Chapter 2
The Conjoint Origin of Proof and Theoretical Physics

Hans Niels Jahnke

2.1 The Origins of Proof

Historians of science and mathematics have proposed three different answers to the question of why the Greeks invented proof and the axiomatic-deductive organization of mathematics (see Szabó 1960, 356 ff.).

(1). The *socio-political thesis* claims a connection between the origin of mathematical proof and the freedom of speech provided by Greek democracy, a political and social system in which different parties fought for their interests by way of argument. According to this thesis, everyday political argumentation constituted a model for mathematical proof.
(2). The *internalist thesis* holds that mathematical proof emerged from the necessity to identify and eliminate incorrect statements from the corpus of accepted mathematics with which the Greeks were confronted when studying Babylonian and Egyptian mathematics.
(3). The *thesis of an influence of philosophy* says that the origin of proof in mathematics goes back to requirements made by philosophers.

Obviously, thesis (1) can claim some plausibility, though there is no direct evidence in its favor and it is hard to imagine what such evidence might look like.

Thesis (2) is stated by van der Waerden. He pointed out that the Greeks had learnt different formulae for the area of a circle from Egypt and Babylonia. The contradictory results might have provided a strong motivation for a critical reexamination of the mathematical rules in use at the time the Greeks entered the scene. Hence, at the time of Thales the Greeks started to investigate such problems by themselves in order to arrive at correct results (van der Waerden 1988, 89 ff.).

H.N. Jahnke (✉)
FB Mathematik, Universität Duisburg-Essen, 45117, Campus Essen, Essen, Germany
e-mail: njahnke@uni-due.de

[1]I would like to thank Gila Hanna and Helmut Pulte for their valuable advice.

G. Hanna et al. (eds.), *Explanation and Proof in Mathematics: Philosophical and Educational Perpectives,* DOI 10.1007/978-1-4419-0576-5_2,

Thesis (3) is supported by the fact that standards of mathematical reasoning were broadly discussed by Greek philosophers, as the works of Plato and Aristotle show. Some authors even use the term "Platonic reform of mathematics."

This paper considers in detail a fourth thesis which in a certain sense constitutes a combination of theses (1) and (3). It is based on a study by the historian of mathematics Árpád Szabó[2] (1960), who investigated the etymology of the terms used by Euclid to designate the different types of statements functioning as starting points of argumentation in the "Elements."

Euclid divided the foundations of the "Elements" into three groups of statements: (1) Definitions, (2) Postulates and (3) Common Notions (Heath 1956). Definitions determine the objects with which the Elements are going to deal, whereas Postulates and Common Notions entail statements about these objects from which further statements can be derived. The distinction between postulates and common notions reflects the idea that the postulates are statements specific to geometry whereas the common notions provide propositions true for all of mathematics. Some historians emphasize that the postulates can be considered as statements of existence.

In the Greek text of Euclid handed down to us (Heiberg's edition of 1883–1888) the definitions are called ὅϱοι, the postulates ἀιτήματα and the common notions κοιναί ἔννοιαι. In his analysis, Szabó starts with the observation that Proclus (fifth century AD), in his famous commentary on Euclid's elements, used a different terminology (for an English translation, see Proclus 1970). Instead of ὅϱοσ (definition) Proclus applied the concept of ὑπόθεσις (hypothesis) and instead of κοιναί ἔννοιαι (common notions) he used ἀξιώμα (axiomata). He maintained the concept of ἀιτήματα (postulates) as contained in Euclid. Szabó explains the differing terminology by the hypothesis that Proclus referred to older manuscripts of Euclid than the one which has led to our modern edition of Euclid.

Szabó shows that ὑπόθεσις (hypothesis), ἀιτήματα (postulate) and άξιώμα (axiom) were common terms of pre-Euclidean and pre-Platonic dialectics, which is related both to philosophy and rhetoric. The classical Greek philosophers understood dialectics as the art of exchanging arguments and counter-arguments in a dialog debating a controversial proposition. The outcome of such an exercise might be not simply the refutation of one of the relevant points of view but rather a synthesis or a combination of the opposing assertions, or at least a qualitative transformation (see Ayer and O'Grady 1992, 484).

The use of the concept of hypothesis as synonymous with definition was common in pre-Euclidean and pre-Platonic dialectics. In this usage, hypothesis designated the fact that the participants in a dialog had to agree initially on a joint definition of the topic before they could enter the argumentative discourse about it. The Greeks, including Proclus, also used hypothesis in a more general sense, close to its meaning today. A hypothesis is that which is underlying and consequently can be used as a foundation of something else. Proclus, for example, said: "Since this

[2] For a discussion of the personal and scientific relations between Szabó and Lakatos, see Maté (2006). I would like to thank Brendan Larvor for drawing my attention to this paper.

science of geometry is based, we say, on hypothesis (ἔξ ὑποθέσεως ἐιναι), and proves its later propositions from determinate first principles … he who prepares an introduction to geometry should present separately the principles of the science and the conclusions that follow from the principles, …” (Proclus 1970, 62).

According to Szabó, the three concepts of hypothesis, aitema (postulate) and axioma had a similar meaning in the pre-Platonic and pre-Aristotelian dialectics. They all designated those initial propositions on which the participants in a dialectic debate must agree. An initial proposition which was agreed upon was then called a “hypothesis”. However, if participants did not agree or if one declared no decision, the proposition was then called aitema (postulate) or axioma (Szabó 1960, 399).

