Analogies in Physics

Analysis of an Unplanned Epistemic Strategy

Von der Philosophischen Fakultät

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktors der Philosophie

Dr. phil.

Genehmigte Dissertation von

Ing. grad. MA Gunnar Kreisel

Erscheinungs- bzw. Druckjahr 2021

Referent: Prof. Dr. Mathias Frisch

Korreferent: Prof. Dr. Torsten Wilholt

Tag der Promotion: 26.10.2020

To my early died sister Uta

Acknowledgements

I could quote only very few by name who have contributed to my work on this thesis, for discussing some of the developed ideas with me or comments on parts of my manuscript. These are in the first place my advisor Mathias Frisch and further Torsten Wilholt, who read critically individual chapters. Much more have contributed by some remarks or ideas mentioned in passing which I cannot assign to someone explicitly and therefore must be left unnamed. Also, other people not named here have supported my work in the one or other way. I think they know who were meant if they read this.

A lot of thanks are due to Zoe Vercelli from the International Writing centre at Leibniz University Hannover improving my English at nearly the whole manuscript (some parts are leaved to me because of organisational changes at the writing centre). So, where the English is less correct Zoe could not have had a look on it.

Of course, all errors and imprecisions remain in solely my responsibility.

Abstract

This thesis investigates what tools are appropriate for answering the question how it is possible to develop such a complex theory in physics as the standard model of particle physics with only an access via electromagnetic interaction of otherwise unobservable objects and their interactions it was investigated what the tools are to do this. The answer is found in the usage of essentially two types of analogies. These two types are explanatory and predictive analogies. Explanatory analogies are needed to link observed new phenomena to accepted scientific theoretical models of well-known phenomena, predictive analogies to find in possible experiments further new phenomena not already observed.

The possibilities of observation and the related conditions are the main reason to use analogies as well in the development of theories, as in the representation of experiments especially in high energy physics, because there we find only charged particles observable via electromagnetic interaction. Other neutral particles with no charge are concludable only by visualization of their assumed tracks filling the gaps between the tracks of charged and therefore in some way observable particles. Today this is only possible with the help of computers evaluating the huge amount of data delivered by extremely complex detectors using the electromagnetic interaction in different ways. This visualization is also a kind of analogy because in fact all the particles are not observable directly, only their electromagnetic interaction if it occurs.

My investigation considered reasons for the following questions: Why in philosophy of science analogies were neglected during about one hundred years? How can analogies be characterized? And at last, as point of origination for my investigation where have analogies played an important role the development of high energy physics from the explanation of the constitution of the atomic core to the postulation of quarks, mainly in original publications? Important in this task is especially the consideration of the context which determines in some way the possible directions of development, however, no direction of development in general.

The intention in my thesis was to shed some more light on the epistemological side of scientific work in physics and how analogies could help enabling this task. The method used was primarily text analysis of mostly original publications of leading physicists.

Keywords: analogy, observation, epistemology, high energy physics

Contents

Abbreviations

CERN	Conseil européen pour la recherche nucléaire; nomination of the European Centre of Nuclear Research
DDI	denotation, demonstration, and interpretation
DN	deductive-nomological
GSW	Glashow, Salam and Weinberg
HD	hypothetical-deductive
HEP	high energy physics
QCD	quantum chromodynamics
QED	quantum electrodynamics
SLAC	Stanford Linear Accelerator

"(The scientist) appears as realist insofar as he seeks to describe a world independent of the acts of perception; as idealist insofar as he looks upon the concepts and theories as the free inventions of the human spirit (not logically derivable from what is empirically given); as positivist insofar as he considers his concepts and theories only to the extent to which they furnish a logical representation of relations among sensory experiences. He may even appear as Platonist or Pythagorean insofar as he considers the viewpoint or logical simplicity as an indispensable and effective tool of his research." (Albert Einstein with the following addition by Wolfgang Pauli jr.)

"*It is easy for me to emphasize with these sentences whereas the thinking in '-isms' is strange, even impossible to me.*

May be the great synthetical power as human being and thinker of Einstein an example also in the physics of the future when they have to weigh up the empirical given and the mathematical-logical structure of a theory." (Pauli jr. 1961, p. 80)

Analogies in Physics

Analysis of an Unplanned Epistemic Strategy

Gunnar Kreisel, Gottfried Wilhelm Leibniz University Hanover, Institute of Philosophy

1 Introduction

In the introductory chapter of a modern German textbook of high energy physics, one can find the following sentences:

1) "The ultimate aim of physical research is the proposal of a theory of matter"[1] (Berger 2006, p. 3).
2) "In relation to our knowledge today, the whole empirical material of sub-nuclear physics is described with astonishing precision by this theory [i.e., the standard model of particle physics]"[2] (Berger 2006, p. 6).
3) "We hope that the interaction of all branches of physics will lead in the future to a theory of matter, which will explain nature not only in the way we find nature today but explains why it is just so and not in another way"[3] (Berger 2006, p. 6).

These sentences characterize in distinct respects the dilemma of modern physics. On the one hand, the aim of physics (first citation) is fulfilled (second citation), however, on the other hand only in a

[1] "Das endgültige Ziel physikalischer Forschung ist die Aufstellung einer Theorie der Materie". (translation by the author)
[2] "Nach unserem heutigen Wissen wird das gesamte Erfahrungsmaterial der subnuklearen Physik durch diese Theorie [das Standardmodell der Teilchenphysik] mit erstaunlicher Präzision beschrieben". (translation by the author)
[3] „Wir hoffen, daß aus dem Zusammenwirken aller Teilgebiete der Physik eines Tages eine Theorie der Materie entsteht, die nicht nur die Natur so erklärt, wie wir sie jetzt vorfinden, sondern auch zeigt, warum alles gerade so und nicht anders ist". (translation by the author)

phenomenal manner (third citation), i.e., the theory describes the empirical phenomena only in an adequate way and does not suggest why they occur, as asked for in the third citation. Therefore, the actual fulfilment of the standard model is not satisfying for physics. But how to find a theory which fulfil the desired explanation? To fulfil this demand is not the task of a philosopher, however, the question leads to another, which, as we will see, lays more in the domain of a philosopher, and perhaps could help the physicists to answer the question how this fulfilment can be achieved. This question addresses the second citation of Berger and reads (in a Kantian way): How is it possible to build such a sophisticated and precise theory as the standard model of particle physics despite the only possible access via electromagnetic interaction? To understand this second question, I will first say something to the physical background and then why it concerns a philosopher of science.

The standard model of particle physics rests on the properties of particles in matter. These can thus be classified into two groups: one with half integer spin, and the other with integer spin. Spin can be imagined as a kind of rotation around its own axis. The half spin particles are called fermions, and these are either elementary particles themselves, or constitute elementary particles. They are considered as 'point-like' (Berger 2006, p. 5). The fermions also consist of two groups, i.e., the quarks and the leptons, which each contains three families of particle pairs with different masses. Each quark pair has a charge of 2/3 respective -1/3 of a charge unit and exists in three 'colours', as this additional property is called. Like the three colours in TV which add to white, the three 'colours' of quarks always combine to a 'white' particle, which are only found in experiments, e.g., a positive charged proton built up by two 2/3 charged quarks and one -1/3 charged quark of the lightest family (these quarks are called 'up' and 'down') or a neutral neutron containing one up-quark and two down-quarks. The proton is one of the six particles with potentially infinite lifetime, i.e., it is stable, whereas the neutron decays in about twenty minutes. The lepton pairs consist of one particle with charge -1 and one particle with charge 0. The lightest particle with charge -1, the electron, and all three particles with charge 0, the neutrinos, are also stable or assumed to be stable[4], respectively.

The particles with integer spin called 'bosons' are the mediators of the different interactions between the fermions. The oldest known and only stable one is the massless photon with spin 1 mediating the electromagnetic interaction described in quantum electrodynamics (QED). The photon in some sense is also the only directly perceivable if light hits our eye. Then there are three massive bosons with spin 1, one with charge 0, and two with opposite charge ±1, mediating the so-called weak interaction. The electromagnetic and the weak interaction are formally united in the theory of electro-weak interaction called GSW-theory after their developing physicists[5]. Further, there are the massless gluons with spin 1 combining two colours and mediating the strong interaction, which causes the inseparable composition of quarks to numerous particles of integer charge. The last and latest found particle in the standard model is the massive Higgs-boson with charge and spin 0, which gives via 'spontaneous symmetry breaking' the weak bosons and the fermions the mass.

[4] 'Assumed to be stable' means, that the neutrinos could change into each other by so-called 'oscillations' but do not change their number.

[5] Glashow, Salam und Weinberg

Nevertheless, there remain a lot of problems in this theory: about twenty arbitrary parameters are needed for the description, some of them contrary to models of cosmology or with the need of tremendous exactness so that the third citation of Berger above is understandable, when he states that the theory is still not satisfying. This reason supports the missing of the outstanding explanation why the particles exist in the way as found and not in another way.

However, from this preliminary description of the standard model of particle physics and its complexity it is still remarkable that it is constructed out of the knowledge produced by the electromagnetic interaction only, i.e., the interaction of photons with charged particles, namely the charged leptons and the integer charged compositions of quarks i.e., mesons and baryons like the proton. All other uncharged particles beside the photon can only be concluded from gaps in the trajectories of charged particles, therefore the question is justified, how it is possible to construct such a theory.

My answer to this question is clear, considering the title of this dissertation – namely by using analogies. However, it is not clear what in detail analogies are, how they work, and why their investigation was neglected by most philosophers so far - with some exceptions in the sixties (e.g., Ernest Nagel, Peter Achinstein, Mary Hesse, see chapter 2 and 4). In fact, the neglect in philosophy of science does not concern analogies alone but the whole context of discovery. Only recently this topic comes more into focus (see Danks & Ippoliti 2018) also in respect to analogies (see chapter 2).

Analogies I insist are part of the building of theories, especially physical ones, which are in the focus here. *More precisely, my thesis is: analogies are indispensable in the development and justification of theories in physics for epistemological reasons*. Of course, analogies cover not all of theory building, but I hope to demonstrate their importance in the following chapters.

First in chapter 2 a short overview should be given over the philosophy of science of about the last hundred years against the background of analogical reasoning and the investigation of epistemic techniques. As mentioned above the investigation of developing hypotheses was not really in the focus of philosophers of science during these hundred years (with only a few exceptions) and I will look for reasons why that was not the case and where instead the focus was laid. For this discussion, the distinction between the context of discovery and the context of justification becomes relevant which was introduced by Reichenbach (Reichenbach 1938, p. 6f). The context of discovery according to Reichenbach is the way the scientist publishes 'his logical reasoning in the foundation of a new theory', whereas the context of justification is 'the rational reconstruction' as part of 'the descriptive task of epistemology' executed by a philosopher of science. 'Rational reconstruction' means that the philosopher replaces in his description the finding and the presentation of a theory by the scientist with a 'logical substitute' consisting of 'justifiable operations' which are 'bound to factual knowledge'. In a modern introduction to the philosophy of science this claim of a normativity argumentation is expressed in the way that: firstly, the cognitive value of a hypothesis depends only on the rational justification of its truth demand' and secondly that the actual history of its discovery is irrelevant and therefore philosophy of science should deal only with rational justification (Schurz 2011, p. 21). Also, in the publication of Schurz, the second premise was criticized as a fallacy, however, in my opinion not in an enough depth because the essential epistemic contribution to a theory lies in the creative act of the hypothesis formulation. Only the creative act leads to a theory and not the investigation of a completed and established theory by checking its logical justification (see chapter 4 and 6). Revision of the logical justification can uncover weaknesses of that theory but not its epistemic achievement. This last is the point of my interest and where analogies come in, analogies perhaps not in all cases but very often.

To set the background for the discussion from another point of view namely that of physics an appropriate part of the historical development in high energy physics (further abbreviated with HEP) is described in chapter 3. The description comprises for the further discussion of the development the postulation of the constitution of the atomic core till that of quarks as components of mesons and baryons, together with all relevant discoveries of new particles and special consideration of analogical reasoning. Of course, not the whole story could be told, this would go far beyond the scope of chapter 3. It contains only relevant issues for understanding the paradigmatic examples much deeper discussed in the following chapter 4 investigating the role of analogies in physical theories. Also, chapter 3 is more oriented on physical textbooks than on historical scrutiny because it is intended to deliver the physical context for the discussion in chapter 4 in a form as physicists understand and teach them in their common practice, and not of the exact use of analogies in the corresponding physical theories, which is spared to chapter 4.

In chapter 4 first a short overview is given about speculations and theories of the structure of matter ending with non-relativistic quantum mechanics, developed in the second decade of the last century. In 1928 Dirac published a relativistic equation later named after him, which laid the foundation for the quantum field theories on which rests the standard model for particle physics mentioned in the beginning. In the same year as an interpretative consequence of this equation, a positive charged anti-particle to the electron called positron, and in 1932 also a further particle later called neutron were detected. This last particle, the neutron, gives the connection to the deeper discussion in chapter 4 because it plays an essential role in Heisenberg’s theory of the atomic nucleus which is analysed in detail there.

However, before this study can be performed, more knowledge must be gained on analogies, which are intensively used in Heisenberg’s theory and that of others. For additional motivation, I will give in chapter 4.1 a short introductory sketch of my own theory how analogies are used in physical theories. My theory is presented in detail in chapter 6.

Then, in chapter 4.2 the characteristic structure and function of analogies in science is investigated. After a short description of a conclusive general account of analogies pointing beyond scientific reasoning presented by Douglas Hofstadter and Emmanuel Sander in 2015, which supports *my thesis of the indispensability of analogies from the psychological side,* in chapter 4.3 I will analyse how the relevant types of analogies were employed in some examples of physical theories on the way to quarks.

The first example is as already mentioned the theory of the structure of the atomic core, published in 1932 and 1933 by Heisenberg in three parts. There he used several analogies to demonstrate that the atomic core can be described by quantum mechanical calculations at much higher energies than in the atomic shell. However, this was possible only if open questions about stability all were laid in the neutrons on the one hand and proton and neutron were taken as two different states of one particle on the other. These two states were later called ‘isospin’ of the nucleon, the name of the assumed particle with the two states. Then, under reasonable assumptions the stability or instability of all known and unknown atomic cores including the decay modes of the unstable cores could be predicted.

The second example by Yukawa, which is clearly less complex in comparison in comparison, is the theory of strong interaction forces modifying the analogous model of the Coulomb forces by an additional component, which reduces the range of strong interaction forces to the size of the atomic core. This suggestion is combined with the prediction of a massive particle, the pion, which was found in the end in 1947.

The third, last and most complex example in my analysis is a collection of papers in the period from 1953 to 1964, mainly by or with participation of Gell-Mann, which ended with the prediction of quarks as constituents of hadrons like the proton and the pion. In the first paper, he used the analogy to isospin to explain the properties of so-called ‘strange particles’. In a second paper together with Pais, Gell-Mann also referred in an analogous way to doublet and triplet states of then well-known nucleons and pions to set in relation the observed strange particles and predict two new ones, which were not already observed. In the third paper, Gell-Mann used in an analogous way a technique which already Heisenberg had introduced, namely the fictitious stepwise switching off weaker forces as electromagnetic and weak interactions, to show the symmetries which then occur in the pure strong interaction. The properties derived from this procedure allowed him to predict allowed and forbidden interactions and an additional new particle, too. All these predicted particles were later found, however, some further predicted particles were not. Two further papers, one together with Feynman, linked the properties of particles with strong interaction with that of weak interactions using a fictitious model of weak interaction between electron, muon and neutrino as model analogue for a symmetry model consisting of the nucleons (proton and neutron) and the strange particle Λ (lambda) forming an octet. Based on these papers, the next and last step to the postulation of quarks in a paper from 1964 was not far away because the triple structure of constituents for the baryons (nucleons plus Λ) occurred already in the paper before. Nevertheless, from the postulation of quarks to the complete standard model of particle physics still it was a long way going, too far away from the focus of this investigation.

The motivation for selecting these examples and their use of analogies was that they cover approximately the essential steps of the development, from the explanation of the atomic core to quarks as important elements of the standard model of particle physics, despite of course these men not being the only players in that game.

So far, the analysis regarding analogies in physics will be realized for theoretical work in chapter 4 but also analogical use in experimentation must be considered. This will be done in chapter 5. Again, examples of the period from the discovery of the neutron in a cloud chamber in 1932, which motivated Heisenberg to present his theory of the structure of the atomic core, till the postulation of quarks in 1964 and their proof of evidence using electronic counters in the discovery of a predicted new particle J/φ in 1974 were selected. However, additional examples, namely the time projection chamber and especially Monte Carlo simulations, must be given because time projection chambers complete a development from visual over logical to a union of both techniques manifesting an analogy of the observable over the unobservable to the observable. This means, simulations copy in some more or less analogical way a simplified version of the theoretical image of reality into a digital computer. The process of visualizing events in electronic detectors and simulations demonstrates a *third thesis of mine that visualization of relations is an essential analogy for understanding phenomena of unobservable objects in the real world*.

With the former chapter all parts of the analysis are together to flesh out the theory sketch of chapter 4 in chapter 6. This theory contrasts with most existing theories in philosophy of science by focussing on the epistemological process of developing physical theories and their testing as suggested by the analysis of the discussed examples in the foregoing chapters, i.e., it is focussed on the context of discovery as condition for justification without ignoring the latter. An essential component in this process is the use of analogies of two different forms in physical theorizing, and as further form mainly the visualization of unobservable phenomena after the model of observable phenomena. Visualization is the primary human method of representation because visual perception has evolved obviously as the most developed sense of human beings.

The details of my theory should be reserved for chapter 6, but I will mention some relevant consequences resulting from the analysis of the investigated examples and which are associated with my theory.

Firstly, only phenomena of observable objects can be explained. However, the use of analogical models can extend explanations to only indirectly observable areas. Indirectly observable means here that there are phenomena which are observable, but the objects which are supposed to cause the phenomena are not e.g., like the atoms, which emit radiation or particles. This procedure then can be continued under appropriate conditions from one indirectly observable area to further indirectly observable areas, like the emission of muons and neutrinos in the decay of pions, or electrons and neutrinos in the decay of muons.

Secondly, the development of these theories depends on the observed phenomena and the analogies used for their explanation in the corresponding theory, and that is the reason for the description of the development of philosophical and physical theories in historical order. This means that the direction of the development in science in the long run is unplanned, and unplanned means that the development in retrospect is dependent from history. If history would have taken another course physics today may have another form, I suppose. There is no guarantee that discoveries made in the past, or that models and analogies applied would have been developed under every possible circumstance.

This finding leads to a last point concepts which are developed by human beings use categories or analogies made by experience with phenomena of the world. Therefore, these categories must be compatible with the phenomena of the world, however, this does not mean that they represent the world in a realistic way because the categories are adapted also to our kind of perception, which was determined by our evolution to survive in an ecological niche and not to generate realistic representations of the world.

For science these three conditions result in the consequence that we have no satisfying explanation for phenomena that have no corresponding analogy (e.g., dark matter, dark energy, or - for a discussed example in chapter 4 - the constitution of neutrons for Heisenberg). Similarly, for theories which are not compatible with each other and where therefore was found no acceptable analogy, we have no satisfying explanation, too (e.g., for quantum gravity).

To summarize, I will argue in the following for three theses:

- *analogies are indispensable in the development and justification of theories*
- *because analogization is the method for categorization as foundation for building of concepts, especially unobservable ones and*
- *visualization of relations uses essentially analogies for understanding unobservable phenomena of the real world*.

In the conclusions I will recapitulate these three theses in the context of the most important supporting reasons.

2 Philosophy of Science and the Scientific Practices in Physics

To show as mentioned in the introduction that the restriction on the context of justification is not enough for understanding the epistemic process of scientific research and especially physics, where the most used examples come from, I will sketch the development of philosophy of science starting with the physicist Mach who maintained a more general point of view in some respect than most philosophers later. Most descriptions of the development of philosophy of science start with the Vienna circle and the Berlin group around Reichenbach because some of their proponents established the most accepted model of the philosophy of science after the Second World War. However, despite the Vienna circle and the Berlin group are mostly mentioned together and later subsumed under the notion 'logical empiricism' their positions were different, as we will see. Because logical empiricism wants to reconstruct scientific theories, analogies play no role. It will be shown in this chapter why this is so.

A severe challenge for this 'standard model' of philosophy of science becomes the revolutionary model of Thomas Kuhn in the sixties of the last century, which drew the attention on the change of theories. However, despite Kuhn's argumentation regarding the change of theories and their understanding, and the importance he laid on the kind of education in physics, in Kuhn's theory analogies occur only in a special meaning: these are what he called 'paradigms'. Paradigms were changed alone during crises caused by enduring anomalous phenomena in science, crises not explainable by the established theories of his time. In contrast, in my view analogies were used all the time in physics.

A consequence of Kuhn's theory was the beginning of a discussion about the relativism of theories and world views and their realism or anti-realism. One of the most influential anti-realists was Bas van Fraassen, who insisted in his theory of 'constructive empiricism' that all that can be demanded in science is empirical adequacy, and that the acceptance of theories has an irreducible pragmatic aspect. Especially the second point has a relation to analogies as we will see, despite analogies are not mentioned by van Fraassen at all.

Another conclusion from constructivist viewpoints and what he calls ‘distributed cognition’ was found by Ronald Giere in his ‘scientific perspectivism’. ‘Distributed cognition’ means the use of diagrams, pictorial representations and physical or abstract models as mental tools for theorizing. ‘Scientific perspectivism’ re-establishes the connection of human perception and scientific theories already found with Mach, claiming that theories, observation, and measurements of scientific instruments depend on the kind of detector and human perception likewise. This point of view was taken up again by van Fraassen to develop his ‘constructive empiricism’ to an ‘empiricist structuralism’. This ‘empiricist structuralism’ built up on a more detailed discussion of representation and its need of a standpoint and reference frame or perspective in Giere’s sense, called by van Fraassen ‘indexical’ on the one hand and structuralism on the other. Structuralism in the view of van Fraassen is the assumption that we cannot know how nature really is, but only represent it in a way that it is empirical adequate to the phenomena (in some respect).

Also, R.I.G. Hughes uses models in his DDI-account of physical theories which represent phenomena as abstract theoretical constructions. DDI stands for denotation, demonstration, and interpretation. According to Hughes a theoretical model can be interpreted as subdivided in three parts: denotation, which defines the used variables, demonstration, which explains the behaviour within the model, and interpretation, which transposes the results back to the world.

All these approaches to physical theorizing deal with the explanation of already formulated and therefore existing theories and their justification but not for the purpose of heuristics, i.e., epistemic tools to find new theories. As mentioned above, only recently some philosophers (e.g., Danks and Ippoliti) paid attention to the process of discovery but only as I think in a focus to limited to give an adequate picture because they mostly assumed the development of hypotheses is only problem solving. Of course, problem solving is one aspect of developing hypotheses, however, observed phenomena may be new and surprising, nevertheless they stimulate in a lot of cases associations to similar already known phenomena which gives an intuition of a theoretical approach, i.e., an analogical model. This sketch of philosophical development in the following should set the background to the discussion of analogies in physics.

2.1 Physicists and Philosophy of Science

The topic of how to develop theories in physics began to gain the attention of researchers starting in the 19th century, when the physicist Ernst Mach developed a view on the philosophy and psychology of research, which is noteworthy to consider still today in some respects, so it is obvious to start with him.

Mach was an experimental physicist, who studied mathematics and physics and was employed since 1895 as a professor of “Philosophy, especially History and Theory of Inductive Sciences” in Vienna (Ganslandt 1995, p. 730). This last position was conferred to him due to his investigations of sensorial perceptions and epistemological questions of science and especially of physics. Mach was deeply convinced that science and especially physics only could throw light on its foundations by research in biology through analysis of the perception by the senses (Mach 2015 [1922], p. VI). Mach’s aim was to connect the view of physics of the world to the sensations of the mind mediated by the perception of the senses, expecting that in the end relations between the facts (Tatsachen) of the world and the sensations could represent all empirical knowledge without metaphysical speculations. However, this is an aim we are so far away from today, as more than hundred years ago and one that presumably never could be reached. Additionally, in my view it is not enough to refer only to sensations because if one assumes a phylogenesis of human beings as Mach did, one must take for granted that there is a world where this phylogenesis take place, even if it cannot be recognized in its reality.

However, intending a functional relationship Mach has contributed a lot of interesting considerations on the process of gaining knowledge in science. There are about seven main points to consider. The first point is the importance of history in the development of science. The second is the entanglement of theory and measuring practice in this development. The third is a limitation on sensorial facts and therefore a refusal of metaphysics. The four last points are his studies on the meaning of a concept, the importance of analogies, the difference between physiological and physical space and time as well as the meaning and value of natural laws subsumed under the notion of empirio-criticism. Thought economy is only a lesser important aspect allowing abbreviation in speech in the meaning of a concept because only ordering and categorization in more and more general and abstract concepts allow orientation in the world in an increasing sophisticated way. As Mach emphasizes, animals also have concepts however in a more rudimentary form.

Mach assumed that human beings like animals learn in the beginning simple concepts, which were extended later to more complicated and abstract ones. However, because of the richer intellect and the cultural achievements of human beings they learn faster and more comprehensive, a quantitative but no qualitative difference (Mach 2015 [1906], p. 74). Essential for this learning of both, human beings, and animals, is the forming of more and more abstract concepts based on these experiences, of course, presumably more complex and abstract in human beings. A concept in the sense of Mach is the generalized categorization of positive or negative experiences, made with objects or situations. These categorizations, sometimes called ides, are used instinctively or intuitively to decide in new situations how to act (Mach 2015 [1906], p. 127). As a result, 'knowledge and error follow from the same psychological sources, only success can divide them. The clear realized error is corrective and also instructive as a positive realization' (Mach 2015 [1906], p. 116). Therefore, a concept is a living 'object', i.e., it changes during the whole lifetime.

One aspect of this change of concepts can be seen by Mach's description of the development of the thermometer, which needed about 225 years from the first attempts to measure the temperature by an air thermoscope presumably in 1592 till the proof of principle in 1817 (Mach 2014, p. 38). During this time, not only the concepts changed several times but also the kinds of measurement (Mach 2014, p. 3ff). The invention of the open air thermoscope was ascribed to Galilei, who claimed this also for himself. This early thermoscope varied in its indication by the actual atmospheric pressure and was therefore not reliable. Additionally, two reliable reference points were needed to define a scale, which was not available for Galilei. About fifty years later the first closed liquid thermometers filled with ethyl alcohol were constructed. Again, about thirty-five years later when Boyle and Mariotte had found the gas laws named after them, a closed air thermometer was built by Sturm and others. However, different scales and reference points were still used by different scientists. This was changed in 1817 by Gay-Lussac with the realization that all (ideal) gases expand at constant pressure in the same way with the temperature and - using an idea of Amontons about hundred years earlier - have therefore their minimal volume at -273 degrees centigrade. These developments in both theoretical and experimental domains show that they are entangled, and none can develop without the other.

Another aspect in the change of concepts is the question how they are formed, especially in science. Mach did not assume that there are essential differences between everyday concepts and those in the sciences. In Mach's view concepts originate from the sensations and their connections, where concepts give the sensations an ordering. 'The aim of concepts is to give our sensual stimulated imaginations the best correspondence to the sensations in the most comfortable and shortest way' (Mach 2015 [1906], p. 144), i.e., what he called 'thought economy'. This 'thought economy' however is no satisfying explanation because it is in no way clear that the claimed correspondence is in fact realized, as we will see. What alone can be claimed and what Mach claimed as well (Mach 2015 [1906], p. 220ff) is that similarities in some relevant aspect in our sensations can be categorized to build up a concept related to that aspect, i.e., analogies are the essential tool to bring order into the world. Relevant in this sense are all aspects, which are learned first from our parents or later from other human beings and also from own experiences needed for one's own survival and welfare. However, in this statement only relevant social and cultural experiences are considered and that is not all what is needed.

As mentioned above, Mach was intensively interested in psychological and physiological questions of perception and participated in its investigation in some degree. From this knowledge he concluded that the space of perception and the geometrical space of physics are not the same (Mach 2015 [1922], p. 282; Mach 2015 [1906], p. 337) the former is much more complicated than the latter, i.e., the latter is a reconstruction out of the combined perceptions of different senses, namely the visual, the tactile and a kind of movement sense and with some restriction for human beings the auditory and olfactory sense. In fact, the auditory and olfactory sense are for many animals much more important than for human beings and they are also combined with the mentioned movement sense by turning the ears or nose in the direction of a noise or a smell. With the movement sense is meant the information of the position and movement of the muscles in the body. The same statement is valid for the perception and physical measurement of time (Mach 2015 [1922], p. 285; Mach 2015 [1906], p. 423). The perception of space and time is dependent on the phylogenetic and ontogenetic development of our mind and body, geometry and chronometry from the comparison of the properties and behaviour of physical bodies, however, heavily formed by the imaginations and concepts developed in physical experience (Mach 2015 [1906], p. 425). Therefore, laws of physics are not laws nature must obey, but rather expectations we have on the properties and behaviour of observed objects in nature (Mach 2015 [1906], p. 449).

In his attempt to reduce all experience on sensorial perceptions, Mach is convinced that only empirical facts are relevant for theories in the sciences. Because empirical facts depend on their historical appearances, the sciences are guided by their history and that the completeness of theories about the world can be formulated only as aim. However, the restriction on empirical facts can be exaggerated as also in the case of Mach. So, Mach insisted on the belief that 'all practical and intellectual requirements are satisfied when our thoughts allow completely to reconstruct the sensorial facts' (Mach 2015 [1922], p. 257). This attitude found its expression in his view that a theory should in the end only describe the connection between sensations in an elementary form. Empirical facts, i.e., all what goes beyond such descriptions, e.g., like explanations based on unobservable atoms, forces, or laws, are to be rejected as metaphysical, if they are not helpful for the reproduction of the sensation-to-fact connection.

As a short summary: Already about hundred years ago a physicist and philosopher has stated the importance of analogies for science and especially physics. However, his successors have laid the focus on another argument of him: Mach's position against metaphysics. For obvious reasons, this anti-metaphysical attitude was also adopted as unifying criterium by the 'Vienna circle' to be considered next.

2.2 Philosophy of Science and Metaphysics

The successor to Mach's chair of "Philosophy, especially History and Theory of Inductive Sciences" was the philosopher Moritz Schlick, the founder of the second 'Vienna circle'. This foundation followed a suggestion of his students Friedrich Waismann and Herbert Feigl (Lorenz 1995, p. 695). Other members of the Vienna circle were the mathematician Hans Hahn, the physicist Philipp Frank and the economist Otto Neurath, who were also members of the first 'Vienna circle' constituted during the lifetime of Mach to discuss Mach's phenomenalism and his principle of thought economy with respect to the conventionalism of Poincaré and the holism of Duhem (for Duhem see also chapter 4.2.1, p. 59ff). Poincaré's conventionalism means that the axioms of geometry or principles of physics are not experimental facts but convenient agreements and Duhem's holism means that not a wrong single scientific theory can be rejected but only the whole system of theories, because the point where the error occurs cannot be located within these theories (Wolters 1995, p. 463). These last three members of the first Vienna circle influenced the philosopher Rudolf Carnap in his phenomenalist approach (Carnap 1974, p. XII) and the later separation of theory and observation language in logical empiricism (Lorenz 1995, p. 695).

Carnap was the most important member of the Vienna circle after the only partly cleared murdering of Schlick by a student of Schlick. In his phenomenalist approach, Carnap tried like the 'elementary sensations' of Mach to connect 'elementary experiences' via a similarity relation as basic relation to constitute basic concepts and to define more abstract concepts by combinations of these basic concepts (Carnap 1974, p. XII). Carnap's aim was to reconstruct a rational and more exact and clearer system of concepts in science using the achievements of modern logic, in contrast to the common more spontaneous and unconscious development of concepts (Carnap 1974, p. X). With other words that means to express the formal structure of a scientific theory and in the end to include all scientific theories in one constitutive system in an abstract way (Carnap 1974, p. 252). This aim was strongly influenced by the 'Tractatus logico-philosophicus' of Ludwig Wittgenstein published in 1921, where Wittgenstein intended to limit human thought with the claim that this limit can only be reached by language from the inner side - beyond this border there is no sense (Wittgenstein 1984, p. 9). Therefore, Carnap argued that propositions constituting the intended unique system can only be true or false; they are true if the proposition corresponds to the basic experiences, otherwise they are false. Propositions which do not refer to this verification principle of meaning ('principle of verifiability') are metaphysical and have no sense, e.g., questions like whether the external world is real or not. In comparison, Wittgenstein maintains that 'the most sentences written about philosophical objects were not wrong but nonsense' (Wittgenstein 1984, § 4.003). Therefore, both mean 'metaphysics is to ban from philosophy' (Carnap 1974, p. XIX) despite a different reasoning. Also, the definition of the main task of philosophy is the same; where Wittgenstein states that 'the purpose of philosophy is the clearing of thoughts' (Wittgenstein 1984, § 4.112), one finds with Carnap 'such a clearing of concepts today often called 'explication' seems to be still one of the most important tasks of philosophy' (Carnap 1974, p. X). Carnap distinguishes in this tradition between 'object questions' and 'questions of logic', logic in the wider sense of modern logic, and argues that 'object questions' fall alone in the responsibility of the empirical sciences and only 'questions of logic' are meaningful philosophical questions (Carrier 2009, p. 24). This differentiation meant that the former, the scientists, consider properties of objects and their relations, whereas the latter, the philosophers, consider concepts and propositions used for the description of these objects. However firstly, this 'linguistic turn' as it was called later has the disadvantage that it limits a concept to a linguistic form, which is too restrictive because a lot of concepts could not be expressed in words without using analogies, e.g., describing how it is to see a red object. Already Mach has accepted (nonverbal) concepts with animals (see p. 17). And secondly, the restriction on 'questions of logic' neglects the interesting epistemic question, how a scientist develops a theory or which concepts and imaginations determine the design of a theory or alternatively an experiment. We will find this limitation as already mentioned again and again in the following decades. One reason is, I suppose that the intended logical reconstruction of an existing theory excludes to get its development into one's focus.

The so-called Berlin group is often viewed as an external part of the Vienna circle and in fact there was some contact between the two groups and they have some views in common, however, there were also several substantial differences in their opinions. Both groups dissolved after Hitler's seizure of power and most members of both groups emigrated into the United States. Today the best known representants of the Berlin group were Hans Reichenbach and his student Carl Gustav Hempel, who both emigrated into the United States, too. Reichenbach was not interested as the Vienna circle in the explication of the language in science but in clarifying the meaning of e.g., physical theories – independent of the interpretation by their creator (Milkov 2011, p. VIII). To do this clarification he was oriented in his influential three books on relativity (English translations: The Theory of Relativity an A Priory Knowledge (1920), Axiomatization of the Theory of Relativity (1924), and The Philosophy of Space and Time (1928)) on the axiomatization program of David Hilbert and not on logic like Carnap (Milkov 2011, p. IX). Both theories of relativity (special and general) were interesting for him as for Schlick, who also wrote a book on relativity[6], because in their view in relation to Newtonian mechanics special and general relativity gave an example how in science (here physics) observational, i.e., empirical facts could be separated from metaphysical positions and the latter eliminated. Metaphysical positions should be eliminated e.g., absolute rest and velocity or equally absolute space and time in special relativity in relation to classical Newtonian mechanics, and

[6] Raum und Zeit in der gegenwärtigen Physik. As Carnap, too: Der Raum – Ein Beitrag zur Wissenschaftslehre

absolute (gravitational) acceleration in general relativity in relation to special relativity (see Friedman 1983, p. 6). In general, Reichenbach claimed that scientific theories have more contributed to knowledge than all philosophical systems so far and therefore these theories must be investigated looking for their individual requirements and principles (Reichenbach 1931, p. 52). In physics e.g., coordination principles are needed which connect a theoretical object, a mathematical entity, to a measurement process. This is not unproblematic as the example of temperature measurement and the long time for its establishment shows (see p. 17f). The main problem is that the theory describes relationships between mathematical objects, but the measurement process is an act in the world. In his first book in 1920 Reichenbach assumed that these relationships must be given a priori e.g., measuring the length by rigid rods, however, contrary to Kant these principles are held as revisable e.g., by changing the method of measurement using the duration of time for a light signal reflected by a distant object (see Friedman 1983, p. 19). Therefore, later in 1924 and influenced by Schlick, Reichenbach changed his notion of ‘coordination principles’ because of this revisability into the notion ‘coordination definitions’, indicating that they are arbitrary and merely conventions (Milkov 2011, p. XIf). After his second book on relativity Reichenbach changed his opinion because of the problems regarding the relationship between theoretical and observational definitions, which led in consequence to a break between him and Schlick.

From the beginning, the Berlin group was more interested in the practice in single sciences than the Vienna circle, who wanted in majority to develop a so-called ‘unity in science’, which meant to find a formal language suitable for applying in all sciences (and humanities?). In contrast, Reichenbach saw the task of philosophy in the analysis of probability and induction problems in the different sciences like logistics, physics, biology, and psychology (Milkov 2015, p. XIV). Indeed, Reichenbach defended the opinion that knowledge was gained by induction and probability (Reichenbach 2011, p. 99). Scientific theories are in this view inductive concluded predictions which are not true or false but more or less probable (Milkov 2011, p. XIV). Induction was understood as setting a prognosis about future events to overcome Hume’s induction problem.

Where Carnap tried a logical (re)construction of the world, Reichenbach saw the task of philosophy as organizing the results of the sciences in a logical and epistemological way as he did in physics. Logical means that propositions of science must be justified, and the concepts must be well defined. Epistemological means that theoretical constructions correspond in a correct way to perceptions and the common physical world (Milkov, 2015, p. XV). To distinguish the task of a scientist and a philosopher he introduced the separation of the context of discovery and of justification, where the former is assigned to the scientist and the latter to the philosopher (Reichenbach 1938, p. 6f). This distinction does not mean that the scientist does not justify her theories but should indicate that the philosopher must reconstruct with more epistemological scrutiny and logical exactness. Here is the reason why most philosophers in the last century are not interested in the way a scientist gained his knowledge namely why philosophers restricted themselves on the reconstruction of already available theories in an in their view more well-founded manner. However, this is not my interest because it neglects just the empirical way a scientist develops his theory. Usually, this interest was discredited by hinting to the example of Kekulé and how he found the structure of the benzol ring by imaging a snake biting in its own tail. But this response takes the process of theorizing much too easy. As we see, here is an analogy at work - the ring of the snake compared to the ring of benzol - and this is a texture found in a lot of other scientific theories as we will see.

As mentioned, Hempel was a student of Reichenbach but also spent one semester in Vienna, where he also participated at the regular meetings of the Vienna group and was impressed by Carnap's book on the logical constitution of the world. Hempel began his dissertation with Reichenbach in Berlin but finished it in 1934, when Reichenbach already had emigrated to Istanbul before later going to the United States. After his dissertation Hempel went to Paul Oppenheim in Brussels (Gethmann 2004, p. 72f). A later result of this collaboration of these two men was their famous paper about scientific explanation from 1948. In this paper Hempel and Oppenheim investigated the characteristics and function of explanation, and in contrast to Mach and others took for granted that description of phenomena in the world is not enough because science must answer not only the question 'what' but also the question 'why' something happens (Hempel & Oppenheim 1948, p. 135).

Hempel and Oppenheim analysed two main parts as 'basic pattern' for an explanation, which they called 'explanandum' and 'explanans' (Hempel & Oppenheim 1948, p. 136). 'Explanandum' means the sentence describing the phenomenon which should be explained and 'explanans' the sentences forming the arguments which explain why that phenomenon happens. The explanans divides according to their analysis in two classes one with conditions, which determine the occurrence of the specific observed phenomenon, and the other a general law, which describes the regularity of a whole class of phenomena of the same kind (Hempel & Oppenheim 1948, p. 137). This pattern resembles intensively on physical laws and their initial conditions, where it come from. However, this pattern should also be valid for other sciences e.g., in chemistry, psychology, sociology, economy, or biology as Hempel and Oppenheim claimed (Hempel & Oppenheim 1948, p. 141). In the view of Hempel and Oppenheim the extension to other sciences, especially the humanities, demands to exclude teleological, motivational, or 'empathic familiar' reasons for an explanation (Hempel & Oppenheim, p. 142ff). This means on the one hand that events in physics or chemistry as well as human acts are not unique in the sense that there are always attendant circumstances that are different and therefore not relevant for the explanation as e.g., complex motivational or teleological aspects. On the other hand, Hempel & Oppenheim also rejected the idea of 'empathic familiarity' i.e., that e.g., the free fall of a body is more familiar than the gravitational law and therefore easier to accept.

In addition to the described pattern some further logical conditions must be fulfilled to guarantee a sound explanation. Firstly, the explanandum must be logically deducible from the explanans. Secondly, as already mentioned the explanans must contain and also use general laws. Thirdly, the explanans must be testable in some way i.e., the laws must be confirmed under the special conditions provided by experiments. Lastly, all sentences of the explanans together must be true (Hempel & Oppenheim 1948, p. 137). A consequence of the last point is as Hempel & Oppenheim also stated that the validity of this condition must not be fulfilled for all times because later tests may falsify it and the explanation must be dropped then. Following these conditions, the basic pattern has the form shown in Table 1 (Hempel & Oppenheim 1948, p. 138).

Table 1: Basic Pattern of Explanation by Hempel and Oppenheim

	C_1, C_2, C_3, …	antecedent conditions	
	L1, L2, L3, …	general laws	explanans
deduction	____________		
	E	description of the empirical phenomenon to be explained	explanandum

Also, Hempel & Oppenheim claim that the same pattern applies not only to explanations but to predictions, too. In the case of explanations, the phenomenon of the explanandum E (see table 1 above) has been observed and the antecedent conditions and the general laws are found in retrospect. In the case of a prediction the explanandum is concluded from the general laws under specified conditions. Thus, a prediction can be used as a test of the theory based on this basic pattern. However, one must consider that especially in common use outside the sciences often incomplete explanations are given, where only some special conditions and no general law are articulated because they are alleged as obvious and therefore must be completed for a full understanding.

Later Hempel called his model of explanation the deductive-nomological or DN model (Hempel 1995 [1962], p. 686), which formed together with the 'double language model' and the theory of confirmation by Carnap as mentioned the widely accepted 'standard model' (see p. 15) of philosophy of science (Carrier 1998, p. 28). The double language model was developed by Hempel and Carnap trying to solve a problem of the principle of verifiability, namely that theoretical concepts are not verifiable by empirical methods, so that they distinguish an observational and a theoretical language as different domains connected by correspondence rules between at least one notion of each domain (Carrier 2004, p. 284). The observational language should describe pure empirical phenomena in a language protocolling directly the phenomena. The theoretical language could define concepts on an axiomatic way e.g., like electrons but must connect some of them to observational notions like tracks. The theory of confirmation by Carnap, which was later took over by Hempel suggested to support the 'creditability of a scientific hypothesis as relative to a given body of knowledge' (Hempel 1998 [1966], p. 457). This means that the hypothesis must not contradict to that knowledge and must fit with its concepts.

Again, in this 'standard theory' analogies play no role. The focus is laid on the reconstruction of science (language) regarding theories in the tradition of the Vienna circle and the Berlin group, however, it considers additionally the possibility of predictions and extends in this way the pure reconstruction. The question of how a scientist gains the knowledge to formulate the general laws and its special conditions, is not considered as relevant for the philosopher just in agreement with Reichenbach. As we will see in chapter 4 this weakness of the theory of Hempel was attacked by Hesse in 1966 in favour of using analogies. However, her critique found not much attention at that time because a few years before Kuhn in 1962 had published his much more influential approach to the development of science in general.

Despite Kuhn's influence I will continue first with Karl Popper, who was at the first glance a philosopher interested on the context of discovery, as the title of his first book 'Logik der Forschung' (Logic of scientific discovery) suggests. Contrary to this assumption, Popper thought that proposition of hypotheses and the context of discovery was a psychological question alone and not a philosophical one because the task of a philosopher was the justification of theories (Popper 1976 {1935], p. 6) in agreement with the Vienna circle and the Berlin group. From the proposed hypothesis should be derived conclusions in a logical and deductive way (Popper 1976 [1935], p. 7), which led to the name hypothetical-deductive (HD) method not to be confused with the DN-theory of Hempel. These conclusions are used to find possible contradictions by comparing the different conclusions, to avoid e.g., tautologies by analysing their logical structure, to compare them with other theories to evaluate their novelties, and of course to apply them in experiments looking for testing their conclusions by falsification or provisional confirmation (Popper 1976 [1935], p. 7f).

Popper was a important philosopher with connections to the Vienna circle, however, he was never an official member. In some respect, Popper was nearer to the views of the Berlin group regarding his interest on real scientific practice than to those of the Vienna circle (Milkov 2011, p. LII) but he was a thinker of his own. Popper tried also like the other both groups to separate science from other non-scientific enterprises but by using the scientific method instead of language for separation (Carrier 2009, p. 29). His most important work was on the 'logic of scientific discovery' published 1935 in German, however, becoming influential only since 1959, when it was translated in English. Popper was convinced that induction rules together with the principle of verification could not be a reliable method for gaining knowledge because an inductive conclusion could verify a theory only till the respective present but may be falsified in the future e.g., for a long time in Europe only white swans were known, so that it was assumed that only white swans exist, until black swans were found in Australia (Popper 1976 {1935], p. 3). Therefore, he insisted that theories could never be verified but only falsified. Together with this falsification principle, Popper argued that every hypothesis is allowed, so long as it is testable by experiment; the more unlikely it is, the more explanatory power it has (Popper 1976 [1935], p. 84).

Popper already saw serious objections against his falsification criterium, namely a strategy of immunisation or by changes of a convention. He recommended to avoid these problems by the decision or methodological rule not to try to save a theory by such a strategy (Popper 1976 [1935], p. 50f) i.e., a not very secure but perhaps the only method.

A main problem of Popper's theory is the reference to 'basic sentences' (Popper 1976 [1935], p. 60ff). A basic sentence is for Popper a singular claim that at some point in space and time an observable event or process takes place (Popper 1976 [1935], p. 69) i.e., a prognosis, which can be accepted as proved or falsified. With this definition, he wants to avoid a psychological reference to perceptions. His reasoning was that perceptions are not objective in the sense that they are intersubjectively provable because a sensation is always subjective and cannot be shared between different subjects. The notion of 'observable' in Popper's claim is that it is one, which cannot be defined using language and is certainly a psychological concept, but it is meant only as explanation of some not definable basic concept like a mass-point in physics and it is also not restricted to visual observation. With reference to other senses, Popper will break the subjective limit, arguing that other perceptions deliver further confirmations of the objects and processes in the material world. However, according to Popper this representation is only meant as an explanation but not a definition and it is only specified by its usage in language because of the reasons given before (Popper 1976 [1935], p. 69). Therefore, a theory is with Popper like with Carnap, Reichenbach, or Hempel an axiomatic logical system without relation to the empirical world (Popper 1976 [1935], p. 41). However, the claim of the basic sentences resulting out of a theory has as such also no relation to the empirical world therefore the possibilities of immunisation or conventionalist changes stay intact. The demand of observability in the basic sentences is in my opinion not suitable to establish the needed connection to the empirical world. For this connection is needed, what in science really takes place, the comparison between events in the real world. The individual sensations about these events can be shared by subjective observation and the talk about these sensations as subjectively similar: 'I see that event A is like event B. You agree?' This statement is possible because both observers compare two perceptions of their own representation of the world and agree on them as comparable or similar. In this way, I suggest a solution is given for the problem of the connection to the world, but not with the basic sentences suggested by Popper. This failing of a solution is a result of Popper 's exclusion of the context of discovery (Popper 1976 [1935], p. 65), which cuts off also the historical context which is relevant for the development of the basic concepts, as we will see in chapter 4. Nevertheless, Popper is convinced that the formulation of hypotheses is dependent from the perspective of a scientist, with perspective characterized as a searchlight (Popper 1994, p. 175), but history plays no role because science is interested in the truth of universal laws valid everywhere in space and time. Because we don't know the truth, wrong hypotheses must be eliminated by falsification (Popper 1994, p. 179) i.e., Popper assumes that science is a rational enterprise for the accumulation of (true) knowledge of the world (see Popper 1994, p. 44f) by exchanging old, falsified theories by new better and often more general ones, which is why he called his view 'critical rationalism'.

2.3 Philosophy of Science and Theoretical Construction

In contrast to Popper's view Thomas S. Kuhn emphasizes the historical context and argues for a historical dynamic of theory change. Further, he challenged the assumption of continuous knowledge accumulation and rationality of science development referring to social factors (Lorenz 2004, p. 504). Kuhn assumes in his theory a general development in three steps, where the first step is passed only one time but the other two could be repeated alternatively with no end.

The first step in Kuhn's theory, called 'pre-normal' or 'pre-paradigmatic' phase, is identified by a concurrence of several different views on a specific domain of phenomena e.g., electricity to mention an example of Kuhn with no satisfying explanation for all the then-known phenomena (Kuhn 1976, p. 28f). None of the different views explain the same phenomena well and none against the same background and, of course, none has the power to persuade the proponents of the other views from their own merits. The pre-normal phase is left only then, when one of the views could be improved in a way to explain also some of the before-unexplained phenomena and showed an advantage against the other, so that it convinced the most scientists working in this domain phase two of the development arises.

The second phase of development is called by Kuhn 'normal' or mature science (Kuhn 1976, p. 37) because it usually lasted a longer time and is dominated by a special 'paradigm' (Kuhn 1976, p. 25). A paradigm comprises an idea of the nature of the appropriate domain, the acknowledgement of solutions of special problems as exemplary and the acceptance of particular quality standards for explanations (Carrier 2009, p. 30f). All scientific work in normal science is governed by this paradigm. The classical Newtonian mechanics and its gravitation theory is an example for a paradigm. Students learn the tradition of the scientific domain indirectly by solving the exemplary problems to get the idea of the governing paradigm because it is important to learn the way of solving such model problems and not the special result of this special problem. This kind of way of problem solving is a tacit knowledge as part of the paradigm. A consequence is that always a solution in a special form exists because the problem is defined by the paradigm and only the special solution is unknown. Kuhn called the work on these problems 'puzzle solving' because the problem solving is like composing the pieces of a puzzle in the right way (Kuhn 1976, p. 50f). A further consequence, in this case social, is that if a solution is not found this indicates a failure of the scientist, not of the accepted theory. All work within normal science is aimed at the completion and refining of the theory and not on its falsification, contrary to the opinion of Popper (Kuhn 1970, p.5). Four reasons are relevant for this attitude according to Kuhn. First analogous to the rules of a puzzle there are the 'explicit claims of scientific laws and statements on scientific concepts and theories' (Kuhn 1976, p. 54). On a more basic level are 'preferred kinds of technical equipment and their permissible application' (Kuhn 1976, loc. cit.). Then, more general are metaphysical and methodological assumptions on the characteristics of nature e.g., that the world consists of formed matter in movement, which must be described by laws on their structure and interaction (Kuhn 1976, p. 55). At last, there are some bindings like the obligations of extending the knowledge on the world (Kuhn 1976, loc. cit.).

