoning will not be accepted (or will be accepted only partially) by the scientific community. However, optimistically, the repeated use of these mathematical concepts will finally lead to accept the ability to reason asymptotically as a natural epistemic tool to be used in our 'explanatory' scientific practice. In other words, a conceptual resource will have permitted the introduction of an intellectual tool. On the other hand, to use an intellectual tool might provoke the introduction of new theoretical concepts, which may act as conceptual resources. Very roughly, I am thinking about Euclidean geometry, where the ability to reason visually on a diagram was a method to discover new geometrical properties and

---

elements which are fixed, and which have been left untouched by the temporal transition, but new elements (new abilities to reason used as intellectual tools) as well. Moreover, there is also the possibility that in the temporal transition some abilities to reason which were used as intellectual tools have disappeared (or better, have been discarded) from this subset. Roughly, this means that an ability to reason which was used as an intellectual tool in a context C at time t, is not used as intellectual tool in the same context at time t_1 (where $t < t_1$). For instance, the ability to reason asymptotically might not be used to explain phenomena in a future context C.

[58] Again, asymptotic techniques might provide just one possible way of analysing the situation. This means that in a different context scientist might have developed a different mathematical formalism in order to study the same physical phenomenon. And to this conceptual apparatus there might correspond an *accepted* way of reasoning that is very different from asymptotic reasoning. The teaching we learn is that contextual factors do play a very important role in our explanatory practice.

led to the introduction of new mathematical concepts. For instance, Marcus Giaquinto offers an example where, using a diagram, we are intuitively led to the discovery of the Pythagoras' theorem [Giaquinto, 2008, p. 32-33]. In that case, it is reasonable to think that the discovery of the Pythagoras' theorem will introduce new mathematical concepts, and those concepts might act as conceptual resources. More precisely, if the ability to reason visually functions as an intellectual tool (i.e. the practice in which the ability to reason visually is involved is considered as explanatory), the new (potential) conceptual resources will have been introduced by such intellectual tool.

Thus the interactions between intellectual tools and conceptual resources are reciprocal, and these influences are such that intellectual tools, as conceptual resources, are not unchangeable but can vary over time.

There is a remark which must be added here. In introducing this last part I have indicated, as an additional payoff of my approach, the fact that my framework sees as extremely favourable the intervention of history of mathematics (and history of science in general) as an instrument to investigate MEPP. This claim can be justified by observing that, to analyse the mutual interactions between conceptual resources and our abilities to reason, or even the different use of conceptual resources and intellectual tools in a mathematical explanation of a phenomenon, we are often demanded to use the history of science as instrument. From the history of mathematics we can learn how MEPP change and what are the differences between two MEPP belonging to very different mathematical contexts. More precisely, we learn how conceptual resources and intellectual tools are used in genuine MEPP (MEPP recognized as such in scientific practice), and how they change or develop through history. The importance of such a historical perspective is clear, for instance, if we consider the illustration of conceptual resources and intellectual tools given in the case of Euler's theorem, in subsection 8.4.4. In order to give a quick analysis of how the use of conceptual resources and intellectual tools changed from Euler's original explanation to the modern one, I had to turn to Euler's original proof (together with a look at his mathematical and scientific practice). To observe that some geometrical conceptual resources are now embedded in the framework of modern algebra required, of course, historical considerations. The historical component was also essential to stress the point that MEPP depend on the context in which they are produced. This is why I claimed that my approach to MEPP can benefit from the intervention of the history of science and the history of mathematics. This idea excellently reflects and even interprets the desire of various philosophers who demanded for a strict continuity between the history and the philosophy of mathematics.

History of mathematics can give us a better comprehension of MEPP, and it should be considered as an important instrument to study conceptual resources and intellectual tools (as someone pointed out in 1976, philosophy of mathematics without history of mathematics is in danger of becoming "empty"). I recognize, however, that to perform such a combined analysis can result in an extremely long and even intricated investigation.

8.7 Concluding remarks

In this last chapter I presented my approach to MEPP. I applied it to a case of MEPP and I suggested how the same schema can be used in the examples of MEPP proposed by the authors studied in part I and part II. Moreover, in my discussion I have proposed the idea that this schema can be generalized. As natural, however, the details of this generalization (and of an application to different cases of MEPP recognized as genuine) require further investigations and assessments.

Intellectual tools are (epistemic) utilities which are employed through our conceptual resources (and, implicitly, our background knowledge which is context-dependent and varies over time and across scientific communities). To use some mathematical concepts (rather than others) to explain the world amounts to using a hammer (together with our ability to use it), rather than a table or our hands, to drive a nail.

Far from the Herculean task of giving any comprehensive story about explanation, the more modest purpose of this final part was to contribute to a debate which is only at its earliest stage and whose development could have strong repercussions on different areas of the philosophy of mathematics and of the general philosophy of science. Optimistically, the notions sketched in these last two chapters might learn from new case studies and approaches to MEPP. In particular, as I stressed before, I still have not offered a robust characterization of conceptual resources and intellectual tools, but only an intuitive one. Further work has to be done to offer an adequate definition of these concepts and their interplay, as further work is required to investigate the role of these concepts in debates of philosophy of science in which the notion of MEPP (and explanation in general) is regarded as playing a central role (for instance, in the debate about the acceptability or not of the EIA, or in that concerning the linkage explanation-understanding).

Again, I have not provided certain answers. Nevertheless I hope to have given some potential directions of analysis, and to have raised a couple of questions which will be useful for the investigations to come. After all, as Russell pointed out in his essay "Logical Atomism":

> [...] we shall be wise to build our philosophy upon science, because the risk of error in philosophy is pretty sure to be greater than in science. If we could hope for certainty in philosophy the matter would be otherwise, but so far as I can see such a hope would be chimerical [Russell, 1924, p. 339]

Bibliography

[Achinstein, 1983] Achinstein, P., *The Nature of Explanation*, New York: Oxford University Press.

[Amit, 1978] Amit, D. J., *Field Theory, the Renormalization Group, and Critical Phenomena*, Singapore: McGraw-Hill.

[Aristotle, BWA 1941] Aristotle, *The Basic Works of Aristotle*, ed. by R. McKeon, New York: Random House.

[Aristotle, CWA 1984] Aristotle, *The Complete Works of Aristotle*, ed. by J. Barnes, Princeton: Princeton University Press.

[Arnold, 1992] Arnold, V. I., *Ordinary Differential Equations*, Berlin: Springer Verlag.

[Artin, 1957] Artin, E., *Geometric Algebra*, New York: Wiley Interscience.

[Aspray *et al.*, 1988] Aspray, W., and Kitcher, P. (eds.), *History and Philosophy of Modern Mathematics*, Minneapolis: University of Minnesota Press.

