Faculty of Philosophy
University of Oxford



Explanation Through Functionalism

A Low-hierarchical Interpretation of Scientific Theories

Candidate 1031742

Word count: 28.741

Thesis submitted in partial fulfilment of the requirements for the degree of

BPhil in Philosophy

July 2020

Abstract

The notion of explanation has been widely used to determine and single out dependencies within the Philosophy of Science. In particular, it is often believed that symmetries have a certain priority over laws on the basis of their explanatory role [E.g. Wigner 1964] or that one of the criteria that are employed to discriminate between laws and accidental generalisations is that laws are supposed to "explain" their instances (E.g. Bird 1998). These stances can lead to different hierarchical proposals, sometimes merely methodological but often with strong metaphysical outcomes.

In this work, I am proposing a functional account of explanation, which refrains from a hierarchical commitment, although recognising the different roles that symmetries, laws and phenomena play within the physical theories. In particular, the form of functionalism I am endorsing is the one first introduced in Cassirer 1910 and Cassirer 1937. Addressing the fact that this form of functionalism has been considered a predecessor of structural realism (French 2014), I shall deal with contemporary issues in the Philosophy of Science (E.g. French and Saatsi 2018, Weatherall 2011) as well as recent versions of Kantianism (Massimi 2011a).

Contents

Introduction 1

1 Hierarchies and Explanations 5

1.1 Symmetry 5

1.1.1 The Notion of Symmetry 5

1.1.2 Dependencies and Hierarchies 7

1.1.3 Symmetries, Laws and Explanation 10

1.2 Scientific Explanation 12

1.2.1 The Structure of Explanation 13

1.3 Explanation as a Functional Structure 16

1.3.1 Beyond the DN Model 18

2 Functionalism 21

2.1 Functionalism 21

2.1.1 From Functions to Laws 22

2.2 A Kantian Approach to the Laws of Nature 25

2.2.1 Objects and Theories 28

2.2.2 Functions of Objects 31

2.2.3 Low-hierarchical Functional Structures 36

2.3 Functionalism and Structural Realism 40

2.3.1 Which Structure? 42

3 Functional Explanation 49

3.1 Towards a Functional Theory of Explanation 49

3.2 Dependencies and Hierarchies 55

3.2.1 Low-hierarchies and Functional Explanation 65

Conclusions 69

References 74

Introduction

In the spectrum of the possible meanings of the word "explanation", the scientific one is most frequently discussed due to its cross relevance in different areas of the Philosophy of Science. As we shall see, endorsing any account of explanation determines a whole conception of the role of laws, the connection between laws and phenomena, the structure of our theories and the nature of the scientific understanding. Quite often, however, the effort to detect and extract the distinctive features of scientific explanation turns out to be influenced by many other implicit and sometimes unconscious assumptions deriving from the wider, common meaning of explanation. Therefore, it is not rare to see *e.g.* psychological and behavioural remarks popping up in a discussion that should be of mere scientific methodology. The problem is twofold: we want to look for the characteristics of the scientific explanation distinguishing it from the wider notion; but on the other hand, philosophers have failed to agreed on the very notion of scientific explanation. For this reason, a preliminary task is to provide a satisfactory account of how explanation works in the context of scientific reasoning. It is only on these grounds that one can then evaluate the difference from the wider use of the term.

The aim of this thesis is to provide the groundwork for a theory of scientific explanation based on a particular view in the Philosophy of Science, namely functionalism. Even if I disambiguate the meaning of the term when I address the topic, it is worth spending a few words on it. The sort of functionalism I advocate is different from the usual understanding of this notion in the contemporary debate. The functionalist project, indeed, is a far-reaching philosophical project elaborated and taken forward by Ernst Cassirer in the first half of the 20th Century. Functionalism is part of a wider *ensemble* of philosophical proposals commonly known as neo-Kantianism. Even if Cassirer's works range across a variety of fields, from the pure History of Philosophy to Philosophical Anthropology, his studies in the Philosophy of Science and early Philosophy of Physics are of the

utmost importance, to the point that he engaged in discussions with the most renowned scholars in the contemporary epistemology, such as Mortiz Schlick and Hans Reichenbach. Unfortunately, however, Cassirer's scientific thought had no successors able to keep this neo-Kantian perspective in the frontline of the epistmological debate, and functionalism gradually faded away.

This work is an endeavour to prove how the theoretical framework elaborated by Cassirer can offer valuable instruments to address the contemporary debate about scientific explanation. Please note that Cassirer did not explicitly address the issue of explanation in his works, apart from some insights and remarks from time to time. This was mainly for historical reasons, since this topic was not considered as central as after Hempel and Oppenheim's 1948 work, a watershed in the whole debate about scientific explanation. Therefore, my study will be based on the functionalist perspective, which I will use to lay the groundwork for a functional account of explanation. There will be, therefore, an overhaul of Cassirer's thought in the light of an integration within the most recent debate.

In order to do so, I have divided my work into three main parts. The first part is a preparatory study of some topics that I will reconsider throughout the following chapters. One crucial issue is the debate about the notion of symmetry and the role that symmetries play in our physical theories (Chapter 1.1). It is quite common to find arguments of various nature concerning the priority of symmetries with respect to laws or dynamical properties of systems, also from the point of view of an explanatory priority. Here I shall introduce the issue of hierarchical and low-hierarchical[1] representation of physical theories and physics in general. Then, I shall move on to the problem of scientific explanation, which I shall mostly analyse considering the Deductive-Nomological model (Chapter 1.2). It is crucial to have a good understanding of the logic and problems of this model as it will be considered in many parts of this work. Here I also introduce the notion of functional explanation (Chapter 1.3).

The second part is entirely devoted to a discussion of functionalism. After a brief outline of the main components of this approach (Chapter 2.1), I shall elaborate on the theoretical premises of this

[1] Also, *non-hierarchical*. I shall use the two terms interchangeably, as I am proposing a depiction that is non-hierarchical from the structural point of view (*i.e.* on the basis of the constituents of our theories, such as principles, symmetries, laws, empirical observations) but admits that, in general, there may be other kind of categorisation based on different parameters.

stance within Kantian thought and provide a more detailed presentation of the philosophical details of functionalism (Chapter 2.2). As Cassirer does not give an entirely straightforward definition of functionalism, I propose three complementary ways of implementing this approach. From our perspecive it is quite natural, while studying functionalism, to compare it to structural realism. This comparison bears several interesting insights that are certainly worth mentioning in order to get a better sense of some outcomes of functionalism. Therefore, in Chapter 2.3 I shall analyse these two views and conclude the general exposition of functionalism.

The third part consists in a more profound analysis of some of the issues, raised in the previous one, concerning the functional theory of explanation. In Chapter 3.1 I shall address a few points related to the DN model and other theories of explanation (the pragmatist and the unificatory views) that may be compared and contrasted with functional explanation. These comparative analyses allow us to tackle the most fundamental features of functionalism, and therefore play a crucial role in my work. From that point on, in fact, I shall discuss some aspects of functionalism that I deem substantially important - in particular about its ontology - through some other pieces of the contemporary literature (Chapter 3.2). In the Conclusions I shall eventually wrap up some loose ends of the whole work.

There are several assumptions underlying the following discussions. Some of them will emerge in due course, when I introduce the relevant issues. However, there is at least one point that drives me to a methodological consideration I will also recall in the final part of this work. It is strikingly odd to notice how there are two opposite yet complementary methodological fallacies in some of the arguments concerning hierarchies and explanation. On the one hand, one can easily find proposals aimed at building hierarchical constructions[2] and then argue that this dependency relation also represents an explanatory dependency, with little or no discussion at all about how this conclusion can be drawn[3]. On the other hand there are accounts of explanation that do not involve at all a discussion of the ontological framework in which that explanation should be integrated. My

[2] This can be done both within a specific theory or as a more general discourse on the priority of some elements of physics with respect to others, *e.g.* the priority of some principles within a specific theory, or in general the higher ontological status of symmetries with respect to laws. In the following I will refer to both of these and disambiguate where necessary.

[3] Among the examples that I discuss take *e.g.* French and Saatsi 2018.

point, which though being a methodological point has substantial implications, is that a theory of explanation can only be built starting from a specific account of the theories, objects and terms we are dealing with. And, *vice versa*, it is not sufficient to establish some sort of priority relation between terms to derive that there is also an explanatory dependency. To avoid these risks, I shall articulate the following work consistently referring both to the side related to the conception of physical theories and to the field of the discussion on scientific explanation.

Chapter 1

Hierarchies and Explanations

1.1 Symmetry

1.1.1 The Notion of Symmetry

As I have highlighted in the Introduction, a preliminary discussion of symmetries lays the groundwork for my following discussion about the concepts of explanation. Before getting into the specific case of physics, let me try to put forward a provisional exposition of the concept of symmetry. What is the peculiar feature characterising both simple reflection symmetries (*e.g.* the axis of symmetry of a bilateral animal) and other analogous but more complex processes? The idea is that a symmetry is a transformation that leaves something unchanged. The nature of this "something", it goes without saying, is everything but unproblematic. We establish some relevant aspects and require that they are unchanged for the symmetry to hold. There is, therefore, a part of the object which is ignored and some other features that are considered relevant and preserved under the operation.

Let us now leave the general idea of a symmetry and try to define it more thoroughly getting to the specific case of physics. What if the object considered in our transformation is a system of laws? A good example to start from is the Galilean case of the ship experiment. If we are in a ship and perform any operation below deck like jumping or observing the motion of objects, we notice that these operations are in no way influenced by the moving of the ship, provided that the motion is uniform *i.e.* the vessel moves with constant velocity and without any fluctuation. It is easier to perform the experiment under deck because there one does not experience the influence of air, which does not move with the ship. This is an everyday observation that we can test in different

contexts, travelling on a car or in a train.

What does it mean from the physical point of view? In our provisional definition, we have said that a symmetry is a transformation that leaves some peculiar features of the system unchanged. In the example of the ship, we have a transformation from a state of rest to a state of uniform motion that preserves the laws of motion. Obviously, these laws are preserved also in the case of spatial and time shift, in addition to a change in the orientation of the system. These transformations are called *Galilean transformations* and can be formally described in the following way. Consider the state of the system as a point in spacetime identified by $(\mathbf{x}, t)$ where $\mathbf{x} \in \mathbb{R}^3$, $t \in \mathbb{R}$. A Galilean transformation is a transformation on the coordinates of this kind

$$(\mathbf{x}, t) \to (R\mathbf{x} + \mathbf{v}t + \mathbf{b}, t + s) \tag{1.1}$$

Here we have expressed all the relevant transformations. $R \in SO(3)$ is the rotation of the reference frame; $\mathbf{v}$ is the velocity vector representing the increase in velocity; $\mathbf{b}$ is a vector in $\mathbb{R}^3$ indicating the spatial displacement; $s \in \mathbb{R}$ is a shift in time.

Newton's laws of motion are Galilean invariant in the sense that they are preserved under Galilean transformations. This is clear when we think that the laws of motion are the same even in the case of rotations, space and time displacements, and uniform velocity boosts. These transformations form a *group*, and it is therefore correct to say that two physical systems are symmetry-related when they are invariant under a group of transformations, in this case the Galilean group.

What we have seen can be generalised in a definition of symmetry that we come across quite frequently[1]. Instead of just Newton's laws, consider a system of differential equations

$$\mathbf{F}(x, \mathbf{y}^{(n)}) = 0 \tag{1.2}$$

where the x are the independent variables and $\mathbf{y}$ the dependent variables, and $\mathbf{y}^{(n)}$ are the derivatives of $\mathbf{y}(x) = \{y_1(x), ..., y_n(x)\}$ up to order n. A common statement is that a symmetry is defined as a transformation of this system mapping solutions to solutions.

It is worth expanding on this a little further. Consider a space $\mathcal{S}_K$ of kinematically possible states

[1] *E.g.* Olver 1995, p. 177.

- all the possible mathematical states of a system, designated by defining its objects, *e.g.* a differentiable manifold and the geometrical fields of the theory. We can narrow it to a subspace $\mathcal{S}_D \subset \mathcal{S}_K$ of the *solutions* of the theory. This is done by picking out the dynamical possible trajectories of the system, which are specified by the relevant dynamical equations. This is why $\mathcal{S}_D$ is also called the *space of the solutions*.

Within this framework, we can reformulate the idea of "mapping solutions to solutions" in the following way. Consider a solution to the system (1.2) $y = f(x) \in \mathcal{S}_D$ and a bijective transformation T: $\mathcal{S}_K \to \mathcal{S}_K$. We define T as a *symmetry* if, for any solution $y = f(x), T(y) \in \mathcal{S}_D$ as well. This means that the space of solutions $\mathcal{S}_D$ is preserved.

Is this first exposition of the concept of symmetry satisfactory? Certainly not[2]. When it comes to looking for a unique definition reflecting all the specific practices in physics, it is quite comprehensible that this amounts to going through a labour of Heracles. We know, in general, what symmetries are in physics, so that it is clear what we are referring to while talking about a spacetime or a variational symmetry. Nevertheless, the definitions that could be put forward in order to make abstraction from specific instances of symmetries are subject to a complex interpretative process[3]. However, my purpose was that of giving a satisfactory grasp of what is the sense of a symmetry. We can extract a definition of symmetry as a transformation on a space (a generic mathematical space, the space of solutions) that preserves some features (physical quantities, other pieces of structure). This is enough for our purposes and allows us to look deeper in the first point we are interested in, namely the common idea of symmetries as a "guide", or as principles endowed with particular explanatory properties.

1.1.2 Dependencies and Hierarchies

Symmetries have always played a key role among the concepts of modern physics. But the interpretation of their relationship with laws has varied significantly. In classical mechanics, symmetries were regarded as properties of systems instantiated only on the basis of the laws[4]. The approach consisted of first discovering the law and then studying it in order to single out its symmetries. Wigner argues

[2] Consider, for instance, the "disaster" (Belot 2013, p. 322) resulting from the combination of the "solution-to-solution" principle with the assumption that solutions are symmetry-related if and only if they are phyisically related

[3] See *e.g.* the different kinds of definitions in Dasgupta 2016.

[4] Brading and Castellani 2007, p. 1347.

that the change in this approach has to be found at first in Einstein's work: «Einstein made the most extensive use of invariance principles to guess the correct form of a law of nature, in this case that of the gravitational law, by postulating that this law conforms with the invariance principles which he postulated»[5]. What Einstein did in his Special Relativity, in fact, was to first postulate a principle requiring that the laws are unchanged under transformations between inertial coordinate systems. The methodological shift is apparent. Einstein assumed the invariance principles and then used it to restrict the results and the form of laws.

More generally, symmetries have played an important heuristic and methodological role in various fields of modern physics. I would like to set aside for now specific instances from particle physics and spacetime theories, and rather take into account the nature of these methodological functions of symmetries. It is quite common, in fact, to say that symmetries are a "guide to reality"[6], and this is certainly true. Let us consider again the Galilean case. The fact that Newton's laws of motion are invariant under Galilean transformations allows us to draw conclusions on some of the objects that could be part of our ontology. For instance, absolute motion has to be excluded since we see that velocity can vary without modifying the form of the laws. Symmetries have a heuristic role in the sense that they allow us to exclude varying features *i.e.* features that should not be regarded as part of the structure that we had required to be unchanged. This kind of reasoning is used in several metaphysical accounts. Consider, for instance, Earman's justification of the principle that, given a spacetime theory T, «Any dynamical symmetry of T is a space-time symmetry of T»[7]. It is using an Occamist norm that we rule out superfluous spacetime structure in order to allow laws of motion to pick out the dynamical possible models in $\mathcal{S}_D$.

This peculiar methodological function that we can ascribe to symmetries has led to conceiving a hierarchical structure, where the criteria to construct this hierarchy can be of different kinds: based on explanation, modality, generality, and so on. The first explicit claim of this kind is Wigner's one. His idea is that, as well as laws are used to explain and foresee events, we should postulate that symmetries constitute a sort of higher category of principles (he uses the expression "superprinciples"[8])

[5] Wigner 1964, p. 997.
[6] Dasgupta characterises this function of invariance principles in terms of a "symmetry-to-reality inference" (Dasgupta 2016, p. 840).
[7] Earman 1989, p. 46.
[8] Wigner 1964, p. 996.

with the same function with respect to laws. This means that these invariant principles allow one to discover more relevant information and relationships between events, *i.e.* discover and understand laws. His words are quite clear: events are the "raw material"[9] for laws, likewise laws are the raw material for symmetry principles. The argument to substantiate this claim refers, as we have seen before, to how symmetries can be used to work with laws, organise them, and discover new ones. Starting from these considerations, Wigner comes to what he himself defines a *hierarchy*: «The progression from events to laws of nature, and from laws of nature to symmetry or invariance principles, is what I meant by the hierarchy of the world around us»[10]. Wigner, actually, does not believe that the route of relying on the connection with conservation laws gives symmetries a priority[11]. This is quite different from a kind of approach that one could have by invoking Noether's theorem. Even if from Wigner's writings it is difficult to have a clear idea of how his hierarchy would look like, apart from this rough three-stages picture, it is quite clear that he is using a criterion of generality and not of explanatory dependancy. Henceforth, when I use the term *hierarchical* as opposed to *low-hierarchical* or *non-hierarchical*, I will refer to this kind of proposals.

The peculiar role of symmetries calls into question their very relationship with laws as well as physical phenomena. It is crucial to make clear that not all the accounts proposing a priority of symmetries over laws construct this relationship in hierarchical terms. In the example that we have just seen, Einstein certainly uses invariance principles assuming them as a methodological instrument to put constraints on the form of the laws. However, this does not amount to saying that he deduces a sort of ontological or metaphysical priority of symmetries. One must make a leap to derive such a conclusion. The point that I am making is that any inference from a methodological or heuristical perspective to an ontological one has to be grounded in further assumptions. At this stage, it is worth briefly taking into account one case in which the priority of symmetries is called into question in different ways, in order to get to the core of the problem of the explanatory dependencies within theories[12].

[9] Wigner 1964, p. 995.
[10] Wigner 1967, p. 30.
[11] See, for instance, the whole discussion in Wigner 1954.
[12] I have already discussed in more detail the problem of the relation between Noether's theorems and conservation laws, as well as Lange's model in a previous BPhil essay and I do not want to bring the whole topic back up [reference omitted for marking anonymity reasons]. However, it is necessary to briefly illustrate the underlying problems to have a better understanding of the relation between laws and symmetries.

1.1.3 Symmetries, Laws and Explanation

Besides the numerous proposals that account for the special role of symmetries in terms of methodological and heuristic instruments[13], there is a particular case in which symmetries are sometimes considered to have priority over laws, namely in the interpretation of Noether's theorems. The argument centres on the fact that Noether's theorem claims that continuous symmetries of the Lagrangian function of a system correspond with conservation laws. The connection between symmetries and conservation laws is a well-known fact in classical mechanics, just think of the conservation of linear momentum under spatial displacements, but Noether's theorems provides a substantive generalisation for the Lagrangian formalism.

