Laws or Models?

A comparative study of two models of scientific explanation

C. Martijn Wubs

referent:	Prof.dr. T.A.F. Kuipers
co-referent:	Dr. A.C.D. van Enter
derde beoordelaar:	Prof.dr. J.D. North

Vakgroep Wetenschapsfilosofie, Logica en Kentheorie
Faculteit der Wijsbegeerte
Rijks*universiteit* Groningen
Augustus 1997

"Die Frage ist nur: Wohin tun wir das Phänomen, wie benennen wir es, wie erklären wir es? Es klingt schulmeisterlich, aber wir Kastalier sind nun einmal Schulmeister, und wenn ich Euer und unser Erlebnis einzuordnen und zu benennen wünsche, so wünsche ich das nicht, weil ich seine Wirklichkeit und Schönheit durch Abstraktion und Verallgemeinerung auflösen, sondern weil ich sie möglichst bestimmt und deutlich aufzeichnen und festhalten möchte. Wenn ich auf einer Reise irgendwo einen Bauern oder ein Kind eine Melodie summen höre die ich nicht kannte, so ist mir das ebenfalls ein Erlebnis, und wenn ich dann diese Melodie sofort und so genau wie möglich in Noten aufzuschreiben versuche, so ist das kein Abtun und Weglegen, sondern eine Ehrung und Verewigung meines Erlebnisses."

fragment from: Herman Hesse, *Das Glasperlenspiel*

Cover illustration: Wassily Kandinsky, *Counter-Gravitation*, Stadt Museum, Mülheim.

Contents

Preface

Vier jaar geleden kampte ik met enige motivatieproblemen voor mijn studie natuurkunde: het is weliswaar een prachtstudie, maar de eerste paar jaar waren nogal dogmatisch van karakter. Heel anders ging het toe bij de twee filosofiecolleges die ik toen al had gevolgd, te weten het college wetenschapsfilosofie bij Gerben Stavenga en het vak Kennis, macht & Moraal dat door Hans Harbers gegeven werd. In beide vakken leerde ik ook eens op een andere manier te kijken naar de natuurkunde. Dat beviel me en daarom besloot ik de avondcolleges te gaan volgen die mij de filosofiestudie hebben binnengetrokken.

Naast alle filosofisch-inhoudelijke vakken heb ik met veel plezier bij Tsjalling Swierstra het college Algemene Vaardigheden[1] met als vervolg het Leeronderzoek gedaan, waar je het maken van teksten als ambacht wordt bijgebracht en waar je leert kritiek te geven op andermans werk en kritiek op de eigen opstellen leert verstouwen en gebruiken. Die vaardigheden waren een must toen ik het voorrecht had om in Cambridge drie maanden bij de vakgroep 'Philosophy of Physics' van Michael Redhead te mogen studeren, waar de 'Groningse' Katinka Ridderbos momenteel haar Ph.D. doet.

In Cambridge was het filosoferen – meer dan in Groningen – geconcentreerd rond allerlei lees- en discussiegroepen. Ook over onderwerpen waarbij de studenten zich niet op glad ijs wilden wagen, daagde Michael Redhead ons uit om stelling te nemen. Zijn vraag "And, Martin, what do *you* think of this?", bij voorkeur gesteld tijdens momenten van geestelijke afwezigheid van de aangesprokene, heeft bij het formuleren van mijn eigen mening in deze scriptie nog als een mantra door mijn hoofd gespookt.

Door voor de vakgroep Wetenschapsfilosofie, Logica & Kentheorie te kiezen, ben ik in de filosofie dicht bij de natuurkunde gebleven. Dat is voor mij een kruisbestuiving geweest: ik ben opnieuw enthousiast geworden voor de natuurkunde door er ook eens met de bril van de filosoof naar te kijken, terwijl andersom mijn ideeën over wetenschapsfilosofie beïnvloed zijn door mijn ervaringen in de natuurkunde. Deze scriptie over wetenschappelijke verklaringen in de natuurkunde is een resultaat van die wisselwerking.

Ik dank Theo Kuipers dat hij, ondanks zijn 'sabbatsjaar' in Wassenaar, mijn eerste begeleider wilde zijn. We hebben met name per e-mail gecommuniceerd, maar zijn 'instantaneous (re-)actions-at-a-distance' heb ik als stimulerend ervaren, zowel in de aanloopfase bij het kiezen van een onderwerp, als tijdens de schrijffase. Ik dank mijn tweede begeleider Aernout van Enter voor de discussies die we voerden en die hij lardeerde met voorbeelden uit de natuurkunde. Mijn derde beoordelaar, prof. J.D. North, evenals mijn twee begeleiders, wil ik bedanken voor de gezwindheid

[1] Ik zou een naamswijziging van dit vak willen voorstellen: zo algemeen zijn deze vaardigheden nu ook weer niet. 'Filosofische Vaardigheden' is weer te specifiek, want ook buiten de filosofie is het handig als je je ideeën helder mondeling en schriftelijk kunt verwoorden. Ik wil daarom de naam 'Academische Vaardigheden' voorstellen, ten eerste omdat de naam de lading aardig dekt en ten tweede vanwege het wervende en politieke karakter ervan: er worden bij Filosofie vaardigheden aangeleerd die (ten onrechte) in curricula van andere studierichtingen zijn ondergesneeuwd.

en nauwkeurigheid waarmee zij mijn werk hebben gelezen en becommentarieerd, en dat in de hete zomermaanden van 1997!

Mijn allereerste twee begeleiders zijn eigenlijk mijn ouders. Hun interesse in mijn verrichtingen en hun enthousiasme voor mijn plan om filosofie als tweede studie te gaan doen – wat mijn studietijd niet bekort heeft – zijn voor mij heel belangrijk geweest. Ik wil ze daarvoor bij deze in mijn moers taal bedanken.

Zonder mijn vrienden, tot wie ik ook mijn broer en zus mag rekenen, had ik het schrijven van nóg een scriptie een heel wat zwaardere inspanning gevonden. Amicae amicique, bedankt! Vrienden zijn er gelukkig niet alleen voor de ontspanning. Ik heb bijvoorbeeld met veel plezier met vrienden van mijn studentenvereniging 'Hendrik de Cock' de 'eloquentia' beoefend, waarbij je aan den lijve kunt ondervinden hoe een opponent met de zaal wegloopt als je in je eigen betoog je publiek onvoldoende in het oog houdt... Ik hoop dat mijn scriptie de lezer aanspreekt!

C.M.W. Groningen, augustus 1997.

Chapter 1

Introduction

In science, theories are constructed in order to explain phenomena that were not understood, or not known before. But what *are* scientific explanations?

Philosophers of science have made models of scientific explanation. In these models, the structure of scientific knowledge is reflected in the *form* of explanations: are explanations based on laws, on theories, on models, or on a combination of them? Secondly, philosophers try to find a theory of explanation to accompany the model: the theory should tell how scientific explanations as described in the model are possible. Theories of explanation are metaphysical theories.

Two models of scientific explanation will be discussed in this thesis: the Deductive-Nomological model and the simulacrum account of explanation. The models differ quite a lot, so that it is interesting and instructive to ask *which of the models is best? And for what reasons?* Even if in the end neither of the models should turns out to be the ultimate model, we will have learnt lot about scientific explanation.

In chapter 2, the Deductive-Nomological model will be introduced, which was articulated by C.G. Hempel. It is a model of explanation that is based on laws of nature, and it has been the received view for some time. At the end of the chapter, I shall mention four points of criticism of the model: the problem of theoretical ascent; the problem of provisos; the question what makes an argument an explanation; and finally, the problem to deal with quantitative arguments.

In chapter 3, I present Nancy Cartwright's criticism of the Deductive-Nomological model, and also her own simulacrum account of explanation, which stresses the important role of models in physics.

In order to keep the discussion clear, I postpone my own criticism until chapter 4, where the four points of criticism that were raised in chapter 2 will be raised again. I discuss how important I think the problems are for both models and suggest some solutions and point at some inconsistencies.

Chapter 5 contains new 'empirical input' in the philosophical discussion: several physical models are introduced, which I use to clarify and to extend the ideas that I expressed in the previous chapter. I shall only use physical models as examples, so that the subject of the thesis becomes a bit less general, and simply because physics is the science I am most familiar with.

In the final chapter 6, the reader can find my conclusions about the two models of explanation, their relative strengths and weaknesses, and the problematic aspects that both models have in common.

Chapter 2

The Deductive-Nomological model of science

In the Middle Ages there were rationally trained university scholars and craftsmen who were good at making instruments. When for social and economical reasons these two groups of people came into contact, modern science was born. Modern science can be characterized by the mutual influence that theory and experiment have on each other (see [43]).

If one wants to explain the success of modern science from the fact that it is a rational activity where theoretical and experimental knowledge are linked, one should be able to describe *how* these two kinds of knowledge are related. One account of this relation is the Deductive-Nomological model (or DN-model), which was articulated by C.G. Hempel [22] and which for some time was the recieved view among philosophers of science.

In this chapter the DN-model is presented. Although it was (and still is) a popular view of science, the model has been criticised as well. I shall focus on some points of criticism which were put forward by deductivists, including Hempel himself. In this way, we can better understand Nancy Cartwright's criticism of the DN-model as discussed in chapter 3, and in what respect her anti-deductivist position differs from the deductivists. Moreover, we can see in what respect her model of science is liable to the same criticism and in what respect it is an improvement.

2.1 Deduction and deductivism

Deduction is a well-known form of logic that already was used and studied by Greek philosophers. Euclid used deductions in his *The Elements*, where geometry is presented as an axiomatized deductive system: a few postulates (or axioms) are used to derive a large number of theorems (or propositions). Aristotle, whose syllogistic logic is well-known, studied the nature of deductive proof: what is the relation between premises and conclusions so that the truth of the premises guarantees the truth of the conclusion? This thesis will be 'Euclidean' rather than 'Aristotelian': we concentrate on the use of deductions in science rather than on proof theory about deductions in general.

Generally, a deduction can be presented as follows:

$$P_1 \,\&\, P_2 \,\&\ldots\&\, P_n \models C.$$

The deduction is valid if the truth of the conjunction of the premises $P_1, \ldots, P_n$ implies the truth of the conclusion C.

There are other forms of 'logic' than deduction, like induction, intuitive reasoning, reasoning by analogy, etcetera. Now *deductivism* can be defined by the statement that:
deduction is the only *acceptable form of logic in the empirical sciences.*
Loosely stated, the general idea behind deductivism is that only in deductions we can be certain about the truth of the conclusion given the truth of the premises, and since science should produce knowledge which is certain, deduction is the only allowed form of logic.

A deductivist must at least show two things: in the first place he must show *how* deductions are used in science, in other words: which kind of statements are used as premises and which as conclusions in scientific reasoning. He therefore has to have a deductive model of scientific knowledge. Secondly, a deductivist must be able to show that deduction alone forms *sufficient* logical apparatus to do science. This is an empirical claim: if there happen to be results in science which are the product of non-deductive reasoning, then these results cannot be considered truly scientific.

Perhaps this second point about the deductivist position is formulated too strongly: it might be that scientists use other modes of reasoning than deduction in order to arrive at results which afterwards can be incorporated into a deductive model of science. Therefore the second point could be mitigated by requiring the deductivist to show that it is possible to make a satisfactory 'rational reconstruction' of science in a deductive model.

This weaker deductivist position concentrates on the knowledge which is the end product of scientific activity; this "Logik der Forschung" can do without the analysis of the role of the individual scientist as a rational agent or of the scientific community as a whole, which one might call the "Logik der Forscher". The deductivist makes the distinction between the "context of discovery" and the "context of justification": for the deductivist it is not important how the scientific results are discovered; calling the results 'scientific' can only be justified by giving them a place in a deductive model. In the next section such a model is introduced.

2.2 The DN-model

The best known deductive model of science is Hempel's Deductive-Nomological model (or DN-model). It relates physical theories to laws of nature and to empirical phenomena. The deductive explanation of a phenomenon F may take the following form:

$$L_1 \& \ldots \& L_n \ \& C_1 \& \ldots \& C_m \models F$$

The laws of nature $L_1, \ldots, L_n$ together with the specific conditions $C_1, \ldots, C_m$ imply the occurrence of the phenomenon F. As a simple example, Kepler's 'area law' (which states that the line between the centre of gravity of a sun-planet system and the planet circling around it, sweeps out equal areas in equal times) plus the specific condition that the celestial object which I am observing through my telescope is say the planet Jupiter, implies that the line joining the sun and this object sweeps out equal areas in equal times.

The observation that the object considered satisfies the area law is *explained* by the conjunction of the law and the specific conditions under which the observation is made. One therefore speaks about the Deductive-Nomological model *of explanation*; a premise is called an 'explanans' and the conclusion is the 'explanandum'. The model can also be used to describe how *predictions* can be made, like "Given the area law and the assumption that the unidentified flying object which I observe in the sky is a planet, then this object must obey the area law". Often phenomena which have been predicted by a theory will be explained by that theory once they have been observed. Hempel stated in his 'structural identity hypothesis' that there is a symmetry between scientific explanation and prediction, which both can be described using the DN-model ([21], p. 364-376). We shall not pursue the question here whether the structural identity hypothesis is right. I want to make two points by mentioning it here. The first point is that a deductivist must show that both predictions and explanations can be rationally reconstructed using deduction as the only form of logic, and depending on the truth of the structural identity hypothesis he needs one or two deductive models. The second point is that in later sections it will prove useful to analyze whether critics object to the DN-model as a model of scientific explanation, as a model of scientific prediction, or both.

The laws of nature entering the deductions often implicitly or explicitly have the form of universal generalizations (e.g. "*All* planets circling around the sun sweep out equal areas in equal times"). The DN-model is sometimes called the *covering law model*, because the laws in the deduction provide universal 'coverage' of the phenomena which can be subsumed under them ("All planets..."). The occurrence of specific phenomena is explained/predicted by the fact that general laws are at work in specific circumstances. There are several interpretations of laws of nature: is a law (necessarily?) true about the world, is it a summary of our knowledge about the world, or is it an inference rule for the scientist? I leave these interpretative issues to later chapters and concentrate on the form of the DN-arguments in which the yet uninterpreted laws occur.

Since the laws are used as premises in the deductions, the deductions do not guarantee the truth of the laws. Then, what makes these laws true? The truth of laws of nature which are universal generalizations can not be verified experimentally, because the truth value of universal generalizations can only be determined if the behaviour of all objects in the universe about which it generalizes are known. This usually is not the case; the law is considered to be true about unknown objects as well. For the deductivist, the truth of the law cannot be inductively inferred from the finite number of cases in which it is not violated. So how can a deductivist use laws which are universal generalizations?

For some laws of nature their truth is implied by the truth of physical theories and other laws by which the laws can deductively explained. These laws are said to have *theoretical support*. This is the DN-model on a higher level:

$$T_1 \& \ldots \& T_n \; \& L_1 \& \ldots \& L_m \& \; C_1 \& \ldots \& C_q \models L'.$$

From the physical theories $T_1, \ldots, T_n$, taken together with the laws $L_1, \ldots, L_m$ and the extra conditions $C_1, \ldots, C_q$, the law L' can be deduced. For instance, Kepler's area law can be deduced from Newtonian mechanics, together with the assumption that the planet orbits around the sun because an attractive central force is keeping them together. However, from Newton's laws it also follows that the range of validity of the area law is limited: the planet not only feels the gravitational force exerted by the sun, but also the (much smaller) gravitational pull by the other planets. As a consequence, the planet's motion around the sun will deviate from the perfectly elliptical orbit. Newton's laws tell you how much deviation is caused by the perturbation. If among the conditions $C_1, \ldots, C_q$ it is stated that the other planets are far away and not too heavy, then in that limited range the area law "holds true in fairly close approximation" ([22], p. 76). Theoretical support is not unconditional.

Not all laws and theories enjoy theoretical support: if scientific knowledge is structured deductively, it must have some axioms. These are the statements in science which are taken to be true. So, what kind of statements are taken to be true in science? In physics, some theories are called 'fundamental' and at least these theories will be taken as axioms; the word 'fundamental' suggests that the truth of these theories somehow guarantees the truth of laws, observations or other theories.

The reasons why the axioms in the DN-model of science are taken to be true is not explained by the DN-model itself. It only presupposes that there are theories and laws that are considered true. In the chapter 3 and 4, much more will be said about this. Here I just remark that the truth of the theories used in a DN argument should not be taken too strictly, since many physical theories which used to be considered true and fundamental have lost this privileged status. This is what happened to Newtonian mechanics after the advent of the theory of relativity. Therefore, if one supposes that the DN-model gives a good account of modern science, then it becomes clear that advocates of the DN-model do not presuppose that the theories which are used as axioms today, must be true and fundamental forever.

The fundamental theories can not be the only axioms in the DN-model. As we have seen, predictions and explanations of phenomena in specific experimental circumstances are deduced with the help of certain conditions which are used as premises and which are not deducible from the fundamental theories. One therefore needs a theory of observation which tells how these statements about specific experimental conditions can be considered true premises. Going back to our example, how do we arrive at the statement that the object we are observing is a planet and that it is Jupiter? Observations can only be used as premises in the DN-model if they can be cast in the form of true statements in some language: experimenters are supposed to tell the truth. The language in which these statements about the observed data are made, is called the *observation language.*

2.3 The DN-model and language

Unlike the logical positivists like Carnap, Hempel does not require that the observation language should do without any theoretical concepts. He requires that the observation language is "antecedently understood" ([22], p. 73), which means that the theoretical concepts of the observation language come from well-established theories which are not put to the test in the experiment considered.[1] Still, in the theory which *is* put to the test, some theoretical entities (electrons, fields) will be postulated the names of which do not occur in the observation language. So there is a *theoretical language* which is different from the observation language.

At first sight there seems to be a gap between the theoretical and the observation statements: How can a theoretical statement about electrons be tested in experiment if experimental findings are formulated using the observation language dictionary, in which the word 'electron' does not occur? The gap between the two languages is closed by *bridge principles.* An example from physics of such a bridge principle is the statement that the quantity 'temperature' of a physical system is to be identified with the mean kinetic energy of the particles which make up the system (see [24]). This bridge principle (plus a few other principles and a few postulates) bridges the gap between statistical mechanics and the laws of thermodynamics. Without bridge principles, statistical mechanics would not explain anything and the same is true about all theories. As in the above example, bridge principles often connect a microscopic theory with a macroscopic theory or law. But not always: the Born postulate of quantum mechanics states that the absolute square of the wavefunction of a particle is the probability density function of measuring it somewhere [15]. This postulate is a bridge principle which relates a microscopic theory to observation.

It is the bridge principles which create the possibility of formulating predictions

[1] I shall not discuss theories of observation here; Hacking [19] is a good read about this topic. He bases his entity realism not on theories, but on the fact that theoretical entities can be manipulated in experiments.