[Aubin *et al.*, 2002] Aubin, D. and Dalmedico, A. D., "Writing the History of Dynamical Systems and Chaos: Longue Durée and Revolution, Disciplines and Cultures", *Historia Mathematica*, 29, p. 273-339.

[Avigad, 2008] Avigad, J., "Understanding Proofs", in Mancosu, P. (ed.), *The Philosophy of Mathematical Practice*, Oxford: Oxford University Press, p. 449-499.

[Baker, 2005] Baker, A., “Are there Genuine Mathematical Explanations of Physical Phenomena?”, *Mind*, 114, p. 223-238.

[Baker, 2009] Baker, A., “Mathematical Explanation in Science”, *British Journal for the Philosophy of Science*, 60, p. 611-633.

[Bangu, 2008] Bangu, S. I., “Inference to the Best Explanation and Mathematical Realism”, *Synthese*, 160, p. 13-20.

[Barnes, 1992] Barnes, E., “Explanatory Unification and the Problem of Asymmetry”, *Philosophy of Science*, 59, p. 558-571.

[Batterman, 2002a] Batterman, R., *The Devil in the Details*, Oxford: Oxford University Press.

[Batterman, 2002b] Batterman, R., “Asymptotics and the Role of Minimal Models”, *British Journal for the Philosophy of Science*, 53, p. 21-38.

[Batterman, 2005a] Batterman, R., “Critical Phenomena and Breaking Drops: Infinite Idealizations in Physics”, *Studies in History and Philosophy of Modern Physics*, 36(2), p. 225-244.

[Batterman, 2005b] Batterman, R., “Response to Belot’s ‘Whose Devil? Which Details?’”. *Philosophy of Science*, 72(1), p. 154-163.

[Batterman, 2010] Batterman, R., “On the Explanatory Role of Mathematics in Empirical Science”, *British Journal for the Philosophy of Science*, 61(1), p. 1-25.

[Bealer, 1982] Bealer, G., *Quality and Concept*, Oxford: Clarendon Press.

[Belnap *et al.*, 1976] Belnap, N. and Steel, T. B., *The Logic of Questions and Answers*, New Haven: Yale University Press.

[Belot, 2005] Belot, G., ‘Whose Devil? Which Details?’, *Philosophy of Science*, 72(1), p. 128-153.

[Benacerraf, 1965] Benacerraf, P., "What Numbers Could not Be", *The Philosophical Review*, 74, p. 47-73.

[Berry *et al.*, 1980] Berry, M. V. and Upstill, C, "Catastrophe Optics: Morphologies of Caustics and their Diffraction Patterns", in Wolf, E. (ed), *Progress in Optics*, Amsterdam: North Holland, vol. 18, p. 257-346.

[Berry, 1987] Berry, M. V., "Quantum Chaology", in *Proceedings of The Royal Society of London A: Mathematical and Physical Sciences*, 413, London: The Royal Society, p. 183-198.

[Berry, 2009] Berry, M. V., "Emergence and Asymptotics in Physics: How One Theory Can Live Inside Another", *Conference on Mathematical and Geometrical Explanations in Physics*, 11th-12th December 2009, University of Bristol.

[Boltzmann, 1964] Boltzmann, L., *Lectures on Gas Theory*, Berkeley: University of California Press.

[Bondy et al., 1976] Bondy, J. A. and Murty, U. S. R., *Graph Theory with Applications*, Amsterdam: North-Holland.

[Braithwaite, 1953] Braithwaite, R., *Scientific Explanation*, Cambridge: Cambridge University Press.

[Bridgman, 1927] Bridgman, P. D., *The Logic of Modern Physics*, New York: Macmillan.

[Bromberger, 1963] Bromberger, S., "A Theory about the Theory of Theory and about the Theory of Theories", in Baumrin, B. (ed.), *Philosophy of Science: The Delaware Seminar*, New York: Interscience, vol. II, p. 79-106.

[Bromberger, 1966] Bromberger, S., "Why-Questions", in Colodny, R. (ed.), *Mind and Cosmos*, Pittsburgh: University of Pittsburgh Press.

[Brumfiel, 1979] Brumfiel, G. W., *Partially Ordered Rings and Semi-Algebraic Geometry*, Cambridge: Cambridge University Press.

[Bueno *et al.*, 2002] Bueno, O., French, S. and Ladyman, J., "On Representing the Relationship between the Mathematical and the Empirical", *Philosophy of Science*, 69(3), p. 497-518.

[Bueno *et al.*, 2011] Bueno, O. and Colyvan, M., "An Inferential Conception of the Application of Mathematics", *Noûs*, 45(2), p. 345-374.

[Butchart, 2001] Butchart, S. J., *Evidence and Explanation in Mathematics*, Ph.D. Thesis, Monash University.

[Callen, 1985] Callen, H. B., *Thermodynamics and an Introduction to Thermostatistics*, 2nd edition, New York: Wiley and Sons.

[Cannon, 1984] Cannon, J. R., *The One-Dimensional Heat Equation*, Encyclopedia of Mathematics and its Applications (No. 23), New York: Cambridge University Press.

[Cao, 2004] Cao, T.U.,"Ontology and Scientific Explanation", in Cornwell, J. (ed.), *Explanations: Styles of Explanation in Science*, Oxford: Oxford University Press, p. 173-196.

[Cartier, 2000] Cartier, P., "Mathemagics (A Tribute to L. Euler and R. Feynman)", in Planat, M. (ed.), *Noise, Oscillators and Algebraic Randomness: From Noise in Communication Systems to Number Theory*, Dordrecht: Springer, p. 6-67.

[Cartwright, 1983] Cartwright, N., *How the Laws of Physics Lie*, Oxford: Clarendon Press.

[Cartwright, 1989] Cartwright, N., *Nature's Capacities and their Measurement*, Oxford: Oxford University Press.

[Cartwright, 1999] Cartwright, N., "Models and the Limits of Theory: Quantum Hamiltonians and the BCS Model of Superconductivity", in Morgan,

M. S. and Morrison, M. (eds), *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press, p. 241-281.

[Cellucci, 2008] Cellucci, C., "The Nature of Mathematical Explanation", *Studies in History and Philosophy of Science*, 39, p. 202-210.

[Chuang, 2001] Chuang, L., "Infinite Systems in SM Explanations: Thermodynamic Limit, Renormalization (Semi-) Groups, and Irreversibility", *Philosophy of Science*, 68(3), p. 325-344.

[Colyvan, 2001] Colyvan, M., *The Indispensability of Mathematics*, Oxford: Oxford University Press.

[Colyvan, 2002] Colyvan, M., "Mathematics and Aesthetic Considerations in Science", *Mind*, 11, p. 69-78.

