In Paolo Mancosu (ed.), The Philosophy of Mathematical Practice [OUP, 2008], Chapter 8

8

Reflections on the Purity of Method in Hilbert's *Grundlagen der Geometrie*

MICHAEL HALLETT

8.1 Introduction: The 'Purity of Method' in the Grundlagen

The publication of Hilbert's monograph *Grundlagen der Geometrie* in 1899 marks the beginning of the modern study of the foundations of mathematics. In the *Schlusswort* to the book, Hilbert explicitly mentions the 'purity of method' question. He states first that his book was guided by 'the basic principle' of elucidating given problems in such a way as to decide whether they can be answered in a '*prescribed way with certain restricted methods*', or not so answered, and then designates this as a general principle governing the search for mathematical knowledge:

> This basic principle seems to me to contain a general and natural prescription. In fact, whenever in our mathematical work we encounter a problem or conjecture a theorem, our drive for knowledge [*Erkenntnistrieb*] is only then satisfied when we have succeeded in giving the complete solution of the problem and the rigorous proof of the theorem, or when we recognise clearly the grounds for the

This paper is a much expanded version of talks given to the Logic Group of the Department of Philosophy at Stanford in April 2004, and to the Annual Meeting of the Association of Symbolic Logic in Pittsburgh in May 2004. An abbreviated talk based on it was given at the HOPOS Biennial Conference in Paris in June 2006. I am grateful to the various participants for discussion and comments, especially Yvon Gauthier, and to Bill Demopoulos and Dirk Schlimm, who suggested helpful modifications to an earlier written version. I am also grateful to my fellow General Editors of Hilbert's unpublished lectures on foundational subjects, William Ewald, Ulrich Majer, and Wilfried Sieg, for discussions on these and other issues over many years. The generous financial support of the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

> impossibility of success and thereby the necessity of the failure. (Hilbert (1899, p. 89), p. 525 of Hallett and Majer (2004))

Hilbert goes on to stress the importance of impossibility proofs:

> Hence, in recent mathematics the question as to the impossibility of certain solutions or problems plays a very prominent role, and striving to answer a question of this kind was often the stimulus for the discovery of new and fruitful domains of research. By way of example, we recall *Abel's* proof for the impossibility of solving the equation of the fifth degree by taking roots, further the establishment of the unprovability of the Parallel Axiom, and *Hermite's* and *Lindemann's* proofs of the impossibility of constructing the numbers e and π in an algebraic way. (*Loc. cit.*, p. 89, p. 526 of Hallett and Majer (2004))

Hilbert now associates the basic principle explicitly with the idea of the 'purity of method':

> This basic principle, according to which one ought to elucidate the possibility of proofs, is very closely connected with the demand for the 'purity of method' of proof methods stressed by many modern mathematicians. At root, this demand is nothing other than a subjective interpretation of the basic principle followed here. (*Loc. cit.*, pp. 89–90; p. 526 of Hallett and Majer (2004))

Hilbert does not say what he means by the search for 'purity of method', nor who the 'many modern mathematicians' are, nor why he thinks the search for 'purity' is subjectively coloured. However, part of what he means by questions of purity is this: one can enquire of a given proof or of a given mathematical development whether or not the means it uses are 'appropriate' to the subject matter, whether one way of doing things is 'right', whereas another, equivalent way is 'improper'. Is it 'appropriate' to use complex function theory in proofs in number theory, as had become common following the work of Riemann and Dirichlet, or transfinite numbers in proofs in analysis or point-set theory?[1] Is it 'right' or 'better' to pursue geometry synthetically rather than analytically? Hilbert's reaction to this question would be to say that neither way is 'right', that each is the right way to do things if certain purposes are kept in mind, and the 'right' approach is to embrace both developments. What is genuinely explanatory for one mathematician might be simply opaque for another, and to insist on one method would be regarded as unjustifiably idiosyncratic, or (as he says here) 'subjective'. No one kind of mathematical knowledge is, in general, superior to another. Again, the 'appropriateness' of proof methods might easily be taken as a matter of personal taste. More importantly, to put

[1] For analysis of the use of transfinite numbers in the original proof of the Heine–Borel Theorem, see Hallett (1979*a, b*).

all these questions on an objective footing requires gathering certain kinds of information. For example, we require answers to questions of the following sort: Can two different ways of developing a theory match each other in the kinds of theorem they can prove? Can a certain theorem *only* be proved by using the extended theoretical means? Are the connections so established accidental, or can one find a deeper theoretical reason for them?[2] To answer such questions as these, as Hilbert suggests, is just to pursue the 'basic principle' set out in the *Grundlagen der Geometrie*. Certainly this is what Hilbert does *say* with respect to the principle's application in the case of his work on geometry:

> In fact, the geometric investigation carried out here seeks in general to cast light on the question of which axioms, assumptions or auxilliary means are necessary in the proof of a given elementary geometrical truth, and it is left up to discretionary judgement [*Ermessen*] in each individual case which method of proof is to be preferred, depending on the standpoint adopted. (*Loc. cit.*, pp. 89–90, p. 526 in Hallett and Majer (2004))

In other words, in geometry, no general decision is to be made as to what is to be preferred and what not. The purpose of the foundational investigation, neutrally stated as it is in the *Grundlagen der Geometrie*, is therefore to assess what we might call the 'logical weight' of the axioms and central theorems. This, in any case, is what there is in abundance in Hilbert's work on geometry. We will come to some examples in due course.