As a rule, participants will introduce into a dialectic debate hypotheses that they consider especially strong and expect to be accepted by the other participants: numerous examples of this type can be found in the Platonic dialogs. However, it is also possible to propose a hypothesis with the intention of critically examining it. In a philosophical discourse, one could derive consequences from such a hypothesis that are *desired* (plausible) or *not desired* (implausible). The former case leads to a strengthening of the hypothesis, the latter to its weakening. The extreme case of an undesired consequence would be a logical contradiction, which would necessarily lead to the rejection of the hypothesis. Therefore, the procedure of indirect proof in mathematics can be considered as directly related to common customs in philosophy. According to Szabó (1960) this constitutes an explanation for the frequent occurrence of indirect proofs in the mathematics of the early Greek period.

The concept of common notions as a name for the third group of introductory statements needs special attention. As mentioned above, this term is a direct translation of the Greek κοιναί ἔννοιαι and designates “the ideas common to all human beings”. According to Szabó, the term stems from Stoic philosophy (since 300 BC) and connotes a proposition that cannot be doubted justifiably. Proclus also attributes the same meaning to the concept of ἀξιώμα, which he used instead of κοιναί ἔννοιαι. For example, he wrote at one point: ”These are what are generally called indemonstrable axioms, inasmuch as they are deemed by everybody to be true and no one disputes them” (Proclus 1970, 152). At another point he even wrote, with an allusion to Aristotle: ”…whereas the axiom is as such indemonstrable and everyone would be disposed to accept it, even though some might dispute it for the sake of argument”(Proclus 1970, 143). Thus, only quarrelsome people would doubt the validity of the Euclidean axioms; since Aristotle, this has been the dominant view.

Szabó (1960) shows that the pre-Aristotelean use of the term axioma was quite similar to that of the term aitema, so that axioma meant a statement upon which the participants of a debate agreed or whose acceptance they left undecided. Furthermore, he makes it clear that the propositions designated in Euclid’s “Elements” as axioms or common notions had been doubted in the early period of Greek philosophy, namely by Zenon and the Eleatic School (fifth century BC). The explicit compilation of the statements headed by the term axioms (or common notions) in the early period of constructing the elements of mathematics was motivated by the intention of rejecting Zeno’s criticism. Only later, when the philosophy of the Eleates had been weakened, did the respective statements appear as unquestionable for a healthy mind.

In this way, the concept of an axiom gained currency in Greek philosophy and in mathematics. Its starting point lay in the art of philosophical discourse; later it played a role in both philosophy and mathematics. More important for this paper, it underwent a concomitant change in its epistemological status. In the early context of dialectics, the term axiom designated a proposition that in the beginning of a debate could be accepted or not. However, axiom's later meaning in mathematics was clearly that of a statement which itself cannot be proved but is absolutely certain and therefore can serve as a fundament of a deductively organized theory. This later meaning became the still-dominant view in Western science and philosophy.

Aristotle expounded the newer meaning of axiom at length in his "Analytica posteriora":

> I call "first principles" in each genus those facts which cannot be proved. Thus the meaning both of the primary truths and the attributes demonstrated from them is assumed; as for their existence, that of the principles must be assumed, but that of the attributes must be proved. E. g., we assume the meaning of "unit", "straight" and "triangular"; but while we assume the existence of the unit and geometrical magnitude, that of the rest must be proved. (Aristotle 1966, I, 10)

Aristotle also knew the distinction between postulate (aitema) and axiom (common notions) as used in Euclid:

> Of the first principles used in the demonstrative sciences some are special to particular sciences, and some are common; . . . Special principles are such as that a line , or straightness, is of such-and-such a nature; common principles are such as that when equals are taken from equals the remainders are equals. (Aristotle 1966, I, 10)

Thus, Szabó's study leads to the following overall picture of the emergence of mathematical proof. In early Greek philosophy, reaching back to the times of the Eleates (ca. 540 to 450 BC), the terms axioma and aitema designated propositions which were accepted in the beginning of a dialog as a basis of argumentation. In the course of the dialog, consequences were drawn from these propositions in order to examine them critically and to investigate whether the consequences were desired. In a case where the proposition referred to physical reality, "desired" could mean that the consequences agreed with experience. If the proposition referred to ethics, "desired" could mean that the consequences agreed with accepted norms of behavior. Desired consequences constituted a strong argument in favor of a proposition. The most extreme case of undesired consequence, a logical contradiction, led necessarily to rejecting the proposition. Most important, in the beginning of a dialog the epistemic status of an axioma or aitema was left indefinite. An axiom could be true or probable or perhaps even wrong.

In a second period, starting with Plato and Aristotle (since ca. 400 BC) the terms axioma and aitema changed their meaning dramatically; they now designated propositions considered absolutely true. Hence, the epistemic status of an axiom was no longer indefinite but definitely fixed. This change in epistemic status followed quite natural because at that time mathematicians had started building theories. Axioms were supposed true once and for all, and mathematicians were interested in deriving as many consequences from them as possible. Thus, the emergence of the classical view that the axioms of mathematics are absolutely true was inseparably linked to the fact that mathematics became a “normal science” to use T. Kuhn’s term. After Plato and Aristotle, the classical view remained dominant until well into the nineteenth century.

Natural as it might have been, in the eyes of modern philosophy and modern mathematics this change of the epistemic status of axioms was nevertheless an unjustified dogmatisation. The decision to build on a fixed set of axioms and not to change them any further is epistemologically quite different from the decision to declare them absolutely true.