[Collins *et al.*, 2004] Collins, J., Hall, N., and Paul, L. A. (eds.), *Causation and Counterfactuals*, Cambridge: MIT Press.

[Cox *et al.*, 1998] Cox, R. T. and Carlton, C. E, "A Commentary on Prime Numbers and Life Cycles of Periodical Cicadas", *American Naturalist*, 152, p. 162-164.

[Da Costa *et al.*, 2003] Da Costa, N., and French, S., *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning*, Oxford: Oxford University Press.

[Daly *et al.*, 2009] Daly, C., and Langford, S., "Mathematical Explanation and Indispensability Arguments", *The Philosophical Quarterly*, 59(237), p. 641-658.

[Decock, 2002] Decock, L., "Quine's Weak and Strong Indispensability Argument", *Journal for General Philosophy of Science*, 33, p. 231-250.

[De Pierris, 2002] De Pierris, G., "Causation as a Philosophical Relation in Hume", *Philosophy and Phenomenological Research*, 64(3), p. 499-545.

[De Regt, 2004] De Regt, H. W., "Discussion note: Making sense of Understanding", *Philosophy of Science*, 71, p. 98-109.

[De Regt *et al.*, 2005] De Regt, H. W. and Dieks, D., "A Contextual Approach to Scientific Understanding", *Synthese*, 144, p. 137-170.

[De Regt, 2009] De Regt, H. W., "The Epistemic Value of Understanding", *Philosophy of Science*, 76, p. 585-597.

[Diestel, 2005] Diestel, R., *Graph Theory*, Series: Graduate Texts in Mathematics, vol. 173, 3rd edition, Heidelberg: Springer-Verlag.

[Dietl, 1966] Dietl, P., "Paresis and the Alleged Asymmetry between Explanation and Prediction", *British Journal for the Philosophy of Science*, 17, p. 313-318.

[Dorato *et al.*, 2011] Dorato, M. and Felline, L., "Scientific Explanation and Scientific Structuralism", in Bokulich, A. and Bokulich, P. (eds.), *Scientific Structuralism*, Boston Studies in Philosophy of Science, vol. 281, Dordrecht: Springer, p. 161-176.

[Dove, 1853] Dove, H.W., *Darstellung der Farbenlehre und Optische Studien*, Berlin.

[Dowe, 2008] Dowe, P., "Causal Processes", in Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy*, fall 2008 edition, http://plato.stanford.edu/archives/fall2008/entries/causation-process/.

[Dretske, 1972] Dretske, Fred., "Contrastive Statements", *The Philosophical Review*, 81(4), p. 411-437.

[Ducheyne, 2009] Ducheyne, S., "Understanding (in) Newton's Argument for Universal Gravitation", *Journal for General Philosophy of Science*, 40, p. 227-258.

[Dummett, 1993] Dummett, M., 'What is Mathematics About?', in Dummett, M., *The Seas of Language*, Oxford: Clarendon Press, p. 429-445.

[Eigner, 2009] Eigner, K., "Understanding in Psychology: Is Understanding a Surplus?", in De Regt, H. W., Leonelli, S., and Eigner, K. (eds.) *Scientific Understanding: Philosophical Perspectives* Pittsburgh: University of Pittsburgh Press, p. 271-297.

[Euler, 1736] Euler, L., *Solutio Problematis ad Geometriam Situs Pertinentis* (E53), in Blanc, C., Grigorijan, A.T., Habicht, W. *et al.* (eds.), *Leonhardi Euleri Opera Omnia*, series I, vol. 7, Basel: Birkhaüser, 1911-1986, p. 1-10. Originally published in *Commentarii Academiae Scientiarum Imperialis Petropolitanae* 8 (1736), 1741, p. 128-140.

[Euler, 1750] Euler, L., *Decouverte d'un Nouveau Principe de Mécanique* (E177), in Blanc, C., Grigorijan, A.T., Habicht, W. *et al.* (eds.), *Leonhardi Euleri Opera Omnia*, series 2, vol. 5, Basel: Birkhaüser, 1911-1986, p. 81-108. Originally published in *Mémoires de l'académie des sciences de Berlin*, 6, 1752, p. 185-217.

[Faye, 1999] Faye, J., "Explanation Explained", *Synthese*, 120, p. 61-75.

[Feigl, 1970] Feigl, H., "The 'Orthodox' View of Theories: Remarks in Defense as Well as Critique", in Radner, M. and Winokur, S. (eds.), *Theories and Methods of Physics and Psychology*, vol. 4 of Minnesota Studies in the Philosophy of Science, Minneapolis: University of Minnesota Press, p. 3-16.

[Feyerabend, 1962] Feyerabend, P. K., "Explanation, Reduction and Empiricism", in Feigl, H., and Maxwell, G. (eds.), *Scientific Explanation, Space and Time*, vol. 3 of Minnesota Studies in the Philosophy of Science, Minneapolis: University of Minnesota Press, p. 28-97.

[Feyerabend, 1965] Feyerabend, P. K., "On the Meaning of Scientific Terms", *The Journal of Philosophy*, 62(10), p. 266-274.

[Feyerabend, 1981] Feyerabend, P. K., "Explanation, Reduction and Empiricism", in Feyerabend, P. K. (ed.) *Realism, Rationalism, and Scientific*

Method: Philosophical Papers, vol. 1, Cambridge: Cambridge University Press.

[Field, 1980] Field, H., *Science Without Numbers: A Defence of Nominalism*, Oxford: Blackwell.

[Field, 1989] Field, H., *Realism, Mathematics, and Modality*, Oxford: Blackwell.

[Frege, 1980] Frege, G., *The Foundations of Arithmetic*, 2nd ed. translated by J. L. Austin, Evanston: Northwestern University Press.

[French, 2010] French, S., "Disentangling Mathematical and Physical Explanation: Symmetry, Spin and the 'Hybrid Nature' of Physical Quantities", unpublished typescript.

[Friedman, 1974] Friedman, M., "Explanation and Scientific Understanding", *The Journal of Philosophy*, 71, p. 5-19.

[Friedman, 1983] Friedman, M., *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science*, Princeton: Princeton University Press.

[Gentner *et al.*, 2001] Gentner, D., Holyoak, K. J. and Kokinov, B. K. (eds.), *The Analogical Mind. Perspectives from Cognitive Science*, Cambridge: MIT Press.

[Giaquinto, 2005] Giaquinto, M., "From Symmetry Perception to Basic Geometry", in Mancosu, P., Jorgensen, K., and Pedersen, S., (eds.), *Visualization, Explanation and Reasoning Styles in Mathematics*, Dordrecht: Springer, p. 31-55.

[Giaquinto, 2008] Giaquinto, M., "Visualizing in Mathematics", in Mancosu, P. (ed.), *The Philosophy of Mathematical Practice*, Oxford: Oxford University Press, p. 22-42.