Falsification is only an intention in the third phase of scientific development, namely in 'its occasional revolutionary parts' (Kuhn 1970, p. 6). However, scientific revolutions are bound on special conditions. These conditions, which lead to the end of a period of normal science, are so-called 'anomalies' (Kuhn 1976, p. 65f). An anomaly is the detection that the expectations on phenomena in nature, which are based on a view determined by the established paradigm of normal science, are not fulfilled. The reaction is an intensive research of the attendant circumstances of the anomaly, which result, when they are understood, in an adaption of the paradigmatic theory. Only then, if the view on the nature of these phenomena has changed, they are scientific facts (Kuhn 1976, p. 66). Anomalies are not only a result of new detected phenomena but are also at work when new theories arise. Theoretical problems like that in the geocentric Ptolemaic system of astronomy, where the needed calculations quickly became more complicated than accurate, opened the mind for possible new solutions like the heliocentric Copernican theory (Kuhn 1976, p. 79ff). Realizing this Kuhn makes the assertion that nly in these cases, where two or more theories like the Ptolemaic, the Copernican and Tycho Brahe's which was a mixture of both other are in concurrence, the proponents of each try to falsify the other. When this conflict is resolved by acceptance of one of the opponents, a new paradigm is established, and a new phase of normal science begins. However, with the new paradigm the view on the scientific domain has changed in an incommensurable way as Kuhn claims (Kuhn 19,76, p. 123). Kuhn compares this change of view with the gestalt-switch in the well-known picture puzzle of a rabbit and a duck. Another consequence of the change of views is that as Kuhn assumes the development of science has as evolution no direction and therefore no progress and approximation to a true picture of nature (Kuhn 1976, p. 171ff).

Kuhn has directed again the attention to the importance of history and its development, however, in a relatively general form. Analogies work finer and, in more detail, as we will see in chapter 4. Additionally, incommensurability and non-cumulativity of knowledge is a kind of perspectivism.

Bas van Fraassen argued like Kuhn against realism i.e., that scientific theories give a true picture of the world or at least approximate to it, a picture as it was taken for granted e.g., by Popper. Van Fraassen distinguishes in his 'constructive empiricism' (van Fraassen 1980, p. 12) two sorts of philosophical studies on the aim and structure of science (van Fraassen 1980, p. 2). Of the two sorts of studies, the first is directed on the content and structure of theories i.e. usually the explanation of phenomena, which are observable (otherwise they were no phenomena) by 'postulating other processes and structures not directly accessible to observation' (van Fraassen 1980, p. 3) and the second on the relation of theories to the world respective to the user of the theories i.e., if the theories are a true explanation of the world especially regarding the unobservable parts or otherwise, when a theory is accepted by the user. Van Fraassen is interested of course on the first kind of the second sort i.e., when a theory is accepted by the user, because he assumes in his standpoint of this time only observable facts of the world as real given, which I do not for reasons discussed later. My interest goes across the two sorts, it is just the relation between the first and the second type of the second sort i.e., how the user develops a theory and what effect the theory has on the user, when she accepts it.

Regarding the relation of theories to the world van Fraassen defended three main positions. Of the three main positions taken by van Fraassen in his 'constructive empiricism' only the first two are relevant in the considered context here because the third deals with statistics and probability in scientific theories, which lays outside the focus of this paper. Van Fraassen's first position is that '*science aims to give us theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate*' (van Fraassen 1980, p. 12, emphasis in original). Empirical adequateness requires according to van Fraassen 'only to give a true account *of what is observable*' (van Fraassen 1980, p. 3, emphasis in original). This statement shifts the question from adequateness to observability and consistently van Fraassen investigates the border between observability and non-observability, a question also relevant for my purposes. Van Fraassen answers this question by reference to the human sensorial abilities i.e., 'X is observable if there are circumstances which are such that, if X is present to us under those circumstances, then we observe it' (van Fraassen 1980, p. 16). As example van Fraassen mentions the moons of Jupiter observed through a telescope and micro-particles in a cloud chamber. The first entities are observable for him because if astronauts would visit Jupiter, they could see them by bare eye, but the micro-particles are not because, if the theory is right, one sees the vapour trail but not the micro-particle. For van Fraassen the moons are observable in principle but not the micro-particle. I agree to this position, however, in a limited historical way, historical in the sense that that differentiation plays a role in the historical development of science as the example of Mach and his rejection of atoms and molecules show. Contrary to van Fraassen a micro-particle is observable for me today in an indirect way. The reason is that in my view the truth condition for observability in case of empirical adequacy is not fulfillable. This means, I can confirm empirical adequacy to my observations, however, this statement must not be true because, if e.g., the standard theory of particle physics is right, then the observation of the Jupiter moon is not true in the sense of observing a simple rigid body. Nevertheless, the description of the Jupiter moon as simple rigid body is approximately empirical adequate in view of a Newtonian model of the system of the moon orbiting around Jupiter. That is, I accept the empirical adequateness condition of van Fraassen, however, not his anthropocentric truth condition of observability because it is not ensured that our senses give a true picture of the world. From this perspective our senses as all theories establishes only an empirical model of the real world, not the world as it is.

The concept of a model was also used by van Fraassen in reference to the semantic view of theories and in contrast to the syntactic view (van Fraassen 1980, p. 41). The older syntactic view e.g., of Hempel or Popper, prescribes that a theory consists out of a deductively organized sequence of sentences formulated in an uninterpreted or syntactic language, which is in a second step connected to empirical situations by correspondence rules (Carrier 2004, p.271). In the semantic view the relation is reverse, theories were constructed by a set or family of models, which were described by the structure of a set of true sentences forming the model (Frisch 2005, p. 6) or, as van Fraassen characterizes it, 'any structure which satisfies the axioms of a theory […] is called a *model* of that theory' (van Fraassen 1980, p. 43, emphasis in the original). However, this is not the kind of model I meant above. By my use of the notion 'model' it is meant something like the model of hydrogen used by Bohr intended to represent phenomena.

The second main position of van Fraassen regards a theory of scientific explanation. Van Fraassen argues that a scientific explanation is a context sensitive answer on a why-question which does not rise the evidence of a theory but supports only the 'desires for informative description' depending on its context, which is just a pragmatic aspect (van Fraassen 1980, p. 156). Van Fraassen points out, in my view more correctly, that explanation is a three-valued relation between theory, facts and context. However, I do not follow him in his assessment of explanation as just a pragmatic aspect because an explanation is addressed to a recipient to support his understanding and that demands to connect the explanandum on familiar and accepted knowledge and has therefore an epistemic component.

About thirty years later van Fraassen integrates Ronald Giere's ansatz of scientific perspectivism in his view, where scientific representation is determined by more than a dual relationship (van Fraassen 2008, p. 28).

"[I]f we think of representation as a relationship, it should be a relationship with more than two components. One component should be the agents, the scientists who do the representing. Because scientists are *intentional* agents with goals and purposes, I propose explicitly to provide a space for purposes in my understanding of representational practices in science. So we are looking at a relationship with roughly the following form: S uses X to represent W for purposes P." (Giere 2006, p. 60)

Whereas Giere as realist introduces perspectivism as bridge between realistic and constructivistic[7] views accepting that 'there is a valid point to the constructivist critique of science (Giere 2006, p. 3) van Fraassen reinforces his view that science has the task (only) to represent observable phenomena in an adequate way (van Fraassen 2008, p. 86). Both, Giere and van Fraassen, agree that measurements in science have a perspectival form (I will come back to this in chapter 5), however, disagree in the consequences on theorizing in science. In van Fraassen's view the representation of a theoretical model is a larger structure used for revealing the structure of observable phenomena embedded in this structure on the one side and for practical appliances on the other (van Fraassen 2008, p. 87f), whereas Giere accepts that scientific claims contain 'some degree of contingency' and sees like Kuhn 'a role for history' to broaden theoretical perspectives (Giere 2006, p. 93f). I suppose that van Fraassen's position is too restrictive particularly as historical studies can also reveal the development of the structure of theoretical models e.g., of used analogies and the reasons for that using.

Nevertheless, it is informative to follow further van Fraassen's discussion of his 'empiricist structuralism' as he calls the refinement of his former 'constructive empiricism' regarding the representation of structure in theoretical models. At first, van Fraassen hold on to the aim of science 'to provide empirically adequate theories about what the world is like' (van Fraassen 2008, p. 87). However, he admits as mentioned 'it may be practically as well as theoretically useful to think of the phenomena as embedded in a larger – and largely unobservable – structure' (van Fraassen 2008, loc. cit.). This larger structure is as van Fraassen argues following proponents around the beginning of the last century like Maxwell, Boltzmann or Hertz only a picture (van Fraassen 2008, p. 197), i.e., 'science abstracts, it presents us with the structural skeleton of nature only' (van Fraassen 2008, p. 205). The problem with this position is that a structure as scientific representation is mathematical and according to Weyl in mathematics no distinction cuts across structural sameness (van Fraassen 2008, p. 208). This statement means not that two isomorphic structures must be indistinguishable in all respects but in respect to the supposed structure. An example from Weyl cited by van Fraassen can illustrate the problem, the representation of the colour space of human colour vision is a region in the projective plane, i.e., they have the same structure. Mathematically they are indistinguishable, but we know that they are different[8] (van Fraassen 2008, p. 209). The situation reminds as van Fraassen does on the story in philosophy of mind regarding Mary, the colours scientist who knows all about colour in scientific terms but has never seen one (van Fraassen 2008, p. 210ff). After some time, she could leave the totally monochrome room she was living till then and saw e.g., a red rose. Only then after this impression she could locate this 'red' in the colour space i.e., the structural knowledge of the colour space must be supplemented by the impression of colours to distinguish it from other isomorphic structures. Van Fraassen's solution to this problem goes exactly this way, i.e., he gives examples for the coordination between e.g., a Euclidean square and a tabletop with sides of equal length and right angles between them (van Fraassen 2008, p. 249). In fact, the relation is somewhat more complicated, the four sides and the angles of the tabletop must be measured if they

[7] Constructivists advocated mostly in sociology, literature and cultural sciences assume that scientific knowledge is some sort of 'social construct' (Giere 2006, p. 2).
[8] In fact, if we colour the plane in the corresponding way also the representation is different, but mathematics knows no colours.

are in fact equal. Then you have a data model, which is to compare with the theoretical model of the Euclidean square and its properties. However, this comparison and the acceptance of the result '(A) that the theory is adequate to the phenomena and the claim (B) that it is adequate to the phenomena as represented, i.e. *as represented by us*, […] *is a pragmatic tautology'* (emphasis in original, van Fraassen 2008, p. 259) depending from the perspectival indexical 'us'. After all I ask, what is the described relation other than an analogy? The tabletop and a Euclidean square have obviously an analogous structure where the Euclidean square is a more abstract realization than the data model and the tabletop. I will discuss this and other functions of analogies more in detail in chapter 4 and the following.

2.4 Philosophy of Science and Models

Also R.I.G. Hughes deals with the subject of models and representation. For Hughes models are in the tradition of Hertz, Duhem, the semantic view of theories and therefore also van Fraassen 'constructs stand[ing] in a particular relation to the world' (Hughes 2010, p. 150). This relation is described by the kind of representation, where Hughes distinguishes several kinds e.g., material, and mathematical representations depending on the model, which represents the phenomena. However, also in the form of modelling Hughes states differences e.g., from the scale over incompatibility even to contradiction. Regarding scale he mentions the viscosity of fluids, which is not like the temperature or pressure of gases reducible to the statistical motion of molecules and atoms but is explained by some kind of friction (Hughes 2010, p. 128). Incompatibility we find in the different approaches to explain the behaviour of atomic nuclei (Hughes 2010, p. 140f, see also chapter 3.1, p. 37 and footnote 10). Contradiction describes Hughes for Einstein's theory of the Brownian Motion from 1905 where molecules in a solution and particles in a suspension were handled like the particles in an ideal gas but the liquid as continuous and homogeneous medium despite that the liquid also consists of molecules (Hughes 2010, p. 148). Einstein pointed for justification of this different handling of the molecules later to the large difference in size between the two types of these molecules of the solution respective suspension on the one side and of the liquid on the other (Hughes 2010, p. 149).

As main structure for the use of models in physics, Hughes suggests the so-called DDI account where the first D stands for denotation, the second for demonstration and the I for interpretation (Hughes 2010, p. 153). However, he looks not at it as recipe for physical theorizing:

"I am not arguing that denotation, demonstration, and interpretation constitute a set of acts individually necessary and jointly sufficient for an act of theoretical representation to take place. I am making the more modest suggestion that, if we examine a theoretical model with these three activities in mind, we achieve some insight into the kind of representation that it provides. Furthermore, we shall rarely be led to assert things that are false." (Hughes 2010, p. 155)

Denotation is following a statement of Nelson Goodman cited by Hughes as 'the core of representation and is independent of resemblance' (Hughes 2010, p. 156). Denotation is used to coordinate the elements of the scientific model with elements of the considered subject, a symbol for that element e.g., t for time or a horizontal regularly subdivided line (a time axis in a coordinate frame, horizontally because time is the independent axis) for the flowing of time. As we see from this last example, resemblance is not necessary because the line has nothing to do with the flowing of time, it is only a symbol. Besides that, we learn from this example how familiar we are already with such conventions as a horizontal axis for the independent one.

Demonstration is intended to show the internal dynamic of the model equally for mathematical or material models. In mathematical models the internal dynamic is realized by a time-dependent differential equation, algebra or geometry. Especially geometrical theorems in the seventeenth century were 'demonstrated' which gave this act the name (Hughes 2010, p. 159f). In material models, the dynamic is mapped on natural processes as the ripples in a tank for the demonstration of light waves. A third more modern type of modelling is the computer simulation. 'Each of them acts as an epistemic engine, generating answers to questions we put to them' (Hughes 2010, p. 159).

Interpretation is the back translation from the model to the subject that, what makes the predictions of the model testable and proves if the theoretical model is empirical adequate. Therefore, 'the requirement of empirical adequacy is thus the requirement that interpretation is the inverse of notation' (Hughes 2010, p. 161). Meant is with this statement that e.g., the ripples in a tank of the former paragraph must be interpreted as an analogous behaviour of light.

Important on this account is that it can be nested, i.e., one can translate the denotations of one model into the denotations of a further, more familiar model, where then the demonstration takes place and translate this in the end back into the interpretation of the first model and then again to the observable phenomena (see Hughes 2010, p. 163). This nesting is nothing other than an analogy because during this nesting one domain of meaning is transposed to and by this compared with another more familiar one i.e., comparation is an analogical form.

Also, in another point I will follow Hughes. His intention is not, as several times annotated before, the reconstruction of scientific theories 'but rather to show how it [i.e., the theoretical practice of physics] is exhibited from the recent and not so recent history of physics, and to draw philosophical lessons from doing so' (Hughes 2010, p. 5). My philosophical lesson is mainly to show how analogies are used in theoretical and experimental practices of more recent physics, i.e., the development and establishment of the concept of quarks. Like Hughes

'I shall examine the theoretical practices of physics as they appear in physics journals, treatises, monographs, and the like. I treat these publications as texts: and thereby cast the philosopher of science in the role of critic.' (Hughes 2010, loc. cit.)

Before doing that, I will make some remarks on recent publications fortunately directed not on the reconstruction of scientific theories, i.e., the context of justification, but on the development of scientific hypotheses, i.e., the context of discovery. They are published in a volume on building theories (Danks & Ippoliti 2018), where the editors in the preface refer to several papers in the book arguing that the context of discovery and justification are not separable and that 'theory building is guided by defensible principles and practices that do not guarantee success, but are also not completely arbitrary' (Danks & Ippoliti 2018, p. VI). This is the view I argue for, too.

Emiliano Ippoliti supposes that the processes of generating and justification of hypotheses 'are distinguished mostly for the sake of conceptual simplicity and for educational purposes' (Ippoliti

2018a, p. 3) and refers to analogies as main counter-example because analogies are rational procedures. Analogies participate as we will see as well in the generation of hypotheses as in the justification of the theories based on these hypotheses and undermine the traditional distinction (by Reichenbach, see p. 21) between the not rational reasonable process of generating hypotheses and the rational process of the justification of theories. Ippoliti then reasons, why whether the justification of hypotheses nor their generation can be pure logical or probabilistic procedures. In the end he suggests a 'heuristic view' on theory building, where theory building is problem solving using rational methods (Ippoliti 2018a, p. 12), e.g., by using induction, analogy or metaphor. In the outline of a theory of 'heuristic logic' Ippoliti indicates analogy and induction as building blocks of the heuristic method, from which others can be derived (Ippoliti 2018b, p. 199f).

Margaret Morrison criticises that the syntactic and semantic view on theories does not really represent the structure, how theories are constructed in science and gives two examples from high energy physics and one interdisciplinary example transferring methods from physics to biology (Morrison 2018, p. 21). In high energy physics she distinguishes a top-down strategy using symmetries and a bottom-up strategy embedding models in a wider framework, where the last example works with analogies. However, these examples illustrate in her eyes only applied strategies but no blueprints.

In the same book, David Danks investigates the possibilities of scientific discovery by intertheoretical reduction (Danks 2018, p. 45). There he argues that discoveries by intertheoretical reduction are possible via constraints not only with theories on different levels of phenomena or structures in both directions (bottom-up or top-down) of the same domain but also with more different theories with a common constraint e.g., causal structures (Danks 2018, p. 53). After all, Danks classifies discoveries by constraint more in phases of normal than revolutionary science (Danks 2018, p. 59) and analogies 'as a more speculative type' of discovery (Danks 2018, p. 45). The way I see analogies is that analogies are not so speculative as Danks supposes in the case of discovery by constraints because in almost all (or really all?) cases as I will show also an existing theory about similar phenomena is the starting point.

Another ansatz is advocated by Carlo Cellucci, who subsumes in his analytic view following the approach of Aristotle and Plato theory building as problem solving (Cellucci 2018, p. 63). In this view 'theory building consists in starting from problems, arriving at hypotheses and deducing solution to problems from them' (Cellucci 2018, p. 68). Hypotheses however are based on inductions, analogies or metaphors as he concedes. As examples Cellucci mentions Kepler as based on analogy, Newton as based on induction, Darwin as based on induction and analogy and Bohr as based on a metaphor.

Only two authors in this book take the focus on analogies directly. One is Monica Uglaglia, who compares the application of mathematics in mathematical physics and by Aristotle as reservoir of analogies for physics and other uses (Uglaglia 2018, p. 120). The reason for Aristotle's view contrary to mathematical physics, where mathematics is used to describe and predict the behaviour of objects, is that problem solving (where mathematics can be used) for Aristotle is not scientific, only a deductive scheme derived from first principles, definitions or derived theorems is scientific (Uglaglia 2018, p. 140). In a similar way, A. Ulazia investigates the mental i.e., von-verbal schema of an eddy or vortex as a special analogical model for e.g., the development of fluid dynamics by Bernoulli and of the dynamics and resistance of flow by Reynolds (Ulazia 2018, p. 145). Whereas analogies in Uglaglia's paper are used by Aristotle similar to Hempel or Popper only outside a justified reconstruction of a theory, they are used in Ulazia's paper in the special form of a vortex as model for the development of a theory. This last is the way I will go in the next chapters however in a more detailed form. But before doing that, I will give some conclusion remarks followed by an overview as announced in the introduction over the development of physics from the theory of the atomic nucleus to the theory of quarks.

2.5 Summary

In the investigations in philosophy of science theory development was neglected for a long time because the context of discovery was seen not as a rational process. Only recently this situation has changed. However, also nowadays in my opinion analogies have not found the attention they earn in the process of the scientific enterprise. The reason is as we will see that its role in the building of concepts also outside of science was not considered in an enough way. It is remarkable but not really astonishing that a physicist, Ernst Mach, more than hundred years ago in some basic insights despite a utopian aim was nearer on a comprehensive conception of the scientific research process in physics including the use of analogies than most philosophers later because he could look on it from its inner and outer side, today one would say in an interdisciplinary manner. Then, after the philosophical renaissance of an historical view on the change of theories by Kuhn also the different views of scientists on their research object came into focus such that a unique relation between the assertions of a theory and the observed phenomena was no more obvious. One solution to this problem was to consider the assertions of the theory as a formal description of a structure and its application on specific phenomena as a realized instantiation or model of this structure. Here one can suspect some evidence of analogies if the same structure is applicable for phenomena in different domains but in my view the direction of association for theory development must be the reverse, first the detection of similarities between different models describing the phenomena, then the generalization in a common structure, i.e., first the analogy, then the development of a theory. How analogies work in this way will be demonstrated more in an overview of a period from Heisenberg's theory of the atomic nucleus to the prediction of quarks in the following chapter and in much more detail after a closer look on types and function of analogies in chapter 4. There we will see that the notion of 'observability' plays a crucial role as already indicated in the discussion of van Fraassen's view on observability (see p. 27f).

3 Analogies in Theoretical Practices

Some remarks on the whole story of the development of knowledge on the constitution of matter should be given here first despite the limitation in this chapter on the steps from the atomic nucleus to the development of quarks. The reason for these remarks is that, in my view development in the sciences is not planned or plannable but only in historical review understandable. After philosophical speculations about atoms already in Ancient Greece, the first understanding regarding them began with the law of multiple proportions in chemical reactions by Dalton in 1809 and 1810, which states that the relation of the quantities of the constituents of a compound are always whole numbered.

This whole numbered relation suggests that matter is not infinitely divisible and only specific numbers of atoms could combine. This different chemical behaviour of the special atoms of a substance enabled later the construction of the periodic table of elements. Also, statistical thermodynamics, the Brownian motion, and Faraday's laws of electrolysis with its elementary charge quantum pointed in that direction.

First insights into the structure of atoms itself gave the irradiation with cathode rays (i.e., electrons) of different energy, showing that atoms have only a small positive charged core within a negative shell, later confirmed by the scattering experiments of Rutherford.

However, a deeper theoretical view into an atom was possible only using results of spectroscopy: the lines found by Fraunhofer within the spectrum of light, absorbed or emitted, respectively, and somewhat later the demonstration of their structure by the formula of Balmer regarding the spectrum of hydrogen. The formula of Balmer and Planck's quantum of action in blackbody radiation enabled Bohr in 1913 to explain the stability and the spectrum of hydrogen by a system of stationary states of definite energies. Then in 1916, Sommerfeld found two further quantum conditions which excluded otherwise classical also possible orbits. The predicted spatial quantization was observed by Stern and Gerlach in 1922. In 1925, Pauli then could explain why the length of a period in the periodic table is double the number of the square of their principal quantum number, i.e., the quantum number introduced by Bohr. Pauli's explanation postulated a 'particular two-valuedness of the quantum-theoretic properties of the electron, which cannot be described from a classical point of view' (Brandt 2009, p. 146). This property was later called the 'spin' of electrons or other particles, now termed fermions. Associated is this property with the condition that 'never be two or more equivalent electrons in the atom for which in strong fields the values of all quantum numbers […] coincide' (Brandt 2009, p. 147). This last part is Pauli's famous 'exclusion principle', valid for all fermions like the electron.

In the same year, Heisenberg published his new quantum mechanics in several papers, in part together with Born and Jordan: a reinterpretation based on matrix calculation of the old quantum theory developed by Bohr and Sommerfeld. The reinterpretation regards the replacement of the position coordinate by an 'ensemble of quantities' (Brandt 2009, p. 154), also valid for smaller differences between stationary states, which corresponds to the emitted or absorbed frequencies of light observed in experiments. One year later, Schrödinger presented his quantum wave mechanics which was demonstrated by him as equivalent to Heisenberg's quantum mechanics and interpreted by Born as probability distributions also in 1926. However, that were not all relevant events of this year because Fermi and independent from him Dirac (and Heisenberg, too) developed a new statistic, called later after the two first mentioned scientists Fermi-Dirac statistics, which is valid for fermions. This statistic combines the Bose-Einstein statistics, developed by Bose and published with the help from Einstein in 1924, with the exclusion principle of Pauli. The Bose-Einstein statistic assumes in difference to the classical statistics, the only former known, that all particles are indistinguishable and therefore invariant against the exchange of them. The Bose-Einstein statistic is valid for bosons like e.g., the photon.

Again, one year later Heisenberg introduced his uncertainty relation, which states that position and momentum of a particle could not be measured at the same time with more precision than in the order of Planck's quantum of action. Together with the probability interpretation of Born and Bohr's complementary principle, which assumes that the wave and particle properties of a particle are complementary to each other, the uncertainty relation comprise the so-called Copenhagen interpretation of quantum mechanics. This is the state where my overview of physical development presented in the following will begin.

3.1 The Theory of the "Strong Interaction" and its Problems

According to Berger (see citation 1 in chapter 1, p. 9) the aim of physics is to explain the constitution of matter. After the great success of explaining the behaviour and spectra of atoms with quantum mechanics and its more sophisticated derivation quantum electrodynamics (QED) the interest focussed on the atomic core. However, some problems were to be solved to explain the constitution of the atomic core. It was known that (1) almost the whole mass is concentrated in the atomic core, that (2) the number of electrons in the atomic shell is equal to the position of the element in the periodic table, however, (3) the weight (mass) of atoms nearly doubles with this number, and that (4) the atomic core was very small in relation to the size of the whole atom (Brandt 2009, p. 66ff). Therefore, for the first two reasons Rutherford (1920, Bethge & Schröder, p. 282) suggested that this atomic core, the nucleus, was put together out of multiple positive charged cores of hydrogen, 'proton' called by Rutherford (Bethge & Schröder 1986, p. 282).

Following from Rutherford's speculation two problematic questions arose for the physicists of this time: (1) what holds the positive charges of the nucleus together, and (2) why do most nuclei have about the double of the mass of the proton? A possible answer to the first question was that there exists a much stronger force than that of the electromagnetic interaction, i.e., the so-called Coulomb force. Various scattering experiments showed that this strong force must decrease extremely fast outside the nucleus. Regarding the second question, already in 1920 Rutherford supposed that the large nucleonic mass was caused by heavy uncharged particles with about the same mass as protons. An alternative suggestion supposed that the nucleus consisted of protons and electrons. But this idea must be dropped for a combination of reasons: the spin of protons and electrons is ½ but the spin of the nucleus of e.g., the seventh element nitrogen is whole numbered, and no nucleus with whole numbered spin can be constructed with both an even number of protons (in this case 14) and an odd number of electrons (Bethge & Schröder 1986, p. 283).

Fortunately, the heavy uncharged particle postulated by Rutherford was found by Chadwick in 1932 by experiments with a cloud chamber. He called it 'neutron' (Brandt 2009, p. 210). The neutron not only solved the problem with the mass because it has nearly the same mass as a proton, but the problem with the even spin of the above-mentioned nitrogen nucleus, too, because the neutron has also half spin, like the proton.

For the account of the strong interaction between protons the existence of the neutron helps, given it can be considered as an elementary particle and the decay of neutrons into a proton and electron (and a neutrino as postulated by Pauli some years before but not published) can be put aside temporarily. In case that the two conditions are fulfilled, quantum mechanics can be applied as

Heisenberg suggested in 1932. Heisenberg investigated the nucleus of the deuterium, the lightest element besides hydrogen, consisting of one proton and one neutron in the nucleus in *analogy* [9]to an ionized hydrogen molecule, i.e., two protons bound by one electron. Quantum mechanics tells for this molecule that the electron must be found with the highest probability between the two protons, compensating their repelling force. Such a compensating force based on the symmetry properties of a quantum mechanical wave function was called “exchange force” (Brandt 2009, p. 224). If one takes proton and neutron for the strong interaction as two states of the same particle, i.e., consider the different charge as the different state and ignore the slightly different mass, then the deuteron (the nucleus of deuterium or heavy hydrogen) obeys the same symmetry conditions as the ionized hydrogen molecule. However, there is a difference: The exchanged particle must be a particle like an electron yet without spin because both proton and neutron have half spin, they are fermions. With the exchange of this hypothetical particle the state of the proton and the neutron would change simultaneously from one to the other, but the spin must not.

However, Heisenberg takes the idea of particle exchange not really as a serious one, “he decided to describe the two possible states (proton and neutron) of the nucleon with the same formalism that Pauli had developed for the two states of the electron spin” (Brandt 2009, p. 225). Because of that *analogy* to the electron spin this description of particle exchange was called ‘isospin’ formalism. The charge in this formalism is given by $Q = ½ + I_3$, if $I_3 = ½$ is attributed to the proton and $I_3 = - ½$ to the neutron. In addition to the stability of the deuteron, this model could explain altogether three aspects: (1) that “the nuclear binding energy is proportional to the number A of nucleons in the nucleus”. (2) “for light nuclei, the number Z of protons and the number N = A – Z of neutrons are approximately equal”. (3) for heavier nuclei, the electrostatic repulsion of the protons forces the nucleus to contain more neutrons to compensate this repulsion (Brandt 2009, p. 225). Nevertheless, this theory does not sufficiently explain the strong interaction because it uses only symmetry conditions and does not really introduce a new force with a potential like that of the electrostatic (Coulomb) field or the gravitational field.

A step in a more promising direction (Brandt 2009, p. 250) was the attempt of Yukawa in 1934. Yukawa suggested a potential in *analogy* to the electrostatic field, however, in a modified form as solution of a modified Laplace equation $(\Delta - \lambda^2)\, U(r) = 0$ instead of $\Delta\, U(r) = 0$. Yukawa’s potential $U(r) \sim e^{-\lambda r}/r$ has a limited range depending on λ as needed and corresponds to a particle with mass $m = \lambda\, h/c$ using the method of second quantization in quantum mechanics. However, his paper, published 1935, received no attention until 1937, when a study by Anderson and Neddermeyer reported the detection of tracks of a new particle in photographs of a magnetic cloud chamber. The magnetic cloud chamber was exposed to cosmic radiation at the height of 4300 m. The found particle showed a mass between that of an electron and that of a proton, matching the prediction of Yukawa (Brandt 2009, p. 251). In the end this particle was called ‘muon’ because, as recognized later, this particle interacts not in a strong way with nucleons and nuclei, rather it behaves like a heavy electron. The particle with the ‘right’ behaviour proposed by Yukawa was found in 1947 by Occhialini and Powell with a slightly larger mass. That new particle was called π-meson or pion (Brandt 2009, p. 301). In

[9] The use of analogies or similarities will be emphasized because the play an important role in the following.

fact, pions were found in three charge states, positive, negative, and neutral as predicted by the theory because the exchange in a nucleus must occur from proton to neutron, from neutron to proton, and between two protons or neutrons, respectively. Nevertheless, properties of the nucleus remained which Yukawa's potential could not describe.

One of these properties is called by the term 'isotope': An isotope is an element (e.g., the noble gas neon), which exists with different nearly whole numbered masses in relation to hydrogen. Already in 1912 J.J. Thomson found that not only do radioactive elements exist in different isotopes but stable elements as well. Aston, too, found in the following years that almost all elements have isotopes (Brandt 2009, p. 91). Therefore, Aston refined the whole-number rule with the condition that it was even better fulfilled, if 1/16 of the atomic mass of oxygen was used as unit rather than the atomic mass of hydrogen. Lenz pointed out in 1920 that the difference between these two masses (hydrogen has about 0.8 per cent more mass) could be interpreted with Einstein's formula $E = mc^2$ as binding energy, which becomes free if a proton was bound in an atomic core (Brandt 2009, p. 91f).

Remarkably, the distribution of isotopes in the periodic table shows some regularity: In some regions with specific proton and/or neutron numbers more isotopes occur than usual. The structure of this regularity was explained independently in 1949 by Goeppert-Mayer on the one side and Haxel, Jensen and Suess on the other with a shell model of the atomic core consisting of protons and neutrons. This shell model is *similar* to that by Pauli of the atomic hull built up by electrons, which was formulated in 1925 (Brandt 2009, p. 317ff). However, there were two decisive differences between these two models the nucleus consists of two types of particles, which build up separate shell structures, and in heavier elements a strong spin-orbit coupling appears. Spin-orbit coupling means that the interaction between the orbit and the spin of the particle is quite large, thus the orbit and the spin of every particle in that shell are not separated any more. This resulted in the effect that exclusively the sum of both is effective (Finkelnburg 1967, p. 103). This is true for nuclei with higher masses. In contrast, the sum of the orbit and the spin is irrelevant in nuclei with lower mass or in the atomic hulls of electrons, there the interaction between the orbits of the electrons on the one side and of their spin (of the outset shell) on the other side are large and the sums of the orbits and the sum of the spins combine separately. Obviously, this model with shells is another model of the nucleus than that of Yukawa.

Some other models for the interaction of the constituents of the atomic core are not mentioned here. The reason for omitting certain models in this context is that the strong interaction is extraordinarily complex[10], so it cannot be described with one potential or one model.

---

[10] The complexity of the strong interaction in the atomic core is demonstrated by the following probable structure (Mayer-Kuckuk 1981, p.170):

(1) There is a central force.
(2) There is a central spin-dependent force.
(3) Forces in the core are in particularly good approximation charge independent.
(4) There is a non-central force (a quadrupole moment not mentioned before).

To find a unique description of such a kind of strong interaction another method / approach has been developed, the so-called S-matrix or scattering matrix. However, this method was not very suitable for bounded states but only for scattering processes (Mandl & Shaw 1993, p. 102). Therefore, the interest shifted gradually from nuclear physics to particle physics, supported by the detection of many new particles which are unknown for the atomic nucleus.

Scattering experiments are the main source of getting information within particle physics, together with the observation of the decay of unstable particles (Berger 2006, p. 43). Both sources of information describe the statistical behaviour of the outcome of experiments regarding observations in relation to a known input. This 'known input' means the relation of the number of input particles to specific numbers of output particles. In scattering experiments, the results were represented as differential or total cross section (Berger 2006, p. 44f). The notion of 'cross section' dates to the experiments of Rutherford in 1911. Then, he explained the great scattering angles of α-particles by central charges in an atom and computed theoretically the probability when a projectile particle interacts with a target particle. The notion 'cross section' is the modern word for what Rutherford called 'probability of scattering' (Brandt 2015, p. 67). However, the modern notion has some more aspects in its meaning. One aspect is for example the change of the direction of particles, another is the creation of new particles. Both will be described shortly in the following:

The differential cross section describes the relationship between the scattering angle on the one hand and the probability depending on the conditions of the experiment on the other hand. Originally, at the time of Rutherford, three conditions were considered as relevant: (1) the energy of the incoming (α-) particles, (2) the number of elementary charges of the target atom (e.g., gold), and (3) the thickness of the target (Brandt 2015, p. 68). Rutherford supposed from the large scattering angles observed that the atomic core was particularly small, i.e., point-like. The source of the projectile particles was a radioactive natural element like radium emanating α-particles, the detector in the 1911 experiment a scintillation screen with a coating made of zinc-sulphide. From the data that Rutherford got out of this arrangement, the probability of scattering for every small region of angles could be computed using Rutherford's scattering formula.

This experiment of Rutherford's team is the ancestor of all modern scattering experiments. The modern experiments are of course much more sophisticated with more efficient detectors and sources. Moreover, a multitude of detectors can be arranged around the collision point and for more efficient use of the needed energy, colliding beams were used. The different detectors now allow not only to register the particles emitted in a special direction, the differential cross section, or particles emitted in all directions but also particles of different kinds of that participating in the collision, i.e., particles newly created in the collision. In Rutherford's case, i.e., where the particles in the output are the same particles as in the input, the scatterings are called 'elastic' scatterings. Those with the creation of new particles are called 'inelastic scatterings'. The evaluation of all possible reactions results in a collision of defined input particles, which summed up over a representative sample of

---

The strong interaction is much stronger than the electromagnetic interaction: The strength is in the order of one (thus the notion "strong") with the consequence that perturbation calculus cannot be used (Bethge & Schröder 1986, p. 285).

collisions and containing both, elastic and all inelastic scatterings; such an evaluation is called the 'total cross section'. Expressed as a probability the total cross section has of course a probability of 1, each outcome of a differential cross section a probability smaller than 1.

Because of the possibility of elastic and inelastic scattering, different detectors are needed to register all the different particles appearing in the experiments. Most of the new created particles have no long lifetime. They decay in the end in stable particles, perhaps over several stages with intermediate short-lived particles or 'resonances', i.e., systems of orbiting elementary particles. As mentioned above, in particle decays also probabilities occur. These probabilities can be expressed as differential and total cross sections, too, depending on possible 'decay channels', i.e., combinations of particles which obey some observed conservation rules or symmetries. Symmetries and conservation rules in physics are like two sides of the same coin: A conservation rule states that a specific property does not change under special transformations, i.e., a specific property is like the symmetry of e.g., a square which does not change its appearance under a rotation of ninety degrees or multiples of these rotations. If all conservation rules or symmetries were found, then a consistent picture of all scattering and decay processes of the strong interaction could be given without any need of an underlying theory which explains why the observed symmetries do exist. Therefore, the question at that time from the point of view of physics reads, what are the symmetries giving an order to the found particles and resonances? In other words: How to classify the detected new particles?

3.2 How to Classify the New Detected Particles?

In contrast to the relatively few other kinds of particles as many as fifteen particles that participate at the strong interaction (called hadrons), eleven anti-particles as well as 33 resonances were known in 1965 (Finkelnburg 1971, p. 329f). The few other kinds of particles are the then known four leptons, i.e., electrons, muons, their corresponding neutrinos with additionally their anti-particles on the one hand and photons on the other hand. Within the hadrons the physicists classified six kinds of particles as mesons like the pions and nine as baryons like the protons or neutrons. In 1978 the total number grew up to 137 baryons and 64 mesons (Lohrmann 1981, p. 118). What are the symmetries to classify (order) such a large number of altogether 201 hadrons?

Five symmetries were in question (Finkelnburg 1971, p. 331f). Within these symmetries charge, spin and parity, i.e., the (positive or negative) invariance against space inversion, were already well known. The two new symmetries baryon number and strangeness will be explained below. However, the values of these symmetries for all known particles and especially their resonances remained still doubtful.

A first attempt to order the number of particles by uniting two into one with two states was already made in 1932 by Heisenberg. He described the proton and the neutron as two states of one nucleon, which allows to express the charge by the new quantum number of „isospin "(= I) using the formula $Q = ½ + I_3$ and assigning $I_3 = ½$ to the proton and $I_3 = -½$ to the neutron (see p. 36). They form an isospin-doublet like the splitting in doublets in atomic spectroscopy, i.e., corresponding to the two spin orientations in alkali atoms obeying the multiplicity rule: multiplicity of the spectral lines is equal to the double of the total spin plus one $m = 2 S + 1$ (see Finkelnburg 1971, p. 105). With these definitions of the isospin, it is obvious to use the same (*analogous*) multiplicity rule to the isospins of the three pions and assign to the pions π^+, π^0 and π^- the values +1, 0 and -1 forming the three states of an iso-triplet. However, for a correct calculation of the charge of the pions this assignment requires that the charge is equal to the third component of the isospin $Q = I_3$. To get a satisfactory formulation in combination with baryons Gell-Mann suggested in 1955 to introduce a new quantum number, the so-called baryon number B (Brandt 2009, p. 343). Baryons should have a baryon number of 1 and mesons of 0 and anti-particles always the opposite sign for both the baryon number and the isospin. Then, Heisenberg's formula for the calculation of the charge reads $Q = B/2 + I_3$. In this way, the number of five particles is reduced to two with two or three states, respectively. Additionally, because the π^+ is the anti-particle of π^- and π^0 is its own, this equation is true for the sum of particles and of anti-particles, too.

Already in 1947, unknown particles were detected by Rochester and Butler on photographs of events in their bubble chamber. These particles did not match to the existing picture of nucleons and pions, taken as a proof of Yukawa's theory for strong interaction (Brandt 2009, p. 307). Moreover, nowadays these particles were known as a neutral and a positive charged K meson. The K mesons must be till then unknown particles because they decayed into pions (π mesons), and their masses could be calculated to 'at least twice or thrice as high as that of the π meson ', thus much less than a proton (Brandt 2009, p. 308). Two years later Powell's group (one of the discoverers of the pion) detected a slow-moving charged particle „observed in photographic emulsion exposed at an altitude of 3460 m on the *Jungfraujoch* in Switzerland" (Brandt 2009, p. 308), which came to rest in the emulsion, and decayed into three charged particles. One of the three particles was identified as a pion because of its nuclear reaction. The other two particles had to be pions, too, or muons regarding their ionization. Considering the assumption that all three particles were pions the mass of the primary particle could be estimated as 925 electron masses. This was the first observation of a decay of a charged K meson into three pions[11]. In the following years not only were also negative charged K mesons discovered but the existence of two quite different kinds of neutral particles, too. One of these neutral particles was the already known K meson, the other a particle much heavier than a neutron, later called Λ^0 (Brandt 2009, p. 311). In one case, the group around Powell detected even a cascade of decays (Brandt 2009, p. 308). The new discovered particle in this cascade was called later Ξ^-. Further particles found in 1953 by Anderson's and an Italian group (Brandt 2009, p. 311) were called Σ^+ and Σ^-. The last four particles Λ^0, Ξ^-, Σ^+ and Σ^- were summarized as hyperons, heavier than a neutron. All these particles behave very strangely insofar as the production process was one of strong interaction, but the lifetime of the particles was typical for weak interactions with a duration of about 10^{-10} seconds.

In 1952 Pais realized from the analysis of the experimental data that these strange particles like K mesons or hyperons can occur only in pairs in an "associated production" and decay only in weak interactions (Brandt 2009, p. 340). Pais formulated a rule which goes, as he later resumed, as follows:

> "Assign a number 0 to all 'old' particles (π, N, γ, leptons) and a number 1 to the new particles Λ, K (the other were not there yet). In any process, first add these numbers for the initial-

[11] This was one of the decays which lead later to the so called ‚fall of parity' in the weak interaction because the same particle should not decay in two or three particles alternatively because of parity reasons.

state particles, then for the final-state particles, these numbers being n_i and n_f. Then in all strong and electromagnetic processes n_i and n_f must be both even or both odd; in weak decays of the new particles one shall be odd, the other even. Thus $\pi^- + p \rightarrow \Lambda + \pi^0$ is strongly forbidden, $\pi^- + p \rightarrow \Lambda + K^0$ is strongly allowed. In general, new particles come in pairs, a mechanism later called 'associated production' (I do not know who coined this term). Electromagnetic decays like $K^0 \rightarrow 2\gamma$ are forbidden. $\Lambda \rightarrow p + \pi^-$, $K^0 \rightarrow 2\pi$, etc., proceed by weak interaction" (Pais 1986, cited in Brandt 2009, p. 340f).

This introduction of a new quantum number was confirmed two years later by the predicted associated production of a Λ^0 hyperon and a K^0 meson in a collision of a π^- meson with a proton. Selection rules like those proposed by Pais were considered beside some other options by several Japanese physicists like Nambu and Nishijima and others at nearly the same time, too (Brandt 2009, p. 341).

In 1953, Gell-Mann extended the above-mentioned isospin picture (*analogous*) to the new particles, postulating those strange particles do possess half-integer isospin with K mesons and integer isospin with hyperons, in contrast to known particles (Gell-Mann 1953, p. 833f). In addition, Gell-Mann postulated that K mesons form isospin-doublets K^+ and K^0 with $I_3 = +½, -½$, and K^- and $\overline{K}^0$ with $I_3 = -½, +½$. However, K^- was taken to be the anti-particle to K^+, and $\overline{K}^0$ to be the anti-particle to K^0. Therefore, particles and anti-particles do not belong anymore to the same isospin-doublet (Brandt 2009, p. 341). In this respect, the K mesons are more like (*analogous to*) nucleons, not like pions (Gell-Mann 1953, p. 834).

With the help of this hypothesis, Gell-Mann was now able to explain the behavior of the particles in interactions by three further postulates (Brandt 2009, p. 341):

(1) Strong interactions conserve isospin, $\Delta I = 0$.
(2) Electromagnetic interactions can change isospin by an integer value.
(3) Weak interactions do not respect isospin at all. (It was found that very frequently $|\Delta I| = ½$.)

The application of these assumptions on hyperons was published in a report together with Pais in 1955 (Brandt 2009, p. 342):

(1) „The Λ^0 has no charged partners. Its isospin assignments are $I = I_3 = 0$.
(2) Two somewhat heavier charged hyperons are given the names Σ^+ and Σ^- and form an isospin triplet with their neutral partner Σ^0 which is predicted to exist [found in 1956].
(3) The cascade particle, by now confirmed, is given the name Ξ^-. In order to explain its weak decay, $\Xi^- \rightarrow \Lambda^0 + \pi^-$, it is assigned a half-integer isospin, contrary to the original assumption that all strange fermions have integer isospin. A neutral partner, the Ξ^0 with $I_3 = ½$, is predicted to exist [found in 1959] and to form an isospin doublet with the Ξ^-, having $I_3 = -½$. "

Still in 1955, Gell-Mann introduced the above-mentioned baryon number as a new quantum number and extended the formula for the charge to $Q = I_3 + B/2 + S/2$. Often, the baryon number and the strangeness number were summed up to a quantity Y called hypercharge. At the same time, the same scheme was independently developed by the Japanese Kazuhiko Nishijima (Brandt 2009, p. 343). With this relationship for the charge, all hadrons could be ordered now in diagrams by mass vs.

charge with higher masses and rising positive strangeness for mesons, and rising negative strangeness for baryons, respectively.

However, in the early 1960s a lot of further strongly interacting particles were detected, particles like Σ^*, K^*, ρ or ω (Brandt 2009, p. 407f). Because of their noticeably short lifetime, they were first called resonances *analogous* to the behavior of an oscillating circuit before it became clear that they are particles decaying via the strong interaction. By evaluating the frequency of events in particular regions of the energy diagram the mass, the total angular momentum and, if it decays via the strong interaction, also the parity could be determined (Bethge & Schröder 1986, p. 303).

The total angular momentum and the parity helped to order the resonances: In *analogy* both to the orbital model of the atom and one of the models of the atomic nucleus, resonances were considered as a particle with constituents orbiting around their common center of mass. Therefore, the total angular momentum is composed out of the angular momentum of the orbit and the spins of the components, like the other two models. Likewise, the states were named (called in the same (*analogous*) way as in the spectroscopy of the atom) with (1) principal quantum number, (2) quantum number of the orbital angular momentum, (3) orientation number and (4) spin (Bethge & Schröder 1986, p. 305). However, it was not clear how to use these models, so different attempts were made to explain the resonances.

One further attempt to find an explanation for the resonances was to order them by plotting their masses against the states of the total angular momentum with the isospin as parameter in a diagram. Especially for mesons this approach turned out to be appropriate. Thus, it was possible to build up diagrams for constant isospins remarkably similar (*analogous*) to the term-scheme of hydrogen (Bethge & Schröder 1986, p. 310). For baryons, the other way around was undertaken. According to an idea of Regge (Bethge & Schröder 1986, p. 313), the total angular momentum has been plotted against the mass. It was found that for the first discovered particles like Σ, Δ or Λ, the higher resonances with higher mass and total angular momentum lie on a straight line with a distance of two units of angular momentum but the same parity. Particles separated by one unit are characterized alternating parities. This behavior of particles allows to predict further resonances. However, whether all resonances obey these regularities remained still unclear in 1986.

A third attempt from Chew and Frautschi in 1961 assumed substructures of nucleons because always hadrons were created again, in all interactions between hadrons[12] (Bethge & Schröder 1986, p. 313). Chew and Frautschi assumed that all hadrons must be treated as equal and all could be created from all. This attempt was known as ‘bootstrap model’ due to its closedness. Then, no particle is more elementary than another. However, the mathematical difficulties of this approach are so huge that it failed.

Nevertheless, the idea of substructures of the nucleons or more exactly of the hadrons together with Gell-Mann’s order of charge vs. mass with its common parameter of strangeness (see p. 41f) became

[12] However, this not the sole reason to assume substructures for the nucleons, one further reason is the complicated interaction between them in the atomic nucleus (Finkelnburg 1971, p. 336), another the anomalous magnetic moments of the proton and neutron (Brandt 2009, p. 345).

the starting point for the postulation of quarks. Gell-Mann's scheme of ordering particles possessing all the same spin and parity as parameter was the approach, which was accepted in the end, as will be shown in the next subchapter.

3.3 Postulating Quarks by Analogy as Solution

"In 1964 the wealth of hadrons was ordered by considering them to be composed of only three different types of particles, called quarks, and their antiparticles. The name was coined by Gell-Mann who was not alone in this search for order but only he contributed to all three stages of this search." (Brandt 2009, p. 408)

The first stage already described above was the introduction of the baryon number and together with Pais the strangeness number (see p. 41). The introduction of these numbers allowed to study hadrons with the same baryon number, spin, and parity as a set with similar regularities (Brandt 2009, p. 408).

The second stage was to use symmetries of the hadrons to order them. Symmetries are unchanged properties of (physical) systems observed if transformations were applied to them. The object to explain for Gell-Mann was the scheme of the eight baryons, the iso-doublet of the two nucleons, the iso-triplet of the three Σ particles, the iso-singlet of the Λ particle, and the iso-doublet of the two Ξ particles, all with spin ½ and positive (i.e., even) parity. Gell-Mann combined these baryons to a 'supermultiplet, degenerate in the limit of a certain symmetry but split into isotopic spin multiplets by a symmetry-breaking term' (Gell-Mann 1963, first version 1961, p. 12). This supermultiplet was ordered in two dimensions with strangeness number vs. the third component of the isospin, i.e., the isotopic spin in the citation above. Fixed parameters are, as mentioned, spin and parity.