(or 'test implications') of a new theory in terms of the 'antecedently available' observation language. These principles are necessary extra premises in the DN-model: they are neither part of the theory nor part of the specific circumstances in which a phenomenon or a law is derived. So, taking the bridge principles explicitly into account in the DN-model, in general the form of a deductive-nomological argument looks like:

$$T_1\& \ldots \& T_n \; \& B_1\& \ldots \& B_m \; \& L_1\& \ldots \& L_p\& \; C_1\& \ldots \& C_q \models L',$$

where the B's stand for the bridge principles needed. The conclusion can be either a law L', for example Ohm's law, or a specific phenomenon, like "The electric current will be 3 mA."

In the section 2.1 I stated that a deductivist at least must make two points clear: he must show how deductions are used in science and he must show that deduction is the only logic needed to analyze (a rational reconstruction of) science. In this and the previous section I introduced Hempel's version of the first point: in his DN-model, scientific knowledge is presented as a deductive network of theories, laws, statements about specific experimental conditions, bridge principles and observations.

Now that we have a model of science which suits the deductivists, the interesting question arises whether all scientific knowledge can be caught in a deductive net, which is the second point that a deductivist should make clear. The best way to operate here seems to be analyzing how deductivists react to critics of the DN-model. Some criticism will be given in the next section.

2.4 The limits of deductivism

The DN-model for some time was the standard view of scientific explanation. Neopositivists even said that the explanatory power of the social sciences is limited because explanations in those sciences seldom are of the deductive-nomological type. But leaving the social sciences to other theses, in this section some criticism on the DN-model as a model for the natural sciences will be discussed. The best way to evaluate the deductivists' claim that the only acceptable form of logic in science is deduction, is to analyze objections by critics of that claim. In this section, four points of criticism will be discussed. After Nancy Cartwright's model of scientific explanation has been introduced in chapter 3, it will be interesting to compare in chapter 4 both models of explanation by the same four points of criticism that are raised in this section.

It is well-known that there has been a lot of criticism of the positivists' distinction between the 'context of discovery' and the 'context of justification', by Kuhn [28], Laudan [33] and others. Below, I introduce a selection of criticisms of the DN-model that I found in *The Limits of Deductivism* [17], by philosophers who think that this distinction of contexts is useful and who try and justify scientific knowledge by way of formal analysis. I shall give some comments, but a more thorough discussion has to wait until chapter 4.

2.4.1 Hempel's provisos

Hempel himself has been a critic of 'his' DN-model [17]. A first point of criticism, which Hempel thinks is quite serious, is the problem of "provisos". Consider the case where the DN-model is used to explain a specific phenomenon in an experiment, for instance the phenomenon that two metal bars which are suspended freely in one plane not far from each other align themselves in one straight line. One need not be a physicist to suspect that magnetism is at work here, so what is the problem in using the theory of magnetism in the DN-model to find the observed behaviour of the bars as a consequence?

The problem is that the theory of magnetism does not guarantee the absence of disturbing factors like external magnetic fields, friction from strong air currents etcetera, which could lead to quite different behavior of the metal bars. Therefore, every time the experiment is performed, the behaviour of the bars can only be explained in a deduction where for every possible disturbing factor a premise is added which states its importance. Even if no other disturbing influences are present, this must be stated in an extra premise, without which the deduction could not be said to explain the specific experiment then and there. The extra premise is an example of what Hempel calls *provisos:*

essential, but generally unstated, presuppositions of theoretical inferences.

The proviso does not belong to the theory in general, but to the specific experiment which the theory should explain. A proviso is different from a *ceteris paribus* clause, because it does not state that 'other things are equal' (equal to what?), but that other things are right.

Provisos have important consequences for the philosophy of science. If apart from the theory a proviso is used to deduce experimental consequences, the theory no longer is falsifiable. That can be understood as follows: the proviso should imply that frictional, electric, gravitational and any other forces are not relevant in the experiment at hand. The proviso therefore can not be formulated in terms of the theory of magnetism alone, so that the bridge principles of that theory do not suffice to formulate the conclusion of the deduction in the antecedently understood observational language alone. And even if the conclusion *could* be formulated in observation language, and the prediction fails to be true in experiment, then one could always blame some unknown proviso, so that the theory never is put to the test. This is what Popper calls a 'conventionalist stratagem'.

2.4.2 Theoretical ascent

Another problem which Hempel [17] sees for the DN-model is *theoretical ascent*, which in some sense is the counterpart of the provisos: provisos are about the theories and laws which are considered unimportant, whereas the problem of theoretical ascent has to do with the theories and laws which are assumed to be 'at work' in the phenomena. Again consider the example of the metal bars which arrange themselves in one line. It was just said that one need not be a physicist to suspect that magnetism is at work. But if it is really that easy, how do we know that we should use the theory of magnetism in order to explain this phenomenon? Can we somehow deduce that the bars are magnets without using the theory of magnetism? It seems that not deductive, but rather some sort of inductive reasoning is used here. It would be a problem for the strict deductivist if he can only use the DN-model after some inductive argument.

2.4.3 Can a deductive model be explanatory?

According to Hempel the DN-model can be used both for prediction and explanation of phenomena. About explanation he states the following:

> "A D-N explanation will have to contain, in its explanans, some general laws that are *required* for the deduction of the explanandum, i.e. whose deletion would make the argument invalid" ([21], p. 338).

Wesley Salmon's criticism of the DN-model is that it can not be used to give explanations, because the only thing a valid deduction shows is that the truth of the theory is *sufficient* for the truth of the phenomenon ([17], p. 95-127). But so long as the truth of the theory is not *necessary* for the truth of the phenomenon, the theory does not explain the phenomenon; if many other theories (known or unknown) might also be used to deduce the same phenomenon, then we can not say that the deduction using the first theory is an explanation.

Salmon's requirement is a strong one. It means that somehow the phenomenon excludes the possibility that another theory can be used to explain it. That were the case if all theories could be located in some abstract 'theory space', where different theories occupy volumes that mutually do not overlap. A theory T in this abstract space would explain a phenomenon if it were possible to locate the phenomenon in one and only one volume, corresponding to theory T. However, sometimes a continuous transition from the one theory to the other is possible, in the sense that a range of phenomena can be deduced from both, with minute numerical differences. In fact, the *correspondence principle* of the physicist Niels Bohr states that such a continuous transition exists between classical and quantum physics (more about this principle in section 5.2.2). The picture of non-overlapping theoretical volumes to locate phenomena in, breaks down in such cases.

The deduction with the theory as a premise and the phenomenon as a conclusion shows that the phenomenon is *necessary*, given the truth of the theory and of the other premises; the theory is only sufficient for the deduction. If you require the theory to be necessary for the deduction to be explanatory, then it seems that you also need another deduction, where the theory is deduced from the phenomenon plus other premises. However, the phenomena usually do not dictate what the theory 'behind' them should look like, or in philosophical parlance: theories are underdetermined by the data which they are supposed to explain. So, if we really require that we can only explain a phenomenon with the help of a theory if the truth of the theory is necessary for the phenomenon, then not many things in the world can be explained by an underlying theory. I think that the more modest philosophical task of analyzing how scientific explanation actually works, is interesting as well.

Billiard ball dynamics can still be explained by Newtonian mechanics (as physicists actually do). Newtonian mechanics is not necessary to explain billiard-ball dynamics (because a relativistic mechanics could explain it as well), but it is sufficient. This brings us to the tricky question how explanations which are based on a theory which is known to be false in some respects, still can be cast in any DN argument.

Perhaps the notion of *potential DN explanation* may be of some use, which Hempel defines as "... any argument that has the character of a D-N explanation except that the sentences constituting its explanans need not be true." ([21], p. 338). He describes two situations where potential explanations can be used:

> "We use the notion of potential explanation, for example when we ask whether a novel and yet untested law or theory would provide an explanation for some empirical phenomenon; or when we say that the phlogiston theory, though now discarded, afforded an explanation for certain aspects of combustion" ([21], p. 338).

This is a bit odd, as we can see from distinguishing three phases in the history of, say, phlogiston theory: when the theory had not yet been tested, explanations based on phlogiston theory were potential explanations. Secondly, when the theory had been tested and accepted, the potential explanations had turned into explanations; finally, when phlogiston theory was rejected, the explanations became potential explanations again.

Especially the last step is quite odd. If Hempel means instead that all the time explanations based on phlogiston theory were only potential explanations, since that theory might be falsified in the future, then we can never tell explanations from potential explanations. In that case, potential explanations cannot solve problems for any model of explanation. It seems natural if we only call explanations potential explanations in the first situation, where the theory has not been tested yet. But if we restrict our definition in that way, then the notion 'potential explanation' is of no use to us if we want to understand how a theory that has *already proved false* in some experiments, still can be used in DN explanations of other phenomena.

I conclude that, contrary to Salmon's views, theories need not be necessary in order for DN arguments to be explanatory. (Mere sufficience may not be sufficient, but necessity is not necessary...) Since 'potential explanations' did not solve it, a remaining problem is that theories that have been refuted in some experiments, still are being used to explain the outcomes of other experiments. They are sufficient for the deduction, but how can it be justified to use them as true premises in a DN explanation? We shall return to this question in chapters 4 and 5.

2.4.4 From black and white to shades of grey?

Another flaw of the DN-model, which is mentioned by J.E. Kyburg ([17], p. 61-94), is the fact that it does not take into account the quantitative character of science. In the DN-model the essence of science seems to be assigning the truth values true or false to statements, whereas in experiments numbers are assigned to quantities. One might object that there is no problem here, because the result of a measurement can be cast in the form of a statement: if quantity Q is measured to have the value v, then the statement "Quantity Q has value v" is a true statement that can be compared with a prediction.

However, the situation is more complicated than that. For instance, if Ohm's law is used to predict an electric current through a copper wire of 3 mA and if 3.02 mA is the measured value, then the statement "The current is 3 mA" apparently is false. Does this mean that the theory is refuted? Probably not. Physics students learn that their measurements are not complete without an estimation of the errors involved, and how errors in a measurement give rise to errors in results which are calculated from the measured values (see for instance [20]). The type of error that is most frequently used is the probable error, which states that the probability that the value of the quantity lies inside the error range is about 67 %. Thus probability enters all statements about observations. Therefore error analysis is relevant for the philosopher who claims that a deductive model with statements in it which are either true or false, can be a *rational* reconstruction of science.

The quantitative character of science does not only come to the fore while doing measurements. In section 2.2 we saw that Newton's laws imply that Kepler's area law holds true only *approximately*, in a limited range defined by qualifying conditions. The DN-model only seems capable of explaining that, given the conditions, a law is either true or false: *tertium non datur*. Since the DN model can not handle laws that hold approximately, doesn't that mean that it is a bad model of scientific explanation? I shall leave this question unanswered until I have presented Nancy Cartwright's answer in chapter 3.

2.5 Conclusion

In this chapter the DN-model was introduced. It was shown that both specific phenomena and general laws of nature can be the conclusions of a DN-argument. The fact that theoretical terms do not occur in the language in which observations are made, calls for a special type of premises in DN-arguments: bridge principles. The DN-model is used both as a model of scientific explanation and prediction, so that it may fail both in its account of explanation and prediction.

The DN-model suffers from the problem of provisos: how can we make a valid deduction from a certain theory if possible influences which are described by other theories can not be ruled out? Another problem is theoretical ascent: how can we know which theory to apply in what situation?

Then it was argued that the DN-model cannot be used to explain phenomena, because it only shows that the theories are sufficient for the phenomena to occur, and not necessary. I objected that this requirement was too strong. The general problem remains what makes a DN explanation explanatory. More specifically, do explanations that use refuted theories fit into the DN model, and if so: what are the reasons to treat the refuted theory as a true premise?

A final point of criticism was that the DN-model can not be adequate, since not deductive logic but rather quantitative arguments seem to be used in scientific explanations, if only because measurements involve large or small errors. Probability and accuracy must be given a place in scientific explanations.

In the next chapter Nancy Cartwright's model of science is presented, which is quite different from the DN-model and which perhaps solves some of the problems of the DN-model presented in this chapter.

Chapter 3

Cartwright's model of science

Things fall apart; the centre does not hold;
Mere anarchy is loosed upon the world...
W.B. Yeats, *The Second coming*

Here is another introductory chapter: after introducing the DN model, I now introduce the work of one of the most fervent critics of that model. Her name is Nancy Cartwright. Section 3.1 is about her criticism of that model. The following section is about the importance of modelling in physics; in section 3.3, Cartwright's simulacrum account of explanation is presented, in which models play a major role. Section 3.4 deals with the question in what sense models can be realistic. Cartwright's metaphysics is discussed in section 3.5 and section 3.6 concludes. Because Cartwright often is criticizing others, I shall not criticize her work too much in order to keep the presentation as clear as possible. So, unless stated otherwise, I present Cartwright's ideas. (I shall have my say in later chapters.) This being said, What's wrong with the DN model?

3.1 Criticism of the DN-model

Cartwright argues that the DN model is a bad model of scientific explanation, and that a close look at 'actual explanatory practice' reveals why. The first reason why there is something wrong with the model is that the laws which are used in scientific explanations often are *false*. Kepler's area law - which we met in chapter 2 - is only valid in a Euclidean universe with only one perfectly spherical planet circling around an equally spherical star. These are not the conditions of our solar system, so the area law is false. Since in the DN model phenomena are explained by showing that they are manifestations of true laws of nature under special conditions, the area law can not be used as a premise in a DN argument about planetary motions.

The area law is not special in that it is used to explain even though it is false: most laws of nature are *ceteris paribus* laws, that is: laws which are only true if certain conditions are satisfied. In practice these conditions often are not fulfilled. The laws are deficient, but they are treated as if they were exceptionless: false laws are used to explain ([4], p. 46)[1]. There may be unknown more complex true laws about the real situation which some day may be discovered, but in the meantime explanations are given anyway, without using these true laws. This shows that

[1]Below, I shall only give a page number if (and only if) I refer to Cartwright's *How the laws of Physics lie.*

explanation does not need covering laws. The conclusion must be that all covering law models, including the DN model, are inadequate models of scientific explanation.

The DN model has two important aspects: it uses laws, and it uses them in deductive explanations. We have just seen that many laws are false, so that Cartwright concludes that the DN model cannot handle explanations based on them. Now we shall see that the DN model is also flawed because many derivations in practice are not purely deductive. In chapter 2, it was shown that the DN model can be applied on a higher level, where laws of nature are explained by more fundamental theories. If the fundamental theories are true about nature, then so are the laws derived from these theories. However, often (mathematical) approximations are made in the derivation. These approximations are not dictated by the facts; they are made to simplify the mathematics, or just to arrive at the law which one sets out to explain. And sometimes, approximations improve the accuracy of the law (p. 107). In all these cases, the approximation is not a deduction. If the theory is right, then the law which was derived after the approximation was made, is false. If the derived law is a good phenomenological law, this only shows that there does not exist a deductive relation between theory, laws and phenomena; moreover, it supports Cartwright's thesis that the empirical *content* of science lies in the phenomenological laws and not in the abstract theories. The latter may have the function of organizing knowledge, but organizing power has nothing to do with truth.[1] And even if the theories are true, "The truth doesn't explain much", (not deductive-nomologically, that is), since the approximations - which are not dictated by the facts - stand in the way between theory and phenomenological law. I shall criticize Cartwright on this point in chapter 4.

A third reason why the DN model is defective, is a metaphysical one: the model presupposes that there are laws to cover every situation. Realists argue that the false simple laws which are used to explain, approximate unknown more complicated laws. Cartwright happens to disagree with realists in that she thinks that the world is 'messy': there is not a law to cover every case. However this may be, Cartwright thinks that unknown laws should not be invoked in order to explain why false laws are used to explain; the DN model should not depend on realist metaphysics, especially not when the model is used to defend realism (p. 126).

The fourth defect of the DN model is that it presupposes realism in yet another way. In the DN model, laws are derived from theories, and phenomena are derived from laws. Thus the DN model reflects a very simple ultra-realist picture of science: fundamental theories make laws of nature true, and the latter make phenomena happen. There are, however, no causal or other non-metaphorical arguments why one law derives from another. We can transfer our attention to language, and formulate theories and laws as statements, and then in the DN-model derive one statement by other statements and logic. But this procedure can not be an argument for realism.

For Cartwright the above arguments sufficiently make clear that scientific explanatory practice can not be explained with the help of the DN model. We shall see that for that reason, she introduces a new scientific world picture which is radically different from a picture based on laws of nature.

[1] Compare this with the 'best system account' of laws of nature (by Mill, Ramsey and Lewis), which says that some statements are called the laws of nature in virtue of the fact that they organize human knowledge about nature best; their small number and great empirical adequacy somehow are in equillibrium.

3.2 On the ubiquity of models in physics

The DN model sketches a picture of science that is incorrect; in practice, explanation is different from stating which laws of nature are at work and which are the specific circumstances of the measurement and the system being measured. The DN model only *seems* adequate, "because it ignores the fact that explanations in physics generally begin with a model." (p. 103). Cartwright gives a broad informal definition of 'model':

> "I think that a model - a specially prepared, usually fictional description of the system under study - is employed whenever a mathematical theory is applied to reality, and I use the term 'model' deliberately to suggest the failure of exact correspondence..." (p. 158).

In as far as (theoretical) physics can be seen as applying mathematical theory to reality, Cartwright states here that models are everywhere in physics. Some examples and consequences of the abundance of models in physics will be given below.

The use of models is widespread and this reduces the status of laws of nature: statements which are used as laws of nature in a DN argument are derived from more fundamental laws or theories *within a model*. Kepler's area law again serves as an example: it can only be derived in a 'model solar system' where the sun and the planet are the only objects present and where spacetime has zero curvature. This is *How the laws of physics lie* and Cartwright goes at pains to get this point across in her book:

> "The primary aim of this book is to argue against the facticity of fundamental laws" (p. 152).

Why would one choose to use models? It is *explanation* itself which requires models, because explanation needs generality and simplicity. Assuming that it exists, a real covering law of a situation becomes both very complex and so specific that it would not be likely to work anywhere else (p. 112). Such a law can not have much to do with explanation.

Secondly, Cartwright gives a *functional* explanation why physicists use models and only a limited number of them: the only way for physicists to communicate about the complex phenomena which they study, is to use a small number of models. This is an external, Kuhnian account of the paucity of models (p. 143). It not only makes clear why models are used which always are a bit besides the point in specific cases, but also why the same models are used throughout physics. The best example is the 'harmonic oscillator model': an object experiences a force that increases linear with the distance to its equilibrium position. The harmonic oscillator pops up as a model for several physical systems in classical and quantum mechanics.

A third reason to work with models is that the facts do not dictate how they should be described. Different models are best for different *purposes*. The same physical system (a laser for instance, see p. 158) will be described by different models, depending on which observable one wants to calculate, or how general a model one wants to use.