On a more general level, we can draw two consequences: First, Szabó’s (1960) considerations suggest the thesis that the *practice of a rational discourse* provided a model for the organization of a mathematical theory according to the axiomatic-deductive method; in sum, proof is rooted in communication. However, this does not simply support the socio-political thesis, according to which proof was an outcome of Greek democracy. Rather, it shows a connection between proof and dialectics as an *art of leading a dialog*. This art aimed at a methodically ruled discourse in which the participants accept and obey certain rules of behavior. These rules are crystallized in the terms hypothesis, aitema and axiom, which entail the participants’ obligation to exhibit their assumptions.

The second important consequence refers to the *universality* of dialectics. Any problem can become the subject of a dialectical discourse, regardless of which discipline or even aspect of life it involves. From a problem of ethics to the question of whether the side and diagonal of a square have a common measure, all problems could be treated in a debate. Different persons can talk about the respective topic as long as they are ready to reveal their suppositions. Analogously, the possibility of an axiomatic-deductive organization of a group of propositions is not confined to arithmetic and geometry, but can in principle be applied to any field of human knowledge. The Greeks realized this principle at the time of Euclid, and it led to the birth of theoretical physics.

2.2 Saving the Phenomena

During the Hellenistic era, within a short interval of time Greek scientists applied the axiomatic-deductive organization of a theory to a number of areas in natural science. Euclid himself wrote a deductively organized optics, whereas Archimedes provided axiomatic-deductive accounts of statics and hydrostatics.

In astronomy, too, it became common procedure to state hypotheses from which a group of phenomena could be derived and which provided a basis for calculating astronomical data. Propositions of quite a different nature could function as hypotheses. For example, Aristarchos of Samos (third century BC) began his paper “On the magnitudes and distances of the sun and the moon” with a hypothesis about how light rays travel in the system earth-sun-moon, a hypothesis about possible positions of the moon in regard to the earth, a hypothesis giving an explanation of the phases of the moon and a hypothesis about the angular distance of moon and sun at the time of half-moon (a measured value). These were the ingredients Aristarchos used for his deductions.

In the domain of astronomy, the Greeks discussed, in an exemplary manner, philosophical questions about the relation of theory and empirical evidence. This discussion started at the time of Plato and concerned the paths of the planets. In general, the planets apparently travel across the sky of fixed stars in circular arcs. At certain times, however, they perform a retrograde (and thus irregular) motion. This caused a severe problem; since the Pythagoreans, the Greeks had held a deeply rooted conviction that the heavenly bodies perform circular movements with constant velocity. But this could not account for the irregular retrograde movement of the planets.

Greek astronomers invented sophisticated hypotheses to solve this problem. The first scientist who proposed a solution was Eudoxos, the best mathematician of his time and a close friend of Plato's. Though the phenomenon of the retrograde movement of the planets was well known, it did not figure in the dialogs of Plato's early and middle period. Only in his late dialog "Nomoi" ("Laws") did Plato mention the problem. In this dialog, a stranger from Athens (presumably Eudoxos) appeared, who explained to Clinias (presumably Plato) that it only seems that the planets "wander" (i.e., perform an irregular movement), whereas in reality precisely the opposite is true: "Actually, each of them describes just one fixed orbit, although it is true that to all appearances its path is always changing" (Plato 1997, 1488). Thus, in his late period Plato acknowledged that we have to adjust our basic ideas in order to make them agree with empirical observations.

I will illustrate this principle by a case simpler than the paths of the planets but equally important in Greek astronomy. In the second century BC, the great astronomer and mathematician Hipparchos investigated an astronomical phenomenon probably already known before his time, the "anomaly of the sun." Roughly speaking, the term referred to the observation that the half-year of summer is about 1 week longer than the half-year of winter. Astronomically, the half-year of summer was then defined as the period that the sun on its yearly path around the earth (in terms of the geocentric system) needs to travel from the vernal equinox to the autumnal equinox. Analogously, the half-year of winter is the duration of the travel from the autumnal equinox to the vernal equinox. Vernal equinox and autumnal equinox are the two positions of the sun on the ecliptic at which day and night are equally long for beings living on the earth. The two points, observed from the earth, are exactly opposite to each other (vernal equinox, autumnal equinox and center of the earth form a straight line). Since the Greek astronomers supposed that all heavenly bodies move with constant velocity in circles around the center of the earth; it necessarily followed that the half-years of summer and winter would be equal.

The Greek astronomers needed to develop a hypothesis to explain this phenomenon. Hipparchos proposed a hypothesis placing the center of the sun's circular orbit not in the center of the earth but a bit outside it (Fig. 2.1). If this new center is properly placed, then the arc through which the sun travels during summer,

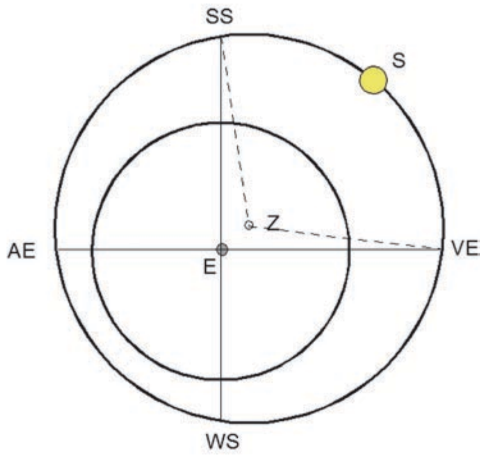

Fig. 2.1 Eccentric hypothesis

observed from the earth, is greater than a half-circle; the anomaly of the sun is explained. Later, Hipparchos' hypothesis was called by Ptolemaios the "eccentric hypothesis" (Toomer 1984, 144 pp).