One can identify two opposite stances on this issue. On the one hand, works as Brown and Holland 2004 aim to deny that the correlation between symmetries and conservation laws account for a priority of the former over the latter. Or, *e.g.* Lange 2007, though acknowledging that this last point about Noether's theorem is correct[14], nevertheless builds a taxonomic structure based on invariance under counterfactuals that would establish a hierarchy in which symmetries have an explanatory priority over laws in the same sense as laws with respect to singular events. The problem is that this whole discussion suffers from several significant flaws deriving from an ill-defined notion of explanation[15]. Smith 2008 emphasises the ambiguities of Lange's plan and the fact that no explanatory dependency can be actually derived by his taxonomic structure[16].

There are, therefore, two points that seem quite confusing. The first is that the notion of explanation that should come into play in these discussions is rather obscure. The second concerns the hierarchies that one may be tempted to build starting from explanatory dependencies. Moreover, so far I have put forward a general exposition of the concept of symmetry and tried to give some examples of the roles that it can play in the physical theories but this discussion is still assuming some crucial terms that have not been explained yet. For instance, what does it mean when we say that there is an *explanatory dependency*? Moreover, how are we characterising the kind of explanation

[13] *E.g.* Ismael and Van Fraassen 2003 argues that symmetries of laws can reveal superfluous theoretical structure in theories' formalisms.

[14] Lange 2007, p. 464.

[15] Also Brown and Holland do not spend much words on the kind of explanatory dependency they are opposing. However, they correctly argue that «The very notion of explanation involved is misguided.» (Brown and Holland 2004, p. 1337).

[16] Smith 2008, p. 326.

that would connect symmetries and laws? There are three players: symmetries, laws and explanation. And we have seen these players acting in contexts that seem different from each other. On what basis is a hierarchy supposed to work? And - this is the question that I will answer negatively - is it possible to construct a proper hierarchy with stronger generalisations (mere universals, laws, symmetries) and different levels of lawfulness?

When we talk about laws, the main problem is that of distinguishing them from other kinds of universal statements that refer to particular cases and facts. The Humean problem - namely whether laws are only generalisations based on repeated observations of events - is the core of the issue. A discussion about the concept of law, however, would divert us from the aim of this work. In the following, I will assume laws in the most neutral terms without endorsing any specific account (Humean, Lewisian, etc.), *i.e.* as generalisations with a peculiar descriptive and predictive status *vis-à-vis* events and single instances. Still it is worth mentioning that, quite often, one of the criteria that are employed to discriminate between laws and accidental generalisations is that laws are supposed to "explain" their instances[17]. Is it correct to say that laws are characterised by this feature?

If the response were affirmative, that would amount to saying that explanation is that relationship that occurs between a law and the phenomena it governs. But it is easy to guess that such an outcome would be too hasty, due to the carelessness with whom the term "explanation" is chosen. Even if it is easy to find cases in which laws are said to explain phenomena, in fact, this is attributable to the ambiguity that the term "explanation" has when it is used in the common language. That a law explains a phenomenon just means that we can trace back that specific instance to the law, but nothing more. The act of tracing back does not capture the concept of explanation as we would like to. It is possible to imagine cases in which we know that certain phenomena are governed by a law (in the sense of a generalisation maintaining a predictive power within a deductive system), even if we would not say that the phenomenon is adequately explained, *e.g.* when there are wider interpretative concerns about the status of the law.

Explanation turns out to be a wider concept which cannot be exhausted in the mere connection between a law and the phenomena it covers. *A fortiori* it is then apparent that without a comprehensive understanding of how scientific explanation works it is also impossible to properly discuss

[17] Bird 1998, p. 28.

the case of symmetries, but more in general the connections between the elements of our theories. Starting from this background I will then address my proposal of functional explanation as a way to build explanatory relationships without a strict hierarchical form.

1.2 Scientific Explanation

This work is based on an assumption that should not be taken for granted, namely that it is possible to focus on explanation in the scientific context as something different from our common way of conceiving it in our everyday life. The point is that explanation has multiple meanings and is used in a variety of contexts. In the vast majority of our everyday conversations, we say that something is explained when we grasp its meaning and it becomes clear to us. In its common sense, therefore, explanation has a vague meaning with an intense psychological flavour. Right because within a scientific context we want to get rid of all the psychological assumptions, I do not think that it would be useful, or even possible, to look for a common meaning that could capture the entire spectrum of possible disambiguations of "explanation", from its everyday use to the scientific one[18]. The first requirement to bring this discussion forward, therefore, is to narrow down the possible meanings and limit ourselves to scientific explanation.

However, even focusing only on science, there are many assumptions of a psychological nature lurking in the background - even if this is usually denied. Let us take, for instance, the case of the "familiarity view", according to which a new observation or phenomenon is explained when it is possible to reduce it to a familiar one. The notion of familiarity should be specified better, but in any case sounds as if it involves subjective elements. To whom is a phenomenon familiar? Which other phenomena can play the role of the familiar ones we should relate the new one to? But, more importantly, is it true that scientific explanation always work in such a way? If so, as Micheal Friedman argues[19], we could not say *e.g.* that quantum phenomena are explained, since they are all but familiar. It is true that the interpretation of these phenomena is often controversial but this does not

[18] I shall discuss Van Fraassen's pragmatic theory of explanation in Chapter 3.1. In a few words: with my functional account I will endorse a context-dependent interpretation involving also pragmatic elements, but this does not lead to Van Fraassen's conclusion that «to call an explanation scientific, is to say nothing about its form or the sort of information adduced, but only that the explanation draws on science to get this information» (Van Fraassen 1980, p. 155). In the following I will thereby work in the direction of a distinctive scientific account of explanation.

[19] Friedman 1974, p. 10.

mean that, when a physicist makes any empirical observation comprehensible within the quantum theory, that observation is not explained.

I deem that scientific explanation should be kept as much as it is possible away from these kind of spurious interpretations and does not have anything to do with assumptions related to our wider conception of explanation. This does not mean, however, that we cannot take into account pragmatic elements. The point is that it is necessary to keep ourselves as close as possible to a model which we can hold to be a good condition for explanation to happen, and then ensure that this model is flexible enough to face the different situations that we find in the scientific enterprise.

The aim of the next sections is that of addressing the advantages and limits of the DN model and give an idea of the concept of functionalism that will be the object of Part 2. As we shall see in Part 3, my point is that the limits of the DN model and its counterintuitive conclusions can be overcome if we integrate it in a wider account of explanation as the functionalist one. In order to do so, first of all we need to evaluate to what extent a general model of scientific explanation such as the DN one is a reasonable *desideratum* and what kind of modification it would need.

1.2.1 The Structure of Explanation

The Deductive-Nomological Model was first put forward by Hempel and Oppenheim in 1948. The idea is that of extracting the essential features that are common to any scientific explanation. According to the model, an explanation is made up of two elements, the *explanans* and the *explanandum*, defined as follows[20].

- *Explanans*: the component including what is supposed to explain the *explanandum*. It is again divided into sentences of laws $L_1, ..., L_r$ and sentences of particular further conditions $C_1, ..., C_k$.

- *Explanandum*: the component including what is to be explained. It is a sentence which expresses those phenomena and laws we are aiming to understand.

It is important to mention that the parts of what is to be explained as well as what is explaining it,

[20] Please note that the peculiarity of the DN model is the way in which *explanans* and *explanandum* are defined, not this twofold division of explanation that instead dates back to Aristotle's *Posterior Analytics* (see also Hempel and Oppenheim 1948, p. 140n.)

as Hempel and Oppenheim themselves point out, are not the phenomenon and the laws in themselves, but their logical counterparts within the explanatory scheme. This is significant to stress that, as I shall illustrate below, the purpose of this model should only be that of capturing the essential components - a sort of skeleton - of the process of explanation.

The nomological constituent is expressed by the *explanans*. It is in fact required that we have some general laws within the set of sentences that we are using to get to the phenomena. Since the DN model requires an initial set of sentences including laws, it is clear that, as I have mentioned earlier, laws have to be established as such via other assumptions in order to be then employed in the model. In addition to the laws, we need to specify further conditions. In fact, since the model aims to give an account of the empirical phenomenon described in the *explanandum*, laws have to be supported by some other elements. Each time these conditions are coupled with the laws, the phenomenon takes place.

Here we get to the second part of the model. The deductive constituent is represented by the fact that, given the set of laws and conditions, a deductive procedure brings to the *explanandum*. The deductive one is a key part of the process, because it seeks to capture the very nature of the explanation process. Here, it is not important if the connection between *explanans* and *explanandum* is strictly causal or statistical - an option which is considered by Hempel and Oppenheim themselves to be in some cases even more adequate than the causal one[21]. The crucial fact is that the deductive component is the one that plays the major role when we draw attention to the role of explanation in the scientific reasoning. Explanation is not only a way to trace back phenomena to a theoretic and observational background able to account for them. The flip side, in fact, is that a model of explanation should also give reasons why we can also expect a certain phenomenon given specific nomic and operational premises, *i.e.* why we can make a prediction.

The DN Model is extremely tempting - Peter Lipton would say *seductive*[22] - if considered as a potential necessary condition for a model of scientific explanation. The concerns arise if we expect it to be a sufficient condition as well. Is it enough to have *explanans* and *explanandum* complying

[21] Hempel 1965, p. 140.

[22] «Virtually every intuitive conception of the nature of scientific understanding finds some echo in the Deductive-Nomological Model. It is perhaps this feature more than any other that account for the model's continuing influence even among those who reject it» (Lipton 1992, p. 694).

with the DN conditions to have a proper explanation? It could be worthwile to consider a counterexample. Peter Achinstein's instance[23] is particularly insightful.

Let us consider a case of a law which, coupled with its conditions, should allow us to explain a certain related phenomenon - or make a prediction. A law of this kind (L_1) would be that a person who takes a certain dose of arsenic dies within twenty-four hours. The relevant condition is that Jones has unfortunately eaten that dose of arsenic (C_1). The observation is that, as a matter of fact, after twenty-four hours Jones is dead. Are we allowed to infer that the arsenic dose has killed Jones? The answer seems to be negative, because Jones has been actually killed in a hit-and-run.

There are at least two remarks that can be done. The first one is that, in my view, the example misses the point due to insufficient details about the conditions within the *explanans*. Much depends on how these conditions are formulated. It is not enough to have just the law about the lethality of arsenic and a general condition by which Jones has taken the dose. In a scientific explanation we usually require that conditions are complete enough - both when we want to trace back a phenomenon to its causes and when we try to make a prediction on the basis of our true premises. A more accurate condition would then be "Jones has taken the lethal dose of arsenic" (C_1) together with "The acid has had the possibility of kicking in" (C_2). The DN model would be obviously and blatantly wrong if it did not put some restrictions over the adequacy of the conditions.

However, the point I would like to focus on is the lesson of this example. The intended one is that the DN model does not provide the sufficient conditions to capture explanation because, even if the *explanans* is composed of true premises, this is not enough to establish *a priori* if the premises can give a proper account of the empirical observation. Still, one could argue that there is another implicit lesson. By affirming that the DN model cannot account for the explanation of Jones' death amounts to saying that it is too loose. In fact, the DN model admits as a correct explanation the fact that Jones had eaten arsenic in the twenty-four hours before his death even if the reason of his death is different. Which means that the model cannot exclude some cases and has the tendency to recognise as explanation a relationship which is of another nature.

In my opinion, it is correct to say that the DN model is not a sufficient condition to claim that there is an explanation. The problem, though, is not in its looseness, but rather in its lack of flexi-

[23] Achinstein 1981, pp. 209-210.

bility. My next example, taken from thermodynamics, addresses this point and clarify what kind of versatility should be required by an account of scientific explanation. In the last part of this work I will then apply the functional account to the relationship between symmetries, laws and other constituents of scientific theory, in order to argue that it is preferable to use a functional connection rather than a hierarchical one.

1.3 Explanation as a Functional Structure

The case I am suggesting here is drawn from thermodynamics. While considering *Boltzmann entropy*, which is a function of the microstates of the system, we do not have a direct interpretation of ρ, unlike the case of *Gibbs entropy* where ρ is a probability distribution on the phase space of that system. One of possible options is that ρ represents the ignorance of an observer, and this stance is sometimes called *subjective entropy*. In particular, it is possible to find remarks about the fact that subjective entropy thus defined lacks explanatory power:

> Consider for example the phenomenon that by thermal contact, heat always flows from the hotter to the cooler body, not the other way round. The usual explanation of this phenomenon is that entropy decreases when heat flows to the hotter body, and the second law exclude that.[24]

Goldstein's argument, in a nutshell, is that if we assign entropy an epistemic status depending on a subject, we are led to ask whether the heat flow holds in the same fashion also in absence of that observer. However, the point we are interested in is about the terms involved in the explanation. According to the excerpt, it is quite apparent that the phenomenon is the heat flow from the hotter to the cooler body, while the second law of thermodynamics is the law that explains it. But this depiction is anything but unquestionable. Both Kelvin and Clausius' formulation of the second law concern cycles in which heat is transferred. These two logically equivalent statements are usually interpreted in terms of entropy. And it is also true that *e.g.* Kelvin's statement shows that a reduction in entropy is not feasible. But it does not stem from this that the approach toward equilibrium is governed by the second law. Brown and Uffink have even argued that the spontaneous approach

[24] Goldstein et al. 2019, p. 13.

to a unique equilibrium state should be considered to be a fundamental principle and called «the Minus First Law»[25]. Depending on the place we give to the fact that heat flow from the hotter to the cooler body, we have two different explanatory frameworks. Heat flow can be held to be only a phenomenon explained by the second law. Or conversely, it is a fundamental principle that lies much deeper than the increase in entropy.

The point I would like to emphasise is that, in this case, explanation specifically depends on the understanding that we have of the whole theory. Hempel's attempt to get rid of the «psychological or pragmatic aspects of explanation»[26], which would be those involved if we make explanation dependent on the wider notion of understanding, is comprehensible and to a certain extent embraceable. It is reasonable for a philosopher of science to focus on the «logical aspects»[27], *i.e.* structure underlying explanation. But the case I have considered is particularly telling that we cannot separate the relationship between the *explanans* and the *explanandum* from the theoretical context they are organised within.

A point that is crucial to stress is that I am not referring to a kind of psychological understanding of that context, as if the context were our mental representation of the relationship between laws and phenomena, and as understanding were only a specific way of grasping those connection by ourselves. When I refer to a theoretical context, I am considering the way laws are functionally organised. It could be argued that this organisation depends on and follows from our way of conceiving theories, and this is undoubtedly correct. But the fact that we organise theories on the basis of our needs does not affect in any way my argument. The point is that we can say that something explains something else only referring to the logical structure that we give to the components of our theories (laws, phenomena, connections between them, priorities of laws over other laws). This is why we can say that the second law of thermodynamics explains the heat flow only by posing it as fundamental in the context of the behaviour of heat. But the same does not holds if the approach toward equilibrium is counted among the fundamental principles of the same theory.

One could make a last objection and say that I have picked a case that may be a double-edged sword, since the status of the second law of thermodynamics is not entirely straightforward and still

[25] Brown and Uffink 2001, p. 529.
[26] Hempel 1965, p. 413.
[27] Hempel 1965, *Loc. cit.*

amenable to multiple interpretation. Therefore, I should have chosen an unambiguous instance of explanation. But I deem that my last remark could reply to this concern. Regardless of whether the approach toward equilibrium is held to be among the posits or the descriptive and phenomenological components of the theories, the notion of explanation always depends on the organisation of the theory and not on an external logical rule which can be just applied independently of the context (as the defenders of the DN model would claim). This example taken from thermodynamics is not an apodictic argument but an example of how we can consider theories on the basis of their interrelated components without establishing a strict hierarchy between them. We will see, in Chapter 3.2, how this insight can be represented through a representation of theories as spheres of connected elements.

1.3.1 Beyond the DN Model

It is quite correct to argue, as Woodward does[28], that the particular status that Hempel gives to laws and phenomena in his account is due to the empiricist tradition. The DN model is lean thanks to the essential role that laws play in the explanation, where laws are considered as regularities in empiricist terms without a significant further discussion. This kind of account has deep roots in a distinguished philosophical tradition. Hempel himself, for instance, recalls this quote by Mill: «An individual fact is said to be explained by pointing out its cause, that is, by stating the law or laws of causation of which its production is an instance»[29]. The fact that Hempel endorses, or at least recognises a debt to this tradition, is telling. From Mill's words it is easy to understand what are the original problems of such a position. The idea that explanation only consists in the fact that we exhibit a connection between an event and a law, is rather poor and inadequate for our purposes. It is the same insufficiency that we also notice in the DN model in its minimal exposition. Laws govern singular phenomena, but do not explain them. Explanation needs a much more subtle interpretation.

However, before turning our attention back to the proposal outlined with the entropy example, I would like to advocate the simplicity of the DN model. So far, we have seen that it cannot be held to play the role of universal scheme of scientific explanation. It just does not capture the distinctiveness

[28] Woodward 2014, Ch. 1.
[29] Hempel and Oppenheim 1948, p. 140n.

and some subtleties of the notion of explanation, and therefore it is not the sufficient condition that someone would like to look for. But why should we look for a sufficient condition? It would be excellent for a philosopher to find a scheme into which every sort of scientific explanation could fit. However, I think that if we turned these *desiderata* into less ambitious ones, we would not lose much and could still get to a satisfactory account of explanation. As I have noticed, the DN model provides a minimal scheme which can be filled with a wide variety of explanatory processes. The main critique one could raise is that it is too simple, which means two things. The first one is that its looseness (as in the case of Achinstein's objection) induces us to accept as explanations also incorrect kinds of reasoning. The second one is that it is too generic and does not manage to capture some unique features of explanation, which is much more complex as we have seen in the entropy case. But what if we kept the DN model as a mere descriptive scheme, rather than a necessary and sufficient condition for explanation to occur?

Bird, quite correctly, writes that the problem is not that of putting forward another model[30]. The strictness of models is the key of the difficulties in capturing explanation. Hempel's work for a description of explanation which is universal and does not rely on «psychological and pragmatic aspects» has an understandable aim. But such a description is feasible only if we do not expect it to account for the whole explanatory process. The DN model can be used also in the thermodynamical case considered above. One must only change the law in the *explanans* depending on the choices. But one must also consider that the particular application of the DN model is just a part of the whole interpretation of the theory. Therefore, it is wrong to argue that, in thermodynamics, "the second law explains the tendency toward equilibrium" as an absolute claim where the second law is the *explanans* and the heat flow the *explanandum*. But it is correct to argue that, given a certain interpretation of the status of the second law, this law plays the role of explaining some phenomena. Explanation, far from being a subjective notion, is a way of describing in a functional fashion the connection between the elements of an interpreted given theory. This certainly amounts to introducing contextual (one may be tempted to say "pragmatic") considerations. However, we shall see in Chapter 3.1 that this is quite different from pragmatic accounts of explanation.