[Gingras, 2001] Gingras, Y., "What did Mathematics do to Physics?", *History of Science*, 39, p. 383-416.

[Goldenfeld, 1992] Goldenfeld, N., *Lectures on Phase Transitions and the Renormalization Group*, Frontiers in Physics, vol. 85, Reading MA: Addison-Wesley.

[Goldstein, 1957] Goldstein, H., *Classical Mechanics*, Reading MA: Addison Wesley.

[Goldstein *et al.*, 2001] Goldstein, H., Poole, C. and Safko, J., *Classical Mechanics*, 3rd edition, New York: Addison Wesley.

[Goles *et al.*, 2001] Goles, E., Schulz, O. and Marcus, M., "Prime Number Selection of Cycles in a Predator-Prey Model", *Complexity*, p. 33-38.

[Goodman *et al.*, 1947] Goodman, N. and Quine, W. V., "Steps Toward a Constructive Nominalism", *Journal of Simbolic Logic*, 12, p. 105-122.

[Greiner *et al.*, 1995] Greiner, W., Neise, L., and Stöcker, H., *Thermodynamics and Statistical Mechanics*, New York: Springer Verlag.

[Gross *et al.*, 2004] Gross, J. L. and Yellen, J. (eds.), *Handbook of Graph Theory*, New York: CRC Press.

[Grove *et al.*, 1985] Grove, L.C., and Benson, C.T., *Finite Reflection Groups*, 2nd edition, New York: Springer.

[Grove, 2002] Grove, L.C., *Classical Groups and Geometric Algebra*, Providence: American Mathematical Society.

[Hafner *et al.*, 2005] Hafner, J. and Mancosu, P., "The Varieties of Mathematical Explanation", in Mancosu, P., Jorgensen, K., and Pedersen, S. (eds.), *Visualization, Explanation and Reasoning Styles in Mathematics*, Dordrecht: Springer, p. 215-250.

[Hafner *et al.*, 2008] Hafner, J. and Mancosu, P., "Beyond Unification", in Mancosu, P. (ed.), *The Philosophy of Mathematical Practice*, Oxford: Oxford University Press, p. 151-179.

[Hales, 2001] Hales, T., 2001, "The Honeycomb Conjecture", *Discrete and Computational Geometry*, 25, p. 1-22.

[Hanson, 1963] Hanson, R. N., *The Concept of The Positron*, New York: Cambridge.

[Hart, 1977] Hart, W. H., "Review of *Mathematical Knowledge* by Mark Steiner", *Journal of Philosophy*, 74, p. 118-29.

[Hellman, 1983] Hellman, G., "Realist Principles", *Philosophy of Science*, 50, p. 227-249.

[Hempel *et al.*, 1948] Hempel, C. and Oppenheim, P., "Studies in the Logic of Explanation" *Philosophy of Science*, 15(2), p. 135-175.

[Hempel, 1965] Hempel, C. G., *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, New York: Free Press.

[Hempel, 1966] Hempel, C. G., *Philosophy of Natural Science*, Englewood Cliffs N.J.: Prentice-Hall.

[Hénon *et al.*, 1964] Hénon, M. and Heiles, C. "The Applicability of the Third Integral of Motion: Some Numerical Experiments", *Astronomical Journal*, 69, p. 73-79.

[Hesse, 1966] Hesse, M. B., *Models and Analogies in Science*, Notre Dame IN: University of Notre Dame Press.

[Hezemans *et al.*, 1991] Hezemans, P., and van Geffen, L., "Analogy Theory for a Systems Approach to Physical and Technical Systems", in Fishwick P. A. and Luker, P. A. (eds), *Qualitative Simulation Modelling and Analysis*, New York: Springer-Verlag, p. 170-216.

[Hintikka *et al.*, 1995] Hintikka, J. and Halonen, I., "Semantics and Pragmatics for Why-Questions", *Journal of Philosophy*, 92(12), p. 636-657.

[Hiskes, 1986] Hiskes, A. L., "Friedman on the Foundations of Space-Time Theories", *Erkenntnis*, 25(1), p. 111-126.

[Holyoak *et al.*, 1997] Holyoak, K. J. and Thagard, P., "The Analogical Mind", *American Psychologist*, 52, p. 35-44.

[Homer, 1964] Homer, W. I., *Seurat and the Science of Painting*, Cambridge, MA: MIT Press.

[Hopkins et al., 2004] Hopkins, B., and Wilson, R., "The Truth about Königsberg", *College Mathematics Journal*, 35(3), p. 198-207. Reprinted in Bradley, R. E. and Sandifer, C. E. (eds), *Leonhard Euler: Life Work and Legacy*, Amsterdam: Elservier, 2007, p. 409-420.

[Hughes, 1997] Hughes, R. I. G., "Models and Representation", *Philosophy of Science*, 64, p. 325-336.

[Hughes, 1999] Hughes, R. I. G., "The Ising model, Computer Simulation, and Universal Physics", in Morgan, M. S. and Morrison, M. (eds), *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press, p. 97-145.

[Hume, 1999] Hume, D., *An Enquiry Concerning Human Understanding*, edited by Beauchamp, T. L., Oxford: Oxford University Press.

[Humphreys, 1993] Humphreys, P., "Greater Unification Equals Greater Understanding?", *Analysis*, 53(3), p. 183-188.

[Jones, 1995] Jones, T., "How the Unification Theory of Explanation Escapes Asymmetry Problems", *Erkenntnis*, 43(2), p. 229-240.

[Jones, 1998] Jones, M. R., "Idealization and Abstraction: A Framework", in Jones, M. R. and Cartwright, N. (eds.), *Idealization XII: Correcting the*

Model; Idealization and Abstraction in the Sciences, Amsterdam: Rodopi, p. 173-217.

[Kim, 1994] Kim, J., "Explanatory Knowledge and Metaphysical Dependence", *Philosophical Issues*, 5, p. 51-69.

[Kitcher, 1975] Kitcher, P., "Bolzano's Ideal of Algebraic Analysis", *Studies in History and Philosophy of Science Part A*, 6(3), p. 229-269.

[Kitcher, 1976] Kitcher, P., "Explanation, Conjunction, and Unification", *The Journal of Philosophy*, 73(8), p. 207-212.

[Kitcher, 1978] Kitcher, P., "Theories, Theorists and Theoretical Change", *The Philosophical Review*, 87(4), p. 519-547.

[Kitcher, 1980] Kitcher, P. "A Priori Knowledge", *Philosophical Review*, 89, p. 3-23.

[Kitcher, 1981] Kitcher, P., "Explanatory Unification", *Philosophy of Science*, 48(4), p. 507-531.