One could admit that models are distortions of reality and still suppose that the DN model gives a good account of how theories and models are related: proponents of the DN model might claim that fundamental theories together with the specific situations described in models, deductively imply the validity of less fundamental laws. Laws which are only true within the model, that is. But three of the four objections to the DN model which were raised in the previous section, also apply in this case: derivations of laws often are not deductive and they use approximation procedures which often improve the empirical adequacy of the resulting law.

Cartwright's objections to the metaphysics hidden in the DN model also applies in this case. So there is not a simple amendment to the DN model, which corrects for the use of models. Rather, the DN model of explanation must be discarded and replaced by a new candidate, which we shall discuss now.

3.3 The simulacrum account of explanation

In practice, Cartwright observes, physicists do not make explanations according to the DN model. She therefore favours a new model of explanation, which is named the *simulacrum account* of explanation. It relates theories, models and phenomena:

> "To explain a phenomenon is to find a model that fits it into the basic framework of the theory and that thus allows us to derive analogues of the messy and complicated phenomenological laws which are true of it" (p. 152).

As this quotation makes clear, the process of explanation involves two steps: finding a model for the physical system, and deriving laws.[1] I shall mainly discuss the first step, because in the second step Cartwright does not make the difference with the DN model explicit.

In an idealized account of how models are introduced, Cartwright distinguishes two stages. The first stages ends with saying "This physical system A behaves *as if* it is a B". For example, "This helium-neon laser behaves *as if* it is a collection of three-level atoms in interaction with a single damped mode of a quantized field, coupled to a pumping and damping reservoir". The name 'simulacrum account' can be understood from this stage, where for the sake of the explanation, A is treated as if it were the B that shows similar behaviour. But how do we arrive at saying "This A behaves as if it is a B"?

On the left of the 'as if'-operator we find what Cartwright calls the 'unprepared description' (p. 133): the system is described using whatever information we have, and all the terms we think necessary for descriptive adequacy. The unprepared description shows an existential commitment: there exists a physical system out there which can be described more or less accurately. Notice that as a contrast with Hempel, Cartwright does not suppose or require that in the description of the system there must be a distinction between theoretical and 'antecedently understood' observation language: the unprepared description may also contain terms from the theory which is put to the test.

Once the unprepared description has brought the phenomenon under a theory, we go to the right hand side of the 'as if'-operator, where we choose the *prepared description* in the same informal way as we chose the unprepared description: it is neither dictated by the facts, nor do principles of the theory tell us how we are to prepare the description. For example, neither the facts nor the theory dictate that the atoms in a laser should get the prepared description of 'three-level atoms'. The reason to choose a particular description lies in the eventual explanatory success, after the second stage has also been completed.

The second stage of finding a model is formal (p. 134): once the description of the system has been prepared, the principles of the theory dictate what equations will govern the system. The kind of principles which link prepared descriptions of systems to the theory are the *bridge principles* of the theory. As an example,

[1] The theory seems to be 'given' here. But theories are not 'given'. Firstly, they must satisfy some very general criteria (like the CPT-theorem) in order to be physical theories. Secondly, the function of the theories is a bit stange in the simulacrum account: some theory is chosen to derive a model with. What determines the choice of that theory, if the theory is not about the system but about the model that is yet to be constructed? More criticism in the next chapter.

Hamiltonians are bridge principles of quantum mechanics; the Hamiltonian of a system contains information about its possible energies, about its time evolution and its symmetries. One such Hamiltonian is the Hamiltonian of a three-level system.[1] If I understand Cartwright rightly, the prepared description of a system, taken together with the bridge principle linking that description to the theory, is what Cartwright later on calls the model of the system within the theory. This completes the story of how a model is chosen for a physical system. The success of the chosen model depends on how well it serves the explanatory purposes.

After the model has been chosen as just described, we do not know yet whether it is has explanatory success. That we can only evaluate after the second step of explanation: the derivation of laws within the model. Of course, the model and its theory are more successful if more laws can be derived. But the nature of the derivations is a problem in the simulacrum account:

> "The simulacrum account is not a formal account. It says that we lay out a model, and within that model we 'derive' various laws which match more or less well bits of phenomenological behavior. But even inside the model, derivation is not what the D-N account would have it be, and I do not have any clear alternative" (p. 161).

Cartwright's criticism of the DN model is partly based on actual explanations that were not deductions; now the same explanations prevent her from using deduction within her simulacrum account. I shall return to this problem in section 4.4.2.

3.4 'Realistic' models

The simulacrum account has been introduced and now it is time to see whether it gives a good account of scientific explanation. In quantum mechanics the word 'model' of a system is used in a narrow sense to denote a Hamiltonian, so that the model can easily be distinguished from the theory. This should make quantum mechanics a simple example of how the simulacrum account of explanation works. But quantum mechanics can teach us more: it is a theory with great predictive power. Does this mean that the models used in quantum mechanics are realistic? And what does the term 'realistic' mean here? Cartwright's answers to these questions are given below.

3.4.1 Example: 'the hydrogen atom'

In order to check whether her story of model choice is realistic, Cartwright turns her attention to quantum mechanics textbooks. She is surprised to find out that physics students in quantum mechanics courses do not learn the art of choosing models for physical systems, the creative process which has just been described in section 3.3. Rather, in textbooks the discussion *starts* with models. Models for which it it seems unlikely that there are real physical systems which behave 'as if' they were these models. For example, there are the Hamiltonians for a 'particle in a box', a 'particle in a square well potential', the 'harmonic oscillator', etcetera (see for instance [15]). Among the model Hamiltonians she discovers one Hamiltonian that seems to be about a real system: 'the hydrogen atom'. But there are essentially two reasons why it is not the Hamiltonian for a real hydrogen atom. Here is the first one:

[1] Note the different definition of the term 'bridge principle', as compared to Hempel's use of the term, which was discussed in section 2.3: for Hempel, a bridge principle connects the theoretical *language* to the observational language, whereas for Cartwright bridge principles are the mathematical equations which belong to the prepared description of the system.

> "Real hydrogen atoms appear in an environment, in a very cold tank for example, or on a benzene molecule; and the effects of the environment must be reflected in the Hamiltonian. What we study instead is a hypothetically isolated hydrogen atom" (p. 137).

The second reason why it is not real hydrogen is even much more important:

> " 'The hydrogen atom' ... is just a name for a two-body system where only the Coulomb force is relevant. Even if the system stood alone in the universe, we could not strip away the spin from the electron. Even less could we eliminate the electomagnetic field, for it gives rise to the Lamb shift even when no photons are present. This two-body system, which we call 'the hydrogen atom', is a mere mental construct."

Cartwright finds out that standard quantum theory is not about the real hydrogen atom, and neither is Dirac's relativistic quantum theory (see [11]), which is more realistic. Still, not only students but also professional physicists (who should know 'better') work with the spinless nonrelativistic 'hydrogen atom' Hamiltonian. This fact can be understood from the simulacrum account of explanation: we choose a model that serves our purposes best, and that model will not always include relativistic effects.

3.4.2 Two meanings of 'realistic'

To prepare for a discussion of realism, I must first mention Cartwright's analysis of how physicists use the term 'realistic'. It will also be important for later chapters. One meaning of the word 'realistic' is about the relation between the unprepared and the prepared description of a physical system: the model is realistic if "it describes the real constituents of the system - the substances and fields that make it up - and ascribes to them characteristics and relations that actually obtain." This is close to the common sense meaning of 'realistic'. In a second sense, it is opposed to 'phenomenological'. A phenomenological treatment of a system

> "... is unrealistic *from the point of view of the explanatory theory*. It gives a theoretical description of the behaviour, but nothing in the model shows what gives rise to this behaviour" (p. 149).

This second sense of 'realistic' has to do with the relation between the model and the mathematics:

> "A fundamental theory must supply a criterion for what is to count as explanatory. Relative to that criterion the model is realistic if it explains the mathematical representation" (p. 150).

Let me clarify this with a simple example that Cartwright does not give but which I think is instructive: riding on a bike can be described using Newtonian mechanics. As every cyclist knows, your low maximum speed should somehow be explained by friction. Although there may be all sorts of sources of friction, often the total friction is modelled by a force which is proportional to the velocity; the constant of proportionality is named the friction constant. The cyclist is modelled as a mass with a force that drives the system consisting of cyclist plus bike. Now is this a realistic treatment of the system? To begin with, is it realistic in the first sense? It is not, because the prepared description of a man by a mass delivering a force is highly idealized. The treatment of the system is not realistic in the second sense either, if only because the friction is treated purely phenomenologically: Newtonian mechanics does not give a clue as to what the numerical value of the friction constant should be, nor why the friction force should be proportional to the velocity. So, however accurate this explanation of bike dynamics is, it is in no way realistic.

As was said earlier, many types of friction can be treated the same way phenomenologically, so that the phenomenological treatment has wide applicability. If in a specific situation one manages to replace a phenomenological term by a bridge principle, then the treatment becomes more realistic in the second sense; the more bridge principles, the more realistic. It can not be applied to as many cases however, so the more realistic treatment has less explanatory power.

In what sense would Cartwright say that the nonrelativistic 'hydrogen atom' model of the previous section is realistic? It is not realistic in the first sense, because a real hydrogen atom interacts with its environment and has a spin, but neither spin nor environment are represented in the model. Is the model realistic in the second sense, with respect to the explanatory theory? I would say that it *is*: it is a very simple model without any phenomenological terms or parameters.[1]

In the next section, we shall see how the two senses in which a model can be realistic are related to metaphysical realism about laws, theories and theoretical entities.

3.5 Three types of realism

Since metaphysical arguments are very important in Cartwright's analysis of scientific explanation, as we have seen here and there in this introductory chapter, I conclude with a section especially devoted to metaphysics. Perhaps Cartwright's metaphysics is linked to her model of scientific explanation in the same way that she says realism is linked to the DN model? As in the rest of the chapter, I intend not to interfere the presentation of Cartwright's ideas with my criticism, which especially in this section I find hard to postpone until chapter 4.

3.5.1 Realism about laws

Cartwright thinks that the explanatory role of so-called laws of nature is limited, because these laws are derived from within a model, so that they are not covering laws of real physical systems. Since nature is messy, no two real situations are the same and therefore every situation has its own covering law. Or better, *perhaps* it has a law, because we can not know about that: such a law would be so complex that at the moment we do not have real covering laws (not even for hydrogen atoms!) and we probably never will. So why assume laws that never have been formulated? The statements which have been formulated and are called 'laws', are not covering laws of any real physical system. Therefore these laws are not real.

In the DN model, many physical systems are 'nomically linked' in that they obey the same law: "All planets circling around the sun sweep out equal areas in equal times". But, for the sake of the argument assuming that every situation can be described by a different covering law, Cartwright finds that these different covering laws do not produce links between different systems. What we have instead is a 'patchwork of laws'. Cartwright proposes quite a radical change of scientific world picture, from a tidy organized world where many physical systems are governed by the same laws of nature, to a messy disorganized world where every physical system has its own law. She baptized her metaphysical attitude towards laws 'nomological

[1]It seems a bit odd to call this simple model of the hydrogen atom realistic. The simplest model of hydrogen is no model at all, and because in no model there are no phenomenological aspects, 'no model' is as realistic as the simple model! This makes no sense, so that perhaps one should require that a model can only be called realistic if it is realistic in both senses.

pluralism'([6], p. 322). The change is so radical that the dramatic motto of this chapter seems appropriate.

3.5.2 Realism about theories

Is the world out there governed by the equations of our physical theories? The simulacrum account of explanation links theory via models to reality. The question of the reality of theories will therefore be answered in two parts: Is reality as described in the models? And secondly, if models are 'real', does their reality carry over to the theories?

Let us consider models first. The common sense meaning of the word 'model' can be described as an 'idealization of part of reality'. In some sense it is a cliché to say that a model is not the real thing. But let us see what Cartwright thinks about it. While investigating the reality of models, it is instructive to recall the discussion about the meanings of 'realistic' among physicists. Here the first meaning is important: a model is more realistic if it is less idealized. It suggests that models can be ordered as more and less realistic. One might think that by steadily improving realistic models, one could arrive at a most realistic model which is real in the metaphysical sense. In her book *Nature's Capacities and their Measurements*, Cartwright argues that this concretization process can not have anything to do with realism:

> "... this kind of process will never result in an even approximately correct description of any concrete thing. For the end-point of theory-licenced concretization is a law true just in a model" ([5], p. 207).
> "After a while, it seems, in any process of concretization, theoretical corrections run out and the process must be carried on case by case" ([5], p. 207).
> "The physics principles somehow abstract from all the different material manifestations . . . to provide a general description that is in some sense common to all, though not literally true of any" ([5], p. 211).

Therefore, even the hypothetical models which are the limits of a concretization process are not real. Cartwright does not refute the cliché that a model can't be the real thing.

Another reason that models are not real, is that

> "... some properties ascribed to the model will be genuine properties of the objects modelled, but others will be merely properties of convenience. (...) Some properties [of convenience] are not even approached in reality. They are pure fictions" (p. 153).

In a concretization process, one would still be left with the 'fictions' in the model, so that a model which contains 'fictions' can never approach the real thing.[1]

There are other arguments why models are not real: there are many situations which are described by the same model. And vice versa, the same system can be described by several models. Therefore the models and the laws derived in it do not literally apply to the part of reality which they are meant to describe. The unreality of models carries over to theories: at best, theories are true about models, because the models were constructed that way (p. 156). Since models (just like laws) do not literally apply to any real system, neither does the theory. Theories which are not literally true of any real system cannot be real.

This being said, Cartwright remarks that there might be exceptions to the rule that theories do not describe real things:

[1] Cartwright argues that probability distributions in classical physics are an example of such 'pure fictions'. I shall not discuss that now.

> “I do not mean that there could never be situations to which the fundamental laws apply. That is only precluded if the theory employs properties or arrangements which are pure fictions (...) One [situation] may occur by accident, or, more likely, we may be able to construct one in a very carefully controlled experiment, but nature is generally not obliging enough to supply them freely” (p. 161).

Thus, leaving exceptions aside, the fundamental laws are too neat and clean to apply to reality.

One could ask whether theories could not somehow side-step the unreality of models. Couldn’t one say that a theory (and the laws which can be derived from it) is literally true about a system, even if one cannot construct a precise model of the system?[1]

In 1989, she writes that there may be one theory for which this is true: mechanics.

> “Among other modern sciences, it is mechanics alone that can claim to be able to derive laws that are literally true. For it assumes that ‘force’ is a descriptive term, and adopts the pretence that all changes in motion are due to the operation of forces” ([5], p. 210).

Mechanics might be true about a system, even if you do not know what forces are working on it. In 1992, she advocates a more radical view: she even calls the belief that laws are at work in situations where we do not have a model *fundamentalism* ([6], p. 316). She gives the example of a 1,000 dollar bill on a windy square: as long as we do not have a model of this situation from which we can derive its erratic motions, we can not say that the situation is governed by fluid mechanics, Newtonian mechanics or any other theory, unless we are prepared to be called fundamentalists. Therefore, the reality of mechanics, as well as any other theory, cannot do without the reality of models.

3.5.3 Realism about theoretical entities

One can be a realist in several ways. John Worrall for instance calls himself a structural realist. He states that at the ontological level, successive theories change drastically, whereas the structures of successive theories do not change much: nature is governed by the same equations [41]. In a way, Nancy Cartwright takes the opposite stand: the equations of the fundamental theories do not apply to nature, but theoretical entities do exist. Here is her ‘creed’ of theoretical entities:

> “I believe in theoretical entities and in causal processes as well. (...) All sorts of unobservable things are at work in the world, and even if we want to predict only observable outcomes, we will still have to look to their unobservable causes to get the right answers” (p. 160).

It is clear that in saying this, Cartwright opposes Van Fraassen’s constructive empiricism, which holds that only observable consequences of a theory should be considered real; the rest is fiction. Cartwright is a bit vague in which theoretical entities precisely she believes. Here are a few of them:

> “... I recognized that real materials are composed of real atoms and molecules with numerically specific masses, and spins, and charges; that atoms and molecules behave in the way that they do because of their masses, spins, and charges ...” (p. 128).

[1] This idea is especially tempting if a similar situation was one of the rare situations where the fundamental laws literally applied...

For me it came as a surprise that after so much anti-realist talk, Cartwright starts to talk as a realist. But I do not think that her realism about entities is inconsistent with her anti-realism about theories: it is hard to model systems consisting of real theoretical entities realistically, again because nature is messy: physical systems are always more complicated than their models. Still, I think that her realism about theoretical entities is a 'disconnected belief', in the sense that she neither presents arguments for her realism nor puts her entity realism to use in her philosophical project.[1]

3.6 Conclusions

In this chapter, Nancy Cartwright's ideas about scientific explanation have been presented: the inadequacy of the DN model and its metaphysical bias show that it should be replaced by another model of explanation, the simulacrum account, which is much more informal. The main point of this account is that everywhere in physics, models are used. The choice of a model is not simply dictated by the facts; physicists choose models for their explanatory purposes, and not to be as realistic as one can be. This has some anti-realist consequences: theories are not true about real systems, but about models. Phenomenological laws give a more accurate description of reality than laws which are derived from a theory in some model.

It took me two chapters to set the stage, but now on the one hand there is Hempel's Deductive-nomological model and on the other there is Cartright's simulacrum account, two rivalling models of scientific explanation. In the following chapters we shall judge ourselves which one is best. It is not just a battle in which one model wins because it is the more accurate description of explanatory practice in science. The battle is more loaded, because of the metaphysics involved.

In the next chapter, I shall discuss in what respect Cartwright's simulacrum account of explanation is liable to the same criticism which was formulated against the DN model in chapter 2, and in what sense it is better or worse. I shall also be critical about some of the metaphysical arguments which are used in the 'battle between models'.

[1] Contrast this with Ian Hacking [19], who bases his entity realism on observations and manipulations in experimental physics.

Chapter 4

Criticism of the two models

Two models of scientific explanation have been introduced. In chapter 2 about the DN model, I briefly discussed four points of criticism that were raised against that model: theoretical ascent, proviso's, its questionable explanatory function, and finally its alleged inability to cope with quantitative arguments. In this chapter, I shall compare both models of explanation by considering those points of criticism again: does the simulacrum account withstand the criticism better? During this test of the models, I shall also consider whether I think the points of criticism create serious problems for the models. I think that especially the problem to deal with quantitative arguments is pressing for both models, for reasons that I shall not mention here, since I am not supposed to give the game away so soon.

As in Cartwright's own writings, in the discussion of this chapter we shall frequently end up in a metaphysical argument; the reader is invited to find out whether this can be done without obscuring the presentation.