[30] «Finding an answer to the problems of explanation does not lie in constructing alternative models of explanation. [...] The idea of a model of explanation is this: if something fits this model then it is an explanation. It is a template against which we can judge something to see whether it should count as an explanation» (Bird 1998, p. 53).

Why should we shift to a functional account? The example of thermodynamics shows that it is quite difficult to build a straightforward picture of a theory. In my opinion, it is not correct to look for a scheme of an explanation hierarchy (a structure where meta-laws explain other laws, laws explain phenomena). A functional conception of the role of laws, for instance, allows us to get rid of hierarchies and just use on a case-by-case basis our minimal explanatory schemes (the DN model or anything analogous) to interpret specific connections of the elements of a theory. However, before assessing more in details what kind of non-hierarchical relation we can determine between symmetries, laws and events, a primary task lies in a further characterisation of the term "functionalism" as I intend to employ it in the following.

Chapter 2

Functionalism

2.1 Functionalism

Functionalism is a recurring concept in the Philosophy of Science and thereby needs to be thoughtfully defined to avoid any ambiguity and potential misconception lurking in the background. I shall primarily refer to two cases in which functionalism comes into play to better highlight the distance from my outlook.

The first one is Eleanor Knox's *spacetime functionalism*[1]. The origin of this proposal lies in the difficulties in defining spacetime. In Knox's account, «once one has analysed the role and understood what fills it», it is possible to get that «a structure that plays the spacetime role in a theory just is spacetime»[2]. For our purposes, it is not relevant what it means to "play the spacetime role". The effective meaning and significance of this role and therefore the identification of such a structure hinges on each specific theory. This stance, therefore, does not dovetail with the dichotomy substantivalism/relationism, *i.e.* whether the structure identified as spacetime has a fundamental ontological status or is derived by its relation with other objects (*e.g* laws, symmetries)[3].

However, what we are interested in is that within this kind of functionalism, we do not start from defining a geometrical structure as spacetime, but rather derive what structure is spacetime on the basis of its role. Functionalism, in this case, refers to the function that a certain object perform

[1] I consider Knox's account as a specific instance of what functionalism can look like, but it is only one among different possible functionalist proposals.

[2] Knox 2017, p. 4.

[3] See Knox 2017, pp. 2-3. Knox herself states that her view is, when maintained in its general terms, «neutral on substantivalism and relationism».

within a theory.

A second circumstance in which it is likely to detect an application of functionalism is in the case of the so-called *functional laws*. Let us start with an example. The equation of state of ideal gases is a relation between some magnitudes (*i.e.* the pressure *p*, the number of moles *n* volume *V*, temperature *T*) of the gas and some parameters represented by the gas constant *R*: $pV = nRT$. An analogous principle is Newton's second law. What characterises functional laws is the fact that they characterise some quantities as functions of other ones, which is the reason why they are sometimes called "quantitative"[4]. The general structure of a functional law is therefore of this kind: "P=f(Q)", where P and Q are the quantities involved. This form of law is particularly interesting in the critiques of some minimal Humean accounts of the laws of nature, claiming that laws are mere generalisations derived from empirical regularities. When the quantities considered range over a continuous set of values, functional laws provide an example of uninstantiated generalisation. It may well be the case, in fact, that we are not able to find an instance for every value of the quantities[5].

In this case, the term "functional" only refers to a connection between quantities ranging over possible values. Therefore, it is the form of the law to be functional. The perspective is internal to the elements governed by a law and not related to the connection between laws. In this discussion we are neither interested in this, nor concerned with the relation between laws and instances, which should be then set aside. Having ruled out this second interpretation, we can proceed towards recovering a third meaning of functionalism.

2.1.1 From Functions to Laws

Substance and Function (1910) is one of the main works of the neo-Kantian philosopher Ernst Cassirer. This work is a very comprehensive and far-reaching research project on the evolution of scientific thought. Even if the date of publication is halfway between Einstein's special and general relativity, and much before the mature formulation of quantum mechanics, it provides the basis for the philosophical interpretation that Cassirer would elaborate later on these new theories. Cassirer's understanding of physics detects the concept of function as the peculiar element underlying scien-

[4] Bird 1998, p. 91. Certainly, this does not imply that any law involving quantities is functional.
[5] Armstrong 1985, p. 22.

tific development, which proceeds from substantivalist stances[6] to more abstract representations where theories are structured in terms of the relations between their elements.

Before further delving into this, some background notions may be helpful to get a better sense of Cassirer's view of physics. The classical Kantian account of nature and science recognises that our experience is based on a ground of objectivity guaranteed by the uniformity of our intuition of space and time, as well as the categories of our understanding[7]. Although it may be easy to confuse Kant's idealism - namely the fact that the construction of our experience is determined by the shape of our sensiblity and understanding - with a kind of subjectivism, this perspective is actually entirely realist. Reality is in fact that construction that is guaranteed in its objectivity by the homogeneity of the form of our experience. Physics and mathematics are built upon that form, and therefore also scientific knowledge has the uniformity that we require. The nature of this structure is mainly relational, since relations constitutes the framework through which we can have experiences, make observations and get specific contents.

Cassirer starts from this perspective but with a particular focus on the problem of the construction of the scientific knowledge and working toward an account which is stripped of any psychologistic vestige. This constitues a first distinguishing feature of his philosophy. Physics and geometry were constant benchmarks for Kant as well, but here the discussion places science as the main subject matter. We are therefore led throughout an investigation whose center shifts to the objects of theories and in general of our scientific representation of the world.

In Cassirer's perspective our scientific experience (both theoretical and empirical) is based on some series of representations which - inasmuch as they are in fact representations - are already organised in a structured fashion that ensures further connections and generalisations. A single instance is never only the plain "individual" that metaphysics opposes to the "universal" when facing the problem of induction. The single case α is always represented within a function $\psi(\alpha)$ expressing its possible logical positions, *e.g.* in a certain temporal sequence, presented in a geometrical

[6] Please note that in this single case I am using the term *substantivalism* in a different way from its current meaning, namely referring to theories hypostatising some physical objects and considering them in a similar fashion to material objects, *e.g.* Carnot's conception of heath (Cassirer 1923 [1910], p. 403). I am not even referring, however, to the early discussion on spacetime substantivalism going on in these years (see footnote 2.2.2).

[7] All these Kantian terms are outlined in the next chapter.

structure, or in a causal connection with other events or instances[8]. I shall analyse this in more depth below. Suffice it to say that whenever we consider, for instance, the motion of a certain object during an experiment, there is no such thing as the object independent from a previous physical characterisation, *e.g.* within a spatial reference frame.

The core of Cassirer's functionalism is in this account of the single representation, which involves the overall scientific knowledge. Let us follow his line of argument.

> Two fields of phenomena A and B are first united each according to a particular law $\psi_1(\alpha_1, \alpha_2, \alpha_3)$, $\psi_2(\beta_1, \beta_2, \beta_3, ...)$; these laws are again connected among themselves by a new relation $\phi(\psi_1, \psi_2)$, until we finally reach the most general relation, which ascribes to each individual factor its definite place with regard to the others. The fundamental form F is analysed for thought into a structure of mutually dependent determinations, which would be symbolically represented, for example, by an expression $F[\phi_1(\psi_1, \psi_2), \phi_2(\psi_3, \psi_4), ...]$[9].

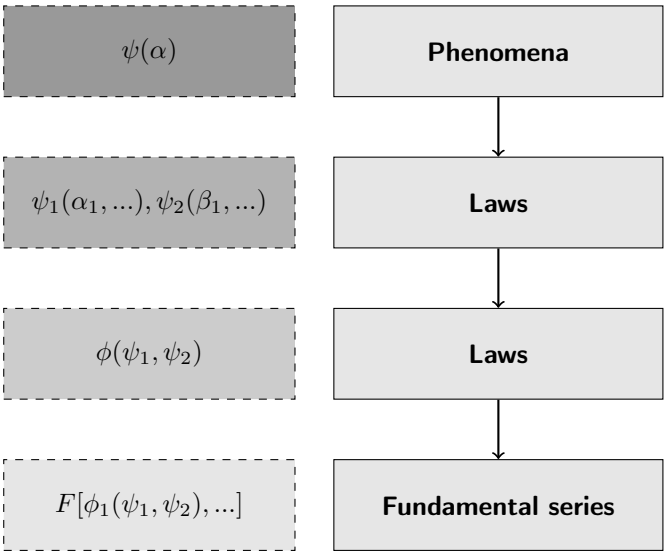


Figure 2.1: Functional organisation

Even if it is presented organised in levels, this is far from being a hierarchical structure. The organisation of laws is flexible and can be modified to meet both theoretical and empirical needs.

[8] Cassirer 1923 [1910], p. 248. See Chapter 2.3.1 for a discussion about the affinity between functionalism and structural realism on this point.

[9] Cassirer 1923 [1910], p. 267.

It may well be that new observations compel one to revise this structure, adapting the series of relations to a more complex description of phenomena[10]. However, this does not mean refuting it nor having to get rid of the whole of it. The fundamental forms ensuring the uniformity of our experience allow us to proceed modifying the network of functional relations. For instance, some phenomena can be arranged under a different governing law, or a law can be valid limited to certain restrictions or phenomena, turning out to be less general than another one. Different degrees of generality are admitted, but this does not single out any hierarchy. The Kantian factor of characterising objects (things, events, laws) *via* an aprioristic causal form does not prevent it from being modifiable depending on empirical bases. In this sense Kantianism does not commit to any strict metaphysical hierarchy in which, *e.g.* some principles always have a certain role within a theory.

If we compare this proposal to the two previous ones, we see that there are huge differences, and it was therefore necessary to disambiguate this meaning of "functionalism". Nevertheless, we can still recognise some common features. Cassirer's functionalism, in fact, on the one hand considers the objects (theoretical and empirical) of a theory on the basis of their function within the theoretical structure, and on the other it actually depicts these objects as included in a mathematical space of elements, functions and functionals.

This model has several virtues and will be explored further down the line, in order to get the basis to tackle the concept of explanation again. Before doing this, though, we need to characterise this functionalism in all its aspects more thoroughly and remove some possible ambiguities. This clarification can only be based on its deepest conceptual assumptions, which all lie within Kantian thought.

2.2 A Kantian Approach to the Laws of Nature

The neo-Kantian theory of the construction of concepts, into which Cassirer delves with a unique concern for scientific concepts, is based upon Kant's *Critique of Pure Reason*. What is crucial for us is the determination of two sources of our knowledge: *sensibility* and *understanding*. Sensibility provides us with the representations of the objects, but we only think (*i.e.* conceptually determine) them through understanding: «Without sensibility no object would be given to us, and without

[10] Cassirer 1923 [1910], Loc. cit.

understanding none would be thought»[11]. Phenomena, therefore, are not given to us through some immediate perceptual faculty, but are determined through the forms of sensibility and understanding. Let us then get into the nitty gritty of the issue and examine what is the nature of these forms.

Sensibility has two pure *a priori* principles: space and time. In Kantian jargon, *pure* means that they do not have any empirical origin nor do they rely on empirical components. *A priori* means that they precede experience, in the sense that they are the conditions for any experience to be valid. Here, what we are concerned with is only the validity of experience, *i.e.* the fact that an experience is possible. For this reason Kant's argument hinges on the condition of the possibility of experience. Kant is suggesting that whenever we have a determined experience, the phenomenon is always extendend in a spatial structure and located in a temporal sequence. His point is purely formal. Any phenomenon, *i.e.* any determined object of our experience has a spatial and temporal form. There are no further specifications other than that, in terms of the specific structure of these forms. We cannot have experience of anything that is not located in these spatiotemporal dimensions. However, these pure intuitions do not suffice to entirely determine the object. For this reason, we need to turn our attention to understanding.

A complete discussion of the functioning of the system of categories and principles that Kant sets out is outside the scope of this work. Suffice it to explain that, in order for a phenomenon to be thought, we need to introduce some principles accounting for how the pure logical categories of our understanding can actually be applied in our empirical knowledge. Kant starts from some purely logical categories and then lays out the way in which these provide us with a conceptual understanding of the world. Principles, in Kant's words, are «rules of the objective use of the categories»[12]. What we are most concerned with is the category of causality and dependence. In its original form it is only a logical function with no legitimacy in our actual knowledge apart from that of providing a general rule of connection of phenomena. This category can be then applied to phenomena only throughout its proper principle, namely that «All alterations occur in accordance with the law of connection of cause and effect»[13]. Experience is possible only insofar as phenomena are connected in a causal relation.

[11] *KrV* A51/B75 (Kant 1998 [1787], p. 193).
[12] *KrV* A161/B200 (Kant 1998 [1787], pp. 284-285).
[13] *KrV* A189/B232 (Kant 1998 [1787], p. 304).

The crucial point is that the level of this whole discussion is merely formal. While defining the pure intuitions and principles, Kant's perspective is entirely transcendental, *i.e.* concerning the conditions of possibility of our knowledge. Neither in the case of space and time nor in that of causality does Kant provide an empirical characterisation of phenomena, for instance by saying what precise causal rule will be experienced. From the very outset, this point hints at the question whether specific causal laws come into play and in what form. It stands to reason that these specific laws - which comprise laws of physics - are related to the causal principle. Kant is aware of the special role that physics play in this discussion and provides some very clear insights. In particular, he spells out that in our scientific depiction of the world there is a «possibility of infinitely manifold empirical laws»[14] which, with respect to our understanding are merely «contingent (cannot be cognized *a priori*)»[15].

The line of argument is clear. The principle that the form of our experience is causal, *i.e.* in order for an event to be an object of our possible experience it has to be arranged in a causal connection, does not anticipate anything about the specific content of that causal structure. All the formal structures we have encountered so far (pure intuitions of space and time, pure principles of understanding) do not anticipate any information about the matter of phenomena. Once we have said that the free fall of a body abides by a causal structure, we do not know anything about the specific law governing that family of phenomena. The only necessary feature for our understanding is causality, but the matter of the physical laws is completely contingent as far as our cognitive faculties are concerned. The transcendental perspective charts a course that can be completed only by our empirical experience, which is made up of all our experimental procedures and theoretical elaboration. It is only relying on specific empirical and theoretical assumptions that we get from the phenomenon of free fall to Newton's law of gravitation.

A significant aspect of this discussion is that, quite frequently, Kant refers to the concept of function as descriptive of the kind of objects outlined in his doctrine of concepts. Concepts themselves are defined as having a functional form, in the sense that a concept is a general rule encompassing a

[14] *KdU* XXXII/5:183 (Kant 2000 [1790], p. 70). I have already discussed in a BPhil Essay why the concept of "empirical law" is somewhat problematic for Kant [reference omitted for marking anonymity reasons]. In this handling we can turn a blind eye on this issue and just assume that the empirical laws consist in the commonly understood scientific laws.

[15] *Ibid.*

sequence of specific values, different from each other: «By a function [...] I understand the unity of the action of ordering different representations under a common one»[16]. More generally, the process of unification of the multiplicity of our experience, on the basis of the pure logical forms of the categories, is functional in an analogous fashion. Any object of our experience is never merely an external entity that we grasp through a passive filter, as sometimes Kant is understood. Insofar as they are objects of our experience, phenomena are already thought in a functional connection.

It may seem that when we move beyond the simple concept defined as above, the notion of function gets vague. Actually, the very procedure of synthesis, *i.e.* the connection of the multiplicity of phenomena under the categories, is entirely functional. We never have the mere x representing an immediate object, but always a certain function $y = f(x, z)$ connecting the multiplicity of phenomena *e.g.* in a causal relationship or in a certain temporal sequence. This is the fundamental meaning of Kant's functionalism.

This functional representation of the world is not just a perceptual depiction or a sort of psychological characterisation. This is apparent in the example of laws of nature. We begin from the assumption that any phenomenon in our world is represented within a certain causal structure, which is in no way psychological or subjective. Our scientific enterprise consists in that of unfolding the empirical content of this structure employing our experimental as well as theoretical methods seeking for those "empirical laws" such as the laws of physics. However, Kant undertakes his whole analysis on the transcendental level, without addressing in significant depth how physics works in concrete terms. With this introduction to the Kantian perspective, we can now delve into the specific issues of the science of nature. It is exactly from this point that Cassirer opens his line of research.

2.2.1 Objects and Theories

In his research, Cassirer is mostly concerned with the "fact" of physics, considered as a peculiar sphere of human knowledge with its own assumptions, terms and procedures. This point of departure is significant, as it indicates that Cassirer balks at considering physics as a discipline only dealing with

[16] *KrV* A68/B93 Kant 1998 [1787], p. 205. This approach can be compared with Frege's definition of a concept as a «function whose value is always a truth-value» (Frege 1960 [1891])[p. 30]. The use of functions to capture the different instances of a certain general concept is the same. However, the extension of the notion of concept and object in Kant and Frege is not equivalent and a further investigation is not essential for the purposes of this work.

pure terms in a "truer" dimension, or representing a "more real" world. Physics cannot be considered only limited to its formal structure, detached from the perspective in which its terms are taken into account in concrete terms, and disregarding the significance of its historical development. These assumptions are anything but some general philosophical overview that overlooks the technical aspects of the specific theories, and are rather crucial to ground a theory of the physical object[17].

Getting down to Cassirer's account of object, we recognise how deep its Kantian roots lie. The physical object is not the pure, standalone being of a straightforwardly realist perspective. Any experiment, procedure and assumption is always given within certain categories and in a specific theoretical structure. The role of the concept of cause is crucial in this regard. When we consider a certain phenomenon, *e.g.* an event produced in an experimental setup constituting our object, the underlying Kantian assumption is that this phenomenon will always be conceived within a certain causal structure. This holds for any observation in the science of nature. Even data that we do not understand or that pose significant threat to our scientific understanding are already to be conceived in some causal relation with the remaining part of our theoretical and observational framework. The point, here, is that this causal prerequisite constitutes the object itself: «We connect a series of different, temporally separate, circumstances by means of a unitary complex of causal rules; and this connection gives the individual, in Platonic language, the seal of being»[18]. The fact that we can say, *a priori*, that any phenomenon lies within a causal framework guarantees that we can expand our research and find theoretical connections with other terms.

As we have seen considering Kant's account of the laws of nature, the point is not that of finding specific laws but rather explaining why we manage to build coherent theories that can account for phenomena, and how we relate the experimental procedures to the whole of a theory. The issue of the relevance of empirical findings is particularly interesting. Cassirer does not introduce a strong ontological distinction[19] between single events and laws. I have mentioned that a common question is how can laws account for events and explain them. Here we find a clue to answer that. As Cassirer argues, «everything significantly factual is already a theory»[20]. This means that we can establish a

[17] By "object", here, I denote in general anything that can be singled out for experimental or theoretical reasons. Therefore, the object can be both an individual item (*e.g.* an elementary particle, a field) and an observed event.