[Kitcher, 1982] Kitcher, P., "Genes", *The British Journal for the Philosophy of Science*, 33(4), p. 337-359.

[Kitcher, 1984] Kitcher, P., *The Nature of Mathematical Knowledge*, Oxford: Oxford University Press.

[Kitcher, 1985a] Kitcher, P., "Darwin's Achievement", in Rescher, N. (ed) *Reason and Rationality in Science*, Washington: University Press of America, p. 127-189.

[Kitcher, 1985b] Kitcher, P., "Two Approaches to Explanation", *Journal of Philosophy*, 82, p. 632-639.

[Kitcher *et al.*, 1987] Kitcher, P., and Salmon, W., "Van Fraassen on Explanation", *The Journal of Philosophy*, 84(6), p. 315-330.

[Kitcher, 1988] Kitcher, P., “Mathematical Progress”, *Revue Internationale de Philosophie*, 42, p. 518-540.

[Kitcher *et al.*, 1989] Kitcher, P., and Salmon, W. (eds.), *Scientific Explanation*, vol. 13 of Minnesota Studies in the Philosophy of Science, Minneapolis: University of Minnesota Press.

[Kitcher, 1989] Kitcher, P., “Explanatory Unification and the Causal Structure of the World”, in Kitcher, P. and Salmon, W. (eds.), *Scientific Explanation*, vol. 13 of Minnesota Studies in the Philosophy of Science, Minneapolis: University of Minnesota Press, p. 410-505.

[Kitcher, 1993] Kitcher, P., *The Advancement of Science*, Oxford: Oxford University Press.

[Klein *et al.*, 1999] Klein, E. and Lachieze-Rey, M., *The Quest for Unity: The Adventure of Physics*, Oxford: Oxford University Press.

[Kleiner, 2007] Kleiner, I., *A History of Abstract Algebra*, Boston: Birkhauser.

[Kneale, 1949] Kneale, W. *Probability and Induction*, Oxford: Clarendon Press.

[Koetsier, 2007] Koetsier, T., “Euler and Kinematics”, in Bradley R. E. and Sandifer, C. E. (eds.), *Leonhard Euler: Life Work and Legacy*, Amsterdam: Elservier, p. 167-194.

[Korsch *et al.*, 2008] Korsch, H. J., Jodl, H.J. and Hartmann, T., *Chaos. A Program Collection for the PC*, Springer-Verlag: Berlin.

[Koukkari *et al.*, 2006] Koukkari, W. L. and Sothern, R.B., *Introducing Biological Rhythms*, New York: Springer.

[Krantz *et al.*, 1971] Krantz, D., Lute, D., Suppes, P. and Wersky, A., *Foundations of Measurement*, vol. I, New York: Academic Press.

[Kuhn, 1970] Kuhn, T. S. , *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

[Kvanvig, 2003] Kvanvig, J. L., *The Value of Knowledge And the Pursuit of Understanding*, Cambridge: Cambridge University Press.

[Ladyman, 2002] Ladyman, J., *Understanding Philosophy of Science*, London: Routledge.

[Lakatos, 1976] Lakatos, I., *Proofs and Refutations*, edited by Worrall, J. and Zahar, E., Cambridge: Cambridge University Press.

[Landau, 1937] Landau, L. D., *Phys. Z. Sowjetunion* 11, 26.

[Landau, 1958] Landau, E., *Elementary Number Theory*, New York: Chelsea Publishing Company.

[Leng, 2005] Leng, M., "Mathematical Explanation", in Cellucci, C. and Gillies, D. (eds), *Mathematical Reasoning, Heuristics and the Development of Mathematics*, London: King's College Publications, p. 167-189.

[Lewis, 1977] Lewis, A. L., "Lattice Renormalization Group and the Thermodynamic Limit", *Physical Review B*, 16(3), p. 1249-1252.

[Lewis, 1973] Lewis, D., "Causation", *The Journal of Philosophy*, 70, p. 556-567.

[Lewis, 1993] Lewis, D., "Causal Explanation", in Ruben, D. H. (ed.), *Explanation*, New York: Oxford University Press, p. 182-206. Originally published in Lewis, D., *Philosophical Papers 2*, New York: Oxford University Press, 1986, p. 214-240.

[Lipton, 2004] Lipton, P., "What Good is an Explanation?", in Cornwell, J. (ed.), *Explanations: Styles of Explanation in Science*, Oxford: Oxford University Press, p. 1-22.

[Lipton, 2009] Lipton, P., “Understanding Without Explanation”, in De Regt, H. W., Leonelli, S. and Eigner, K. (eds.), *Scientific Understanding: Philosophical Perspectives*, Pittsburgh: University of Pittsburgh Press, p. 43-63.

[Liu, 2001] Liu, C., “Infinite Systems in SM Explanations: Thermodynamic Limit, Renormalization (Semi-) Groups, and Irreversibility”, *Philosophy of Science*, 68(3), p. S325-S344.

[Lute *et al.*, 1990] Lute, D., Krantz, D., Suppes, P. and Wersky. A., *Foundations of Measurement*, vol. 3, New York: Academic Press.

[Lyon *et al.*, 2008] Lyon, A. and Colyvan, M., “The Explanatory Power of Phase Spaces”, *Philosophia Mathematica*, 16(2), 227-243.

[Mach, 1893] Mach, E., *The Science of Mechanics: A Critical and Historical Exposition of Its Principles*, translated by McCormack, T. J., Chicago: The Open Court Publishing Co.

[Maki *et al.*, 2009] Maki, U. and Marchionni, C., “On the Structure of Explanatory Unification: the Case of Geographical Economics”, *Studies in History and Philosophy of Science Part A*, 40(2), p. 185-195.

[Malament, 1982] Malament, D., “Review of Field’s Science Without Numbers”, *Journal of Philosophy*, 79, p. 523-534.

[Mancosu, 1996] Mancosu, P. (ed.), *Philosophy of Mathematics and Mathematical Practice in the Seventeenth Century*, Oxford: Oxford University Press.

[Mancosu, 1999] Mancosu, P., “Bolzano and Cournot on Mathematical Explanation”, *Revue d’Histoire des Sciences*, 52, p. 429-455.

[Mancosu, 2000] Mancosu, P., “On Mathematical Explanation”, in Grosholz, E. and Breger, H. (eds), *The Growth of Mathematical Knowledge*, Dordrecht: Kluwer, p. 103-109.

[Mancosu, 2001] Mancosu, P., "Mathematical Explanation: Problems and Prospects", *Topoi*, 20, p. 97-117.

[Mancosu *et al.*, 2005] Mancosu, P., Jorgensen, K., and Pedersen, S., (eds.), *Visualization, Explanation and Reasoning Styles in Mathematics*, Dordrecht: Springer.