Another hypothesis competing with that of Hipparchos was the "epicyclic hypothesis" of Appolonios of Perge (third century BC; see Fig. 2.2). It said that the sun moves on a circle concentric to the center of the universe, however "not actually on that circle but on another circle, which is carried by the first circle, and hence is known as the epicycle" (Toomer 1984, 141). Hence, the case of the anomaly of the sun confronts us with the remarkable phenomenon of a *competition of hypotheses*. Both hypotheses allow the derivation of consequences which agree with the astronomical phenomena. Since there was no further reason in favor of either one, it didn't matter which one was applied. Ptolemaios showed that, given an adequate choice of parameters, both hypotheses are mathematically equivalent and lead to the same data for the orbit of the sun. Of course, physically they are quite different; nevertheless, Ptolemaios did not take the side of one or the other.

Hence the following situation: The Greeks believed that the heavenly bodies moved with constant velocity on circles around the earth. These two assumptions (constancy of velocity and circularity of path) were so fundamental that the Greeks were by no means ready to give them up. The retrograde movement of the planets and the anomaly of the sun seemed to contradict these convictions. Consequently, Greek astronomers had to invent additional hypotheses which brought the theory into accordance with the phenomena observed. The Greeks called the task of inventing such hypotheses "saving the phenomena" ("σώζειν τὰ φαινόμενα").

The history of this phrase is interesting and reflects Greek ideas about how to bring theoretical thinking in agreement with observed phenomena (see Lloyd 1991; Mittelstrass 1962). In written sources the term "saving the phenomena" first appears in the writings of Simplikios, a Neo-Platonist commentator of the sixth

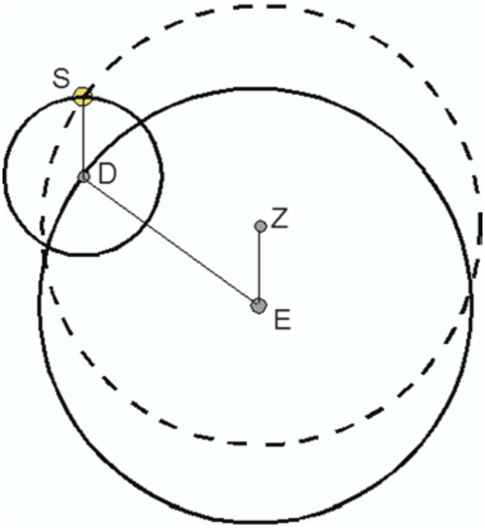

Fig. 2.2 Epicyclic hypothesis

century AD, a rather late source. However, the phrase probably goes back to the time of Plato. Simplikios wrote that Plato made "the saving of phenomena" a task for the astronomers. But we have seen that Plato hit upon this problem only late in his life; it is much more probable that he learnt about it from the astronomers (e.g., Eudoxos) than vice versa. It seems likely that the phrase had been a terminus technicus among astronomers since the fourth century BC.

A number of philosophers of science, the most prominent being Pierre Duhem (1908/1994), have defended the thesis that the Greeks held a purely conventionalist view and did not attribute any claim of truth to astronomical hypotheses. They counted different hypotheses, like the excentric or epicyclic hypotheses as equally acceptable if the consequences derived from them agreed with the observed phenomena. However, the Greeks in fact never questioned certain astronomical assumptions, namely the circularity of the paths and the constancy of velocities of the heavenly bodies, attributing to them (absolute) truth.

Mittelstrass (1962), giving a detailed analysis of its history, shows that the phrase "saving the phenomena" was a terminus technicus in ancient astronomy and stresses that it was used only by astronomers and in the context of astronomy (Mittelstrass 1962, 140 ff). He questions Simplikios' statement that Plato posed the problem of "saving the phenomena" and contradicts modern philosophers, such as Natorp (1921, p. 161, 382, 383), who have claimed that the idea of "saving the phenomena" was essential to the ancient Greek philosophy of science. According to Mittelstrass, only Galileo first transferred this principle to other disciplines and made it the basis of a general scientific methodology, which by the end of the nineteenth century was named the "hypothetico-deductive method."

Mittelstrass is surely right in denying that "saving the phenomena" was a general principle of Greek scientific thinking. He can also prove that the phrase was used explicitly only in astronomy. Greek scientific and philosophic thinking was a mixture of different ideas and approaches; there was no unified "scientific method." Nevertheless, Mittelstrass goes too far in strictly limiting to astronomy the idea that a hypothesis is evaluated through the adequacy of its consequences. As Szabó (1960) has shown, such trial was common practice in Greek dialectics and was reflected in early meanings of the terms aitema, axioma and hypothesis, meanings that the terms kept until the times of Plato and Aristotle. The procedure of supposing a hypothesis as given and investigating whether its consequences are desired abounds in Plato's dialogs. Thus, the idea underlying the phrase "saving the phenomena" had a broader presence in Greek scientific and philosophical thinking than Mittelstrass supposes. Besides, Mittelstrass did not take into account Szabó's (1960) study, though it had already been published.

Hence, I formulate the following thesis: The extension of the axiomatic procedure from geometry to physics and other disciplines cannot be imagined without the idea that an axiom is a hypothesis which may be justified not by direct intuition but by the adequacy of its consequences, in line with the original dialectical meaning of the terms aitema, axioma and hypothesis.