4.1 Theoretical ascent revisited

Which theory should we apply and which can be neglected in a specific situation? As we saw in chapter 2, Hempel was worried because the DN model of explanation does not model these first steps in the explanatory process: only after one theory has been chosen and others neglected, a deduction with that one theory as a premise can be represented in the DN model. In this section and the next one, I try to find out whether this should count as a big problem for the DN model and secondly, whether the simulacrum account of explanation solves it.

In a DN argument some physical theories are used as a premise and others are not. Theories which are used in the explanation of a phenomenon must be considered as true premises, or else the conclusion is false. But truth is not enough: Einstein's theory of general relativity may be true, but that does not mean that it can be used to explain the phenomenon of superconductivity. Evidently, in order to explain the phenomenon, the true theory must also be 'at work' in the physical system that exhibits the phenomenon. We can only apply the theory of magnetism to a system consistsing of a metal bar to which iron filings cling, if the bar can be *identified* as a magnet. However, the theory of magnetism plus the observation language ('metal bar') cannot be used in a deduction to show that the metal bar indeed is a magnet. And if deduction is the only acceptable form of logic in the empirical sciences, then the identification process constitutes a logical gap in scientific explanation.

The extra requirement that a theory not only is true, but also applicable, is not made explicit in the DN model. In principle, this creates the possibility that we make predictions using a valid deduction and a true theory, which have no relevance to the system under study. To put it dramatically, what can be the use of a deductive model of explanation in which all premises may be true about a system, while the deductively true conclusion in fact is false since the theory does not apply? Does that model truly capture the essence of explanation? In a worst case scenario, the false prediction might even refute a true theory ... After this emphatic reintroduction, I shall discuss two solutions to the problem of theoretical ascent below.

The first solution consists in 'repairing' the incomplete deduction by adding an extra premise, that applies the theory to the specific situation: *"If theory T is true, then it applies to phenomenon Ph of system S."* This way, the truth of the theory is separated from its applicability in a particular situation. This applicability premise is a bridge principle from a theory to an *individual* physical system. One could object to this way of patching up that it is begging the question: in the DN argument we are trying to explain the phenomenon by the theory, and even before we know that the theory can do that, we must use it as a premise. However, we are not begging the question, since the extra premise is not the prediction that we are trying to derive: the prediction is not a statement about applicability, but a quantitative law or phenomenon. If the law or phenomenon is indeed observed, then this strengthens the whole deductive argument, including both the theory and the extra premise about the relevance of the theory: one can indeed say that the theory explains the phenomenon. On the other hand, if the predictions and observations do not agree, then at least one premise is false. But the conclusion need not be that the theory is false, because it might also be that it is a true theory that does not apply in the situation at hand.[1] In that case, the theory is still considered true, but not applicable where it was considered applicable before.

Of course, the above procedure is methodologically suspect, because it can degenerate into a 'conventionalistic stratagem' (see Popper [35], p. 82-84): one can always assume auxiliary hypotheses that prevent the refutation of theories. Therefore, if the conclusion is that the theory is true but not applicable, then this must have consequences: one should explain *why* the theory is not applicable, or the situation should be branded an anomaly. The latter means bad publicity for the theory, but not always its refutation (see Lakatos [32], p. 94). Thus, by adding the extra premise of applicability, the logic of explanation is not refuted if Lakatos' criticism of Popper is right. The extra premise saved the DN model, but as a consequence DN arguments can not be judged separately anymore. This will be discussed in more detail in sections 4.2 and 4.3.

This first solution of the problem of theoretical ascent by an additional premise can be related to earlier work by Hempel. The subsumption of the phenomenon under the theory, or, to say it the other way round, the hypothesis that the theory applies to a specific case, is as bold a conjecture as the conjecture that the theory is true. The extra premise in the deduction is special in that its truth value is not known until the observation has been made. Hempel considered DN arguments of this type: he calls them *self-evidencing* arguments:

> "... arguments of the form (D-N) in which the information or assumption that [the explanandum] is true provides an indispensable part of the only available evidential support for one of the explanans statements ..." ([21], p. 372)

But should one call DN arguments containing an applicability premise a self-evidencing argument? Hempel restricts the use of the term:

[1] We do not consider the possibility that other premises than the theory and the applicability premis are false, because these cases are not interesting here.

> "It might be held that the actual occurrence of the explanandum event always provides some slight additional support even for an explanans whose constituent sentences have been accepted on the basis of independent evidence, and that in this sense every D-N explanation with true explanandum is in some sense self-evidencing; but we will apply this appellation to an explanatory account only if (...) the occurrence of the explanandum event provides the only evidence, or an indispensable part of the evidence, available in support of the explanans-statements" ([21], p. 372-373).

I think we can say that the problem of theoretical ascent shows that (almost) every DN argument becomes self-evident: from the fact that iron filings stick to a bar we cannot deduce that it is a magnet. Therefore, in a DN argument about that bar, we need a premise that the theory of magnetism is true and another that it applies to the bar; the fact that the explanandum really happens provides an indispensable part of the evidence of the applicability premise.

My analysis seems to support Nancy Cartwright in her crusade against fundamentalism: the problem of theoretical ascent shows that a theory can only be considered 'at work' in a phenomenon if there exists a (self-evidencing) explanation of that phenomenon, based on that theory. There is a subtle difference, however. One can still be a metaphysical realist and believe that laws apply universally. At the same time this realist may need the applicability premise because he is *ignorant* about the applicability of a law in the situation at hand. For that reason, the problem of theoretical ascent can be analyzed without discussing metaphysics. Theoretical ascent does not support Cartwright's metaphysical pluralism.

One might object that if we apply a theory in a DN argument in order to explain a phenomenon, we already implicitly assume that the theory is applicable, so that it is not necessary to add an extra premise to the argument: the truth of the theory is stated explicitly as a hypothesis, whereas its applicability in a specific case is an implicit hypothesis. But that would create the possibility that the explanandum is not observed (is false), while all (explicit) premises are true: that could happen if you leave out the identification of the metal bar by a magnet out of the argument. If it happens, we either must change the implicit assumption that the theory applies into an explicit one, or blame one of the true premises. Therefore, I think it is better to make the applicability-premise explicit right from the start.

One could also say that the applicability premise is not necessary, because by the truth of the theory we *really* mean the truth of the theory in the situation at hand. But this option is not available, unless we are prepared to admit that we never refute theories, but rather theories in specific situations. However, this would have the radical consequence which one could call a 'patchwork of theories': every physical system has its own theories to be formulated and tested. Again, I think that the problem of theoretical ascent can be analyzed without resorting such an extreme metaphysical position.

The second solution to the problem of theoretical ascent is only half a solution: it consists of showing that theoretical ascent is a non-problem *for explanation.* It *is* a problem for prediction, though. In the first solution, prediction and explanation were treated on a par, but there is a difference between them that is important now: if you predict, you do not know yet whether the prediction is right, whereas if you explain a phenomenon by a theory, you already know that the phenomenon has successfully been predicted by the theory. If the theory has not been sucessful, then you simply do not have an explanation. We therefore need not worry about magnet-look-alikes to which we naively apply the theory of magnetism: our predictions fail, so there is no explanation. Perhaps we find out that the magnet-look-alikes are not real magnets and on the basis of that make a new prediction. But also if we really do not know what goes wrong with our predictions, that is: in case we have an anomaly, there is no problem for explanations: anomalies (by definition) can not be explained. We can only light-heartedly declare theoretical ascent a non-problem if we consider explanation a retrospective activity that follows successful prediction.

The problem of theoretical ascent raised when discussing the DN model of explanation, and until now have we considered ways of solving the problem within that model. However, since in chapter 3 Cartwright's simulacrum account of explanation was presented as an alternative, we shall now see whether theoretical ascent also is a problem in that model of expanation.

In the simulacrum account of explanation, both theory and phenomena are used in a first informal step in the explanatory process, the creation of an 'unprepared description' of a system (section 3.3). For this description, all possible theoretical terms can be used if we think them necessary for 'descriptive adequacy'. In her description of the explanatory process, Cartwright does not show any worries about theoretical ascent: just like in the DN model, it seems evident which theory should be used in the explanation. But it can be a problem to identify a physical system as a magnet, or as a helium-neon laser. Cartwright presupposes that some rough identification of the system is unproblematic: she does not say: "This system behaves *as if* it is a helium-neon laser", but rather "This He-Ne laser behaves *as if* it is a collection of three-level atoms, etcetera." On the left hand side of the 'as if'-operator we do not only find what Cartwright calls an existential commitment, but also an identification of a physical system *as a* such-and-such:

> "There are well-known cases in which the 'as if'-operator should certainly go all the way in front: the radiating molecules in an ammonia maser behave *as if* they are classical electron oscillators. (...) The classical electron oscillators are undoubtedly fictions. But, even in cases where the theoretical entities are more robust, I still wanted to put the 'as if' operator all the way in front" ([4], p. 129).

It seems that Cartwright's belief in theoretical entities prevents the 'as if' operator from going even more 'to the front'. Cartwright seems to be saved by her metaphysics, but that is not so: theoretical entities not only are supposed to exist, but also to be identifyable. Cartwright does not show *how* theoretical entities are identified. Hempel may also believe in theoretical entities, for example in magnets. But the problem of theoretical ascent is how in a concrete case, a system can be identified as such. Cartwright does not consider theoretical ascent a problem, but it is not a smaller problem for her account of explanation than for the DN model. Now we turn our attention to provisos, which are another problem for the DN model, and perhaps also for Cartwright's model?

4.2 More about provisos

Provisos are another problem for the DN model, as we have seen in section 2.3: what reasons do we have to assume that theories which we do not consider relevant for the explanation *really* do not matter? The motions of iron filings in the presence of a magnet can not be described by the theory of magnetism alone if strong air currents are present. The 'theory of wind currents' might be irrelevant after closing the window, but other theories may still be at work.

If we try to predict a phenomenon by a theory, we assume that other theories are irrelevant. As we did in when considering theoretical ascent, we could make this implicit assumption explicit and formalize the problem of provisos by adding an extra premise in every DN argument: *"Other theories are irrelevant"*. The extra premise is a strong assumption, and its truth value can only be determined *after* the comparison of prediction and observation: provisos are another reason why all DN arguments are self-evidencing arguments (in the sense as described in the previous section). In fact, the whole analysis of theoretical ascent can be used here for provisos. Again, the threat of a conventionalistic stratagem is looming, but for the same reasons as mentioned in relation to theoretical ascent, I do not think it is a real threat. The problem of provisos can also be solved partially by arguing that it is a non-problem for explanation.

Are provisos a problem for Cartwright? In the simulacrum account, we do not need an extra premise to prevent false predictions from causing true theories being refuted, because the model in the simulacrum account can play that role. If a prediction fails, then we must choose another model that serves our predictive purposes better. It is the model that is refuted, not the theory. But this easy way of dealing with provisos does not imply that the simulacrum model of explanation is the superior one.

The reason is that empirical testing of theories becomes problematic, because we can always blame the model. This is the 'simulacrum version' of a conventionalistic stratagem. In Cartwright's analysis, the route to the conventionalistic stratagem is open: theories are not about reality but about models, models are distortions of reality and they are successful if they serve their purpose; if the model does not serve the purpose, you need another one. However, if nothing else could be learnt from the failure of a model than that some other model should be tried, how can we find a better model next time? In order to arrive at another model, at least the prepared description of the system must change, and perhaps the unprepared description changes as well (see section 3.3). But does the failure of one model influence the construction of the next one? And how much choice do we have in model choice, given a theory, a purpose and a physical system? And - this is especially important in relation to provisos - will the same theory be used in the next model or will other theories be considered as well? *Only if* these are unanswerable questions, if model choice defies all logic, then Cartwright's conclusion follows that we can always blame the models and never test a theory. This is precisely what Cartwright needs for her central thesis that theories are about models and not about reality. However, if choosing models is not as arbitrary as Cartwright suggests, then perhaps theories *can* sometimes be tested by experiments.

I disagree with Cartwright that the only ingredients for model choice are the theory and the system itself; models stand in relation to other models in the same and in other theories. Models do not come and go so easily without consequences for the theory. Therefore, the failure of a prediction in a model not only has repercussions for that model, but also for the theory to which it has been adapted. I shall have more to say about this in the next section, which is about theories of explanation.

4.3 How is explanation possible?

In section 2.4 I argued that one should not put forward too stringent requirements for an argument in order to be explanatory. Let us rather grant that scientific explanation exists and study how it can be characterized. After that, it is time to ask the Kantian question how scientific explanation is possible. The philosopher who studies scientific explanation needs both a model and a 'theory': in the first place he needs a model that gives an empirically adequate account of how explanation works. All information that is used to explain should get a place in the model; the logic of the explanatory argument should also be captured in the model of explanation. In the second place, he needs a 'theory' of scientific explanation that goes beyond description: the theory must show how explanation, as described in the model, is possible.[1] I introduced the theory in brackets here, because the theory of scientific explanation will not be a physical, but probably a metaphysical one. In this section, we compare the two models of explanation with regard to their empirical adequacy and the theories of explanation accompanying them.

For a realist about laws of nature, it is clear why the DN model of explanation works: in nature laws are at work. These laws make the natural phenomena happen. If we know the laws and the specific circumstances of the physical system in which they are at work, we can deduce (predict and explain) which phenomena show up in the system. Thus, metaphysical realism about laws of nature can be a 'theory of explanation' for the DN model. For a realist about laws who is also a 'normative' deductivist, it is also clear how explanation looks like: if you want to explain something, there is no other option than to try and find a general law that covers the situation, and to subsume the situation under that law by using the law in a deductive argument about the situation. Because of this normative aspect, the DN model is a descriptively adequate model of scientific explanation. I leave it as an open question whether the DN model can get 'theoretical support' from other theories of explanation than metaphysical realism about laws.

Cartwright criticizes both the DN model and realism about covering laws. In the first place, she argues that many scientific arguments are not deductive and do not use laws. For these reasons, the DN model gives an inadequate description of explanation. In the second place, Cartwright shows that the laws that are used in scientific explanations are not literally true of the situation to which they are applied. Therefore, she says, realist metaphysics about laws can not be based on DN explanations, not even when they are successful. Conversely, if real covering laws either do not exist or can not be known, then this suffices to show why the DN model can not explain explanation. So far for Cartwright's criticism of the DN model. In the next section we shall study more thoroughly how she drives the wedge between the DN model and a possible metaphysical basis. Now it is time to study the empirical adequacy and theoretical support of the simulacrum account of explanation.

In the first place, does the simulacrum account give an accurate description of actual scientific explanations? This question can best be answered by considering a great many explanations in physics. Here, I shall stick to a more general discussion, but in chapter 5 some physical models will be discussed. Cartwright's model of explanation links theory through models to specific situations. I think she is right in stressing that scientific explanation can only be understood if we understand the function of models in physics. Models are everywhere in physics. Cartwright describes how physical systems are being modelled, but modelling has already started before that: even to *call* a part of reality a 'physical system' means that you make a model in which you neglect the (details of) physical interactions of that part of reality and others, gravitational attraction for example. The simulacrum account of explanation is empirically more adequate than the DN model, in the general sense that the former makes explicit that our theoretical reasoning in the first place is about models of reality.

However, I do not think that Cartwright is sufficiently clear about the process of model choice. I expressed some of my criticism in the previous sections. Here I shall suppose that theoretical ascent and provisos are not a problem: we know which one theory we should use in a given situation. Then the process of modelling begins. The first stage consists of making an unprepared description, using whatever information we have about the theory and about the system. In the second stage, the

[1] To be clear, I do not mean that this theory of explanation must show why the physical theory that is used to explain, *exactly* has the form that it has. Philosophical theories that predict the form of physical theories are rare. A philosophical theory of explanation should rather explain how in general theories and laws can be used to explain.

prepared description of the system is made, in the form of mathematical equations and relations. Then the model is finished and laws can be derived in it. Cartwright remarks that often the same models are used, for instance the harmonic oscillator model in quantum theory. Now theories have great explanatory power if they can be used to explain a lot by using only a few models. Therefore, if you strive for great explanatory power of your theory, then you should not choose new models all the time. But Cartwright's analysis of model choice does not make clear *how* it can be that one often arrives at old models for new situations, because in her analysis only one theory and one system are used to construct the model of that system. The multiple use of models can only be explained, I think, if also the stock of already existing successful models plays a role in model choice. This would be the case when you first try one of the existing models in a new situation. This procedure of repeated application of (almost) the same model in the same theoretical framework, is what Oliver Darrigol [9] calls a 'horizontal analogy'. There are also 'vertical analogies', for instance when a quantum theoretical version of a classical model is constructed (see Brigitte Falkenburg [13] for an excellent discussion). Cartwright does not account for these links between different theories and models in physics, thereby presenting physics as a patchwork of dissociated theories and models, more than it really is.

Cartwright, in her attempt to show that theories are not real, also overlooks the fact that models and laws not only appear in between theory and experiment, but also without the support of theories. Especially in new areas of physics, physicists usually *start* with making models and laws to explain the phenomena. For example, Kepler formulated his laws (including the by now well-known area law) before the advent of Newtonian theory.[1] Cartwright is in trouble both when these laws are examples of the phenomenological laws that in her view form the content of scientific knowledge, and when they are not.

First suppose they are 'phenomenological'. Then they are excellent examples of phenomenological laws that can be derived from a theory in a two-body model of the solar system. But that is not a problem for Cartwright, because she could say that this shows that the theory is only true in a distorted model of reality. The problem for Cartwright is that the theory also shows the limits of applicability of Kepler's laws, by accounting for the observed influence of gravitational interaction with other planets in a perturbative way. Via perturbation theory, the theory extends beyond the model and this way corrects the phenomenological law. This contradicts Cartwright's thesis that the content of scientific knowledge lies in the phenomenological laws, and questions her claim that theories are not as real as phenomenological laws. Another problem for Cartwright is that Kepler's laws are very neat and clean mathematical relations between quantities describing planetary motions, so that if they are phenomenological, they are a counterexample to her claim that phenomenological laws are always 'messy' or 'complex'. Secondly, if Cartwright on the other hand would say that Kepler's laws are not phenomenological laws, then what kind of laws *are* they, given that they were not derived from abstract theories? The simulacrum account tells an incomplete story about model choice and a questionable story about phenomenological laws.

The simulacrum account also tells an incomplete story about the derivation of laws which are made within the models. Cartwright herself says that the derivations of laws in models often is not deductive, but she has not an alternative, as we saw in chapter 3. This is another reason why the simulacrum account is an incomplete model of explanation. Cartwright would agree with me that the model is incomplete, but I think it has more gaps than she would admit.

Is there a theory of explanation behind the simulacrum account? The main

[1] For the history of (celestial) mechanics, see for instance Cohen [7] or Dijksterhuis [10].

objective for Cartwright to write *How the laws of physics lie* [4], was to argue against the standard conception of scientific explanation. She showed that the standard view must be revised, by pointing at the role that models play in explanations in physics. But does her 'anti-philosophy' have a metaphysical basis? In the remainder of this section, we shall study how Cartwright's ideas about that changed in the course of time.