[18] Cassirer 1923 [1910], p. 294.

[19] See Chapter 2.2.2 for a merely explicative distinction between statements.

[20] Cassirer 1956 [1937], p. 35.

relationship between the experimental fact and the theoretical framework right because the single event has already the logical form that enables it to be integrated into or read according to some specific theory. Here I prefer to defer the point of the relation between events, laws and explanation to the following chapters while spending a few more words on the role of the objects of experiments.

In particular, my previous remarks may have brought to mind the various philosophical approaches claiming that any observational statement is theory-laden. In different fashions, these stances have shed light upon the fact that the events we consider in our experiments and observations are always incorporated in a wider structure comprising various theoretical material: operational and procedural assumptions, specific theories, logical functions. A common ancestor of these views is Pierre Duhem's *The Aim and Structure of Physical Theory* (1906)[21]. His work comprises a subtle analysis of the role of experiments in physics. In particular, Duhem highlights the fact that the experimental and empirical facts are not merely observative phenomena that we try to match with certain theoretical elements in order to confirm some assumptions. From the very outset, rather, experiments are theoretically interpreted, so that the phenomena provided in our tests are viewed in the light of the theories that we are using in that experimental setup.

The employment of devices itself is subject to a theoretical interpretation. The example that Duhem suggests is that of the galvanometer[22]. This instrument measures the intensity of a current that flows from the positive to the negative pole of a battery through a coil placed into a magnet. In its turn, the coil is connected through a spring to a bar that acts as a pointer on a screen indicating the degrees of possible deviation. When the current flows through the spring it generates a magnetic field that, opposed to the magnet, twists the coil causing a deviation of the pointer. Duhem points out that, in itself, the variation of the pointer has no significance. Since we use that deflection to calculate the current intensity, the degrees of the angle of variation of the pointer provides us with the value requested only applying the laws of electromagnetism. The process of generalisation that fol-

[21] An instance of these more recent views is, for instance, Quine's epistemological holism in the form of the so-called *Duhem-Quine thesis* (Quine 1951). However, note that Cassirer's reception of Duhem is quite specific and he does not take a hardline on holism. When Cassirer emphasises the role of experimental assumptions, this is only intended to consider all the theoretical elements involved in the characterisation of the physical objects, including those related to the execution of an experiment. However, this never drifts to an empiricist derivation of these theoretical elements. The line of thinking Cassirer is pursuing would rather be more congruent with Friedman's relativised *a priori* and the associated historicist critique to Quine (see Goldberg 2009, pp. 262-264, Friedman 2001, Ch. 2.1, 2.2 *inter alia*).

[22] Duhem 1982 [1906], p. 154.

lows, allowing us to derive the value of the current intensity from any galvanometer, is also a process of symbolic interpretation (*i.e.* a translation into abstract, mathematical terms) of the experimental event through a certain physical theory.

It is no coincidence that Duhem is also one of Cassirer's main sources when it comes to discussing the value of experiments and empirical findings[23]. Here, Duhem's remarks lead to substantial results in the conception of a physical object. In Cassirer's words, the concept of thing «separated from the whole of the empirical connections by which it is accredited, has no significance»[24]. This is a very strong stance, as it is the ontological status of the thing itself to be at stake. A point that *prima facie* seems to involve merely the form of our experimental results actually calls into question the status of objects in the physical sciences. Here Cassirer's project transcends Kant extending that conception of the object built throughout categories and conceptual structures to the construction of objects within exact sciences.

So far, I have focused on the role of experiments and empirical findings with regard to the underlying theoretical structure. The importance of our scientific experience for Cassirer is unquestionable[25]. However, if we finished this analysis here, Cassirer would appear as a sort of pragmatist only concerned with a reassessment of the role of experiments within the definition of the objects, but with no particular insights into the most significant problems of physics. Therefore, we need to focus our attention on the whole problem of the ontological status of physical objects, and we shall see that the experimental elements have to be integrated into a wider theory of objects. Let us then get back to the concept of function and assess what kind of contribution it can provide to this discussion.

2.2.2 Functions of Objects

The above remarks on the status of experimental facts points toward the problem of the object of physics in general. We are now on a different level from that of a perceptual conceptualism criticising the "myth of the given"[26]. The object we are considering is no more the mere perceptual result

[23] See *e.g.* Cassirer 1923 [1910], p. 144, 281.

[24] Cassirer 1923 [1910], p. 305.

[25] Criticism «recognizes no other and no higher objectivity than that, which is given in experience itself and according to its conditions» (Cassirer 1923 [1910], p. 278).

[26] Even if the perspective is certainly analogous, even considering the role of Kant. See Sellars 1956.

of our sense, but is conceived within a highly-conceptualised scientific framework. The shift that Cassirer operates consists in extending to scientific knowledge the Kantian concept of the object as given through a specific functional connection. In Kant, this means that we have no experience of an immediate object x, but only of a functional object $y = f(X)$. The function represents the logical connection that constitutes the object *per se*, where the logical connection can be the spatiotemporal representation or the causal relationship within a chain of events. When we analyse an event in physics, we proceed in a further specification of that functional connection. The object, for instance, is represented in some spatial coordinates that can be merely formal or descriptive of the experimental setup.

In Cassirer's depiction, the term *functional* denotes the characterisation of the object as represented within certain relationships. These relationships are not something that get added onto an alleged "real thing", but rather are what constitute the thing and the concept of reality as such. However, it may seem that this characterisation of functionalism is still somewhat vague. The problem is that Cassirer himself does not present a clear definition of function encompassing the whole range of examples he discusses. I have therefore identified three constitutive elements of functionalism that can help us in bringing this discussion forward.

(i) Formal

A certain physical object can be described on the basis of different variables and parameters. Let us consider as our object a particle involved in some kind of physical process. As we have seen in Chapter 2.2.1, the individual particle will be already considered with respect to some spatial and temporal parameters, *e.g.* in a three-dimensional spatial coordinate system and at a certain point in time t_0. Cassirer calls this object $f(x)$ to indicate the fact that, from its most primitive stage, it is in a certain experimental or theoretical reference frame[27]. If this object is considered as part of an event as a spatial displacement, it is then considered in a wider functional connection $\psi_1[f(x)]$. This can be, for instance, the equation of motion used to represent its displacement. The choice of the law

[27] Such a framework should be conceived in the most general terms and this dependency on some theoretical assumption holds even if *e.g.* we describe the motion of this object in a coordinate-independent way. Moreover, please not that x here still refers to an object (as we have seen in the Kantian thought) and not to a coordinate system. Similarly, ψ in the next passage indicates a law.

through which we want to describe the event depends again on our experimental and theoretical purposes. We may use kinematic equations involving only quantities that can be defined on the basis of the spatial and temporal coordinates, such as velocity and acceleration. Or we may decide to consider a functional connection $\psi_2[f(x)]$ represented by a dynamic equation of motion that take into account quantities such as mass. In this example, I am referring to the possibility of describing the motion of an object in a kinematical or dynamical representation, depending on our aim. The connections that are made by means of laws or different descriptive parameters characterise the object as such, and represent the most fundamental meaning of Cassirer's functionalism.

(ii) Conceptual

The concept of object in physics can also be wider and not necessarily limited to a pure formal characterisation. It is very common that objects that physicists have identified through experiments or hypothesis. The example that Cassirer suggests is that of the concept of atom[28]. It is widely known how the atom has represented very different entities even from the Greek philosophy. However, we are only concerned with the more recent development. In general, the atomic theory postulates the existence of a "minimum" in the composition of matter, identified with the term "atom". Our understanding of this discrete structure of matter underwent radical changes throughout the 19th and 20th centuries. At the early stages of its modern formulation, the atomic theory was opposed by theories that recognised energy as the basic constituent of matter and regarded atomism as a scholastic doctrine to overcome[29]. It is clear that Planck's law, moving that concept of fundamental minimum to the emission of energy, ushers in a new phase of the atomic theory and the relationship between matter and energy. If we then compare, *e.g.* Rutherford's atomic model to the classification of elementary particles provided by the standard model, we can see how much the concept of the minimum constituent of reality has changed.

This represents quite closely how science proceeds under steps that change the content of some concepts thanks to empirical findings. Cassirer dwells on this case to analyse how an object can change and actually have different content from time to time[30]. If reality was only to be conceived

[28] Cassirer 1956 [1937], Ch. IV.II.
[29] *E.g.* See the discussion on the failure of Ostwald's energetics in Cassirer 1956 [1937], pp. 141-142.
[30] This point is very close to how the issue of continuity of reference through theory change is addressed in works as

on the basis of the experimental evidence, we should say that almost any object that was employed at some point in a scientific theory was not real, and that almost every scientific hypothesis - excluding the ones we are currently working with - is false. We are not interested in saying that the concept of atom, or other concepts identifying the minimum constituent of matter, are wrong or false on a case-by-case basis. Science does nor postulates objects in order to affirm that they are the ultimate reality. Science employs some concepts, like those of the atomic theory, to proceed and confirm its hypothesis through empirical evidence. The objects can be considered as existing only referred to a certain theory, which means that it is determined by a certain structure of relations. The laws of a theory determine which objects are included and considered as real. The functional element lies in this point. If we consider the concept of the minimum consituent of matter, the empirical content of this concept can change and be adapted according to theoretical and experimental development. Different atomic models provide the material content of this conceptual function, but the functional form of the concept accounts for the continuity of these models and their validity in different frameworks.

Am I saying that we should give up on objectivity, or just refrain from giving to our scientific object any dependability? Certainly not. The fact that the concept of a certain object can change its empirical content depending on the state of the art of each scientific enquiry does not imply «it does not ask whether the whole of experience is objectively true and valid,— for this would presuppose a standard that could never be given in knowledge,— but it only asks whether a special, particular content is a permanent or transitory part of this whole»[31].

(iii) Operational

First and foremost, it is necessary to disambiguate the term "operational". Here I am far from any operationalist stance regarding the meaningful scientific objects only those that can undergo

e.g. Hardin and Rosenberg 1982. Such a compatibility may seem incongruous, since Hardin and Rosenberg advocate a realist stance. However, I agree with Laudan 1984, p. 159 that this putative realism is enfeebled to the point where their reference is no more the realist's one: «If empirical success [...] is not co-variant with genuineness of reference, then there is no reason whatever to believe that the world is the way our most successful theories say it is». Interestingly, Laudan regards Hardin and Rosenberg's causal theory of reference as "functionalist" (Laudan 1984, pp. 160-161). He only refers to the fact that their theory of reference considers objects of outmoded theories as still referring if there is some explanatory continuity with corresponding concepts in the new theories. There is no mention of Cassirer.

[31] Cassirer 1923 [1910], p. 278.

a process of empirical confirmation *via* measurements or analogous feasible practical operations[32]. The way I employ the term "operational" refers to the role that a certain concept plays within both theoretical and empirical operations that define its status. To clear up any confusion, we can refer to what we have encountered in Knox's functionalism (Chapter 2.1). The proposal of considering spacetime the structure that plays the role of spacetime is akin to the characterisation that Cassirer gives to the physical object. Interestingly, he puts forward the same example of spacetime in his *Einstein's Theory of Relativity* (1921)[33]. Actually, Cassirer's wider concern is related to the evolution of the concept of relativity and the coherence of some theorical premises of Einstein's work with the Kantian perspective. In particular, the relativisation of the concepts of space and time proceeded in the same direction as Cassirer's interpretation of Kant. In Cassirer's view, Einstein's renowned claim that the postulate of general covariance «takes away the last remnants of physical objectivity from space and time»[34] confirms the revision of the Kantian notion of space and time that he was pursuing[35].

This third meaning of functionalism arises from the role that some classical physical objects play in General Relativity (GR). The meaning of the metric field in GR is completely different from theories as Special Relativity where we deal with flat spacetime represented by the Minkowski metric field. In GR, the metric field has its associated dynamical equations, and the fact that the line element with its ten independent components can also represent the gravitational field poses crucial questions regarding the difference between metrics, matter fields and force[36]. In Cassirer's account, these three terms that have always been considered three separate objects now change completely their meaning. We can still talk about space, energy and matter, but now they are functionally connected through the new role that metrics play. The very same objects, therefore, change their mathematical representation and acquire new functional connections even being defined through

[32] Here I refer to Bridgman's operationalism (see Chang 2019). In my previous analysis of Duhem's stances, some remarks may have sounded similar to Bridgman. However, despite misleading analogies, Cassirer would not have endorsed such an extreme position. Methods of measurement are one of those theoretical elements involved in the characterisation of objects, but Cassirer would never establish them as the defining criterion of concept(*viz.* Bridgman)

[33] Re-published as an appendix to Cassirer 1923 [1910].

[34] Einstein 1920 [1916], p. 97.

[35] This is meant to be in an anti-substantivalist fashion. However, note that the discussion about relationism/substantivalsim of spacetime structures was quite primitive and therefore Cassirer's enthusiasm about Einstein's claim should be taken with a grain of salt. The fundamental point here is the fact that the relativisation of space and time perfectly fit in his way of understanding the Kantian perspective.

[36] Cassirer 1923 [1910], pp. 397-398.

the same concepts as before. It is the role that each physical object plays in the theory that lead us to call it *e.g.* "matter field" or "spacetime", and not a unique and ultimate characterisation.

Cassirer's remarks are surprisingly akin to a whole line of argument developed in the most recent literature and well represented by Rovelli's example of the action of gravitational waves[37]. The idea that gravitational waves can destroy the rock of Gibraltar similarly to what the electromagnetic field can do, should lead to a revision of the distinctions between different fields. Rovelli's proposal is to use the term "spacetime" for the differentiable manifold only, while "matter" for all the other fields with dynamical components. It is clear, in this example, that the classical distinction between energy and matter gets more and more blurred. What is the ontological outcome of this point? In Cassirer's view, this example substantiates the point that in physics there is no such thing as an object defined *per se*, a standalone item whose properties completely characterise it: «the scientific concept of a definite object comprises in its ideal completion not only the totality of its properties given here and now, but also the totality of necessary consequences, that could develop from it under definite circumstances»[38]. These "definite circumstances" are the relations that determine the status of the object.

2.2.3 Low-hierarchical Functional Structures

The three possible meanings of functionalism I have suggested are a way to sort out different uses of the term that we find in Cassirer's works. However, I do not take the fact that Cassirer does not disambiguate providing a unique definition of functionalism as a fault. What he is mostly concerned with, in fact, is analysing and singling out the different ways in which physics considers its objects, without claiming to give a unique definition of it. It would be at odds with functionalism itself to expect to reduce the whole of physics to a single form of treating its objects. His point is that, regardless of which specific approach we take, things in physics exist on the basis of their position within a system (experimental, formal, conceptual, etc.). What we are interested in, therefore, is not to establish whether there is a ultimate system that describes nature in its entirety, but only to develop and adapt our theories in light of our experimental or theoretical findings, adapting concepts

[37] Rovelli 1997, p. 193.
[38] Cassirer 1923 [1910], p. 294.

and objects to get more and more reliable depictions of nature[39].

The premises of this, as I have pointed out emphasising the Kantian roots of this perspective, lie in the fact that every object is already contained into a connection that makes it amenable to a causal representation. Insofar as the object as such is found within some framework of representation, events and phenomena can be represented into a theory with laws, assumptions and peculiar procedures. The generalisation, therefore, is not something that we acquire in a second phase, starting from this standalone objects with pure individual characters. Objects are already generalised, as they are conceived in a network of causal connections. This is the reason why it is not a miracle if laws can describe and account for singular events. There is no such thing as a strict individual event. Each object is already generalised and connected in some general representational framework.

This stance is quite the opposite of an inductive perspective. The problem with representations of nature based on induction is that they inherently tend to engender hierarchical classifications. Phenomena are captured through different inductive processes and we then get to wider and wider generalisations. It is quite easy to build a hierarchy based on the different degrees of generality involved. The scheme I have proposed in Chapter 2.1.1, rather, represents the way in which functional connections are established to build a theory or a comprehensive theoretical structure. However, there is no reason to claim that such a construction has a hierarchical form based on an intrinsic ontological difference between, for instance, different kinds of generalisation.

Nevertheless, this perspective does not prevent us from differentiating the components of a theory on different bases. It is important to explore how such a classification can work, in order to rule out a strictly hierarchical option. For instance, in *Determinism and Indeterminism in Modern Physics* (1937) Cassirer sets out a threefolded stucture of scientific knowledge, whose constituents are the following[40]:

- *Statements of the results of measurements*. With these statements we proceed from the sensible representations to the physical statements. Data are given in a determination that serves as material for the physical enquiry.

[39] «We are not concerned with estimating the absolute value of the system in its totality, but with establishing a difference in value among its particular factors» (Cassirer 1923 [1910], p. 278).

[40] Cassirer 1956 [1937], Ch. 3-5.

- *Statements of laws*. Phenomena of our empirical observations are governed by statements that connect them in a functional connection. These statements are based on a "if-then" structure.

- *Statements of principles*. Laws are organised, correlated to each other via principles which are «not themselves laws, but rules for seeking and finding laws»[41].

An instance of principle considered by Cassirer is the principle of stationary action. In general, the principle states that the path of a system evolving between two points of the configuration space is the one for which the action is stationary, *i.e.* $\delta S = 0$. Within different theories, the principles is formulated with proper parameters and quantities. In Lagrangian mechanics, considering a Lagrangian function $L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)$ Hamilton's principle states that the trajectories of the Hamiltonian action $S = \int_{t_0}^{t_1} L\, dt$ are stationary, *i.e.* $\delta S = 0$. We can then derive from them the Euler-Lagrange equation of motion, as well as imposing the variation of the action to be zero in the Einstein-Hilbert action we obtain Einstein fields equations. In our empirical evidence, we observe the motion of systems and describe it as governed by those equations of motion.

One may be tempted by concluding that we are being offered a pyramidal system with different levels of generality built one on the other, downplaying the role of principles as mere abstract systematisations drawn from the fundamental empirical levels. However, such a reading relapses into the inductive reasoning that functionalism is opposing[42]. Even in this tripartite portrayal, the point we are pivoting on is the theoretical nature of any factual element cited above. Phenomena are neither the basis of a theoretical construction nor something derived from higher principles on the basis of a vague aprioristic perspective. If we get back to the example of the principle of least action, in fact, we see that it can be both considered as a common description of the phenomena of natural motion and as an instrument to the systematisation of the equations of motion, without any one-way derivation being in place. The image that Cassirer suggests is that of a sphere of interrelated elements, rather than a pyramid. Phenomena, Laws and Principles are connected in a functional whole with no strictly hierarchical structure[43].

[41] Cassirer 1956 [1937], p. 52.