[Mancosu, 2005] Mancosu, P., "Visualization in Logic and Mathematics", in Mancosu, P., Jorgensen, K. and Pedersen, S. (eds.), *Visualization, Explanation and Reasoning Styles in Mathematics*, Dordrecht: Springer, p. 13-30.

[Mancosu, 2008a] Mancosu, P. (ed), *The Philosophy of Mathematical Practice*, Oxford: Oxford University Press.

[Mancosu, 2008b] Mancosu, P., "Mathematical Explanation: Why it matters", in Mancosu, P. (ed.), *The Philosophy of Mathematical Practice*, Oxford: Oxford University Press, p. 134-150.

[Mancosu, 2008c] Mancosu, P., "Explanation in Mathematics", in Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy*, fall 2008 edition, http://plato.stanford.edu/archives/fall2008/entries/mathematics-explanation/.

[Massey, 1968] Massey, G. J., "Hempel's Criterion of Maximal Specificity", *Philosophical Studies*, 19, p. 43-47.

[Maxwell, 1857] Maxwell, *The Scientific Letters and Papers of James Clerk Maxwell*, edited by Harman, P. M., Cambridge: Cambridge University Press, 1990.

[Maxwell, 1865] Maxwell J.C., "A Dynamical Theory of the Electromagnetic Field", *Philosophical Transactions of the Royal Society*, 155, p. 459-512.

[May, 1979] May, R.M., "Periodical Cicadas", *Nature*, 277, p. 347-349.

[McGinn, 2004] McGinn, C., "What is It Not Like to Be a Brain?", in Cornwell, J. (ed.), *Explanations: Styles of Explanation in Science*, Oxford: Oxford University Press, p. 157-172.

[McMullin, 1978] McMullin, E., "Structural Explanation", *American Philosophical Quarterly*, 15, p. 139-148.

[McMullin, 1985] McMullin, E., "Galilean Idealization", *Studies in the History and Philosophy of Science*, 16, p. 247-273.

[Melia, 2000] Melia, J., "Weaseling Away the Indispensability Argument", *Mind*, 109(435), p. 455-479.

[Melia, 2002] Melia, J., "Response to Colyvan", *Mind*, 111, p. 75-79.

[Mischel, 1966] Mischel, T., "Pragmatic Aspects of Explanation", *Philosophy of Science*, 33, p. 40-60.

[Morgan *et al.*, 1999a] Morgan, M. S. and Morrison, M. (eds.), *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press.

[Morgan *et al.*, 1999b] Morgan, M. S. and Morrison, M., "Models as Mediating Instruments", in Morgan, M. S. and Morrison, M. (eds.), *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press, p. 10-37.

[Morrison, 1990] Morrison, M., "Unification, Realism and Inference", *British Journal for the Philosophy of Science*, 41, p. 305-332.

[Morrison, 1992] Morrison, M., "A Study in Theory Unification: The Case of Maxwell's Electromagnetic Theory", *Studies in History and Philosophy of Science Part A*, 23, p. 103-145.

[Morrison, 1998] Morrison, M., "Modelling Nature: Between Physics and the Physical World", *Philosophia Naturalis*, 35, p. 65-85.

[Morrison, 1999] Morrison, M., "Models as Autonomous Agents", in Morgan, M. S. and Morrison, M. (eds), *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press, p. 38-65.

[Morrison, 2000] Morrison, M., *Unifying Scientific Theories*, Cambridge: Cambridge University Press.

[Morrison, 2002] Morrison, M., "The One and the Many: the Search for Unity in a World of Diversity", *Studies in History and Philosophy of Modern Physics*, 33, p. 345-355.

[Nagel, 1961] Nagel, E., *The Structure of Science. Problems in the Logic of Scientific Explanation*, New York: Harcourt, Brace and World.

[Newton-Smith, 2000] Newton-Smith, W. H., "Explanation", in Newton-Smith, W. H. (ed.), *A Companion to the Philosophy of Science*, Oxford: Blackwell, p. 127-133.

[Nickles, 1973] "Two Concepts of Intertheoretic Reduction", *Journal of Philosophy*, 70(7), p. 181-201.

[Norton, 2007] Norton, J. D., "Causation as Folk Science", in Price, H. and Corry, R. (eds), *Causation, Physics, and the Constitution of Reality*, New York: Oxford University Press, p. 11-44.

[Oddie, 1982] Oddie, G., "Armstrong on the Eleatic Principle and Abstract Entities", *Philosophical Studies*, 41, p. 285-295.

[Panza, 1995] Panza, M., "From Nature that Economizes to Generous Forces: the Principle of Least Action between Mathematics and Metaphysics, Maupertuis and Euler, 1740-1751", *Revue d'Histoire des Sciences*, 48(4), p. 435-520.

[Panza, 2001] Panza, M., "A Proposito dell'Applicazione dei Metodi Matematici nelle Scienze Umane. Esempi Storici e Riflessioni Epistemologiche", *Lettera Pristem*, 42, p. 31-42.

[Panza, 2002] Panza, M., "Mathematization of the Science of Motion and the Birth of Analytical Mechanics: A Historiographical Note", in Cerrai, P., Freguglia, P. and Pellegrini, P. (eds.), *The Application of Mathematics to the Sciences of Nature. Critical moments and Aspects*, New York: Kluwer, p. 253-271.

[Panza *et al.*, 2010] Panza, M., and Sereni, A., *Il Problema di Platone*, Roma: Carocci Editore.

[Peano, 1898] Peano, G., "Analisi della Teoria dei Vettori", *Atti della Accademia delle Scienze di Torino. Classe di Scienze Fisiche, Matematiche e Naturali*, 31, p. 513-534.

[Pfeuty *et al.*, 1977] Pfeuty, P., and Toulouse, G., *Introduction to the Renormalization Group and to Critical Phenomena*, New York: Wiley and Sons.

[Pincock, 2004a] Pincock, C., "A Revealing Flaw in Colyvan's Indispensability Argument", *Philosophy of Science*, 71(1), p. 61-79.

[Pincock, 2004b] Pincock, C., "A New Perspective on the Problem of Applying Mathematics", *Philosophia Mathematica*, 12(2), p. 135-161.

[Pincock, 2007a] Pincock, C., "A Role for Mathematics in the Physical Sciences", *Noûs*, 41(2), p. 253-275.

[Pincock, 2007b] Pincock, C., "Mathematical Idealizations", *Philosophy of Science*, 74, p. 957-967.

[Pincock, 2011a] Pincock, C., "On Batterman's On the Explanatory Role of Mathematics in Empirical Science", *British Journal for the Philosophy of Science*, 62(1), p. 211-217.

[Pincock, 2011b] Pincock, C., "Modeling Reality", *Synthese*, 180, p. 19-32.