The starting point for the 'unitary symmetry', as Gell-Mann called it, was a fictitious (*analogous*) model of leptons and their anti-particles. He chose leptons as a model because there were particles, a charged one and a neutral one, with spin ½ and negative parity, electron and neutrino, showing a symmetry like the isospin on the one side. On the other side a third lepton exists, the muon, with the same properties as the electron besides its greater mass. This mass difference of the leptons has no explanation still today. However, Gell-Mann assumed that an *analogous* 'equally mysterious mechanism' (Gell-Mann 1963, p. 19) is at work breaking the supersymmetry of the baryons into the iso-multiplets of the baryons mentioned before. In his lepton model he coupled it to the charge of the electron. Gell-Mann writes:

"If we now 'turn off' the μ-e mass difference, electromagnetism, and the weak interaction we are left with a physically vacuous theory of three exactly similar Dirac particles with no rest mass and no couplings. This empty model is ideal for our mathematical purposes, however, and is physically motivated by the analogy with the strongly interacting particles, because it is at the corresponding stage of total unitary symmetry that we shall introduce the basic baryon mass and the strong interactions of baryons and mesons." (Gell-Mann 1964, p. 19)

He then reduced the model temporarily on only two states, electron and neutrino. For two states the transformations are well known, they are the transformations of isospin or its underlying model of ordinary spin states of electrons. This model is mathematically a representation of the group SU(2), the special unitary group of two dimensions with the properties Tr $\sigma_i\sigma_j = 2\delta_{ij}$, $[\sigma_i, \sigma_j] = 2i\, e_{ijk}\, \sigma_k$, $\{\sigma_i, \sigma_j\} = 2\delta_{ij}\, \mathbf{1}$, i,j,k = 1,2,3[13]. Thereby, Gell-Mann considered two kinds of infinitesimal transformations, one where both states, electron, and neutrino, were multiplied by the same phase factor corresponding to the conservation of the states $\mathbf{1} + i\partial\Theta\, \mathbf{1}$. In the second kind these states were multiplied with each of three phases multiplied with one of the three Pauli isospin matrices, leaving the products of the phase factors invariant $\mathbf{1} + i\sum\partial\Theta_k\, \sigma_k/2$, k = 1,2,3. This second transformation corresponds to charge independence or isospin symmetry.

In a further step, the muon must be introduced in this model. To do this Gell-Mann extended the group from SU(2) to SU(3). SU(2) is a group with two, SU(3) a group with three independent elements. Further, SU(2) has four transformation matrices, i.e. the square of the two states consisting of the unit matrix and the three Pauli matrices; in the same way SU(3) must have nine transformation matrices. One of these matrices is the three-dimensional unit matrix; the other eight matrices must be traceless like the Pauli matrices and can be constructed in an *analogous* way. Gell-Mann called them λ1, …, λ8.

Table 2: Transformation Matrices of the Group SU(3)

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ 0 & 0 & \frac{-2}{\sqrt{3}} \end{pmatrix}$$

Analogous to the case of SU(2), these matrices have the properties Tr $\lambda_i\lambda_j = 2\delta_{ij}$, $[\lambda_i, \lambda_j] = 2i\, f_{ijk}\, \lambda_k$, $\{\lambda_i, \lambda_j\} = 4/3\, \delta_{ij}\, \mathbf{1} + 2\, d_{ijk}\, \lambda_k$, i,j,k = 1, …, 8. The f_{ijk} are real and totally antisymmetric like e_{ijk}, and the d_{ijk} are real and symmetric. The transformation of the second kind reads then $1 + i\sum\partial\Theta_k\, \lambda_k/2$, k = 1, …, 8.

The first three λ matrices correspond to the three σ matrices which show that the isospin is still conserved, ignoring the symmetry with the muon. New is the existence of a second diagonal matrix, namely λ_8. This matrix commutes with three first λ matrices and is therefore like strangeness or hypercharge, which distinguishes the singlet of the muon from the doublet of the electron and the neutrino.

"Now turning-on of the muon mass destroys the symmetry under λ_4, λ_5, λ_6, and λ_7 (i.e., under the 'strangeness-changing' components under the 'unitary spin') and leaves lepton number, 'isotopic

[13] Tr: trace; σ_i: Pauli spin matrices, i: imaginary unit, δ_{ij}: Kronecker symbol, [,]: commutator, e_{ijk}: total anti-symmetric tensor, { , }: anti-commutator, **1**: 2-dimensional unit matrix

spin', and 'strangeness' conserved. The electromagnetic interactions (along with the electron mass) then break the conservation of λ_1 and λ_2, leaving lepton number λ_3, and strangeness conserved. Finally, the weak interactions allow the strangeness to be changed (in muon decay) but continue to conserve the lepton number n_l and the electric charge $Q = e/2\ (\lambda_3, + \lambda_8/\sqrt{3} - 4/3\ n_l)$. […]

We see that the situation is just what is needed for the baryons and mesons." (Gell-Mann 1963, p.22)

We have seen above (p. 36) that the nucleons have isospin $I_3 = \pm½$. Therefore, their representation must correspond to the fictitious doublet electron and neutrino with $I_i = \sigma_i/2$. Other particles can be built up by superpositions of these particles, which transforms like the original doublet electron and neutrino. The same transformation rule is true for the anti-particles, too. Thus, the anti-particles of the fictitious leptons, positron and the (in this case) negative of the anti-neutrino, also form a doublet. The product representation in the model of the leptons and their anti-particles then departs into a singlet $(e^+e^- + \bar{\nu}\nu)/\sqrt{2}$ (= $\bar{e}e/\sqrt{2}$ for short) with isospin 0 and a triplet $e^+\nu = ½\ \bar{e}\ (\sigma_1 - \sigma_2)\ e$, $(e^+e^- - \bar{\nu}\nu)/\sqrt{2} = 1/\sqrt{2}\ \bar{e}\ \sigma_3\ e$, and $\nu e^- = ½\ \bar{e}\ (\sigma_1 + \sigma_2)\ e$ with isospin 1.

Further, following Gell-Mann, this model must be extended to three states and to the operator $\underline{F}$ with spin matrices $F_i = \lambda_i/2$. Here, however, arises one difficulty, because the transformation of the fictitious anti-leptons under the operator $\underline{F}$ is not equivalent to the transformation of the leptons. This inequality arises because 'when we go from the leptons to the anti-leptons the eigenvalues of the electric charge, the third component of the isospin, and the lepton number all change sign and thus the eigenvalues of F_8 change sign' (Gell-Mann 1963, p. 25). No similarity transformation changes this fact. Therefore, Gell-Mann called the representation for leptons 3 and the different representation for anti-leptons $\underline{3}$.

Then Gell-Mann introduced a new set of "particles" L_α with exactly the same transformation rules as the leptons under unitary spin and applied the same procedure as used for the isospin and the doublet e for the construction of the components of the operator $\underline{F}$. As a first result, Gell-Mann got a one-dimensional representation $\bar{L}e/\sqrt{3}$ with all components $F_i = 0$, called $\underline{1}$ by him.

Also, by *analogy* to the triplet with isospin 1 above, the components $\bar{L}\lambda_i e/\sqrt{2}$ with I = 1, …, 8 must be formed. The result is (~ is used as "transforms like") $\Sigma^+ \sim ½\ \bar{L}(\lambda_1 - i\lambda_2)e \sim D^+\nu$, $\Sigma^- \sim ½\ \bar{L}(\lambda_1 + i\lambda_2)e \sim D^0e^-$, $\Sigma^0 \sim 1/\sqrt{2}\ \bar{L}\ \lambda_3\ e \sim (D^0\nu - D^+e^-)/\sqrt{2}$, $p \sim ½\ \bar{L}(\lambda_4 - i\lambda_5)e \sim S^+\nu$, $n \sim ½\ \bar{L}(\lambda_6 - i\lambda_7)e \sim S^+\ e^-$, $\Xi^0 \sim ½\ \bar{L}(\lambda_6 + i\lambda_7)e \sim D^+\mu^-$, $\Xi^- \sim ½\ \bar{L}(\lambda_4 + i\lambda_5)e \sim D^0\mu^-$, $\Lambda \sim 1/\sqrt{2}\ \bar{L}\ \lambda_8\ e \sim (D^0\nu + D^+e^- - 2\ S^+\ \mu^-)/\sqrt{6}$. The last terms in each of the eight expressions illustrate the *analogy* of the doublet D^+, D^0 and the singlet S^+ to the anti-particles e^+, $\bar{\nu}$ and μ^+. The electric charge, the isospin and the hypercharge of the multiplets are as expected for the baryons und form a degenerated supermultiplet called $\underline{8}$. Supposing a perturbation by a mass difference like that between the electron and the muon decomposes the supermultiplet to the known multiplets of the baryons. However, a particle representing the singlet $\underline{1}$ was never found; it would have to be "very heavy and highly unstable" if it exists, Gell-Mann predicted (Gell-Mann 1963, p. 28).

In the same paper, Gell-Mann derived the multiplets of the pseudo scalar (isospin = 0) and vector mesons (isospin = 1) from the product representation $\underline{8}$x8 starting from the consideration that the mesons must have $\underline{F}$-invariant couplings. With a strong Yukawa coupling then the transformation of the bilinear form $\underline{8}$x8 must be computed. However, in the year 1964 Gell-Mann suggested a simpler algebraic way to construct the mesons and the baryons, too (Gell-Mann 1964). He did not consider the eight baryons as fundamental, but a triplet t formed by a doublet u, d and a singlet s *analogous* to his fictitious leptons. With these triplets and a basic neutral baryon b^0 similar as in the procedure shown above, singlets **1** and octets **8** of baryons by ($b^0 t\bar{t}$) or higher multiplets like **1**, **8**, **10**, $\overline{\mathbf{10}}$, **27** by ($b^0 t\bar{t}t\bar{t}$) can be built up. Even a still simpler construction could be reached If, contrary to the leptons, non-integral charges were allowed. The basic baryon could be abandoned if the following values were assigned to the constituents of the triplet: spin ½, baryon number 1/3 and charge u = 2/3, d = -1/3, s = -1/3, respectively. Gell-Mann called these constituents ‘quarks’. A baryon would then be formed by three quarks, mesons by a combination of a quark and an anti-quark. Again, these combinations could be extended by quark – anti-quark – pairs. However, in both cases only the lowest combinations could be observed. At a conference in 1964 Gell-Mann announced for the decuplet **10** a new particle, later called Ω^-, which was built up by three s quarks (Brandt 2009, p. 411). This is the sole ground state in a set otherwise only consisting of resonances.

Thus, Gell-Mann gets the following tables of baryons, mesons and vector mesons built up by quarks (only observed ones and for L = 0, modified from Bethge & Schröder 1986, p. 322ff and called by its modern names).

Table 3: Baryon Octet with Spin1/2 and Charge +1

Baryon	Spin	Parity	Charge	Isospin		Up	Down	Strangeness	Mass
octet	S	P	Q	I	I_3	u	d	s	MeV/c
p	1/2	+1	+1	1/2	+1/2	2	1	0	939
n	1/2	+1	0	1/2	-1/2	1	2	0	939
Λ	1/2	+1	0	0	0	1	1	1	1115
Σ^+	1/2	+1	+1	1	+1	2	0	1	1192
Σ^0	1/2	+1	0	1	0	1	1	1	1192
Σ^-	1/2	+1	-1	1	-1	0	2	1	1192
Ξ^0	1/2	+1	0	1/2	+1/2	1	0	2	1315
Ξ^-	1/2	+1	-1	1/2	-1/2	0	1	2	1315

Table 4: Baryon Decuplet with Spin 3/3 and Parity +1

Baryon	Spin	Parity	Charge	Isospin		Up	Down	Strangeness	Mass
decuplet	S	P	Q	I	I_3	u	d	s	MeV/c
Δ^{++}	3/2	+1	+2	3/2	+3/2	3	0	0	1232
Δ^+	3/2	+1	+1	3/2	+1/2	2	1	0	1232
Δ^0	3/2	+1	0	3/2	-1/2	1	2	0	1232
Δ^-	3/2	+1	-1	3/2	-3/2	0	3	0	1232
Σ^{*+}	3/2	+1	+1	1	+1	2	0	1	1385
Σ^{*0}	3/2	+1	0	1	0	1	1	1	1385
Σ^{*-}	3/2	+1	-1	1	-1	0	2	1	1385

Ξ^{*0}	3/2	+1	0	1/2	+1/2	1	0	2	1530
Ξ^{*-}	3/2	+1	-1	1/2	-1/2	0	1	2	1530
Ω^-	3/2	+1	-1	0	0	0	0	3	1672

Table 5: Meson octet with Spin 0 and Parity -1

meson	Spin	Parity	Charge	Isospin		Up	Down	Strangeness	Mass
octet	S	P	Q	I	I_3	u, $\bar{u}$	d, $\bar{d}$	s, $\bar{s}$	MeV/c
π^+	0	-1	+1	1	+1	1, 0	0, 1	0,0	135
π^0	0	-1	0	1	0	(-)1, 1	1, 1	0	135
π^-	0	-1	-1	1	-1	0, 1	1, 0	0	135
$\eta_8 \approx \eta$	0	-1	0	0	0	(+)1, 1	1, 1	0	549
K^+	0	-1	+1	1/2	+1/2	1, 0	0, 0	0, 1	498
K^0	0	-1	0	1/2	-1/2	0, 0	1, 0	0, 1	498
$\bar{K}^{0-}$	0	-1	0	1/2	+1/2	0, 0	0, 1	1, 0	498
K^-	0	-1	-1	1/2	-1/2	0, 1	0, 0	1, 0	498

Table 6: Meson Singlet with Spin 0 and Parity -1

Meson	Spin	Parity	Charge	Isospin		Up	Down	Strangeness	Mass
singlet	S	P	Q	I	I_3	u, $\bar{u}$	d, $\bar{d}$	s, $\bar{s}$	MeV/c
$\eta_1 \approx \eta'$	0	-1	0	0	0	1, 1	1, 1	1, 1	958

Table 7: Meson Octet with Spin 1 and Parity -1

meson	Spin	Parity	Charge	Isospin		Up	Down	Strangeness	Mass
octet	S	P	Q	I	I_3	u, $\bar{u}$	d, $\bar{d}$	s, $\bar{s}$	MeV/c
ρ^+	1	-1	+1	1	+1	1, 0	0, 1	0,0	770
ρ^0	1	-1	0	1	0	(-)1, 1	1, 1	0	770
ρ^-	1	-1	-1	1	-1	0, 1	1, 0	0	770
ω^0	1	-1	0	0	0	(+)1, 1	1, 1	0	783
K^{*+}	1	-1	+1	½	+1/2	1, 0	0, 0	0, 1	891
K^{*0}	1	-1	0	½	-1/2	0, 0	1, 0	0, 1	891
$\bar{K}^{*0-}$	1	-1	0	½	+1/2	0, 0	0, 1	1, 0	891
K^{*-}	1	-1	-1	½	-1/2	0, 1	0, 0	1, 0	891

Table 8: Meson Singlet with Spin 1 and Parity -1

Meson	Spin	Parity	Charge	Isospin		Up	Down	Strangeness	Mass
singlet	S	P	Q	I	I_3	u, $\bar{u}$	d, $\bar{d}$	s, $\bar{s}$	MeV/c
ϕ^0	1	-1	0	0	0	0, 0	0, 0	1, 1	1020

In the years 1974, 1977 and 1995 new particles were detected built up by new quarks (Brandt 2009, p. 439ff). These quarks were much heavier than the other three quarks and formed primarily particles like the first detected J/ψ particle. The J/ψ particle is the ground state of a system called charmonium. Charmonium is called *analogous* to positronium, which is a bound state of an electron and a positron. Charmonium is a bound state of the new charm quark and its anti-particle, too. Both show an *analogous* spectrum as a hydrogen atom, the charmonium, however, at much higher energies.

Table 9: Today known quarks (Bethge & Schröder 1986, p. 322; mass: Berger 2006, p. 300)

Quark	Charge	Isospin		Strangeness	Charm	Bottom	Top	Mass
	Q/e	I	I_3	S	C	b	t	MeV/c
u	2/3	1/2	1/2	0	0	0	0	5
d	-1/3	1/2	-1/2	0	0	0	0	8
s	-1/3	0	0	-1	0	0	0	150
c	2/3	0	0	0	1	0	0	1200
b	-1/3	0	0	0	0	-1	0	4500
t	2/3	0	0	0	0	0	1	174000

Some hard problems with hadrons and especially baryons constituted by quarks remained. First, free quarks were not found in any single experiment. Second, combinations of two or four quarks could not be observed, only combinations of a quark and an anti-quark or three quarks. Third, the baryons shown in the tables above all have positive parity, i.e., they have a symmetric total wave function, a state which is not compatible with Pauli's exclusion principle, i.e., that the total wave function of a system of fermions must be anti-symmetric if that principle stays valid (Berger 2006, p. 260).

The last problem was tackled by Greenberg with the suggestion of introducing another statistic than the Fermi-Dirac statistics called para statistics which allowed "three identical quarks to share one state" (Brandt 2009, p. 434). One year later 'in 1965 Han and Nambu attributed a new, three-valued property to quarks' (Brandt 2009, p. 434). Thus, each of the three quarks could have its own value of the new property and obey Fermi-Dirac statistics. Because this new property is invisible in experiments it was soon called 'colour' mixing like (*analogous* to) red, green and blue to a white (invisible) colour. Also, mesons could be explained with this model: The combination of a colour and its anti-colour sum up to white, too. With the assumption that hadrons always must be colourless (white) the second problem could also be solved (Berger 2006, p. 245). Other colourless combinations than the two mentioned are not possible.

Alone the colourless combination associated with the other properties is antisymmetric. This is the reason why no singlet of baryons exists (see p. 45). The wave function of that singlet with three quarks constructed by Gell-Mann together with the additional property 'colour' would be antisymmetric, resulting in a total wave function inclusive colour which would be symmetric. However, that contradicts the Pauli principle (Berger 2006, p. 270).

The two aspects of the third problem, the existence and linkage of quarks, could only be solved together because of the first problem. The hypothesis of quarks could only be supported by these effects if a theory of the force between the quarks has effects outside the hadron. Such a theory must be non-abelian, i.e., two successive transformations do not commute, because several fermions are involved, not only one (Brandt 2009, p. 426). Actually a lot of years ago in 1954 a non-abelian theory was developed in (*analogous*) generalization of quantum electrodynamics (QED) by Yang and Mills (Brandt 2009, p.325f), a so-called gauge theory based on the group SU(2). This type of theory was successfully applied to the weak interaction and its unification with QED by Glashow, Salam and Weinberg (Brandt 2009, p. 327ff). Originally Yang and Mills aimed to handle the strong interaction, though. This original theory of Yang and Mills motivated Fritzsch to try to apply it to the interaction of quarks. However, only when he met Gell-Mann they found together a successful extension to $SU(3)_C$ describing the colour force, later called quantum chromodynamics (QCD) in *analogy* to QED (Brandt 2009, p. 435). Therefore, in QCD there is an octet of massless gauge bosons, contrary to QED where is only a single massless gauge boson and the 'electroweak' theory of Glashow, Salam and Weinberg (GSW), where besides the massless photon additionally exist three very massive bosons because of their short range (Berger 2006, p. 13). A singlet of gauge bosons for the strong interaction does not exist because a colourless gluon could not distinguish different colours of quarks and thus could not make sure that only in sum do colourless bound quarks occur (Berger 2006, p. 247). The gauge bosons in QCD are called gluons, derived from 'glue', and carry both a colour and another anti-colour. For instance, "a red quark emitting such a gluon [red and anti-blue] turns blue; an anti-quark with the colour anti-red absorbing it becomes anti-blue. In this way a meson, for instance, a π^+ composed of a u quark and a $\overline{d}$ anti-quark, stays colourless while its constituents continuously change colour by exchanging gluons." (Brandt 2009, p. 435f).

A peculiar property of the non-abelian theory of QCD was that the interaction strengthens with higher energies and therefore distances, too, contrary to the case of QED. This property was later called asymptotic freedom and explains why no free quarks were found (first problem, see p. 48).

The property with external effect, which allowed experimental confirmation mentioned above, was the predicted ratio between e.g., the produced quark – anti-quark pairs and muon – anti-muon pairs in the annihilation of an electron and a positron. The predicted ratio of production agreed with the measurement results only if the factor 3 of the three colours were considered (Brandt 2009, p. 436).

3.4 Summary

In this overview of the development in high energy physics starting from the theory of the atomic nucleus to the constitution of quarks, we find in all mentioned theories a lot of cases using analogies for the reasoning of their hypotheses from authors like Heisenberg and Yukawa to Gell-Mann. The role of these analogies consists mainly in the support of the model explaining the new phenomena in a theory by reference to an already existing and accepted explanatory model for then well-known phenomena, if necessary, with extensions in the new model e.g., as in the extension of the Laplace equation by Yukawa. So then, why would analogies be used so often and no other kind of reasoning, one may ask. The answer is, while not explicitly mentioned in this chapter but understandable if one looks at the different applications of analogies in this chapter, that the considered objects are not directly observable in the model in focus of that theory or in the model referred to. Only specific phenomena are observable which allow to conclude the similar behaviour of the postulated objects. This observation gives a reason to investigate the role of analogies and their application in much more detail in the following chapter.

4 The Role of Analogies

Analogies are not a specific tool only in science, in contrast they are generally used in daily life in countless various situations and as such not exotic. However, considerations of analogies in science are rare. The first philosophical investigation in the second half of the last century was made by Mary Hesse in 1966, besides a short acknowledgement by Ernest Nagel (Nagel 1968 [1961], p. 107ff) and some more detailed remarks by Peter Achinstein (Achinstein 1971 [1968], p. 203ff) comparing them with models. Hesse's work was intended as criticism against the dominating account of explanation in scientific theories of that time, the deductive-nomological or DN account of Carl G. Hempel, an exponent of logical empiricism presented in chapter 2. Hesse argued that analogies are important for the explanation of 'new and unfamiliar' phenomena with 'familiar and intelligible' terms (Hesse 1966, p. 1) on the one side and for the prediction of new phenomena on the other. She distinguished three relevant types of analogies positive, negative, and neutral ones and described their function in the modelling of theories.

After not very satisfactory attempts for characterizing analogies only Bartha started in 2010 a new consideration of the usage of analogies in science as a 'refinement' of Hesse's theory. Regarding their structure, there he introduced four types of analogies. Only two of them, explanatory and predictive analogies, are relevant for physical theories. In the following, these analogies should be integrated in a model of theorizing in physics, especially regarding high energy physics.

4.1 Epistemological Model of Gaining Knowledge in High Energy Physics

In agreement with Pickering, I will defend the thesis that analogies are supposed to be 'at the basis of all that transpired' in high energy physics (Pickering 1984, p. 407). I agree and see the reason in case of high energy physics in the fact that the objects of investigation are unobservable. As mentioned at the end of the former chapter analogies allow as we will see in this chapter the extension of experiences made in observable areas into unobservable ones and further. Unobservability in this context means merely that human beings have no direct access with their senses; unobservability means not that there is no access at all. There was no phenomenon if there were no access because there was no observation. Access in case of high energy physics is mediated by electromagnetic interaction[14]. Electromagnetic interaction itself can cause actions or phenomena which are perceptible by our senses e.g., by observing the yellow light of natrium in the flame of a Bunsen burner or as in most cases via special measuring equipment e.g., the diffraction of white light by a prism. Therefore, such a limited access makes the investigation harder, however, not impossible.

My interest in this section is to find a model which explains the theoretical practice in HEP. The question addressed is how theories can be developed that treat highly distinct levels of scale of

[14] The electromagnetic interaction mentioned here does not obey classical electrodynamics because the emission of light must have a direction as Einstein explained during a talk in 1908 (Pauli 1961a, p. 86).

objects such as the atomic nucleus and quarks despite the limited access to the objects of investigation.

Generally, it may be assumed that theorizing in physics has two main tasks to explain: observed phenomena on the one hand and the prediction of new phenomena based on the theoretical explanation of the already known phenomena on the other hand. Explanation can mean in this context, whether to associate the phenomena in scope with already existing knowledge about similar phenomena and to describe their relationship regarding what is in common and what is different on the one hand. Otherwise, it can mean if it is enough to describe the regularities of the phenomena in an adequate way as some assumed on the other. In contrast, in the sixties of the last century the majority seemed to vote for the last option in view of quantum physics as Hesse claimed (Hesse, 1966, p. 5).

Predictions are conclusions for events in the future or laws drawn from the conditions of the theoretical relationship presented in the explanation of a theory. The predicted new phenomena then could be tested by a suitable designed experiment. Observed phenomena can either be a result of a prediction (e.g., the pion) or a result of an unexpected discovery (e.g., the kaon). So, three steps of development follow from one observed phenomenon to a next new observed phenomenon. Each of the three steps (in the following called sub-steps for a reason explained in a moment) between two different observed phenomena from the explanation of the first observed phenomenon, a related prediction of a new phenomenon, and the experimental verification or falsification of the predicted phenomenon form a single but wider step on a “ladder of development”. The “ladder of development” is used as a metaphor for the development here of HEP from quantum mechanics to the standard model of particle physics containing e.g., quarks. So far, this intertwining process is a rough sketch of the scientific enterprise in HEP as described in some other respects also by Pickering, Galison, Traweek or Hughes from various perspectives (Pickering 1984, Galison 1997, Traweek 1988, Hughes 2010). Thus, I suppose the succession of theoretical and experimental process steps leads to the standard model step by step like a ladder, of course only in retrospect and not in such a direct way as sketched here. In contrast, the goal of the standard model as it was reached since then was not in mind of the physicists in the beginning, because no one could foresee it. Instead, the aim of each theoretical or experimental step as an overall goal was to find an explanation for the constitution of matter (see the citation three in the introduction, p. 9). Only in the hindsight the direction of that development could be rated as satisfying or not.

I want to point out that it is a process guided by a so-called piecemeal methodology as Popper called it (Popper 1974, p. 51). ‘Piecemeal’ means that the direction of the development in the end is not fixed as in biological evolution, too. However, in contrast to evolution the separate steps have a starting direction (an explanatory theory for some observed phenomena). If the step is successful, a new step follows, if not, a step in a new direction will be tried till one succeeds. In this manner I suppose, a development takes place step by step like a ladder consisting each of the described three sub-steps i.e., explanatory analogy, predictive analogy, and experimental confirmation towards an explanation of the constitution of matter. The standard model of particle physics is an intermediate step of this enterprise (see citation two of Berger in the introduction, p. 9). Unsuccessful or nonpromising offshoots will die out over the time because of contradictory experimental results or outstanding acceptance in the community of physicists. A theory could only demonstrate the adequate description of the phenomena because of its non-phenomenological surplus structure not if it is true or false.

Nevertheless, it is assumed that in nature exists some "central order"[15] as Heisenberg called it (Heisenberg 1969, p. 251). "Central order" means a guiding principle that allows to form elementary particles, light, atoms, and molecules etc., comparable perhaps with the notion 'form' of Aristotle in contrast to 'substance'. The 'form' is meant as all that we can know whereas the 'substance' like the Kantian 'thing in itself' or the Aristotelian 'essence' is outside of possible epistemological reach. The form is an abstract structural model of the way nature behaves constructed by the abstracting process of historical development in science started in the beginning from directly observable phenomena. These phenomena were then addressed by a theoretical model after some time. After acceptance and proving of reliability of this explanation this model can serve as a model for an explanation of phenomena which seem to be similar in some respect but are not as observable as the phenomena used for modelling before. This process can be repeated in ever more abstraction and non-observability. However, as mentioned, what these only indirectly observed phenomena are in 'substance' remains nevertheless unclear I suppose.

> The general scheme of a "ladder of development"[16] seems to be a possible description of the overall development process of modern physics but not of the single steps. The restriction on modern physics appears because of the only indirect access to the objects of interest. The limitation of access forces all the three sub-steps of the process within one step of the ladder (the explanation, the prediction, and the experiment), to introduce special techniques of practice to carry out, two theoretical and one experimental. The two theoretical techniques will be handled in this chapter, the experimental techniques, and their role in chapter 5. In chapter 6 the theory will be worked out in more detail.

The theoretical techniques are the usage of analogies. Analogies are the only possibility to postulate unobservable objects by comparing phenomena and properties of known objects, with similar phenomena and properties of at the time of postulation unknown objects, as I will show in the next section. The behaviour of the known objects then explains the similar behaviour of the unknown objects. In general, the phenomena between the two compared domains are not identical such that additional assumptions must be made. These assumptions may extend the theoretical model in a way that predictions of new phenomena become possible. In both these steps analogies are useful. In a third step, the predicted phenomena were experimentally tested. If the tests are successful, the theoretical model would usually be accepted, and new phenomena can be taken into focus. This means a new step on the ladder can be climbed up by using a new (or extended) analogy to explain the new phenomena. Analogies in this context are as we will see the vehicle to extend well known structures from an established domain of knowledge to an unknown domain of observed

---

[15] „zentrale Ordnung" (my translation)

[16] following the notion of "Schild's ladder' for the procedure to progress from one point to the next in general relativity (Misner; Thorne; Wheeler 1973, p. 249)

phenomena of nature with assumed similar structure. The known structure serves via analogies as model in explaining and for predicting the unknown structure. Using an analogy is a method especially for theorizing on structures unobservable directly, i.e., for objects with limited access as in HEP.

Analogies are an epistemic tool to develop scientific explanatory hypotheses on the one hand and to support predictions on the other. The idea for the hypothesis will be supported by a 'prior association' (Bartha) in relation to an already known and accepted 'source model', which reminds in some respect to the observed new phenomena and therefore can explain them as belonging to a similar conceptual category. An analogy is intended to connect the new phenomena to the established knowledge in science and to put them at the right place within that corpus. The connection between the source and the target model can be very loose, e.g., only pictorial, or much stronger in a formal way of structural similarity. At the same time, the analogy can direct the attention to new experimental questions or hypotheses but also with some danger of misleading. Because of the repeated use of analogies, one can take them as an unplanned strategy, unplanned in two senses: one in intentional sense, the other in historical sense. In the first, the intentional sense, it is usually unplanned to look for a special analogy which a research hypothesis can be derived from. Mostly it goes the other way around, some observed phenomena remember to already be known phenomena which can be used as model for the description of the new phenomena. The other, the historical sense, means that the direction of development of a theory or the succession of theories is not planned and cannot be planned. To understand this assertion in the context of theory development and justification and the related questions, several steps of investigation are needed.

The sketched model of theorizing in HEP is based on the epistemic power of analogies. Therefore, it must be explained what analogies are (in the next section) and further how they work in HEP (in the last section of this chapter).

4.2 Characterization of Analogies

In this section I will discuss three important aspects of analogies, first why they are indispensable, then the formal structure of analogies and at last what analogies can do in case of unobservable phenomena or more correctly, observable only under difficult conditions. Let me start with the question of indispensability.

4.2.1 On the Indispensability of Analogies in Science

It is no surprise that analogies are widely used in everyday human thought and in science already in ancient times (Bartha 2013, p. 1). So, Hofstadter describes its omnipresence in the following comment:

"If analogy were merely a special variety of something [namely reasoning] that in itself lies way out on the peripheries, then it would be but an itty-bitty blip in the broad blue sky of cognition. To me, however, analogy is anything but a bitty blip—rather, it's the very blue that fills the whole sky of cognition—analogy is everything, or very nearly so, in my view." (Cited in Gentner, Holyoak, and Kokinov 2001, p. 499)

Hofstadter's characterization of analogy as (nearly) 'everything' in Gentner et al. is rather vague, yet it gives a good idea where analogical reasoning comes from. It is a kind of association based on regularities observed in the world. Learning and experience consists of comparing new impressions with already perceived ones that need to fit coherently in one's knowledge base or otherwise to reject them (see e.g., Elgin 2005).

On the role of analogies in scientific thinking, Hesse was the first modern author who raised constructive questions in her five essays under the title 'Models and Analogy in Science' (Hesse 1966). It is important here to recognize that analogies are not to have without models despite that Hesse does not mention this fact explicitly. However, it may be more understandable that analogies and models belong together, if one makes it clear to oneself that in an analogy one thinks of an object or a system of objects in reference to another, where the latter serves in some respect as model for the other. An analogy refers to a source domain, which fixes the context on which the analogy is based and points to a target domain, which is compared with the source domain. Therefore, Bartha defines an analogy as: 'is a comparison between two objects, or systems of objects, that highlights respects in which they are thought to be similar' (Bartha 2013, p. 1). In my opinion, Bartha's definition is quite not complete because the reference objects are not the objects or systems of objects themselves, but some imagined abstract model of these objects, which possess only the limited properties needed for the reference. This last point is that an analogy can be a dynamical object, which changes with the interests and its purposes.

In everyday thought, these objects or models in analogies are not really specified; they result from the context but can fail in some cases and generate then a misunderstanding. In science, the models are described more thoroughly. Nevertheless, here the same problem can occur. Therefore, some scientists want to refrain from the use of models and possible analogies, but most scientists do not. For example, Hesse asked whether analogies are indispensable for scientific understanding and argued for a link between analogical inference and its role for the development of concepts for the first time (Bartha 2010, p. VII). Hesse defends in her theory the use of analogies (or metaphors[17]) based on models against the DN account of scientific theories introduced by Hempel and Oppenheim dominating in the sixties (see chapter 2.2, p. 22f). In the DN (deductive-nomological) account, explanation consists of "a deductive subsumption of the explanandum [i.e., what has to be explained] under principles which have the character of general laws [i.e., nomos, Greek for law]" (Hempel 1998, p. 686). Hesse criticized this account because in her view "the deductive model of scientific explanation should be modified and supplemented by a view of theoretical explanation as metaphoric redescription of the domain of the explanandum" (Hesse 1966, p. 157). This statement means that in respect to theoretical science a model should replace the law because it is a more flexible object than the static law and an analogy should replace the deduction for the same reason. This flexibility is needed because explanation "impl[ies] an account of the new and unfamiliar in

[17] Hesse uses 'analogy' and 'metaphor' in the same sense. In contrast, Cellucci distinguishes analogy and metaphor reasoning that the further expresses similarity, but the latter creates similarity by comparing different subjects (Cellucci 2018, p. 71). I do not distinguish them, too, but use it in the sense of an analogy.

terms of the familiar and intelligible, […, and] does […, not] involve only a correlation of data according to some other criteria, such as mathematical economy or elegance" (Hesse 1966, p. 1).

Therefore, analogies obviously are not only relevant for discovery as we will see but also for justification of theories, namely its explanation. Problematic with the DN account is that a law as general principle is accepted and justified without any information how this justification is reached; it has no relation to its history. The circumstances of its discovery are part of this history. Contrary to that account a model has a history of background assumptions and associations, which compare these both with observed new and unfamiliar phenomena. The aim of this comparation is to check if they are analogous to the already known and familiar phenomena, such that the same or a similar description could apply to it, i.e., that it can explain it. However, Hesse emphasizes that this analogy or metaphor is no one-way street, the context of the new domain can affect also the view on the originating model of the explanation, changing the former context of background information and associations (see Hesse 1966, p. 162f).

Hesse's theory distinguishes three types of theory to describe the different implications: formal theories, conceptual models and material-analogy models. The first types also called mathematical models are simply empirical generalisations (regularities) like the law of Boyle and Gay-Lussac, stating that at constant temperature the product of volume and pressure for a gas is constant, i.e., they contain only observational predicates. Those theories have only weak predictivity because they are only weakly falsifiable, i.e., only in the case that it is no more valid in the next occurrences under the same conditions. The second type, the conceptual models, contain additionally theoretical predicates, which may be not interpreted in all cases as e.g., the predicates in the wave theories (see p. 58). This type is strongly predictive and therefore strongly falsifiable because new correspondences between observable predicates can be predicted within this type with chance of verification or falsification. Both types, however primarily the second, correspond to the law demanded in the DN account. The third type then is the type Hesse argues for - for the use of models based on material analogies. These models add to the strong predictivity and falsifiability of the second type, a justification 'by choice-criteria which appeal to the models as empirical data' (see Hesse 1966, p. 39f and 129).

Hesse introduces arguments from two books to illuminate her position more in detail and to explain basic notions of her theory: of 'La Théorie physique' from Duhem in the second edition as a proponent for a view like Hempel's[18] and 'The Elements' from Campbell (Campbell 1920) as a proponent of her own view. Both mentioned physicists have a contrary point of view, the first more from a theoretical standpoint, the second from an experimental one. Hesse refers with her arguments in a fictitious dialogue to a modern 'Campbellian', who defended presumably against a modern 'Duhemist' the necessity of analogous models not alone in the formulation of theories but also as essential part to develop their predictive power (Hesse 1966, p. 4f).

[18] Duhem was intensively discussed in the 'Vienna circle' and influenced mainly after the disintegration of the group and the migration of a lot of its members into the USA the development of logical empiricism (Schäfer 1998, p. XV*f). I use the German translation of the first edition (Duhem 1998).

Hesse explains her usage of the notion of 'model' and 'analogy' with an example from Campbell regarding the dynamical theory of gases (Hesse 1966, p. 8). In this example, billiard balls in random motion are taken as a model for a gas. Billiard balls are not identical in all respects to the molecules of the gas. For example, it must not be assumed that the particles of the gas are coloured like the billiard balls. However, it is assumed that the gas particles behave dynamically in an analogous way as the billiard balls, i.e., they have motion and impact. These properties, motion, and impact are called positive analogies of the model. The property of colour, which is assumed for billiard balls and not for gas molecules, is called a negative analogy of the model. Negative analogies are not relevant for the model (if they do not contradict some positive analogies).

For some of the properties of the model it is not known if they represent a positive or negative analogy. These properties are the relevant ones, called neutral analogies. Neutral analogies are important especially in science because they allow to predict a behaviour that has not been observed already. The model, and therefore the theory, is wrong if the prediction in a test is shown not to be fulfilled and must be altered or rejected. In the sense described by Hesse, the model contains the billiard balls and the positive, negative, and neutral analogies in relation to the gas molecules. An additional remark is needed with this presentation of analogies by Hesse. At least one negative analogy is essential for the presentation of an analogy, namely what distinguishes the source and the target domain in a relevant way. If there were no difference, then these two domains would fall together and there were no analogies but two instances of the same domain. A further clarifying remark is here not found in the literature, there are at least two levels of analogies, one on the level of the properties or analogical arguments (i.e. the positive, negative, and neutral analogies), and one on the level of the models (i.e. the target model is analogue to the source model), i.e. the motion of the billiard balls is analogous to the motion of the gas molecules for the first level and the system of billiard balls is analogous to the system of gas molecules for the second level. The last is the level of model analogy inclusive its context or, as Hesse it would name it, of the metaphor.

The former example from Hesse respective Campbell was used to define the usage of 'model' and 'analogy' by Hesse. Analogies in models can refer to equal or similar properties (e.g., mass of the billiard balls and molecules in the example above) or to equal or similar structures (e.g., mathematical description of the behaviour of the billiard balls and molecules) (Frigg and Hartmann 2012, p. 9). Some remarks about models are relevant here also in respect that they will play a role in the forthcoming discussion. Generally, a model is a representation of some other objects or structures. In science Frigg and Hartmann state that they "perform two fundamental different representative functions" (Frigg and Hartmann 2012, p. 3). First, a model can represent some part of the world as model of phenomena or data, and second, a theory as interpretation of the laws and axioms of a theory. Both functions are not exclusive for the same model, differential is only the focus of interest. In the first case, the interest is directed to regularities of the world, which should be represented by the model (billiard balls as molecules in a gas, however, also the molecules of the gas as imagination of the phenomena observed in connection with a gas in the real world). In the second case it is directed to the interpretation or instantiation of a theory in the semantic view of scientific theories (billiard balls and molecules as two model systems in kinetic theory, see Chapter 2.3, p. 28 for the explanation of the semantic view of theories). In case of the semantic view, models are on the higher level of model analogy. My use of the word ‘model’ refers to the first meaning because in my view the DDI account of a theory by Hughes (see chapter 2.4, p. 30f) is a better model for a theory in physics especially in HEP than e.g., the semantic view.

Back to Hesse’s theory, I have to explain the notion of ‘material analogy’ (Hesse 1966, p. 68) and additionally the former not mentioned notion of ‘formal analogy’ which corresponds to the horizontal respective vertical co-occurrences in the following table, which refers to a further example by Hesse, the wave model of water, sound, and light, respectively (Hesse 1966, p. 11, Table 10).

Table 10: Variables and parameters in the wave model of water, sound, and light by Hesse

Water waves	Sound	Light
Produced by motion of water particles	Produced by motion of gongs, strings, etc.	Produced by moving flame, etc.
Properties of reflection	Echoes, etc.	Reflection in mirrors, etc.
Properties of diffraction	Hearing round corners	Diffraction through small slits, etc.
Amplitude	Loudness	Brightness
Frequency	Pitch	Colour
Medium: Water	Medium: Air	Medium: “Ether”

In this example, the properties of the dynamical behaviour of water are compared with the corresponding kinds of behaviour of sound and light. All three phenomena are structurally described by a wave, mathematically formalized in a wave equation. The phenomena were produced by the motion of a generating object (gongs, strings, flame, or in case of water more correctly than in Hesse’s work by motion of water particles originated by wind or an infalling stone). All produced waves show the properties of reflection and diffraction and have corresponding parameters like amplitude and frequency. At first sight, only their medium is different (the essential negative analogy, see p. 56 above), i.e., the analogy seems to be nearly perfect, but not so in the positions of the fictitious Campbellian and Duhemist introduced above. The difference of the positions of the Campbellian and Duhemist according to Hesse is characterized by their notion of theory: whereas for the Campbellian the model of water is needed to explain that of sound and light, the Duhemist starts immediately with the mathematical model of a wave. In the following, I will modify some entries in the original table of Hesse to clarify the different points of view by introducing the remarks just made, and I will add a column for the mathematical description (Table 11).

Table 11: Variables and parameters in the wave model of water, sound, and light by Hesse, supplemented and modified by Kreisel

Water waves (Hesse modified by Kreisel)	Mathematical wave (Kreisel)	Sound (Hesse)	Light (Hesse)
Produced by motion of wind, thrown stones, boats, etc.	Variation of a parameter x by time	Produced by motion of gongs, strings, etc.	Produced by moving flame, etc.
Properties of reflection on walls, etc.	Properties of reflection by inverting parameter x	Echoes, etc.	Reflection in mirrors, etc.
Properties of diffraction at obstacles	Property of diffraction	Hearing round corners	Diffraction through small slits, etc.
Maximal Height of the wave	Amplitude	Loudness	Brightness
Number of waves per length unit	Frequency	Pitch	Colour
Medium: Water	No special medium	Medium: Air	Medium: “Ether”

My corrections are intended to show that in Hesse’s version (Table 10) the properties of water waves are mixed with notions of the mathematical model like amplitude and frequency on the one hand and unobservable properties as the motion of water particles itself on the other. The unobservable motion of water particles in my version is replaced by observable generating sources like wind or thrown stones, more comparable with the corresponding descriptions with sound or light in the two other columns. The two mixed kinds of parameters are separated in two columns, the first for water and the second for the mathematical model. The order of these two columns should reflect the historical occurrence of these models, perhaps the models of sound and light must come also before the mathematical model. In fact, the mathematical model could not have been developed before differential calculus i.e., not before Newton and Leibniz.

Regarding the formal analogies the example shows that the co-occurrence of the predicates in each column is connected by phenomena in the same medium they are, as Hesse argues, causal related in a wider sense. In a wider sense means for Hesse there ‘is at least a tendency to co-occurrence’ (Hesse 1966, p. 77) of the properties in the different columns of the table, i.e., if reflections and diffractions occur with water waves, then also echoes and hearing around corners occur with sound, and reflections and diffractions occur with light - or in other words if a set of properties occur in one medium it also occurs in another i.e. the properties of a column are connected in some sense. However, material analogies in the view of Hesse are the more important ones because they establish as demanded in the third type of theories a kind of justification for the choice of the model. Material analogies 'are pretheoretic analogies between observables such as properties of sound like pitch which are analogous to properties of light like colour which enable predictions to be made from a model' (Hesse 1966, p. 68). 'Pretheoretic' in this case means there is a theory of sound with e.g., pitch standing for the frequency of sound; but colour at the 'pretheoretic' time is only an observation with properties like bending on a prism, where white light is spread into different colours. This observation is 'pretheoretic' in the sense that one does not need a theory to observe this fact and because it is observable directly as phenomenon in the world, it is called 'material'. A formal theory containing such a model must be available to interpret colour as the frequency of a wave like sound or water waves for demonstrating the correspondence between properties of the known model and the new domain of observations which should be explained. These correspondences are the material analogies, which build up the formal analogy of the theory. To summarize, a theory in Hesse's view contains a 'causal' vertical structure based on observable 'pretheoretic' i.e., material phenomena, which serves as model (e.g., sound) for building up horizontal analogies to observed similar phenomena in a new domain (e.g., light) and concluding from these horizontal analogies to an analogy also in the formal structure. This is roughly my understanding of Hesse's approach to which I agree so far.

However, I do not agree to Hesse's statement that Campbell's and her more explicitly formulated theory of the need of models and analogies is a direct challenge for Duhem (Hesse 1966, p. 3). Hesse assumes that Campbell had 'Duhem among others in mind mounting his attack' (Hesse 1966, loc. cit.) according to a footnote in Campbell (Campbell 1920, p. 151) despite that Hesse concedes that Duhem was not mentioned by Campbell. The attack of Campbell was in contrast directed among others on the Continent against Mach, who was explicitly mentioned by Campbell (Campbell 1920, p. 7, 140, 152, 222 – 224). In the footnote cited by Hesse, only a preference of British and American scientists to theories using models (i.e., Hesse's third type) and of Continental physicists to mathematical theories (i.e., Hesse's second type) were stated as an 'interesting fact' however not in the somewhat unfavourable manner as in Duhem (Duhem 1998, p. 67ff). Duhem distinguishes British on one hand and Continental and especially French thinkers on the other as 'broad and weak' and 'strong and narrow', respectively.

Nevertheless, it is interesting to have a look at Duhem's theory because it leaves an explanatory gap, which in my opinion must be filled with models and analogies between them. There it must be held in mind however that Duhem argues against models but not against analogies (see Duhem 1998, p. 123ff). Hesse refers in respect to Duhem to the distinction between discovery and justification[19] and argues that in Duhem's view models are only a psychological aid for the discovery and are irrelevant for the justification of a theory as widely accepted at the time of her writings (Hesse 1966, p. 4). This matches in part with Duhem's view. However, theoretical physics is mathematical physics for Duhem because mathematics is the only science in which 'errors can be easily avoided and if they occur are easily to detect' (Duhem 1998, p. 139). Therefore, models and especially mechanical models have no place in this conception. This conception starts with a thorough definition of the entities used in the

[19] See the different kind of meaning of these notions as introduced by Reichenbach (chapter 2.2, p. 21)

theory and justification of the hypotheses on which the following deductions are based (Duhem 1998, p. 99).

To be more precise, Duhem defines a physical theory as: "A theory in physics is not an explanation. It is a system of mathematical propositions, which can be derived from a little number of principles and with the purpose to represent a set of experimental laws in a simple, complete, and precise way" (Duhem 1998, p. 20f). The only criterion for a theory's accuracy is the adequateness to experience (Duhem 1998, p. 22) which is the same as later for van Fraassen (see chapter 2.3, p. 27).

For Duhem four steps are needed to propose a theory. First, those properties should be chosen out of the set of physical properties to represent which are assumed to be simple whereas all other were assumed as combinations or sets of the simple ones. These selected properties should be represented by mathematical symbols and numbers, i.e., physical quantities fixed by adequate measuring methods. In this way symbols are only denotations for the selected properties but have no natural relation to them (Duhem 1998, p. 21). This first step corresponds apparently to the act of denotation in the DDI account of Hughes (see chapter 2.4, p. 30f). The second step of Duhem consists of the free formulation of some hypotheses about the mathematical relations between the defined physical quantities of the first step. Free formulation means here that one must avoid logical contradictions only within or between the hypotheses but has no other limitations. The formulated relations must be mathematical because of the demand of a theory as a 'system of mathematical propositions'. Then, at the third step, 'the different principles or hypotheses were connected by the rules of mathematical analysis' (Duhem 1998, p. 21). Again, the only limitations are that there occur no logical contradictions. Therefore, the system of propositions does not need to map the reality. This step together with the second is comparable to the demonstration part of the DDI account however much less limited (see chapter 2.4, p. 31). Only in the last, the fourth step consequences from the hypotheses of the theory on the one side and observations and measurements in experiment must agree on the other side. Duhem then called the postulated mathematical relations experimental laws, comparable to the laws in the DN account. To reach the agreement, the definitions and methods of measurement are important because they deliver a kind of 'dictionary' for the translation of the physical quantities of the theory into the expected results of the experiment. In the DDI account this translator step is called interpretation (see chapter 2.4, p. 31). For Duhem water waves, sound waves and light waves in Hesse's example above are all 'interpretations' of the same (mathematical) theory similar to the semantic view. This means mathematically one can express all three interpretations of the wave theory in a simplified form for plane waves by one equation $\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2}-\frac{\partial^2}{\partial x^2}\right)\boldsymbol{u}=0$ with $\boldsymbol{u}=y(x,t)$ for water, $\boldsymbol{u}=u_0e^{i\omega(t-\frac{x}{c})}$ for sound and $\boldsymbol{u}=\boldsymbol{u}_0e^{i(\omega t-\boldsymbol{kx})},\frac{\omega^2}{c^2}=\boldsymbol{k}^2,\boldsymbol{uk}=0$ for light, i.e. only the interpretation of $\boldsymbol{u}$ changes. They all fall under the same 'natural classification' of the wave theory as Duhem calls it (Duhem 1998, p. 20) even though these waves are different phenomena.

This is a convincing and logical argumentation, I think. However, there is a severe problem with it: a mathematical formulated theory is possible only if all needed ingredients are available at that time. Therefore, Duhem argues that historically the (mathematical) formulation of the theory 'lies in the air' and 'arise in him [the scientist] without any help' (Duhem 1998, p. 342), i.e., he ignores in some way the context of discovery. On the other side, he emphasises the importance of history for the development of physics (Duhem 1998, p. 364ff). The contrast between these two statements makes the problem. On the one hand, Duhem accepts that the historical development of physics (and mathematics) determines, what kind of theory is possible at that time, but on the other hand he ignores the immediate assumptions which make sense to the scientist and allow him to formulate his theory. It does not matter here, that in fact the theory 'lays in the air' in the sense that the same theory is independently developed from more than one scientist at the same time because each one must make his own assumptions. These assumptions are usually motivated by observations of similar already known phenomena and it is obvious to try a similar explanation i.e., to try a model analogous to the known and accepted theoretical model which explains well known phenomena. However, Duhem excludes models from his view of theorizing in physics although models allow to set phenomena in different domains in relation to each other, i.e., to build up phenomenological or as Hesse called it 'material' analogies. And only these analogies can demonstrate the similarities and differences between the phenomena. At the same time, the history of physical science also shows that verbal descriptions and explanations for physical events occur first before any work of abstraction generating a mathematical formalization can happen (see e.g., Huygens, Newton etc., p. 57f below). The type of theorizing Duhem describes corresponds to a phase of development in physics which Kuhn would describe as 'normal physics' of a mature science (see chapter 2.3, p. 26).

However, tentative moves in a new field do not allow such a procedure. There, in the new field (not only during revolutions as with Kuhn because the basic paradigm must not change as we will see in chapter 4.3.2, p. 86, where e.g., Gell-Mann develops an explaining model of hadrons in 1955 before formalizing it in 1961 and 1962) working with models as Campbell suggested is needed, I suppose.