[42] Please note that here I'm not referring to the probem of induction in Popper's terms, but rather in the strict Humean sense of the generalisation from recurring events to laws of nature.

[43] As Massimi says, it is «a purely 'architectonic' distinction» (Massimi 2011b, p. 7). Also French holds that this arrangement is non-hierarchical (French 2014, p. 79n).

A last remark that I would like to make on the risks of having a pyramidal view involves a perspective that may seem authentically Kantian. Massimi 2011a addresses the data-to-phenomena inference issue, investigating the meaning of the distinction between data and phenomena. Massimi's aim is to object both to externalist and conventionalist perspectives by endorsing a halfway Kantian stance.

Her starting point is a drastic stance against any metaphysics of ready-made phenomena, a position that is certainly consonant with Kant. In particular, Massimi bases the distinction from data on the idea that phenomena are "conceptualised appearances", in the sense that they are appearances subsumed a spatiotemporal and causal structure. There is a transformation from data to phenomena that follows several steps[44]. First of all, data are recorded via multiple experiments that deliver results as relative frequencies of occurrencies. Data are then processed both through specific parameters and analytical models to output graphs. Only after this stage, we need to introduce a specific causal concept to explain the employment of new parameters accounting for the results of the analysed data.

The problem is that in this depiction, it sounds like there is a process of filtering out raw material through schemes, which is quite the opposite of Kant. The transcendental perspective is not about a procedure of conceptualisation. We have nothing but phenomena to deal with, which is the reason why the categories of our understanding *constitute* the phenomena rather than stepping in at a later stage. In Massimi's account, it is a process made up of sequent levels of abstraction and elaboration of the raw data: «we make them [Phenomena] by first ascribing certain spatiotemporal properties to appearances [...] height), and then by subsuming them under a causal concept»[45].

The distance from Cassirer is apparent. In his functionalism, we have seen that it is only because every phenomenon is already subsumed under a causal connection that we can make it fit in a wider theoretical structure. Massimi, instead, admits only raw data coming from our experiments as the basis of a gradual and subsequent processing. In this sense, she is performing a data-to-phenomena reasoning, but there is nothing Kantian about that. The fact that *eventually* the material we work on is conceptualised is in no way Kantian. Rather, it is compatible with most forms of non-simplistic

[44] Massimi 2011a, p. 110.
[45] Massimi 2011a, p. 109.

empiricism.

For instance, Massimi contests McAllister 2004, which holds that Galileo's study of free-fall just a process of polishing the data via removal of the accidental occurrences. Massimi argues that the phenomena are not the result of this sophisticated noise reduction, because it would lack the spatial, temporal and causal framework that the scientist applies to the data. This is certainly true, but if we stuck to this scenario, we would not understand *e.g.* why these phenomena can be eventually made coherent in a theory or how they can evolve into a historical perspective. Her stance does not differ so much from what she tries to dispute, as it only makes that polishing process a bit less conventional and more dependent on the categories of our understanding. I deem that, even if Massimi's premises have an explicit Kantian flavour, eventually she ends up in a middle ground in which Kantianism loses much of its strength.

These last remarks aim to clarify the reason why I keep underlying that, even if it admits classifications and categories, functionalism is non-hierarchical. This is crucial for what follows in Chapter 3.2. However, before assessing how explanation may fit into this, it can be useful to briefly explore how functionalism can face some other contemporary issues in philosophy of science.

2.3 Functionalism and Structural Realism

The notion of "structuralism" encompasses a variety of views across linguistics, philosophy, psychology, anthropology and sociology. In general philosophical term, we can call "structuralist" those stances portraying a system of knowledge (*e.g.* theories and texts) in which the status of each individual term is determined only by the position within the system. Such a definition sounds everything but exhaustive, despite revealing that the plan of being comprehensive is thwarted by the multiplicity of disciplines in which we can find some structuralist approach. However, if we buy into this definition, we cannot help but regarding Cassirer as a structuralist. Focusing on this discussion of scientific theories, it is clear that phenomena, experiments and generalisations become meaningful only on the basis of their position within an appropriate theoretical structure. The peculiar point in a Kantian perspective is that the existence itself of these objects depends on their belonging to a structure of relations.

It is no wonder that Cassirer has been considered as a forerunner of contemporary structural

perspectives. Gower, for instance, underlines how both Cassirer and Schlick have come up with forms of structural realism even through very diverse routes characterised by even quite harsh arguments about several issues such as the nature of space and time in the context of the theory of relativity[46]. In particular, Gower locates the common structural element in the role that the *a priori* logical forms play in the constitution of objects. The conditions of objectivity of our experience are ensured by these *a priori* elements, such as the aforementioned causal network. The relations established within these theoretical structures ground the objectivity of physical objects. The result is an «austere version of structural realism»[47] where reality consists in «the structure of any possible mathematical physics yelding objective knowledge»[48]. The sort of generality that Gower ascribes to Cassirer may be disputed. In particular, from these last few words it seems that the mere fact of being amenable to a mathematical formulation guarantees the objectivity of a theory and the reality of its structure[49]. However, it is correct that Cassirer's conception of objects is structural according to Gower's understanding of structuralism.

Another recent analysis of the relation between functionalism and structural realism can be found in French 2014. French's exposition of the neo-Kantian perspective is overall precise and gets to the key points of the issue. I am setting aside a more detailed account of these points and I shall put it forth in the next chapter. Suffice it to mention two elements that French regards as being distinguishing in Cassirer's structuralism[50]. The first is the fact that objects become dependent on relations, and therefore it is the structure displayed by this relations to have an ontological priority. The second point is that objectivity is not pivoting on the objects anymore, but rather derives from the structure itself. Laws, principles and theoretical assumptions constitute the basis to say that we are moving across an objective ground.

Gower and French's works certainly offer interesting suggestions to further examine the similarities between Cassirer's functionalism and structural realism. However, the fact that some points

[46] Gower 2000, pp. 101-102.

[47] Gower 2000, pp. 94-95.

[48] Gower 2000, p. 94.

[49] It is true that the formal structure grounds the objectivity of our scientific experience. Nevertheless, the mathematical formulation has to be supplied with other theoretical assumptions (*e.g.* those related to the execution and interpretation of experiments) in order to genuinely provide us with objects. Cassirer's passage quoted (Gower 2000, p. 94) should be interpreted with this *caveat*. This "highly generalised" (*Loc. cit.*) structure is therefore always inclusive of assumptions related to the empirical sphere.

[50] French 2014, p. 99.

of contacts can be found in retrospect does not amount to saying that there has been a specific influence or derivation from the first to the second. Assessing this is a task for serious historical study. What we are mostly interested in is rather carrying through with both those suggestions taking into consideration some specific philosophical problems.

2.3.1 Which Structure?

Besides the general analogies that one can trace between Cassirer's functionalism and structuralism, what we are most concerned with in the following is how these two stances cope with some specific issues. In particular, it may be interesting to first consider the two problems constituting the matrix from which structural realism is developed as an alternative to stricter forms of realism.

With his 1989 paper *Structural realism: The best of both worlds?*, John Worrall is the trailblazer for the contemporary discussion about structural realism. His work pivots on two crucial arguments of the debate on scientific realism and expounds structural realism as the approach able to overcome these criticalities.

- No-miracle argument

The reason why our scientific theories work, providing us with coherent depictions of the world, is everything but self-evident. Theories are, all things considered, human constructions to understand of phenomena employing both an explanatory and predictive instruments. The agreement between our theories and the world is not to be taken for granted. However, in our common practice we realise how theoretical structures adequately represent phenomena ensuring the effectiveness of natural sciences. In and of itself, this correspondence has sometimes been called miraculous[51]. When we shift our focus from this correspondance to the world itself, it would be equally miraculous that our theories work (*i.e.* are explanatory and predictive) if we did not admit that the description they provide is true. This is a strong argument in favour of realism, as Putnam puts it: «The positive argument for realism is that it is the only philosophy that doesn't make the success

[51] «The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve» (Wigner 1960, p. 14). On a sidenote, it is intriguing that both Worrall and Wigner mention the same examples of mathematical depictions, *e.g.* the Lamb shift, quantum mechanics, and Newton's law of gravitation.

of science a miracle»[52]. The no-miracle arguments amounts to saying that the effectiveness of our scientific theories entails the reality of some aspects of the world they describe.

- Pessimistic meta-induction

The realist stance loses much of its strength when we consider the historical development of scientific theories. It is apparent that all the former descriptions of natural phenomena that have been employed throughout the ages are now outdated and replaced by the most recent ones[53]. The induction lies in the fact that we should expect our current theories to be replaced exactly in the same way. With respect to this argument, therefore, the fact that our theories provide a true (or approximately so) description of reality is quite unconvincing. We are thus encouraged to seek some other way to account for the discontinuity that we come upon while comparing the series of these descriptions[54].

Structural realism (SR) indicates quite a broad ensemble of proposals shifting the perspective of mainstream realism from the objects to the structures and relations characterising scientific theories. The problem of theory-change is addressed from the side of the conservation and change of mathematical and theoretical structure across theories, setting aside the much more problematic issue of the modification of the objects to which one scientific theory is committed. From the perspective of the structural realist, therefore, we both preserve the assumption that the explanatory power of theories is non-miraculous and the observation that substantive changes occur in the development of scientific theories.

There are two wide categories of SR[55]. *Epistemic Structural Realism* (ESR) limits its analysis to the epistemological issue, drawing attention to the point that what we know are specific structures without access to the ontological side. It is a stance which is realist on theory structures and agnostic about objects. *Ontic Structural Realism* (OSR), on the other hand, considers the ontological question about the status of structures and physical objects. It goes without saying that we are

[52] Putnam 1975, p. 73.

[53] This claim may be more nuanced, of course. One may respond that Newtonian mechanics *qua* predictive physical theory has not been replaced. Rather, what has been replaced is the ontological picture of the world associated with Newtonian mechanics. However, here I am making a general point about the modifications occurring across theories and not committing to a specific stance about theory-change.

[54] Worrall 1989, p. 109.

[55] For this distinction, see Ladyman 1998.

concerned with the latter group of proposals of structural realism, since this whole discussion is primarily ontological and Cassirer's arguments consider all the scientific objects *qua* objects requiring an ontological analysis.

In general terms, structural realism and Cassirer's functionalism address the same theoretical problems. Cassirer as well develops a proposal able to both account for the objectivity of our scientific descriptions and theory-change processes. However, these first remarks do not entail that the two positions come to the same conclusions. A first point that one should always bear in mind in this regard is the fact that Cassirer addresses the problem of reality from an entirely different perspective. Within the Kantian thought, no-miracle arguments sound ill-founded. The problem is not that there are two layers, reality and our description of reality, that we join progressively adapting the descriptive part in order to get a better comprehension of the world. Rather, scientific knowledge has a constitutive function of its objects. As soon as we broach the matter of knowledge introducing *a priori* structures that constitute the framework for our entities, the problem of the comparison between these entities and our way of knowing them vanishes.

However, this kind of comparison cannot be kept on the mere analogical level, there are at least two more issues that can help us figuring out to what extent Cassirer's functionalism is compatible with (some form of) structural realism.

1) Structure

From a general perspective, one can find important analogies about how structure is described in Cassirer and OSR. However, as soon as we get to the details of the different options within OSR, a hasty analogy may take us on a slippery slope. A significant problem involves the conception of reality. Cassirer, as I have explained, is concerned with the objectivity of our scientific knowledge. As long as we manage to elaborate our theories on this objective ground, we should give up on the question about the reality of the elements we are considering. I will discuss eliminativism about objects in the next point, but the wider problem of the reality of the structures[56] should first be addressed. As French emphasises, the structural realist may want to maintain some base level of reality,

[56] I am referring to how SR conceives structures in terms of their mathematical and theoretical form.

being unhappy with Cassirer's eliminativist conception of objectivity[57]. If we forego considering the individual objects as real, it is still possible to ascribe reality to laws or mathematical structures. The realist perspective would then be maintained recognising that structures and not objects are real. This outcome certainly is at odds with Cassirer, since his functionalism does not call for an independent reality but only objective representations. From the functionalist's perspective, then, this kind of structural realism is just a form of realism shifting its reference from objects to relations.

As far as the specific structures involved, a detailed survey on the correspondence between functionalism and specific forms of OSR is beyond the scope of this enquiry. The fact that functionalism focuses on the theoretical relations and structures to determine and lay out objects is not sufficient and we would need a more specific analysis of how different proposals of OSR conceive structures. However, it may be worthwile making a brief example. Ross, Ladyman and Collier 2007 puts forward a form of strucural realism defining reality as follows: «To be is to be a real pattern»[58]. The concept of real pattern comes from Dennett 1991, and is based on the assumption that when we admit objects in certain frameworks, it is very rare to look for a *bit map* description of the object: what matters is not that this object can be decomposed in a one-to-one correspondence with the basic entities admitted by the theory (in this sense Dennett uses the concept of *bit map*). Quite the opposite, we admit the not-fundamental object as long as such bit map description is not available but our operation ensures a faithful and economical description of that object within the framework. In Dennett's words: «A pattern exists in some data - *is real* - if there is a description of the data that is more efficient than the bit map»[59]. This form of SR conceives reality as the system of relations between patterns. Most interestingly, this proposal has a pluralist stance about structures and does not focus on some preferred ones, *e.g.* quantum mechanics[60], which is very close to Cassirer's view. There is no preferential structure, and therefore no objects with some priority in a certain domain of discourse. This stance, called "rainforest realism"[61], just like Cassirer's general conception of sciences, is open to the difference of admitted objects that may occur across various theoretical

[57] French 2014, pp. 99-100.
[58] Ross, Ladyman, and Collier 2007, p. 226.
[59] Dennett 1991, p. 34 (*italics mine*).
[60] See the criticism of French's view in Ladyman 2019, p. 2: French has a non-ecumenical approach fostering quantum physics as a fundamental structure to the detriment of special sciences ontologies.
[61] For a hint to the origin of this name see Ross, Ladyman, and Collier 2007, p. 233-234.

domains (branches of physics, scientific disciplines)[62].

2) Eliminativism

As I have outlined above, the peculiarity of Cassirer's stance - and a substantial result of his Kantian legacy - is the fact that it is the existence of the objects in our theories to be determined on the basis of their functional connections. Objects end up being only those entities built in nodes of relations and are, therefore, ontologically derived. The strong structuralist flavour of this approach is undeniable. The theoretical structure of our theories determines those entities that we require in order to obtain a coherent depiction of phenomena, also able to anticipate predict regular events.

However, this kind of eliminativism about objects is disputed also within structural realism. For instance, Esfeld and Lam 2011 provides a non-eliminativist account of objects in a form of "moderate" ontic structural realism based on the assumption that if properties and relations are the modes of existence of objects, then we must admit the existence of fundamental physical objects. Making this point orthogonal to the problem whether individual objects are distinguishable, Esfeld and Lam maintain an ontology of thin objects[63].

Unlike these moderate options, we should regard Cassirer as holding a strong eliminativist position. French and Ladyman 2003 examines Cassirer's theory of object emphasising a crucial point. While, from the realist's standpoint, we start from the object with its properties and then regard it as integrated within a frame of relations, in the functionalist perspective laws and structures come first and on these basis objects are singled out[64]. Accordingly, while considering the concept of *e.g.* electron, this is inspected with purely Kantian arguments[65]. The quantization of the elementary charge e let us consider the electron as a «"determinate object"; but it not sufficient for substantialization and hypostasis of the electron»[66]. Both in Cassirer and eliminativist form of structural realism, the objects to which our theories commit are given *via* their dependence from the defining structure.

[62] «Some of these patterns are indeed conceptualized as individuals in some special sciences, while simultaneously not being so conceptualized by other special sciences making projections at other scales of resolution» (Ross, Ladyman, and Collier 2007, p. 229).

[63] Esfeld and Lam 2011, p. 152. Esfeld explicitly recalls the notion of "thin objects" of French and Ladyman 2011, p. 30.

[64] French and Ladyman 2003, pp. 39-40.

[65] Cassirer cites *KrV* A258/B341: «And whatever we can cognize only in matter is pure relations [...] But there are among these some self-sufficient and persistent ones, through which a determinate object is given to us» (Kant 1998 [1787], p. 379).

[66] Cassirer 1956 [1937], p. 182.

These two points reveals on the one hand why structural realism may find in Cassirer's functionalism an important touchstone, on the other that functionalism - far from being just a tale from the "lost history of structuralism"[67] - can adequately address some crucial problems in contemporary philosophy of science. I could have raised many other points that may lead to fruitful discussions. For instance, the arguments employed in Ross and Ladyman 2007[68] to claim that OSR can fit in the semantic approach to scientific theories hold similarly also for functionalism. Cassirer does not provide a complete account of scientific theory, with a definition of theory and an ordered depiction of its constituents. Therefore, we may assess the relation of functionalism with the semantic view only after a major review of Cassirer's interpretation of how theories work. However, so far, I have outlined and introduced the crucial elements to revert to the problem of explanation and evaluate how functionalism can address it.

[67] French 2014, p. 65.
[68] Ross and Ladyman 2007, Ch. 2.3.3.

Chapter 3

Functional Explanation

3.1 Towards a Functional Theory of Explanation

So far, the only account of explanation I have considered as adequate is the DN model. My point is that the DN model provides a minimal framework through which we can read different kinds of explanation with some common essential features. Apart from the *explanans* / *explanandum* pairing, which dates back to Aristotle's *Topics*, we are interested in the way in which both the nomological and the deductive components are arranged. I deem that this form of the *explanans* as comprised of laws and auxiliary conditions constituting a deductive basis is crucial. However, I have argued that the lack of flexibility of the DN model is also its main weak spot. It is in fact quite easy to argue, as many critics do[1], that it does not suffice for a procedure to satisfy the DN model in order to be immediately explanatory. The problem is that some proposals to integrate the DN model with context-based constraints run the risk of reintroducing psychological and subjective elements in disguise, which the DN model was supposed to eliminate.

The proposal of a functionalist account such as Cassirer's proceeds in the direction of providing a scheme like the DN model with the necessary constraints. In this framework, in fact, we explain a phenomenon when it is included within an ordered whole of relations, getting explanation as a precise and specific way of ordering and relating those functions. Here there is no psychologistic element, since the fact that a certain structure is explanatory does not depend on our subjective understanding of it, but rather on the objective relationships constituting a theory. We can therefore

[1] Kitcher and Salmon 1989, Ch. 1.1.

use a deductive-nomological basis to single out the specific form of the explanatory relationships, but then arguing that it is the functional connection with the entirety of the elements of a theory that actually provides us with an explanation[2].