In [4], Cartwright does not supply a theory of explanation. She mentions situations where both the simulacrum account and the DN model fail:

> "The best theoretical treatments get right a significant number of phenomenological laws. But they must also tell the right causal stories. Frequently a model which is ideal for one activity is ill-suited to the other, and often, once the causal principles are understood from a simple model, they are just imported into more complex models which cover a wider variety of behaviour. (...) But I do not have a philosophic theory about how it is done. The emphasis of getting the causal story right is new for philosophers of science; and our old theories of explanation are not well-adapted to do the job. We need a theory of explanation which shows the relationship between causal processes and the fundamental laws we use to study them, and neither my simulacrum account nor the traditional covering-law account are of much help" ([4], p. 161-162).

The theory that Cartwright tries to find is a metaphysical theory of explanation. I think that for Cartwright it will be hard to find such a theory, without disclaiming some of her own criticism. I shall explain why I think so. First I shall reformulate the problem. In the context of justification[1], a model of a physical system is tested by comparing predictions in the model with experiments. In the context of discovery, Cartwright gives functional, Kuhnian explanations why physicists use the same model in many situations, which is a fact, she claims, that has anti-realist consequences. Now she is looking for a metaphysical theory of explanation that somehow should link the two contexts: it must explain why the convenient procedure of using the same models over and again is justified in experiments. How, then, can Cartwright put her metaphysics to work in such a theory of explanation, while at the same time upholding that the multiple use of models has anti-realist consequences?

But let us not stop at this rhetorical question, but try to answer it. Cartwright is a realist about theoretical entities, but not about theories. If she wants to find a *realist* theory of explanation, it stands to reason to base that theory of explanation on her entity realism. As a first shot, one could try to explain why a model that could be used in a simple case, can also be used in a more complex case, from the observation that in both cases the system under study consists of the same kind of real particles (theoretical entities). But I do not have a clue how to pass from theoretical entities to models, without using theories. Cartwright can *not* explain the double applicability of the model from the fact that the two situations are described by the same theory, because in her view, theories are not about reality, but rather about models. The fact that the same theory describes the same model is not informative, and does not address the question why that model can be used more than once in reality. From this first failed attempt I conclude that it must be very hard to side-step Cartwright's anti-realism about theories and models that she defends so strongly. I think Cartwright is in an impossible position to find a realist theory of explanation that solves the problems she wants it to solve.

In her second book, *Nature's Capacities and their Measurement*, Cartwright again studies the modelling process, which she then calls 'the problem of material

[1] Cartwright herself does not make the distinction between the two contexts, but I think I can use the distinction in the presentation of her views without distorting her ideas.

abstraction'. The main purpose of the book is to argue that physical systems have *capacities* to do things and that capacities are real.[1] Unfortunately, the reality of capacities does not help to formulate a theory of explanation that shows how material abstraction can work:

> "I think we cannot understand what theory says about the world nor how it bears on it, until we have come to understand the various kinds of abstraction that occur in science, as well as the converse processes of concretization that tie these to real material objects. (...) I look for an explanatory structure that will make sense of the methods of abstraction and idealization that are at the heart of modern science, and mere laws are not enough for this difficult job. What kind of different explanatory structure is required?" ([5], p. 211-212).

Cartwright leaves the question open; the theory of explanation she needs can not be found in her second book either.

In her [6], Cartwright fights against fundamentalism: a universal realist metaphysics should be based on more than local explanatory success:

> "I am prepared to believe in more general theories when we have direct empirical evidence for them. But not merely because they are the 'best explanation' for something which seems to me to need no explanation to begin with. (...) Theories are successful when they are successful, and that's that. If we insist to turn this into a metaphysical doctrine, I suppose it will look like metaphysical pluralism" ([6], p. 322).

Cartwright's 'patchwork-of-laws' metaphysical pluralism could be based on explanatory success. However, giving a theory of explanation would be just the other way round: explaining the success of scientific explanations with a metaphysics. But could metaphysical pluralism be used as a theory of explanation?

It would be an ultra modest theory, because it would not claim more than "this explanation, based on this particular law, is successful in this particular situation, because this is the one and only situation in which that law is true." The fact that the law applies to only one case makes the argument viciously circular.[2] The explanatory power of metaphysical pluralism as a theory of explanation is very limited indeed; straightforward metaphysical realism about laws does a better job, at the expense of not being literally true in any specific case. Thus, it is tempting to apply Cartwright's analysis of scientific explanation to 'metaphysical explanation' as well: she does not put metaphysical pluralism to use as a theory of explanation, because if it were true, then *the truth doesn't explain much*.

4.4 Towards a model of quantitative explanations?

Let me recall from the previous section that there are quantitative reasons to call a part of reality a 'system': the interaction with its environment must be small compared to internal interactions. It is not literally true that the system is isolated, but often it is a convenient simplification and a good approximation to treat it as such. Only quantitative arguments can justify why something can be treated as an isolated system. The model is not true about the system, but it must be good enough to obtain the desired accuracy of your predictions.

[1] Capacities are about properties, not about individuals. Real capacities are not simply the real properties of individual real theoretical entities: "... the property of being an aspirin carries with it the capacity to cure headaches. What the capacities of individuals are is another, very complex, matter" ([5], p. 141).

[2] If the law were universal, the circularity would be benign, because the law should be able to explain other phenomena in 'similar cases' to which the law applies.

The definition of a what counts as physical system in a concrete case, illustrates the predominant role of quantitative arguments in physics. Accuracy is an important element of any prediction or explanation. Therefore, a model of scientific explanation should also be able to deal with quantitative arguments, including arguments about the accuracy of predictions. In this section, we shall see that in this respect, both the DN model and the simulacrum account are incomplete models of explanation. I shall also give examples where I think the DN model can be stretched a bit to incorporate so-called self-consistent arguments, that are widely used in physics.

4.4.1 Approximately true covering laws?

Section 2.4.4 sketched the problem how to get from measurements with finite accuracy, to true statements that can be used as premise in a DN argument. The use of quantitative statements in the DN model is problematic in yet another sense. Hempel uses laws, of which he knows they are only approximately true, as premises in DN arguments. The law covers some cases better than others and a continuous transition seems possible from situations to which the law applies (almost) exactly, to situations where it does not apply at all. It seems necessary to add to the DN model a theory that tells what is meant by a covering law that is approximately true: for example, does anything in reality correspond to an approximately true law?

Cartwright also speaks in terms of truth and falsity of laws. For Cartwright, 'covering laws' are not laws that cover many cases, but rather laws that cover individual cases in every respect. One should be aware of this meaning shift when comparing Cartwrights and Hempel's writings! Covering laws, Cartwright says, do not literally apply, and *therefore they are not real.* For Hempel, covering laws need not be literally true in order to apply to a specific case. Cartwright only shows that covering laws *in her sense* are unknown; Hempel can easily show that there are many well-known examples of covering laws in his sense, but if someone wants to be a realist about these approximately true laws, one has to show how that is possible.

The strength and the weakness of Cartwright's argument is its wide range: it applies equally well to cases where a law is a very inaccurate description as to a case where the law gives rise to predictions which are correct only up to the tenth decimal place; in neither case the law is literally true, so it is false, so it can not be real. I am dissatisfied with this argument, which I call an 'inverted no miracle' argument. The 'no miracle' argument for realism is that the great predictive success of a physical theory cannot be a miracle, so that the theory must be real (see [30] and [36]). The 'inverted no miracle' argument that Cartwright uses here, runs as follows: every predictive failure of a law about a system, *however infinitesimally small*, suffices to show that the law is false about that system and therefore is not real. I do not like this argument, because it is too cheap. I think that an adequate realist response can easily be formulated.

A realist about laws might give the following response: "I disagree with you that I cannot be a realist about laws because the predictions based on them never are precisely realized in experiments. You may be surprised to find out that I would even *stop* being a realist about laws when my predictions turned out dead right! The reason is that -just like you- I know that I derived the law within a model. I treated the physical system as isolated from the rest of reality, whereas in experiments there are always small interactions with the rest of the world, which are governed by laws that I may or may not know. I believe that the law would have been true about the part of reality that we both call the 'physical system', *if that part had been totally isolated*, also because predictions turn out better the better the system is isolated.

That cannot be a miracle; this law must be real. In that quantitative sense, I can be a realist about *ceteris paribus* laws that govern the constituents of the physical world, without being unrealistically optimistic that the law I know gives predictions that are exactly right." This is how a realist might attack Cartwright's inverted no miracle argument against realism. The main purpose of this response was to show that not all realists about laws are vulnerable to Cartwright's criticism. Of course, a naïve realist about laws, who believes that real laws dictate outcomes of experiments with minute precision, has no adequate response to Cartwright.

By the same 'inverted no miracle' argument, Cartwright argues that the models in her simulacrum account are not real, and a fortiori that physical theories are not real. She makes even stronger claims, as we have seen in section 3.5.2: the processes of material abstraction and concretization "will never result in an even approximately correct description of any concrete thing". Since Cartwright is well aware that there are successful theories with which one can make very accurate predictions, I assume that she is not concerned with accuracy here. But neither can she be concerned with (convergent) realism, since in virtue of her inverted no miracle argument, approximately correct descriptions - if they existed - could have nothing to do with realism. So I conclude that I do not know what Cartwright means here.

A quote from the physicist Richard Feynman may shed new light on the discussion of approximately true laws. He argues that approximately true laws not only are important because somehow they describe what approximately goes on in nature, but also because precisely their approximate character can teach something new about nature:

> "The moons of Jupiter appeared to get sometimes eight minutes ahead of time and sometimes eight minutes behind time, where the time is the calculated value according to Newton's laws. It was noticed that they were ahead of schedule when Jupiter was close to the earth and behind schedule when it was far away, a rather odd circumstance. Mr Roemer,[1] having confidence in the Law of Gravitation, came to the interesting conclusion that it takes light some time to travel from the moons of Jupiter to the earth, and what we are looking at when we see the moon is not how they are now but how they were the time ago it took the light to get there. (...) when a law is right it can be used to find another one. If we have confidence in a law, then if something appears to be wrong it can suggest to us another phenomenon. If we had not known the Law of Gravitation we would have taken much longer to find the speed of light (...). This process has developed into an avalanche of discoveries, each new discovery permits the tools for much more discovery" ([14], p. 22-23).

This example by Feynman illustrates how approximately true explanations can be seen as very useful *partial explanations*; the difference between the predicted and the observed phenomena, and especially the regularity in this difference, call for a *residual explanation*. Feynman speaks about laws that are 'right', without being literally true. It is hard to pin down what it means for such laws to be 'right', but from the quotation it is quite convincing that science could not get started if laws were discarded for the reason that they do not give true predictions. Likewise, I would say, a model of scientific explanation would not get started if partial explanations can not exist in it; and a theory of explanation should somehow be able to justify these partial explanations.

[1] Olaus Roemer, 1644-1710, Danish astronomer.

4.4.2 Shaky derivations?

Now we turn from laws to derivations. Cartwright argues against the DN model that actually, derivations often are not deductions. Secondly, the derivations often involve assumptions that are made just in order to find the desired law:

> "Where approximations are called for, even a complete knowledge of the circumstances may not provide the additional premises necessary to deduce the phenomenological laws from the fundamental equations that explain them. Choices must be made which are not dictated by the facts (...) and different choices give rise to different, incompatible results" ([4], p. 106-107).

I think that often derivations are not as arbitrary as it might seem, and I shall illustrate this with the following example.

It is quite usual that a law L can only be derived from a theory T by assuming that a parameter p is small, $p \ll 1$, let's say.[1] If we do not have any a priori information about that parameter, Cartwright would say that we just choose the parameter to be small, a choice that is not dictated by the facts. I would rather say that it is a *self-evidencing* deduction (see section 4.1), where the premise "$p \ll 1$" is found to be true only after L has been observed. The conclusion that "$p \ll 1$" is assumed just to arrive at L is too rash for another reason as well, because there may be independent ways of testing that premise. As in section 4.3, I think that Cartwright forgets that every derivation in principle can be related to others.

The logic of the derivation is improved if it is not only self-evidencing, but also *self-consistent*. Self-consistent arguments are widely used in physics and in our example it may take the following form: the premise "$p \ll 1$" can be used in the derivation of L, and vice versa, once the law has been derived, L can be used to derive that $p \ll 1$. A derivation that is both self-evidencing and self-consistent, is quite convincing. A derivation that is self-consistent but not self-evidencing, is also helpful. It suggests that in the situation at hand the parameter p is not small enough to arrive at the law. In that case, the self-consistent derivation is just circular reasoning, but it gives the new information about the parameter p. A self-evidencing but not self-consistent derivation suggests that he parameter p is small, but that apart from the law L more information is needed to derive that fact. And in the final case, if the derivation is neither self-evidencing nor self-consistent, then one should derive another law than L, based on the assumption that another parameter than p is small. This shows that there is more between simple deduction and opportunistic derivation than Cartwright suggests.

Cartwright has more arguments why derivations of laws by theories are shaky. She claims that abstract theories are not to be taken real, because often predictions improve by approximating the theory. As an example, she discusses the derivation of the exponential decay law from quantum mechanics ([4],p. 113-118): there is not a rigorous theoretical derivation in quantum mechanics that excited states decay exponentially to lower states: some terms are neglected in arriving at the exponential decay law. Cartwright assumes that since experimentally no differences with exponential decay can be found, the exponential decay law is the true phenomenological law. She concludes that the law improves upon approximation, which shows that quantum mechanics is not literally true, so that it is not real:

[1] As an example, if perturbation theory is used in quantum mechnics in order to calculate energy shifts, then one must assume that the perturbative potentials are much smaller than the energies of the system. Usually, perturbation theory is applied right away and from the fact that the energy shifts are small compared to the energy differences between levels of the unperturbed system, one concludes with hindsight that perturbation theory could be applied.

> "... the exact answer will be a close approximation to the exponential law we are seeking; a close approximation, but still only an approximation. If a pure exponential law is to be derived, we had better take our approximations as improvements on the initial Schrödinger equation, and not departures from the truth" ([4], p. 118).

However, since the predicted differences between the rigorous and the approximated law can be estimated to be smaller than the experimental accuracy, it is *mistaken* to assume that "a pure exponential law is to be derived". It remains an open question whether the approximation leads to an improved law or not.

It is a pity that Cartwright does not discuss contrary cases, where a theory shows the limits of applicability or accuracy of an earlier phenomenological model or law, like BCS-theory did in superconductivity (see for instance [8]), and Newtonian mechanics did for Kepler's laws. Even if the theories are only true in a model, these theories improved predictions about real systems. A debate between realists and anti-realists can become interesting, when anti-realists who do not like the 'no miracle' argument explain why these success stories of science need not lead to some form of realism.

In this section, I have tried to show that a quantitative science will not often arrive literal truth and that a model of explanation should be able to deal with approximately true laws. Moreover, I have argued that deriving laws from theories need not be as arbitrary as Cartwright claims: if one accepts self-evidencing and self-consistent derivations, then much more derivations in physics are sound than Cartwright believes. I think that the DN model can be stretched to allow for self-consistent and self-evident explanations, but perhaps a strict deductivist does not accept that. Cartwright forgets that every assumption that is made in a derivation might be tested in other experiments; a parameter of a system that is small in one experiment, should also be small in another experiment under similar conditions: derivations are not isolated from each other. Finally, I have shown that the fact that 'the exponential decay law' can not be derived rigorously from quantum mechanics, does not show that quantum mechanics improves upon approximation. This result undermines one of Cartwright's arguments against realism about theories.

4.5 Conclusions

In this chapter, I presented my criticism of both the DN model and Cartwright's simulacrum account of explanation. I also presented some possible solutions. For example, I related the problems of theoretical ascent and of provisos by showing in sections 4.1 and 4.2 that both problems could be interpreted as one problem: they both make DN arguments self-evidencing arguments, which is a less strict form of deduction that Hempel also discusses within the DN model. Some would even say that self-evidencing arguments even are not deductive, but I disagree with that.

In section 4.2, I compared the DN model and the simulacrum account of explanation and asked whether they were empirically adequate and whether they are supported by a theory of explanation. I showed that Cartwright was looking for a metaphysical theory that did that job. I also showed that she did not find such a theory and also gave reasons why she was in an impossible position to find one. Neither entity realism nor metaphysical pluralism proved good candidates as theories of explanation.

Section 4.4 was about quantitative aspects of explanation: for Hempel it does not seem to be a problem that covering laws often are only approximately true; the DN model should therefore be extended to include arguments about the accuracy of predictions; metaphysical realism about approximate laws seems a problematic position. Cartwright argues that it is an *untenable* position. For Cartwright, approximately true covering laws are a *contradictio in terminis*; covering laws that strictly speaking are false, cannot be real. I also showed that many approximations that at first sight do not seem to be 'dictated by the facts', are not as arbitrary choices as Cartwright suggests, for two reasons: the 'choice that is not dictated by the facts' in these self-evidencing arguments may be put to the test in other situations, and secondly, approximation procedures often are self-consistent, so that observed laws 'argue for their own derivation'.

After writing this chapter full of criticism, I think that many attacks by Cartwright and others against the DN model, can successfully be repelled, especially if self-evidencing arguments are accepted in the DN model. However, I think that the widespread use of models in physics calls for a model of explanation in which a prominent place is given to 'models'. Cartwright's simulacrum account is such a model, but it is partly incomplete and partly mistaken. It is especially mistaken, I think, in that it assumes that links between models and theories do not exist. In the next chapter, I shall discuss several models in physics. Each of them will shed some light on the question what an improved model of explanation will look like.

Chapter 5

A gallery of models

Until now, the discussion of scientific explanation has been rather abstract. At the end of chapter 4, I concluded that a good model of scientific explanation should discuss the use of models. In this chapter, I should like to introduce a few models that are used in physics, for a number of reasons: in the first place, I should like to focus on how physical models are actually used in explanations, so that this empirical input can be useful. Secondly, the term 'model' is widely used throughout physics. It does not always have the meaning that Cartwright attaches to it. If a model of scientific explanation is to be constructed in which physical models occur, then it should be clear what 'model' means. If explanation in physics involves several types of 'models', then a model of scientific explanation should be able to cope with them. Thirdly, I want to make some points about the use of models that Cartwright does not mention. For example, she neglects the importance of mathematical tractability in the choice of models. And she does not discuss (potential) truth approximation of successive models, which I shall do in section 5.5. Fourth, this is an excellent opportunity to drag some interesting examples from physics into the philosophy of science that I have not come across there yet; Kepler's 'area law' has served its purpose. It is time for fresh examples!