The Greeks set up a range of possible hypotheses in geometry and physics with a variety of epistemological justifications. For example, Euclid's geometrical postulates were considered from antiquity up to the nineteenth century as evident in themselves and absolutely true. Only the parallel postulate couldn't claim a similar epistemological status of direct evidence; this was already seen in antiquity. A possible response to this lack would have been to give up the epistemological claim that the axioms of geometry are evident in themselves, but the Greeks didn't do that. Another way out would have been to deny the parallel postulate the status of an axiom. The Neo-Platonist commentator Proclus did exactly that, declaring the parallel postulate a theorem whose proof had not yet been found. However, this tactic was motivated by philosophical not mathematical reasons, though the problem was a mathematical one. Besides, Proclus lived 700 years after Euclid; we do not even know how Euclid himself thought about his parallel postulate. Perhaps Euclid as a mathematician was more down-to-earth and was less concerned about his postulate's non-evidency.

A second example concerns statics. In the beginning of "On the equilibrium of planes or the centers of gravity of planes," Archimedes set up seven postulates; the first reads as:

> I postulate the following: 1. Equal weights at equal distances are in equilibrium, and equal weights at unequal distances are not in equilibrium but incline towards the weight which is at the greater distance." (Heath, 1953, 189)

This postulate shows a high degree of simplicity and evidence and, in this regard, is like Euclid's geometric postulates. The other postulates on which Archimedes based his statics are similar; they appear as unquestionable. Statics therefore seemed to have the same epistemological status as geometry and from early times up to the nineteenth century was considered a part of mathematics. However, during the nineteenth century statics became definitively classified as a subdiscipline of physics. Meanwhile, the view that the natural sciences are founded without exception on experiment became dominant. Hence, arose a problem: Statics had the appearance of a science that made statements about empirical reality, but was founded on propositions apparently true without empirical evidence. Only at the end of the nineteenth century did E. Mach (1976) expose, in an astute philosophical analysis, the (hidden) empirical assumptions in Archimedes' statics, thus clarifying that statics is an empirical, experimental science like any other.

As a final example, consider the (only) hypothesis which Archimedes stated at the beginning of his hydrostatics ("On floating bodies"):

> Let it be supposed that a fluid is of such a character that, its parts lying evenly and being continuous, that part which is thrust the less is driven along by that which is thrust the more; and that each of its parts is thrust by the fluid which is above it in a perpendicular direction if the fluid be sunk in anything and compressed by anything else. (Heath, 1953, 253)

Archimedes derived from this hypothesis his famous "law of upthrust" ("principle of buoyancy") and developed a mathematically sophisticated theory about the balance of swimming bodies. Obviously, the hypothesis does not appear simple or beyond doubt. To a historically open-minded reader, it looks like a typical assumption set up in a modern situation of developing a mathematical model for some specific aim – in other words, a typical hypothesis whose truth cannot be directly judged. It is accepted as true insofar as the consequences that can be derived from it are desirable and supported by empirical evidence. No known source has discussed the epistemological status of the axiom and its justification in this way. Rather, Archimedes' hydrostatics was considered as a theory that as a whole made sense and agreed with (technical) experience.

Considered in their entirety, the axiomatic-deductive theories that the Greeks set up during the third century BC clearly rest on hypotheses that vary greatly in regard to the justification of their respective claims of being true. Some of these hypotheses seem so intuitively safe that a "healthy mind" cannot doubt them; others have been accepted as true because the theory founded on them made sense and agreed with experience.

In sum, ancient Greek thinking had two ways of justifying a hypothesis. First, an axiom or a hypothesis might be accepted as true because it agrees with intuition. Second, hypotheses inaccessible to direct intuition and untestable by direct inspection were justified by drawing consequences from them and comparing these with the data to see whether the consequences were desired; that is, they agreed with experience or with other statements taken for granted. Desired consequences led to strengthening the hypothesis, undesired consequences to its weakening. Mittelstrass (1962) wants to limit this second procedure to the narrow context of ancient astronomy. I follow Szabó (1960) in seeing it also inherent in the broader philosophical and scientific discourse of the Pre-Platonic and Pre-Aristotelean period.

2.3 Intrinsic and Extrinsic Justification in Mathematics

From the times of Plato and Aristotle to the nineteenth century, mathematics was considered as a body of absolute truths resting on intuitively safe foundations. Following Lakatos (1978), we may call this the *Euclidean* view of mathematics (in M. Leng's chapter, this volume, this is called the "assertory approach"). In contrast, modern mathematics and its philosophy would consider the axioms of mathematics simply as statements on which mathematicians agree; the epistemological qualification of the axioms as true or safe is ignored. At the end of the nineteenth century, C. S. Peirce nicely expressed this view: "… all modern mathematicians agree … that mathematics deals exclusively with hypothetical states of things, and asserts no matter of fact whatever." (Peirce 1935, 191). We call this modern view the *Hypothetical* view.