A theory can only be acceptable according to Campbell and Hesse "if it is an explanation of phenomena, and this satisfaction implies that the theory has an intelligible interpretation in terms of a model, as well as having mere mathematical intelligibility" (Hesse 1966, p.4). Therefore firstly, you cannot start with a mathematical formulation but must have a model before, which establishes the connection to physics. This model demonstrates the circumstances of the occurrence of the phenomena considered in a more or less abstract but familiar way. The familiarity is needed for acceptance, to integrate the theory describing a physical system into one's own knowledge base (see chapter 4.2.1, p. 54). Otherwise, the system remains a strange and unexplained object if one is not in such a sophisticated mode based on decades of experience, which Duhem's account requires in my opinion. Secondly, Hesse claims with Campbell that a theory "is not a static museum piece, but is always being extended and modified to account for new phenomena" (Hesse 1966, p. 4). This statement should express the change of the underlying model (here the movement of water molecules respective air particles) to adapt it on the properties of a new or altered domain (the movement of air particles generating sound or the imagined movements of light in the 'ether'). Most of the observed properties stay the same or will be similar (e.g., reflection or refraction) however some new ideas occur (the transversal motion of the water wave is changed to the longitudinal motion of a sound wave to explain that it spread out in all directions, not only on a surface and later back to the transversal motion of light[20]). These new ideas produce perhaps new properties which are observable and therefore testable by experiment and can help to establish the acceptance of the new theory. However, Hesse's Duhemist (i.e., a fictitious person with Duhem's position) argues against this:

> "*I have an[other] objection to your account of the genesis of a theory of sound. You seem to imply that there are no sorts of theory-construction going on here. First there is the theory of water waves, which is arrived at by making a hypothesis about the propagation of disturbances, expressing this in mathematical language, and deducing from it the observed properties of water waves. There is no mention of any analogies or models here. But in the case of sound it is said that one-to-one correspondences between properties of water and properties of sound are set up first, and then the mathematical wave theory is transferred to sound. This may well be the way in which theories are often arrived at in practice, but you have said nothing to show that reference to the water model is essential or that there is any difference in principle between the relations of theory and observation in the two cases. Both theories consist of a deductive system together with an interpretation of the terms occurring in it into observables, and from both systems can be deduced relations which, when so interpreted, correspond to observed relations, such as the law of reflection. This is all that is required of an explanatory theory. You have implicitly acknowledged it to be sufficient in the case of water waves, and it is also sufficient in the case of sound waves. If we had never heard of water waves, we should still be able to use the same information about sound to obtain the same result. The information consists of the observed production of sound by certain motions of solid bodies, the relations between the magnitudes of these motions and the loudness of the sound, and between lengths of strings and pitch of note, and the phenomena of echoes and bending.*
>
> *All of these can be deduced from a mathematical wave theory with appropriate interpretation, without mentioning the water-wave model, and, what is more important, without supposing that there is anything connected with the transmission of sound or light which is analogous to water – that is, without supposing there are "hidden" motions of particles having the same relation to these observed properties of sound or light that the motions of water particles have to the properties of water waves. In fact, it would be very misleading to suppose any such thing, because some of the further consequences derived from a theory of water waves turn out not to be true if transferred, by the one-to-one correspondence, to sound and light transmission*". (Hesse 1966, p. 16f)

Hesse's Duhemist argues in this statement against the need of models or analogies and for the formulation of a mathematical wave theory as sufficient for deriving every kind of wave theory, for water, sound, and light, respectively. However, already in the case of water (Huygens 1890, p. 11, or any other fluid as with Newton, see Newton 1999, p. 353ff) waves are described verbally without a formalized mathematical wave theory and of course sound and light in analogy to water and sound, respectively (see loc. cit.). The order of the description of the wave phenomena follows the degree of

[20] In fact, light waves were modelled by Huygens as longitudinal waves in a medium called 'ether' historically (Huygens 1890, p. 11 and 17).

direct observability of the phenomena. In case of water the waves themselves with their properties are observable, in case of sound the properties of the assumed wave only are observable in a medium like ‘air’, assumed to behave like a fluid, in case of light similar properties are observed also, however, additionally properties not similar to fluids (polarization) and there is no observable wave and no observable medium so one is assumed, namely the ‘ether’. Of course, all this can be derived from a mathematical wave theory but only if one has, as we will see, an image of what a wave is, and this image is learned from a model and its behaviour. Usually in our times, as in the case of Duhem also, this fact is covered by the omnipresent familiarity with wave phenomena but, nevertheless, such an image is needed to accept the mathematical wave theory as an abstract model for wave phenomena. For an explanation to this claim look at the wave equation at page 55 and realize that to understand it one must first assume that ‘t’ stands for the time, which is not a mathematical, but a physical quantity as is ‘x’ for a length. Therefore, the interpretation of these quantities is already fixed in a special way and not uninterpreted as e.g., the members of a mathematical group, which are only defined by the relation between them.

Further, the formulation of solutions of the wave equation introduces already more or less periodic time-dependent functions, which are limited by some initial conditions. Where does these functions and conditions come from? Of course, from the former experience with waves. In fact, this is the reason one can throw away some details of the discovery context (namely the special circumstances where any wave occurred) however not the discovery context as a whole because the fact of the discovery has changed the perception of associated phenomena in all the future (namely the mental image what a wave is, which is not automatically covered by a mathematical equation). This conceptual imagination is the source of the explanatory and predictive power transferred by an analogy. In consequence, in table 2 above, the mathematical model must come always after the description of a model and again after the extension and adaptation to a new domain, and only in the review a ‘natural classification’ of different analogical phenomena under a common mathematical model is possible and from this perhaps a further interpretation. From this just described point of view both, Duhem and Campbell cover parts of the scientific practice in physical theorizing. This ambiguity seems to occur also in Hesse’s statement that ‘the protagonists [the Duhemist and the Campbellian] finally agree, fairly amicably, to differ’ (Hesse 1966, p. 5).

Moreover, as Michael Friedman argues, empirical laws rest beside formal structure also on physical principles (coordination principles, see chapter 2.2, p. 21) as relative a priori conditions for a physical theory (Friedman 2001, p. 71), where in case of Kuhnian scientific revolutions also ‘distinctively philosophical reflection takes places’ (Friedman 2001, p. 105). These physical principles must be introduced (often implicit and unexpressed) into the theory and it will be done by an analogy, which transports these principles by reference to a source model possibly with the need of modifications. Therefore, using an analogy constitutes a special kind of reasoning for the transfer of concepts with the consequence that it is also important for the justification of theories. I mean that with my argument above that mathematical formulation is not enough because one must know before its application, what a wave is.

Nevertheless, there is a danger in using models, namely in taking the wrong model for explaining some phenomena because no one knows in the beginning if predictions can be confirmed. The use of models however can limit the possibilities. E.g., in the case of wave theory, the wave function ***u*** can take an infinite number of analogical forms, a model limits this number to all which are compatible with the model and gives in this way a guideline for predictions i.e., neutral analogies, despite there being no guarantee that the prediction proves as equivalent to the observed phenomena. However, this is a little disadvantage against the need to try countless alternatives. From this view, I think, models and analogies are indispensable.

4.2.2 Formal Aspects

Following Hesse, Bartha especially considered the usage of analogies in science, however, the way I see it, primarily in refining the formal and justificatory aspects of Hesse's theory. Bartha noticed himself that he focussed on Hesse's second and third essay in her book, on first the structure and validity of analogies and second their justification (Bartha 2010, p. 40). However, before I review Bartha's own theory his criticism of Hesse's account should be discussed.

Regarding formal aspects already Hesse distinguished horizontal and vertical relations of analogies corresponding to the representation in table 10 and 11 above (see p. 57 and 58). The horizontal relations, which she called 'material analogies', are that of similarity, i.e., the positive, negative, and neutral analogies described at p. 56. The vertical relations, called 'formal analogies', show the analogical dependencies between the different properties of the two or more compared domains, or with other words, different 'interpretations of the same formal theory' (Hesse 1966, p. 68), i.e., formal similarities as the relation of pitch of sound and colour of light to e.g., refraction in both domains.

Remarkable is that Hesse prefers the 'material' analogies as first criterium of four in her argumentation for a good analogy. Her justification for this preference is that 'material' analogies rely on 'observable' or 'pre-theoretic' similarities (Hesse 1966, p. 69) between source and target domain, whereas a 'formal' analogy needs more knowledge on the target domain than usually is available during the development of a theory. Bartha criticizes this preference as the first of three points because in his view often a formal analogy is the starting point of an analogical inference even though he holds his own view on analogies as a 'refinement' of Hesse's view (Bartha 2010, p. 35). I agree with this critique of Bartha, especially to the extension of the target domain. Nevertheless, Hesse is right that her material analogies are fundamental because they contain the observations which enable the comparisons. However, the explanatory power rests more on the formal analogies because these allow to transfer the structure of the source to the target domain and to check if it matches. Secondly, Bartha sees no justification for Hesse's preference of material analogies (Bartha 2010, p. 43). This point of criticism can be opposed with Hofstadter & Sander's claim that all analogies and therefore also all scientific analogies are rooted in 'pre-theoretic' concepts (Hofstadter & Sander 2014, p. 520ff). The third and most important point for Bartha is that '[o]ften, an analogical argument is employed to *extend* our theoretical knowledge of the target domain' (Bartha 2010, p. 43, his emphasis). In my opinion, Bartha's critique is only partly justified. As we will see, Bartha's own theory will give an argument because analogies come in different forms. As mentioned to the first point of Bartha, first a check of a match of the model structure is needed. Then, in a second step, if the check was positive, the focus should be directed on known but still unexplained phenomena e.g., on polarization in case of light with wave theory. Only these phenomena call for an extension of the model and an extension of our theoretical knowledge existing in more than only the transfer of the model structure.

Hesse's second criterium that the relation between the different properties must be causal, i.e., essentially the same causal relations must hold in the source and in the target domain, (Hesse 1966, p. 87) was criticized by Bartha as too restrictive again (Bartha 2010, p. 43). I agree with Hesse that essentially the same relation must hold in both domains, and I agree with Bartha that to allow only causal relations is too restrictive. For example, in a mathematical model there are no causal relationships, nevertheless, such models occur in analogies, as especially Duhem demands, in different interpretations of a mathematical system (see p. 60). However, Hesse's understanding of 'causal' is a very wide one and may cover also mathematical relationships, i.e., 'causality' in her usage means 'at least a tendency of co-occurrence' of most of the same properties as e.g., diffraction and refraction in the wave example on both sides of the analogy.

Thirdly, Hesse claims that negative analogies should not be part of relevant properties or relations of the source or target domain because negative analogies contradict the similarities of positive or neutral analogies (Hesse 1966, p. 91). However, if one takes this literally, water waves and sound waves have then in the different medium a negative analogy. This shows that this claim is also too restrictive as Bartha criticizes (Bartha 2010, p. 44) because, if there were no negative analogy of the medium, both domains of the analogy would fall together and there would no analogy but only one domain i.e., an identity relation, which demands no different view on the one or the other. On the other side, difficulties like these could be resolved if such differences were taken as irrelevant for the analogy and the similarities as identical, as Hesse seems to demand as last criterium (Hesse 1966, p. 70f). This simplifying assumption guarantees then the validity of that analogy but may be too general as Bartha remarks (Bartha 2010, p. 41). However, I think Bartha has misunderstood Hesse here. I think she wanted to show that the compared properties are identical like two cats to the other, despite different colours, different length of the hair etc. they would be identified both as cats.

Generally, Bartha seems to be the first who grapples with Hesse's account of analogous reasoning intensively. His aim is to constitute a general theory of analogous or 'parallel reasoning', as he also calls it. His 'ultimate goal [is] to understand analogical reasoning' (Bartha 2010, p. 4). My critique on Bartha is that, as we will see, he restricts his refinement of Hesse's theory too much on the formal characteristics of hers because he focusses only on her chapters 2 and 3 regarding criteria for plausibility and possibilities of justification (Bartha 2010, p. 40) and does not give consideration enough to the power of the explaining metaphor in her account.

Earlier alternative attempts in philosophy were made to reduce the justification of analogous arguments either to deductive (e.g., determination rule of background information by Davies and Russell or missing premise by Weitzenfeld) or to (enumerative) inductive arguments (e.g., single-case induction by Hesse, Hume, Keynes, and Cartwright or sampling arguments by Mill or Harrod), or to argue for scepticism against them altogether (Agassi).

The assumption of Russell and Davies enables 'to convert an analogical inference in a valid deductive argument' (Bartha 2010, p. 46) in some cases with additional background information. Russell and Davies introduce to realize this a determination rule as function of the background information (Davies & Russell 1987). The determination rule should combine all relevant positive analogies needed to determine deductively the hypothesis. However, this requires that all relevant background information must be known because only then the application of this rule is justified, which is only exceptionally the case.

The other advocate for conversion of analogies into a deductive argument, Weitzenfeld, assumes in an analogical argument more generally a missing premise without requiring it to be part of the background knowledge (Weitzenfeld 1984). He concludes in the end of his argumentation that this missing premise 'is based on plausibility arguments' (Weitzenfeld 1984, p. 148). However, this is a result with no more reliability than the analogy in the beginning because also analogies start with a plausible assumption.

More generally, the main problem with reducing the justification of analogies to a deductive argument is that the justification cannot lie in the deductive part but in the quality of the premises as the problems above show. However, this quality could not be improved by pulling it into the deduction.

The second mentioned strategy for justification is to reformulate it as an enumerative induction. Both most popular attempts were generalizations from a single case and treat it as a kind of sampling argument (Bartha 2010, p. 48). A single case can work for Cartwright under special circumstances with 'sufficient control of the materials and [when] our knowledge of the requisite background assumptions is secure' (Cartwright 1992, p. 51). Then it could be generalized, e.g., for a chemical compound soluble in water it could be generalized that all such compounds are soluble or, as another case, if an essential property belongs to one exemplar of a natural kind then it belongs to all. However, this type of induction obviously is too restrictive for the use of analogies.

The second kind of reduction to an inductive argument takes the 'acknowledged similarities' as 'statistically relevant evidence for further similarities' (Bartha 2010, p. 50). Such a sampling argument is e.g., Harrod's 'fair sample argument' (Harrod 1956), which assumes that the part of shared properties within all known properties defines a probability which should not deviate much from the same probability of all, including the unknown properties. However, the search for one or more of the unknown similarities, usually needed in an analogy as hypothesis, is not supported by this procedure. Besides that, each additional considered property changes the calculated probabilities, which is not plausible because it contradicts the starting assumption.

Russell's more sophisticated version (Russell 1988, p. 257) considering additionally the determination rules mentioned above in the context of background information (see above) does not solve the problem of changed probabilities by adding further perhaps irrelevant features and assuming that then the list of features is complete. The same argument concerns Harrod's version, too. Regarding reducing the justification for inductive arguments, Mach already concedes that comparing relevant abstractions to irrelevant features is reminiscent of similar cases of induction. However, this is, as he claims, an error because the attention in an analogy is focussed on the abstract categorization of stable features and relations of the compared phenomena, not on repeating the same conclusion (Mach 2015, p. 313). Therefore, a reduction of justification to induction must fail.

A third strand of arguments referring to analogies are sceptical regarding their justification. Agassi rejects analogical arguments as either deductively valid, redundant, or ad hoc (Agassi 1964) or later as 'methodological essentialism' and therefore false (Agassi 1988). In the first case, Agassi assumes an analogical argument as showing nothing other than a deductively valid transfer of two isomorphic systems of laws. All other analogies in his view are generalizations of already known features, which makes the analogy redundant or based on a single case and therefore ad hoc. The problem with this argumentation is that it rests on the assumption of mere generalization as a method and not on the consideration of the plausibility of an analogy (Bartha 2010, p. 53). 'Methodological essentialism' in the second objection means that there are natural kinds (essentialism) and that 'human intuition reliably identifies natural kinds' (methodological essentialism, Agassi 1988, p. 417). However, there are analogies e.g., used in law, where no natural kinds are needed (Bartha 2010, p. 54), i.e., Agassi's argument in his second objection excludes too much.

The best way to justify analogical arguments seems to be a kind of a priori justification as tried earlier by John Maynard Keynes, too. However, Keynes's strategy resting it on the more fundamental principle of conditionalized probabilities of positive analogies to increase the conclusion (Keynes 1921, p. 258ff) fails because it works only for perfect analogies, i.e., if no negative analogies occur (Bartha 2013, p. 49f). The reason is that, if a non-trivial negative analogy occurs, the probability of the hypothesis does not increase because then source and target domain are not compatible.

Bartha favours a different approach, based on a modal extension of a symmetry principle as e.g. articulated by van Fraassen, that "problems which are essentially the same must receive essentially the same solution" (van Fraassen 1989, p. 236). The modal extension reads "if problems *might* be essentially the same, then they *might* have essentially the same solution" (Bartha 2013, p. 50). The first modality (the first '*might*' in the former sentence) is intended to be satisfied by the conditions of Bartha's own theory of analogous reasoning presented in the following, the second ('*might*') by a *prima facie* plausibility of a hypothesis which makes it reasonable to investigate because it may be correct. Much broader is the claim of Hofstadter & Sander that without concepts no thinking and without analogies no concepts are possible (Hofstadter & Sander 2014, p. 17), which is described in the next paragraph. This is the approach also proposed by Mach and sketched already above (see chapter 2.1, p. 17), which seems to me the most fundamental and therefore most satisfying.

Hofstadter & Sander argue that analogies are the other side of categorizations, starting from the beginning of speech learning like 'MUM' for your own mother, 'mum' for the mothers of other children to 'mother' as mothers of human beings and animals, over concepts like 'airport' as 'port' for airplanes and the sense of proverbs and ultimately, sophisticated concepts like the theory of elementary particles, the standard model of particle physics. The learning consists of extending categories like 'MUM' or 'elementary particle' by using these concepts as model for the occurrence in a similar situation and adapt it on this situation. This account is in the words of Hofstadter & Sander at 'the heart of thinking' (Hofstadter & Sander 2014, subtitle of their book) to which I agree, however, so general that for the purpose of explaining physical theorizing it must be made more precise e.g., by looking at Bartha's theory for scientific analogical arguments.

Bartha's own theory of analogous reasoning, the 'articulation model', rests on two principles called 'prior association' and 'potential for generalization', respectively (Bartha 2010, p. 98). 'Prior association' means that there must exist a strong connection of similarities between the source domain and the target domain and a further hypothetical similarity valid in the source domain, which is assumed to hold in the target domain, too. These similarities correspond to the positive and neutral analogies of Hesse, respectively. The similarities of the analogies between the source and the target domain are horizontal relations as Hesse called them and which her focus lies on. In contrast to that, Bartha lays more weight on the vertical relations of the properties within each domain because this allows to distinguish relevant and irrelevant properties and equivalently good and bad analogies (Bartha 2010, p. 93). Bartha's prior association must suggest that the observed similarities between source and target domain occur not by chance but regularly in some way and have their reason in a similar structure of dependencies between the properties of them. This suggestion makes the association *prima facie* plausible and establishes a potential for generalization, i.e., support for extending the assumption of the hypothesis to the target domain. The general structure of an analogical argument consists therefore in the two steps 'prior association' and 'potential for generalization'. Its differentiation in horizontal and vertical relations can be arranged in a scheme like Hesse's, which is easy to remember (see Table 12).

Table 12: Scheme of an Analogical Argument

Source domain (S)		Target domain (T)
P1(S)	->	P1(T)
P2(S)	->	P2(T)
…		…
Pn(S)	->	Pn(T)
N1(S)	->	N1(T)
N2(S)	->	N2(T)
…		…
Nn(S)	->	Nn(T)
Q(S)	->	Q(T)

P stands in this scheme for a positive analogy, N for a negative analogy, and Q for the hypothesis. Negative analogies are of three kinds: one which is irrelevant (the colour of the billiard balls in the gas example), one which is relevant and accepted (the different media of waves, these differentiate source and target domain), and one which is relevant and not accepted (these kinds lead to the rejection of the analogical argument). Negative analogies can be valid in the source domain and not valid in the target domain or vice versa. Also, different vertical relationships are possible. These vertical relationships are supposed to be valid in the source and the target domain. Bartha distinguishes four types of vertical relations depending on their function as analogical argument and further, two different subtypes of inference with two of these types (Bartha 2010, p. 96f).

In the first type, the prior association for the analogy goes from the positive analogies to the hypothesis and is called a *predictive analogy*. It is called *predictive* because the fact stated in the hypothesis of the source domain is a consequence of known facts in the positive analogies and, therefore, it seems to be plausible that an analogous hypothesis is valid in the target domain. This analogy can occur in a deductive or inductive form. In the deductive form, the features of the positive analogies entail those of the hypothesis in contrast to the inductive form. A typical deductive example is the mathematical wave theory in the view of Duhem. This mathematical model is, as usual in mathematics, not restricted in its application to water waves because a mathematical model catches only the formal structure of relations between some parameters, not their meaning (interpretation) in a special case. This is the argument Hesse uses in the role of the Duhemist in her discussion of the differences between the three physical models of water, sound, and light (see p. 55f). The wave equation in general is of the form $\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right)\mathbf{u}(t,\mathbf{x}) = 0$ with different functions $\mathbf{u}$ depending on its application on water waves, sound, or light. The inductive form will appear in the next subchapter (see e.g., chapter 4.3.3, p. 96 with Gell-Mann's analogous conclusion from two nucleons to three baryons).

In the second type, the positive analogies can be concluded from prior associations for the hypothesis, i.e., in the opposite direction as with the predictive analogy. This second type is called an *explanatory analogy*. It is *explanatory* when a prior association for a hypothesis explains some phenomena in the source domain, it is plausible then that a similar hypothesis in the target domain explains corresponding phenomena with similar properties there. Also, this type of analogy can occur in the deductive and the inductive form. The behaviour of the billiard balls as model molecules in a gas is an example for an explanatory analogy. Properties like the momentum of a billiard ball are transferred to a molecule with the effect that e.g., particles with a higher momentum reflected on a wall increase the observed pressure of that gas in a constant volume and therefore the resulting temperature.

The third type is called functional analogy and the association goes in this case from the positive analogies to the hypothesis and vice versa, i.e., in both directions. "Functional analogies are used to infer similarities in function from similarities in form" (Bartha 2010, p. 133) following the motto: form follows function and function follows form. Bartha gives biological and archaeological classifications as example. This is a type which can be found only in inductive form. In physics this type seems not to occur.

The last type or correlative analogy has no direction of association at all and only inductive form. The following example for an inductive analogical system can be found in Bartha (Bartha 2010, p. 96f). In a historical animal experiment from 1934 by Schaumann, mice developed S-formed tails caused by a new substance meperidine. It was well-known that this happened, too, when mice were given morphine, a strong narcotic. Therefore, meperidine must also have a narcotic effect, Schaumann concluded. With a second analogy of the same type, Schaumann transferred this result also to humans. Today this substance is still used as effective painkiller (Demerol). This is also a type which may not occur in physics.

The first two formal types of analogy correspond to the first two steps of theory development sketched in the introducing subchapter of this chapter and are used in the next subchapter to characterize the analogies which lead such physicists as Heisenberg, Yukawa or Gell-Mann to their theoretical achievements. However, before doing this a third and last essential fact regarding the theorizing towards the introduction of quarks is to consider what makes analogies indispensable in this process for a further, special reason: the role of observability, which divides the scientists at the end of the 19th and the beginning of the last century, too. Observability was strived often in this subchapter; however, it was not in the direct focus, but now it will be.

A concluding remark should be made: besides these more philosophical approaches considered before, much work was done in developing computer programs in artificial intelligence, which tried to formalize analogous inference in special cases or based on a principle of systematicity (Bartha 2010, p. 88f). However, this is not in the focus and will therefore not be deepened here.

4.2.3 Observability

We find, returning to the example of wave theory and its relation to water, sound, and light waves, that the model of the water waves possesses properties which are almost completely observable (i.e., all mentioned properties except the water molecules themselves). What is meant are properties as reflection on walls or refraction at obstacles, which are used to demonstrate the relationships of the observed phenomena, at least in times where computers or films and videos for such demonstrations were not available. However, also today it seems to me that such a demonstration with water would give a more realistic feeling what waves do or can do than a film, video or computer or a mathematical model (Try it at the sea!).

With good reason Kuhn, and Mach and Duhem decades before, emphasizes the importance of learning by practice and examples because only then a student develops skill and expertise (Kuhn 1976, p. 60f). A not neglectable part of that skill remains tacit knowledge, which seems to be the basis for the 'feeling' or intuition mentioned before. Also, Hofstadter & Sander give the example of the three types of waves and pick up the picture of the sea to illustrate its analogical power (Hofstadter & Sander 2014, p. 286ff). Once one has an imagination (a thought model) of the relationship between the observable phenomena and the unobservable particles of the water or another fluid, one can describe it as e.g., Newton did in his 'Principia' (Newton 1999, p. 356ff).

A description of the propagation of sound as pulses through the air analogue to a pendulum was also given by him (Newton 1999, p. 360ff). He was able to calculate the speed of sound, however, it corresponds not to the observed speed; it was about a factor of a sixth too low. In consequence, his calculation of the wavelength was too high. There are calculations in Newton's presentation, however, no complete mathematical wave theory was given despite him being one of the inventors of differential calculus needed for a wave theory. Light was only mentioned in common with sound (Newton 1999, p. 360). Also, Huygens gives a description of light in analogy to sound, published in

1678 (Huygens 1890, p. 11). However, a full mathematical formulation of a wave theory was also not to be found there. These findings continue in historical textbooks of physics respective philosophy of nature (e.g., Krüger 1763, p. 179 and 430) till about 1788, when Lagrange's 'Mechanique analytique' was published by his widow.

There, Lagrange gives a wave equation for an infinitely small amplitude in two dimensions of a fluid (water) in a shallow basin with even horizontal ground and explains that this is analogous to the density changes of the air with sound (Lagrange 1797, p. 550ff). At the same place, he criticized Newton's account with water oscillations in a bent tube as 'unnatural', whereas he explains his own description of ups and downs in the water as perfect analogy to the density fluctuations of sound. Lagrange was aware that his calculation was a solution of the general equations of hydrodynamics under much simplified conditions. However, his calculation was not more exact than Newton's.

Only in 1818, it becomes clear to Laplace that the pressure differences in a sound wave becomes greater because of accompanying temperature changes and therefore the speed of sound increases in the proportion of the specific heat constant at constant pressure to the specific heat constant at constant volume (Mach 2014, p. 207). About ten years later, observations of the Weber brothers show that surface waves on water are much more complicated phenomena which were not described by computable solutions of the hydrodynamical equations at that time (Baumgartner 1829, p. 219f).

The reason for these complications in formulating a common wave theory is that the wave lengths, frequencies, and amplitudes of sound waves are not observable with our senses in contrast to water waves. Also, the wave lengths for sound are much shorter. Observations of the properties of the new phenomenon (sound) that correspond to already known phenomena (water waves) are needed to reach the conclusion of a wave phenomenon (Table 11; column 3). In this situation, the problem arises that not all features of sound corresponding to waver waves are indeed observable. Still, it is true that both the loudness and the pitch, which correspond to the amplitude and the frequency, are clearly to be heard. However, both parameters must be measured in relation in an indirect way because the sound wave itself is unobservable.

Therefore, observational data, needed to formulate an adequate wave equation, were a long time not available. I will mention further the derivation of the laws of motion of the planets around the sun by Kepler to demonstrate the problem of a mathematical formalization out of well-known data. Kepler worked four years with the data of the orbit of Mars to find out that it has the form of an ellipse - and he could use the excellent observation data of Tycho Brahe. Therefore, a mathematical formalization of the relationship between observed data is impossible without an imagination of what you are looking for. Lastly, the formulation of a formal wave theory was impossible before the invention of differential calculus.

The problem for the observation of light is much worse. Of course, the wave itself is not observable, too, and the brightness and colour corresponding to amplitude and frequency seem to be observable in an easy way but only if one can interpret them in the way just mentioned, which was not the case in that time. In fact, the theory of light existed in two variants a long time. The Newton variant (emanation or corpuscle theory) assumed that light was the emanation of a type of corpuscles, which are produced in interaction with the eye and other matter. The Descartes and Huygens variant (undulation theory) on the other hand assumed that light is a longitudinal wave phenomenon in a fluid called ‘ether’. The first variant could explain the transmission of kinetic energy and refraction, the second could explain easily interference and partly double refraction. Both theories allow partly a mathematical formulation of their properties, however, both have problems with the account of polarisation.

The undulation or wave theory of light was accepted as correct explanation only after an experiment of Arago, confirming in turn the explanation of Fresnel and Young that coloured edges of a shadow result from the different diffraction of light with different wavelengths at the edges of the body that cause the shadow. Brightness and colour could be interpreted as amplitude and frequency of light only from that time on (1815). This explanation by Young and Fresnel helped to understand another property of light: that the wavelengths of light were extremely short because only in this case diffraction at the borders occur. Fresnel improved the undulation theory assuming transversal oscillations, however, retained the account of the ether as a fluid, which was criticized by Poisson. Maxwell’s explanation of light as an electromagnetic wave seemed to solve all existing problems in the end (Christiansen 1903, p. 356). I will not follow the history further here, but I will emphasize with Sommerfeld that ‘the water waves have served since ancient times as general model of a wave theory in physical imagination, despite they are much more complicated than acoustical or optical waves’ (Sommerfeld 1978, p. 151).

The last example, the two concurrent models of explanation of light mean that the mathematical model could not be the starting point of an explanation. In contrast, a model containing observables was needed to explain the phenomena by an analogous interpretation. This source model could alternatively be the wave model of water or the wave model of sound. In fact, both models do not explain all the phenomena observed with light e.g., the water model does not explain diffraction and the sound model does not explain polarization (as the water model does not either). Therefore, a modification of the mathematical model was needed, the introduction of transversal waves. The introduction of transversal waves is a modification which is not observable in a direct way, only via the application of special crystals on a beam of light. Additionally, the measurement of the frequency became possible in an indirect way with thin plates, which used the phenomena of reflection and extinction by superposition. The possibility of this measurement was enabled by the already known value of the speed of light. The corresponding amplitude must be calculated out of the measurement of the intensity, which is the square of the mean amplitude. It was assumed that light emitted from a bright surface spreads in all directions in space and decreases therefore with the square root of the distance from the increasing surface.

The ever more complicated situation with the media water, air and ‘ether’ resulted from the fact that water is visible, i.e., observable, air is not visible, however measurable e.g., via the air pressure, i.e., indirectly observable, and ether is not observable at all. Later, negative analogies were found which contradicted the hypothesis of an observable ‘ether’ (as happened through the experiments by Michelson and Morley) and this hypothesis was dropped finally. However, the long-lasting assumption of an allegedly non-observable ‘ether’ demonstrates the impressive power of the imagination in a model which needs a certain medium analogous to water waves or sound. In contrast, the pure mathematical model is obviously not dependent of a medium.

These historical details demonstrate that regularly an analogy to one or more simpler models with more observable features did support developing an explanation of a new phenomenon. From a clear and simpler understandable starting point more abstract models can be developed with less direct observable features if the simpler model is accepted as reliable beforehand. In case of light, Baumgartner writes: “A light wave is the same in vibrating ether what in vibrating air was called a sound wave”[21] (Baumgartner 1829, p. 277). Nevertheless, Baumgartner is sure that “we are in the darkness regarding the nature of light; we cannot perceive it with any sense, certainly it makes other objects visible, but could not be seen itself: however, we are compensated for that by an exact knowledge of its modes of operation and their laws”[22] (Baumgartner 1829, p. 276). These two statements indicate in a short way that analogous models help to explain new phenomena especially in unobservable regions as in quantum physics.

Therefore, Table 11 is to read from left to right. The model of water waves is (nearly) completely observable. Moreover, the model of water waves opens the opportunity to develop a mathematical model (when the mathematical methods were available). The mathematical model gives a framework generally to be applied to other wave phenomena. For this application, an interpretation of the positive, negative, and neutral analogies between the phenomena of the different domains of application is needed. The interpretation becomes harder if some properties or phenomena are not observable, as with sound and even more with light. Exactly this situation occurs even more drastically in case of the composition of the atomic core and of elementary particles like the nucleon, explained in subchapter 4.3.1.

Analogies can, following Hofstadter & Sander (Hofstadter & Sander 2014, p. 525ff), shortly be described as building connections between former unconnected however in some respect similar conceptual categories, which in turn allows to extend the concept of the former separated categories a level higher in a more abstract category with a new understanding. Thus, the role of analogies in physics is to explain phenomena and predict new ones, also and especially in unobservable regions, based on already known and accepted types of models. In case of not already contained but observed phenomena, they must be extended. From this extension logically new (novel) observable phenomena might arise, phenomena which again are not already observed. These novel phenomena allow to set up a new predictive analogous model, which then can be experimentally tested.

To summarize, two types of analogies, explanatory and predictive ones, are needed to build up a theory of physics to explain already known phenomena in a new domain, to extend this theory with already observed phenomena in the new domain and to predict then unobserved phenomena which in turn can support or contradict the theory by experimental results. This claim will be demonstrated

[21] “Eine Lichtwelle ist dasselbe im vibrierenden Äther, was in der vibrierenden Luft eine Schallwelle genannt wurde.“ (my translation)

[22] Über die Natur des Lichts sind wir ganz im Dunkel; wir können es durch keinen Sinn wahrnehmen, es macht uns zwar andere Gegenstände sichtbar, kann aber selbst nicht gesehen werden: dafür sind wir aber durch die genaue Kenntniß der Gesetze seiner Wirkungsweise entschädigt.“ (my translation)

by some examples in the following subchapter for the development from the explanation of the atomic nucleus to the postulation of quarks.

4.3 From the Atomic Nucleus to Quarks

The step from a wave model of light, as in the example in the last section to the atomic nucleus starting in this section, is a heavy one. The whole development of electrodynamics, special relativity and quantum mechanics lies in between with increasingly unobservable facts and increasingly sophisticated models to explain them. However, here is not the place to discuss them because the focus lays on the role of analogies in the theoretical practice of physics leading to the hypothesis of quarks and including the former steps by examples would be a too long way to go. To demonstrate the application of the analogies characterized in subchapter 4.2 based on the sketched theoretical developmental of subchapter 4.1 on the theoretical path to quarks, I will describe three examples: the model of the atomic nucleus developed by Heisenberg, the model of the strong force developed by Yukawa and of course the several steps to the model of quarks developed by Gell-Mann (and independently by Arnold Zweig but not considered here).

4.3.1 Heisenberg's Theory of the Atomic Nucleus and Yukawa's Theory of Force

I will start with the analogy in Heisenberg's hypothesis about the constitution of the atomic nucleus in 1932 (Heisenberg 1932a). In the section before I emphasized that the necessity to use analogies is forced by the limitations of observability and Heisenberg was aware of that fact. In his autobiographic memoirs, he described, referring to a remark of Einstein that theory decides what is observable, his solution to the problem of the observability of electrons (Heisenberg [1969] 2010, p. 96f). There, he states that the orbit of an electron in an atom is not observable although the track of the same particle in a cloud chamber is. Heisenberg solves this problem of different behaviour with his now well-known uncertainty relation. Remarkable is his reasoning because it prefers the unobservability. He writes that the track of an electron in a cloud chamber consists of an imprecise chain of little drops of water. However, these drops are much larger than the electron. Therefore, he argues the track can only give an imprecise impression of the location and the velocity of the electron in agreement with the experimental situation. Heisenberg calculated the minimal uncertainty of the observation as the product of the uncertainties of location and momentum (the last means the product of mass and velocity), which cannot become smaller than the Planck constant. This is Heisenberg's uncertainty relation. Indirectly, he uses in this problem also an analogy; in Bartha's nomenclature an explanatory analogy (see subchapter 4.2.2, p. 69). This analogy compares the orbit of an electron with the track of an observable object, namely the drops of water, generated by the unobservable electron. First Heisenberg formulates the hypothesis that electrons are unobservable in the orbit and concludes that they must be unobservable in the track of the cloud chamber, too. The positive analogies, which secure that both unobservable particles are electrons, are suppressed here. However, for a physicist these analogies are obvious because of the theoretical construction of an atom and the experimental confirmation via the Zeeman effect. With the splitting of the spectral lines in a magnetic field on the one hand and the deviation of the tracks in a magnetic field on the other, the charge to mass ratio could be calculated as the same and therefore both as the same kind of particle. Using these positive analogies, the explanation would run as just described.

In the following, however, I will not enumerate the positive analogies belonging to an explanatory analogy all the time, only if needed for understanding.

Heisenberg introduces in his hypothesis about the constitution of the atomic nucleus again an explanatory analogy, however now explicitly. As described in chapter 3.1, page 36, he compares the nucleus of deuterium with the positive ion of a hydrogen molecule. The orbits of electrons in an atom are unobservable as mentioned in the last paragraph. Therefore, this finding applies of course to the hydrogen molecule and much more to the nucleus of an atom, too. On the other hand, according to Einstein a theory decides what is observable. Einstein meant with this that 'only theory, i.e., the knowledge of the laws of nature, allows to conclude from the sensorial perception to the underlying phenomenon' (Heisenberg 2010, p. 80), i.e., an observation is theory-laden. In this way the theory must explain observable phenomena in a plausible manner. For the constitution of atoms and molecules this task was successfully solved in the years before Heisenberg's papers on the atomic nucleus were published despite the problem of non-observability of the atoms and molecules itself. This was possible because the spectra of the atoms and molecules, their frequencies and intensities, were observable - and already observed. Rutherford's model of the atom in analogy to a planetary system, an explanatory analogy with the atomic nucleus as sun and electrons as planets (see chapter 3.1, p. 36), inspired Bohr to improve this model by demanding that only orbits are allowed which are whole numbered multiples of Planck's constant because from atoms (here the hydrogen atom) light was emitted only with discrete frequencies. Bohr's model explained the stability of atoms and the discrete spectra; however, the frequencies did not correspond to the frequencies of the electron orbits. Only in the limit of extremely high excitation the emitted frequencies and the orbit frequencies correspond to each other (Heisenberg [1958] 2007, p. 57).

On the other hand, the strange double nature of light, sometimes as wave, sometimes as particle - depending on the experiment, had weakened the aversion to live with undecidable facts. Thus, in 1925 the idea arises to describe the electron orbit not as orbit with location and velocity but with amplitude and frequency of a Fourier expansion (Heisenberg [1958] 2007, p. 58; Heisenberg 1925), which maps better to the observable phenomena. One year later, Schrödinger found an equivalent formulation based on wave functions (Schrödinger 1926a), which allowed in an easier way to extend the theory from one to many body problems (Heisenberg 1926). That was the starting point to establish in principle the description of the structure and the related phenomena of all the atomic elements in the periodic table and the molecules inclusive solid-state physics built up by them (Finkelnburg 1967, p. 204). "On base of quantum mechanical calculations made predictions earn in this area of physics absolute confidence"[23] (Finkelnburg 1967, p. 205). Therefore, the ionized hydrogen molecule just could be the right model for a similar quantum mechanical object as Brandt stated in his presentation of Heisenberg's theory on the nuclear forces (Brandt 2010, p. 224). The binding of two protons by one electron was plausibly explained by quantum mechanical calculations despite its difficulties and accepted by confirmation through observations of the spectra in experiments (e.g., Hylleraas 1931). The calculation showed that the wave function must be invariant against rotations around the axis between the two protons and stay unchanged against the exchange

[23] "die auf Grund quantenmechanischer Rechnungen gemachten Voraussagen verdienen in diesem Bereich der Physik *unbedingtes* Vertrauen" (my translation)

of the protons, i.e., it is symmetric to the centre of the molecule. The whole system would shrink if the force between the two protons would be much stronger - e.g., to the size of an atomic nucleus if the force were strong enough. This would be a possible explanatory analogous model for the nucleus of deuterium, the hydrogen atom with about double the mass of hydrogen, if the charges and masses were exchanged.

Both models of the analogy, the hydrogen molecule and the deuterium nucleus, were unobservable, and the formal structure of the calculation alone was transferred. Nevertheless, there remains a problem, such a model would be not stable because the electron must move so fast because of the uncertainty relation that it could not stay in an orbit in such a system. The picture changed when Chadwick identified the neutron as a new particle in 1932 (see chapter 3.1, p. 35). Two months later, Heisenberg postulated that atomic nuclei were built up by protons and neutrons without the participation of electrons (Heisenberg 1932a, p. 1). Deuterium is the only case which could be calculated directly because of its analogy to the ionized hydrogen molecule. The system of deuterium consists of one proton and one neutron, where the neutron is thought of as a composition of a proton and an electron despite Heisenberg' awareness of problems with that idea. First, such a composition could not be stable for the same reason mentioned before, the electron would not stay in the neutron because of the uncertainty principle. Second, such an electron must be a special kind of electron because it must have zero spin. However, Heisenberg preliminarily accepts this situation, leaving the solution of the problem for the future because of the explanatory power of his assumption as we will see.

In fact, the former description by Brandt is not complete, especially regarding the used analogies. Heisenberg refers in his depiction of the deuterium nucleus ("Urey'sches Wasserstoffisotop") to the 'helium problem' (Heisenberg 1932a, p. 4), which introduced the 'exchange forces' mentioned by Brandt (see chapter 3.1, p. 36) as name for the exchanged energy between two equal coupled oscillators. The 'helium problem' is the fact that helium like other atoms with more than one electron in the outer shell were not correctly described by (Bohr's) quantum theory of that time on the one hand and that helium exists in two not combining forms (ortho and para helium) on the other hand (Heisenberg 1926, p. 426ff). Heisenberg showed by a quantum mechanical calculation to solve this problem that two electrons, oscillating on a line on both sides of an opposite charge located in the middle, behave in the same way as the classic oscillators in (explanatory) analogy to the coupling of two equal classic harmonic oscillators. The coupling of classical harmonic oscillators demands that both have nearly the same resonance frequency, independent from the energy of the whole system. Then, the system can mathematically be taken apart in two uncoupled oscillators if the exchange energy is a quadratic function of the coordinates: one oscillator with a higher energy and frequency than in the uncoupled case, one with a lower energy and frequency. In quantum mechanics the equality of two systems of a chemical element is regularly fulfilled so that this decomposition is always possible. Additionally, frequency and energy are coupled in a linear way, i.e., E = hν. Therefore, the energies of the stationary states are equal to

$$\mathrm{H}_{\mathrm{n}_1,\mathrm{n}_2} = \frac{h(\omega_0+\lambda)}{2\pi}\left(n_1+\frac{1}{2}\right)+\frac{h(\omega_0-\lambda)}{2\pi}\left(n_2+\frac{1}{2}\right), n_1, n_2 \in \mathbb{N}$$

because of the quantum condition E = hν.

In the second term of the decomposition in the equation above, the two oscillators swing in phase (symmetric case), in the first term they swing in counter phase (in a phase shift of π, i.e., anti-symmetric case). In the symmetric case of the first term with n_1 only an electric moment occurs with the consequence that transitions to other states are only possible by changes of n_1 by steps of one unit. However, also higher momenta can contribute to the radiation with smaller values.

In the first case, the second term can contain n_2 only in even steps because in the anti-symmetric case the state of one oscillator fixes that of the other. Thus, two term systems arise, which do not interact. A split of every term in four somewhat different terms arises if one considers the magnetic moment additionally to the electric moment of the electrons. The split into two noninteracting term systems survives but in a different form. At first sight, it seems that it is not determined which term system is realized in quantum mechanics, the first, the second, or both. However, Heisenberg argues that only one term system can be realized in nature determined by the counting rule of the Bose-Einstein statistics in combination with Pauli's exclusion principle, now known as Fermi-Dirac statistics[24]. The realized system in case of helium is the term system consisting of a singlet (para helium, symmetric in space and with spin of both electrons in opposite directions, i.e., anti-symmetric regarding the spin) as ground state and with excited states. Alternatively, a triplet (ortho helium, anti-symmetric in space and symmetric spin, i.e., spin of the electrons in the same direction) occurs with only excited states. The other possibility, singlet and triplet states exchanged in their position, cannot be found in nature due to the validity of the exclusion principle, which does not allows electrons in equivalent orbits with all the same quantum numbers. Additionally, the number of possible systems is reduced from two to one in congruence with the Bose-Einstein statistic.

The importance of this result is its generality because it assumes only harmonic oscillators and the quantum condition. Moreover, the scheme can be extended to any number of systems. The scheme can also be used as explanatory analogical model for the ionized hydrogen molecule described by Brandt with exchange of the positive and negative charges and half the core charge. That was what Hylleraas does in his publication (Hylleraas 1931, p. 743) and again, with the assumptions introduced by Heisenberg in his paper from 1932, for the atomic nucleus of deuterium (see pp. 76 in this and 36 in chapter 3.1).

The assumptions by Heisenberg regarding the atomic nucleus are that in first approximation proton and neutron are two states of one particle, neglecting the charge and the slightly different mass and leaving the question of stability of the neutron for the future. Heisenberg thought before the discovery of the neutron and even somewhat later those electrons are constituents of the atomic nucleus because beta decay emits electrons. He assumed that proton and electron are put together to form the neutron, which calls for the question of stability. However, the problem of the spin of nitrogen (see chapter 3.1, p. 35) demands that the spin of the neutron must be one half, therefore, the spin of the electron as part of the neutron must be zero, which would be a curious kind of

[24] Fermi has developed this statistic a short time before the publication of Dirac using a quantum mechanical formulation. Heisenberg has developed this statistic independently also a short time before as Dirac noted (Brandt 2009, p. 175f).

electron. So, all problems were shifted into the neutron and the combined system could be handled by quantum mechanics (Heisenberg 1932a, p. 1). The calculation method of harmonic oscillators mentioned in one of the former paragraphs is more general than only reduced to deuterium and could be applied to other atomic nuclei, too. Therefore, Heisenberg could give a general formula (the Hamiltonian), in principle applicable to all atomic nuclei describing their composition (Heisenberg 1932 a, p. 3):

$$H = \frac{1}{2M} \sum_k p_k^2 - \frac{1}{2} \sum_{k>l} J(r_{kl}) \left(\rho_k^\xi \rho_l^\xi + \rho_k^\eta \rho_l^\eta\right) - \frac{1}{4} \sum_{k>l} K(r_{kl}) \left(1 + \rho_k^\zeta\right)\left(1 + \rho_l^\zeta\right) + \frac{1}{4} \sum_{k>l} \frac{e^2}{r_{kl}} \left(1 - \rho_k^\zeta\right)\left(1 - \rho_l^\zeta\right) - \frac{1}{2} D \sum_k \left(1 + \rho_k^\zeta\right).$$

All terms depend on three space coordinates r = (x, y, z), on the spin σ^z of each particle in direction z in the nucleus and a fifth number ρ^ζ which can take two values, +1 for the proton and -1 for the neutron. Further transition values

$$\rho^\xi = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \rho^\eta = \begin{vmatrix} 0 & -i \\ i & 0 \end{vmatrix}, \rho^\zeta = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}$$

occur also in the formula additionally to the two values +1 and -1 in explanatory analogy to the Pauli matrices because of the needed exchange of proton and neutron. This analogy to the two-valued possible orientation of the spin of electrons represents in the same way the two states of a nucleon either as proton or as neutron, today known as 'isospin'. Thus, the explanatory analogy is used to describe the same behaviour of a new phenomenon.

The first term in the formula of the Hamiltonian above stands for the kinetic energy of the particle (proton or neutron), the second for the exchange energies of two particles, the third for the attraction between any two neutrons, the fourth for the Coulomb repulsion of any two protons and the last for the mass defect of the neutrons. Here, in this formula in the second term, Heisenberg makes the reference to the ionized hydrogen molecule, which Brandt mentioned as analogy to the binding of proton and neutron: "Putting neutron and proton into a distance comparable to the dimension of an atomic nucleus an exchange ["Platzwechsel"] of the negative charge will take place – in analogy to the H_2^+ ion"[25] (Heisenberg 1932a, p. 2). From this statement follows that the binding energy is much higher than in the hydrogen molecule ion because of the smaller distance and that proton and neutron are bound by a particle with negative charge like an electron in such a way that proton and neutron exchange their position 'periodically in a continuous way'[26] (Heisenberg 1926, p. 421).

Obviously, in the naming of Bartha, this is an explanatory analogy, the hypothesis of the binding of proton and neutron is supported in analogy to the hydrogen molecule ion by the argument that there is exchanged an electron-like charge without spin (a boson). In contrast to the example of an atom with two electrons (helium), described in Heisenberg's paper from 1926, in this case the proton and neutron 'change' their place and not the electrons. However, the explanation scheme of energy exchange between harmonic oscillators is general enough to 'take this burden'. The word 'change'

[25] „Bringt man Neutron und Proton in einen mit Kerndimensionen vergleichbaren Abstand, so wir d - in Analogie zum H_2^+-Ion - ein Platzwechsel der negativen Ladung eintreten". (my translation)
[26] "periodisch in kontinuierlicher Weise" (my translation)

was set in hyphen because it was supposed in this picture that the particles do not change their place but only the states as proton or neutron by the exchange of a spin-less electron.

For the third term, Heisenberg uses a similar explanatory analogy, the analogy between the neutral hydrogen molecule with its two electrons on the one side and two neutrons on the other side. These two neutrons are thought as before as composed out of a proton and a spin-less electron. In analogy to the hydrogen molecule, it is assumed that the binding between two neutrons is a little bit weaker. This assumption is essential for the constitution of atomic nuclei as we will see. Both terms, the third and the second, should have a radius of action of about 10^{-12} cm only because of the size of the nucleus.

The fourth term stands for the usual repulsion of the positive charges of the protons. Here no new analogy is needed.

The last term is an expression for the mass defect of the neutrons. This means that the neutrons are bound somewhat weaker to the nucleus than a proton. This fact will be important for the explanation of the stability of specific nuclei of different elements and their isotopes.

Heisenberg could explain qualitatively the constitution of all atomic nuclei interpreting this equation, not only the deuterium nucleus. But deuterium has one advantage: it was the only nucleus for which in reference to the helium problem the wave function could be given directly:

$\psi(r_1\varrho_1^\varsigma, r_2\varrho_2^\varsigma) = \varphi(r_1 r_2)\left[\alpha(\varrho_1^\varsigma)\beta(\varrho_2^\varsigma) \pm \alpha(\varrho_2^\varsigma)\beta(\varrho_1^\varsigma)\right]$ with $\alpha = \delta_{\varrho,1}, \beta = \delta_{\varrho,-1}$.

Analogous to the helium problem, the symmetric solution (that with the positive sign) must be the correct one for attraction and therefore for stability. This symmetry is a spatial one. However, the whole wave function must be anti-symmetric because of the exclusion principle of Pauli. Therefore, the spin must be symmetric (a triplet), as found in measurements, because of the anti-symmetry of the isospin (a singlet).