[Pincock, 2011c] Pincock, C., "Philosophy of Mathematics", in French, S. and Saatsi, J. (eds.), *The Continuum Companion to the Philosophy of Science*, New York: Continuum, p. 314-332.

[Pincock, 2011d] Pincock, C., "How to Avoid Inconsistent Idealizations", unpublished typescript, version May 1, 2011.

[Pitt, 1988] Pitt, J.C. (ed.), *Theories of Explanation*, New York: Oxford University Press.

[Polya, 1954] Polya, G., *Mathematics and Plausible Reasoning*, vol. I, Princeton: Princeton University Press.

[Polya, 1968] Polya, G., *Mathematics and Plausible Reasoning*, vol. II, Princeton: Princeton University Press.

[Portides, 2008] Portides, D., "Models", in Psillos, S. and Curd, M. (eds), *The Routledge Companion to Philosophy of Science*, New York: Routledge, p. 384-295.

[Presser, 1974] Presser, H. B., "Temporal Data Relating to the Human Menstrual Cycle", in Ferin, M., Halberg, F., Richert, R.M. and Vande Wiele, R. (eds.), *Biorhythms and Human Reproduction*, NewYork: Wiley, p. 145-160.

[Putnam, 1971] Putnam, H., *Philosophy of Logic*, New York: Harper and Row.

[Quine, 1951] Quine, W., "Two Dogmas of Empiricism", *The Philosophical Review*, 60(1), p. 20-43.

[Railton, 1981] Railton, P., "Probability, Explanation, and Information', *Synthese*, 48, p. 233-256.

[Railton, 1989] Railton, P., "Explanation and Metaphysical Controversy", in Kitcher, P. and Salmon, W. (eds.), *Scientific Explanation*, Minneapolis: University of Minnesota Press, p. 220-252.

[Redhead, 2004] Redhead, M., "Discussion Note: Asymptotic Reasoning", *Studies in History and Philosophy of Modern Physics*, 35, p. 527-530.

[Reichenbach, 1958] Reichenbach, H., *The Philosophy of Space and Time*, New York: Dover.

[Reichl, 1998] Reichl, L. E., *A Modern Course in Statistical Physics*, 2nd edition, New York: Wiley and Sons.

[Resnik *et al.*, 1987] Resnik, M., and Kushner, D., "Explanation, Independence, and Realism in Mathematics", *British Journal for the Philosophy of Science*, 38, p. 141-158.

[Resnik, 1997] Resnik, M., *Mathematics as a Science of Patterns*, Oxford: Oxford University Press.

[Richardson, 1995] Richardson, A., "Explanation: Pragmatics and Asymmetry", *Philosophical Studies*, 80, p. 109-129.

[Rood, 1881] Rood, O., *Students' Textbook of Color; Or, Modern Chromatics, with Applications to Art and Industry*, New York: D. Appleton and Company.

[Rouse Ball, 1892] Rouse Ball, W. W., *Mathematical Recreations and Problems of Past and Present Times* (later entitled *Mathematical Recreations and Essays*), London: Macmillan.

[Ruben, 1990] Ruben, D. H., *Explaining Explanation*, London: Routledge.

[Rudin, 1953] Rudin, W., *Principles of Mathematical Analysis*, New York: McGraw-Hill Book Co.

[Rueger, 2000] Rueger, A., "Robust Supervenience and Emergence", *Philosophy of Science*, 67(3), p. 466-489.

[Russell, 1913] Russell, B., "On the Notion of Cause", *Proceedings of the Aristotelian Society*, 13, p. 1-26.

[Russell, 1924] Russell, B., "Logical Atomism", in Marsh, R. C. (ed.), *Logic and Knowledge*, London: Allen and Unwin, 1956, p. 323-343.

[Saatsi, 2011] Saatsi, J., "The Enhanced Indispensability Argument: Representational versus Explanatory Role of Mathematics in Science", *British Journal for the Philosophy of Science*, 62(1), p. 143-154.

[Sabatés, 1994] Sabatés, H., "Problems for Kitcher's Account of Explanation", *Philosophical Issues*, 5, p. 273-282.

[Sachs *et al.*, 1988] Sachs, H., Stiebitz, M. and Wilson, R. J., "An Historical Note: Euler's Königsberg Letters", *Journal of Graph Theory*, 12(1), p. 133-139.

[Salmon, 1971] Salmon, W., "Statistical Explanation", in Salmon, W. (ed.), *Statistical Explanation and Statistical Relevance*, Pittsburgh: University of Pittsburgh Press, p. 29-87.

[Salmon, 1984a] Salmon, W., *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

[Salmon, 1984b] Salmon, W., "Scientific Explanation: Three Basic Conceptions", *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 2, p. 293-305.

[Salmon, 1989] Salmon, Wesley C., "Four Decades of Scientific Explanation", in Kitcher, P. and Salmon, W. (eds.), *Scientific Explanation*, vol. 13 of Minnesota Studies in the Philosophy of Science, Minneapolis: University of Minnesota Press, p. 3-219.

[Sandborg, 1998] Sandborg, D., "Mathematical Explanation and the Theory of Why-Questions", *British Journal for the Philosophy of Science*, 49, p. 603-624.

[Schurz, 1999] Schurz, G., "Explanation as Unification", *Synthese*, 120, p. 95-114.

[Scriven, 1959] Scriven, M.,"Explanation and Prediction in Evolutionary Theory". *Science*, 30, p. 477-482.

[Scriven, 1962] Scriven, M., "Explanation, Prediction and Laws", in Feigl, H. and Maxwell, G. (eds.), *Scientific Explanation, Space, and Time*, vol. 3 of Minnesota Studies in the Philosophy of Science, Minneapolis: University of Minnesota Press, p. 170-230.

[Seidenberg, 1954] Seidenberg, A., "A New Decision Method for Elementary Algebra", *Annals of Mathematics*, 60, p. 365-374.

[Sernesi, 1993] Sernesi, E., *Linear Algebra*, London: Chapman and Hall.

[Shai et al., 1999] Shai, O., and Preiss, K., "Graph Theory Representations of Engineering Systems and their Embedded Knowledge", *Artificial Intelligence in Engineering*, 13, p. 273-285.

[Shapiro, 1997] Shapiro, S., *Philosophy of Mathematics: Structure and Ontology*, Oxford: Oxford University Press.

[Shapiro, 2005] Shapiro, S. (ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*, Oxford: Oxford University Press.

[Shields, 2007] Shields, C., *Aristotle*, London: Routledge.

[Sintonen, 1999] Sintonen, M., "Why Questions, and Why Just Why-Questions?", *Synthese*, 120, p. 125-135.

[Steiner, 1975] Steiner, M., *Mathematical Knowledge*, Ithaca: Cornell University Press.