The advantages of this perspective will be explored further in these sections. The aim of this last chapter is that of assessing how the functional account that I have introduced in the previous sections can be used as a framework for scientific explanation. An important clarification, which I briefly mentioned earlier, consists in the fact of remembering that Cassirer does not provide us with a theory of scientific explanation. In his works, he assumes explanation as a common term that does not require a particular further examination. Nevertheless, it is quite easy to get a sense of how he conceives explanatory relations between theories from his discussion of how theories are structured. The main point that I am highlighting is that explanation turns out to be the relation itself between the constituents of a theory. In what sense do we say that an event instantiates a law? Let us not consider the event as the individual that has to be traced back to the law with some inductive procedure. Rather, if we consider the *ensemble* of the empirical elements (observations, experiments, collections of singular events) as a constituent of a theory, already connected with its other constituents - principles, symmetries, laws, theoretical assumptions - then we understand in what sense a law explains the event. The explanatory relation is that connection of the event (and class of events) within the theoretical structure of the theory. Such a framework, however, needs to be bolstered by a thorough examination of some relevant arguments in the discussion about scientific explanation. This will be the purpose of the first section of this chapter.

The second issue concerns the low-hierarchical representation of scientific theories (that I have also sometimes called non-hierarchical). This issue is deeply rooted in the functional perspective itself, and is certainly one of the conceptual premises of the framework of explanation that I am suggesting. Indeed, we can only establish that sort of explanatory connection between symmetries, principles, laws and events only if we first conceive theories in this low-hierarchical structure. What I am arguing is that, in the functional framework, we can certainly establish explanatory relations within a theory. We cannot, however, draw the conclusion that these relations entail a correspond-

[2] «If we represent mathematically a given totality of observations by the "superposition" of several series, we do not indeed thereby increase our knowledge of the absolute and transcendent causes of the process; but we have raised ourselves to a new type of knowledge» (Cassirer 1923 [1910], pp. 263-264).

ing ontological priority of some elements over other ones. Therefore, the last part of this chapter will address the issues related with hierarchies between the constituents of physical theories. Please note that in this discussion I am not denying that one cannot establish levels based on different parameters and criteria, *e.g.* a higher level of generality for laws than for singular events. The point I am making is that one cannot build these levels on the explanatory dependencies that we can single out in a theory. Explanation is not a criterion for a taxonomical classification, but rather a functional connection between terms playing different roles within a physical theory.

For the moment, I would like to focus on some issues and examples that may help us giving more concreteness to Cassirer's model. A first point we can briefly address is the fact that within this structure we can no longer worry about those theories that conflict with familiarity views[3], as I have noticed above. It does not matter, in fact, whether or not quantum phenomena are explained in terms of a connection to more familiar concepts that make us understand them. Such a perspective carries along dangerous traces of psychologism and compels us to conclude that the more counterintuitive physical theory are not explanatory at all. In contrast, building the explanation of certain phenomena on the functions of the theories and series of theories allows us to get rid of any reference to our subjective understanding of laws and observations.

Another significant example that may help us to understand functional explanation is the flagpole case. This is one of the standard counterexamples to the DN model[4], aiming to prove that the deductive-nomological structure cannot capture the real meaning of scienific explanation. The argument runs as follows. Consider a flagpole positioned on a flat ground in a sunny day. Due to the sun rays, the flagpole casts a shadow on the ground. We need two pieces of information to know the length of the shadow: the height of the flagpole and the elevation of the sun above the horizon. Given these data, we can say that the shadow has a certain length because of the height of the flagpole and the solar incidence angle. This is a case of explanation that can perfectly fit into the DN model. The nomological component of the *explanans* consists of the laws of geometrical optics (in particular, the rectilinear propagation of light through a material with constant refractive index), to which we add the further condition about the height of the flagpole and the information about

[3] See Chapter 1.2.
[4] Salmon and Kitcher ascribe it to Bromberger (Kitcher and Salmon 1989, p. 47).

the position of the sun. Then we have a rigorous deduction of the length of the shadow with some basic trigonometry. We seem to have a straightforward case of scientific explanation. The problem is that the DN model can be employed exactly in the same way to describe the opposite explanatory procedure. If we decide to consider as the nomological part the same laws of optics but change the other two elements so that we have the length of the shadow as a condition and the heigth of the pole as the *explanandum*, this can still fit in the DN model as a correct scientific explanation.

The usual argument to illustrate what seems, *prima facie*, a fallacy is the fact that the DN model cannot account for causal relationships. One may also put it in terms of temporal sequences[5], but the fundamental reason to reject this counterintuitive outcome is the fact that the flagpole causes the shadow and not *vice versa*. However, I do not think that this point should be acceptable in a philosophical theory of explanation. This resorting to causation is actually only a way to reintroduce a psychologistic element under false pretences. The problem is that, from a common point of view, we think to this sort of explanation as a nonsense. We are used to explanations that, to a certain extent, rely on causal elements. However, these considerations should not play any role in the definition of scientific explanation. We can perfectly imagine a situation (*e.g.* a computer simulation or another kind of research based on satellite images) in which, knowing the laws of optics and the shadows projected on the ground of a city we have to get the information about the height and position of each building. In this case, it would be legitimate to argue that the height of the buildings are explained by the length of the shadows.

We have already seen in Chapter 1.2.1 that the DN model admits as explanation also derivations that are quite different from our common way of conceiving explanation. And this is to a large extent due to the lack of constraints that would adapt it to various kinds of explanation. However, the flagpole example is not a counterargument at all. In order to understand this point we have to shift to functionalism. In the functionalist account of explanation, we have an explanatory relation when there is a connection between relevant part of a theory, including the empirical observations that we are making. In the flagpole case, we have some information about the pole, the shadow and the position of the sun, and all these data are connected through a theoretical context comprised of

[5] «The height of the flagpole explains the length of the shadow because the interaction between the sunlight and the flagpole occurs before the interaction between the sunlight and the ground. [...] the temporal relation is wrong» (Kitcher and Salmon 1989, p. 103).

laws and other assumptions. Among these assumptions there may be the fact that the explanatory relation that we are looking for proceeds from the length of the shadow to the height of the building, in the case we are performing *e.g.* some kind of study on satellite photos. As long as there is a functional connection between some theoretical elements and other experimental results or empirical data, we can identify an explanatory relation. From this functionalist perspective, therefore, it is perfectly fine to argue that the direction of explanation is twofold, and the flagpole objection vanishes.

Before moving on to the discussion about the non-hierarchical form of functionalist explanation, there are two last issues related to the wider theories of explanation that we may take into account. A first point that can bring in an important disambiguation is that of the pragmatic account of explanation. The crucial point made by Van Fraassen is that explanation is not «a relationship between a theory and a fact», but rather «a three-term relation, between theory, fact, and context»[6]. When we ask a "why-question", *i.e.* we ask for an explanation, the context determines both the *explanandum* out of a contrast class (containing the alternatives to the explanandum) and the relevance relation between the *explanans* and the *explanandum*. The outcomes of this perspective are particularly interesting. For instance, in the tower example (an alternative to the flagpole), Van Fraassen draws the same conclusions as me. It is the context, in fact, that determines the explanatory relation and therefore the direction of explanation from the tower to the shadow or *vice versa*. However, the general conclusions that Van Fraassen draws are quite different. The major difference is about the fact that Van Fraassen understands the context in a much wider sense than functionalism, ending up in claiming that there is nothing distinctive in scientific explanation as the scientific component is not in the form the explanatory argument but rather in the field where we get the data needed for the explanation[7]. Conversely, the context that functionalism admits is only the context of the specific scientific theory we are working with[8]. This is the reason why one cannot properly talk about pragmatism while interpreting functionalism. While in fact the explanatory relations depend on the specific connections of elements within a theory or between more theories,

[6] Van Fraassen 1980, p. 153.

[7] See quotation in Chapter 1.2n (Van Fraassen 1980, p. 155).

[8] The context can also admit the multiple theories that we can relate through functional connections (see Weatherall's example below) or the connection of a theory with empirical data as in the flagpole example.

the contextual element is limited to this sphere and not to a wider notion of explanation involving assumptions *e.g.* outside the scope of science.

On a sidenote, when Hempel writes that he works towards excluding «psychological or pragmatic aspects of explanation»[9], he is associating the pragmatic element with subjectivity. If pragmatism means that explanation is based on psychological understanding or subjective intents, Hempel is perfectly right. However, this is not the case of functionalism. We shall see in the next section an example discussing the status of inertial and gravitational mass in Newtonian and General relativity. The point is that we can explain the distinction between inertial and gravitational mass in the Newtonian theory of gravitation with a derivation from General relativity (where gravitational mass does not make sense). This kind of derivation is entirely rigorous and does not rely on any psychological sense of explanation. Nevertheless, we introduce contextual elements as the different understanding of the concept of mass in the two theories. Or, in a different way, we refer to a context when we aim to provide an explanation of some specific experimental results within a certain theory. One can appreciate, however, that these last examples are quite different from Hempel's conception of pragmatism[10].

As a second point, one may have thought to Friedman's unificatory account in some passages of this work. It is therefore worth spending a few words[11]. In particular, one may think to unificationism when I say that the functional account of explanation is based on the connection that we establish between the constistuents of a theory. Certainly there is a familiar flavour when Friedman talks about the "global view" that one should endorse, defining scientific explanation as that unificatory connection that reduces the number of independent phenomena (let us assume the word *phenomena* in a general sense, including also laws governing sets of singular phenomena)[12]. First of all, though, functionalism is not limited to a claim about the systematisation of principles, laws, empirical findings in more general functional connections, but also with all the conceptual transformations that we have when we develop new theories. Moreover, most importantly, functionalism

[9] Hempel 1965, p. 413, see also Chapter 1.3.1.

[10] As Van Fraassen correctly claims about the reasons of this rejection of pragmatism, «the positivists tended to identify the pragmatic with subjective psychological features» Van Fraassen 1977, p. 150.

[11] Even if the most substantive development of unificationism was made in Kitcher 1989, my aim is not to assess the whole discussion about unificationism, but rather only to mention some points relevant to the comparison with functionalism.

[12] Friedman 1974, p. 18.

does not identify a one-directional direction of explanation based on the unification of phenomena under wider theories, but rather builds more complex connections of terms also allowing mutual explanation relationships within theories. It seems, therefore, that the unificatory account is limited to a narrow part of what functionalism aims to capture.

3.2 Dependencies and Hierarchies

The cases I have analysed in the first part of this chapter give us a better understanding of how an account of functional explanation based on Cassirer's thought can fruitfully engage with the contemporary debate. However, from the beginning of this work the whole discussion has always connected two focal points: the notion of explanation and the structure of physics. One cannot properly examine the former without addressing the latter. What I would argue is that while analysing Cassirer one can derive quite straightforwardly a theory of explanation only starting from his depiction of scientific theories. This suggests that any notion of explanation is not only deeply rooted in every specific account of how physical theories work, but it is actually a direct consequence. If we reconsider the scheme in Chapter 2.2.3 we see that Cassirer conceives physics as comprising three broad categories of statements: statements of principles, statements of laws, statements of the results of principles. There are, therefore, three categories each one playing a peculiar role and covering the wide spectrum of what we admit in our physical theories. I have already underlined that this distinction does not imply any kind of hierarchy between these terms. Even if I discuss this in more detail in the last section of this chapter, it is important to notice that the results of measurements play the same role as laws and symmetries in this depiction. We can find the reason for this in the fact that for Cassirer the distinction between theory and experience is blurred. The empirical results are part of the wider sphere of a theory, in fact Cassirer goes as far as to say that «everything significantly factual is already a theory»[13]. This is a peculiar element that we should always bear in mind while considering claims such as "laws explain events". In what sense should laws explain singular cases? Explanation is something more than just the fact that a laws describes and predicts series of regular events. However, as soon as we assume the perspective according to which experimental results have a theoretical structure as well, we can read this explanatory relation as a plain functional connection.

[13] See again Chapter 2.2.1.

The functional depiction of theories also gets around other problems of generalisation and induction-based generalisation. One of the difficulties in a Humean conception of laws is the fact that there is no guarantee that the laws that we draw from our empirical observation have any kind of coherence and can be integrated in wider theoretical systems. More generally, the whole question of how we have an evolution of theories and a historical development of science is something that needs a much deeper account of how laws are connected to the rest of scientific theories and how concepts can evolve. Another crucial point, in fact, is that of the evolution of concepts (and corresponding objects) in different theoretical contexts. The comparison with Structural realism I have made in Chapter 2.3 has layed the basis for this issue, enabling us to know how to face the problem of the continuity of concepts through theory-change. However, it may be helpful to address James Weatherall's perspective on this topic (Weatherall 2011), as he provides a clear instance in which we have an explanation of the change in concepts through different theories.

Weatherall starts from an example from Newtonian physics and asks why it is the case that, within the context of the Newtonian theory of gravitation, inertial mass m_i is equivalent to gravitational mass m_g in spite of being theoretically different. What he wants to achieve is an understanding of the kind of explanation that one can obtain in this case and compare it with some models, in particular the Deductive-nomological and the Unificatory account. To get to the nitty-gritty of his argument, the explanation runs as follows. Within the framework of General Relativity, m_g coincides with m_i (Einstein's equivalence principle) or, to be more accurate, it makes sense only to consider m_i. Why then are we still distinguishing between m_g and m_i in the Newtonian theory of gravitation? Suffice to say that when we consider Newtonian theory as a limit of GR, one can then derive the equivalence of m_g and m_i. This is done through an intermediate stage consisting in the passage from Newtonian gravitation to Newton-Cartan theory, which is based on a curved spacetime that does not account for m_g, just as in the case of GR.

His first remark is against the applicability of the DN model to this explanation. Even if there are some very general elements that would make the explanation fit in a deductive-nomological structure, I agree that in a case like this the DN model does not make entirely perspicuous why we have an explanation - assuming, as I am doing, that this is actually an explanation[14]. The main

[14] Weatherall 2011, p. 434.

reason for this is the fact that the DN model would not capture the explanatory element consisting in the transition through the concepts of mass in the two theories, but would rather only provide some very general conditions for that explanation to happen. However, I am mostly interested in what makes the previous line of argument explanatory in Weatherall's account. The crucial point is the passage in which we perform the limiting procedure through which we recover Newtonian theory from GR. In broad outline, this limit works by associating a solution of Newtonian gravity (in particular, Newton-Cartan theory) with a solution of GR[15]. By doing this we have connected the two theories so that considering gravitation as a force within Newtonian theory we then say that $m_g = m_i$. This claim is critical for us: «The explanatory work, then, is done by presenting the details of the relationship between the two theories»[16].

Our functionalist approach accounts for this explanation in terms that are quite close to Weatherall's. Consider the Newtonian theory of gravitation as a certain functional system of laws $\psi_1(\alpha_1, ...)$, $\psi_2(\beta_1, ...)$ governing some phenomena as described above. For the sake of clarity, we can mention the law of universal gravitation (ψ_1) and the theorems and equations for central forces (ψ_2). General Relativity is then another theory of this kind $\phi(\psi_1, \psi_2)$. The distinction between m_g and m_i, which only makes sense within the former theory, is maintained as valid and explained by turning to the limiting procedure that we have seen. It is this functional connection that makes Newtonian theory valid within its domains and let us put into correspondence the concepts of mass in the two theories. In this case, therefore, the kind of functional connection that we have identified is coherent both with the formal and conceptual meaning of functionalism that I have outlined in Chapter 2.2.2. The limiting procedure allows us to get through a mathematical procedure from the concept of mass in Newtonian theory to mass in GR. If we consider, instead of the groups of laws ψ_1, ψ_2 objects that are among the constituents of Newtonian theory, namely m_g (γ_1) and m_i (γ_2), then we have the limiting procedure providing us with an explanatory connection of this kind $\phi[\psi_i(\gamma_1, \gamma_2)]$. Which means that the gravitational and inertial masses are explained in the context of General Relativity.

Weatherall's last thoughts are about the evolution of theories and the role that explanation

[15] This is done by widening the lightcones, as described in Weatherall 2011, p. 443.
[16] Weatherall 2011, p. 437.

plays in connecting old and new, more comprehensive theories. I entirely agree with his critique against Kuhnian perspectives, which instead insist to establish an incommensurability of theories[17]. A framework that can account for the continuity of theories, in fact, is much more faithful to the historical development of physics and depicts quite carefully the relationship between different theoretical structures[18]. The kind of functionalism I have put forward works in this direction. Not suprisingly, Cassirer uses the principles of Newtonian mechanics[19] as an example not of «absolutely unchanging dogmas», but rather as «the temporarily simplest intellectual "hypotheses," by which we establish the unity of experience»[20]. In the event of findings that need a conceptual revision, the unity of experience represented by fundamental series of the form $F[\phi_1(\psi_1, \psi_2), ...]$ will be maintained coherent modifying the functional structure, *i.e.* establishing new connections and explanatory relations between laws and events.

So far, I have stressed that this functional tree is not hierarchical, since it needs to be always flexible to revisions and modification. One can argue that these two points are not necessarily incompatible: we may think of a structure which is ordered in different levels of generalisation or explanatory dependencies and nonetheless be potentially subject to change. This is certainly true, and in fact Cassirer's structuralism conceives theories precisely in this fashion. I would not call such a system strictly hierarchical. As I have explained in Chapter 1.1.2, what I mean by hierarchy is a fixed ontological structure where certain elements *e.g.* symmetries have a place established through a one-way relation with other elements.

However, the example of Newtonian mechanics and General Relativity gives evidence of different levels of generality. It may be tempting, therefore, to think that some sort of hierarchy based maybe on the explanatory relations can be still established. It is useful to consider some proposals going in the direction of an explanatory priority of symmetries over laws. This discussion is also important to understand in greater depth the role of counterfactuals in scientific explanation. The next step will be that of assessing to what extent it is legitimate to distinguish symmetries from other

[17] Weatherall 2011, p. 439. On this point, one may correctly argue that mathematical continuity between theories does not entail conceptual commensurability. However, both Weatherall and I are criticising a radical understanding of incommensurability, which can be easily overcome in a framework such as the functionalist one, able to account also for significant changes between theories (see the three components of Cassirer's functionalism in Chapter 2.2.2).

[18] Besides capturing the core of the explanatory process, unlike the DN model.

[19] Remember that *Substance and Function* was published in 1910.

[20] Cassirer 1923 [1910], p. 268.

laws of nature, and which meaning we assign to the latter. The functional paradigm will provide us with a new ground of interpretation.

In order to best understand what one means by saying that symmetries *explain* laws of nature - a choice I reject, opting for a functional connection between the two elements - it is worth considering French and Saatsi 2018. The authors' aim is to use Woodward's account of counterfactual-dependence for the case of symmetries, in order to justify an explanatory priority over the laws and phenomena they would cover. Apart from the fact that French and Saatsi try to get rid of the causal structure underlying Woodward 2003, the use of counterfactuals is the same.