In general, all nuclei with the same number of protons and neutrons must be the most stable because of the symmetry of the second term of the Hamiltonian, neglecting the last three terms at the first approximation. This fact corresponds to the observation that most atomic nuclei have about the double of the weight as their charge indicates in relation to a proton (or hydrogen nucleus). Then, the combination of two neutrons should be the most stable considering the spin and the exclusion principle. However, this is not the case because of a reason discussed by Heisenberg at the end of his first paper on the constitution of atomic nuclei (Heisenberg 1932a, p. 6). The reason is that the binding of a proton and a neutron gains more energy than the binding between two neutrons if, as assumed, the second term is larger than the third. Then, a nucleus with two neutrons would decay by emitting an electron (β-decay). Therefore, the most stable nucleus, considering the minimal condition of the second term, is the helium nucleus with two protons and two neutrons and with isospin and spin zero (the first nucleus with a closed shell). In contrast to elementary particles with half spin, compositions of elementary particles built up by an even number of them (e.g., in atoms) could have a symmetric wave function – therefore they are bosons.

The assumed higher weight of the second term, the exchange term (“Platzwechselterm”), in relation to the last three terms results in an overweight of the neutrons over the protons in more massive atomic nuclei because of the rising of the Coulomb repulsion of the protons. If, however, in all cases the energy gain of a proton is higher than that of a neutron, the neutron decays and emits an electron. This process will continue so long as that condition holds, i.e., the proportion between neutrons and protons is larger than an upper limit fixed by the energy gain of proton vs. neutron.

The binding of the constituents of a helium nucleus is larger than the binding to the other constituents within a heavy atomic nucleus. Therefore, If the proportion between the number of neutrons and protons sinks below a lower limit, which is determined by the Coulomb repulsion of the protons, a helium nucleus, i.e., an α-particle, will be emitted. Both processes limit the region of stable nuclei, only those nuclei are stable which are near enough to the minimum function of the energy in relation to the proportion of neutrons and protons. In his second paper to the constitution of atomic nuclei, Heisenberg argues that in case of an even sum of the number of protons and neutrons, nuclei with also an even number of both protons and neutrons are more stable than those with each an uneven number because of the stabilizing properties of the helium nuclei as part of heavier nuclei (Heisenberg 1932b, p. 156ff). Heisenberg mentioned only four stable nuclei of this second kind (Heisenberg 1932b, p. 159). Nuclei with an uneven sum show not such a difference of their behaviour besides the effect that heavier nuclei with an even number of protons are somewhat more stable (Heisenberg 1932b, p. 159).

The subject for all three papers on the constitution of atomic nuclei by Heisenberg was of course the discussion of the unclear and inconsistent properties of the neutron. Whereas in the first paper Heisenberg assumes for the setup of the Hamiltonian of atomic nuclei only that the neutron is a fermion with half spin setting stability and constitutional problems aside (Heisenberg 1932a, p. 1f), he calculated in the second paper two contradictory results for the binding energy of the electron on a proton as part of a neutron (Heisenberg 1932b, p. 163f). The value of the binding energy according to the mass defect was calculated as about 137 times the value according to the extension of the neutron. The last value was also that which was observed in scattering experiments. Especially in the third paper, Heisenberg compared two possibilities regarding the stability of neutrons, one as not destroyable elementary particle, the other as composed of electron and proton (Heisenberg 1933, p. 594ff). He argues that the first option would be possible, however it would demand electrons in the nucleus and makes some further problems as no longer a valid correspondence between statistics, spin and mass. Therefore, Heisenberg opts in his whole argumentation for the second possibility. Nevertheless, there stay problems, too, especially regarding the stability of the neutron against β-decay. Bohr suspected that because of the continuity of the β-spectrum, the energy conservation is violated (Heisenberg 1933, p. 596). This was curious because a possible solution by Pauli was suggested earlier (Pauli 1930) but published later than Heisenberg’s papers only in 1934. Pauli was a fellow student of Heisenberg and they stayed in close contact all the time. Pauli’s solution was the prediction of a particle hard to discover, later called neutrino by Fermi. (The evidence for the neutrino lasted till 1956). It is hard to imagine that Heisenberg did not know about this prediction.

However, new theoretical particles were hardly accepted at this time as Fermi had to experience in 1934, when his paper explaining the β-decay considering the neutrino was refused by 'Nature' as 'abstract speculation' (Brandt 2010, p. 232). The situation changed for Heisenberg only when the positron was discovered also in 1932. He asked himself if the neutron could be a composition of proton and electron or the proton a composition of neutron and positron otherwise. And further: does it really make sense to say that electrons or positrons are present in the nucleus? Perhaps there are no smaller non-decomposable components. In the end, there is no more splitting, only changing of energy into matter, Heisenberg asked himself (Heisenberg 2010, p. 159).

Heisenberg used as demonstrated several analogies in his explanation of the constitution of the atomic nucleus. All these analogies helped to establish an equation whose mathematical interpretation fixed the qualitative structure of atomic nuclei (the shell model), in general, setting aside the problems of the constitution of the neutron. I think, most essential for his success was the very general explanatory analogy of the harmonic oscillator, which assumes that all atoms, molecules, and atomic nuclei have a type-identity, i.e., are individually indiscernible, as their spectra show. This is a main difference to particles in classical physics. However, this is only an indirect conclusion because particles in or below atomic dimensions are not directly observable[27].

Heisenberg's construction of analogies in his theory is a remarkable complex one. The structure of the nuclei is decomposed in relations between always two nucleons, either two equal or two un-equal even though nearly all atomic nuclei consist of more than two nucleons. This is possible only because the binding energy per nucleon also called 'mass defect' is nearly constant. This fact is due to the observation of a quasi-linear relation between the mass defect and the number of nucleons. However, these are not all the analogies used in the papers. An additional one is the predictive analogy that there are further yet unknown stable or unstable elements and isotopes, which could be calculated from Heisenberg's model. Also, the stability or instability could be predicted (Heisenberg 1932a, p. 7f and 1932b, p. 159f). Here, we find really the typical combination of explanatory and predictive analogies assumed in subchapter 4.1 (see p. 451).

The most important unsolved problems in the explanation above by Heisenberg refer first to the characterization of the binding between the nucleons and second to the properties of the neutron. A suggestion of Yukawa solved the first problem, again by using an analogy (Brandt 2010, p. 250). It was clear, as Heisenberg assumed, that the binding force must fall rapidly to zero at about the radius of a nucleus i.e., the range of that interaction force must be limited to about this radius. The only known force law in the earlier thirty years was the Coulomb or equivalently the gravitational law determined by the Laplace equation

$$\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\psi(\boldsymbol{r}) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi(\boldsymbol{r}) = 0$$

[27] Modern pictures of moved atoms on a grid of atoms as in raster tunnel microscopes are an indirect access, too. These pictures are constructed by a computer from measurements of voltages between an extraordinary thin needle and the atomic grid. The moved atom is hold and released by difference in the voltage. From these values the picture is calculated and visualized.

and a potential inversely proportional to the distance r as solution. This means the electromagnetic (and gravitational) interaction become weaker but have an unlimited range in principle. However, a new force must fall much more rapidly to limit the range to about the size of a nucleon. The steepest known function is the exponential function. Therefore, Yukawa suggested as solution a potential U(r) ~ $e^{-\lambda r/r}$. For this result the Laplace equation must be modified to

$$\nabla \cdot \nabla \psi(\boldsymbol{r}) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \lambda^2\right) \psi(\boldsymbol{r}) = 0.$$

Till this point, the analogy in Yukawa's theory is an explanatory analogy with the hypothesis of the exchange of a boson with the properties described by Heisenberg. It is analogous to a force law with well-known properties and modified to meet the limited range. In a next step, Yukawa adds a predictive analogy applying the so called second quantization to the field U(r) which could be written as a sum of matter waves of a particle with mass $m_U = \lambda$ h/c, h being Planck's constant and c the speed of light in vacuum (Brandt 2010, p. 250). Yukawa obtains a value of about 200 times the electron mass with the radius of the proton as value for the inverse of the parameter λ for m_U. This hypothesis results from de Broglie's (explanatory) analogy taking a particle as a sum of matter waves. Again, here we find the combination of explanatory and predictive analogies.

4.3.2 Theories on Strange Particles

A step on the ladder of scientific development can as mentioned on page 51 in subchapter 4.1 also start with the discovery of unexpected phenomena, as in the case of the strange particles in contrast to the theoretical predicted particle of Yukawa and the therefore expected discovery of the pion. The first to contribute to the understanding of such unexpected particles was Pais in 1952 (see chapter 3.2, p. 40f). In his attempt to find a model explaining this strange behaviour of the then observed particles K^+, K^0, K^- and Λ^0 (at that time all called V particles), he assumed in (explanatory) analogy to the nucleons and pions and their reactions that the lighter K particles are bosons and the heavier is a fermion. The latter heavier particles occur also as positive and negative charged particles (Pais 1952, p. 664). Pais investigated possible combinations of reactions between the old and the new particles first in production and then in decay, supported by the former and the further assumption of an analogous interaction between the new particles analogous to the nucleon-nucleon-pion-coupling. Pais concludes that the evenness or oddness of the sum of all particles before and after the collision does not change because on the one hand the production of strange particles in a strong interaction is initiated by collisions of normal (old) particles and, on the other, generally the production of pairs of strange particles was observed (Pais 1952, p. 664, see also Chapter 3.2, p. 40f). During the decay, however, which is because of the longer lifetime a weak decay, the evenness or oddness must change (Pais 1952, p. 667, see also Chapter 3.2, p. 40f). This rule, called 'even-odd rule', is the hypothesis of an explanatory and a predictive analogy at the same time. The analogy is explanatory because the rule is derived from the behaviour of the well-known 'old' particles and it is predictive because it describes possible interactions with new particles not already observed. Observation in this case means that tracks of charged particles in a cloud chamber could be observed with charges and masses calculated by the density of the drops in the tracks and the deflection of the track by magnetic fields (see chapter 5.1, p. 104). One of the possible reactions was observed two times two years after Pais's publication, namely the reaction π- + p -> Λ_0 + K_0 (Brandt 2009, p. 341).

A year later, Gell-Mann supposed that the isospin formalism can be applied to the strange particles in the following form (see chapter 3.2, p. 41) based on Pais[28] and a suggestion of Peaslee, who assumed that charge independence 'may extend to the new unstable particles as well' (Gell-Mann 1953, p. 833):

- Strong interactions conserve isospin, ΔI = 0.
- Electromagnetic interactions can change isospin by an integer value.
- Weak interactions do not respect isospin at all. (It was found that very frequently |ΔI| = ½.)

Gell-Mann suggested to fulfil these conditions in contrast to Pais, who assumed in some not clear way that the isospin has value one and negative parity (see footnote above), alternatively the hypothesis that contrary to the 'old' particles the new fermions should have integral isospin and the bosons half integral isospin. The relation to Pais's ideas is unclear also because according to Gell-Mann his alternative became clear to him by a slip of the tongue on a talk about the curious slow decay of the new particles (Gell-Mann 1994, p. 370f). Independent from this question, Gell-Mann's suggestion led him to the conclusion that the fermions (in Gell-Mann's paper called V_1^+, V_1^0, V_1^-, today Σ^+, Λ^0, Σ^-) should form a triplet and the bosons (in Gell-Mann's paper called τ^+ and $V_4^0 = \tau^0$, today K^+ and K^0) a doublet, with each also existing as anti-particle (Gell-Mann 1953, p. 833f). Gell-Mann could show with these assumptions altogether that production and decay of the new particles fit perfectly to the observed phenomena and that Pauli's exclusion principle is warranted. Curious is that the z component of isospin is stronger conserved than the charge of the participating particles while nevertheless the total charge of each side of a reaction is conserved (Gell-Mann 1953, p. 834).

We see that besides an explanatory analogy regarding the bosons in relation to the nucleons here an unnamed explanatory analogy is used, built up by parts of the combination of the two independent models offered by Pais and Peaslee, and intended to bring the theoretical model, namely the positive analogies in agreement with the observed phenomena. This was only possible because both models possessed features (potentially positive analogies) inspiring Gell-Mann besides others, which are not satisfactory (Peaslee's model could not exclude rule (2) above in a satisfying way, objections to Pais's model could not clearly be grasped as mentioned). None of the postulated objects in all three models could be observed directly, only the tracks of the charged particles could be observed indirectly in the way described above (see p. 74). This illustrates how analogies can help to develop more satisfying theoretical models of unobservable objects. A predictive analogy was not involved in this model.

This was only a first step of theory development in the case of Gell-Mann. Together with Pais, who was invited to hold a talk at a conference and asked Gell-Mann to contribute as co-author, he wrote a report 'mostly (except for the particle names) already presented in an earlier unpublished report' by him (Brandt 2009, p. 342). Four alternative theoretical models on two argumentation lines were

[28] Gell-Mann refers to Pais's paper from 1952 and with a somewhat unclear description to an unpublished paper.

compared in this report to explain the long lifetimes of the new particles and their consequences on possible further particles. It was published in the proceedings of the conference (Gell-Mann & Pais 1955, p. 342ff).

The first and only model of the first line by Fermi and Feynman (Gell-Mann & Pais 1955, p. 343ff), unpublished before, assumes the new particles to be resonances of e.g., nucleons and pions with a high angular momentum (i.e., a compound of these particles, see chapter 3.2, p. 39). An objection to this model was that it allows some reactions which are not observed so far in a sufficient number to accept it as single account. This was the first line of argumentation based on 'the strong variation of energy with matrix elements' of the S-matrix (Gell-Mann & Pais 1955, p. 345, see also chapter 3.1, p. 38 for the S-matrix).

The second line with three models takes up the strand of argumentation of Pais and Gell-Mann mentioned in the paragraphs directly above, i.e., associated production of the new particles in a strong interaction and slow decay in a weak interaction. This phenomenon is like the pion and the muon, which are produced also in a strong interaction and decay in a weak interaction (Gell-Mann & Pais 1955, p. 345). However, the slow decay only is an indication of this strand of arguments because a kind of production is thinkable where it can be concluded that θ^0 and its anti-particle are not identical[29] and which could not be excluded (Gell-Mann & Pais 1955, p. 346).

Here, an analogy of the strange particles to the well-known particles nucleon and pion was introduced. In the following discussion of the paper this analogy appears as a paradigmatic case of an explanatory analogy, as the following longer citation shows (Gell-Mann & Pais 1955, p. 346f):

"Now we have seen in approach (B) [the second strand of argumentation] severe restrictions are imposed on which reactions may be fast. There is at least the requirement of associated production, and there may be more rules as well. The theoretical physicist cannot rest content, if he is following up this point of view, until he has connected these new and phenomenological rules somehow with more familiar physical concepts. Apparently, we have to deal with selection rules that hold true for the class of strong interactions but are violated by the very weak decay interactions. A somewhat similar situation occurs, of course, in the domain of pion and nucleon physics. There is the principle of charge independence or conservation of isotopic spin [isospin] which hold to be true for the strong interactions between pions and nucleons, but which is violated in certain ways by electromagnetic effects and perhaps also, for all we know, by the very weak interactions associated with β-decay. One is tempted therefore, first, whether charge independence apply to somehow the new particles, and second, whether it is connected with the new selection rules. [This last sentence is the hypothesis of the analogy.]

[29] This is an old name for the kaon K^0 for which its anti-particle is not identical really, however, the K^0 can change in its anti-particle called 'oscillation' (see e.g., Berger 2006, p. 147ff). Therefore, a beam of kaons is today interpreted as a superposition of both kinds changing with time and would be seen at the time of the publication of Gell-Mann and Pais as two particles, one shorter living, decaying primarily in two pions (the θ^0), and a longer living, decaying in three pions (called τ^0). A more carefully investigation of this curious behaviour led to the discovery of the non-conservation of the charge-parity-invariance of the weak interaction. However, this is not the focus here.

If charge independence extends to the new particles we would expect to observe charge multiplets among them (similar to the nucleon-doublet and the π-triplet), that is, sets of particles with different charges but approximately the same masses and similar strong interactions. There is some evidence for this [the positive analogy of the first point]: the striking similarity of the masses of the θ^0 and $\tau^\pm$ mesons and apparent existence of both a positive and a negative hyperon which we shall call $\Sigma^\pm$ with the decay schemes [here is an omitted footnote with reference to observations by C.C. Butler published in the same proceedings],

$\Sigma^\pm$ -> N + $\pi^\pm$ + (~115 MeV)

$\Sigma^\pm$ -> P + π^0 + (~115 MeV).

If it should turn out that the spins or parities of the θ^0 and $\tau^\pm$ are different then the evidence would be greatly weakened [defeating possible negative analogy]; of course if a τ^0 and a $\theta^\pm$ should also exist, with the same mass, then there would be evidence of two charge multiplets lying remarkly close together by energy [further positive analogy of the first point in the hypothesis].

We know that; besides charge multiplets, the conservation of isotopic spin should yield selection rules. Following approach (B), we will want to know whether the new rules we need are among those given by charge independence. This is impossible, however, unless we introduce some new feature [extension of the old model in the new domain]. The reason is that in the pion-nucleon system [the old model] isotopic spin selection rules are violated by electromagnetic effects whereas the rules that we need for the new particles must not be violated except by very much weaker interactions. [omitted footnote with reference to Peaslee's paper mentioned above, see p. 83]

From the present point of view the situation looks therefore as follows: there are three kinds of interactions, [1][30] the strong coupling responsible for pion production and absorption, the copious associated production of the new particles and the binding of hyperons in nuclear matter; [reference to observations of C.F. Powell published in Nature 173 (1954), p. 469, and P. Ciok, M. Danysz & J. Gierula in Nuov. Cim. 11, p. 436] [2] electromagnetic interactions and the effects of mass differences within charge multiplets; [footnote in original: We will not take a stand here on the question whether or not such mass differences are of electromagnetic origin.] [3] very weak interactions inducing hyperon and boson decays and possibly similar in nature to β-decay interactions.

Correspondingly there appear at least three classes of selection rules or conservation laws: the first consists of absolute laws like the conservation of charge and the conservation of baryons (nucleons and hyperons) which are not violated by any interactions. The conservation of isotopic spin constitutes a second class: this law is presumably valid for interactions of type [1], but is violated by interactions of type [2] and possibly of type [3] as well. Finally the third class comprises the new laws required for the metastability of the new particles, such as associated production. These laws must be valid for interactions of types [1] and [2] and violated by interactions [3]. [positive analogy of the second point in the hypothesis]"

[30] This is the enumeration in the original publication.

Three different kinds of assumptions were made to fulfil the technical demands of the extended positive analogies mentioned in the citation above, corresponding to the three variations of the strand of associated production.

The first of three these model variations by Pais assumed that the extended isospin has three components; an absolute value labelling a multiplet, an angular momentum corresponding to the different charges of a multiplet and a parity needed for the laws of the third class. This means the isospin space of rotations is extended to also allow reflections, which are demanded to hold in strong interactions and securing metastability for new particles (Gell-Mann & Pais 1955, p. 348). As consequence, neither pions nor photons can change the isospin. Assigning even parity to the old and odd to the new particles can explain most however unfortunately not all new particles. For the cascade particle (in the same paper later called Ξ^-, see also chapter 3.2, p. 40) this approach fails. Presumably, this approach is meant by the unclear description in Gell-Mann's paper of 1953 (see above, p. 83).

The second strand is Gell-Mann's own attempt to extend the model of nucleons and pions with the new particles obeying the stated rules. He suggests relaxing the 'usual connection between the charge Q and the Z-component I_Z of the isotopic spin (isospin) for members of a charge multiplet', i.e. the formulae Q= I_Z +1/2 for nucleons respective -1/2 for their anti-particles and Q= I_Z for pions was extended to Q= I_Z for Λ^0 (I=0 singlet) and $\Sigma^\pm$ and a hypothetical Σ^0 (I=1 triplet) and Q= I_Z -1/2 for Ξ^- and a hypothetical Ξ^0 (I=1/2 doublet) (Gell-Mann & Pais 1955, p. 348). Thus, we have three conserved quantum numbers Q, I, and I_Z as with Pais if we assume with Gell-Mann that the hyperons are metastable for interactions of type [1] (strong) and [2] (electromagnetic and mass-difference effects). Perhaps the idea of three quantum numbers was what Gell-Mann got from Pais in his paper from 1953. It is remarkable that he predicts two new particles in his account. To understand that this is a predictive analogy (hypothesis associated from positive analogies) one must remember that the I=1 triplet of Σ -particles is an analogy to e.g., the triplet of pions and the I=1/2 doublet of Ξ -particles is an analogy to the nucleons. More general, it is an analogy to the theoretical multiplet structure introduced by Heisenberg about thirty years ago (see p. 76f).

It should be mentioned that both I and I_Z change with ±1/2 in weak decays of hyperons and only Ξ^0 can decay also into a Σ^0 by emitting a photon.

Still open in this strand is the question of bosons. Gell-Mann assumes for these an isospin doublet B^+ and B^0 with corresponding anti-particles like the case of the nucleon proton and neutron (Gell-Mann & Pais 1955, p. 349; again, an explanatory analogy). He used B as name because it was not clear at that time if the bosons θ^0 and $\tau^\pm$ belong to the same family or if they form two families (Gell-Mann & Pais 1955, p. 350, see p. 84 and footnote 29).

The third strand in this line is again from Pais. There, he extends the isospin space to four dimensions with two independent isospins of the ordinary type (i.e., that of the nucleons), but no reflections occur (Gell-Mann & Pais 1955, p. 350). For nucleons, the second type is always zero and because of $|I_1\text{-}I_2|$=1/2 one should have Q= I_{1z}+I_{2z}+1/2 and should assign (0,1/2) to Λ, (1,1/2) to Σ, and (1/2,1) to Ξ. It can be shown that these assignments are analogous to the atomic shell model in the following way $^2S_1/2$ <-> (1/2,0), $^2P_1/2$ <-> (0,1/2), $^2P_3/2$ <-> (1,1/2), $^2D_3/2$ <-> (1/2,1) etc., so that the Λ will become a doublet like the nucleons, the Σ and Ξ sextets, both with one double positive charged particle. It gives some evidence at that time for such last-mentioned particles, (reference in the paper to G. Ascoli (1953), Phys. Rev. 90, p. 1079). A lot of more particles and decays by emitting photons occur in this model, which could distinguish between the last two models.

All four strands of argumentation for explaining the metastability of the new particles try to extend the accepted model of the nucleons and pions by a positive analogy comprising the new particles, too. However, the last three strands were based more on the (material) properties of the applied and accepted model of nucleons and pions.

The next (successful) step of development of his extended model of old and new particles was also made by Gell-Mann. He clarified in a summary of his model (Gell-Mann 1956) first the concepts of particle and anti-particle, then of the four groups of observed particles as baryons (nucleons, hyperons and their anti-particles), leptons (muons, electrons, neutrinos and perhaps fermionic K particles[31]), mesons (pions and heavier bosons) and at last photons. Considering this, Gell-Mann specified the types of interaction as strong (and rapid) for the interactions between baryons, mesons, and their anti-particles 'responsible for the nuclear forces and the production of mesons and hyperons in high energy nuclear collisions' (Gell-Mann 1956, p. 849), as electromagnetic (with intermediate lifetime) for interactions of charged particles with real or virtual photons, and at last as weak for slow decays (β-decays, that of muons, of hyperons and K-mesons) and the absorption of muons in matter. Then, he used in an unnamed formal analogy to Heisenberg the method of the so called 'approximation' to demonstrate the effect of the three types of interaction (see p. 79 in chapter 4.3.1). In first approximation only the strong interaction is active, then in the second the electromagnetic interaction is switched on and at last the weak interaction. Of course, this is an explanatory analogy but an unusual one, fitting not so well in Bartha's classification because its object is not the type of argumentation but the application of a method. Therefore, it seems that it supports the wider concept of analogies by Hofstadter & Sander (and Mach) than only analogical reasoning does, as with Bartha (and Hesse).

Back to Gell-Mann's summary, he describes in first approximation i.e., with only effective strong interactions, the concept of charge independence. Charge independence is also a concept introduced by Heisenberg used in his isospin (see p. 77). Charge independence means that in the definition of Heisenberg the total isospin I i.e., the amount is conserved in a multiplet as in the doublet of proton and nucleon, whereas the charge changes with one unit and with the charge the value of the Z-component of the isospin. The isospin in a doublet is one half and using the formula for a multiplet 2I+1 we get the two components of the doublet. Thus, we get the relation $Q/e=I_z+n/2$ where n stands for the sum of particles minus anti-particles. This formula is valid for all old particles, inclusive atomic nuclei with the centre of charge equal to n/2 for all multiplets of nucleons and pions.

In the first approximation, i.e., only the strong interaction switched on, it is assumed that all particles of a multiplet have the same mass. This changes with the second approximation, when the electromagnetic interaction is switched on, then the masses differ slightly within a multiplet. With

[31] In fact, fermionic K particles do not exist.

the third approximation all known phenomena are possible, like β-decay and the decay of pions into muons and neutrinos.

In the extended model with strange particles in first approximation it is also assumed that charge independence is strictly conserved as with nucleons and pions, however, the centre of charge is modified to $Q/e=I_z+n/2+S/2$ with integral S varying with the multiplets and S=0 for nucleons and pions (this is the hypothesis). Gell-Mann named S as strangeness because only for these S is not zero (Gell-Mann 1956, p. 852). The strangeness must stay conserved also in the second approximation i.e., in electromagnetic interactions because they cannot decay rapidly in ordinary particles i.e., nucleons or pions. They must have been produced with the sum of strangeness equal to zero because of the associated production. It is important, however, that virtual dissociation of the new particles is provided according to Yukawa's hypothesis like that of a neutron into a proton and a pion cloud (see p. 81f). Only then, the so-called minimal coupling of the electromagnetic field meets the demands of the conservation of strangeness i.e., the substitution $\frac{\partial}{\partial x_\mu} \rightarrow \frac{\partial}{\partial x_\mu} - iQA_\mu(x)$ like that in the Schrödinger equation in an atom (an explanatory analogy between unobservable objects) (Gell-Mann 1956, p. 853). In the third approximation, of course, strangeness is not conserved anymore.

From these conditions (the hypothesis of an explanatory analogy) a singlet Λ^0 (I=0), a triplet $\Sigma^\pm$ with a hypothetical Σ^0 (I=1), a doublet Ξ^- with a hypothetical Ξ^0 (I=1/2) and two meson doublets θ^+, θ^0 and τ^+, τ^0 and their anti-particles can be derived (see p. 83, the positive analogies to nucleon and pion multiplets). The Λ^0 must be a singlet because no other charged particle with about this mass was ever found and the Σ particles are about 150 MeV heavier. The Σ particles must be a triplet because no double charged particles with about this mass were found. For the Ξ^- particle there are two possibilities to build up a multiplet because there were no positive or double charged particles found. First, it can be a singlet with strangeness minus three or a doublet with a neutral particle as mentioned above. However, together with a new and simple rule that the strangeness could only change with one unit in agreement with the observations, one must assign strangeness minus two and therefore one gets a doublet. For the heavier mesons, the only possibility is to assign a strangeness plus and minus one and a doublet to them because of their lifetime and that there exist no multiple charged mesons of this type. The two families θ and τ were assumed because of their different decay behaviour into two or three pions, respectively, as mentioned above. Gell-Mann justified also the metastability of the mentioned particles, however, this will not be repeated here. It is obvious that always the model of multiplets developed by Heisenberg as positive analogy is applied here (see p. 76f).

Based on these observed and postulated particles and their properties, Gell-Mann predicted a lot of allowed and forbidden interactions (Gell-Mann 1956, p. 859ff). This is a predictive analogy, where from the positive analogies of the observed and expected properties of the particles the hypothesis of the interaction processes is assumed. I will not list all Gell-Mann's interaction processes here but will mention the hypothesis of a hyperon Z^+ with strangeness plus one, which, if it exists, could be bound to a nucleus in a so called 'hyperfragment' instead of a heavy meson (only mesons could bind, however, not their anti-particles because of their opposite charge) or as already known a Λ^0 particle (Gell-Mann 1956, p. 862). This hypothesis shows that the full structure of the possible multiplets was still not understood at that time because hyperons with positive strangeness were never found. However, a further prediction was a hyperon with strangeness minus three called Ω^-, made in an appendix to the paper and mentioned in principle in connection with the Ξ particles. Both these hypotheses assume that, if there were a triplet and two doublets with a strangeness one unit higher respective lower, there should be a singlet with again one unit higher or lower. However, only the last, the Ω-, would be found later which results in the end in a somewhat different structure as we know now (see Table 4, p. 46f). In case of the mesons the same consideration led to the assumption of two singlets ω^+ and ω^- with strangeness plus and minus two, which were also not found. Only another meson was later found, also called ω and also a singlet but neutral and with a mass in the range predicted[32] (Berger 2006, p. 264).

In a further step to understand the structure of the multiplets and to include a lot of further new particles discovered in the meantime, Gell-Mann used field theoretic methods to investigate symmetries of the baryon-meson system (Gell-Mann 1962). It was based on a former developed simpler model (Gell-Mann 1957) despite the field theoretic approach came under criticism at that time (Gell-Mann 1962, p. 1068). This criticism argues that there are no 'elementary' particles, and all particles are on the same footing (the so-called 'bootstrap' approach).

A symmetry means here that e.g., proton and neutron are indistinguishable against exchange and behave in the same way in first approximation with electromagnetic and weak interaction switched off as introduced by Heisenberg with isospin (see chapter 4.3.1, p. 77). In the simpler basic model Gell-Mann used the (explanatory) analogy of the strong and electromagnetic interaction with the nucleons to postulate a very strong (VS) and a moderate strong (MS) interaction for both the nucleons and the new particles Ξ, Σ and Λ known in 1957. With only the VS interaction on and the MS interaction switched off, these eight particles form a 'supermultiplet' i.e., one particle with eight degenerated i.e., undistinguishable states. This 'supermultiplet' splits into the multiplets described above (see p. 86) when the MS interaction is switched on. Gell-Mann suggested that the VS coupling between the baryons is mediated by the pions and the weaker MS coupling by the K-particles (Gell-Mann 1957, p. 1299). The doublet of the Ξ particles is directly analogous to the nucleons with this assumption. For the Σ and Λ particles, Gell-Mann uses the trick to introduce fictitious particles $Y^0 = (\Lambda^0 - \Sigma^0)/\sqrt{2}$ and $Z^0 = (\Lambda^0 + \Sigma^0)/\sqrt{2}$, which each were combined with one of the charged Σ particles to build up two doublets in analogy to the nucleons. Thus, all baryons, including the strange ones, are thought to be coupled in the same way to the pions via the VS interaction i.e., they are symmetric. Estimating the MS interaction is somewhat more difficult because it is asymmetric. However, Gell-Mann argues (Gell-Mann 1957, p. 1300) the remaining interactions behave like that of the nucleons with pions because there are no interactions possible between the two fictitious particles Y^0 and Z^0. This leads to the result that the interaction with K mesons is weaker (has a smaller cross section) than with pions but they are otherwise similar. In a footnote Gell-Mann mentions the relation $(m_N + m_\Xi)/2 = (3m_\Sigma + m_\Lambda)/4$, which gives some information on the strength of the MS interaction,

[32] Additionally, two further neutral meson singlets with a mass in the same range (η and η') and one somewhat heavier (φ) were found later.

investigated in the paper from 1962 in more detail (Gell-Mann 1962). However, the former parts of the paper must be considered to understand the meaning of this formula before discussing it.

4.3.3 The Way to Quarks

The aim of the paper is 'to clarify the meaning of such possible symmetries [introduced in the paper of 1957), for both strong and weak interactions' (Gell-Mann 1962, p. 1068) or in other words in this last paper Gell-Mann extended the model of weak interactions between leptons to weak interactions between leptons and hadrons in several analogical steps. To do this Gell-Mann used a field theoretic approach despite it was called by some as 'a misleading encumbrance' (Gell-Mann 1962, loc. cit.). Gell-Mann defends his way of handling the fact that the introduced field operators are well defined independent from the field theoretic formulation. To avoid the subtleties of the field theoretic derivation I will cite here the abstract from this article (Gell-Mann 1962, p. 1067), which gives a short summary of this argumentation, and which will be explained where necessary in the following:

"The system of strongly interacting particles is discussed, with electromagnetism, weak interactions, and gravitation considered as perturbations. The electric current j_a, the weak current J_a, and the gravitational tensor $\theta_{\alpha\beta}$ are all well-defined operators, with finite matrix elements obeying dispersion relations. To the extent that the dispersion relations for matrix elements of these operators between the vacuum and other states are highly convergent and dominated by contributions from intermediate one-meson states, we have relations like the Goldberger-Treiman formula and universality principles like that of Sakurai according to which the ρ meson is coupled approximately to the isotopic spin. Homogeneous linear dispersion relations, even without subtractions, do not suffice to fix the scale of these matrix elements; in particular, for the nonconserved currents, the renormalization factors cannot be calculated, and the universality of strength of the weak interactions is undefined. More information than just the dispersion relations must be supplied, for example, by field-theoretic models; we consider, in fact, the equal-time commutation relations of the various parts of j_4 and J_4. These nonlinear relations define an algebraic system (or a group) that underlies the structure of baryons and mesons. It is suggested that the group is in fact U(3)xU(3), exemplified by the symmetrical Sakata model. The Hamiltonian density θ_{44} is not completely invariant under the group; the nonvariant part transforms according to a particular representation of the group; it is possible that this information also is given correctly by the symmetrical Sakata model. Various exact relations among form factors follow from the algebraic structure. In addition, it may be worthwhile to consider the approximate situation in which the strangeness-changing vector currents are conserved and the Hamiltonian is invariant under U(3); we refer to this limiting case as 'unitary symmetry'. In the limit, the baryons and mesons form degenerate supermultiplets, which break up into isotopic multiplets when the symmetry-breaking term in the Hamiltonian is 'turned on'. The mesons are expected to form unitary singlets and octets; each octet breaks up into a triplet, a singlet, and a pair of strange doublets. The known pseudoscalar and vector mesons fit this pattern if there exists also an isotopic pseudoscalar meson χ^0. If we consider unitary symmetry in the abstract rather than in connection with a field theory, then we find, as an attractive alternative to the Sakata model, the scheme of Ne'eman and Gell-Mann, which we call the 'eightfold way'; the baryons N, Λ, Σ, and Ξ form an octet, like the vector and pseudoscalar meson octets, in the limit of unitary symmetry. Although the violations of unitary symmetry must be quite large, there is some hope of relating certain violations to others. As an example of the methods advocated, we present a rough calculation of the rate of $K^+ \to \mu^+ + \nu$ in terms of that of $\pi^+ \to \mu^+ + \nu$."

Let me discuss the statements in the order they occur. The strong interaction is the dominating one because of its strength and for this reason the other interactions can be handled as perturbations of the strong interaction. Well-defined operators guarantee that every variable which the operator works on has a unique value i.e., that the operator is a function. Dispersion relations describe the relation between the real and the imaginary part of e.g., a complex wave function or in this case a scattering amplitude and can be used to compute the unknown or unobservable part of this relation by the other already known. The matrix elements refer to the S-matrix (see chapter 3.1, p. 38).

The real or imaginary parts of the wave functions in question are integrals over the time where for reasons of causality only values of the past should have an effect. If then the integral's Fourier transformation, i.e., the development of the integral into the contained frequencies, goes in the limit fast enough to zero with the frequency going to infinity, i.e., the development is highly convergent, then we have the expected relation between real and imaginary part and one can calculate one from the other. With this condition and the assumption that the mesons, thought as mediating the strong interaction, are the same particles also decaying in the β-decay of atomic nuclei, one can connect the meson states with so-called 'currents' or 'densities' (Gell-Mann 1962, p. 1968). To understand what is meant by 'currents' and 'densities' I must go back a bit and introduce some knowledge of weak interactions neglected so far. Nevertheless, I must start with quantum electrodynamics (QED) because the first version of the weak interaction was developed in analogy to QED.

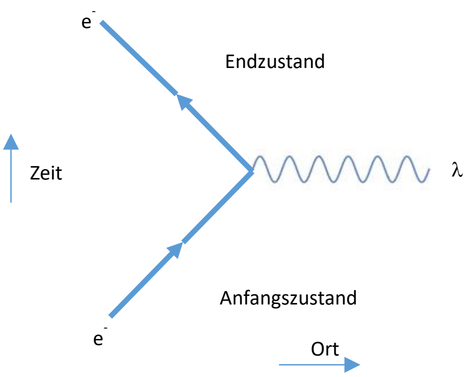
There are three energy domains in quantum physics, the first in the range of several eV (electron volt, an energy measure adapted to the quantum domain), applicable to the processes dealing with atoms and molecules handled by quantum mechanics and later by QED, the second in the range of MeV (millions of eV), applicable to processes of the atomic nucleus as described by Heisenberg, and third in the range of GeV (billions of eV), applicable to elementary particles, i.e. in the domain of HEP. To do this investigation in HEP scattering experiments were executed. The results of these scattering experiments were subsumed in the so-called S matrix (scattering matrix, see chapter 3.1, p. 38), which contains theoretically all possible combinations of outcoming particles depending on the incoming particles before their interaction takes place, actively prepared for the experiment and therefore well-known. A scattering experiment is like firing a cannon ball on a target (e.g., another incoming cannon ball with the same velocity in opposite direction) and registering always all fragments of all such collisions. Each type of collision event can be characterized by an entry, an element in the S matrix. The relation between an incoming and an outgoing particle is called a current after the model of electric current where the electric circuit must be closed, i.e., the number of electrons leaving the source on one pole must be the same as the number of the returning electrons on the other pole.

The current in the electric example is generated by some difference of voltage, which is the source of an electric field (applied to an electron this gives the energy of some eV). This electric field has accelerated the electron, which theoretically then can fly free with constant velocity and therefore constant energy without needing the electrical field (corresponding to the second law of Newton).

This flying electron (i.e., moving in relation to the observer) is also a current and, because a current is always accompanied by a magnetic field, the current is in fact called electromagnetic current. However, more interesting is the reaction if an outer electric (or magnetic) field is present because then the electron changes direction and velocity. Then we have the incoming electron, the interaction with the electric field, and the outgoing electron with changed momentum and energy. In QED, this electric field is quantized, i.e., it consists of photons which behave like particles. The interaction between two charged particles like electrons outside and protons within the atomic core via a photon governs all reactions between atoms and molecules. An atom can be imagined with some restrictions like our solar system with the sun (the protons together with the neutrons) in the centre and the planets (the electrons) orbiting around it (an imagination used by Bohr in his first model of the hydrogen atom however with the modification of discrete orbits). This is the first domain in the range of some eV.

The interaction term takes the form of $-e\bar{\varphi}(x)\gamma_\alpha\varphi(x)A^\alpha(x)$ where -e stands for the coupling constant (here the negative charge of the electron), $\bar{\varphi}(x)\gamma_\alpha\varphi(x)$ for the current, $\bar{\varphi}(x)$ for the outgoing (or generated), $\varphi(x)$ for the incoming (or annihilated) particles (both electrons in this case) and $A^\alpha(x)$ for the electromagnetic field (the photons). The factor γ_α is a set of four 4x4-matrices introduced in the Dirac equation and constructed out of the three Pauli matrices and the two-dimensional unit matrix, which relates the different components of the incoming and outgoing particle description (two spinors[33]) to each other building up the current. The electric field is coupled by the coupling constant to the current. The interaction description in QED represented as photon can be thought of now as applied to the incoming electron as acceleration or to the outgoing electron as slowing down or read in the opposite direction as the interaction of a positron, the positive charged anti-particle to the electron, with the photon. However, a more complicated picture is possible, which we explain when we have introduced Feynman graphs (see figure 1): -e -> -e + λ

Figure 1: Interaction of an electron with a photon



Since Feynman one can illustrate the interaction term with a graph. The two continuous lines represent the current, i.e., the incoming particle below and the outgoing particle above, what means the time is going up, the wavy line represents the photon. In QED, the particles are electrons (arrow

[33] A spinor is a special type axial vector with half the angle of an axial vector. An axial vector is a special product of two polar vectors. Polar vectors are entities defined by a value and a direction in a space. Axial and polar vectors differ in their behaviour in transformations.

up) or positrons (arrow down, anti-particles have the opposite charge and opposite values of some other conserved properties), which are fermions with half valued spin (imaginable with some restrictions as a rotation around their axis). Positrons were detected in 1932 in the same year as the neutron (Brandt 2009, p. 210ff and 214ff) and the neutrino as anti-particle to the anti-neutrino 1956 (Brandt 2009, p. 350ff).

This model of interaction was used by Fermi 1933 (in Italian) respective 1934 (in German) to develop a theory of β-decay replacing the generated photon in QED by the simultaneous creation of an electron and a neutrino if a neutron, as Heisenberg's theory of the structure of the atomic nucleus postulated, would transform into a proton (Brandt 2009, p. 232, Fermi 1934). Thus, the interaction term of the QED was replaced by the term $G(\overline{\phi}_p(\mathrm{x})\gamma_a\phi_n(\mathrm{x}))(\overline{\phi}_e(\mathrm{x})\gamma_a\phi_\nu(\mathrm{x}))$ where G is a coupling constant like electric charge, $(\overline{\phi}_p(\mathrm{x})\gamma_a\phi_n(\mathrm{x}))$ the term substituting the electron current with $\overline{\phi}_p(\mathrm{x})$ and $\phi_n(\mathrm{x})$ as wave function of a proton respective neutron, and $(\overline{\phi}_e(\mathrm{x})\gamma_a\phi_\nu(\mathrm{x}))$ the term substituting the electromagnetic field with $\overline{\phi}_e(\mathrm{x})$ and $\phi_\nu(\mathrm{x})$ as wave function of an electron respective neutrino[34]. In contrast to the electromagnetic field two components interact here, an electron and a neutrino instead of only one photon. The opposite process changing a neutron in a proton by annihilating an electron and a neutrino is described by the complex conjugate of the former term. A similar but more probable process is the catching of an electron from the K shell of an atom changing a proton into a neutron and emitting a neutrino. In 1947 Pontecorvo noted that instead of an electron perhaps a meson with half spin, i.e., a particle today called muon, could take its place. However, the muon must decay then according to Fermi's process into an electron and two neutrinos (Pontecorvo 1947, p. 246). This assumption was confirmed as probable in 1949 by Steinberger (Steinberger 1949, p. 1136). In consequence of these results, Fermi's theory was revised by Tiomno and Wheeler indicating but not guaranteeing that the coupling constant for the muon decay and β-decay is identical and the coupling has vector form[35] (Tiomni & Wheeler 1949, p. 152) as already used in the formula above. They assumed that the coupling is universal because of all these similar forms of decay.

The next step in the development of the weak interaction was related to the puzzle about the mysterious mesons θ and τ; they seem to be nearly identical, however, decay in two or three pions, respectively, i.e., seem to have different parities (see chapter 4.3.2, p. 84f). This fact led Lee and Yang to the hypothesis that the weak interaction violates the conservation of parity (Lee & Yang 1956, p. 254), which was confirmed by an experiment of the team of Wu (Wu et.al. 1957, p. 1413ff). One year later they postulated a new theory that neutrinos exist only with one spin direction, compatible with parity violation in weak interactions and, therefore, only two of four components of the spinor[36] are

[34] This is not the original formulation by Fermi (see Fermi 1934, p. 165), however, with this formulation it is easier to understand the analogy.
[35] The form of coupling is possible only in five forms or combinations of them, as scalar staying unchanged under space inversion, as (polar) vector like a distance including its direction changing into the opposite direction, as pseudo vector (or axial vector) defined as product of two vectors staying unchanged in its direction, a pseudo scalar defined as scalar product of a polar and pseudo vector changing its sign under space inversion, and at last as tensor defined as a multilinear form of vector products changing its sign depending on the degree of the multilinear form.
[36] A spinor is like a vector but transforms in another way.

unequal to zero (Lee & Yang 1957, p. 1671). On the following page they defined also, valid till today, that the neutrino has a spin parallel to its momentum and the anti-neutrino a spin opposite to its momentum. Additionally, also violation of charge conservation occurs because of the relation between particles and their anti-particles, so that only the combination of parity and charge conjugation together are valid. Lee and Yang introduced to express these conditions in the mathematical formulation a projection operator $\frac{1}{2}(1-\gamma_5)$[37] into the formula $\frac{G}{4}(\overline{\phi}_p(\mathrm{x})\gamma_a(1-\gamma_5)\phi_n(\mathrm{x}))(\overline{\phi}_e(\mathrm{x})\gamma_a(1-\gamma_5)\phi_\nu(\mathrm{x}))$ which selects only particles with a left-handed spin because only these participate in the weak interaction. Alternatively, an operator $\frac{1}{2}(1+\gamma_5)$ can be introduced, which would pick out the experimentally excluded particle with opposite spin. The terms $(\overline{\phi}_p(\mathrm{x})\gamma_a(1-\gamma_5)\phi_n(\mathrm{x}))$ and $\left(\overline{\phi}_e(\mathrm{x})\gamma_a(1-\gamma_5)\phi_\nu(\mathrm{x})\right)$ are called 'charged currents' in analogy to the 'electromagnetic current' in QED, however, charged because of the change of charge from neutron to proton or (anti-)neutrino to electron, respectively. Of course, both 'currents' must be equal because of the charge and other conservation rules. However, Feynman and Gell-Mann found another explanation for the only two components of the spinor describing neutrinos and anti-neutrinos (Feynman & Gell-Mann 1958, p. 193f). They argued that the usual four components of the spinors in the Dirac equation are not necessary but only two describing the spin if the equation is changed like the Klein-Gordon equation into an equation of second order. Then, the sign of the energy depends on the particle, electron or positron, respectively, and one gets the same relation for these particles as Lee and Yang for the neutrino (Feynman & Gell-Mann 1958, p. 193). Therefore, they assumed that all wave functions should have the factor $\frac{1}{2}(1-\gamma_5)$, anti-particles the other sign; and the interaction is valid as well for β-decay and the decay of muons i.e., is 'universally' valid. This theory was called 'V-A theory' of weak interaction because the part $(\overline{\phi}_p(\mathrm{x})\gamma_a\phi_n(\mathrm{x}))$ behaves like a vector and the part $(\overline{\phi}_p(\mathrm{x})\gamma_a\gamma_5\phi_n(\mathrm{x}))$ like a pseudo or axial vector. Feynman and Gell-Mann argued that the vector part is conserved in analogy to the electromagnetic current, where the coupling constant e to the electromagnetic field is the same for all charged particles despite that nucleons can emit virtual pions (Feynman & Gell-Mann 1958, p. 196). To achieve this, they assumed that an additional current term $i[\varphi^* T_+ \nabla_\mu \varphi - (\nabla_\mu \varphi) T_+ \varphi]$ of the pions occurs, which generates further possible weak interaction processes like the decay of π^- into π^0+e+$\bar{\nu}$ (a predictive analogy). One can 'imagine that the interaction is due to some intermediate (electrically charged) vector meson of very high mass' (Feynman & Gell-Mann 1958, loc. cit.). For the axial vector part, however, it remained unclear if it could be arranged in an equivalent way (Feynman & Gell-Mann 1957, p. 197). One year later, Taylor remarked that the pion could not decay in an electron or muon and a neutrino if the axial vector part is conserved because the product of the momentum of the pion with the axial vector part must then be zero (Taylor 1958, p. 1216). From this fact, Goldberger and Treiman concluded that the axial vector is not conserved and derived a dispersion relation which was too large to agree with experimental results (Goldberger & Treiman 1958, p. 1478f). However, Gell-Mann could show that the altering (the divergence) of the non-conserved axial vector part at low energies is 'dominated by contributions from intermediate one-meson states' (see citation of the abstract of

[37] Lee & Yang and Feynman & Gell-Mann defined the γ matrices in a somewhat different way, but I have harmonized them in my formulations to the modern notation.

Gell-Mann 1962 on my p. 90f) and that the relation between the axial vector part and the vector part at low energies is the Goldberger-Treiman relation. This relation fixes only their relative strengths and not their absolute values, therefore, another method must be used to fix them.

The method Gell-Mann applied is the evaluation of equal-time commutation relations, which because of their non-linearity have solutions only for special values. The equal-time commutation relations are well-known from the electronic spin, i.e., the fact that the vector or wedge product according to the commutation rule of two spinors gives a third one orthogonal to both other. Gell-Mann extended this scheme by another component with similar rules. As mentioned, spinors are like vectors, however, transform in another way than vectors (see footnote 36 on p. 93). Normally in case e.g. of rotations in three dimensional space, the system of vectors or spinors, respectively, is closed i.e. there are not more than three orthogonal vectors $[I_i, I_j]= I_iI_j - I_jI_i = i\varepsilon_{ijk}I_k$, which describe the whole space but, because rotations are limited to the surface of a sphere, the description can be reduced to two appropriate values (usually the absolute value I and one basic vector or spinor usually the third I_3). A combination of more than one of such systems, e.g. of two as in case of the vector and the axial vector part of the weak currents, results in a mixed system like the two oscillators of Heisenberg (see chapter 4.3.1, p. 76f) which is also closed, i.e. it has together with the already described commutation rule the form $[I_i, D_j] = [D_i, I_j] = i\varepsilon_{ijk}D_k$ and $[D_i, D_j] = i\varepsilon_{ijk}I_k$ where the combinations $I_+\equiv(I+D)/2$ and $I_-\equiv(I-D)/2$ are two commuting angular momenta. These two angular momenta fix the scale of the coupling because quantum field theory demands that they are quantized over a definite ground state. Of course, if the vectors, spinors, or some of their values are time dependent, the rules are valid only if they were considered at the same time. The connection to symmetry groups is established by the fact that the transformations (rotations in case of the vectors) mathematically form a group, an object which e.g., behaves like the properties of a square. Rotations of a square with integer multiples of ninety degrees or reflections on the middle axes or diagonals do not change the form i.e., the form is conserved, and all these transformations do not leave the square i.e., the group of transformations is closed. In the same way the system of Gell-Mann above works. However, mass terms disturb this closing, so to speak distort the square in the example before, they lead to the two angular momenta mentioned before fixing the scale by quantifying the 'distortion'. This is the general picture Gell-Mann paints.

For the components of the well-known isotopic spin current I_{ia} we then get $I_i=-i\int I_{ia}d^3x$ and because of the conservation law $\partial_a I_{ia}=0$ we have $\dot{I}_i=-i\int \partial_a I_{ia}d^3x=0$. Additionally, for $x\neq x'$ we must demand $[I_{ia}(\mathrm{x,t}), I_{ja}(\mathrm{x',t})] = -i\varepsilon_{\mathrm{ijk}}I_{ka}(\mathrm{x,t})\delta(x-x')$. Applied to the case of the weak current of the leptons as in the equations on page 94, we find $I_a^{(l)} = (\overline{\phi}_\xi(\mathrm{x,t})\tau\gamma_a\phi_\xi(\mathrm{x,t}))$ with $\xi=(e,\nu)$, τ for the Pauli matrices, and $I^{(l)} = -i\int I_a^{(l)}d^3x$ for the vector part respective $P_a^{(l)} = \left(\overline{\phi}_\xi(\mathrm{x,t})\tau\gamma_a\gamma_5\phi_\xi(\mathrm{x,t})\right)$ and $D=-i\int P_a^{(l)}d^3x$ for the axial vector part obeying the commutation rules of the former paragraph.