Mathematical proof underwent a foundational crisis at the beginning of the twentieth century. In 1907, Bertrand Russell stated that the fundamental axioms of mathematics can only be justified not by an absolute intuition but by the insight that one can derive the desired consequences from them (Russell 1924; see Mancosu 2001, 104). In discussing his own realistic (or in his words "Platonistic") view of the nature of mathematical objects, Gödel (1944) supported this view:

> The analogy between mathematics and a natural science is enlarged upon by Russell also in another respect … He compares the axioms of logic and mathematics with the laws of nature and logical evidence with sense perception, so that the axioms need not necessarily be evident in themselves, but rather their justification lies (exactly as in physics) in the fact that they make it possible for these "sense perceptions" to be deduced: … I think that … this view has been largely justified by subsequent developments. (Gödel 1944, 210)

On the basis of Gödel's "Platonistic" (realistic) philosophy, the American philosopher Penelope Maddy has designated justification of an axiom by direct intuition as "intrinsic", and justification by reference to plausible or desired consequences as "extrinsic" (Maddy 1980). According to Maddy, Gödel posits a faculty of mathematical intuition that plays a role in mathematics analogous to that of sense perception in the physical sciences:

> … presumably the axioms [of set theory: Au] force themselves upon us as explanations of the intuitive data much as the assumption of medium-sized physical objects forces itself upon us as an explanation of our sensory experiences. (Maddy, 1980, 31)

For Gödel the assumption of sets is as legitimate as the assumption of physical bodies, Maddy argues. Gödel posited an analogy of intuition with perception and of mathematical realism with common-sense realism. If a statement is justified by referring to intuition Maddy calls the justification *intrinsic*. But this is not the whole story. As Maddy puts it:

> Just as there are facts about physical objects that aren't perceivable, there are facts about mathematical objects that aren't intuitable. In both cases, our belief in such 'unobservable' facts is justified by their role in our theory, by their explanatory power, their predictive success, their fruitful interconnections with other well-confirmed theories, and so on. (Maddy, 1980, 32)

In other words, in mathematics as in physics, one can justify some axioms by direct intuition (intrinsic), but others only by referring to their consequences. The acceptance of the latter axioms depends on evaluating their fruitfulness, predictive success and explanatory power. Maddy calls this type of justification *extrinsic justification*.

Maddy enlarged on the distinction between intrinsic and extrinsic justification in two ways. First, she discussed perception and intuition (1980, 36–81), trying to sketch a cognitive theory that explains how human beings arrive at the basic intuitions of set theory. There she attempted to give concrete substance to Gödel's rather abstract arguments. Second, she elaborated on the interplay of intrinsic and extrinsic justifications in modern developments of set theory (1980, 107–150). Mathematical topics treated are measurable sets, Borel sets, the Continuum hypothesis, the Zermelo–Fraenkel axioms, the axiom of choice and the axiom of constructability. She found that as a rule there is a mixture of intrinsic and extrinsic arguments in favor of an axiom. Some axioms are justified almost exclusively by extrinsic reasons. This raises the question of which modifications of axioms would make a statement like the continuum hypothesis provable and what consequences such modifications would have in other parts of mathematics. Here questions of weighing advantages and disadvantages come into play; these suggest that in the last resort extrinsic justification is uppermost.

Maddy succinctly stated the overall picture which emerges from her distinction:

> … the higher, less intuitive, levels are justified by their consequences at lower, more intuitive, levels, just as physical unobservables are justified by their ability to systematize our experience of observables. At its more theoretical reaches, then, Gödel's mathematical realism is analogous to scientific realism.

Thus Gödel's Platonistic epistemology is two-tiered: the simpler concepts and axioms are justified intrinsically by their intuitiveness; more theoretical hypotheses are justified extrinsically, by their consequences. (Maddy 1980, 33)

In conclusion, until the end of the nineteenth century, mathematicians were convinced that mathematics rested on intuitively secure intrinsic hypotheses which determined the inner identity of mathematics. Extrinsic hypotheses could occur and were necessary only outside the narrower domain of mathematics. This view dominated by and large the philosophy of mathematics. Then, non-Euclidian geometries were discovered. The subsequent discussions about the foundations of mathematics at the beginning of the twentieth century resulted in the decisive insight that pure mathematics cannot exist without hypotheses (axioms) which can only be justified extrinsically. Developments in mechanics from Newton to the nineteenth century enforced this process (see Pulte 2005).

Today, there is a general consensus that the axioms of mathematics are not absolute truths that can be sanctioned by intuition: rather, they are propositions on which people have agreed. A formalist philosophy of mathematics would be satisfied with this statement: however, modern realistic or naturalistic philosophies go further, trying to analyse scientific practice inside and outside of mathematics in order to understand how such agreements come about.

2.4 Implications for the Teaching of Proof

As we have seen, the "Hypothetical" view of modern post-Euclidean mathematics has a high affinity with the origins of proof in pre-Euclidean Greek dialectics. In dialectics, one may suppose axioms or hypotheses without assigning them epistemological qualification as evident or true. Nevertheless, at present the teaching of proof in schools is more or less ruled by an implicit, strictly Euclidean view. When proof is mentioned in the classroom, the message is above all that proof makes a proposition safe beyond doubt. The message that mathematics is an edifice of absolute truths is implicitly enforced, because the hypotheses underlying mathematics (the axioms) are not explicitly explained as such. Therefore, the hypothetical nature of mathematics remains hidden from most pupils.

This paper pleads for a different educational *approach to proof based on the modern Hypothetical view* while taking into account its affinity to the early beginnings in Greek dialectics and Greek theoretical science. This approach stresses the *relation between a deduction and the hypotheses* on which it rests (cf. Bartolini Bussi et al. 1997 and Bartolini 2009). It confronts pupils with situations in which they can *invent* hypotheses and *experiment* with them in order to understand a certain problem. The problems may come from within or from outside mathematics, from combinatorics, arithmetic, geometry, statics, kinematics, optics or real life situations. Any problem can be become the subject of a dialog or of a procedure in which hypotheses are formed and consequences are drawn from them. Hence, from the outset pupils see proof in the context of the hypothetico-deductive method.