The problem of introducing a few physical models is that they function in a wider range of physics that I am not able to review (here). The function of the models in this chapter is similar to the function of models in a fashion show: the purpose is not that you learn to know her (or him) well, although she (or he) may seem interesting in her own right. The model is there to show clothes, and likewise a glance at the physical model will serve to show aspects of models in scientific explanation. Often the same fashion model walks the catwalk several times, to present different clothes. In this respect too, the analogy with physical models in this chapter will be striking.

5.1 Heuristic models in nuclear physics

Many models in physics do not function as a bridge between a theory and physical systems in the way that Cartwright describes. For example, it may be the case that in a new area of physics theories are not available yet. In that case, the new field of research is explored by models. It may also be that there *is* a theory with which the phenomena in principle should be explained, but no one knows how to do it, because it is mathematically too complicated; there is a *computation gap*, as Redhead [38] puts it. In the latter situation, heuristic models may be used that are relatively independent from the theory. In this section, we shall discuss two heuristic models of atomic nuclei. Since they are models of the same systems, we can see how different models serve different purposes.

5.1.1 The liquid drop model

Physicists believe that in principle all properties of nuclei can be explained by quantum theory. However, since nuclei consist of many particles (nucleons), it is practically impossible to explain important nuclear properties 'ab initio' from quantum mechanics. An important question about nuclei is their stability: how can we understand which nuclei are stable and which are unstable (radio-active)? If you want to know why some nuclei can fall apart, you must study what forces keep a nucleus together. For that reason, physicists try to compute nuclear *binding energies* that prevent nucleons from falling out of the nucleus.

Since the nucleons do not collapse into each other, there must be a repulsive force if the nucleons are near each other, and since they stick to each other there is an attractive force at somewhat larger distances. For this reason, and because from experiments it is found that the number of nucleons per unit volume is nearly independent of the size of the nucleus, one can make an analogy between nuclei and a liquid drop, because drops have all these properties too. In the simple *liquid drop model* for nuclei, one tries to find nuclear analogues for the forces in a liquid drop. At first, no quantum mechanics will be involved.

Denote the number of nucleons by A and the number of protons by Z. The other nucleons are neutrons and there are $A - Z$ of them. The bold conjecture of the model is that the binding energy B is a function of A and Z only, and the analogue with the liquid suggests it has the following form:

$$B(Z, A) = a_V A - a_S A^{2/3} - a_C Z^2/A^{1/3}, \tag{5.1}$$

where the three terms at the right hand side are the so-called *volume term* (which is attractive), the *surface term* (the nuclear analogue of surface tension in a liquid, repulsive) and the *Coulomb term* (repulsive), respectively. This model, as it stands, gives several false predictions. In the book Nuclear and Particle Physics that I use for this section, by W.S.C. Williams, we then read:

> "The effects we must add in extending the so far too naïve model are quantum mechanical" (see [40], p. 57).

So, although quantum theory does not directly help you in predicting nuclear binding energies, it can help you in improving a simple heuristic (non-quantum) model. From quantum mechanical considerations, two terms are added to the formula for the binding energy, and the only thing that I shall say about them is that they are called the *asymmetry term* and the *pairing term.* Thus, we end up with a binding energy of a 'quantum liquid drop' as a function of A and Z, consisting of five terms:

$$B(Z, A) = a_V A - a_S A^{2/3} - a_C Z^2/A^{1/3} - a_A (A - 2Z)^2/A \pm a_P A^{-1/2}. \tag{5.2}$$

After the model has been corrected by the theory, the values of the parameters $a_?$ must be determined by fitting the equation to the experimental binding energies. That's how the model is applied to *and changed by* reality. Whatever the values of the parameters, the binding energy will come out as a smooth function of A and Z. However, experimental evidence shows that for small nuclei ($A < 20$), the binding energy is a rapidly varying function. From this general consideration, the liquid drop model is considered only correct for nuclei with $A > 20$ and only the experimental values of these bigger nuclei are used in the fit. This way, physicists improve the accuracy of their predictions by choosing the limits of applicability of their model. They find a fit that is not entirely correct, because also for larger nuclei the experimental binding energy as a function of A shows some bumps. But still, the resulting formula for the binding energy, which is called the *semi-empirical mass-formula*, gives a prediction of the binding energy of hundreds of nuclei, using only five free parameters.

The semi-empirical mass formula can also be used to *explain* the value of the binding energy of a particular nucleus that is characterized by nucleon number A and proton number Z. The explanation has the form of a deduction: given the formula for the binding energy and the numbers A and Z of a nucleus, an approximately correct binding energy can be deduced. But before that, the fit of the coefficients by the correct experimental values could also be called a deduction (where a criterion for 'best fit' is used as one of the premises). Thus, a particular binding energy is deduced from a formula, the coefficients of which were deduced from all binding energies. I would not call this strictly deductive explanation. Nor would I say that filling in a mathematical formula is an explanation: I think that it can only count as an explanation if all terms in the formula are supplemented by their physical meaning ('surface energy', etcetera). Is the formula a law of nature? It is not correct for light nuclei and neither for some nuclei where the binding energy has a sudden 'bump'. I would not call it a law, but rather a summary of nuclear effects that have been identified thus far. I do not think that a strict deductive model is ideal to represent explanations with the semi-empirical mass formula. Some nuclear effects that have been left out of the liquid drop model, for example the 'bumps'. This and other effects are accounted for in another nuclear model, that will be discussed now.

5.1.2 The shell model

There is another simple model that explains many nuclear properties. For example, it explains why for some numbers A and Z, the binding energy is extra high, so as to produce a 'bump' in the binding energy function. For these so-called *magic numbers*, the nuclei are extra stable. Magic numbers also occur in atomic physics: atoms of rare gases are very stable and inert, and their number of electrons are also called 'magic'. Using the so-called Pauli exclusion principle from quantum mechanics, one can reason that nucleons do not collide very often. Since the same is true for electrons in atoms, and since that fact led to the successful *atomic shell model*, one can assume that an analogous model can be made for nucleons. In the *nuclear shell model*, the nucleus is described as a single nucleon that is confined by a smoothly varying potential that is the average of the effect of all other nucleons. In other words, the complicated nucleus, consisting of many particles, can be treated as if it is just one particle in a potential. Qualitatively, quantum mechanics could justify the model, but it can not be used to predict the precise form of the potential that the single nucleon feels. This potential is chosen by trial and error: try a potential, compute the quantum mechanical eigenvalues for a particle in that potential, using the Schrödinger equation, and from the ground state energies of the nuclei determine the magic numbers of your model (in a way that I shall not discuss now). If your predictions fail, then try another potential. After some trial and error, it turns out that, just as in atomic physics, the only way to find the right magic numbers is to take the *spin* of the particle into account, by adding a so-called *spin-orbit interaction* term to the potential. The shell model is a heuristic model that can be used to predict many nuclear properties. Besides the magic numbers, also the spin of nuclei in their lowest energy state can be successfully predicted and explained.

5.1.3 Models have their purposes

As promised, for non-physicists the discussion of two models for atomic nuclei probably was not entirely clear. But at least it should have been clear that we have glanced at two nuclear models that have quite different starting points: the one is an analogy with a liquid drop, the other with an atom. We have seen that there are some theoretical or experimental reasons to make these analogies.

Both models have been adapted to the new situation to which they were applied. I do not know whether Cartwright would still call this a multiple use of the same model. However, I do not think that a theory of explanation is needed that explains why these models can be used outside nuclear physics as well. I think that *quantum* theory can be used both to explain the limits of aplicability of the models, and to justify the analogies.

The reasons why the analogies are made, for the main part belong to the 'context of discovery' of the new application. However, I do not mean to say that it is unimportant how models are chosen. The two examples from nuclear physics show that Cartwright's analysis of models choice neglects the influence that models from other areas of physics can have. We have seen how the same models have been used in different areas of physics. This strengthens the view that I expressed on several occasions in chapter 4. Connections between different areas of physics show up in the context of discovery. The sole *justification* of the model in the new situation, however, lies in the accuracy of its new predictions: that's how one puts the analogy to its mettle.

Of course, Cartwright is right in that different models serve different purposes, but I remember being puzzled during a course on nuclear physics that two models that give such a different 'picture' of the atomic nucleus, both can give insight in its behaviour. Williams says more or less the same:

> "The shell model and the liquid drop model are so unlike that it is astonishing that they are models of the same system. Of course, each model has limited objectives and limited successes. Neither model contains any assumptions that will allow, firstly, the prediction that for $A > 20$ the central nuclear density is almost constant, or secondly, the value of that density" ([40], p. 155).

Since both heuristic models are successful, neither of the models should be thought of as an exact replica of reality. The fact that there are *two* incompatible models, prevents us from believing the models too much. We should also be cautious if only one model is available.

A more rigorous treatment of nuclear properties is the *theory of nuclear matter*, that starts with information about the pairwise potentials between the nucleons, and attempts to predict nuclear properties, thereby also justifying the chosen average potential in the shell model. [1] The computations in the theory of nuclear matter are much more complicated than the ones in the above two models. That brings us to two other reasons to have models: you can explain some nuclear properties before you have rigorous results, and secondly, results from the models can be set as 'goals' for the more complicated theory.

The heuristic models have properties that are to be taken seriously in the yet unknown theory; which properties of the models should be taken seriously (as 'goals') and which not, cannot be decided logically. Redhead [38] remarks that at this point, there is an irreducibly intuitive element in science. It is precisely this intuitive step that Cartwright wishes to explain in a theory of explanation (see section 4.3).

[1] The theory of nuclear matter is a *more* fundamental, but not a fundamental theory, because the protons and neutrons that make up the nucleus are themselves made up of quarks.

5.2 Solvable models in quantum electrodynamics

In this section, we shall learn by example and find new arguments why the simulacrum account and the DN model are incomplete models of explanation. We shall pay extra attention to *mathematical* aspects of models: predictions and explanations based on models involve mathematical calculations. We shall briefly discuss how mathematical considerations can influence model choice, and we shall see how and why exactly solvable but unrealistic models are used.

5.2.1 Model choice and two-level atoms

The theory that physicists use to describe the interaction of light with matter, is quantum theory. But in the book *Optical resonance and two-level atoms* by Allen and Eberly about this matter, we can read:

> "It is impossible to discuss the interaction of collections of atoms with light exactly. It is impossible to treat even one atom's interaction with light exactly" ([2], p. 28).

This sounds disappointing, but let us see what precisely can be meant here. By an exact treatment it is meant that the solutions can be found of the quantum mechanical equations that descibe the interaction, *without making any approximations*. The phrase that an exact discussion is impossible does not mean that with quantum theory we can write down the 'true equations' that govern a system, which turn out to be impossible to solve. We are talking about models instead, and Cartwright is right in stressing that. To be concrete, when studying atom-light interaction, the atomic nucleus will not be modelled more realistically than as a single point, with electric charge and spin, but without internal structure. From the previous section we know that this model would not do in nuclear physics. But a more realistic model for the nucleus is quite unnecessary in order to explain optical phenomena of atoms. An 'exact' treatment of a system usually is physical nonsense, but an exact treatment of a model of a system makes sense.

It is quite easy to write down a Hamiltonian (model) of an atom, in which the known interactions between its electrons, and between the electrons and both the atomic nucleus and the light are all taken into account. One can believe that in principle, all optical properties of the atom can be explained using this Hamiltonian. Cartwright, however, denies this 'principle': only if you have actually computed an optical phenomenon with it, you can say that the model explains anything. The problem is that it is impossible to compute exactly any optical property of the system with this realistic Hamiltonian. In that sense, an exact discussion of the interaction between light and an atom is impossible. We therefore need to make various approximations. The first kind of approximation will be the choice of a simpler Hamiltonian. An enormous simplification that is often made in optics, is to treat the atom as a *two-level system*. In the previous section, I have criticized that the simulacrum account of explanation tells an incomplete story about the choice of models. In the following discussion about the choice of a two-level model, I shall argue that the simulacrum account is incomplete in yet another sense.

Quantum mechanics tells and measurements show that atoms have many discrete energy levels. However, atoms with only two energy levels do not exist in nature. What, then, are the reasons to treat real atoms as if they had only two energy levels? If light is absorbed by an atom, the atom literally makes a 'quantum jump' from a lower energy level to a higher one. Atoms can capture light better if the energy of the light is nearer to the energy difference between two such levels. If the light energy equals such an energy difference, then the absorption of the light is called *resonant*. Now, if the incoming light is nearly monochromatic (one colour, or equivalently, one energy, like in the case of laser light), and if there is only one pair of energy levels of the atom for which the light absorption is (nearly) resonant, then only these two energy levels need to be considered when computing, predicting and explaining the light absorption. That's why the atom can be treated as a two-level system (see [2]).

There are two points that I should like to make about this model. In the language of the simulacrum account, we have just attached a prepared description (two-level atom) to an unprepared description (a specific atom with known experimental values of its energy levels). Cartwright claims about this stage:

> "... no principles of the theory tell us how we are to prepare the description. We do not look to a bridge principle to tell us what is the right way to take the facts from our antecedent, unprepared description, and to express them in a way that will meet the mathematical needs of the theory. The check on correctness at this stage is not how well we have represented in the theory the facts we know outside the theory, but only how successful the ultimate mathematical treatment will be" ([4], p. 134).

If her claim were right, then explanation only starts once the model has been prepared. The theory would not have much to do with explanation in that case. However, in the choice of the two-level model, we have seen how the theory, the facts about the specific atom, and the explanatory purpose (optical resonance) together were used *to explain the choice of the model!* Theories are much more important in explanations than the simulacrum account shows, because the quantitative explanation of a phenomenon in a model is preceded by a qualitative explanation of the choice of that model.

5.2.2 Theory choice and spontaneous emission

We stay in the realm of optics and continue using the two-level model of the atom. If we shine light with the atomic resonance energy on an atom, then it can absorb the light while going from the lower to the upper energy level. The reverse process is also observed: in the absence of light, an atom in a higher energy state can decay to its lower energy state, while emitting light with energy equal to the energy difference of the two states. This process is called *spontaneous emission*, and its explanation requires a quantum treatment of the light, which was not necessary when considering absorption. Allen and Eberly write the following introduction to a chapter about quantum electrodynamics and spontaneous emission:

> "For the most part optics is insensitive to the finest details of the electromagnetic field. The possibility that quantization of the radiation field has important consequences is usually deliberately overlooked in both experimental and theoretical work. There are, however, two situations in which field quantization cannot be ignored. The first situation is encountered when one discovers an optical problem, almost no matter how idealistically stated, which can be solved exactly within the framework of quantized field theory. Such problems are rare enough to have an importance of their own, virtually independent of their relevance to possible experimental work. The second situation arises when an optical interaction causes an observed effect that can not be quantitatively accounted for except by a quantized field theory" ([2], p. 152)

The authors give examples of both situations: the best example of the first type was provided by Jaynes and Cummings, who showed that a single two-level atom obeys solvable equations of motion when interacting with a single-mode quantized radiation field [25]. The process of spontaneous emission is an example of the second type. The specific examples are not important now, but the two situations that were defined in the quotation are very interesting.

Let us consider the second situation first. What can we learn from it? In optics, usually the light is treated classically, whereas the matter with which it interacts is treated quantum mechanically. This is the 'semi-classical' approach.[1] The reason to treat the light classically is that it is mathematically simpler, while the predictions do not deviate much from a quantum treatment. At least, this is assumed as long as the classical predictions are right. From the outset, we are not sure whether the classical theory is applicable, but we know that is *is*, once the predictions turn out right in experiments. This sounds like a nice example of a self-evidencing argument (compare with section 4.1): when cast in the form of a DN argument, the truth value of the premise that states that a classical treatment of the light is applicable to the situation, is uncertain. As we shall see further on, there is independent evidence for the applicability in this case. In this example, here is a bigger problem than theoretical ascent for the DN model of explanation.

The problem has been discussed in other situations in sections 2.4.4 and 4.4.1: the classical theory must be considered a true premise for the sake of the DN argument. We know that in some other situations, the classical theory is refuted. It is useless to say that classical theory is "false but sometimes applicable", because that would not work in a DN argument. Neither can we solve the problem by deducing that the classical treatment of light is 'made true' by quantum mechanics. 'True' quantum mechanics for example can not be used to justify the use of 'false' classical mechanics in explanations about the dynamics of macro-objects like marbles and bulldozers, not only because the derivation in practice would be too complicated, but also because there are fundamental problems in applying quantum mechanics to macro-objects.[2] The classical treatment of the light must be considered true for the sake of the DN argument, but it is not clear what more is meant by this than that the treatment is applicable. It seems that, in practice, truth of a theory is not a prerequisite for using it in an explanation. Classical physics can still be used to give successful explanations, whereas in other areas it has been refuted. If you want to use it in a DN explanation, it has to be judged 'true' in that case. The judgements 'true' or 'false' are too crude to be used in a model of scientific explanation, not only at the level of laws that explain phenomena, as Cartwright showed, but also at the level of theories. The DN model is not subtle enough to handle these explanations.

We stop thinking how classical theory and quantum theory can (not?) both be true. Let us focus instead on a range of situations where both theories are supposed to be *applicable*, according to Bohr's *correspondence principle* between classical and quantum physics. In atomic physics, the principle tells that the lower the energies of the light, the more similar the predictions using classical and the quantum physics will be (see [13], p. 34). The correspondence principle is a principle of continuity of applicability between classical and quantum physics. If physical theories make

[1]The Hamiltonian that describes the intraction of light with matter contains 'light terms', 'matter terms' and 'light-matter-interaction terms'. Often, either the matter or the light are simplified considerably, giving a *light picture* or a *matter picture*, respectively. In the light picture, the interaction between light and matter is described as light that scatters in a material, whereas in the matter picture, the interaction is described as a perturbation by an external electromagnetic field of the atoms that make up the material. In principle, both pictures are equivalent, because both are pictures of the same processes. In practice, however, the equivalence is broken by the introduction of different simplifying assumptions in the two pictures (see [31]).

[2]These fundamental problems are captured in the famous Schrödinger's cat paradox: can a macro-'object' like a cat be in a quantum mechanical superpostion of states? And when and where does the wave-function collapse into one of its eigenfunctions? See for instance Rae [37] or Auyang [3].

a patchwork, then at least some of these patches are connected. The continuous transition between classical and quantum treatments makes it possible that not only models, but also theories are chosen that serve the explanatory purposes best. It stands to reason to try the simplest (classical) treatment first.

I do not want to overstate the importance of applicability. From the fact that sometimes the simplest applicable theory is used to explain, one should not conclude that physicists stop theorizing once their present theories are applicable. For example, theoretical physicists have not turned from quantum field theory to string theory because quantum field theory was not applicable in some experiments. The reason is that one of the four forces in nature, gravity, cannot be incorporated in quantum field theory in a consistent way. This fact does not seem relevant to any experiment to which quantum field theory is applied, because gravity is a so much weaker force than the others that it only becomes important on the experimentally utterly unattainable so-called Plack-scales. Here theoretical consistency is the driving force for doing research, not applicability to experiments. This ends the discussion of the first situation that was quoted at the beginning of this section.