By setting proton and neutron fields instead of electron and neutrino fields as Gell-Mann does, the algebraic structure can be transferred to a field-theoretic model of baryons and mesons. The symmetry group here is the group SU(2) with two parameters, the same as that of the electronic spin, a symmetry introduced to the nucleons already by Heisenberg with the isospin. This is the first step of the extension of the explanatory analogy from leptons to hadrons. The progress here is that, as we will see, the nucleons are integrated in a wider framework of explanation.

Expressed in the model used by Gell-Mann, the weak currents in the field theory of baryons and mesons $I_a = \left(\overline{\phi}_N(\mathrm{x,t})\tau\gamma_a\phi_N(\mathrm{x,t})\right)$ with $N = (p, n)$ and I = $-i\int I_a d^3x$ for the vector part and $P_a = \left(\overline{\phi}_N(\mathrm{x,t})\tau\gamma_a\gamma_5\phi_N(\mathrm{x,t})\right)$ and D=$-i\int P_a d^3x$ for the axial vector part, respectively, are constructed out of proton and neutron fields like that of the leptonic currents out of electrons and neutrinos above.

So far, we have only considered the interaction terms consisting of the weak currents and their commutation rules. However, to specify the scale of the interaction additionally, the Lagrangian *L* or the Hamiltonian *H*, respectively, must be taken into account, which describe the energy contained in the system i.e., if I= I^++ I^- and D were conserved, they would commute with *L* and *H*. Then, because H transforms as isoscalar, I must be zero, and because I is associated with the total isospin quantum number, and I^+ and I^- are commuting angular momenta, they must be equal with isotopic spin quantum numbers (i_+,i_-). However, the mass of the nucleons destroys the commutativity so that *H*=*H*(0,0) – u_0. The simplest model with these properties is according to Gell-Mann a Fermi-Yang model where 'the pion is a composite of nucleon and anti-nucleon (Gell-Mann 1962, 1072). To avoid singularities, Gell-Mann replaces the direct four-fermion coupling by a massive neutral vector meson field B^0 leading to the Lagrangian $L = -\overline{N}\gamma_a\partial_a N - \frac{(\partial_a B_\beta - \partial_\beta B_\alpha)^2}{4} - \frac{\mu_0{}^2 B_\alpha B_\alpha}{2} - ih_0 B_\alpha \overline{N}\gamma_a N - m_0\overline{N}N$. Only the last term, the mass term, prevents commutation of I and D with both *L* and *H*. Therefore, the Hamiltonian departs in the two terms *H*= *H*(0,0) – u_0, the first commuting with I. Thus, B^0 belongs to (0,0), while $N_L \equiv (1 - \gamma_5)$ N/2 belongs obviously to (1/2,0) and $N_R \equiv (1 + \gamma_5)$ N/2 to (0,1/2), such that u_0 results in (1/2,1/2). Therefore, we get four components of the representation to (1/2,1/2) obeying the commutation relations $[I_i,u_0] = 0$, $[D_i,u_0] = v_i$, $[I_i,v_j] = i\varepsilon_{ijk}v_k$ and $[D_i,v_j] = i\delta_{ij}u_0$.

So far, only the two nucleons are considered in the scheme in analogy to the leptons and it must be extended in an analogous way to three particles, which was tried by a model of Sakata for a special case.

Therefore, the next step is the application of the Sakata model where additionally to the nucleons the Λ particle is considered corresponding to the addition of the muon to the leptons[38]. The analogy can be observed directly by comparing the two Langrangian densities for the baryons:

$$L = -\bar{p}\gamma_\alpha\partial_\alpha p - \bar{n}\gamma_\alpha\partial_\alpha n - \bar{\Lambda}\gamma_\alpha\partial_\alpha\Lambda - \frac{1}{4}\left(\partial_\alpha B_\beta - \partial_\beta B_\alpha\right)^2 - \frac{1}{2}\mu_0^2 B_\alpha B_\alpha - ih_\alpha(\bar{p}\gamma_\alpha p + \bar{n}\gamma_\alpha n + \bar{\Lambda}\gamma_\alpha\Lambda)B_\alpha - m_{0N}(\bar{n}n + \bar{p}p) - m_{0\Lambda}\bar{\Lambda}\Lambda$$

and the leptons, respectively:

$$L = -\bar{\nu}\gamma_\alpha\partial_\alpha\nu - \bar{e}\gamma_\alpha\partial_\alpha e - \bar{\mu}\gamma_\alpha\partial_\alpha\mu - 0(\bar{\nu}\nu + \bar{e}e) - m_\mu\bar{\mu}\mu.$$

[38] There is considered only one neutrino in this picture, not different neutrinos for the electron and the muon. The existence of different neutrinos was not obvious at that time, whether theoretically nor experimentally.

The Lagrangian describes the difference between the kinetic and potential energy, which summed up (integrated) over the whole curve of the movement must be zero because of energy conservation. The mass of the neutrinos and the electrons is approximately set to zero here. Besides that, the electromagnetic properties differ because the third element with the baryons is neutral whereas it is negatively charged with the leptons. More important, however, in analogy to the mass formulae mentioned on page 85 above is a separation of the Lagrangian into three parts $L = \bar{L} + L' + L''$: "$\bar{L}$ stands for everything except the baryon mass terms, while L' and L'' are given by $L' = (2m_{0N} + m_{0\Lambda})(\bar{N}N + \bar{\Lambda}\Lambda)/3$, $L'' = (m_{0N} - m_{0\Lambda})(\bar{N}N - 2\bar{\Lambda}\Lambda)/3$" (Gell-Mann 196, p. 1074). The second part L' gives the mean of the masses of all three particles p, n, and Λ, while L'' gives the mass-splitting of these particles. If now the third part L'' is neglected (switched off, see p. 87), the Lagrangian is totally symmetric and therefore invariant under the three-dimensional unitary group U(3). A unitary group is a group like the rotational group, which does not change the total value of the vectors or spinors but only the relation of their components, i.e., like the equation $\cos^2\alpha+\sin^2\alpha=1$, which is always valid for all angles α. The same is true if the equation expressed in complex numbers with the scalar product (cos α+i sin α)(cos α – i sin α)=1 or if a scalar product of two quantities $a.b=\Sigma a_i.b_i$ with $a=\Sigma a_i$, $b=\Sigma b_i$, i= 1..n and $(a.b)^2=1$ is inserted: (cos α + i sin α a.b) (cos α – i sin α a.b)=1. In principle, this is the strategy applied by Gell-Mann to extend the group. The group U(3) = U(1)xSU(3) is the product of a group U(1) with SU(3). SU(3) was already described in chapter 3.3, p. 44 together with the eight generating matrices. U(1) is of course a one-dimensional group, which is generated by a multiplication of all components of a unit matrix with the same factor. Together, these nine matrices obey the commutation rules $\mathrm{Tr}\, \lambda_i\lambda_j = 2\delta_{ij}$, $[\lambda_i, \lambda_j] = 2i f_{ijk} \lambda_k$, $\{\lambda_i, \lambda_j\} = 4/3\, \delta_{ij}\, 1 + 2\, d_{ijk}\, \lambda_k$, i,j,k = 0, …, 8, i.e., the same as in chapter 3.3, p. 44 but extended by λ_0=(2/3)1/2 and with f_{ijk}=0 if one index is zero and d_{ijk} with an additional nonzero matrix element equal to (2/3)1/2 if any index is zero and the other two are equal. Thus, in analogy to the two-dimensional nucleonic case, which was extended from the leptonic case, the invariance under U(1) gives the conservation of the baryon current $i\bar{b}\gamma_a b = i\bar{n}\gamma_a n + ip\gamma_a p + i\Lambda\gamma_a\Lambda = F_{0\alpha}$ and one gets $j_a = (\sqrt{2}F_{0\alpha} + F_{8\alpha} + \sqrt{3}F_{3\alpha})/2\sqrt{3}$ for the electric current and $F_{1\alpha} + iF_{2\alpha} + F_{4\alpha} + iF_{5\alpha}$ for the vector weak current. According to Gell-Mann, the same last term is valid for the axial vector currents, resulting in summary to a group U(3)xU(3) = U(1)xU(1)xSU(3)xSU(3) i.e., a left- and right-handed baryon number and unitary spin, too (Gell-Mann 1962, p. 1076). In the axial vector current each component is multiplied with the fifth Dirac matrix $\gamma_5=i\gamma_4\gamma_1\gamma_2\gamma_3$, which projects the components of a Dirac spinor to only left- or right-handed spins. The vector and the axial vector part of the current can be combined, just as in the former two-dimensional leptonic case, to $I = I^+ + I^-$ and $D = I^+ - I^-$ where I^+ and I^- commute with each other, however, I and D do not. As a reminder, because I_+ and I_- obey the parity relation $PI^{\pm}P - 1 = I^{\mp}$, they commute, and we have, in extended analogy to the two-dimensional case with nucleons alone, the relations $F_i = F_i^+ + F_i^-$ and $F_i^5 = F_i^+ - F_i^-, i = 0,..,8$ with the same property.

As in the case of nucleon and anti-nucleon, the Hamiltonian $\bar{H}$, derived from $\bar{L}$ defined above (p. 95), is completely invariant under the group. u_0 is set equal to L' because L' corresponds to the mass term in the two-nucleon-case. Then, we have in u_0 eighteen commutations of F_i and F_i^5 at equal times with u_i and v_i resulting in the relations $[F_i, u_j] = if_{ijk}u_k$, $[F_i, v_j] = if_{ijk}v_k$, $[F_i^5, u_j] = -id_{ijk}v_k$, and $[F_i^5, v_j] = id_{ijk}u_k$ with the f_{ijk} and d_{ijk} from chapter 3.3 (see chapter 3.3, p. 44) analogous to the case with two nucleons. Further, u_8 is proportional to λ_8 because L'' is constructed like λ_8 and we have $H = \overline{H} - u_0 - cu_8$ with c of the order of $(m_{0N} - m_{0\Lambda})/m_{0N}$.

Without the mass-splitting term L'', the remaining Langrangian $= L + L'$ is 'completely symmetrical in p, n, and Λ' (Gell-Mann loc. cit.). Thus, an invariant group with three elements rather than two as in the case of the nucleons alone must occur. However, if L'' is 'switched on' the symmetry breaks up into multiplets with each 2I+1 components, i.e., into the already known multiplets in the octet and a singlet (see p. 83). Thus, we get for a meson, which can dissociate into a baryon b and its anti-particle $\bar{b}$, the singlet $\bar{b}b=\bar{p}p+\bar{n}n+\bar{\Lambda}\Lambda$ and for the octet $\bar{b}\,(\lambda_1-i\lambda_2)/2\,b=\bar{n}p$, $\bar{b}\,\lambda_3/\sqrt{2}\,b=(\bar{p}p-\bar{n}n)/\sqrt{2}$, and $\bar{b}\,(\lambda_1+i\lambda_2)/2\,b=\bar{p}n$ with I=1 and S=0, as well as $\bar{b}\,(\lambda_4-i\lambda_5)/2\,b=\bar{\Lambda}p$, $\bar{b}\,(\lambda_6-i\lambda_7)/2\,b=\bar{\Lambda}n$ with I=1/2 and S=+1, $\bar{b}\,(\lambda_4+i\lambda_5)/2\,b=\Lambda\bar{p}$, $\bar{b}\,(\lambda_6+i\lambda_7)/2\,b=\Lambda\bar{n}$ with I=1/2 and S=-1 and at last $\bar{b}\,\lambda_8/\sqrt{2}\,b=(\bar{p}p+\bar{n}n-2\bar{\Lambda}\Lambda)/\sqrt{6}$. Gell-Mann then predicts, based on the known pseudoscalar mesons π, K and $\bar{K}$ with spin 0, which fit to this scheme very well, an eighth meson with its decay modes, which he calls χ^0 (today called η), and further vector mesons with spin 1 additional to the already known.

In a similar way, Gell-Mann tried to construct baryons by the same algebra. He suggested a structure of $3\text{x}\bar{3}=8+1$ or $3\text{x}3\text{x}\bar{3}=3\text{x}1+3\text{x}8=3+3+6+15$ but must concede that there remain problems, partly because of missing information, partly because of missing candidates for filling out the gaps. Therefore, he argues again for his more abstract model with unitary symmetry, namely the 'eightfold way'. All particles in the octet should have the same spin and parity and 'the masses should obey the sum rule analogous to (6.3) [i.e., for the pseucoscalar and vector mesons of the former paragraph $(m_K + m_{\bar{K}})/2 = (3m_\chi + m_\pi)/4$; $(m_M + m_{\bar{M}})/2 = (3m_\omega + m_\rho)/4$]: $(m_N + m_\Xi)/2 = (3m_\Sigma + m_\Lambda)/4$, which agrees surprisingly well with observations, the two sides differ by less than 20 MeV' (Gell-Mann 1962, p. 1080). This last formula was that already cited at p. 89 and indicates that the mass differences nearly agree with the difference of strangeness. This approach leads to the representation $8\text{x}\bar{8}=8\text{x}8=1+8+8+10+\overline{10}+27$ with 1 and 8 already discussed, 10 consisting of a triplet with Y=0, a doublet with Y=-1, a quartet with Y=+1 and a singlet with Y=-2 (Y=n+S, see p. 84), $\overline{10}$ has the opposite signs and 27 consists of a singlet, a triplet and a quintet with Y=0, and pairs of each doublets, triplets and quartets with Y=±1. Also, this approach was not quite correct, as we will see, because the last two representations, a singlet, and a second 8 representation were never found. To explain that only one octet and the decuplet exists, the colour symmetry must be considered (see chapter 3.3, p. 48f), however, that should not be deepened here.

In the last considered paper presented in this analysis, Gell-Mann postulates the 'quarks' (Gell-Mann 1964, p. 214). In this paper, he pointed out that, if the abstract model of the 'eightfold way' was right, the fundamental group should be SU(8) instead of SU(3). He refers to his model from 1961 (reprinted in 1963), described in chapter 3.3 (see p. 43ff), and the paper from 1962 (see p. 90ff). In both papers, he studied a triplet, built up by a doublet of electron and neutrino and proton and neutron, respectively, and a singlet of muon and hyperon, respectively, which constitute the supermultiplet of eight particles. In this new paper, Gell-Mann claims that all baryons in an octet and an additional singlet can be constructed out of a neutral baryon together with the aforementioned triplet and its anti-triplet and mesons with the triplet alone and its anti-triplet, respectively. However, in the same paper, he also suggested a simpler model if non-integral values for the charge were allowed and called them ‘quarks’ after a poem in ‘Finnegan’s Wake’ by James Joyce. The same octet can alternatively be built up by each three quarks in a baryon and a quark and an anti-quark in a meson analogous to the symmetrical Sakata model in the paper from 1962 and without the neutral baryon. This was the birth of the concept of quarks as constituents of hadronic particles, which was at the same time also suggested by Zweig, independently (Brandt 2009, p. 412). Today, quarks as basic constituents of baryons and mesons are fully accepted as part of the standard model of particle physics.

I suppose, from this sketch of the development of ideas starting from the constitution of the atomic nucleus till the constitution of hadrons, it is convincing that analogies in both forms, in explanatory and predictive form, play an essential and formative role in physical theorizing. As we have seen as well, a combination of explanatory analogies was used in the example of Heisenberg and in the example of Gell-Mann to suggest one or more concurring models. These models describe known observable phenomena on the basis of unobservable generalized assumptions, which were derived from corresponding more familiar models describing already better known and experimentally well-established phenomena. In some cases, as in the not conserved axial vector current or the predictions of the ‘eightfold way’, the model must later be adapted to the experimental results. And of course, the analogical models are not the only foundation for physical theorizing but an important one, allowing as flexible tool the alteration and adaption to the conditions of new but similar phenomena in the world acceptable by its links to already familiar experiences. Analogies serve not only for the task to overcome theoretical obstacles but also to overcome the limit between the visible and invisible in experimental practice, which I will try to show in the next chapter.

4.4 Summary

After the realization in chapter 3 that analogies play a role in physical theorizing, they are investigated in chapter 4 in more detail leading to a characterization of analogies in several respects and the demonstration of the usage of different types of analogies. For better understanding of the role of analogies, the chapter was introduced by a short sketch of a theory of physical development, which will be presented in detail in chapter 6. Then, it was shown that analogies are indispensable in two respects for physical theorizing i.e., they are indispensable in epistemological respect as method for the connection of observed but unexplained phenomena to already known and accepted theoretical explanations of familiar phenomena. This connection serves as vehicle to integrate the new phenomena into the background conditions of well-established knowledge, taken for granted, in one direction and to revise the old theory from the perspective of the new phenomena in the other. This means the analogical transfer and extension of a theory in a new domain also affects the view on the former theory. Thus, analogies have an important effect in the justification of theories by embedding the explained new phenomena into the landscape of accepted knowledge and background assumptions. But analogies are indispensable in psychological respect as well because they are an essential method for the categorization of all experiences perceived by the senses. Categorization and its dynamics originated by each new experience is the basis for orientation in the world and in this way a precondition for the survival in a changing but not chaotic environment.

Two types of analogies are relevant in physical theorizing formally: the explanatory and the predictive analogy introduced under these names by Bartha. The explanatory function of analogies, which have another direction of association as other types of analogies (from the hypothesis to the positive material analogies, see chapter 4.2.2, p. 69), serves also as justification for the validity of the theory because it expands the knowledge by the supporting connection to existing and accepted knowledge. In contrast, the predictive analogy, where the association goes the other way around (from the positive material analogies to the hypothesis), gives an opportunity for the testing of the theory by experiments, which is one possibility in the context of discovery for the finding of new phenomena or objects concluded from the extension of the explanatory analogy. Other possibilities for discovery will be discussed in chapter 6.

The main reason, however, for the usage of analogies in modern physics is, as demonstrated in the three discussed examples in chapter 4.3, the extension of the limits of observability resulting in more and more abstract theories. These are based on partly observable phenomena, which are concluded from the assumed behaviour of unobservable objects on the one hand and models drawn in several historical steps of ongoing development and abstraction from the observable behaviour of observable objects on the other. Those analogical abstractions construct a network of trust for the behaviour of the postulated objects in high energy physics with the limited access only via the electromagnetic interaction. However, how the electromagnetic interaction was developed and used in the experimental practice establishes a further analogy, which will be analysed in the following chapter.

5 Experimental Practices and the Role of Limited Access

In chapter 5 I will start with the cloud chamber, an instrument which could make visible tracks of charged particles for the first time. Originally, the cloud chamber was invented for another purpose, namely, to imitate the generation of coronas, fogs, and clouds in nature. Only that Wilson, the inventor of the cloud chamber, worked at the Cavendish laboratories of J.J. Thomson at that time, brought him in contact with questions on the structure of matter and the assumption of the existence of ions, i.e., charged atoms. The exposition of the cloud chamber to the former discovered X-rays and radioactivity produced ions in the chamber and condensed water around them forming tracks. This was the first visualisation of unobservable moving objects like electrons or α-particles, an analogical representation of the track of an assumed particle which itself was not observable.

A similar visualization, however, with a completely different technique was invented by the Austrian physicist Marietta Blau. She found a way to sensitize photographic emulsions, such that they were able to depict tracks from charged particles with low energy which was till then possible only with cloud chambers. Her greatest success was to find an atomic nucleus emitting several electrons the first time. This and the possibility to derive the energy of the particle from the length of the track encouraged Powell and Occhialini to expose a new emulsion by Ilford to cosmic radiation. With this equipment, they found among other particles the pion in 1947 as mentioned above.

A third method on this path of visualizing events was the bubble chamber, invented by Glaser in 1953. However, the greatest successes with the bubble chamber were achieved by Alvarez and his group, who built among some smaller ones a huge bubble chamber with a length of 72 inches. This last-mentioned bubble chamber was the instrument which gives support to the strangeness concept and the quark hypotheses.

A totally different and a long-time underestimated technique, turning away from visual representation in the beginning, was the use of electronic counters. First invented by Geiger and Rutherford, the single wire inside the counter was replaced by many parallel wires in different directions in a 'multiwire proportional chamber' making possible to get the trajectory coordinates in the above-mentioned experiments by Ting and Richter and their groups, who found the J/φ particle. Cherenkov counters, used by Ting and shower counters used by both registering the generated light of electrons, were used additionally, and were analysed by computers together with the trajectory data of the multiwire proportional chambers. The analysis by computers, combined with triggering of the data registration of the trajectories in multiwire chambers, enabled a trend back to visualization. Culminations point of this development was the 'time projection chamber', which enabled the registration of a three-dimensional pictorial representation of events.

In contrast to the visualization of particle trajectories as analogies to the unobservable ones in nature, Monte Carlo simulation on digital computers was considered as 'artificial reality' corresponding to the discreteness and random behaviour of nature, i.e., a coarse-grained picture of real events. This view also affected the later development of hybrid detectors, like that used at CERN and the Fermilab, where 'Monte Carlo simulations were part of the planning, construction, and analysis phases' (Galison 1997, p. 668).

This way of visual analogization of mostly unobservable physical phenomena will be discussed in the following more in detail.

5.1 The Cloud Chamber and Electronic Counters

The experimental practices of course must deal with the same problems of unobservability as physical theorizing and one may assume that analogies also are relevant there, perhaps even in a much harder form. However, as we will see, the kind of analogy is different here, but its origin has to do with the limits of access, as well. To show this, we must have a look again on history, now of the development of experimental practices in high energy physics.

Because the focus is limited to high energy physics, we will start with a technique which was able to prove elementary particles and their interactions. However, this technique was not developed for this purpose in the beginning. In the beginning, the interest of C.T.R. Wilson was limited to the condensation of fog and clouds, motivating him 'to reproduce the beautiful optical phenomena of the coronas and glories I had seen on the mountaintop' of Ben Nevis in 1894 (Wilson 1960, p. 166, cited after Galison 1997, p. 91) and the electrical phenomena and damages occurring during an electric storm nine months later at the same place (Galison 1997, p. 90). Wilson had a job at the Cavendish Laboratories at that time, but he had no time for research because he was working at full capacity to demonstrate physics for medical students, so he began a notebook where he wrote down his findings and assumptions (Galison 1997, p. 91). Wilson was not the first to try this task, so he could review some results of experiments in the literature.

When he was able to make his own experiments, these results, especially from his Scottish countryman Aitken, inspired Wilson to build a very similar apparatus with a pump to lower the pressure of air entered through a wet filter, which saturates the incoming air with water before the valve is closed (Galison 1997, p. 94). The theory behind this arrangement was that lowering the air pressure of saturated air lowers the temperature during the expansion and causes a supersaturation. With this equipment, Wilson could demonstrate the coloured rings observed at the mountains of Ben Nevis. However, a main difference between his apparatus und that of Aitken was that Wilson filtered the incoming air to remove all dust. Aitken assumed that dust was the relevant source for the condensation in nature, but Wilson wanted to analyse some other possibilities for the origin of condensation, which were known but not understood (Galison 1997, p. 96). One origin for supersaturation and condensation was a fast and high expansion, another the electrification of the air. This last origin was what Wilson began to investigate motivated by a theory of J.J. Thomson, the chief of the Cavendish laboratory, which claimed that small drops of water in the air must evaporate because of a high vapor pressure caused by the surface tension if not an inhomogeneous electric field decreases the vapor pressure (Galison 1997, p. 98f). A first result in 1895 was the determination of the critical expansion ratio at a specified initial temperature of 16.7 ˚C where condensation begins. In the next years, he exposed his 'cloud chamber' to the just discovered X rays and a little later to the uranium rays, discovered by Becquerel, and found that both intensified the generation of drops at the same expansion ratio. Wilson concluded from these findings that always small numbers of ions are present in the air because both types of rays produce ions but without changing the expansion ratio (Galison 1997, p. 100). In further investigations, he could demonstrate four regions of condensation separated by three different expansion ratios, caused, as he assumed, by two different types of ions. This assumption was specified by J.J. Thomson suggesting that positive ions were less caught than negative ones during expansion (Galison 1997, p. 101) and confirmed about half a year later by Wilson, introducing as modification a brass needle hold at constant potential from the top of the chamber. This potential drove away the positive ions and concentrated the fog near the needle with the potential attached in one direction but in the other direction it distributed the fog in the whole chamber because the electromagnetic field was not strong enough to catch the ions (Galison 1997, p. 102). For Wilson, this fact allowed immediately to explain why, as already known, a potential gradient existed between the negative charged earth and the ionosphere.

However, a problem caused him to give up his cloud chamber experiments for some time and to continue with experiments with an electroscope. This problem was that ultraviolet light, but no electric field, could produce condensation, which was not compatible with ions as source of condensation. This fact led Wilson to the assumption that the condensation process in his cloud chamber was artificial and not the same process as in nature (Galison 1997, p. 104f). Wilson's investigations with the electroscope would have caused him to detect cosmic rays if the radioactivity under the earth had not led him astray (Galison 1997, p. 107). So, Hess detected them in 1912 from a balloon at the height of 5000 m also with electroscopes (Brandt 2009, p. 79).

Only when he wanted to investigate the growth of droplets in thunderstorms, he returned to cloud chamber experiments. There, he applied, besides a vertical tube, a technique taken from Worthington's study of splashes, namely high-speed photography illuminated by the light of sparks.

Wilson was an excellent photographer, experienced over long years, so, he could make pictures of tracks of droplets produced by β, X, γ, and α rays for the first time. Between these pictures 'on one occasion in addition to ordinary thread-like rays, one large finger-like ray was seen, evidently a different form of secondary ray – giving rise to enormously more ionization than even ordinary [alpha] ray' (cited from the notebook of Wilson by Galison 1997, p. 112). Shortly after that event, Wilson published in April 1911 his first paper on tracks named 'On a Method of Making Visible the Paths of Ionising Particles through a Gas' (Galison 1997, loc cit.) with a few photographs (Brandt 2009, p. 71). One year later, he submitted a 'comprehensive paper […], which already contained a photograph of the track of an α particle which, in two distinct plates, changes its direction because the particle suffers Rutherford scattering' (Brandt 2009, p. 71). A further year later, Wilson's cloud chamber was produced in series by an instrument maker (Galison 1997, p. 116).

Remarkable on the work of Wilson are two aspects, firstly that 'from his earliest work in 1895 to his last ruminations on thunderclouds when he had turned ninety, Wilson was riveted by the phenomena of weather' (Galison 1997, p. 73f) and secondly that he 'indeed made the invisible visible' (Galison 1997, p. 114). The first point means that Wilson was never really part of the atomic research community because of his different aims, despite the essential contribution which made his apparatus to the investigation of the structure of matter. The second point will become the focus of this chapter: It is the analogy between the observed track and the imagined trajectory of a postulated particle obeying as assumed at least similar rules of movement like directly observable particles. I will not decide if these postulated particles are real or not, it is enough that the theoretical explanation is, as van Fraassen formulates it, 'adequate to the phenomena'. Informative for me is the story of visualization techniques from the cloud chamber, photographic emulsions and the bubble chamber via non-visual logical techniques of counters and statistics back to a visualization of these last techniques with the help of computers, or by Monte Carlo simulations totally on computers. I will discuss the way of this development in the last part of this chapter, when I have presented the difficulties caused by the experimentation with invisible objects and the additional limitation that they were only accessible by electromagnetic interaction.

As mentioned, Wilson was the first who observed tracks caused by invisible particles. However, he was not the only scientist who worked with the cloud chamber at this time. Because he was employed at the Cavendish laboratory under J.J. Thomson and later under Rutherford, other members of the laboratory and also Thomson and Rutherford themselves used the chamber. Already in 1898 i.e., three years after Wilson's first experiments, Thomson measured the elementary charge assuming that each droplet in the chamber contained only one charged ion (Galison 1997, p. 103). The relation of the charge to the mass of 'electrons', as these particles in cathode rays were called later, was determined by Thomson in the year before. This new experiment determined a value of the elementary charge by estimating the number of droplets in relation to a fixed charge. Because the amount of water in the chamber is fixed, the number of droplets depend on from its size; and the size can be calculated from their velocity falling in the gravitational field using Stokes' law. Stokes' law says that the force on the droplet is equal to a constant, the viscosity of the fluid, here the air, the radius of the droplet and the velocity, so that the radius can be calculated by the velocity. A similar, but improved method was used later by Millikan during his measurement of the elementary charge.

However, Rutherford used a technique which could not visualize tracks either. Together with Geiger, who invented this technique, Rutherford could observe a single alpha particle in 1908 for the first time (Brandt 2009, p. 54f). This technique was the electric counter named after Geiger and constructed as a long cylinder filled with air at a low pressure of about forty percent of the atmospheric pressure. In the centre of this cylinder, a wire was spanned, which was set under a voltage of about 1000 Volts producing an inhomogeneous electrical field. in the front of this cylinder, a small hole is cut covered with mica, through which alpha rays emitted from a radium probe could penetrate. Always, when an alpha particle comes in, it ionizes the air inside and the electric field accelerates the electrons in direction of the wire so fast that these produce additional ions and these further ones etc. In this way an amplification of the alpha particle passage occurs which could be measured by an instrument outside. So, every time when an alpha particle enters the cylinder an event is counted. Therefore, this instrument was called Geiger counter. The possibility to detect single particles was a profound reason for the acceptance of atoms as constituents of matter.

Some years later in 1912, both Geiger and Rutherford improved the equipment by registering the pulses of the counter on a fast-moving photographic film allowing higher rates of particles to register. One year later, now in Berlin, Geiger improved his counter again by replacing the wire in the centre by a sharp needle, which also allowed to register beta rays i.e., electrons because of the much higher amplification of this construction. With the Geiger needle counter and the cloud chamber, now two somewhat concurring techniques were available, which could register the invisible particles of matter: one, which is easier to control and can use the significance of statistics (see Galison, p. 434) and the other, as we will see in the next paragraph, with the immediate persuasiveness of an optical impression but limited to 'golden events' i.e., a few interesting phenomena among masses of photographs.

After Wilson's detection of tracks in the cloud chamber, the first successful application of the evaluation of them was that Rutherford together with Blackett discovered the 'disintegration' of nitrogen by fast alpha rays which was in fact an 'integration' (Galison 1997, p. 119). Because there were only two characteristic tracks after the interaction, it was clear that the alpha particle was absorbed, 'transmuting' nitrogen in an oxygen isotope and emitting a proton. This was a later so-called 'golden event' i.e., only 18 of about 400 000 photographed collisions were not elastic but transmutations. That means a 'golden event' is a new one, never observed before among a huge mass of other events. Characteristic for the observed tracks means that all tracks can be evaluated by the number of originated droplets. A track with few droplets in greater distance can be identified with a beta ray (electron), more in smaller distance with a proton, a thick track to an alpha particle and more like that. Influenced by a magnetic field, the tracks are curved, and the momentum can be calculated. By taking stereographic photographs the momenta of all particles moving in all directions can be reconstructed (Brandt 2009, p. 71f). Thereby it was assumed that the particles obey the Newtonian or Lorentz force laws, respectively, and energy, momentum and mass were conserved.

In contrast, events in the counters can be controlled in some way and their results are supported by statistical considerations. This should be demonstrated in the following examples from 1925 and 1929. In 1925, Geiger and Bothe used two needle counters to execute the first so-called coincidence experiment, an application pointing the way for the future of all experiments using several electronic detectors (Brandt 2009, p. 143). This experiment was intended to show that the Compton effect was right. Compton had demonstrated in 1923 with scattered X rays the so-called Compton effect, measured by a Bragg spectrometer, that a photon carries a momentum (Brandt 2009, p. 130f). A Bragg spectrometer is the combination of a crystal, scattering the ray at the regular ordered atoms of the crystal structure, and an ionization chamber, the forerunner of the Geiger counter without the central wire or needle and therefore no amplification. In the Compton effect, the wavelength of the ray changes with the scattering angle.

In the following, Wilson and Bothe observed tracks of recoiled electrons in the cloud chamber. However, a theory from Bohr, Kramers and Slater, later called BKS theory, assumed for agreement with Maxwell's theory of electromagnetism that all oscillations occur in the atoms only and energy and momentum is conserved only in the long run. Then, between electrons and photons there would be no coincidence, but the experiment of Geiger and Bothe showed that Compton was right, prompting Bohr, Kramer and Slater to withdraw their theory. The experiment itself consisted as mentioned of two neighbouring needle counters but separated by a thin window of aluminium. This window was chosen in such a way that photons could pass but electrons were recoiled. Parallel to the window, a collimated pure X ray produced scattered electrons, which were counted by the counter on the same side as the X ray. Geiger and Bothe looked for counts of the other counter, immediately produced with the first counter indicating that the photon collided with a further electron in the second counter. The results were recorded on a photographic film and evaluated by the simple rule that the relation between coincident counts within 1 ms and singular counts of only one counter of longer time intervals was calculated. The result was five to one for the coincidences and thus confirming Compton. Here, we see a simple example of the theoretical interplay between the use of counters and statistics.

With the old Geiger counter, developed together with Rutherford, there was a problem however it would be considered as permanently unreliable because if its sensitivity was raised by raising the voltage an autonomous discharge occured. This could be 'quenched' by e.g., coating the central wire with a bad conductor, as Müller suggested after investigation of the problem in 1928, but the signal then was no more proportional to the ionization (Brandt 2009, p. 56).

Another result of Müller was that the unreliability was caused by cosmic rays if not shielded by thick walls (Galison 1997, p 440). Therefore, the Geiger-Müller or GM counter, as it was called afterwards, became the preferred instrument for the research of cosmic rays. One of the first experiments of this kind and a further application of the coincidence method was the discovery of the nature of cosmic rays in 1929. In this experiment by Bothe and Kohlhörster, two GM counters in a thick shield were placed above another, separated by a 5 cm thick gold block. In spite of the massive shielding, they observed many coincidences, so that they concluded: "To this whole evidence we believe we must give the interpretation that cosmic radiation, at least so far as it manifests itself in phenomena observed up to now, is no γ radiation but a [charged] corpuscular radiation" (Bothe & Geiger 1929, p. 271, cited in Brandt 2009, p. 143).

In the same year, Bothe began to use electronic tubes to evaluate the coincidences of two Geiger counters, too. Again, Bothe and Becker used a Geiger counter in 1930 when they irradiated light elements like lithium, boron and beryllium with alpha rays emitted by polonium. Unusual, however, was that the resulting rays were much lesser absorbed by lead than all then known gamma rays. Nevertheless, they assumed their nature as extremely hard gamma rays (Brandt 2009, p. 210)., Joliot and I. Curie observed in following investigations with these hard rays that material like paraffin, which contains hydrogen atoms and was attached in front of an ionization chamber or later a cloud chamber, emitted protons, accelerated by these rays. Only Chadwick, when he heard from these experiments in 1932, concluded that the mysterious rays consist of neutrons, particles which were speculatively postulated by Rutherford in a lecture in 1920, but forgotten in the meantime (Brandt 2009, p. 210f). Because of the statistical problems with nitrogen (see chapter 3.1, p. 35), he assumed already the neutron to be an elementary particle.

In the same year, 1932, Anderson detected a further particle, the positron, also by application of a cloud chamber (Brandt 2009, p. 214f). This cloud chamber was equipped with a highly efficient magnetic field generator which curved the tracks of charged particles traversing the chamber. Anderson took photographs from tracks of cosmic rays and found surprisingly besides the expected negative electrons also positive particles in the same frequency. He and Millikan first assumed these positive particles to be protons, but Anderson began to have doubts because some of these particles show the same few droplets per length unit as the electrons, which does not fit to a proton. Therefore, Anderson put a lead plate inside the chamber and found by the stronger curvature of the track above the lead plate that the traversed particle came from downwards. It came, as described in a former paper from Anderson and Millikan, from a nucleus knocked out by the gamma rays of cosmic radiation. At the same time, Blackett and Occhialini from the Cavendish laboratory published also a paper reporting positive electrons (Brandt 2009, p. 217). They also used a cloud chamber, however, equipped with two counters, one above and one underneath the chamber, which triggered in coincidence the expansion of the chamber and took a photograph.

In this way, the electronic and visual technique was combined for more efficiency. This combination secures that only events were documented which fulfil the condition that both counters were released i.e., reduces the number of undesired photographs. Nevertheless, the search for 'golden events' remains because, also when both counters were released, the photograph taken could show an irrelevant event. Irrelevant means here pictures of already known phenomena - relevant are only new till then unknown events or those that the experimenters are looking for, events, for which the experiment was set up. These conditions, with or without triggering counters or with the use of cloud chambers or other visualizing techniques, respectively, stay valid for several further decades and also for the next experimental setting.

A few years after the discovery of the positron, Anderson and his assistant Neddermeyer mounted their magnetic cloud chamber on a trailer to transport it to the Pike's Peak, Colorado, at the height of 4300 meters to observe cosmic rays (Brandt 2009, p. 251f). There they took a lot of photographs from ‘electromagnetic showers’. Electromagnetic showers are cascades of electrons, originated by a high energetic photon, which produces in the strong field of an atomic nucleon a pair of an electron and a positron, which in turn cause by ‘bremsstrahlung’ further photons, which knock out further electrons etc. and thus generate a cascade of particles. Bremsstrahlung is the effect that a charged particle like an electron in a medium moves faster than the light speed divided by the refractive index, causing the particle to lose energy and therefore velocity by radiation of photons. This happens according to the Maxwell equations, describable in analogy to the acoustic shock wave of a projectile, known as Mach’s cone (Brandt 2009, p. 246).

To determine the energy loss of the particles in the electromagnetic shower, Anderson and Neddermeyer ‘placed a bar of platinum of 1 cm thickness in the middle of their chamber’. In this investigation, protons were excluded in their analysis from the beginning. The other observed particles fall into two groups, one with a behaviour as expected from electrons or positrons losing energy proportional to their energy in the beginning, the other with only a small energy loss and seldom as part of a shower. These last particles of the second group have a mass between electrons and protons, they claimed. Street and Stevenson confirmed that assertion from the observed droplet density in the same year, 1937 (Brandt 2009, p. 252).

Unfortunately for Anderson and Neddermeyer, Kunze had published in 1933 already a cloud chamber photograph of this unknown particle, also registering the droplet density between proton and electron and thus confirming the assumption made for the penetrating radiation by Bothe and Kolhörster in 1929 (see p. 105 above). Only the discovery of this new particle, first called mesotron and later meson, made the scientists aware of the theory of strong interaction which Yukawa had published in 1934 (see chapter 3.1, p. 36f, and chapter 4.3.1, p. 81f respectively). However, as also mentioned already, this was not the particle Yukawa suggested because it shows no strong interaction and, therefore, was renamed later to muon.

All the expertise gained by cloud chamber photographs was collected in 1940 in an ‘Atlas typischer Nebelkammerbilder’ published by Gentner, Maier-Leibnitz and Bothe with ‘pictures of beta rays in magnetic fields, of Anderson’s extraordinarily clear glimpse of a positron, and of photoelectrons from X rays and others scattered by gamma rays. There were electron pairs produced by gamma rays, bursts of alpha rays, and of course, the by-then famous pictures of Wilson’s first alpha-scattering events. In the later sections of the book were pictures of protons being hit by neutrons and, perhaps most typically, the just-discovered splitting of the uranium nucleus. Together, these images were supposed to tutor the eye, to train the growing cadre of cloud chamber physicists to recognize the typical, and therefore to underscore the new. By the second edition (1954), Gentner, Maier-Leibnitz, and Bothe had added to their atlas a host of cloud chamber pictures taken at accelerators […]. Together these volumes became standard reference works for cloud chamber particle hunters’ (Galison 1997, p. 121f).

This citation shows the importance of the visual evaluation of tracks because these give a direct impression of the imagined phenomena, despite that they are not observable in a direct way. The visual analogy to the movement of an observable particle allows the formulation of a similar behaviour of unobservable objects. This is also the reason that more visual techniques were developed than non-visual ones, as the following example shows.

5.2 The Predomination of Visual Techniques

The real particle Yukawa suggested in his theory of the strong interaction, the π meson or pion, was found in 1947 by another visual technique, namely the so-called photographic method (Brandt 2009, p. 300f). That photographic emulsions were sensitive to light, X rays and rays from radioactive substances was known for a long time beginning in the middle decades of the nineteenth century, however, they could not compete with the cloud chamber because they were not sensitive enough. Photographic emulsions could not depict electrons of a few MeV. Besides that, other problems occurred in the early times like unstable and to thin emulsions, ‘fading images, inconsistent tracks, distorted trajectories, and published opposition to its validity’ (Galison 1997, p. 152), which made its application questionable. However, for the investigation of cosmic rays the photographic emulsion has a decisive advantage in comparison to the cloud chamber: the tracks for a particle with the same energy are much shorter because of the emulsions’ higher density and, therefore, more interactions with nuclei (Brandt 2009, p. 300). For that reason, the Austrian Marietta Blau and her collaborators developed a new emulsion, which exposed for five months in 1937 at 2300 m height showed a proton track with an energy equivalent to a length of 6,5 m in air, a length not found until then. The most spectacular object they found however were ‘contamination stars’ of up to nine particles emanating from a single point, their ‘golden event’ never seen before (Galison 1997, p. 154). This ‘star’ must have been the destruction of a heavier nucleus in the emulsion elements into at least an alpha particle and several protons, caused by an energy of cosmic radiation much higher than that from radioactive sources. Because she was Jewish, all further research of her suffered from problems, originated by the Nazis, and their consequences, so that she could not gain a further success with her emulsion technique.

However, Powell, a student of Wilson, heard about the emulsions of Blau and began intensively to work with them because, familiar very well with cloud chambers, he supposed “[w]e believe it possible to press the analogy [of the emulsion to the Wilson chamber] further in the sense that the spatial orientation of a track in the emulsion can be determined” (Powell & Fertel 1939, p. 117, cited in Galison 1997, p. 167). Powell introduced over the years several changes in the emulsion method. First, he applied photographic plates made by commercial enterprises, primarily Ilford and later also Kodak. Second, he employed a lot of so-called ‘scanners’, women with no physical expertise but the patience and care that the thousands of fine-grained tracks demanded for their evaluation to be scanned by a microscope. Third, he follows the suggestion of Occhialini, who met Powell’s group, to raise the content of silver bromide in the emulsion to a higher degree (Brandt 2009, p. 300). Plates with this new emulsion produced by Ilford were exposed by the group (further enforced by the Brazilian Lattes and Powell’s student Muirhead) in 1947 to cosmic rays at the Pic du Midi in the French Pyrenees in a height of 2800 m. On these plates, they found mesons, identified by their relatively low density of the track, disintegrating in a visual star of particles, when they came to a stop after the slowdown in the emulsion.

However, they observed in the same year other mesons remarkably decaying also into mesons with an intermediate mass between protons and electrons. For confirmation of the evidence of two different meson types, the group made further expositions on the higher Mount Chacaltaya in the Andes at 5500 m and published the results also in 1947, giving the two different mesons the names π and μ (Brandt 2009, p. 301). The names stuck but the μ meson was later shortened to muon because it does not participate in the strong interaction. The muon is a kind of heavier electron. The π meson (pion) is the particle with the properties Yukawa predicted in 1934 (see chapter 3.1, p. 36f and chapter 4.3.1, p. 81f, respectively). Pions, generated by an accelerator the 184-inch cyclotron in Berkeley, were detected by Lattes also in the same year, 1947, using the emulsion technique. Only in 1950, the full decay chain of the pion in a photographic emulsion was observed, first into a muon and then in an electron. This was possible because only then were emulsions with the needed sensitivity available (Brandt 2009, p. 302). Generally, it was obvious from their definite energy, that both, the pion and the muon decay, are decays into two particles (see e.g., in contrast the beta decay in chapter 4.3.1, p. 80).

We have just seen that photographic emulsions could be used with analogous results as cloud chambers. In some respect emulsions are even superior to cloud chambers because of the higher density of their medium. Nevertheless, both techniques are applied simultaneously, and a reconstructed magnetic cloud chamber, equipped with two trigger counters and a lead bar in the centre, found in the year before new and unexpected particles. These unexpected particles were published in 1947 by their discoverers Rochester and Butler and called V-events there because the first observed event was a 'fork' of two opposite charged particles moving down from a common starting point (Brandt 2009, p. 307). Apparently, an uncharged particle in cosmic rays or generated by them decayed into the two charged particles. In the same year 1946 again, a further fork was observed, however, this time originating from a charged particle and decaying into a charged and an uncharged particle as the direction change indicates. From the droplet density, it could be concluded that the charged particles should be pions. Obviously, in both cases the visual phenomenon of the track was interpreted as an event analogous to the behaviour of observable colliding particles even though none of them was really observed. My claim is, this result is due to the explanatory power of the just sketched analogy, as I will argue at the end of the chapter again in more detail.

Only about two years later in 1949, a further event of a similar kind was observed. This time, Powell's group using photographic emulsions at 3460 m height at the Jungfraujoch in Switzerland found a charged particle slowing down to rest and decaying into three charged particles (Brandt 2009, p. 308). One of the three particles was a pion because of its reaction with a nucleon, the other two also pions or alternatively muons. The originating particle of the decay has therefore a mass of at least 925 electron masses i.e., more than that of three pions. A further event of this kind was reported by the Imperial College in London one year later and then the observed events increased. Also, in 1950 Anderson's group published observations of 'about 30 cases of forked tracks' in the cloud chamber 'exposed in Pasadena and at an altitude of 3200 m on White Mountain in California' (Brandt 2009, p. 309). In this publication, the particles are called V particles i.e., V^0 and $V^{\pm}$, respectively. In the meantime, the group of Rochester and Butler built a new cloud chamber, able to be taken apart for transport, and observed in 1951 on the Pic du Midi at 2800 m 36 neutral and 7 charged V particles.

However, the positive decay tracks of four of the neutral V particles showed a peculiarity: they turned out to be protons (Brandt 2009, p. 310).

Obviously, there existed two different neutral V particles, a heavier V^0_1 and a lighter V^0_2, which was confirmed by observations in a cloud chamber with minimized distortion by a group led by Thompson. Also, for 21 charged V particles, besides the confirmation of their nature as heavy mesons, one new event occurs, namely a cascade of V particle decays where a negative charged V particle decayed in a neutral V particle and a pion, as Rochester and Butler published in 1952. The V particles are the so-called strange particles because of their different production and decay process, mentioned already several times before (see e.g., chapter 3.2, p. 40f or chapter 4.3.2, p. 82). They are assumed to be produced in a strong interaction because of an interaction of cosmic rays with nuclei, but to decay in a weak interaction because of a lifetime of more than 10^{10} times of that in strong interactions. In 1954 a naming scheme for these particles was agreed: K for mesons and Y hyperons i.e., particles heavier than a neutron (Brandt 2009, p. 311).

The year before, tracks produced by another visual technique were published (Brandt 2009, p. 329). This technique, the bubble chamber, was supposed to combine the advantages of the cloud chamber and of the photographic emulsion and avoid their disadvantages. Cloud chambers had problems detecting complicated structures like the V events because of the low density of their filling with saturated air, and photographic emulsions could not fill a larger volume than some litres. The placing of plates of e.g., lead inside the cloud chamber does not solve the problem because events inside of this block could not be observed. The inventor of the new technique, Donald Glaser, made a lot of attempts with cloud chambers as well as with photographic emulsions before he tried to develop the bubble chamber. In a specific way, the bubble chamber is a cloud chamber with a denser medium, namely a liquid. This liquid must be heated in the closed vessel over the boiling point, of course, combined with the raising of the inner pressure over the outer atmospheric pressure according to Boyle-Mariotte's law (see chapter 2.1, p. 17). If then the pressure is suddenly lowered, the liquid will start boiling. But shortly before it begins to boil, cosmic rays or the beam from an accelerator can cause boiling around the trajectory of the ionizing particle and thus produce a track of bubbles in the liquid. In the first bubble chamber built by Glaser, the liquid was diethyl ether boiling at 34,6˚C at atmospheric pressure but heated to about 140˚C, which he assumed to be the right temperature for minimum-ionizing particles. In later models, he uses isopentane and liquified xenon, a noble gas, which could register high energy photons by their production of electron-positron pairs (Brandt 2009, p.330) However, there is one problem with the bubble chamber: its tracks fade much faster than in the cloud chamber and therefore the bubble chamber could not be triggered in such a short time that a picture could be made. That was a disadvantage for research on cosmic rays but an advantage for the use with accelerators, because with the latter unwanted tracks disappeared soon. Because Glaser was not interested to work with accelerators, he left physical research for molecular biology in 1962 but the bubble chamber succeeds him.

The difference in the assessment of suitability of the bubble chamber for cosmic rays and beams from accelerators lays in the control of the circumstances. Cosmic rays cannot be controlled and therefore the desired 'golden events' i.e., phenomena not already observed or confirming theoretical predictions, could not be planned; one could only wait for their occurrence. In contrast, the beam of an accelerator could be controlled in some respect and additionally has a much higher intensity, mostly called 'luminosity' as with light. An accelerator was available at the University of California in Berkeley. This last argument was the reason for the group of Luis Alvarez after some intermediate steps to build there a huge 72-inch bubble chamber filled with hydrogen. Hydrogen as filling has the advantage that always a proton as the collision partner for strong interactions is determined. On the other hand, hydrogen was a dangerous material, as all knew from the catastrophe of the zeppelin 'Hindenburg' in 1937. So, building such a huge bubble chamber filled with liquid hydrogen under high pressure and at extremely low temperatures was an enormous technical and organisational challenge, not considering the tremendous costs of about 2.5 million dollars. The technical challenge was, besides the extreme dimensions, to find material which could resist the high pressure and the low temperatures, which called for an intensive participation of technicians and engineers in an extent never needed before in physics. This heterogenous composition of the working group was a new organisational challenge, too (Galison 1997, p. 340ff). Additionally, the organisational challenge was even greater for the evaluation of the vast numbers of photographs. I will come back to this point a little bit later because first I must explain why the bubble chamber must be so huge.

The reason for its dimensions was that it was designed for the observation of the special decay chain of the strange particles, discovered in the years before and described in several paragraphs above. The considered decay chain started by the collision of a negative pion, delivered by the Bevatron accelerator at Berkeley, and a proton (hydrogen nucleus) in the bubble chamber producing the two neutral strange particles Λ^0 (identical with V^0_1) and θ^0 (which is identical with V^0_2, τ and K^0 as known only later after the fall of parity, see chapter 4.3.2, p. 83). Both particles are not observable because of their neutrality, but the length of their tracks until their decay and the scattering angle between them could be calculated from the energy and momentum of the collision partners, pion and proton, and the lifetime of the strange particles. At the end of these tracks, the strange particles decay into a negative pion and a proton (the decay of the Λ^0) and two oppositely charged pions (the decay of the θ^0), respectively. For possible observations of other assumed strange particles, the needed length of the bubble chamber of 50 inches was extended to 72 inches (Galison 1997, p.363ff).