There are mathematical and pedagogical reasons for this approach. The *mathematical* reasons refer to the demand that instruction should convey to the pupils an *authentic and adequate image of mathematics* and its role in human cognition. In particular, it is important that the pupils understand the differences and the connections between mathematics and the empirical sciences, because frequently proofs are motivated by the claim that one cannot trust empirical measurements. For example, students are frequently asked to measure the angles of a triangle, and they nearly always find that the sum of the angles is equal to 180°. However, they are then told that measurements are not precise and can establish that figure only in these individual cases. If they want to be sure that the sum of 180° is true for all triangles they have to prove it mathematically. However, for the students (and their teachers) that theorem is a statement about real (physical) space and used in numerous exercises. As such, the theorem is true when corroborated by measurement. Only if taking into account the fundamental role of measurement in the empirical sciences, can the teacher give an intellectually honest answer to the question of why a mathematical proof for the angle sum theorem is urgently desirable. Such an answer would stress that in the empirical sciences proofs do not replace measurements but are a means for building a network of statements (laws) and measurements.

The *pedagogical* reasons are derived from the consideration that the teaching of proof should explicitly address two question: (1) What is a proof? (2) Where do the axioms of mathematics come from?

Question (1) is not easy and cannot be answered in one or two sentences. I shall sketch a genetic approach to proof which aims at explicitly answering this question (see Jahnke 2005, 2007). The overall frame of this approach is the notion of the hypothetico-deductive method which is basic for all sciences: by way of a deduction, pupils derive consequences from a theory and check these against the facts. The approach consists of three phases, a first phase of *informal thought experiments* (Grade 1+); a second phase of *hypothetico-deductive thinking* (Grade 7+); and a third phase of *autonomous mathematical theories* (upper high school and university). Students of the third phase would work with closed theories and only then would "proof" mean what an educated mathematician would understand by "proof."

The first phase would be characterized by informal argumentations and would comprise what has been called "preformal proofs" (Kirsch 1979), "inhaltlich-anschauliche Beweise" (Wittmann and Müller 1988) and "proofs that explain" in contrast to proofs that only prove (Hanna 1989). These ideas are well-implemented in primary and lower secondary teaching in English-speaking countries as well as in Germany.

In the second phase the instruction should make the concept of proof an explicit theme – a major difficulty and the main reason why teachers and textbook authors mostly prefer to leave the notion of proof implicit. There is no easy definition of the very term "proof" because this concept is dependent on the concept of a theory. If one speaks about proof, one has to speak about theories, and most teachers are reluctant to speak with seventh graders about what a theory is.

The idea in the second phase is to build local theories; that is small networks of theorems. This corresponds to Freudenthal's notion of "local organization" (Freudenthal 1973, p. 458) but with a decisive modification. The idea of measuring should not be dispersed into general talk about intuition; rather we should build small networks of theorems based on empirical evidence. The networks should be manageable for the pupils, and the deductions and measurements should be organically integrated. The "small theories" comprise hypotheses which the students take for granted and deductions from these hypotheses.

For example, consider a teaching unit about the angle sum in triangles exemplifying the idea of a network combining deductions and measurements (for details, see Jahnke 2007). In this unit the alternate angle theorem is introduced as a hypothesis suggested by measurements. Then a series of consequences about the angle sums in polygons is derived from this hypothesis. Because these consequences agree with further measurements the hypothesis is strengthened. Pupils learn that a proof of the angle sum theorem makes this theorem not absolutely safe, but dependent on a hypothesis. Because we draw a lot of further consequences from this hypothesis which also can be checked by measurements, the security of the angle sum theorem is considerably enhanced by the proof.

Hence, the answer to question (1) consists in showing to the pupils by way of concrete examples the relation between hypotheses and deductions; exactly this interplay is meant by proof.

Question (2) is answered at the same time. The students will meet a large variety of hypotheses with different degrees of intuitiveness, plausibility and acceptability.

They will meet basic statements in arithmetic which in fact cannot be doubted. They will set up by themselves ad-hoc-hypotheses which might explain a certain situation. They will also hit upon hypotheses which are confirmed by the fact that the consequences agree with the phenomena. This basic approach is common to all sciences be they physics, sociology, linguistics or mathematics. We have seen above that the Greeks already had this idea and called it "saving the phenomena." The students' experience with it will lead them to a realistic image of how people have set up axioms which organize the different fields of mathematics and science. These axioms are neither given by a higher being nor expressions of eternal ideas; they are simply man made.

References

Aristotle. (1966). *Posterior Analytics*. English edition by Hugh Tredennick. London/Cambridge: William Heinemann/Harvard University Press.

Ayer, A. J., & O'Grady, J. (1992). *A dictionary of philosophical quotations*. Oxford: Blackwell Publishers.

Balacheff, N. (1991). The benefits and limits of social interaction: The case of mathematical proof. In A. J. Bishop, E. Mellin-Olsen, & J. van Dormolen (Eds.) *Mathematical knowledge: Its growth through teaching* (pp. 175–192). Dordrecht: Kluwer.

Bartolini Bussi, M. G. (2009). Experimental mathematics and the teaching and learning of proof. In *CERME 6 Proceedings*, to appear.