The second situation in which the quantization of the electromagnetic field can not be ignored, is even more interesting: if you can find a model in quantum electrodynamics that can be solved exactly, then it should be studied, whether it is relevant for experiments or not. It will prove worthwhile to study the use of such models more closely.

Why is it interesting to study exact models that are not directly relevant for experiments? The answer that perhaps they will be relevant in the future, can not be the whole story. Let us consider the role of exact models in quantum mechanics first. In the first place these models are interesting, because they show how the energy levels and wave functions of the system depend on the parameters in your model. This gives an understanding of how the different parameters have different influences on the results. It can help you in finding trial wave functions in more realistic and more complex models, so that perhaps also the latter models can be solved exactly. In that case, the exact model has *in*direct experimental relevance.

A second reason to study exact models, is that more complicated models with complicated Hamiltonians can be 'solved' perturbatively, by treating them as simple exactly solvable Hamiltonians plus a perturbation. In the quantum mechanics textbook by Gasiorowicz, the study of perturbation theory is motivated as follows:

> "There are few potentials $V(r)$ for which the Schrödinger equation is exactly solvable, and we have already discussed most of them. We must therefore develop approximation techniques to obtain the eigenvalues and eigenfunctions for potentials that do not lead to exactly soluble equations" ([15], p. 255).

The exactly solvable unrealistic model plus perturbation theory give approximately correct analytical results for more realistic ones.

The two reasons to study exact but unrealistic models are of the same kind. Cartwright argues that theories acquire their meaning through their models. I should add to this, in turn, that the models acquire their meaning through their solvability. Not only for pedagogical reasons do we find exactly solvable models in quantum mechanics textbooks. In section 3.4.1, I mentioned that Cartwright was surprised to find out that only so few Hamiltonians were introduced in the textbooks, most of which did not seem to have much to do with realistic systems. She does not mention that the models introduced in the textbooks are precisely those few models that can be solved exactly. On the other hand, she argues that the explanatory power of quantum mechanics stems from the wide use of a small number of simple models. If she discusses the same few models here (which I think she does), then this shows that it is not simply *simplicity* of models that produces explanatory power, but more precisely, their *solvability.* The physicists' focus on solvable models is in accordance with Cartwright's claim that the most realistic model does not necessarily serve all purposes best. She just did not mention that solvability is an important basic purpose of (all?) models, that is presupposed if the model is to serve other purposes.

In the next section, more functions of exact models will be mentioned, especially in the realm of theoretical physics.

5.3 Special solutions and impoverished theories

In the previous section, we found that unrealistic exact models can be interesting owing to their possible indirect relevance to explanations in experiments. However, the question why exactly solvable models can also have "an importance of their own, virtually independent of their relevance to possible experimental work" has not been dealt with yet. Where do they occur? And what is their function?

This kind of models can be found in modern theoretical physics. During the twentieth century, physical theories have become more general and at the same time mathematically increasingly complicated. For that reason, realistic models within these theories that are mathematically tractable, are so rare, that physicists would not get any feeling or understanding for these theories if they stubbornly decided to try and solve realistic situations only. Therefore, physicists try to solve unrealistic situations as well. I shall discuss three strategies: 'special solutions', 'tinker toy models' and 'impoverished theories'.

5.3.1 Special solutions

In Newtonian theory, gravitation could be descibed using only one field equation (the Poisson equation), whereas in general relativity it requires highly nonlinear equations for a ten-component tensor field. Einstein did not tinker with Newtonian theory of gravitation in order to arrive at a theory that explained some anomalies of the latter.[1] Rather, he constructed a theory that satisfies some very general requirements, the *principle of equivalence* and the *principle of general covariance*, which I shall not discuss here (see [34] or [39]). The the result is a general but mathematically complicated theory with equations that describe how the mass-and-energy distribution in spacetime influences the structure (e.g. the curvature) of spacetime, and vice versa. This mutual influence makes the equations highly nonlinear and difficult to solve. It is interesting to study what possible solutions the theory can have and what kind of new properties show up in these solutions to the equations (curved spacetime, black holes, etcetera). In the first place, this gives understanding of the complicated theory, for how much can one understand a theory consisting of equations that have never been solved in any case? Secondly, the new properties might also reveal properties of the physical world.

The problem becomes easier, and perhaps solvable, if one looks for *special solutions*, for example solutions that represent spacetimes with certain symmetries. The first exact solution of the Einstein equations, which was obtained by Schwarzschild in 1916, is one of this type: Schwarzschild looked for the *static*, *spherically symmetric* gravitational field in the *empty spacetime* surrounding some *massive spherical object.* The Schwarzschild solution would be interesting for the sole reason that it is *a* solution to the Einstein equations. However, it is even much more interesting, since the special conditions imposed upon the Einstein equations make a good model for a star.

[1] Einstein once remarked that this method was "too cheap" ([3], p. 24-25).

The function of the model is twofold: it helps to go from the messy concrete physical system to a tidy material abstraction of the former, where small and specific (and therefore irrelevant) outside influences and other uninteresting individualities of the system are neglected; secondly, the model helps to go from complicated equations of a general theory to simple equations for a special case. As a material abstraction, the model is a simplification *by generalization* of concrete situations. At the same time, the model is a simplification *by specialization* of the general equations.

It is an art to find a model, because it must satisfy contrary requirements: it must be detailed enough to be able to explain phenomena in your system, while it must be mathematically simple enough to be solvable. Cartwright does not discuss the second kind of simplification, which is a pity, because it limits the choice of models in modern theoretical physics considerably. It also gives a reason why physicists are interested in symmetries: it may be easier to test a mathematically complicated theory by physical systems which have certain symmetries, because equations governing the models of these systems will be easier to solve.[1]

5.3.2 Toy models and impoverished theories

We do not always have the luck to find a model that satisfies both requirements. Especially when solvable models of real systems are not available, models are constructed that satisfy only the second requirement: it produces simpler equations, but it can not be thought of as a mathematical abstraction of a real system. These models are called *(tinker) toy models*.[2] I shall not discuss an example of a tinker toy model here; the Schwarzschild model would have been a tinker toy model, if it had not been the model of a star. Empirical testing of theories by definition does not involve toy models, but of course physicists will do their best to find a system where a solvable toy model can be applied, because it would make another test of the theory.

There is another way of getting simple equations than by using toy models: instead of looking at special solutions for the theory, there may be no other option than to approximate the equations in order to make them mathematically tractable, thereby making an *impoverished theory* (the term was coined by Redhead ([38], p. 147). In order to make the distinction clear, let me give a simple example of both: If we try to solve the Schrödinger equation for a funny Hamiltonian, that Hamiltonian is a toy model within quantum mechanics, whereas if we start to simplify the Schrödinger equation, we have impoverished quantum mechanics. In practice, the difference may not always be so clear, but let us now consider a clear case of an impoverished theory.

Theoretical physicists who study *dynamical mass generation* (see [27] or [18]) in quantum electrodynamics (or QED), try to find the masses of fermions (one of which is the electron). They try to find them from the equations of motion that govern the dynamics of the fermions and their interaction with photons. In principle, the mass can be determined from the solutions to these nonlinear equations. If a nonzero mass is found, then it can be said to be generated by the fermion dynamics, hence the name 'dynamical mass generation'. It is a pity that the equations are much too complicated to be solved.[3] Therefore, extra assumptions have to be made.

[1] Perhaps its focus on testing general theories by simple models explains why theoretical physics has grown apart from engineering, where detailed models are necessary to describe and construct complex systems. (This view is held by both the physicist Lagendijk and the engineer Tennekes in a recent debate [12]).

[2] Other names for toy models are 'theoretical models' (Redhead [38]) or 'study models' (Groenewold [16]).

[3] Tacitly, I moved from analytical to numerical solutions: the equations from which dynamical mass generation should be found, probably will never be solved analytically. They also defy numer-

As a first 'impoverishment', one does no longer consider quantum electrodynamics in four-dimensional spacetime (QED_4), but its younger brother QED_3 in three-dimensional spacetime (one time plus two space dimensions).[1] The equations are also simplified by a second assumption, that there are N types ('flavours') of fermions instead of the three flavours in the real world. This generalization is a simplification, because when N is large enough, the equations get simpler. But even then, the equations are still intractable, so that additional approximations have to be made. Depending on the type of extra approximation (or 'Ansatz'), different solutions can be obtained. Partly for that reason, there has been a long standing controversy whether in QED_3 dynamical mass generation occurs for all values of N, or only for N smaller than some critical value.

It is hard say which properties of solutions in QED_3 will also be properties of solutions of the more complicated and more interesting QED_4. That is the big problem for every impoverished theory, and learned intuition seems to play a role when deciding which properties of the impoverished theories are artefacts of the approximations, and which are properties of the 'full' theory as well. The main reason to study QED_3 is that it teaches you something about the process of dynamical mass generation in general. The increase in understanding someday may result in an increased tractability of the equations of QED_4, but nothing can be said with certainty here.

The existence of the 'computation gap' that forces the study of impoverished theories, also creates problems of justification: for many years, theorists have studied dynamical mass generation, while the experimental input was limited. New approaches, assumptions and calculation techniques are justified for *mathematical* reasons. On the one hand, one can say that it is dangerous to do theoretical physics far from experiments, where predictions and explanations form the context of justification of theories. That may be true, but on the other hand it shows that there is also a *mathematical context of justification* in theoretical physics that does not depend of scientific explanations and that influences the choice of models in a way that Cartwright does not discuss.

In this subsection, we have seen that the usage of the term 'model' in theoretical physics is different from the meaning that Cartwright gives to it. We have seen three ways of dealing with mathematically intractable theoretical equations: the nicest way is to find a model of a system where (through 'simplification by specialization') the equations become tractable. If such a system is not available, one can study a toy model, in order to learn to know the theory better. Or finally, the theoretical equations can be simplified by hand, resulting in an impoverished theory. When going from the first to the third option, the empirical justification for making the equations solvable decreases.

5.4 Truth approximation in realistic models with 'disorder'

In section 3.4.2, I discussed Cartwright's finding that physicists use the term 'realistic' for models in two senses: in the first sense, a realistic model is the opposite of an idealized model; in the second sense, it is opposed to a phenomenological model, in which terms occur which are not explained by the theory that describes the phenomena. We have seen in section 3.5.2 that Cartwright does not think that a model can ever give a *real* description of an individual system; in this section we

ical solution. I shall not discuss whether analytical and numerical calculations give a qualitatively different kind of explanation.

[1] QED_3 is not only an impoverishment of QED_4, since it can be used as a model in superconductivity.

concentrate on realistic models, without addressing the question whether the most realistic model will be real in the metaphysical sense.

In section 5.2.1 I agreed with Cartwright that the most realistic model does not always serve all purposes best, if only because it is much easier to write down a realistic model than to solve its equations. Although at a certain time not always the most realistic model will be chosen for explanation, one could ask whether in the course of time models will be used that are more realistic (in either sense), and if so, for what reasons. Cartwright does not address these questions herself; she does not discuss 'progress in science', or (on a smaller scale), changes in explanatory purposes of models. One way of making sense of the coming and going of theories and models in the course of time, is to show that successive theories and models are (potentially) closer to the truth (see Hettema & Kuipers [23]). The purpose of this section is to study whether three successive physical models are increasingly realistic in Cartwright's sense, and whether they form a case of (potential) truth approximation as defined by Hettema & Kuipers. Potential truth approximation can be seen as a theory of explanation, that does not show how individual explanations 'are possible', but rather shows how later explanations can be increasingly successful.

5.4.1 Models without, with static and with dynamical disorder

Of the three models from optics that I shall present now, I studied the third one extensively for my master thesis in physics [42], so I hope not to get too 'physical'.

Isolated molecules that can be described as two-level molecules, will only absorb light of one frequency, ω_0, say. If light of that frequency is absorbed, the molecule goes from a lower to a higher energy state, as was discussed in section 5.2.1. The molecule is then said to be 'excited'. However, if the molecule is part of a collection of molecules that interact with each other, then the optical properties of such a 'molecular aggregate' are different from the properties of isolated molecules. For instance, an excitation can hop from one molecule to another, or the excitation will be 'shared' by several molecules ('delocalized') [26]. This is the case in some bacterial light antenna systems, where light is absorbed by chlorophyl molecules, which together form a circle. Each molecule in such a ring interacts strongly with its two neighbouring molecules and much less strongly with the others, so that usually only 'nearest-neighbour interaction' J is taken into account in models. In the simple model that I have described to you now, light absorption A is a function of three parameters:[1] the optical transition frequency ω_0, the nearest-neighbour interaction J and the number of molecules N:

$$A_1 = A_1(\omega_0, J, N). \tag{5.3}$$

According to this model, the aggregate absorbs light of only one frequency $\omega_0 + 2J$, which can be seen by solving the Schrödinger equation for the model.

However, experimentally one finds that light is absorbed not just at one frequency; the light absorption shows a peak around the frequency $\omega_0 + 2J$ with finite width. This can be understood from the fact that all molecules in the aggregate may be identical, but their optical transition frequencies may be influenced by differences in the environment. One can conclude that the physical system considered was too small, so that part of what was called the environment until now, should be incorporated in a more complex model of a larger system. However, in complex systems like bacteria, a 'system-environment cut' has to be made somewhere in

[1] For the physicists: the magnitude and the circularly symmetric orientation of the molecular dipole moments also determine the light absorption characteristics, but we do not consider changes in these parameters now.

order to be able to calculate anything. For that reason, the model of the aggregate can be extended by a variable Δ, which is a measure (e.g. the standard deviation) of the disorder in the optical transition frequencies of the molecules. In this second model, the light absorption by the aggregate is a function of *four* variables:

$$A_2 = A_2(\omega_0, J, N, \Delta). \tag{5.4}$$

It is assumed that the disorder in the transition frequency is static, and caused by static differences in the environments of the different molecules.

Disorder is not always static. Especially when molecular aggregates operate at room temperature in a fluid (like bacterial antenna systems do), the environment is continually changing and moving on a time scale similar to the 'hopping time' of excitations; experiments indicate that *dynamical disorder* must be considered instead [1]. This can be modelled by introducing a correlation time τ_c and assuming that correlations of the fluctuating disorder decay exponentially, at a rate given by $\lambda = 1/\tau_c$. [1] In this third model, the absorption is a function of *five* parameters:

$$A_3 = A_3(\omega_0, J, N, \Delta, \lambda). \tag{5.5}$$

Are these three models increasingly realistic? The sole difference is the modelling of the disorder. The models are increasingly realistic in the first sense: supposing that we use all three of them as models of the bacterial antennae, it is a gross idealization to consider the ring of chlorophyl molecules as isolated from the rest of the bacteria, as is done in model A_1; the parameter Δ in $A_{2,3}$ accounts for this disturbing interaction, while λ in A_3 models the fact that the disturbing interactions may change in time, which makes the description even less idealized. (However, one is still idealizing a lot if the interaction of a bacteria with one of its antennae is described by two parameters only.)

Are the models increasingly realistic in the second sense? From the point of view of the explanatory theory (quantum mechanics), the value of Δ or λ are not explained. They are stochastic variables, not theoretical entities. The models contain an increasing number of phenomenological parameters (0, 1 and 2, respectively). I would say that in the second sense of the word, the models become less realistic and more phenomenological.[2] As a conclusion, the triple of models $\{A_1, A_2, A_3\}$ is both increasingly and deceasingly realistic, in the sense of being less idealized and more phenomenological, respectively.

5.4.2 Realistic models and truth approximation

Irrespective of which model is the most realistic, it is a fact that A_1 is a special case of A_2, which in turn is a special case of A_3:

$$A_1(\omega_0, J, N) = A_2(\omega_0, J, N, \Delta = 0) = A_3(\omega_0, J, N, \Delta = 0, \lambda = 0), \tag{5.6}$$

or in words: if in A_2 the amplitude of the disorder caused by the environment is $\Delta = 0$, then we have retrieved model A_1, *without disorder*. Similarly, if in A_3 the decay rate of the correlations of the fluctations is put to zero (or in other words: the correlation time is infinite), then we have retrieved model A_2, *with static disorder*. If there is no independent evidence for the values of the disorder parameters Δ and λ, then they are free parameters. Then the model A_2 can always produce a better

[1] If one wants to know whether fluctuations caused by the environment *really* decay exponentially, one again has to take the complex environment explicitly into account, which we do not want.

[2] Cartwright only mentions another situation, in which a model is called more realistic than another, since it has the same number of parameters, of which a smaller number are phenomenological.

fit to the experimental data than A_1, because it has an extra free parameter; for the same reason, A_3 produces a better fit than A_2 to the same experimental data.[1]

Is the triple $\{A_1, A_2, A_3\}$ an example of potential truth approximation? By the property (5.6), one sees that the three models constitute a 'concretization triple', as described in [29], (p. 247-251). In that respect the situation is completely analogous to the paradigmatic case of truth approximation: the transition from the theory of ideal gases to Van der Waals theory of gases.

As a first shot, let us suppose that the three optical models constitute an example of potential truth approximation. If the progress of science can be explained from the fact that successive theories come closer to the truth, then why do not physicists choose the route closer to the truth? For example, the model A_3 can be concretized by introducing static or dynamical disorder in the interaction J between neighbouring molecules as well. Another possible concretization is taking a combination of both static and dynamical disorder of the optical transition frequency ω_0 into account. In a real antenna system, probably all these types of disorder occur, although some are more important than others. Many more types of possible disorder can be thought of, so that many concretizations can be constructed, thereby potentially approximating the truth.

This is not what happens: physicists do not use models with one hundred disorder parameters. The model A_1 is extended via A_2 to A_3 for experimental reasons: if theorists did not somehow take the environment into account, they would predict narrow absorption lines where if fact broad peaks show up. This forms enough evidence that the model A_1 is too simple to explain the phenomena, and therefore the somewhat more complicated model A_2 was introduced. The suggestion from experiments that dynamical fluctuations affect the antenna systems, called for the model A_3, so that we get a better fit and gain a better understanding of the optical properties of these systems.

The model with the smallest number of parameters that explains the phenomena reasonably well, is the one that is chosen; the 'cost' of another parameter must 'pay off'. It is true that a model with one parameter more will always produce a fit to the experimental data that is at least as accurate (if all parameters in the models are free). But the 'marginal cost' of the extra parameter must be smaller than its 'marginal profit' of more accurate prediction. The 'cost' consists at least of increased mathematical complexity and decreased 'trust' in the system-environment cut. Especially if you model the influence of the *environment* on a system, you shouldn't introduce too many free parameters, because the number of parameters indicateds (loosely stated) the importance of the environment: if you need to describe it with so many parameters, then clearly the environment is so important that you have chosen your system too small. For that reason, the number of parameters by which the environment will be modelled will remain small in actual explanations. That's why scientists cherish their simple modelling methods of outside influences.