[Steiner, 1978a] Steiner, M., “Mathematical Explanation”, *Philosophical Studies*, 34, p. 135-151.

[Steiner, 1978b] Steiner, M., “Mathematics, Explanation and Scientific Knowledge”, *Noûs*, 12, p. 17-28.

[Steiner, 1983] Steiner, M., “Mathematical Realism” *Noûs*, 17, p. 363-385.

[Steiner, 1998] Steiner, M., *The Applicability of Mathematics as a Philosophical Problem*, Cambridge: Harvard University Press.

[Suárez, 1999] Suárez, M., “The Role of Models in the Application of Scientific Theories: Epistemological Implications” in Morgan, M. and Morrison, M. (eds.), *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press, p. 168-196.

[Suárez, 2003] Suárez, M., “Scientific Representation: Against Similarity and Isomorphism”, *International Studies in the Philosophy of Science*, 17, p. 225-244.

[Suppes *et al.*, 1989] Suppes, P., Krantz, D., Lute, D. and Wersky, A., *Foundations of Measurement*, vol. 2, New York: Academic Press.

[Sus et al., 2010] Sus, A., and Molinini, D., “Causality and Explanation in Physics, Biology and Economics”, *The Reasoner*, 4(4), p. 60-61.

[Taber, 2001] Taber, K. S., “When the Analogy Breaks Down: Modelling the Atom on the Solar System”, *Physics Education*, 36(3), p. 222-226.

[Tamás and Márton, 2006] Tamás, T., and Márton, G., *Chaotic Dynamics: An Introduction Based on Classical Mechanics*, Cambridge: Cambridge University Press.

[Tappenden, 2005] Tappenden, J., “Proof Style and Understanding in Mathematics I: Visualization, Unification and Axiom Choice”, in Mancosu, P., Jorgensen, K. and Pedersen, S. (eds), *Visualization, Explanation and Reasoning Styles in Mathematics*, Dordrecht: Springer, p. 147-214.

[Targ, 1987] Targ, S. M., *Elementi di Meccanica Teorica*, Edizioni MIR.

[Tarski, 1951] Tarski, A., *A Decision Method for Elementary Algebra and Geometry*, 2nd edition, Berkeley: University of California Press.

[Taylor, 2000] Taylor, P., "What Ever Happened to Those Bridges?", *Australian Mathematics Trust*, http://www.amt.canberra.edu.au/koenigs.html.

[Tong, 1994] Tong, J., "Kummer's Test Gives Characterizations for Convergence or Divergence of all Positive Series", *The American Mathematical Monthly*, 101(5), p. 450-452.

[Toulmin, 1963] Toulmin, S., *Foresight and Understanding*, New York: Harper and Row.

[Trout, 2002] Trout, J. D., "Scientific Explanation and the Sense of Understanding", *Philosophy of Science*, 69, p. 212-233.

[Tuomela, 1980] Tuomela, R., "Explaining Explaining", *Erkenntnis*, 15, p. 211-243.

[Tymoczko, 1998] Tymoczko, T. (ed.), *New Directions in the Philosophy of Mathematics*, Princeton: Princeton University Press.

[Urquhart, 2008a] Urquhart, A., "The Boundary Between Mathematics and Physics", in Mancosu, P. (ed.), *The Philosophy of Mathematical Practice*, Oxford: Oxford University Press, p. 573-587.

[Urquhart, 2008b] Urquhart , A., "Mathematics and Physics: Startegies of Assimilation", in Mancosu, P. (ed.), *The Philosophy of Mathematical Practice*, Oxford: Oxford University Press, p. 588-620.

[Van den Dries, 1988] Van den Dries, L., "Alfred Tarski's Elimination Theory for Real Closed Fields", *The Journal of Symbolic Logic*, 53, p. 7-19.

[Van Fraassen, 1977] Van Fraassen, B. C., "The Pragmatics of Explanation", *American Philosophical Quarterly*, 14, p. 143-150.

[Van Fraassen, 1980] Van Fraassen, B. C., *The Scientific Image*, Oxford: Oxford University Press.

[Van Fraassen, 1985] Van Fraassen, B. C., "Salmon on Explanation", *Journal of Philosophy*, 82, p. 639-651.

[Von Neumann, 1947] Von Neumann, J., "The Mathematician", in Heywood, R. B. (ed.), *The Works of the Mind*, Chicago: University of Chicago Press, p. 180-196.

[Weber, 1996] Weber, E., "Explaining, Understanding and Scientific Theories", *Erkenntnis* 44, p. 1-23.

[Weber *et al.*, 2002] Weber, E. and Verhoeven, L., "Explanatory Proofs in Mathematics", *Logique et Analyse*, p. 179-180.

[Weyl, 1973] Weyl, H., *The Classical Groups. Their Invariants and Representations*, 2nd edition, Princeton: Princeton University Press.

[Whittaker, 1904] Whittaker, E.T., *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Cambridge: Cambridge University Press.

[Wigner, 1960] , Wigner, E. P., "The Unreasonable Effectiveness of Mathematics in the Natural Sciences", *Communications on Pure and Applied Mathematics*, 13(1), p. 1-14.

[Williams, 1995] Williams, K. S., "The Ecology, Behavior and Evolution of Periodical Cicadas", *Annual Review of Entomology*, 40, p. 269-295.

[Wilson, 1971] Wilson, K. G., "Renormalization Group and Critical Phenomena. I. Renormalization Group and the Kadanoff Scaling Picture", *Physical. Review B*, 4(9), p. 3174-3183.

[Wilson, 1982] Wilson, K. G., *The Renormalization Group and Critical Phenomena*, Nobel lecture, 8 December 1982.

[Wilson, 1986] Wilson, R. J., "An Eulerian Trail through Königsberg", *Journal of Graph Theory*, 10(3), p. 265-275.

[Wilson, 1996] Wilson, R. J., *Introduction to Graph Theory*, 4th edition, New Jersey: Prentice Hall.

[Wittgenstein, 1975] Wittgenstein, L., *Lectures on the Foundations of Mathematics, Cambridge 1939*, edited by Diamond, C., Chicago: University of Chicago Press.

[Woodward, 2003] Woodward, J., *Making Things Happen: a Theory of Causal Explanation*, Oxford: Oxford University Press.

[Wright, 2000] Wright, C., "Cogency and Question-Begging: Some Reflections on McKinsey's Paradox, and Putnam's proof", *Philosophical Issues*, 10, p. 140-163.

[Wright, 2002] Wright, C., "Anti-Sceptics Simple and Subtle: Moore and McDowell", *Philosophy and Phenomenological Research*, 65, p. 330-348.

[Yoshimura, 1997] Yoshimura, J., "The Evolutionary Origins of Periodical Cicadas During Ice Ages", *American Naturalist*, 149, p. 112-124.