Bartolini Bussi, M. G., Boero, P., Mariotti, M. A., Ferri, F., & Garuti, R. (1997). Approaching geometry theorems in contexts: from history and epistemology to cognition. In *Proceedings of the 21st conference of the international group for the psychology of mathematics education (Lahti)* (Vol. 1, pp. 180–195).

Duhem, P. (1994). *Sozein ta Phainomena. Essai sur la notion de théorie physique de Platon à Galilée*. Paris: Vrin (Original publication 1908).

Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht: Reidel.

Gödel, K. (1944). Russell's Mathematical Logic. In P. A. Schilpp (Hrsg.) *The Philosophy of Bertrand Russell* (pp. 125–153). New York: Tudor. Quoted according to: P. Benacerraf, & H. Putnam (Eds.) Philosophy of mathematics, Englewood Cliffs 1964.

Hanna, G. (1989). Proofs that prove and proofs that explain. In *Proceedings of the thirteenth conference of the international group for the psychology of mathematics education* (Vol. II, pp. 45–51). Paris: PME.

Heath, T. L. (Ed. and Transl.) (1953). *The works of Archimedes*. New York: Dover.

Heath, T. L. (Ed. and Transl.) (1956). *The thirteen books of Euclid's elements*. Second edition revised with additions. New York: Dover.

Jahnke, H. N. (2005). A genetic approach to proof. In M. Bosch (Ed.) *Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education, Sant Feliu de Guíxols 2005*, 428–437.

Jahnke, H. N. (2007). Proofs and Hypotheses. *Zentralblatt für Didaktik der Mathematik (ZDM), 39*(2007), 1–2, 79–86.

Kirsch, A. (1979). Beispiele für prämathematische Beweise. In W. Dörfler, & R. Fischer (Eds.) *Beweisen im Mathematikunterricht* (pp. 261–274). Klagenfurt: Ho¨lder-Pichler-Temspsky.

Lakatos, I. (1978). A renaissance of empiricism in the recent philosophy of mathematics. In I. Lakatos (Ed.) *Philosophical papers* (Vol. 2, pp. 24–42). Cambridge: Cambridge University Press.

Lloyd, G. E. R. (1991). Saving the appearances. In G. E. R. Lloyd (Ed.), *Methods and problems in Greek science* (pp. 248–277). Cambridge: Cambridge University Press.

Lloyd, G. E. R. (2000). Der Beweis und die Idee der Wissenschaft. In J. Brunschwig, & G. E. R. Lloyd (Eds.) *Das Wissen der Griechen. Eine Enzyklopädie* (pp. 238–258). München: Wilhelm Fink.

Mach, E. (1976). *Die Mechanik. Historisch-kritisch dargestellt* (Unveränderter Nachdruck der 9. Auflage, Leipzig 1933, 1. Auflage 1883). Darmstadt: Wissenschaftliche Buchgesellschaft.

Maddy, P. (1990). *Realism in mathematics*. Oxford: Clarendon Press.

Máté, A. (2006). Árpád Szabó and Imre Lakatos, or the relation between history and philosophy of mathematics. *Perspectives on Science, 14*(3), 282–301.

Mittelstrass, J. (1962). *Die Rettung der Phänomene. Ursprung und Geschichte eines antiken Forschungsprinzips*. Berlin: Walter de Gruyter & Co.

Mancosu, P. (2001). Mathematical Explanation: Problems and Prospects. *Topoi, 20*, 97–117.

Natorp, P. (1921). *Platos Ideenlehre. Eine Einführung in den Idealismus*. Leipzig: Meiner. 1921[2].

Peirce, C. S. (1935). The essence of mathematics. In C. Hartshorne, P. Weiss (Eds.) *Collected Papers of Charles Sanders Peirce* (Vol. III, pp. 189–204). Cambridge: Harvard University Press.

Cooper, J. M. (Ed.) (1997). *Plato: Complete works*. Cambridge/Indianapolis: Hackett Publishing Company.

Proclus (1970). *A Commentary on the First Book of Euclid's Elements*. Translated, with Introduction and Notes, by Glenn R. Morrow. Princeton: Princeton University Press.

Pulte, H. (2005). *Axiomatik und Empirie. Eine wissenschaftstheoriegeschichtliche Untersuchung zur Mathematischen Naturphilosohie von Newton bis Neumann*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Russell, B. (1924). Logical Atomism. In J. M. Muirhead (Ed.), *Contemporary British Philosophy, first series* (pp. 357–383). London: George Allen & Unwin.

Szabó, Á. (1960). *Anfänge des Euklidischen Axiomensystems*. Archive for History of Exact Sciences 1, 38–106. Page numbers refer to the reprint in O. Becker (Ed.) (1965) Zur Geschichte der griechischen Mathematik (pp. 355–461). Darmstadt: Wissenschaftliche Buchgesellschaft.

Toomer, G. J. (Ed. and Transl.) (1984). *Ptolemy's Almagest*. London: Duckworth.

van der Waerden, B. L. (1988). *Science awakening I, Paperback edition*. Dordrecht: Kluwer.

Wittmann, E. C., & Müller, G. (1988). Wann ist ein Beweis ein Beweis. In P. Bender (Ed.) *Mathematikdidaktik: Theorie und Praxis. Festschrift für Heinrich Winter* (pp. 237–257). Berlin: Cornelsen.

Springer

http://www.springer.com/978-1-4419-0575-8

Explanation and Proof in Mathematics
Philosophical and Educational Perspectives
Hanna, G.; Jahnke, H.N.; Pulte, H. (Eds.)
2010, VIII, 294 p., Hardcover
ISBN: 978-1-4419-0575-8