Cartwright's answer that simpler models are chosen for generality does not work here, because the model with one hundred free parameters is even more general (more widely applicable) than the simple one. She also mentions that scientists use simple models in order to be able to communicate with each other. That is another reason to use Occam's razor. Popper argued that one should choose the simplest possible hypothesis in order to explain the phenomena, because that hypothesis can be falsified better ([35], p. 136). That seems another adequate explanation, which is complementary to the ones I just presented.

[1] On the other hand, if a model is extended by introducing an extra phenomenological parameter, with a numerical value that has been obtained from independent evidence (other experiments), then the new fit need not be more accurate than the fit of the old model with the experimental data.

We have seen a few reasons why the concretization process is not carried on as far as one could in the realm of 'disorder models'. If, as we have assumed, the triple $\{A_1, A_2, A_3\}$ is an example of truth approximation, then we have just seen some reasons why the truth is not approached indefinitely, even if concretizations can be thought of that are a means to that end. If experiments roughly can be explained by simple disorder models, then concretized models with more parameters can be constructed, but they will not be accepted.

However, there is a difference between the triple of models that I considered and the paradigmatic example of the successive gas theories: in both triples, successive theories are more realistic in the first sense: they become less idealized. However, the disorder models do not become more realistic in the second sense, because the concretizations are phenomenological.[1] Nobody assumes that the model A_3 is true, even if it captures a lot of the phenomena. In reality, the disorder is caused by a wealth of different processes between the system and its environment; the disorder terms are an efficient way of dealing with them. For these reasons, I am not sure whether a phenomenological concretization of a model should be called closer to the truth than the original model, even if it is evident that by fitting the model, more accurate predictions can be found.

5.5 Conclusions

In this chapter several interesting physical models have made their catwalk in a show that had several purposes. We have seen that the term 'model' is used in many more ways than Cartwright does; for her, every description of physical systems is a model, in the sense that it is a simplified and distorted description. The models in this chapter were models in a narrower sense: they are explicitly called models, like the 'liquid drop model'.

In the previous chapters, two models of scientific explanation were introduced and criticized. In this chapter they could be tested, for if they were good models of scientific explanation, then they should be able to account for explanations based on explicit physical models. Here, I shall summarize what each model showed about explanation; in the next chapter, more general conclusions will be drawn.

Two heuristic nuclear models were discussed: the 'liquid drop model' and the 'shell model'. Both models were analogues of models from other parts of physics, which showed that more input is needed for model choice than one theory plus the system to be modelled. The liquid drop model has an interesting relation to quantum mechanics: originally, it is a non-quantum model, but since the analogy between a liquid drop and an atomic nucleus did not take us very far, the model was corrected quantum mechanically, resulting in a surprisingly accurate model of nuclear binding energies. The shell model is also heuristic, but closer to quantum mechanics, in the sense that the nucleus is descibed quantum mechanically as one nucleon trapped in a potential caused by all other nucleons. The models have different types of successes. A more fundamental theory of nuclear matter has the task of explaining the surprising successes of these very different, simple models. Contrary to Cartwright, I do not think that a theory of explanation should explain the multiple use of models; in nuclear physics, a partial explanation can already be given by quantum theory. The search for a more fundamental theory using information from simple models, involves intuitive (not deductive) steps.

In the realm of optics, we have seen that very realistic models of even simple systems defy solution, so that simple models are chosen. Cartwright argues that simple models are chosen for their generality, whereas I argued that solvability may

[1] The Van der Waals model is *not* just a phenomenological concretization of the ideal gas law, I would say.

be more important. In the discussion of the choice of the two-level atom model, we saw that the theory, together with information about the system that is studied, can be used to explain the choice of a model. Cartwright calls this stage of model choice 'informal', whereas I think that model choice can be defended by more than ultimate predictive success.

In the explanation of many optical experiments, the light is treated classically, whereas the atoms and molecules on which it shines, are provided with a quantum model. Cartwright is right in that different models are chosen for different purposes, but even theories can sometimes be selected for a purpose. Especially where theories overlap more or less continuously, one is free to choose the simplest one. In these cases it is obscure what is meant when in a DN argument a theory is called true: does 'true' mean 'applicable', is truth context -dependent, or ...?

The quantum mechanics textbooks present the exactly solvable models, but not only for pedagogical reasons: a model must be solvable (exactly or perturbatively) in order to be used in an explanation. Often modern physical theories involve complicated mathematical equations, like in general relativity or quantum electrodynamics. These are very general theories, so that it is not a problem to find situations in which the theory *in principle* should apply. The bottleneck for theoretical physicists is to find models in which the theoretical equations become simpler so that they can be solved. Therefore, it is the solvable models that are used over and over again in new situations. Often, equations get simpler if one looks for a special solution with certain symmetries. Therefore, in practice, theories are tested by applying them to systems that have these symmetries. If such systems cannot be found, one can resort to toy models or impoverishment of the theory.

The three optical models with and without disorder were found to be a concretization triple, as defined in the truth approximation literature. In successive models, new phenomenological parameters are introduced, which produce extra accuracy of predictions. In practice, extra accuracy is balanced against disadvantages of a larger number of phenomenological disorder parameters. It remains to be seen whether a phenomenological concretization triple should count as a (potential) truth approximation. At least structurally, phenomenological concretizations cannot be distinguished from concretizations that are not phenomenological.

Chapter 6

Conclusions

After the battle between two models of scientific explanation has been fought in this thesis, it is time to see whether one of them has proved superior. Both of them have suffered serious blows and neither of them, I think, is an adequate model of scientific explanation.

To consider the DN model first, Cartwright argued that laws of nature are not strictly true, and still they are used as true premises. I share this objection with Cartwright: laws are to be given a place in any model of scientific explanation, but it remains unclear what it means for a law to be true, while giving only approximately correct explanations. Below, I shall return to Feynman's remark that a law can be 'right', though only giving a partial explanation. The fact that in optics sometimes both classical and quantum theory of light can be used in explanations (section 5.2.2), which as a whole cannot both be 'true' theories, shows that the 'truth' of premises in the DN model must be taken *cum grano salis*. The DN model is incomplete in that it presupposes that all explanations involve theories and laws, whereas in many areas of physics heuristic models are being used. These models are not supposed to be true but at most efficient accurate descriptions, where sometimes a theory might be used to improve the predictions. Explanations based on such heuristic model do not seem to involve laws of nature.

Cartwright's criticism that derivations often are not deductions, as the DN model requires, is less convincing. I refuted her claim that a particular prediction in quantum mechanics improves upon approximation. I showed that some approximations which Cartwright calls 'not dictated by the facts', are not as arbitrary as she would have it be: experiments do not stand on their own, so that independent testing may be possible of the assumptions needed in order for the approximations to be valid. I also argued that some approximations that Cartwright would call 'not dictated by the facts' are self-consistent derivations, which makes them much more plausible. Her criticism that derivations can not be described by a deductive model, forms a serious problem for her own simulacrum account.

Both proviso's and theoretical ascent are a problem, for both models of explanation. I showed how the deductive nature of an argument can be threatened by the two problems, by adding an 'applicability premise' to the explanans. This was not a solution of the problems, since it merely made the problems explicit by dragging them into the deduction. I showed that if there is no independent evidence available for the applicability premise, then the argument becomes what Hempel calls a self-evidencing argument.

Cartwright does not discuss theoretical ascent or the problem of proviso's for her account of explanation. In the simulacrum account, all models are self-evidencing: if a model does not serve its ultimate explanatory purpose, then just make another model. This struggle for success of models in itself does not yet explain why there can be successful models at all. Cartwright needs these throw-away models for her thesis that empirical testing of abstract theories is problematic. Models, she says, are constructed using theory and experiment in such an informal way that their sole justification lies in their ultimate explanatory success. I used the two-level atom model in resonance optics to show that in the informal stage of model choice, the theory can be used (and is used in practice) to defend and to explain the choice of a model, so that not only the model but also the theory can be put to the test in experiments.

For Cartwright, theories are true in models of systems and not about the systems themselves. By steadily improving the model, via abstraction and concretization, one can get more accurate descriptions, but models will never arrive at literal truth, because models are in a sense "common to many systems but not literally true of any". Literal truth is about singular systems, including their outside influences. If truth is defined that way, I think everyone must agree with Cartwright that literal truth can never be obtained by improving models. For her, a model can only be real if it gives a literally true description. However, – if the reader allows me this one new thought in the conclusions – for someone who believes that there is a metaphysical sense in which the models are 'common' to different systems, a 'real' model does not necessarily give a 'literally true' description about a singular system in the sense described by Cartwright. What real hydrogen atoms have in common, is what they have in common with a hypothetically isolated hydrogen atom. The hypothetically isolated atom by definition does not exist, but that does not imply that it can't be used as a *description* of what real atoms have in common. In that sense, the model of the hypothetically isolated hydrogen can be real. Of course I am wrong here if any description *by definition* can't be real.

Cartwright argues against theory realism, because theories are true about distortions of reality called models. Therefore, the content of scientific knowledge is not in abstract theories, but in phenomenological laws. However, if it is true that Kepler's laws are phenomenological, then there are reasons to assume that the content of scientific knowledge lies in theories, since Newtonian theory can be used to compute perturbatively how other planets disturb Keplerian planetary motion. I think that theory plus model even have more content, because the theory is more general and from the model one can see what effects are neglected in the derivation.

Cartwright wants to find a theory of explanation that tells why some models that can be used in one case can also be used in another. I argued that in some cases a *physical theory* partly can do the job. Especially for Cartwright, it is hard to find a theory of explanation. She would be right that theories are not put to the test in experiments if these theories are not used to justify the choice of a model. But if models are only justified by their success, then it is hard to explain the success of one model in two situations. I think it is more interesting to find a model and a theory of explanation which show how the several types of physical models that I have discussed, including heuristic models, can all be fitted in.

What makes an explanation explanatory? Theories of explanation should answer this question, and realism about laws could provide such a theory. However, if Cartwright is right that laws that are not literally true, cannot be real, then realism about laws cannot serve as the metaphysical basis of the DN model. Cartwright has not presented a metaphysical basis for her own simulacrum account. Her metaphysical pluralism would not work as a theory of explanation. Cartwright is an entity realist, but this is a 'disconnected' belief: I could not find situations where she uses it to support other ideas, or where she presents arguments for their reality. In particular, entity realism is not used to support a theory of explanation.

Quantitative arguments abound in physics, which in a way are more subtle than true or false statements. A prediction may have the form "The current is 5 mA", but it is not simply true or false. All predictions have a finite accuracy, and so have measurements. If neither the predictive nor the measurement error can be estimated, there is no way to refute a prediction. That's one reason why accuracy is more relevant for explanation than truth. Sometimes, one can use both use classical or quantum theory in explanation: they will give almost the same predictions. Again, 'true predictions' are not required for explanation, accuracy is sufficient. The more accuracy the better, of course. Since the DN model uses true or false premises, it cannot handle quantitative arguments well.

Not only should a theory of explanation be able to take *quantitative* arguments into account. It should also explain what a *partial* explanation is, as we saw in chapter 4: Feynman argued that the failure of some laws lead to the discovery of new ones, like the apparent failure of the law of gravitation lead to the discovery of the speed of light: the law of gravitation was the larger part of the explanation of the phenomena; the finite speed of light formed a large part of the *residual explanation.* If new laws are found by assuming that known laws are right but inaccurate, then we should not conclude too soon that laws that are not literally true, are false.

For Cartwright, metaphysics does not involve quantitative arguments: laws that give accurate predictions give false predictions, so they are not real. I have called this the 'inverted no miracle argument' against realism. One can object that predictions about physical systems can not be tested experimentally, because these systems are not isolated in reality. This response is similar to my story about the isolated hydrogen atom in these conclusions. Cartwright's patchwork-of-laws picture of science depends on the inverted no miracle argument: every small difference in experimental outcome implies that two situations are to be described by different laws.

Cartwright thinks that the use by physicists of a limited number of simple models can be explained from the fact that simple models are more general, so that they have more explanatory power and physicists understand each other better. I argued that solvability of models plays an important role. Only solvable models are used to explain. Where theories are used that involve mathematically complicated equations, it is an art to find a model (preferably of a real system) for which the equations are simpler. The rarity of solvable models explains their multiple use in explanations. I also discussed toy models. They are not models of any concrete system, but they give understanding of a theory that is supposed to describe reality. Impoverished theories are supposed to do the same, but the logic of the situation is quite complicated.

In this thesis, I did not present a new model of scientific explanation. Very generally, I have argued for a more unified picture of science, one that has more 'connections': in the DN model, all theories and laws are either true or false, whereas I argued for a more quantitative picture: logically inconsistent theories give numerically almost the same predictions in limiting cases. The quantitative nature of explanations connects many patches of theories and laws. Also, in the simulacrum account, models are not justified by the theories in which they function, nor do the models relate to other models. I have argued that these relations do exist. As a consequence, abstract theories *can* be related to experiments.

Bibliography

[1]

[2] L. Allen and J.H. Eberly. *Optical Resonance and Two-level Atoms.* Dover, New York, 1975.

[3] S.Y. Auyang. *How Is Quantum Field Theory Possible?* Oxford University Press, 1995.

[4] N. Cartwright. *How the laws of physics lie.* Clarendon Press, Oxford, 1983.

[5] N. Cartwright. *Nature's Capacities and their Measurements.* Clarendon Press, Oxford, 1989.

[6] N. Cartwright. Fundamentalism vs the patchwork of laws. In D. Papineau, editor, *The Philosophy of Science*, Oxford readings in philosphy, pages 314–326. Oxford University Press, 1996.

[7] I.B. Cohen. *The Birth of a New Physics.* Penguin Books, London, 1992.

[8] M. Cyrot and D. Pavuna. *Introduction to Superconductivity and High-Tc Materials.* World Scientific, Singapore, 1992.

[9] O. Darrigol. *From c-numbers to q-numbers: the classical analogy in the history of quantum theory.* Berkeley, 1992.

[10] E.J. Dijksterhuis. *De Mechanisering van het Wereldbeeld.* Meulenhoff, Amsterdam, 1950.

[11] P.A.M. Dirac. *The Principes of Quantum Mechanics.* Oxford University Press, fourth edition, 1958.

[12] R. Dobbelaer. Debat: Ad Lagendijk versus Henk Tennekes, wetenschap versus techniek. *Natuur & Techniek*, 65(6), 1997.

[13] B. Falkenburg. Modelle, Korrespondenz und Vereinheitlichung in der Physik. *Dialektik*, pages 27–41, 1997.

[14] R.P. Feynman. *The Character of Physical Law.* M.I.T. Press, Cambridge, Massachusetts and London, England, 1967.

[15] S. Gasiorowicz. *Quantum Physics.* Wiley, New York, 1974.

[16] H.J. Groenewold. The model in physics. In H. Freudenthal, editor, *The Concept and the Rôle of the Model in Mathematics and Natural and Social Sciences*, pages 98–103. 1961.

[17] A. Grünbaum and W.C. Salmon, editors. *The Limitations of Deductivism.* Pittsburgh Series in Philosophy and History of Science. University of California Press, Berkeley and Los Angeles and London, 1988.

[18] V.P. Gusynin, A.H. Hams, and M. Reenders. *Phys.Rev. D*, 53, 1996.

[19] I. Hacking. *Representing and Intervening*. Cambridge University Press, 1983.

[20] E. Hardtwig. *Fehler- und Ausgleigsrechnung*. Bibliographisches Institut Hochschultaschenbücher-Verlag, Mannheim, 1968.

[21] C.G. Hempel. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. The Free Press, 1965.

[22] C.G. Hempel. *Philosophy of Natural Science*. Foundations of Philosophy Series. Prentice-Hall, Englewood Cliffs, N.J., 1966.

[23] H. Hettema and T.A.F. Kuipers. Sommerfeld's atombau: a case study in potential truth approximation. In T.A.F. Kuipers and A.R. Mackor, editors, *Cognitive Patterns in Science and Common Sense; Groningen Studies in philosophy of science, logic, and epistemology*, volume 45 of *Poznan studies in the philosophy of the sciences and the humanities*, pages 273–297. Rodopi, Amsterdam and Atlanta, 1995.

[24] K. Huang. *Statistical Mechanics*. Wiley, New York, 1987.

[25] E.T. Jaynes and F.W. Cummings. Comparison of quantum and semi-classical radiation theories with application to the beam maser. *Proc. I.E.E.E.*, 51, 1963.

[26] V.M Kenkre and P. Reineker. *Exciton Dynamics in Molecular Crystals and Aggregates*. Springer-Verlag, Berlin, 1988.

[27] M. Koopmans. *Dynamical Mass Generation in QED3*. PhD thesis, University of Groningen, 1990.

[28] T.S. Kuhn. *The Structure of Scientific Revolutions*. The University of Chicago Press, Chicago and London, 1970.

[29] T.A.F. Kuipers. Structures in science; heuristic patterns based on cognitive structures, 1992/1993. Syllabus.

[30] A. Kukla. Antirealist explanations of the success of science. *Phil.Sci.*, 63 (Proceedings):S298–S305, 1996.

[31] A. Lagendijk and B.A. van Tiggelen. Resonant multiple scattering of light. *Phys. Rep.*, 270(3), 1996.

[32] I. Lakatos. *The methodology of scientific research programmes*, volume I of *Philosophical Papers*. Cambridge University Press, 1978.

[33] L. Laudan. *Progress and its problems*. Routledge & Kegan Paul, London and Henley, 1977.

[34] A. Pais. *'Subtle is the Lord...': The Science and the Life of Albert Einstein*. Oxford University Press, 1982.

[35] K.R. Popper. *The Logic of Scientific Discovery*. Hutchinson, London, 1972.

[36] S. Psillos. Scientific realism and the 'pessimistic induction'. *Phil.Sci.*, 63 (Proceedings):S306–S314, 1996.

[37] A.I.M. Rae. *Quantum Physics: Illusion or Reality?* Cambridge University Press, 1986.

[38] M.L.G. Redhead. Models in physics. *Brit.J.Phil.Sci.*, 31:145–163, 1980.

[39] L. Sklar. *Philosophy of Physics.* Dimensions of Philosophy Series. Oxford University Press, 1992.

[40] W.S.C. Williams. *Nuclear and Particle Physics.* Oxford University Press, 1992.

[41] J. Worrall. Structural realism: the best of both worlds? In D. Papineau, editor, *The Philosophy of Science*, Oxford readings in philosophy, pages 139–165. Oxford University Press, 1996.

[42] C.M. Wubs. Optical line shapes of dynamically disordered molecular aggregates, 1997. Master thesis.

[43] E. Zilsel. The sociological roots of science. *The American Journal of Sociology*, 47(4), 1941/1942.