Now I will come back to the vast number of photographs to evaluate. To cite Alvarez regarding all visual types of visual detectors i.e., cloud chambers, photographic emulsions and bubble chambers, 'suffer from a common difficulty which is not present in counter experiments. Each event must be studied and measured individually' (Alvarez 1955, p. 18, cited in Galison 1997, p. 372). Because of the high luminosity of accelerators, the data of one day 'keep a group of cloud chamber analysts busy for a year' (Alvarez in Galison 1997, loc. cit.). As solution, he suggested a similar process as Powell introduced in his group i.e., the employment of 'scanners' (see p. 106), however, with technical assistance. The whole process consisted of five steps: firstly, scanning by non-physicists (always women!) for interesting events, secondly preparing the selected events for measurement by a physicist, thirdly measuring the tracks with a track following machine and punching it on an IBM punch card by a non-physicist, fourthly fitting the tracks by computer to a spatial curve and at last checking the computer output by a specialist for acceptance or rejection (Galison 1997, p. 373).

Nevertheless, this process is relatively time consuming. Therefore, Powell and Hough in Europe at CERN favoured another method of data processing for bubble chamber pictures. They suggested an automatic scanning machine which scanned the photographic films by a so-called 'flying spot' and delivered its output directly to a computer. Human aid was only needed for the selection of relevant areas on the film (Galison 1997, p. 385f). A further solution was the possibility to send the photographic films for evaluation to other groups of scientists (Galison 1997, p. 413f), which raises the question of who the scientist is which executed the experiment. However, this is a question not to discuss here.

Physically the yield of the bubble chambers contains two strange hyperons, Σ^0 in 1956 and Ξ^0 in 1959, predicted by Gell-Mann and Nishijima (see chapter 3.2, p. 41), and a lot of resonances i.e., stimulated states of particles with higher mass and angular momentum (see chapter 3.1, p. 39) by the group of Alvarez (Galison 1997, p. 411ff). A further strange hyperon Ω^- was found in 1964 by a group of Samios in Brookhaven, predicted also by Gell-Mann (see chapter 4.3.2, p. 89; Brandt 2009, p. 411). To illustrate the problems of interpretation of the bubble chamber pictures, I will describe the observed interactions in the production and decay of the Ξ^0 and Ω^-. Both particles were created by an interaction of a K^- meson and a proton, giving a charge sum of zero and a strangeness of one from the K^- meson. Because the Ξ^0 in the first case has strangeness two and strange particles were created in pairs, besides the Ξ^0 a K^0 meson must occur, however, both particles are not observable because they have no charge. However, from their momentum and their lifetime, the supposed way can be calculated. At the endpoint of one of these ways, the K^0 meson decays into a positive and a negative charged pion which are observable. But the Ξ^0 decays in two neutral particles Λ^0 and π^0 again, because in a decay the strangeness could change only by one. Only the decay of the Λ^0 in a negative charged pion and a proton at the end could be observed again.

The decay chain of the Ω^- is even more complicated because of the strangeness three. Together with the Ω^- a K^+ meson is created, which are both observable, but also a K^0 meson, which is not observable, but decays in two oppositely charged (indirectly) observable pions. The Ω^- decays in a non-observable Ξ^0 and an observable π^-. The further decay chain is as described above but with a little difference, namely that the short-lived π^0 decays into two photons, which create each an observable electron-positron pair. However, these decay chains are only some occurring interactions among others, so that a lot more tracks are in a photograph which complicate the evaluation and must be excluded from the reconstruction.

5.3 Visual Techniques with Electronic Counters

In contrast to these visual techniques, new particles could be detected also with non-visual techniques, as the discovery of the anti-proton in 1955 and the anti-neutron in 1956 shows. For these discoveries, a system of magnets and counters were used (Brandt 2009, p. 347). Because the track of the particles could not be observed, it must be carefully excluded that other particles were counted accidently.

Anti-protons have the opposite charge as protons but have apart from that the same properties. Therefore, the negative charged particles, created from collisions with a copper target in the Bevatron, which had alone the energy to produce anti-protons at that time, were deflected by two magnetic fields for the selection of particles only with the right momentum. They were then bundled by magnetic lenses consisting of three quadrupole fields behind and before the first and second magnet, respectively. Additionally, the velocity of the particles would be measured by the time the particles needed to travel through a defined path. To do that two scintillation counters, counters triggered by light when ionizing particles traverse, were mounted in a definite distance to measure the time difference together with an oscilloscope, which also could show the time difference to take a photograph. An additional velocity measurement was taken by a Cherenkov counter, which only detected light at the specified angle matching the expected velocity region of the particle. A final scintillation counter registered if the particle correctly passed the Cherenkov counter and only if all conditions were fulfilled was the oscilloscope triggered and the photograph automatically taken. With this equipment in fact 60 events of anti-protons based on the interaction $p+p \rightarrow p+p+p+\bar{p}$ were found by Segrè's group in 1955 and published in the same year.

The proof for evidence of anti-neutrons was a bit more complicated because the chargeless neutron is not observable. An anti-neutron could be detected only based on the interaction $p+\bar{p} \rightarrow \bar{n}+n$ by an anti-proton prepared by the previously described equipment and then passing a sandwich of a lead glass target in a Cherenkov counter between two scintillation counters. If it was guaranteed that an anti-proton entered the target and the Cherenkov counter gives a large signal, but the two scintillation counters do not, then it was sure that the annihilation of an anti-neutron generated the signal. This experiment was executed by another group besides Segrè's also at Berkeley (Brandt 2009, p. 349). Additionally, annihilations of anti-protons and anti-neutrons using photographic emulsions and in bubble chambers were observed. The essential difference between the discovery of these particles and observations using visual techniques will be discussed at the end of this chapter. Before doing that the further development of the counter techniques should be investigated.

The next step was the development of spark chambers. The difference to the counters used before is that spark chambers could localize an event, while a counter could not. A counter could give only a yes/no decision, whereas a spark chamber could additionally show where the event occurs inside. The spark chamber is constructed out of parallel metal or conducting glass plates, which after some intermediate constructive steps could be triggered and render a track of a charged particle (Galison 1997, p. 478f). It worked with a high-pressure noble gas filling and high voltage square pulses applied with alternating sign to the plates. An advantage of the spark chamber was a longer sensitive time of several microseconds and the possible clearing of not wanted tracks by an erasing field with the best suitability for experiments with accelerators like bubble chambers, but not so expensive to build and to maintain (Galison 1997, p. 483).

The most remarkable result achieved with spark chambers, which showed the competitiveness to bubble chambers, was an experiment planned by Schwartz in 1960 'to test weak-interaction theory at high energies' (Galison 1997, p. 485). The V-A theory of weak interaction by Feynman and Gell-Mann (see chapter 4.3.3, p. 94) discussed the idea that an intermediate heavy vector meson may mediate the interaction between the four participating fermions. Because of the assumed high mass of the intermediate meson, high energies are needed to decide about its existence, which the then available accelerators could not deliver. However, Feinberg proposed some other processes which could answer this question i.e., if a muon decays in an intermediate meson and a neutrino, then there are two possibilities for the further reaction. On the one hand, if there is only one type of neutrino, a photon can be emitted, and the intermediate meson could recombine with the neutrino to an electron. But, if there are two different neutrinos, one for the electron and one for the muon, then this could not happen (Brandt 2009, p. 398f). Because the decay of a muon into an electron and a photon could not be observed, this speaks for the second option, which is the correct option as we know now. To show that the assumption of a muon neutrino, which cannot create an electron, is correct, Schwartz suggested the experiment with a spark chamber. Bubble chambers could not fulfil this task because they could not slow down neutrinos enough because of its filling with light hydrogen, in contrast to the filling of the spark chamber with a heavy noble gas. The experiment itself, published in 1962, used charged pions with high energies, emitted from the accelerator at Brookhaven, which after 10 m hit a shielding wall of also 10 m thickness to absorb all strongly interacting particles as well as charged leptons but let pass the neutrinos. The ten combined spark chambers behind this shielding had a weight of one ton and were surrounded by anti-coincidence counters and by a shielding of 13,5 m iron to exclude other particles than neutrinos from outside. Therefore, the only particles entering the spark chamber could be muon neutrinos according to the reaction ν_μ+n -> p+μ^- if there were found only muons and some conditions are met simultaneously. These conditions are: ‘(a) trigger counters indicate one or more charged particles within the detector, (b) the absence of signals from the anti-coincidence counters made sure that no charged particles entered from the outside, (c) the timing was right for the synchrotron [the accelerator] to produce particles’ (Brandt 2009, p. 399). If there were found muons and electrons in equal parts, then only one kind of neutrinos would exist, but that was not the case.

This was the experiment which gave fresh impetus to the counter fraction of the experimenters resulting in similar problems as with bubble chambers, namely the task of scanning and evaluating vast numbers of photographic films. To avoid this strange task, several attempts were made to avoid it by using e.g., television cameras or audio recorders to get an electronic form of the signal, which can be treated by a computer (Galison 1997, p. 493) as well for ‘data acquisition’ as for ‘check and control’ functions or ‘sample computation’ (Galison 1997, p. 491). ‘Data acquisition’ means here storing the signals of the counters on magnetic tape, ‘check and control’ the monitoring of experimental parameters for supervision, and ‘sample computation’ the direct interaction between the spark chamber and the computer with the computer functioning ‘as variable logic element’. Especially the last two points demonstrate an advantage of the logical over the visual techniques of the bubble chamber, expressed by Lindenbaum as: “Because of the almost immediate on-line computer data processing feature, this complicated counter system which would normally be inherently blind is given a remarkable degree of vision. One can now see almost instantaneously the progress of the experiment and make standard checks whenever desired” (Lindenbaum 1963, p. 300, cited in Galison 1997, p. 493). Galison continues: “Here we have a multiple meaning of the restoration of sight. At one level, the blindness of counters refers to their inability to deliver a visual record; they cannot produce the images of a cloud or bubble chamber. But at the same time, Lindenbaum’s ascription of vision is to the logic experimenters themselves. As never before the experimenters could ‘see’ the progress of their experiments – where ‘seeing’ meant being able to visualize what was happening as it happened and to exert control over the events proceeding in front of them [omitted reference to a picture]. Once again, we find an attempt to combine vision and manipulability by combining features of image and logic" (Galison 1997, p. 493). This statement expresses in the best way, how the analogy between the observed processes and the imagination of mechanical collisions in the mind of the experimenter works.

However, evaluation by computer could '*exclude* certain events from consideration' and therefore prevent discoveries, as was criticized (Galison 1997, p. 497, emphasis in the original text). Nevertheless, it was tried by various attempts to improve the equipment of the logic techniques and its evaluation with electronic means. As Galison formulated it: "At stake, I submit, is an example of a general process by which largely delimited components and procedures retain some aspects from their original meanings, gain new ones, and lose others as they pass from one set of technical uses to another" (Galison 1997, p. 504). Is this not the characterization of an analogy model with other words, if one compares it with the description of negative, positive and neutral analogies made by Hesse (see chapter 4.2.1, p. 56)? I think it is. But in some cases, an analogy could also hinder a new development as has happened with the multiwire proportional chamber. Most scientists thought that a multiwire array of wires in a chamber, constructed in analogy to a three-dimensional array of individual proportional counters, would influence each other, so that they must be isolated against each other by extra wires (Galison 1997, p. 511ff). In fact, if in the laboratory a negative pulse was sent along one wire, a positive pulse was created in the adjacent wire.

However, as Charpak and his technician Bouclier discovered this was not true for an avalanche of electrons moving towards one wire because this avalanche itself produced also a pulse in the adjacent wire which cancels out the pulse originated by the first wire, so that each wire works as a single 'cell'. This was the reason that the additional isolating wires could be abandoned, and the localization could be remarkably improved. This was the birth of the multiwire proportional chamber but the development by Charpak does not stop there. The relevant progress should be discussed after the presentation of a famous experiment with the multiwire proportional chamber where another important development played a role.

This experiment is in fact a double experiment, realized by different techniques at nearly the same point of time. One was executed by the group of Ting at the proton synchrotron in Brookhaven using the multiwire proportional chambers for registering the tracks, besides other detectors. The aim of Ting's group was to find possibly further heavy neutral vector mesons i.e., with spin one and negative parity, besides the three already known ρ^0, ω^0, and φ^0 (Brandt 2009, p. 439f). As neutral particles with spin one like the photon, they can decay into an electron-positron pair and, therefore, the experiment behind the target of beryllium hit by the protons was arranged symmetrically. The trajectories were led first through three magnets bending the trajectories. Cherenkov counters before and behind the last magnets, and at the end a lead-glass shower counter, ensured that only electrons and positrons were counted because only these cause Cherenkov radiation respective electromagnetic showers in matter. Multiwire proportional chambers for the trajectory coordinates were posited at four points: one before the last Cherenkov counter and another three behind it and before the lead-glass shower counter at the end, oriented in different directions and in defined distance to get the different coordinates. With this experiment Ting and his group detected a new particle in a narrow mass interval of about 3.1 GeV, called J by them in their draft paper from November 1974.

At the same time, they heard from the opposite process to that detected by themselves, observed at Stanford by the group of Richter i.e., the production of the same new particle in the collision of an electron and a positron (Brandt 2009, p. 441f). Also in this experiment, the multiwire proportional chamber was used to locate the trajectories of charged particles, however, in a new form. Generally, the whole construction as well of the accelerator as the detector was new. That was the important development I announced some paragraphs before. The accelerator was a positron-electron storage ring, brought to operation with this state of technology in 1972 after a development time of more than ten years at different locations in the world. In this ring positrons and electrons rotate in opposite directions and were brought to collision at a defined point, where the detector was located. This makes it possible to enclose the collision point with the detector; a measurement in (nearly) all directions could be implemented. In this way, the new detector was constructed over a length of several meters, concentric around the beam including the collision point. Starting from the beam pipe, it was surrounded first by trigger counters, then by four cylindrical multiwire proportional counters oriented in different directions to measure the coordinates of the charged particles. Outside the multiwire proportional counters were arranged again trigger counters for coincidence measurement with the inner ring of trigger counters, a large coil for a magnetic field parallel to the beam. Then, outside the coil shower counters to record the energy of photons and electrons, an iron yoke for leading back the magnetic flux was mounted and at the outer shell multiwire proportional chambers for the registration of muons which could get through the iron. All the signals from the counters were evaluated and visualized by computer programs. With this detector 'complex events could be registered and analysed in a way that had been possible before only with bubble chambers.

This goal was achieved in a collaborative effort: Richter had asked three groups from Berkeley, led by Chinowsky, Goldhaber, and Trilling, and another group from SLAC [Stanford Linear Accelerator], headed by Perl, to join in. The Stanford-Berkeley collaboration designed and built the *Mark I* detector, which became the model for the detectors of most later collider experiments. Several collaboration members had worked with bubble chambers before. That experience helped in the development for track reconstruction and for the visual display of the whole events' (Brandt 2009, p. 441f). This visualization means that, despite counter techniques are used, the tracks were reconstructed in analogy of their appearance in bubble chamber experiments, which again are visualizations of unobservable objects.

Richter and his group called their new particle ψ because they did not know about the experiment of Ting's group at that time. After some dispute about the naming J/ψ was accepted in the end. A further justification for the name ψ was found a year later in a further new particle ψ', which decays into ψ and two charged pions, forming in the visualization the letter ψ. The nature of the new particles was open in the beginning but then they were identified as two states of the combination of a new quark 'charm' (c) and its anti-quark ($\bar{c}$) forming a system called 'charmonium' like 'positronium', a short-lived system of an electron and a positron. The properties of this new quark, already predicted by theorists several years before but not intended to be found by the experimenters, were besides its much higher mass like that of the quark u but also its strangeness like the quark s. The quarks c and s form a doublet like the quark u and d and were called the second generation of quarks. The third generations in the standard model of particle physics ‘top’ (t) and ‘bottom’ (b) are found because of their again higher masses only several years later as well as the third generation of leptons the ‘tauon’ τ and its neutrino ν_τ for the same reason.

However, despite Mark I in its principal construction was the model for most further detectors, technical improvements were introduced in detail: a new instrument was inserted, developed by Charpak and already mentioned at page 115 in this chapter, namely the drift chamber. The drift chamber was in principle constructed like the multiwire proportional chamber, but also the drift time of a trespassing particle was evaluated. The drift time is the time which an electron, knocked out of a gas atom in the chamber by the trespassing particle, needed to drift from the track to the next sense wire. This time is proportional to the distance from the origin point of the electron (Galison 1997, p. 513f). In the early drift chambers designed from Charpak, the electric fields in the chamber were oriented parallel to the beam and the wire planes perpendicular to the tracks. This meant that the electric and the magnetic field were also perpendicular to each other, which led to difficult conditions for evaluating the position out of the drift time (Galison 1997, p. 560ff). But Nygren found by using parallel electrical and magnetic fields that the drift time was ‘strictly proportional to the distance’ (Galison 1997, p. 560ff) and the drifting electrons were collimated by spiralling around the magnetic lines, which makes the resolution with a factor of ten higher (Galison 1997, p. 470).

With this second generation of drift chambers, called time proportional chambers, in 1979 experimental results could achieved which successfully supported the predictions of the QCD (see chapter 3.3, p. 49): namely the annihilation of colliding positrons and electrons into two or three jets of hadrons according to the expectation that quark – anti-quark pairs fly in opposite directions or decay in three gluons perpendicular to the beam with such high energies that additional hadrons were produced by the colour field in between (Brandt 2009, p. 446ff). These two or three jet events, respectively, were measured directly in the complex new types of detectors, evaluated and stored in the computers and visualized on monitors or printers. Of course, these events are not the only ones occurring but the so-called ‘background’ of huge numbers of other registered events, classified as known, were filtered out, despite the danger of throwing away events which were wrongly classified as trash.

Another aspect of visualization is the use of simulations, applied always in the design of the large detectors (as well as accelerators). Simulations construct an artificial reality in theoretical as in experimental categories (Galison 1997, p. 691). This artificial reality is or at least should be, of course, an analogy to the situation it should simulate, with no large difference whether the simulated phenomenon could be observed or not. Thus, if e.g., a new detector should be tested, it could be fed with real events and the output compared by an evaluation program with stored data already available from similar events, and thus simulate the expected data from that detector. As further example, an evaluation program could be tested also by stored data of already evaluated and known events by comparing its output with the output of the evaluation program used before. For predicted phenomena e.g., possible decay chains, the expected output data of a detector for such an event could be made by hand and fed into the evaluation program to see if the expected result would occur.

Additionally, simulations could also be used to discriminate real events from the background, i.e., Monte Carlo simulations of expected event data e.g., to judge if hadron showers created by the interaction of neutrinos with a nucleon in the bubble chamber are not caused by neutrons in the shielding around the bubble chamber (Galison 1997, p. 749f). Monte Carlo simulations are the generation of event data based on (pseudo-)random[39] numbers. With the random numbers in this considered case, the expected neutron events including location and time were generated randomly and compared with the observed real events. Only if more real events were observed than calculated neutron events one can conclude that in fact the assumed neutrino events occurred. Moreover, the fraction of the real surplus events can show the frequency of the neutrino events. This example stands for a lot of other applications of Monte Carlo simulations used in the physical practice of experimentation. This kind of 'visualizing' of the unobservable neutral particles in interactions became an important and indispensable tool in high energy physics, mixing experimentation and simulation based on theoretical models.

In this chapter I have sketched the development of techniques of experimentation with unobservable objects by analogical visualization starting with the cloud chamber, via photographic emulsions, the bubble chamber and the visualization of counter results with the help of computers, to the 'visualization' of not trackable neutral objects e.g., by Monte Carlo simulations. Under 'analogical visualization' I understand the visual representation e.g., of tracks like that of the cloud or bubble chamber if the unobservable theoretical objects like the electron, protons or charged strange particles behave in a similar way as observable solid bodies in classical mechanics. Despite that, the evidence of the objects creating the tracks is never confirmable. In contrast, the counters were in the beginning far from the intention and much more from the possibilities of visualizing its measuring results. Nevertheless, the most modern types of detectors in high energy physics all allow us to visualize the observed events on monitors and printers with the help of computers, an aim and expectation which was formulated already at a time when it was out of reach:

"Concerning electronic detectors I want to relate what I heard Kowarski say at CERN in the early 1960s. At that time Kowarski […] headed the Data Division and supervised the development of a machine designed to automatically measure bubble-chamber pictures and transfer the data to an electronic computer. He said that the initial process in the bubble chamber was electric, the formation of ions along a path of a particle. That was followed by the thermodynamic process of bubble formation, the optics of photography, the chemistry of development, the electromechanics and optics of measurement, the encoding of data on punched cards and, finally, the entry of these data into a computer, where they were dealt with electric processes. He expected that eventually, also for complicated events with many tracks that at the time could be resolved only in the bubble chamber, detector systems would be available that used electric signals throughout, from the initial

[39] Pseudo-random means not really random but originated by an algorithm which seem to be random, algorithms like choosing the places behind the comma of an irrational number as π one after the other.

ionization to the final computing. Because of the advances in detector and computer technology that has indeed happened" (Brandt 2009, p. 334).

What were the reasons for that aim and these expectations? A hint in my opinion is given by the fact that most new elementary particles were discovered by visual techniques as the following citation from 1959 shows: "As Powell and his colleagues pointed out, by 1959 there were 20 particles in the zoo. Though the 'moment of discovery' of the electron is itself arguably undefinable, the authors assign it to the fluorescent screen technique, while they put the neutron down to the ionization chamber. Of the remaining 18 particles, 15 were located using visual techniques: 7 with the cloud chamber (e^+, μ^+, μ^-, K^0, Λ^0, Ξ^-, and Σ^+), 6 with nuclear emulsions (π^+, π^-, anti-Λ^0, Σ^+, K^+, and K^-), and two in the bubble chamber (Ξ^0 and Σ^0). Only three had been found using purely electronic techniques (anti-n, anti-p, and π^0): the power of the image to establish the reality of particles remained dominant." (Galison 1997, p. 230 after Powell, Fowler & Perkins 1959, p. XIV)

Why were so many particles discovered by visual techniques and only a few by non-visual counter techniques? The answer regarding the anti-particles of proton and neutron I think is obvious because they are anti-particles of already known particles. This means that all properties of the expected particle are precisely known, all one must do is to design an experiment which find it and to exclude the possibility of an error. That is, one does not expect to find complex new events difficult to analyse but wants only to confirm the existence of the otherwise well-known particle. Therefore, a visualization of the track is not needed interactions are directed in a way that they confirm the expectations. A surprise would be if no confirmation would be found. The conclusion is clear: visualization is needed for the identification and analysis of new and unexpected or unusually complex events like the discovery of the muon or the decay chain of strange particles such as the Ξ^0 and the Ω^-. Why then is visualization necessary in such situations despite being is obvious that the visualization is only an analogous description of the assumed processes happening in a region not accessible to the human eye?

I suppose the reason is that only through visual representation that human beings are able to discover and understand new or complex phenomena. This assumption contrasts to the often-advocated position that thinking mainly happens by using words, but I think those who thought about that were misled by their activity, which of course was made using words intended to formulate their thoughts for representation. However, ask yourself: how you do you dream, in pictures or in words? Or, when you try to understand a relationship, do you try to express it in sentences, or do you try to Imagine it in a picture or in a visualizing process? So, I think that thinking in pictures is the more original mode of thought for human beings and it happens not by chance that the first science developed was mechanics. In mechanics, what happens could directly be controlled and observed, step by step. Therefore, classical mechanics had a formative influence on the development of science. For the first time, phenomena in nature could be explained without the help of a god or other external forces. In consequence, phenomenological thermodynamics would be reduced to the statistical mechanics of not observable particles, i.e., atoms and molecules, by the analogical transfer of the mechanics of observable bodies to small unobservable ones. But also in electrodynamics where, as the name indicates, moving bodies, the electrons, play are role which are combined with the imagination of a wave analogous to water or sound waves as the interplay of moving points first with and later without a medium, a mechanical model is taken as the basis in the beginning. With the same mechanical models of thought, quantum mechanics starts in the beginning but introduces corrections to consider the discontinuities. Also, in high energy physics, the collisions of particles with targets or between two beams show an analogical behaviour. There is always a glimpse of the mechanical picture in the explanations and the experimental interpretation despite all modifications as the dualism of wave and particle. The visual wave model of water was already referred to in the explanation of the non-visual wave model of sound, so that it seems that almost alone through visual representation human beings are able to discover and moreover to explain new or complex phenomena.

Now, when explaining the massive dominance of visual techniques in the HEP experimentation practice, its role in compensating the limited access to the objects intended to observe must be considered. To do this I will look again on the decay chain of the neutral Ξ^0 particle, which is totally not observable indeed (see p. 112). How could one know that a Ξ^0 particle occurred? Firstly, it is taken for granted in the theoretical background that the Newtonian laws of motion and the conservation laws of energy and momentum as well as the Lorentz law of the electromagnetic force are valid and preserved. Secondly, it is accepted as true that charged particles cause ionized tracks in the detector space, which enable to determine its charge, its movement, and its mass[40]. This assumption gives the starting point for the observed decay chain. A K^- meson beam produced by the Bevatron accelerator and prepared with a definite energy was directed into the bubble chamber (Alvarez et.al. 1959, p. 215). From this beam, the track of a single meson could be observed because of the charge. If it collides with a proton assumed at rest and unobservable because contained in a neutral hydrogen molecule, the track of meson disappears if the collision creates two neutral particles. But then at first nothing could be observed because both created particles are neutral, only the decay of the neutral particles into oppositely charged ones could be observed. Therefore thirdly, on vertices of so-called V^0 events must be looked for (see p. 111). If there is only one such event corresponding to a K^- meson, the Ξ^0 could not have occurred because the Ξ^0 has strangeness -2 as postulated by the theory of Gell-Mann and Nishijima, and strange particles could only be produced in pairs. Therefore, because the K^- meson has strangeness -1, a K^0 meson must be produced in associative production and both strange particles decay in a V^0 event. However, double V^0 events by two strange particles could also be created by the collision of a π^- with a proton, so that angles and momenta of the created charged particles must be carefully determined. The angle and the momenta of one of the V^0 events were consistent with the K^0 meson decay but not with a Λ^0 decay. One could however not exclude before the investigation of the other V^0 event that, together with the K^0 meson, a Σ^0 hyperon as result of the collision of a π^- with a proton was created. An indication that in fact a Ξ^0 was produced was that the second V^0 event, which was consistent with the decay of a Λ^0 hyperon, and the track of the K^- meson were not coplanar (Alvarez et.al. 1959, p. 216). This nonplanarity means that there must be a further interaction in between with several possibilities, which must be excluded one after the other and only one remains. These possibilities are, besides

[40] I leave out here the methods used for calculating these quantities to focus more on the general scheme despite further assumptions are needed here as e.g., the validity of special relativity.

the Ξ^0, pure chance, an elastic collision with a hydrogen atom or the production of a Σ^0 hyperon with following beta decay of the Λ^0 hyperon. But all these possibilities could be excluded by kinematical or statistical reasons, respectively, so that the production of Ξ^0 and its decay into a neutral pion and the Λ^0 hyperon remains.

Obviously, the finding of the Ξ^0 hyperon was a complicated process, a succession of visual and mathematical steps, which could reconstruct unobservable parts of the decay chain beginning with the creation of two strange and because of their neutrality unobservable particles. The existence of these unobservable particles could only be concluded by theoretical reasons, based on symmetry considerations. The structure of these symmetries was essentially derived from analogies to already discovered and better observable particles and their structures bridging the gap of limited or even no access to the postulated objects. Thus, the theoretical generalisation of recognized structures mainly by analogy can compensate the restricting role of limited access.

5.4 Summary

One could say that high energy physics began in 1911 by two events, one discussed in this chapter and the other in chapter 3.1. The first, is of course, the visualization of the track of a charged particle by Wilson's cloud chamber using the ionizing power of the electromagnetic interaction by the fast moving charged particle, the other Rutherford's explanation of the charge distribution in an atom based on his scattering experiments with alpha particles. The latter is important because all experiments in HEP, as we have seen, are scattering experiments. The former is also important because all these experiments since the invention of Wilson's cloud chamber, but with some exactly reasonable exceptions, use a technique for the visualization of the particle tracks occurring in scattering experiments despite the postulated particles themselves not being observable. These techniques rest in fact first on the ionizing power of a fast-moving charged particle but later, with the introduction and further development of electronic counters, on the evaluation of their non-visual output by computers and visualization of these results in the same way as by the former techniques. This visualization of the tracks by analogy, produced by theoretically postulated charged or even uncharged unobservable particles, defined by some attributed properties like momentum, mass, spin, isospin and charge and their moving governed by the laws specified in special relativity, constitutes an explanatory picture adequate to the observed phenomena but with no warranty for its reality. Nevertheless, the adopted course using analogies to build up an explaining model is only workable because of the discussed limitations of observability if one does not want to limit oneself like Mach to pure descriptive methods, which are in my view like bookkeeping and not fruitful for the further development of physics.

So, we have essentially three types of analogies in use in physics, explanatory and predictive analogies in theorizing and analogical visualization in experimental practice, which are important contributions to a theory of development in modern physics.

6 On a Theory of Development in Modern Physics

In chapter 4.1 I gave a short sketch of a theory of theory development in physics. Now, after the discussion of the historical development of physical theories and experimental practices and the importance of the usage of analogies herein I will work out that sketch in this last chapter in much more detail and derive some consequences from it.

6.1 A Model of Development in Modern Physics

My model of development has seven steps, discussed in the following sections.

6.1.1 Step One: Observation of Phenomena

Theory development in physics based on observations in experiments has a long tradition for over three hundred years now and a lot of theories from philosophers and physicists were formulated to explain how physical theories were developed. Some of these theories were discussed in chapter 2 and 4.2 but analogies do not play a prominent role in any of them, except in the more sociological presentation of high energy physics by Pickering. As well, the consideration of the context of discovery in contrast to the context of justification find the necessary attention only in the last few years. However, the circumstances of the development of a new theory may perhaps 'lie in the air' in some respect as e.g., Duhem characterizes his description of the context of discovery (see chapter 4.2.1, p. 60f). Therefore, the context of discovery is negligible. But this judgement is only in agreement with his esteem of the historical development of science if science results in the end in a true description of the real world: only then it would be irrelevant how the description of the phenomena is found. Otherwise, as I argue for history is not only relevant for the development of science but determines its way for several reasons. I will come back to these reasons in the discussion following the presentation of my theory.

6.1.2 Step Two: Prior Association

Physical theories do not start out of the air and also are not simply problem solving, they always have a history where the same or similar phenomena came into focus. As the investigation on high energy physics in the former chapters show, they usually rest on previous similar theories explaining already known phenomena in an accepted way on the one hand and on extensions which incorporate observed new yet unexplained phenomena not covered by the accepted theory on the other. The new phenomena may be discovered during an observation by chance, a result of an experiment which was expected but not explained in a theory already or e.g., as an experiment to decide between different concurring theories and thus to give reason for a more developed theory. To demonstrate how the development of theories start in the beginning, I will review some of the discussed examples of the former chapters.

I will begin with the result of an experiment which was expected but not already explained, namely the discovery of the neutron in 1932 (see chapter 3.1, p. 35). A heavy particle with a mass like the proton was predicted by Rutherford as conclusion from the relation between the charge and the mass of atomic nuclei but this does not explain what holds the nucleus together i.e., the discovery of such a particle was expected but not explained. The explanation in an extended way was given in the same year 1932 by Heisenberg in his theory of a strong interaction between proton and neutron using some analogies to the binding of molecules. Heisenberg's theory explains more because it also gives reasons for stable and unstable nuclei and the form of their decay.

In contrast, a totally unexpected discovery was that of the strange particles (see chapter 3.2, p. 40), which did not fit into the theoretical model based on Heisenberg's and Yukawa's explanation of the strong interaction between protons and neutrons but gives reason to analyse the conditions of their occurrence and to formulate the assumption of the 'associated production' of strange particles. From this assumption and Gell-Mann's generalizing modification in 1953, four different theoretical models of explanation were developed, presented in 1955 on a conference, which compete for confirmation (see chapter 4.3.2, p. 84ff). The first theory by Fermi and Feynman assumed the strange particles to be resonances of nucleons and pions. The second and third theory by Pais assumed that the isospin space is extended to three dimensions consisting of conserved numbers for multiplet, angular momentum and parity and four dimensions consisting of two independent isospins of the ordinary type, respectively. Instead, the fourth theory by Gell-Mann assumed as in his former theories the conservation of the three quantum numbers Q, I and I_z but with the modified rules from 1953. Gell-Mann's theory predicted two particles not observed before, the other theoretical models several more, and only Gell-Mann's assumption was confirmed in the experiments. Therefore, Gell-Mann was motivated to clarify the relations of his theory to electromagnetic and weak interactions in an extended version of his theory in 1956 (see chapter 4.3.2, p. 87f).

These examples demonstrate that new phenomena come in different ways[41] into the focus of theoretical physicists and that the examples presented in no way exhaust the possibilities. However, in all these cases similarities to already explained phenomena were realized corresponding to Bartha's 'prior association' as one of two principles in his theory of analogical reasoning (see chapter 4.2.2, p. 68). In the examples above, the similarities refer in the case of Heisenberg's theory of the atomic nucleus to the binding of molecules by electrons even it is obvious that the scale of the model and therefore the involved forces differ in several orders of magnitude. Nevertheless, the same quantum mechanical model seemed to be applicable.

In the second example, the prior association was a similarity to the pions and nucleons and their interaction because the strange particles were also produced in a strong interaction $\pi^- + p \rightarrow \Lambda + K^0$ conserving isospin and decay in a weak interaction $\Lambda \rightarrow p + \pi^-$, $K^0 \rightarrow 2\pi$ changing isospin as Pais assumed from his analysis. However, when Gell-Mann has postulated his modification of the isospin assignment, four different prior associations were created in the mind of the involved scientists. Fermi and Feynman assumed the new particles to be resonances i.e., the binding of a pion and a nucleon with higher angular momentum like the stimulation of an electron in an atom to a higher orbit ignoring thereby the possible association of particles with a new property. In contrast to that, the originators or modifiers of the associated production, respectively, Pais and Gell-Mann, followed their former associations again and tried to specify the isospin in an extended form but in different ways. Especially, Pais demonstrates that there are a lot of different possibilities the prior association can lead to offering two different possibilities. Gell-Mann seemed to have a clearer imagination from the beginning already because he offered only one suggestion which he improved further in following steps and proved them to be the mostly adequate to the phenomena in the future.

[41] i.e., observation of expected but unexplained phenomena, unexpected phenomena, or during a deciding experiment between concurring theories (see first paragraph of chapter 6.1.2, p. 122).

6.1.3 Step Three: Explanatory Analogy

The similarities recognized in the prior association already lead to the next step in my model of physical theorizing, namely going back for an explanation to an accepted model of explanation for already known phenomena. For the theory of the atomic nucleus from Heisenberg, the used models are as mentioned above the molecules of hydrogen as well in its ionized as in its neutral state. The binding in these molecules is described as a quantum mechanical model where the two protons forming the nucleus of the two hydrogen atoms were orbited by one or two binding electrons, respectively, with opposite charge to the protons. The binding is generated by the electromagnetic attraction force of the opposite charges of the protons and the one electron or the two electrons, respectively. Because of the repulsion force between the two protons and the attractive force of the electron(s) to both protons an equilibrium in the distances arises with the electron(s) staying most probably in the region between the protons. If one assumes that the distances are drastically reduced and the exchanged electron-like particle has no spin, then the system of proton and neutron or of two neutrons were suggested as analogous to the described ionized or neutral hydrogen molecules.

The associated production in the second example was on the one hand a result of the analysis made by Pais but on the other based on the theory of Heisenberg and Yukawa on the interaction between pion and nucleon. As described in the paragraph before, the spinless but massive 'electron' i.e., the pion, mediates in this theory the interaction between the nucleons. That means e.g., a negative pion was absorbed by a proton creating a neutron and in the following the neutron decays into a proton, an electron, and a neutrino. A charged pion usually decays into a muon and a neutrino. During the production via strong interaction the isospin is conserved, whereas during the decay the isospin changes like the case with strange particles. Gell-Mann changed in his 1953 paper the theoretical attribution of the value of strangeness to mesons and hyperons, but this does not change the processes in a way that the associated model for the analogy must be changed.

The described explanatory analogy motivated by the prior association corresponds at the same time with the 'denotation' step of Hughes' DDI account. The explanatory analogy is characterized by the association of an analogy between the model of the source domain and the model of the target domain, which lead to the assumption that also the positive analogies of properties are transferable to the target domain (see chapter 4.2.2, p. 69). The first D in the DDI account of Hughes stands for denotation which uses on the one hand the established notions of the source domain for a description of the target domain. But also if, as in the case of strange particles, own names are given for them, the properties of the source model give an impression of those of the target model because a lot of a priori principles are transferred. On the other hand, denotation transfers the observed phenomena i.e., the tracks of the interacting particles and their observed or measured properties, into its description as already known or as new particles. Usually, the known particles were identified and denoted by the experimenters, whereas the new and unknown particles could only be identified by experimenters using their observed properties if they were expected from a theoretical prediction. Otherwise, as in the case of the strange particles, experimenters could only determine their properties and leave their explanation to the theorists.

In the four attempts for explanation of the paper of Pais and Gell-Mann only the model of Fermi and Feynman does not refer to the source model of the prior association described before. Instead, the prior association for the source model is the reference to a short-lived system of already known particles, here the pion and proton with a high angular momentum like the electron in a stimulated orbit of an atom or the behavior of an oscillating circuit (see chapter 3.1, p. 39). As we see here, despite a totally different prior association and in consequence a different approach for explanation with a different source model, also in this case an explaining analogy is introduced, which gives a lot of background information: one could say the source model of the analogy sets the stage for the development of the theory in consideration.

6.1.4 Step Four: Extension of the Explanatory Analogy

The three other attempts are all based on the isospin model of Heisenberg and Yukawa which describes proton and neutron as two states of one particle exchanging a pion as mediator of the strong interaction force expressed by the force law of Yukawa (see chapter 3.1, p. 36 and chapter 4.3.1, 81f) as in principle also the approach from Fermi and Feynman. However, one difference must be noticed, while the latter stays in a well-known model because other resonances are already known, the other approaches must extend the isospin model in the form already mentioned above. These extensions, whatever their direction may be, characterize the next step of my model of theory construction i.e., the adaptation and extension of the source model in the analogy to the new phenomena that occur in the target model. In Bartha's 'articulation model' of analogical reasoning, these possible extensions form the 'potential for generalization' as second principle besides his 'prior association' (see chapter 4.2.2, p. 68). However, these extensions are also needed because only then a connection to the new phenomena could be made which explains them. This explanation or 'demonstration' is also part of the second D of the DDI account by Hughes, the demonstration of the relations between the different properties of the new phenomena.

In case of Heisenberg's theory of the constitution of atomic nuclei, this demonstration consists of the transfer of the quantum mechanical calculation from the ionized hydrogen molecule to the deuterium core consisting of a proton and a neutron and its generalization to combinations of two or more protons and neutrons including the interaction of two equal particles. These combinations and their interactions require an extension to different kinds of energetic factors, a new strong one, which holds nucleons i.e., protons and neutrons together called exchange energy, one for the kinetic energy, one for the Coulomb repulsion, an attractive energy factor between the neutrons and a factor for the mass defect for the neutrons (see chapter 4.3.1, p. 77f). With this equation and some assumptions on their relative proportions, Heisenberg could explain the constitution of all known stable and unstable nuclei including if and why they are stable or not.

In a similar manner, the approaches of Fermi and Feynman as well of Pais and Gell-Mann work. Fermi and Feynman could explain with their ansatz the larger masses and higher angular momenta of their assumed resonances, although it later turns out that their approach was not tenable. Also, the two hypotheses of Pais to extend the isospin formalism by an additional angular momentum and an additional independent isospin component, respectively, goes in a similar direction and were justified in the first view because it is not directly obvious, which in direction the source model must be extended, when not much is known about the new particles. That the explaining extension delivered by Gell-Mann was accepted in the end was not settled from the beginning, and that Pais offered two attempts may be initiated by the fact that his first solution could not explain the cascade particle later called Ξ^-. Interesting in the second attempt of Pais and that of Gell-Mann is that both try to keep their former coordination of the value of strangeness to the fermions i.e., ½ by Pais and 1 by Gell-Mann. For Gell-Mann already, other factors seem to play a role, namely the relation of the masses between the different strange particles and the structure analogies of the multiplets of pions and nucleons (see the citation in chapter 4.3.2, p. 84ff). From these structure analogies he concludes the rules which governs the behaviour in strong, electromagnetic and weak interactions. These factors and rules, respectively, guide him for the next step in my model of theory production which belongs in my opinion as second part also to the demonstration in the DDI account of Hughes or in Bartha's terminology to a predictive analogy.

6.1.5 Step Five: Predictive Analogy

To remember, according to Bartha in a predictive analogy the association runs from the positive analogies to the hypothesis (see chapter 4.2.2, p. 69). In Gell-Mann's theory this means that the baryons should exist in two doublets, a singlet, and a triplet. One doublet consists out of the proton and neutron, the well-known nucleons, the singlet out of the hyperon Λ^0, the triplet out of Σ^+, Σ^-, and Σ^0, and the second doublet out of Ξ^- and Ξ^0. Except of the nucleons, all other particles are strange particles and two of them Σ^0 and Ξ^0 were not already known but predicted for symmetrical reasons. As well-known, the nucleons and the pions within its multiplet have only small mass differences and therefore the same should be true within the multiplets of the strange particles. The hyperon Λ^0 was the lightest hyperon and no other hyperon with a comparable mass was ever found, so that it must be a singlet. The next heavier two Σ hyperons have comparable masses but opposite charges. As with the nucleons, their charge should differ only with one unit as with the pions. Therefore, a neutral Σ^0 should exist analogous to the pions. Of the second doublet only the Ξ^- was known but it should have a partner as symmetrical counterpart to the doublet of the nucleons i.e., a Ξ^0 particle with about the same mass as the Ξ^- particle should exist, too. The first explanatory model by Pais was probably ruled out already for further considerations because it could not explain the Ξ^- hyperon. In the two other explanatory models by Fermi and Feynman as well as that by Pais more possible particles were predicted, e.g., in the Fermi-Feynman model exists no theoretical limit, in Pais second model the Λ^0 hyperon becomes a doublet and the Σ and Ξ hyperons even sextets. With these different predictions, a decision could be made between the explanation models and their predictions based on experimental results.

For the nucleus building theory of Heisenberg, the predictions consist of the stability or instability, respectively, of the atomic nuclei including the isotopes with different numbers of neutrons independent from the fact if these isotopes are already known or not. However, Heisenberg's theory could do even more than that, it could determine for instable nuclei the kind of their decay by construction of decay chains with a definite order of α-, β-, and γ-decays (see chapter 4.3.1, p. 79f). Only very few cases it could not decide. The predictive part of Yukawa's theory was that he took the exchange particle in the calculations of Heisenberg serious and that he could calculate the order of its mass from the size of the nucleus. Also in these cases, the predictions should be tested by experiments investigating the different isotopes, the decay chains, and with Yukawa's prediction to find the assumed particle with a matching mass and the other expected properties.

6.1.6 Step Six: Experimental Confirmation

This last sentence and the analogous at the end of the paragraph before mark the transition to the next but not last step of my theory of theory construction: to transfer the predictions in an arrangement appropriate for executing an experiment. That step is what Hughes in his DDI account calls 'interpretation' and is the meaning of the I in DDI. Interpretation is in some sense the reverse process to denotation: translating the theoretical description of the predicted particles into observable phenomena, usually again done by experimenters referring to the properties of the predicted particles or suggested experiments from the theorists, respectively.

For the theory of Heisenberg on the constitution of atomic nuclei predicted isotopes of a specific element could be separated e.g., by using mass spectroscopy. Mass spectroscopy is based on the fact that different isotopes of an element have different masses resulting from a different number of neutrons but all with the same number of protons. A mixture of isotopes found in nature could be heated and ionized and accelerated into a beam which is then sent through a strong magnetic field. This magnetic field deflects the direction of the isotope ions depending on its mass, so that they can evaluated in a counter. The relation between the counts of these counters determines the proportion relative of the isotopes. Decays with the production of β- and γ-rays could also be detected by counters (see chapter 5.1, p. 104), α-rays by scattering experiments or later after the development of Wilson's cloud chamber by its visible tracks recognizably by their high density (see chapter 5.1, p. 104). Depending on the type of decay, the atomic nucleus changes its properties: with α-decay it lowers its position in the periodic table by two places e.g., out of uranium to thorium, with β-decay it raises its position by one place e.g., from nitrogen to oxygen and with γ-decay it stays at its place but changes its spin. From the results of these experiments and their suitable interpretation one can conclude if the registered isotopes and their possible decays correspond to the prediction of the theory or not.

In Yukawa's theory, the predicted particle must be found according to a mass between proton and electron. Such a particle was found in 1937 in cosmic ray showers with a magnetic cloud chamber, but it shows a wrong property, namely traversing thick layers of matter i.e., there is no strong interaction. A track of the right particle, later called π-meson or pion, was found with a photographic emulsion in 1947.

On the page before regarding the theories for the explanation of strange particles, it was already stated that one can decide between them by their predictions. Whereas Pais's first theory was not compatible with the evidence of found particles, the other stayed in concurrence. From the three concurrent theories that of Gell-Mann was the easiest to test because its predictions were the most determined with the smallest number of particles for resonances on the one side and the strange particles on the other. The second place had Pais's second theory and the most difficult was that of Fermi and Feynman because also strange particles decay in protons and pions. The difference lies in the typical decay chains. Mainly for that purpose i.e., the confirmation of the strangeness concept, a new visual measuring instrument was built: Alvarez' 72-inch bubble chamber (Brandt 2009, p. 333), which was designed for the observation of the production of unobservable neutral strange particles and their decay into observable and therefore identifiable particles (see chapter 5.2, p. 111).

6.1.7 Step Seven: Acceptance of the Theory

From these examples is to conclude that the interpretation, as well of the theories as preparation for the experiments as of the results of the experiments if they correspond to the predictions, is not always easy. But only if all the interpretations agree in their results as e.g., in the case of Gell-Mann's theory of the symmetries of the extended isospin including the strange particles, it could be accepted, and its remained two competitors dismissed. With the acceptance or rejection, the last step of my theory of theory construction in physics is reached, but not the development of physics because in the experiments new phenomena could be discovered or questions, which were not explained by the existing theories, remain open, so that a new cycle of theory production and proving experiments can begin.

For better understanding and overview, I will present my theory in the following also in a condensed form:

Summarized Model of Theory Construction

- Observation of new phenomena (discovery)
- Realization of similarities to other phenomena (prior association)
- For explanation going back to an accepted model of explanation (source model) of these other phenomena associated by the similarities (explanatory analogy) (denotation)
- Adaptation and extension of the source model to the new phenomena (generation of the target model i.e., the potential of generalization) (demonstration)
- Investigation if the properties of the extended model contain possible new phenomena by comparing them with the properties of the source model (predictive analogy) (demonstration)
- Test of the predictions of the adapted and extended model by experiment (interpretation), perhaps observation of new phenomena -> next step
- Acceptance or refutation of the target model as new theory

In my theory of theory construction in physics, analogies are an important aspect, so that it seems to be important as well to recall the background and the reasons for this assessment. I will look at four points in connection with analogies in the order of abstractness: observability, visualization, history, and metaphysics.

6.2 Analogies and Observability

Concerning observability, I maintain that originally only observable objects in the world can be explained because only could be recognized by their behaviour. To understand this claim, several aspects must be considered. Firstly, explanation means as I understand it to demonstrate relations between phenomena in the world which allows other human beings to understand these relations as adequate to the observed phenomena and accept them for the own actions if confronted with these or similar phenomena. Therefore, one is interested only then in explanations if the considered phenomena are relevant for her in some situations. A description of phenomena only for classification, as Duhem asked for, seems to be not enough because this does not help for decisions how to act in relevant situations if this classification is not bound to experiences, which connect them to former done or observed similar actions. That means a classification alone delivers no context. One could argue that such relevant situations are part of common life but not of science. However, also science and the results of scientific activities happen in common life, also when these activities refer not directly to common life, they have a relation to it and scientific explanations have consequences in technology and common life as e.g., the effects of climate change or the wave theories for aerodynamics. Thus, scientific explanations have relevance for the orientation in the world.

Secondly, 'observable' means originally 'directly perceptible' with one's own senses. Perception with one's own senses is the source for experiences from the beginning in one's life, also when later e.g., in school and in science more abstract experiences dominate. E.g., the experience of the own mother, I will call her MUM, is one of the first experiences. This is the person a child can trust. Later, when it begins to play with other children, the child will learn by observation that other children also have persons they call MUM. The child must extend its concept of MUM. It knows by similarity what a MUM is and can transfer it to the other MUMs i.e., the MUM becomes mum. Somewhat later the child may learn that also animals have a mum, etc.

In a similar way other concepts can be learned, the 'dictionary' in the mind grows. The growth goes in different directions, in the extension of concepts, in compound concepts like 'airport', in concepts of concepts and others. In this way the concepts can become more and more abstract and lead to the development of more sophisticated observations, which result in the end in science. Thus, all concepts also in science rest in principle directly or indirectly on experience. However, also there the observations begin with directly in their relational details observable phenomena. Such phenomena are the movements of rigid bodies despite an unobservable component already, namely the force which originates the movement. However, this unobservable component can by analogy replaced by the imagination of the force which oneself uses when e.g., throwing a spear or hammering on an anvil[42]. Classical mechanics could develop evaluating observations as the free fall or the movement of celestial bodies. Here is a remark needed regarding my theory of theory development in physics: an objection was that e.g., Newton's theory on gravitation does not use analogies to explain the equation of the gravitational force. In one respect this is correct, but this does not mean that Newton refrains from analogies as we will see in a moment. Also, Duhem criticises a contradiction between Kepler's laws and Newton's laws of motion and gravity arguing that Kepler and in his derivation Newton, too, assume the mass in the centre at rest, which is not compatible with his gravitational law because it has as consequence that the centre also must move (Duhem 1998, p. 255f). Duhem suggests that not induction and generalization from observations were responsible for Newton's result as assumed reason for gaining knowledge but ingenious intuition by choosing the right solution

[42] Not surprisingly this analogical concept lead in its origin first to the animation of forces in nature like spirits and gods before a force became an appearance in nature without an animation.

from an unlimited number of possibilities (Duhem 1998, p. 264). However, Newton's intuition was stimulated and, in its possibilities, limited by two available models of motion i.e., the laws of celestial motion without any notion of force by Kepler on the one hand and the laws of falling bodies obeying the terrestrial force of gravitation on the other.