Again, the scheme is based on *explanans* and *explanandum*. In what sense does the *explanans* explain the *explanandum*, rather than vice-versa? Woodward identifies the main feature of explanation as consisting in «systematic patterns of counterfactual dependency»[21]. This dependency rests on the fact that, assuming a previous explanans-explanandum connection, if the *explanans* is changed, this change is reflected on the the phenomenon to explain. This link can be represented as a what-if-things-had-been-different question. If we consider any physical phenomena governed by a law[22], we expect it to change depending on the covering law. To speak more generally, the modification of the initial conditions represented by the *explanans* determines a change in the explained phenomena. Here we detect the asymmetry between the two sides of the explanation, and understand why a pattern based on counterfactual suppositions (what-if-things-had-been-different questions) establishes a dependency. In this perspective, in fact, it is not the case that if we pick a different phenomenon we get a resultant changing in the covering law[23].

The toy example that French and Saatsi put forward consists in a balance where two weights are in state of equilibrium. The natural way of addressing this case is by saying that there is a symmetry between the forces applied on the two sides of the balance so that $T_1 = -T_2$, *i.e.* the net torque on the pivot is equal to zero. As a consequence, the balance is in a state of equilibrium. Here we can apply the description based on counterfactual-dependence and argue that while it is possible to

[21] Woodward 2003, p. 191.

[22] I am assuming the term *govern* in its most general and intuitive meaning.

[23] With respect to physical symmetries, this is very similar to Wigner's approach in the example of the modification of Newton's law that does not correspond to any modification in the homogeneity and isotropy of space and time Wigner 1954, p. 437.

modify the net torque and accordingly have a change in the equilibirium state, the opposite procedure is not possible (this is the what-if-things-had-been-different question). In fact, if we wanted to change the equilibrium state we should modify the forces involved, and therefore act on the net torque anyway. Therefore, this asymmetry amounts to an explanatory dependency of the state of equilibrium from the state of the torques acting on the balance[24]. As this case involves a clear causal relation between the two elements, this line of argument is then extended in the following examples also to continuous symmetries without causal factors in the explanatory process. The last example considered is that of the correlation between Noether's theorem and conservation laws. French and Saatsi accept Brown's argument that, in itself, Noether's theorem only accounts for a connection between conservation laws and corresponding symmetries of the Lagrangian[25]. However, the argument in support of the explanatory priority of symmetries over laws should not rely on Noether's theorem. The starting point should be the fact that the change in the symmetries of the Lagrangian of a system results in a modification of the dynamical properties of the system[26].

However, the whole discussion becomes intricate when we get to an explicit characterisation of this causal dependency. While asking why we should think of a counterfactual dependence in the case of dynamical symmetries, French and Saatsi claim that «physicists often give explanatory priority to symmetries over conservation laws [because] in analogous applications of Noether's theorem to particular subsystems of the universe [...] the explanatory priority is transparent», and therefore this is extended «to symmetries of the laws covering the whole universe»[27]. This does not sound quite satisfactory. We are still wondering how the passage from the counterfactual dependency to the explanatory one is performed. One can value the fact that, unlike Lange 2007, French and Saatsi do not try to build a metaphysical account of how symmetry should be conceived as meta-laws. Their proposal, therefore, is not focused on a hierarchical representation of symmetries, laws and events. Rather, the aim of their work is that of capturing explanatory relations through counterfactuals. Moreover, it is very interesting how they overcome the fact that Noether's theorem does not provide a one-directional explanatory relation. They consider the theorem as a wider framework in

[24] French and Saatsi 2018, p. 188.

[25] See French and Saatsi 2018, p. 197 and Chapter 1.1.3.

[26] The example given is that of the changing in the potential of a system and the subsequent modification in the moving bodies (French and Saatsi 2018, p. 199).

[27] French and Saatsi 2018, p. 200.

which, due to the connection established between symmetries and conservation laws, allows us to pose the what-if-things-had-been-different question in the specific cases.

The problem of this approach is that it assumes the fact that a counterfactual dependency corresponds to an explanatory dependency without questioning it. In the toy example, the authors claim, «there is a natural sense in which the equilibrium depends on symmetry»[28]. They then follow the same line of argument, as we have seen, for the case of the relation between symmetries and conservation laws. However, when they get to the point of upholding the identification of the counterfactual modal asymmetry with explanation, one cannot find a satisfactory argument. I deem that the reason for this lack lies in an assumption that we notice also in Lange 2007. In particular, while building his hierarchical system, Lange claims that he wants to prove that «the relation between ordinary (i.e., first-order) laws and the facts they govern (a relation involving counterfactuals) may be reproduced one level higher: as a relation between symmetries and the ordinary laws they govern»[29]. The aim of these kind of arguments based on counterfactuals is precisely that of establishing the same relationship between laws and facts in the higher level of the connection between symmetries and laws. However, we have seen that the concept of explanation captures a broader range of connections. For instance, let us consider again Weatherall's example of the fact that we can derive the Newtonian distinction of inertial and gravitational mass starting from General relativity. Nobody would object that this is a form of explanation. Nevertheless, it does not seem to me that the modal argument can capture this connection at all. This is because from a functional perspective we define explanatory relations as those connections that we single out within a system that may consist of symmetries, laws, empirical observations (*i.e.* collections of singular events). The study of the modal dependencies between these constituents is a different issue and does not necessarily entail that these modal connections have also a significance in terms of explanatory relations. Explanation is much more than a list of modal connections, it is a complex procedure of integration of theoretical and empirical elements in view of demands of clarification, systematisation and development of scientific theories. The lack of a clear argument in favour of the identification of counterfactual dependencies with directions of explanation in French and Saatsi 2018 is a proof of

[28] French and Saatsi 2018, p. 189.
[29] Lange 2007, p. 1.

my last considerations. Of course I am not arguing that this approach is worthless. Actually, it is particularly important to set out relevant connections that may be used also for explanatory purposes. My point is that, in order to obtain a comprehensive account of explanation we should move away from the strictly modal perspective.

In these last remarks I have recalled functionalism and the conception of scientific explanation based on functional connections. An interesting way of reconsidering Cassirer's perspective in the light of the points discussed in this chapter is by focusing a bit more on this point through a relevant position in the contemporary literature. Weatherall 2017, in particular, addresses the issue of the correlation between principles of a theory through a comparison of General relativity and Newtonian gravity. The starting point is the discussion about inertial motion, which is governed by the law of inertia in Newtonian physics and the geodesic principle in GR. While in the first theory we say that a body maintains its state of rest or uniform motion (constant velocity) if no external forces act upon it, in GR we say that the spacetime trajectories of the body are geodesics. However, the status of this principle is everything but unambiguous. For instance, the fact that both in GR and Newton-Cartan theory one can prove the geodesic principle as a theorem poses some intriguing questions about the explanatory role of the principle. First of all, the fact that one can prove the geodesic principle also in Newton-Cartan theory deprives GR of its alleged superiority with respect to the explanation of inertial motion. However, the crucial point is whether we can actually say inertial motion is actually explained in GR or Newton-Cartan theory.

Weatherall's point is that if we take a traditional stance with respect to the foundation of physical theories, the fact that the geodesic principle can be proved should induce us to move it from the level of the fundamental principles of the theory (take GR or Newton-Cartan) to a lower level of theorems and derived principles. Weatherall calls this view of the foundations of a theory "a two-tiered system"[30], because it is built in such a way that we characterise a theory as divided into two levels of fundamentality, with some principles and assumptions in the higher and other derived theorems in the lower one. The problem is that there are cases in which the geodesic principle can be used to derive information about matter and energy conditions which, in turn, should be among

[30] Weatherall 2017, p. 29.

the fundamental assumptions of our top-tier.

Is it correct to hold such a hierarchical view? Weatherall deems it is not. Rather, he puts forward a "puzzle-ball conjecture". In this view, we consider the geodesic principle and all the other theorems, principles and assumptions involved in the foundation of a theory as intertwining puzzle pieces in a puzzle-ball. One does not have levels of fundamentality, but rather a whole sphere of principles where each modification of one of these assumptions affects the whole structure of the theory foundations.

> Instead of thinking of the foundations of a physical theory as consisting of a collection of essentially independent postulates from which the rest of the theory is derived, one might instead think of the foundations of a theory as consisting of a network of mutually interdependent principles — a collection of interlocking pieces, as in [a] spherical puzzle[31].

This should strike us as remarkably similar to Cassirer's example that we saw in Chapter 2.2.3. It is worth reading the passage in which Cassirer suggests this picture.

> If we chose a spatial analogy for the structure of physics, we must not liken this structure to a pyramid resting on a broad base of immediately given and independent "facts", rising gradually from this and ending in a highest point, perhaps in a simple "cosmic formula". [...] Physics accordingly is to be compared [...] to the "well-rounded sphere" with which Parmenides described his "universe". [...] This cannot and must not prevent us, to be sure, from ascribing to it a determinate structure, a higher and a lower order of elements; but this logical distinction of the elements must not be interpreted as suggesting the possibility of their actual separation, their independent existence[32].

There are some fundamental differences between the two views, of course. First of all, Weatherall only refers to the foundations of theories, while Cassirer's point concerns the totality of what relates to a theory, including also the empirical part: collections of data, experimental results and

[31] Weatherall 2017, p. 31.
[32] Cassirer 1956 [1937], p. 35.

observations. His spherical view, therefore, is based on the three elements that we saw in Chapter 2.2.3 and discussed at the beginning of this section: principles, laws and empirical results. Despite having this broader scope, the way in which Cassirer considers the connections between these constituents is strikingly similar to Weatherall. For instance, the non-modularity of principles within theories, *i.e.* the fact that one cannot change a principle without affecting the other parts of the theory, is also advocated by Cassirer[33]. Also Weatherall's remarks about Quine assist us in understanding this approach. We cannot modify the peripheral beliefs and maintain a central core unaffected, as any modifications in a theory have radical implications on the other parts, without a hierarchical distinction between most and less fundamental ones. Weatherall and Cassirer's approach is more structural rather than holistic in Quine's sense[34].

So how does explanation fit in all of this? Weatherall claims that he is not providing a theory of explanation here. His point, rather, is that when we argue that certain principles as the geodesic principle have an explanatory role, we should not build this role around the fact that there is a deductive priority between two elements. It is not the case that the geodesic principle is not explanatory only because it can be derived from other assumptions. Explanation is not a one-way derivation of elements from higher-level principles, but rather the connection of different elements in a theoretical structure where the direction of explanation is based on contextual elements.

Weatherall briefly mentions that this approach is very similar to the perspective of Feynman 1965. In this work, Feynman endorses what he calls a "Babylonian view", opposed to a "Greek view"[35]. The Greek view is a conception of rigorous science based on several axioms or fundamental principles from which one can derive the whole body of a theory. The Babylonian view, conversely, is a non-hierarchical conception where the different elements constituting a theory are connected in such a way that the theory, rather than being a chain of derivations, is precisely the whole of these relations. Certainly, this is just an intriguing analogy and Feynman himself does not expand these insights. However, all this discussion offers us the possibility of getting back to the issue of hierarchies and exploring a few more points about functional explanation.

[33] He denies the «possibility of imagining the higher layers removed without destroying the bottom layer or even altering it essentially» (Cassirer 1956 [1937], p. 35).

[34] Weatherall 2017, pp. 35-36.

[35] Feynman 1965, pp. 46-47.

3.2.1 Low-hierarchies and Functional Explanation

The last three cases that I have analysed are extremely relevant as they call into question three crucial issues of our discussion.

- Continuity through theory change, which we have explored in particular through the relation between functionalism and structural realsim.
- The explanatory relations between different constituents of theories (principles, symmetries, laws, events), which I have addressed from the perspective of a structural arrangement of physical theories.
- The low-hierarchical (or non-hierarchical) account of those theories.

At the beginning of the previous section I highlighted how these last two issues are deeply interrelated. I deem that, by focusing a bit longer on these last points we can set out a few thoughts about the outsanding issue of this work. First of all, after the last discussion of symmetries, it is interesting to go back to Chapter 1.1. Symmetries have been the starting point of our work because they are quite often called into question to build hierarchical models. Wigner's point of view represents one of the most extreme positions on this topic. However, also Marc Lange builds a model with very strong ontological intents, aiming to account for a hierarchy based on counterfactual dependencies. In my discussion of French and Saatsi 2018 I pointed out that, in my opinion, the problem of these approaches is the fact that they are very often based on an illicit extension to the scope of explanation. Let us set aside Wigner, as his proposal is nothing more than an outline an therefore does not contain enough material to get a precise sense about what kind of hierarchy he proposes. However, in Lange we find the same kind of improper extension that we have encountered in French and Saatsi 2018. His whole endeavour consists in proving through a structure of modal dependencies that a «symmetry principle constitutes a requirement on other laws rather than a byproduct of them»[36]. Regardless of whether this objective has been achieved, why should we derive from it that symmetries also possess an explanatory priority? As long as the connection between these hierachical models and a theory of explanation is thoroughly illustrated, there is no plausible reason to take this connection for granted.

[36] Lange 2007, p. 458.

Please note that I am not arguing that there is no difference between *e.g.* laws and accidental regularities or series of events. I consciously decided not to engage with the problem of defining physical laws as it would have been outside the scope of this work. However, with the exception of naive regularity theories, any theory of scientific laws accounts for a difference between laws and accidental regularities. The same thing holds for the difference between symmetries and laws. Cassirer himself distinguishes very clearly the statements of principles, among which we can list symmetries, from laws. He actually uses a terminology that may raise concerns about a relapse in a sort of hierarchical view: «Principles do not stand on the same level as laws, for the latter are statements concerning specific concrete phenomena. [...] [Principles] refer not directly to phenomena but to the form of the laws according to which we order these phenomena»[37]. Cassirer accepts precisely that function that we mentioned in several points in Chapter 1.1[38], namely the fact that symmetries act as guides to discover relevant physical features of systems. The reason why in Cassirer we do not find an ontological conclusion from this is that he falls neither in the fallacy of deriving theoretical priorities from mere heuristical procedures nor drawing ontological conclusions involving the notion of explanation. My point about this whole issue is that if one wants to build any hierarchical model characterised by levels with different explanatory power, this cannot avoid a preliminary account of how explanation comes into play and is integrated in the model.

What I have done in my proposal, instead, is precisely to first propose an ontological model and then build a theory of explanation upon that. The functional account of explanation that I have put forward relies on the ontology of functionalism, which has precise characters that I have outlined in Chapter 2.2. This ontology hinges on two fundamental features: (i) a low-hierarchical representation of physical theories, (ii) an eliminativist conception of physical objects. Both these features depend heavily on the Kantian premises of Cassirer's thought. Cassirer starts from this philosophical background but extends it in such a way that he builds a very lean way of conceiving apriorism[39] and addressing physics. Obviously, this ontology may be further characterised, even if I decided to restrict this analysis to the aspects relevant to the issue of explanation. However, it is clear

[37] Cassirer 1956 [1937], p. 52

[38] To refer to the relevant literature, Ismael and Van Fraassen 2003, , Wigner 1964, Earman 1989 *inter alia*. Obviously this also includes French and Saatsi 2018.

[39] As he writes, an apriori that «must be understood in a purely methodological sense» (Cassirer 1956 [1937], p. 74).

that this ontology should not be defined realist in the strict sense. Cassirer does not refer to reality, but to objectivity. And objectivity is not an attribute of the world, but rather of our knowledge of the world (another apparent result of the Kantian premises of this approach). From this perspective, what we have is not reality *per se* but a theoretical construction whose objectivity stems from the relations through which we have our experience. I do not want to expand on this point. I am not interested in a comprehensive discussion of Cassirer's ontology, because the points through which I have characterised this ontology are sufficient to develop the relative theory of explanation. For these last considerations, therefore, we can spend again a few words on the first of two features I have mentioned above.

Why am I laying emphasis on the point of the low-hierarchical character of this conception of physics? There are several reasons that I have explored and will also briefly recapitulate in the Conclusion. However, a last point I would like to focus on concerns again the historical evolution of physics. In the last section I have briefly mentioned the fact that Weatherall distances himself from a perspective in which the puzzle-ball would be part of a web of beliefs and therefore it could be modified in some of its parts, even if in different measure depending on the distance from the core of the structure. However, so far I have always talked about the fact that Cassirer's spherical representation of theories and physics in general accounts for the flexibility needed to understand the evolution of concepts and the structural continuities through theory change. One may therefore ask which one of the two views I am supporting. This point is crucially important. In my opinion, the very fact of conceiving theories as spherical structures - in the sense I have outlined above - is what ensures the continuity of explanation between theories and the evolution of their structure. Non-modularity, *i.e.* the fact that we cannot change a part of a theory without affecting the whole system of related theoretical elements, is not at odds with theory change. Indeed, it is precisely the requisite that we need in order to maintain explanatory continuities. Non-modularity is not a constraint in terms of the flexibility of theories, but rather the guarantee that we can always trace back theories to previous theoretical structures, relate objects in our current ontology with concepts employed in previous ones, and so on and so forth. What do we really obtain from an hypothesis such as Wigner's, namely that of modifying the laws of motion and proving the fact that the isotropy of space and time is unchanged, in order to prove that symmetries have a higher hierarchical position?

It is an interesting thought experiment, but one from which we cannot derive much relevant information *e.g.* to understand the explanatory relations between symmetries and laws. Conversely, a non-hierarchical view allows us to use the functional connections between the elements of a theory or the concepts and structures across theories in order to get explanatory relations and therefore a better grasp of the theories themselves.

Again, these last remarks demonstrate to what extent the account of explanation that we endorse depends on specific ontological assumptions. I am not using "ontological" in a strong, *i.e* metaphysical sense, but rather to indicate the structure of our scientific theories and the way in which objects, terms and functions are employed. From this perspective, it is clear how a low-hierarchical account of theories and physics inherently leads to a different explanatory depiction from a hierarchical one.



Conclusions

The work I have done so far aims at assessing how functionalism can provide an account of explanation in the light of the recent debate about this subject. To my knowledge, this is a novel perspective on the issue. In order to deal with the different themes and problems intertwining within this issue, I worked on several levels. There are a few topics constituting the backbone of this project that I addressed from different perspectives throughout this discussion. Instead of going through the whole work and providing a summary, I think that the best way to draw the conclusions is by considering these topics individually.

1. *Hierarchies*. The usual representation of a hierarchy is a pyramid, or some sort of oriented line. The first part of this work starts from a discussion of the notion of symmetry right because I think that it is one of the elements most likely to be found in a hierarchical depiction of theories. My aim is not that of addressing the specific problems related to the interpretation of symmetries, but to understand what kind of hierarchical structures can be build to relate symmetries with laws and events. This priority can be based on strong metaphysical assumptions or on different considerations about the role of symmetries *e.g.* as guides to understand the behaviour of dynamical systems, systematise and discover laws. Here I have considered different hierarchical accounts. On the one hand, one can find Wigner 1964 and Lange 2007 who propose quite rigid systems. On the other hand, French and Saatsi 2018 work on Woodward's model of counterfactual dependency but without drawing strong metaphysical conclusions. As opposed to this pyramidal views, I suggest a spherical representation of physical theories derived from Cassirer 1956 [1937] and Weatherall 2017. The different constituents of theories are principles, symmetries, laws, assumptions and collections of empirical results. These elements may certainly be ordered with some criteria, which is why I write about a

low-hierarchical model. However, from the point of view of the explanation dependencies this representation of theories is entirely *non-hierarchical*.