My claim is that the combination and extension of these both source models lead Newton to the analogous or, as Duhem would express it, 'symbolic and approximate' (Duhem 1998, p. 258) target model of the gravitational law. As I see it, Newton knows of the relation i.e., the analogy between these models because he concludes in proposition III, theorem III and the following scholion of his 'Principia' that the centripetal force directed to the centre of another body adds to any other motion inclusive accelerations of this other body, so that the steady covering of an area always could be taken as indicator of the centre. That is Kepler's first law read in the opposite direction (Newton 1999, p. 62f). The point, where the critique on my theory was correct, is that Newton gives no explanation for the origin of the gravitational force because he does not want to make hypotheses. But also Heisenberg has in his theory of the atomic nucleus all problems concentrated in the constitution of the neutron and to leave its resolving for the future. This means, a theory could leave open problems where no analogy gives an idea for their solution.

Also, in the time of Newton observations of the movement of nonrigid bodies i.e., fluids and gases like the air come into focus. To explain their behaviour, assumptions about their constitution were necessary. Therefore, in analogy to systems of rigid bodies, fluids were handled like consisting of small mutual adjustable particles. With the help of this concept, the motion of waves in water and air could be described. Also, for light the same kind of wave as in the air was suggested. The properties of the used analogies suggested then the idea that there must exist an invisible medium called 'ether' for transportation of light waves in the same way as for water waves exist the observable water and for sound waves the only indirectly e.g., via compression effects observable air. However, in this case the explanatory power of the source models of the analogies was misleading as we know now: waves could also spread without an explicit medium i.e., in vacuum.

What this fact demonstrates, however, is that analogies allow one to extend explanations to indirectly observable objects and properties. Of course, they sometimes can cause an error as later results then show. Such extended explanations based on analogies runs all the way through the development of physics, as in the next step of development regarding thermodynamics, too. In the beginning, the flow of heat was seen like an invisible and imponderable fluid, an invisible and imponderable substance flowing from one body to another extending the observation of fluids to an unobservable object. In the same way as the masses of fluids, it was assumed that the quantity of heat was conserved and in combination with mechanical work the sum of both will be resulting in the principle of energy conservation. For energy conservation another property of heat must also considered, the radiation of heat, which is often accompanied by light or more precisely vice versa. This prompts observers to the assumption of a similar nature of both but does not correspond to the model of a fluid and stays a problem not solvable at the beginning of the nineteenth century (Baumgartner 1829, p. 454ff).

At the same time, two other phenomena in nature were considered in analogy to fluids but with a further complication (Baumgartner 1829, p. 458ff and 501f). These phenomena were magnetism and electricity, and the complications were that magnetism and electricity both occurred in two kinds, the magnetism with always two poles of a magnet and electricity separable in two bodies which either attract each other or repel each other. Therefore, it was assumed that there are in both cases two invisible and imponderable fluids, which divides always down to the smallest particles i.e., atoms in case of magnetism. In case of electricity, it was somewhat more complicated: the number of quantities of electricity was assumed as usually the same in electrical uncharged matter but different in charged matter. Regarding electricity, other researchers assumed instead only one fluid with a surplus or deficit of it in the charged case. Again, we have explaining analogies here for phenomena based on unobservable postulated objects. Under these circumstances of numerous different unobservable substances, it is understandable that physicists like Mach or Duhem argue against the consideration of unobservable objects in general and for the restriction on the description and classification of phenomena.

However, only a few years after Baumgartner's book Ampere published a mathematical theory on the electrodynamical phenomena where invisible forces act on invisible electric current elements (Duhem 1998, p. 262) and magnetism is explained by 'moved electricity' (Sommerfeld 1977, p. 112). I will not refer here to the used analogies but will emphasize that here only objects observable in their actions i.e., current elements are used to explain the properties and behaviour of other indirectly observable objects i.e., other current elements. This chain of extensions from observable objects and their behaviour via less observable objects and their behaviour to lesser observable objects and their behaviour allows to extend the limits of observation and explanation by analogies more and more because regularities in nature seem often to obey similar patterns which could be used for further explanations. One reason for this fact seems to be that evolution prefers organisms best adapted to its environment i.e., such which could respond to the occurring regularities and irregularities in the world in the best way. In physics, the observed regularities have started with the physical properties of rigid bodies in the beginning. Nevertheless, their observed results and concluded principles govern still the interactions of the colliding of more or less unobservable particles as e.g., electrons, protons, neutrons, mesons or quarks show. The participating forces are modelled like gravitation, measurable in the electromagnetic case or postulated and concluded only in strong or weak interactions. They are constructed in a similar form as the gravitational force and still not really explained as their original model since Newton's time, too. Therefore, it may be observations can tell us only something about the structures of the regularities in nature but not about its causes. Some information, how we get the physical abstractions of the observed regularities of phenomena in the world, may give our sticking on the evaluation of visual representations of the executed physical experiments described in chapter 5.

6.3 Analogies and Visualization

In chapter 5 we followed (after there were made first appliances of scintillation screens and ionization chambers for the optical indication of canal, cathode or X rays) the development of visual techniques for track presentation of charged particles with the cloud chamber, photographic emulsions, the bubble chamber and in the end with electronic counters and detectors evaluated and visualized by computers. The visualization of the tracks was characterized as an analogy to the postulated but directly unobservable and only by ionization detectable charged particles (see chapter 5.1, p. 103). By these techniques decay chains of particles could be made visible and even the gaps caused by neutral particles could be filled by calculation of the expected properties of these neutral particles in their production and their decay.

But a question remains: what motivates physicists to represent the interactions of the unobservable particles by visualized analogies and why is visualization preferred against other kinds of representation? Nersessian and Gooding[43] answered the first part of this question in the example of Faraday that he was against Newton's conception of action at a distance and found the inspiration of his 'lines of force' in the observation of the at that time well-known behaviour of iron filings in a magnetic field ordering in lines between or surrounding the poles of a magnet (Nersessian 1984, p. 43). For Gooding this means that 'mental processes and material manipulations are complementary: the agency whereby observers construct the images and discourse that convey new experience embraces both' (Gooding 1990, p. XV). This account assumes an intertwining of experimentation and theorizing of 'a concrete visual image' for the illustration of the notion of 'continuous, progressive transmission of the action and of vague analogies between electric magnetic actions and 'known' progressive phenomena'. The intertwining was initiated by Faraday's discovery of electromagnetic induction and his attempt to explain how a change in the magnetic force can generate a current whereas a constant force cannot (Nersessian 1984, p. 145). Here we have an analogous situation as I have discussed regarding the tracks of particles in HEP. The iron filings illustrate the behaviour of an unobservable object, namely the magnetic force. However, the intertwining does not really answer the question above. In fact, that was not really the aim of Nersessian and Gooding, their aim was to explain the construction of meaning in science as their book titles show.

My answer to the first part of the question above is, going back to the former paragraphs on the observability of phenomena: it is obvious that modern science began with classical mechanics where rigid bodies were observed directly at night in the heaven moving around the sun or other celestial bodies or as falling bodies on the earth. The tremendous success of the Newtonian theory of dynamics and gravitation was for a long time the leading model for scientific representation, as well in its accuracy as in its generality of meaning and applicability, so that their general structure was tried to transfer also to other domains of science as biology or even humanities. In some of its essential parts e.g., regarding Newton's second (F=ma) and third law (action=reaction) it is valid today. Therefore, mechanics had a forming influence on scientific development. Further, theoretical reasons suggested as constituents of matter atoms and molecules. Third, the tracks observable in the cloud chamber and the following visual technique before electronic detectors arises look like the trajectories of particles, and fourth, theoretical calculations of properties associated with the tracks show compatibility with particles with specific properties. All these indications seam to speak for the existence of particles generating tracks by ionization and support the representation of the results of electronic detectors as visualized tracks, too. However, one cannot conclude that these results show

[43] Both, Nersessian and Gooding, are strong advocates for the consideration of the context of discovery in the investigation of the scientific enterprise, too, but not already mentioned in chapter 2.

the evidence of these particles because this would require that an alternative independent access to them were possible to confirm the assumptions.

For the second part of the question, one must look on the evolution of mankind. The visual system is the most developed sense of human beings. However, this could not be the only answer for the preference of visual representations. Another reason is the manipulatory skill and the high developed brain. Together they compose an efficient problem-solving system, which finds in another combination with other senses not the same efficiency. Therefore, I think visualization is the most preferred method to represent complex relationships. This is also true for mathematical relationships because a written equation is also a visual representation.

6.4 Analogies and History

Here, it is appropriate to come back to the history of physical development. I have sketched above some episodes from the beginning of the development of physics and physical theories. Now, I will change the focus and will show that the observed phenomena and the used analogies determine the direction of the development of science in some restricted sense. To do this, I will resume the examples of the beginning of this chapter and their analogies and will then look back to prove if the development occurred could be predicted in the long run. Only if this last case could happen, the direction of the development of physics in history could be assumed as determined. Determined has therefore following Hacking two meanings, one regarding the 'form of knowledge [which] represents what is held to be thinkable, to be possible, at some moment in time' (Hacking 1999, p. 170) similar to Friedman's relative a priori (see chapter 4.2.1, p. 63). The other is what Hacking calls 'inherent-structurism', which should express the position 'that the world may, of its own nature, be structured in the ways in which we describe it' (Hacking 1999, p. 83). Both meanings are relevant here, the first says that principles which are taken as granted and which are determined by the analogies used in the development of theories are not fixed for all time, the other assumes an 'inherent structure' in the world which may predetermine the direction of possible development in science. I will argue for some determination in the first meaning but will distinguish in the second meaning two parts. The first part in this second meaning is that I also assume an inherent structure in the world (see remark in chapter 4.1, p. 52) but will not agree to the determination of development in the long run. The reason for this distinction is that of course some regularities are observed in nature, and I suppose that these are an indispensable requirement for the possibility of life but on the other hand only these regularities are observable, still never the objects or substances they rest on. For the observation of these objects or substances an interaction with some regularities is needed which could be recognized but this could not explain the object behind. The object behind remains as an unknown Kantian 'thing in itself'.

Supporting this last view in a less thorough form seems to show the following citation from Rutherford in his 'Bakerian Lecture' on the 'Nuclear Constitution of Atoms' regarding the prediction of the neutron:

"On some circumstances, however, it may be possible for an electron to combine much more closely with the H nucleus [the hydrogen nucleus, i.e., the proton], form a kind of neutral doublet. Such an atom would have novel properties. Its external [electric] field would be practically zero, and in consequence it should be able to move freely through matter. Its presence would probably be difficult to detect by the spectroscope, and it may be possible to contain it in a sealed vessel. On the other hand, it should enter readily the structure of atoms, and may either unite with the nucleus or be disintegrated by its intense field, resulting possibly in the escape of a charged H atom or an electron or both" (cited after Brandt 2009, p. 210).

This last-mentioned process was not found at that time despite some trials by the Cavendish group to find it and then probably forgotten in the years after. Only ten years later, Bothe and Becker in Berlin found an unexpected radiation by irradiating the light elements beryllium, lithium and boron with alpha particles from a polonium source. They assumed this radiation to be high energetic gamma rays although with some peculiar properties. Whereas Bothe and Becker used a Geiger counter for their experiment Curie and Joliot in Paris applied an 'ionization chamber connected to a very sensitive electrometer' in their further investigation (Brandt 2009, loc. cit.). But in contrast to Bothe and Becker, Curie and Joliot put a blocking material with a high density of protons i.e., paraffin between the source and the detector registering a large rise of the radiation. They also assumed a high energy gamma radiation produced by the Compton effect on the protons in the hydrogen nuclei. Problematic was the needed high energy for such a gamma radiation of 50 MeV never seen before from radioactive sources. In 1932 Curie and Joliot gave a report of their findings in Paris which motivated Chadwick, who knew about Rutherford's theory and heard of the report, to repeat the experiments and concluded from their results that the radiation consisted of the neutrons predicted by Rutherford. Does this process of discovery of a new particle support the assumption of a determination of physical development in the long run? I think it does not.

The reason lies in the way how the process of development for the explanation of the constitution of the atomic nucleus continued. Of course, Rutherford had predicted some properties of the neutron in the right way such as the free motion through matter and that it might unite with the nucleus, but he had given no argument what such a united nucleus would hold together. As mentioned in chapter 3.1, p. 35 and 4.3.1, p. 76, it required that the neutron must have spin zero or one in contradiction to the results of the spin measurements of nitrogen which demands that the spin of the neutron must be one half as with proton and electron alone. If the development were determined, then the possibilities to choose from were limited but the theory of Heisenberg, developed after the discovery of the neutron, does not directly attack the problem of the force which holds the atoms together. In contrast, his approach started with assumptions on symmetries between the proton and the neutron and considered the question what happens when both particles were exchanged. For this consideration Heisenberg viewed on proton and neutron as one particle with two states bound by an exceptional electron with no spin, which exchanges at the same time the charge between them. This picture was supported by the analogy of the ionized hydrogen molecule, which has the same structure but in a much larger dimension and corresponds to a lower binding energy on the one hand and the two possible states of the electron on the other. This was a direction of development Rutherford could not foresee because both the binding by an exchange energy and the two-valuedness of the electron were not known in 1920 where only Bohr's model of the hydrogen atom was available.

Much more in the case of the structure of the multiplets of strange particles, where four concurring models were discussed by Pais and Gell-Mann, the direction of development was open because all were supported by analogical models with well-established ingredients and only the first model of Pais had some evidence against it. In review, one could argue Gell-Mann's model was the most simple and economical but that is not very convincing because the model of Fermi and Feynman was more economical in another respect, it needs no new principles for explanation like associated production or something like that.

Another argument for the determination of physical development could be that the development goes in the direction of truth but how could you know if that what you found is the truth. Thus, you cannot be sure about the direction of the development and, therefore, it cannot be planned because always phenomena can occur which contradict the adopted direction because the only knowable structure is that of the interaction (see p. 133 above).

Nevertheless, there seems to exist a belief that it can inferred from the observable to the unobservable from the beginning of science development in the seventeenth century. Namely Newton as well as his contemporaries extended the properties of rigid bodies as extension and impenetrability to the invisible parts of them in the same way as it was observed in the visible region. Newton states this claim in his third rule of the third book in his 'Principia' as follows:

"The qualities of bodies which admit neither increase nor decrease of intensity, and which are found to belong to all bodies within reach of our experiments are to be esteemed the universal qualities of all bodies whatsoever" (cited after McMullin 1978, p. 7; see also Newton 1999, p. 380).

McMullin calls this inference transduction after McGuire and classifies it as a form of induction but more restricted insofar as it concludes 'from particular instances to other instances *of the same kind*' (McMullin 1978, p. 15, italics in the original). Transduction in this sense is not an analogy because of the restriction to instances of the same kind.

According to Newton, this rule 'is the foundation of all philosophy' (cited after McMullin 1978, p. 18) because if these conclusions could not be made then no physics were possible. Physics rests on regularities (as life does but without universal validity) which are universally valid because otherwise no reliable order of events was possible, no experience of this order, no method to record this order and formulate it in rules collected in physical science. Essential is not the type of universal quality but that there is one as e.g., impenetrability shows: indispensable for Newton, but dropped in quantum theory. Friedman calls this kind of requirements relative a priori principles (see chapter 4.2.1, p. 63 and p. 133 above).

As just this last example indicates, such qualities can change, and it is not determined in which way they change. The way they change is levelled by using analogies modifying and extending the presumed source models I suppose but this causes not a determination in the long run as the just mentioned example shows. Especially, Mach and in some respect also Duhem and some others have tried to abandon such universal rules concluding on unobservable objects by restricting science to only the description and classification of observable phenomena alone. They rejected every trial to explain phenomena by unobservable objects (e.g., Mach 2014, p. 115; Duhem 1998, p. 20f; this is of course also a general rule) but could not convince the majority of scientists. However, if they had reached their goal what direction of development physics would have taken? The kinetical gas theory was explicitly rejected by Mach and all quantum theories perhaps would not be developed. Spectroscopic phenomena would only be coordinated to the chemical elements but speculations on the constitution of them and its consequences for the order of the periodic table of elements were not possible. I also think suggestions on the fusion reactions in the sun and other stars were missing. Considering these hypothetical consequences, I think it is obvious to conclude that the direction of development in physics and other sciences is dependent from the history but open in the future. Nevertheless, the immediate acceptance of an explanatory theory is directed by the used analogies which gives the framework for thought and supports the experimental results obtained from successful predictions and in the long run limited trends set by (temporarily valid) universal rules as such formulated by Newton. To summarize, development in science is more like open biological evolution than planned development.

6.5 Analogies and Metaphysics

One aspect with a relation to the universal rules mentioned in the paragraphs before remains to be discussed, the relation between analogies and metaphysics. Metaphysics means nothing observable i.e., it means conversely the concepts which are developed by human beings using categories made by experience with phenomena of the world but in abstraction from them. Hofstadter and Sander argue convincing in my view that categories are built up by analogies (Hofstadter & Sander 2014, p. 36f) e.g., in the form sketched above with the analogical extension of 'MUM' (your own mum) via MUMs of your friends, of all human beings to the general category of mothers of all animals including human beings (see chapter 4.2.2, p. 67). Thus, the mothers of all animals including human beings are phenomena in the world but the category of these is an abstract notion formed by the concept of 'all mothers'. In physics mass, energy or motion are also such concepts and Newton's rule three cited above as well. Also, as we have seen with the concept of impenetrability, it must be compatible with the observed phenomena i.e., in this case rigid bodies and its parts but not be realistic from the actual point of view. However, already at Newton's time there were counter examples excluded from the used concept because different fluids or gases could mix i.e., could penetrate each other but were excluded from this concept assuming that the primordial particles of the fluids or gases could mix but not penetrate each other. Again, one can conclude that these categories must be compatible with the considered phenomena but not realistic in the sense of a true picture of the world. These examples motivate the same suspicion also for other concepts used in physics and more: if our quantum theoretical picture of the constitution of matter is more realistic than that of Newton's time then we cannot trust the concepts based on the phenomena perceived by our senses and the extrapolation according to Newton's rule three may lead to a non-realistic concept. What can we trust then? Is scepticism the last resort?

I do not think so because a better way out is to combine the different views of the different sciences into a coherent worldview assuming that their different approaches show e.g., to consider what Duhem[44] has pointed out in his critique on the so-called 'experimentum crucis' (Duhem 1998, p.

[44] Similar is Elgin's (see chapter 4.2.1, p. 54) or alternatively Neurath's position that science should be like a boat which is changed in its construction during its voyage on the sea: Es gibt keine tabula rasa. Wie Schiffer

266f). An ‘experimentum crucis’ is an experiment which should decide between two theories corresponding better to the phenomena or not, respectively. Duhem argues that an ‘experimentum crucis’ is not possible because it is not decidable which hypothesis in the whole system of propositions in a theory is wrong but only that the result of the experiment is not supported by the theory. The consequence from that result is that all scientific knowledge must be combined into one consistent holistic picture where findings from one science support findings in another e.g., where also knowledge is considered on the perception and evaluation of sensorial data of animals. This picture may be criticized as constructivism because coherence could not replace a realistic worldview but that is the best to achieve. The consequent comparison within different sciences with observed phenomena and the results of their experiments should prevent a too one-sided constructivism because different sciences take a different point of view which may uncover contradictions and incompatibilities between them, and which could then be resolved by new approaches.

Of course, at least physical theories, but if Hofstadter and Sander are right that ‘analogies are at the heart of all thinking’, then in fact all theories are essentially based on analogies for explanation and prediction. If this assumption is right then one would expect that phenomena with no satisfying explanation suffer under the lack of a corresponding association for a source model which allows to formulate a hypothesis or neutral analogy supporting conclusions to positive analogies in the target domain (see chapter 4.2.2, p. 69). An historical example for this expectation of a lacking explanation is the constitution of neutrons for Heisenberg in his theory on the composition of the atomic nucleus where he could explain this composition but shifting all the open problems to the neutron. The reason for the unsolved problem with the neutron at that time was the beta decay, which emits electrons with no sharp momentum. That was not compatible with the conservation of energy and momentum. Therefore, Heisenberg, perhaps under the influence of Bohr, considered to drop the demand of energy and momentum conservation in the atomic nucleus associated with a heavy conceptual loss in physics. The ‘despairing solution’ by Pauli, as he called it himself, was somewhat later his famous neutrino hypothesis, the at that time unthinkable existence of a new unobservable particle explaining the observed continuous spectrum of the electron, produced by the variable distribution of energy and momentum on two instead of one particle.

Two contemporary examples for such problems of explanation are the so-called ‘dark matter’ and ‘dark energy’. ‘Dark matter’ is the assumption of additional invisible i.e., ‘dark’ matter in far galaxies and galaxy clusters for which calculations show a much higher dynamics than compatible with visible matter like stars and interstellar dust increasing with the size of the galaxy. Explanatory background is the theoretical principle of energy conservation in a closed system which result in a relation of potentials i.e., gravitational to kinetic and rotational energy, of two to one. In galaxies and galaxy clusters, however, the spread of the rotational energy of the stars measured via redshift and the calculated potential of the visible stars and dust clouds differ in a relation up to four hundred to one which needs, if restricted as assumed to gravitational effects, much additional masses for

---

sind wir, die ihr Schiff auf offener See umbauen müssen, ohne es jemals in einem Dock zerlegen und aus besten Bestandteilen neu errichten zu können. Nur die Metaphysik kann restlos verschwinden (Neurath 2009 [1932], S.401). To the last sentence in the citation however I do not agree because a priori principles are always metaphysical.

explanation (Kapferer 2018, p. 33ff and p. 45). The potentially existing invisible masses are orders of magnitude away for a satisfactory explanation till today despite intensive search. Thus, the concept of dark matter with only gravitational interaction as attempt to explain the dynamical behaviour is not successful currently. In some sense this concept is comparable to that of the neutrino but for the neutrino later evidence was found that the suggested concept is fruitful, however, not so for dark matter today.

A similar case is the so-called 'dark energy', also postulated in cosmology, and with a nature which is also not understood (Liddle & Loveday 2008, p. 86). 'Dark energy' is needed for explanation of the measured overall redshift of all observable matter which shows that the expansion of the universe is accelerating, and which is possible only with an energy surplus comparable to an exploding gas where the higher temperature and pressure in relation to the surrounding air drive the expansion.

In both these cases, dark matter and dark energy, where the nature of these phenomena is not understood and hypothetical explanations using the notion 'dark' are given I suppose that according to my theory the suitable analogies for an explanatory concept are not found today. I suppose phenomena which have no corresponding suitable analogy have no satisfying explanation.

However, I will go a step further: phenomena which have explanations not compatible with each other have no satisfying explanation as the following example of quantum gravity will show.

Since about fifty years it was tried to combine the two most formative theories of the twentieth century, general relativity and quantum theory, in a common theory called 'quantum gravity' but with not really successful because they contradict each other in some respect. Both are extraordinarily successful in different regions, general relativity in cosmology at very large ranges and quantum theory in the constitution of matter at very small ranges. Both imply to describe all physical phenomena in the world, but general relativity assumes a continuous description of the world governed by energy density according to Einstein's famous equation $E=mc^2$ and quantum theory a discrete one governed by Planck's action quantum h. Also different are the handling of time and the relation to background conditions, where in general relativity the four coordinates of spacetime have no explicit direction but in quantum theories for time they have, where general relativity is background independent and quantum theory dependent from background conditions. This makes it extremely difficult to construct a theory which overcomes these difficulties and I suppose one will only succeed if one finds a structure as source model combining the different properties of both theories in a way where they do not contradict but coexist depending from the point of view. Such a concept will then allow to set up analogies which will lead to general relativity and quantum theory as the two ends of a horn I suggest.

6.6 Summary

In this chapter the sketched theory of development in modern physics in chapter 3 was developed in detail. My theory consists of seven steps:

- Observation of new phenomena (discovery)
- Realization of similarities to other phenomena (prior association of Bartha)

- For explanation going back to an accepted model of explanation of these other phenomena (explanatory analogy of Bartha) (supports the denotation of Hughes)
- Adaptation and extension of the source model to the new phenomena (target model i.e., the potential of generalization) (corresponds to one part of the demonstration of Hughes)
- Research if the property of the extended model contains possible new phenomena by comparing them with the properties of the source model (predictive analogy of Bartha) (corresponds to another part of the demonstration of Hughes)
- Test of the predictions of the adapted and extended model by experiment (interpretation part of Hughes is needed for the test), perhaps observation of new phenomena -> next step
- Acceptance or refutation of the target model as new theory.

Four aspects concerning this theory were discussed after its presentation: first the meaning and importance of the limits of observability for the analogies during the development of physics, second the relevance of visualization of theoretical models in experimental results, third the effect of history on the direction of the development by analogical models in short and long timescales and at last the relation between analogies and metaphysics.

Regarding observability, I claim that no extension of theories into regions of non-observable objects and their behaviour is possible without analogical models. This statement means that the behaviour of observable objects in several development steps with ever more abstraction were transferred to the only partly or even only indirectly observable behaviour of unobservable objects. Thereby, it is assumed that the evidence of these objects and their similar behaviour to the observable and familiar objects is supported by principles derived from assumptions about the regularities in nature. In the evaluation of experimental results, these basic assumptions lead to the visualization of the measurement values in a way which the theoretical models suppose i.e., as tracks and collisions of postulated particles.

Because of the massive influence of analogies on the formation of modern physics, the direction of development in history depends on the kind of association and the resulting selection of the basic models in analogies. Thus, analogies affect the direction of development over short times, but this must not be the case in the long run because it is not foreseeable if and when the influence of specific analogies and their basic principles change.

Especially the mentioned basic principles or 'universal rules' but also unobservable objects are extensions of empirical founded concepts out over the border of experience i.e., they are in some sense metaphysical. This view has some consequences in the meaning of analogies in physical science as e.g., in case of 'dark matter' or 'dark energy' where a lack of a suitable analogy hinders the development of a satisfying explanation or in the case of general relativity and quantum theories where perhaps a more foundational analogy may overcome the incompatibility of these theories.

7 Conclusions

Starting point in this analysis of the role of analogies in physics was the question how it is possible that a limited access only by electromagnetic interaction (see chapter 1, p. 10) can produce such a complex theory as the standard model of particle physics with a heterogeneous set of invisible and partly even in fact unobservable particles including their different kinds of interactions. Of course, this was a longer process over about fifty years taking several steps with a lot of work on smaller theories and their confirmation by experiment till its acceptance in the end. Already after a few studies of the development process of these smaller theories, it became obvious that analogies played an essential role what motivated to investigate how analogies work in theory building and why they are indispensable for scientific research in physics. A closer look showed that analogies could be found everywhere during the history of physics and that the knowledge in physics developed from nearly completely observable phenomena in all its details to more and more abstract theories describing phenomena which were explained by the more or less observable phenomena associated with the behaviour of totally unobservable objects. Within this way to more and more unobservable objects, the formulation of the standard theory of particle physics is one of the newest results. Because of the starting question of this investigation, this development was analysed in more detail but restricted till the detection of quarks to limit the scope.

The outcome of this investigation is on the one hand a theory of theory development in physics, based on two types of analogies i.e., explanatory and predictive ones, which differ according to Bartha in their direction of association from positive analogies to the hypothesis or vice versa, respectively. The outcome on the other hand describes the conditions of observability which drive the road to more and more abstraction in the used concepts. Explanatory and predictive analogies are in the most cases intertwined and not distinct because they rest on the same basic source and target models, the difference is their function and by this also their direction of association. The usable concepts are limited by the historical development because the associated models applied in the analogies form the ideas, intended to adapt them on the target model for explanation of the observed phenomena. This limitation could be expressed also in the form: what is not in the concepts or their combination could not be thought. The other limitation is, as already mentioned, the observability but this limitation could be broken up by extension of the concepts. However, they could not be assumed to be realistic in the sense that they demonstrate the world as it is because these extensions of concepts are developed and learned by experience of phenomena and their abstraction into an unobservable region. The consequential conclusion is that development in physics is historical determined on the one side but not plannable in the future on the other.

In relation to the experimental side of the scientific enterprise in high energy physics could be noticed a predominance of visualization techniques. Beginning with the cloud chamber and ending with the modern detectors at the large colliders at CERN, the evaluation of the experimental results happens in form of track representation and their ramifications. I suppose that is due to the dominance of visual perception in the human sensorial system, which enables to realize the structural relationship of objects by their behaviour in the easiest way visually. May be that research on the perception of not visual dominated organisms may help to develop new concepts of representation or imagination.

Philosophically I think, the influence of the rejection of metaphysics at the turn to the twentieth century and the distinction between the context of discovery as task of the scientist and the context of justification as task of the philosopher in the following years was not positive alone. Despite the doubtless success to bring more clearness in the function of language in science and the recognition of the theory-ladenness of the scientific enterprise, it hindered the cognition process in respect to the gaining of knowledge. The reason is that science without metaphysics in the form of relative *a priori* principles is impossible and the distinction between the context of discovery and justification excluded philosophy over decades from essential questions. Only recently the context of discovery came into the focus of some philosophers again. But also, these philosophers, so my impression, do not assign the important role to analogies they deserve also when the most refer in some way to analogies as basic tool for problem solving. But the consideration of problem solving as focus jumps to short because the explanation of observed new or not explained old phenomena must not be always the problem. Sometimes perhaps phenomena were even not recognized as extraordinary but in contrast an analogy can direct the attention on them. So, in my opinion analogies have more a fundamental role than problem solving, especially, if Hofstadter and Sander are right that analogies and categories, respectively, are at the heart of all thinking. They can close the gap to all unobservable objects and their behaviour represented by the observable phenomena, with no difference if these objects really exist or are only imagined as model. In the end, it is enough if they could help to orient oneself in the world and curiosity never goes weary or does not drive us into a catastrophe.

8 References

Achinstein, P. (1971 [1968]). Concepts of Science - A Philosophical Analysis. Baltimore, London, The Johns Hopkins Press.

Agassi, J. (1988). Analogies Hard and Soft. Analogical Reasoning - Perspectives of Artificial Intelligence, Cognitive Science, and Philosophy. D. Helman. Dordrecht, Kluwer Academic Publishers**:** 401 - 419.

Bartha, P. (2010). By Parallel Reasoning - The Construction and Evaluation of Analogical Arguments. Oxford, New York, Oxford University Press.

Bartha, P. (2013). Analogy and Analogical Reasoning. Stanford Encyclopedia of Philosophy. E. N. Zalta. Stanford, Stanford University**:** 1 - 69. URL = http://plato.stanford.edu/archives/fall2013/entries/reasoning-analogy/

Baumgartner, A. (1829). Die Naturlehre nach ihrem gegenwärtigen Zustande mit Rücksicht auf mathematische Begründungen, dritte umgearbeitete und vermehrte Auflage. Wien, J.G. Heubner.

Berger, C. (2006). Elementarteilchenphysik - Von den Grundlagen zu den modernen Experimenten, 2. Auflage. Berlin, Heidelberg, New York, Springer Spektrum.

Bethge, K. and U. E. Schröder (1986). Elementarteilchen - und ihre Wechselwirkungen. Darmstadt, Wissenschaftliche Buchgesellschaft.

Brandt, S. (2009). The Harvest of a Century - Discoveries of Modern Physics in 100 Episodes. Oxford, Oxford University Press.

Carnap, R. (1979 [1928]). Der logische Aufbau der Welt. Frankfurt am Main, Berlin, Wien, Verlag Ullstein.

Carrier, M. (2009). Wege der Wissenschaftsphilosophie im 20. Jahrhundert. Wissenschaftstheorie - Ein Studienbuch. A. Bartels and M. Stöcker. Paderborn, mentis Verlag**:** 15 - 44 [361]

Cartwright, N. (1992). Aristotelian Natures and the Modern Experimental Method. Inference, Explanation, and Other Frustrations. J. Earman. Berkeley, University of California Press**:** 44 - 71.

Cellucci, C. (2018). Theory Building as Problem Solving. Building Theories - Heuristics and Hypotheses in Sciences. D. Danks and E. Ippoliti. Cham, Springer International Publishing**:** 63 - 79.

Christiansen, C. and Müller, J.J.C. (1903). Elemente der Theoretischen Physik - zweite, verbesserte Auflage. Leipzig, Johann Ambrosius Barth.

Curd, Martin; Cover, J.A., Ed. (1998): Philosophy of Science – The Central Issues. W.W. Norton & Company, New York, London.

Danks, D. (2018). Richer than Reduction. Building Theories - Heuristics and Hypotheses in Sciences. D. Danks and E. Ippoliti. Cham, Springer International Publishing**:** 45 -62.

Danks, D. and E. Ippoliti, Eds. (2018). Building Theories - Heuristics and Hypotheses in Science. Studies in Applied Philosophy, Epistemology and Rational Ethics. Cham, Springer International Publishing.

Davies, T. R. R., S. (1987). A Logical Approach to Reasoning by Analogy. IJCAI87: Proceedings of the Tenth International Joint Conference on Artificial Intelligence. J. McDermott. Los Altos, CA, Morgan Kaufmann**:** 264 - 270.

Duhem, P. (1998 [1908]). Das Ziel und die Struktur physikalischer Theorien - Autorisierte Übersetzung von Friedrich Adler - Mit einem Vorwort von Ernst Mach. Hamburg, Felix Meiner.

Elgin, C. Z. (2005). Non-foundationalist Epistemology: Holism, Coherence, and Tenability. Contemporary Debates in Epistemology. M. Steup and E. Sosa. Boston, Blackwell**:** 156 - 167.

Fermi, E. (1934). "Versuch einer Theorie der ß-Strahlen." Zeitschrift für Physik **88**: 161 - 177.

Feynman, R. P. and M. Gell-Mann (1958). "Theory of the Fermi Interaction." Physical Review **109**(1): 193 - 198.

Finkelnburg, W. (1967). Einführung in die Atomphysik. Berln, Heidelberg, New York, Springer-Verlag.

Friedman, M. (1983). Foundations of Space-Time Theories - Relativistic Physics and Philosophy of Science. Princeton, Princeton University Press.

Friedman, M. (2001). Dynamics of Reason. The 1999 Kant Lectures at Stanford University. Standord, Center for the Study of Language and Information (CLSI Publications)**:** XIV, 141.

Frigg, R. and S. Hartmann (2012). Models in Science. Stanford Encyclopedia of Philosophy. E. N. Zalta. Stanford, Stanford University**:** 1 - 46. URL = http://plato.stanford.edu/archives/fall2012/entries/models-science/

Frisch, M. (2005). Inconsistency, Asymmetry, and Non-Locality - A Philosophical Investigation of Classical Electrodynamics. Oxford, New York, Oxford University Press**:** XIII, 212.

Galison, P. (1997). Image and Logic - A Material Culture of Microphysics. Chicago, The University of Chicago Press.

Ganslandt, H. R. (1995). Mach. Enzyklopädie Philosophie und Wissenschaftstheorie. J. Mittelstraß. Stuttgart, Weimar, Verlag J.B. Metzler. **2:** 730 - 732.

Gell-Mann, M. (1953). "Isotopic Spin and New Unstable Particles." Physical Review **92**: 833 - 834.

Gell-Mann, M. (1956). "The Interpretation of the New Particles as Displaced Charge Multiplets." Nuovo Cimiento **IV**(2): 848 - 866.

Gell-Mann, M. (1957). "Model of Strong Couplings." Physical Review **106**(6): 1296 - 1300.

Gell-Mann, M. (1962). "Symmetries of Baryons and Mesons." Physical Review **125**(3): 1067 - 1084.

Gell-Mann, M. (1964). "A Schematic Model of Baryons and Mesons." Physics Letters **8**(3): 214 - 215.

Gell-Mann, M. (1994). Das Quark und der Jaguar. Vom Einfachen zum Komplexen - Die Suche nach einer neuen Erklärung der Welt. München, R. Piper Verlag.

Gell-Mann, M. and Y. Ne'eman, Eds. (1963). The Eightfold Way - A Review - With A Collection of Reprints. Frontiers in Physics - A Lecture Note & Reprint Series. New York, Amsterdam, W.A. Benjamin.

Gell-Mann, M. and A. Pais (1955). Theoretical Views on the New Particles. Proceedings of the Glasgow Conference on Nuclear and Meson Physics. E. H. M. Bellamy, R.G. Oxford, Pergamon Press**:** 342 – 352.

Gentner, D., et al. (2011). The Analogical Mind: Perspectives from Cognitive Science. Cambridge, MA, MIT Press.

Gethmann, C. F. (2004). Hempel. Enzyklopädie Philosophie und Wissenschaftstheorie. J. Mittelstraß. Stuttgart, Weimar, Verlag J.B. Metzler. **2:** 72 - 74.

Gooding, D. (1990). Experiment and the Making of Meaning. Dordrecht, Boston, London, Kluwer Academic Publishers.

Hacking, I. (1982). "Do We See Through a Microscope." Pacific Philosophical Quarterly **63**: 305 - 322.

Hacking, I. (1999). The Social Construction of What? Cambridge, MA; London, Harvard University Press.

Harrod, R. F. (1956). Foundations of Inductive Logic. London, Macmillan.

Heisenberg, W. (1925). "Quantenmechanische Umdeutung kinematischer und mechanischer Beziehungen." Zeitschrift für Physik **33**(1): 879 - 893.

Heisenberg, W. (1926). "Mehrkörperproblem und Resonanz in der Quantenmechanik." Zeitschrift für Physik **38**(6 - 7): 411 - 426.

Heisenberg, W. (1932a). "Über den Bau der Atomkerne. I." Zeitschrift für Physik **77**(1 - 2): 1 - 11.

Heisenberg, W. (1932b). "Über den Bau der Atomkerne, II." Zeitschrift für Physik **78**(3 - 4): 156 - 164.

Heisenberg, W. (1933). "Über den Bau der Atomkerne. III." Zeitschrift für Physik **80**(9 - 10): 587 - 596.

Heisenberg, W. (2007 [1958]). Physik und Philosophie. Stuttgart, S. Hirzel Verlag.

Heisenberg, W. (2010 [1969]). Der Teil und das Ganze - Gespräche im Umkreis der Atomphysik. München, Zürich, Piper Verlag.

Hempel, C. G. (1998 [1962]). Two Basic Types of Explanation. Philosophy of Science - The Central Issues. M. Curd and J. A. Cover. New York, London, W.W. Norton & Company**:** 685 - 694.

Hempel, C. G. (1998 [1966]). Criteria of Confirmation and Acceptability. Philosophy of Science - The Central Issues. M. Curd and J. A. Cover. New York, London, W.W. Norton & Company**:** 445 - 459.

Hempel, C. G. and P. Oppenheim (1948). "Studies in the Logic of Explanation." Philosophy of Science **15**(2): 135 - 176.

Hertz, H. (1956 [1894]). The Principles of Mechanics - with an Introduction of Robert Cohen and an preface of Hermann von Helmholtz. New York, Dover.

Hesse, M. B. (1966). Models and Analogies in Science. Notre Dame, University of Notre Dame Press.

Hofstadter, D. and S. Emmanuel (2014). Die Analogie - Das Herz des Denkens. Darmstadt, Wissenschaftliche Buchgesellschaft.

Hughes, R. I. G. (2012). The Theoretical Practices of Physics - Philosophical Essays. Oxford, Oxford University Press.

Huygens, C. (1890 [1678]). Abhandlung über das Licht. Worin die Vorgänge bei seiner Zurückwerfung und Brechung und besonder bei der eigenthümlichen Brechung des isländischen Spathes dargelegt sind. E. Lommel. Leipzig, Verlag von Wilhelm Engelmann**:** 114.

Hylleraas, E. A. (1931). "Über die Elektronenterme des Wasserstoffmoleküls." Zeitschrift für Physik **71**(11 - 12): 739 - 763.

Ippoliti, E. (2018). Building Theories: The Heuristic Way. Building Theories - Heuristics and Hypotheses in Sciences. D. Danks and E. Ippoliti. Cham, Springer International Publishing**:** 3 - 20.

Ippoliti, E. (2018). Heuristic Legic: A Kernel. Building Theories - Heuristics and Hypotheses in Sciences. D. Danks and E. Ippoliti. Cham, Springer International Publishing**:** 191 - 212.

Kapferer, W. (2018). Das Rätsel Dunkle Materie - Auf der Suche nach dem Unsichtbaren. Berlin, Springer-Verlag.

Keynes, J. M. (1921). A Treatise on Probability. London, Macmillan.

Krüger, J. G. (1763). Naturlehre. Halle, Carl Hermann Hemmerde.

Kuhn, T. S. (1970). Logic of Discovery or Psychology of Research? Criticism and the Growth of Knowledge - Proceedings of the International Colloquium in the Philosophy of Science, London, 1965, volume 4. I. Lakatos and A. Musgraver. London, Cambridge University Press. **4:** 1 - 23.

Kuhn, T. S. (1976). Die Struktur wissenschaftlicher Revolutionen. Frankfurt am Main, Suhrkamp Taschenbuch Verlag.

Lagrange, J.-L. d. (1797 [1788]). Analytische Mechanik. aus dem Französischen mit einigen Anmerkungen und erläuternden Zusätzen von Friedrich Wilhelm August Murhard. W. A. Murhard. Göttingen, Vandenhoeck & Ruprecht.

Lee, T. D. and C. N. Yang (1957). "Parity Nonconservation and a Two-Component Theory of the Neutrino." Physical Review **105**(5): 1671-1675.

Lee, T. D. and C. N. Yang (1956). "Question of Parity Conservation in Weak Interactions." Physical Review **104**(1): 5.

Liddle, A. L., Jon (2008). The Oxford Companion to Cosmology. Oxford, New York, Oxford University Press.

Lohrmann, E. (1981). Hochenergiephysik. Stuttgart, B.G. Teubner.

Lorenz, K. (2004). Wiener Kreis. Enzyklopädie Philosophie und Wissenschaftstheorie. J. Mittelstraß. Stuttgart, Weimar, Verlag J.B. Metzler. **4:** 695 - 698.

Lorenz, K. (2004). Kuhn. Enzyklopädie Philosophie und Wissenschaftstheorie. J. Mittelstraß. Stuttgart, Verlag J.B. Metzler. **2:** 504 - 507.

Mach, E. (2014 [1896]). Die Prinzipien der Wärmelehre. Hamburg, Severus Verlag.

Mach, E. (2015 [1906]). Erkenntnis und Irrtum - Skizzen zur Psychologie der Forschung. Villingen-Schwenningen, Nexx Verlag.

Mach, E. (2015 [1922]). Die Analyse der Empfindungen - und das Verhältnis des Physischen zum Psychischen. Norderstedt, Verlag der Wissenschaften.

Mandl, F. and G. Shaw (1993). Quantenfeldtheorie. Wiesbaden, Aula-Verlag.

Mayer-Kuckuk, T. (1979). Kernphysik - Eine Einführung. Stuttgart, B.G. Teubner

McMullin, E. (1978). Newton on Matter and Activity. Notre Dame, University of Notre Dame Press.

Milkov, N. (2011). Hans Reichenbachs wissenschaftliche Philosophie. Ziele und Wege der heutigen Naturphilosophie. N. Milkov. Hamburg, Felix Meiner Verlag**:** VII-XLIV.

Milkov, N. (2015). Die Berliner Gruppe des Logischen Empirismus. Die Berliner Gruppe - Texte zum Logischen Empirismus von Walter Dubislav, Kurt Grelling, Carl G. Hempel, Alexander Herzberg, Kurt Lewin, Paul Oppenheim und Hans Reichenbach. N. Milkov. Hamburg, Felis Meiner Verlag**:** IX - LXI.

Misner, C. W. et al. (1973): Gravitation. W.H. Freeman and Company, New York.

Mittelstraß, J., Ed. (2004). Enzyklopädie Philosophie und Wissenschaftstheorie Band 1. Enzyklopädie Philosophie und Wissenschaftstheorie. Stuttgart, Verlag J.B. Metzler.

Mittelstraß, J., Ed. (2004). Enzyklopädie Philosophie und Wissenschaftstheorie Band 2. Enzyklopädie Philosophie und Wissenschaftstheorie. Stuttgart, Verlag J.B. Metzler.

Mittelstraß, J., Ed. (2004). Enzyklopädie Philosophie und Wissenschaftstheorie Band 3. Enzyklopädie Philosophie und Wissenschaftstheorie. Stuttgart, Verlag J.B. Metzler.

Mittelstraß, J., Ed. (2004). Enzyklopädie Philosophie und Wissenschaftstheorie Band 4. Enzyklopädie Philosophie und Wissenschaftstheorie. Stuttgart, Verlag J.B. Metzler.

Morrison, M. (2018). Building Theories: Strategies not Blueprints. Building Theories - Heuristics and Hypotheses in the Sciences. D. Danks and E. Ippoliti. Cham, Springer International Publishing**:** 21 - 44.

Musgrave, A. (1985). Realism versus Contructive Empiricism. Images of Science. P. M. Churchland and C. A. Hooker. Chicago, University of Chicago Press**:** 197-221.

Nagel, E. (1968 [1961]). The Structure of Science - Problems in the Logic of Scientific Explanation. London, Routledge & Kegan Paul.

Nersessian, N. J. (1984). Faraday to Einstein: Conctructing Meaning in Scientific Theories. Dordrecht, Boston, London, Kluwer Academic Publishers.

Neurath, O. (2009 [1932]). Protokollsätze. Wiener Kreis - Texte zur wissenschaftlichen Weltauffassung von Rudolf Carnap, Otto Neurath, Moritz Schlick, Philipp Frank, Hans Hahn, Karl Menger, Edgar Zilsel und Gustav Bergmann. M. Stöltzner and T. Uebel. Hamburg, Felix Meiner Verlag**:** 399 - 411.

Newton, I. (1999 [1687]). Die mathematischen Prinzipien der Physik. Berlin, New York, Walter de Gruyter.

Pais, Abraham (1952): Some Remarks on the V-Particles, in: Physical Review **86**(5): 663 – 672.

Pais, Abraham (1986): Inward Bound – Of Matter and Forces in the Physical World. Oxford University Press; Oxford.

Pauli, W. (1961). Relativitätstheorie und Wissenschaft. Aufsätze und Vorträge über Physik und Erkenntnistheorie. W. Pauli. Braunschweig, Friedrich Vieweg & Sohn**:** 76 - 80.

Pauli, W. (1961a). Albert Einstein in der Entwicklung der Physik. Aufsätze und Vorträge über Physik und Erkenntnistheorie. W. Pauli. Braunschweig, Friedr. Vieweg & Sohn**:** 85 - 90.

Pickering, A. (1984). Constructing Quarks - A Sociological History of Particle Physics. Chicago, The University of Chicago Press.

Pontecorvo, B. (1947). "Nuclear Capture of Mesons and the Meson Decay." Physical Review **72**(3): 246-247.

Popper, K. R. (1976 [1935]). Logik der Forschung. Tübingen, J.C.B. Mohr (Paul Siebeck).

Popper, K. R. (1994). Alles Leben ist Problemlösen - Über Erkenntnis, Geschichte und Politik. Darmstadt, Wissenschaftliche Buchgesellschaft.

Quine, W. V. (1998 [1951]). Two Dogmas of Empiricism. Philosophy of Science - The Central Issues. M. Curd and J. A. Cover. New York, London, W.W. Norton & Company**:** 280 - 301.

Raith, W., Ed. (2003). Bestandteile der Materie, 2., überarbeitete Auflage. Bergmann/Schäfer - Lehrbuch der Experimentalphysik. Berlin; New York, Walter de Gruyter.

Reichenbach, H. (2011). Ziele und Wege der heutigen Naturphilosophie. Fünf Aufsätze zur Wissenschaftstheorie, Felix Meiner Verlag.

Reichenbach, H. ([1938]). Experience and Prediction - An Analysis of the Foundations and the Structure of Knowledge. Chicago, University of Chicago Press**:** 428.

Russell, S. (1988). Analogy by Similarity. Analogical Reasoning - Perspectives of Artificial Intelligence, Cognitive Science, and Philosophy. D. Helman. Dordrecht, Kluwer Academic Publishers**:** 251 - 269.

Schäfer, L. (1998). Einleitung: Duhems Bedeutung für die Entwicklung der Wissenschaftstheorie. Duhem, Pierre (1908): Ziel und Struktur der physikalischen Theorien, Nachdruck der autorisierten Übersetzung von Friedrich Adler, mit einem Vorwort von Ernst Mach. L. Schäfer. Hamburg, Felix Meiner Verlag**:** IX* - XLVI*.

Schrödinger, E. (1926). "Quantisierung als Eigenwertproblem (Erste Mitteilung)." Annalen der Physik **79**(4): 16.

Schrödinger, E. (1963). Die Wellenmechanik. Stuttgart, Ernst Battenberg Verlag.

Schurz, G. (2011). Einführung in die Wissenschaftstheorie. Darmstadt, Wissenschaftliche Buchgesellschaft.

Sommerfeld, A. (1977). Elektrodynamik. Thun, Harri Deutsch.

Sommerfeld, A. (1978). Mechanik der deformierbaren Medien. Thun, Farnkfurt am Main, Harri Deutsch.

Steinberger, J. (1949). "On the Range of Elcetrons in Meson Decay." Physical Review **75**(8): 9.

Tiomno, J. and J. A. Wheeler (1949). "Energy Spectrum of Electrons from Meson Decay." Review of Modern Physics **21**(1): 144 - 158.

Traweek, S. (1988). Beamtimes and Lifetimes - The World of High Energy Physics. Cambridge (Mass.), London, Harvard University Press.

Uglaglia, M. (2018). “Take the Case of a Geometer…” - Mathematical Analogies and Building Theories in Aristotle. Building Theories - Heuristic and Hypotheses in Sciences. D. Danks and E. Ippoliti. Cham, Springer International Publishing**:** 119 - 143.

van Fraassen, B. C. (1980). The Scientific Image. Oxford, Clarendon Press.

van Fraassen, B. (1989). Laws and Symmetry. Oxford, Clarendon Press.

van Fraassen, B. (2008). Scientific Representation. Oxford, New York, Oxford University Press.

Weitzenfeld, J. S. (1984). "Valid Reasoning by Analogy." Philosophy of Science **51**(1): 137 - 149.

Wittgenstein, L. (1984). Tractatus logico-philosophicus. Tractatus logico-philosophicus, Tagebücher 1914-1916, Philosophische Untersuchungen. Frankfurt am Main, Suhrkamp Taschenbuch Verlag. **1:** 7-85.

Wolters, G. (1995). Konventionalismus. Enzyklopädie Philosophie und Wissenschaftstheorie. J. Mittelstraß. Stuttgart, Weimar, Verlag J.B. Metzler. **2:** 462-465.

Wu, C. S., et al. (1957). "Experimental Test of Parity Conservation in Beta Decay." Physical Review **105**(4): 1413-

1415.