2. *Explanation*. The account of explanation I put forward is based precisely on those assumptions about the non-hierarchical model of theories. If we endorse a spherical view of theories, in fact, the whole explanatory relations are built between the terms consituting the sphere (*i.e.* the components of our theoretical constructions). Therefore, explanation is neither a mere logical deduction nor a pure reproduction of another kind of hierarchy on the explanatory level. Rather, explanation represents all those processes in which we reciprocally relate objects and structures through mathematical and conceptual connections. The nature of these connections is specified by functionalism. The functional connections that we build between theoretical assumptions, experimental results, and all the sphere of elements that constitutes physical theories is what provide us with relevant explanatory relations. To establish these connections is the core of scientific explanation. More generally, I discussed several different proposals of scientific explanation in order to get a better sense of the problems and the real issues at stake, *e.g.* in the case of pragmatism in Van Fraassen 1980. Among these alternative views, the one that I focused on the most is the DN model. I had a somehow sympathetic approach to this model, as I think it still provides an useful tool to address the structure of explanation from a general point of view. Even though, as I argued, this needs specific constraints and modifications, it is an excellent model to compare and contrast with other accounts of explanation.

3. *Methodology*. The methodological issue is crucial. These first two points, the (non-) hierarchical form of theories and explanation, determine each other. This means that, while working on an account of explanation, both these levels should be kept in mind at the same time. Two considerations derive from this. The first one is that it is not possible to build a hierachical system and then simply take these dependencies as corresponding to explanatory dependencies as well. I argued that French and Saatsi 2018 is a clear example of this first mistake. On the other side, it is likewise impossible to work towards an account of explanation without first considering the underlying ontological assumptions. For this reason, in my work I

have first addressed the fundamental issues about Cassirer's ontology and then tried to build a theory of explanation.

4. *Theory-Change and Structures*. Besides the spherical structure that I have outlined, which focuses on the statical aspect of how to characterise a theory in its components, one can also consider the evolution of theories over time. Functionalism provides both an instrument to interpret the evolution of theories from a structural perspective and account for the continuity between concepts and objects of different theories. As of this latter point, Weatherall 2011 suggests an excellent example and interpret the continuity between the gravitational and inertial mass in Newtonian gravitation and General relativity in a way that is very similar to what functionalism aims to do. As far as the first issue is concerned, I discussed how functionalism can address an argument such as the pessimistic meta-induction in a similar manner as structural realism. I deem that this analogy can be particularly fruitful, as there are some common theoretical premises (and, accordingly, outcomes) that can help us in the analysis of both functionalism and structural realism. However, I have stressed the fact that one should consider each form of structural realism in detail in order to precisely understand the terms of such an analogy. There are certainly some striking common points (Ross and Ladyman 2007), but a more careful analysis leads us to discover that *e.g.* Cassirer's perspective is closer to Ladyman than to French's when it comes to a stricter characterisation of structures.

5. *Kantianism*. Neo-Kantianism could be taken as meaning everything and its opposite. However, I deem that Cassirer's thought has a central position among the interpretations of Kant of the last century right because of his focus on physics. We have seen that functionalism is directly derived from Kant's thought, first of all from the ontological point of view. The functional (eliminativist) conception of objects, the form of our scientific experience as well as the role of causality in the determination of lawfulness within events: all these crucial elements in Cassirer's functionalism have a direct interpretation in Kantian terms. However, I tried to make clear that Kantianism is not a stack of statements concerning the influence of our epistemic or operational structures in the scientific enterprise. What I argued in my discussion of Massimi's work is that Kantianism should be analysed in the most rigorous manner,

especially since the whole functionalist project depends on a proper understanding of these premises.

Certainly, this discussion only provides the groundwork for a functional project that should involve several other topics in the Philosophy of Science. For instance, one more aspect that may be explored regards the notion of law itself. As I have mentioned before, a common way of conceiving laws of nature is by saying that laws explain the events they govern. This amounts to saying that when we observe a certain event, we explain it and answer the why-question "Why does it happen in this way?" by referring to the relevant law. An event, on the contrary, is just an individual instance and does not explain a law, in the sense that we cannot answer a why-question of the kind "Why does this law work in this way?" by simply showing the event. An example that is widely analysed in the literature is that of the gold and uranium spheres[40]. Let us consider the following statements:

- All gold spheres have a mass smaller than 100,000 kg.
- All uranium spheres have a mass smaller than 100,000 kg.

I do not want to discuss whether the second statement can be properly considered as being a law. However, one cannot help but notice that the main difference between these two claims is the fact that the second has a nomological component that makes it necessary, while the first one is only an accidental generalisation. While, for the first statement, we could certainly imagine that somewhere in the universe there is a sphere with a mass larger than 100,000 kg, in the second case the critical mass of Uranium prevents it from being possible. We know, in fact, that due to the density and atomic properties of the specific material, any sphere of Uranium-235 with a mass larger than 52 kg undergoes a criticality accident and therefore a chain reaction. A common argument used to account for this difference, in fact, is based on counterfactuals. While a nomological statement "supports" its counterfactual, an accidental generalisation does not. For this reason we can imagine the larger gold sphere but not the larger uranium one.

However, this point can also be addressed from a functionalist perspective. In fact, the difference that we can actually identify in these two sentences is that the second one has a connection with

[40] See *e.g.* Van Fraassen 1989, p. 27, Salmon 1989, p. 15, and also Chapter 1.1.3.

a further assumption deriving from the atomic theory, namely the value of the critical mass of Uranium. There is, therefore, a specific theoretical context connected to the series of regular events, *i.e.* the regular lack of observations of uranium spheres larger than 52 kg. There are three components involved: the regular observations of no such uranium spheres, the corresponding generalisation and the atomic properties of Uranium. It is only when we bring our observed regularities under a functional connection of this kind that we can assess the lawfulness of the corresponding statement. From this perspective it may be correct to say that a law explain its instances, not in the sense that there is an actual direction of explanation from laws to events, but rather than this generalisation provides a functional connection of events to other factual and theoretical assumptions within a physical theory. The absence of these theoretical connections in the case of the gold sphere, on the contrary, is the reason why we do not have any law involved nor explanatory process. Certainly, this is only an insight that should be developed further, because *per se* the line of argument I have outlined does not address the point of the necessity of laws that, in turn, requires a deep discussion. In the light of these considerations, however, I suggest that an expansion of this work could fruitfully address the problems related to the status of scientific laws.

Let me conclude by addressing concerns that may arise while discussing this proposal of functional explanation. One may think that, despite the examples that I have put forward, functionalism could be not precise enough to really capture the notion of explanation. Why should this functional connection actually correspond to an explanatory relation? One can claim that some of these connections are not explanatory at all. This is certainly true. However, pay attention to the fact that I have not stated that the functional connection corresponds to the relation of explanation, but rather that we have an explanation when we build these functional relations. I am not providing a sufficient condition, but a necessary one. The fact that explanation occurs depends on contextual factors, *e.g.* the fact that we are looking for a connection between two concepts in distinct theories, or that we are working towards a different systematisation of theorems, principles and assumptions within a specific theory. This last remark may sound unsatisfactory. However, I do not think that the Philosophy of Science should necessarily look for those sufficient conditions. Explanation is a complex notion that cannot be captured *via* some formal structure, as the DN model was intended to operate. Therefore, either we accept an account of explanation that does not distinguish between

science and other areas of knowledge, as in Van Fraassen 1980, or we look for a theory of explanation that provides the instruments to interpret and understand the ways in which we have explanation in science. This is what the functionalist account aims to do, rather than providing some sufficient conditions. Again, these considerations are rooted in the underlying conception of physics. If one regards physics as a more or less systematic series of laws and observations, then the DN model with some suitable constraints may work. However, my whole argument, also through Weatherall's perspicuous analysis, aims to prove how the whole process of systematisation of theories is complex and subject to many different parameters. Feynman's "Babylonian view" is an interesting way of understanding a discipline that does not proceeds only through deductive arguments. Rather, much of the work in physics is done through the association and construction of new concepts, the reinterpretation of the role that the constituents of theories have within certain theoretical frameworks, the adaptation and modification of mathematical and conceptual structures. It would be pointless to argue that, whenever we have one of these modifications that also change the explanatory dependencies, this means that the former actually were no real explanations. A new way of arranging theorems and principles or a new connection between different theories is a new explanation in the same manner as the previous one, now replaced or interpreted in another sense.

The variety of themes and issues that scientific explanation induces to discuss is the proof of its central importance in the Philosophy of Science. I would say that, by discussing explanation, we can dialogue with the spirit of the scientific enterprise and the deepest questions of the journey of our reason in this world.

Bibliography

Achinstein, Peter (1981). "Can There Be a Model of Explanation?" In: *Theory and Decision* 13.3, pp. 201–227.

Armstrong, David Malet (1985). *What is a Law of Nature?* Cambridge: Cambridge University Press.

Belot, Gordon (2013). "Symmetry and Equivalence". In: *The Oxford Handbook of Philosophy of Physics*. Ed. by Robert Batterman. Oxford: Oxford University Press, pp. 318–339.

Bird, Alexander (1998). *Philosophy of Science*. London: Routledge.

Brading, Katherine and Elena Castellani (2007). "Symmetries and Invariances in Classical Physics". In: *Philosophy of Physics*. Ed. by Jeremy Butterfield and John Earman. Amsterdam: Elsevier, pp. 1331–1367.

Brown, Harvey and Peter Holland (2004). "Dynamical vs. variational symmetries: understanding Noether's first theorem". In: *Molecular Physics* 102.11-12, pp. 1133–1139.

Brown, Harvey and Jos Uffink (2001). "The Origins of Time-Asymmetry in Thermodynamics: The Minus First Law". In: *Studies in History and Philosophy of Modern Physics* 32.4, pp. 525–538.

Cassirer, Ernst (1923 [1910]). *Substance and Function - Einstein's Theory of Relativity*. Chicago: Open Court Publishing Company.

— (1956 [1937]). *Determinism and Indeterminism in Modern Physics*. New Haven: Yale University Press.

Chang, Hasok (2019). *Operationalism*. url: `https://plato.stanford.edu/entries/operationalism/#OperDefiDoNotExhaMean`.

Dasgupta, Shamik (2016). "Symmetries as an Epistemic Notion (Twice Over)". In: *British Journal for the Philosophy of Science* 67.3, pp. 837–878.

Dennett, Daniel (1991). "Real Patterns". In: *The Journal of Philosophy* 88.1, pp. 27–51.

Duhem, Pierre (1982 [1906]). *The Aim and Structure of Physical Theory*. Princeton: Princeton University Press.

Earman, John (1989). *World Enough and Space-Time*. Cambridge MA: MIT Press.

Einstein, Albert (1920 [1916]). *Relativity: The Special and the General Theory*. London: Methuen.

Esfeld, Michael and Vincent Lam (2011). "Ontic Structural Realism as a Metaphysics of Objects". In: *Scientific Structuralism*. Ed. by Alisa Bokulich and Peter Bokulich. Dordrecht: Springer, pp. 143–159.

Feynman, Richard (1965). *The Character of Physical Law*. Cambridge MA: MIT Press.

Frege, Gottlob (1960 [1891]). "Function and Concept". In: *Translations from the Philosophical Writings of Gottlob Frege*. Ed. by Peter Geach and Max Black. Oxford: Basil Blackwell, pp. 21–41.

French, Steven (2014). *The Structure of the World*. New York: Oxford University Press.

French, Steven and James Ladyman (2003). "Remodelling Structural Realism: Quantum Physics and the Metaphysics of Structure". In: *Synthese* 136, pp. 31–56.

— (2011). "In Defence of Ontic Structural Realism". In: *Scientific Structuralism*. Ed. by Alisa Bokulich and Peter Bokulich. Dordrecht: Springer, pp. 25–42.

French, Steven and Juha Saatsi (2018). "Symmetries and Explanatory Dependencies in Physics". In: *Explanation Beyond Causation*. Ed. by Alexander Reutlinger and Juha Saatsi. New York: Oxford University Press, pp. 185–205.

Friedman, Michael (1974). "Explanation and Scientific Understanding". In: *The Journal of Philosophy* 71.1, pp. 5–19.

— (2001). *Dynamics of Reason*. Stanford: Center for the Study of Language and Information.

Goldberg, Nathaniel Jason (2009). "Historicism, Entrenchment, and Conventionalism". In: *Journal for General Philosophy of Science* 40, pp. 259–276.

Goldstein, Sheldon et al. (2019). *Gibbs and Boltzmann Entropy in Classical and Quantum Mechanics*. url: `https://arxiv.org/abs/1903.11870`.

Gower, Barry (2000). "Cassirer, Schlick and 'Structural' Realism, the Philsophy of the Exact Sciences in the Background to Early Logical Empiricism". In: *British Journal for the History of Philosophy* 8.1, pp. 71–106.

Hardin, Clyde and Alexander Rosenberg (1982). "In Defense of Convergent Realism". In: *Philosophy of Science* 49, pp. 604–615.

Hempel, Carl (1965). *Aspects of Scientific Explanation*. New York: The Free Press.

Hempel, Carl G. and Paul Oppenheim (1948). "Studies in the Logic of Explanation". In: *Philosophy of Science* 15.2, pp. 135–175.

Ismael, Jenann and Bas C. Van Fraassen (2003). "Symmetry as a guide to superfluous theoretical structure". In: *Symmetries in Physics*. Ed. by Katherine Brading and Elena Castellani. New York: Cambridge University Press, pp. 371–392.

Kant, Immanuel (1998 [1787]). *Critique of Pure Reason*. Ed. by Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.

— (2000 [1790]). *Critique of the Power of Judgment*. Ed. by Paul Guyer and Eric Matthews. Cambridge: Cambridge University Press.

Kitcher, Philip (1989). "Explanatory Unification and the Causal Structure of the World". In: *Scientific Explanation*. Ed. by Philip Kitcher and Wesley Salmon. Minneapolis: University of Minnesota Press, 410–505.

Kitcher, Philip and Wesley Salmon (1989). *Scientific Explanation*. Minneapolis: University of Minnesota Press.

Knox, Eleanor (2017). "Physical Relativity from a Functionalist Perspective". In: *Studies in History and Philosophy of Modern Physics*, pp. 1–7.

Ladyman, James (1998). "What is Structural Realism?" In: *Studies in History and Philosophy of Science* 29.3, 409–424.

— (2019). "Introduction: Structuralists of the World Unite". In: *Studies in History and Philosophy of Science* 74, pp. 1–3.

Lange, Marc (2007). "Laws and meta-laws of nature: Conservation laws and symmetries". In: *Studies in History and Philosophy of Modern Physics* 38, pp. 457–481.

Laudan, Larry (1984). "Realism Without the Real". In: *Philosophy of Science* 51, pp. 156–162.

Lipton, Peter (1992). "The Seductive-Nomological Model". In: *Studies in History and Philosophy of Science* 23.4, pp. 691–698.

Massimi, Michela (2011a). "From Data to Phenomena: a Kantian Stance". In: *Synthese* 182, pp. 101–116.

— (2011b). "Structural Realism: a Neo-Kantian Perspective". In: *Scientific Structuralism*. Ed. by Peter Bokulich Alisa; Bokulich. Berlin: Springer, pp. 1–24.

McAllister, James (2004). "Thought experiments and the belief in phenomena". In: *Philosophy of Science* 71, pp. 1164–1175.

Olver, Peter J. (1995). *Equivalence, Invariants, and Symmetry*. Cambridge: Cambridge University Press.

Putnam, Hilary (1975). *Mathematics, Matter and Method*. Cambridge: Cambridge University Press.

Quine, Willard Van Orman (1951). "Two Dogmas of Empiricism". In: *The Philosophical Review* 60, pp. 20–43.

Ross, Don and James Ladyman (2007). "Scientific Realism, Constructive Empiricism and Structuralism". In: *Everything Must Go*. New York: Oxford University Press, pp. 66–129.

Ross, Don, James Ladyman, and John Collier (2007). "Rainforest Realism and the Unity of Science". In: *Everything Must Go*. New York: Oxford University Press, pp. 190–257.

Rovelli, Carlo (1997). "Halfway Through the Woods: Contemporary Research on Space and Time". In: *The Cosmos of Science*. Ed. by John Earman and John Norton. Pittsburgh: University of Pittsburgh Press.

Salmon, Wesley (1989). *Four Decades of Scientific Explanation*. Pittsburgh: University of Pittsburgh Press.

Sellars, Wilfred (1956). "Empiricism and the Philosophy of Mind". In: *Minnesota Studies in the Philosophy of Science* I, pp. 253–329.

Smith, Sheldon (2008). "Symmetries and the explanation of conservation laws in the light of the inverse problem in Lagrangian mechanics". In: *Studies in History and Philosophy of Modern Physics* 39, pp. 325–345.

Van Fraassen, Bas (1977). "The Pragmatics of Explanation". In: *American Philosophical Quarterly* 14.2, pp. 143–150.

— (1980). *The Scientific Image*. New York: Oxford University Press.

— (1989). *Laws and Symmetries*. New York: Oxford University Press.

Weatherall, James Owen (2011). "On (Some) Explanations in Physics". In: *Philosophy of Science* 78, 421–447.

— (2017). "Inertial Motion, Explanation, and the Foundations of Classical Spacetime Theories". In: *Towards a Theory of Spacetime Theories*. Ed. by Dennis Lehmkuhl, Gregor Schiemann, and Erhard Scholz. New York: Birkhäuser, pp. 13–42.

Wigner, Eugene (1954). "Conservation Laws in Classical and Quantum Physics". In: *Progress of Theoretical Physics* 11.4-5, pp. 437–440.

— (1964). "Events, Laws of Nature, and Invariance Principles". In: *Science* 145.3636, pp. 995–999.

— (1967). "The Role of Invariance Principles in Natural Philosophy". In: *Symmetries and Reflections*. Bloomington: Indiana University Press, pp. 28–37.

Wigner, Eugene Paul (1960). "The Unreasonable Effectiveness of Mathematics in the Natural Sciences". In: *Communications on Pure and Applied Mathematics* XIII, pp. 1–14.

Woodward, James (2003). *Making Things Happen*. New York: Oxford University Press.

— (2014). *Scientific Explanation*. url: `https://plato.stanford.edu/archives/win2014/entries/scientific-explanation/`.

Worrall, John (1989). "Structural Realism: The Best of Both Worlds?" In: *Dialectica* 43, pp. 99–124.