The *Grundlagen* itself was immediately preceded by a long series of lectures held in 1898/1899, which Hilbert entitled 'Elemente der Euklidischen Geometrie'. These lectures contain most of what is novel in the *Grundlagen*, but they contain also many more philosophical and informal remarks, and are very differently arranged from the presentation in the *Grundlagen der Geometrie*. The notes for the lectures exist in two different forms: 110 pages of Hilbert's own notes, and then a beautifully executed protocol of the lectures following the notes very carefully (in German, an '*Ausarbeitung*'), which Hilbert had commissioned from his first doctoral student at Göttingen, Hans von Schaper, whose own field of research was analytic number theory, specifically the Prime Number Theorem.[3] Towards the end of these notes, there are several passages which are clearly the origin of the citations from the *Grundlagen* given above.

[2] Recall the emphasis in Hilbert's paper of 1918 on the '*Tieferlegung der Fundamente*'. See Hilbert (1918).

[3] For a more detailed description of the relationship between the 1898/1899 lectures and the *Grundlagen* of 1899, as well as that between the notes and the *Ausarbeitung*, see my Introduction to Chapter 4 in Hallett and Majer (2004).

At the end of the main part of the *Ausarbeitung*, there is the following statement about unprovability results:

> An essential part of our investigation consisted in *proofs of the unprovability* of certain propositions; in conclusion, we recall that proofs of this kind play a large role in modern mathematics, and have shown themselves to be fruitful. One only has to think of the squaring of the circle, of the solution of equations of the fifth degree by extracting roots, Poincaré's theorem that there are no unique integrals except for the known ones, etc. (Hilbert (*1899, p. 169), p. 392 of Hallett and Majer (2004))[4]

In his own notes for these lectures, Hilbert writes:

> The subjects we deal with here are old, originating with Euclid: the principle of the proof of unprovability is modern and arises first with two problems, the squaring of the circle and the Parallel Axiom. <[*Interlineated addition:*] Thus, solution of a problem impossible or impossible with certain means. With this is connected the demand for the purity of method.> However, we wish to set this as a modern principle: One should not stand aside when something in mathematics does not succeed; one should only be satisfied when we have gained insight into its unprovability. Most fruitful and deepest principle in mathematics. (Hilbert (*1898/1899, p. 106), p. 284 in Hallett and Majer (2004))

Moreover, earlier in his notes Hilbert had written the following (the particular example concerned will be discussed in Section 8.4.1; for the moment the details do not matter):

> Thus here for the first time we subject *the means of carrying out a proof* to a critical *analysis*. It is modern everywhere to guarantee the *purity* of method. Indeed, this is quite in order. In many cases our understanding is not satisfied when, in a proof of a proposition of arithmetic, we appeal to geometry, or in proving a *geometrical truth* we draw on *function theory*.[5]

But Hilbert immediately goes on to say this:

> Nevertheless, drawing on differently constituted means has frequently a *deeper and justified* ground, and this has uncovered beautiful and *fruitful relations*; e.g. the prime number problem and the $\zeta(x)$ function, potential theory and analytic functions, etc. In any case, one should never leave such an occurrence of the mutual interaction of different domains unattended. (Hilbert (*1898/1899, p. 30), p. 237 of Hallett and Majer (2004))

Thus, even while he acknowledges the epistemological disquiet behind many purity questions, Hilbert admits that 'impure' mixtures might point

[4] Bibliographical items with the date preceded by an asterisk were unpublished by Hilbert.

[5] This remark has a special interest in view of one of the examples I will present later, namely that concerning the analysis of the proof of the elementary Isoceles Triangle Theorem.

to something important and deep. Thus, the first lesson which might be drawn is that a standard 'purity' question ('Is this means necessary to this end?') is often an occasion for a *foundational* investigation, and this is not carried out to show that a certain kind of preferred knowledge is or is not sufficient, but rather for reasons of *mathematical* productivity and logical clarity, and that answering such possibility and impossibility questions was frequently the occasion for opening up 'new and fruitful domains of research', as happened, say, with Abel's investigation.

The heuristic value of unsolved problems was stressed in a powerful way by Hilbert in his famous lecture on mathematical problems (Hilbert, 1900*c*) held a little over a year later. Hilbert stresses that the lack of a solution to a problem might well raise the suspicion that it is indeed insoluble as stated, or with the expected means, and that one might then seek a demonstration of the relevant *unprovability*. Such a demonstration for Hilbert counts as 'a fully satisfactory and rigorous solution' of the problem at hand (for example, the problems concerning whether we can prove the Parallel Axiom or square the circle), 'although in a different sense from that originally intended'; see Hilbert (1900*c*, p. 261), English translation, p. 1102. Settling such problems in this way is, as Hilbert says, in no large part responsible for the legendary mathematical optimism he displays, i.e. for his 'conviction' of the solvability of every mathematical problem. Hilbert says of the achievement of impossibility proofs:

> It is probably this remarkable fact alongside other, philosophical, reasons which gives rise in us to the conviction (shared by every mathematician, but which, at least hitherto, no one has supported by a proof) that every definite mathematical problem must necessarily be susceptible of exact settlement, be it in the form of an answer to the question as first posed, or be it in the form of a proof of the impossibility of its solution, whereby it will be shown that all attempts must necessarily fail. …
>
> This conviction of the solvability of every mathematical problem is a powerful spur in our work. We hear within us the perpetual call: *There is the problem; seek its solution. You can find it by pure thought, for in mathematics there is no ignorabimus!* (Hilbert, 1900*c*, pp. 261–262)

This is not merely a peculiarity characteristic of mathematical thought alone, but rather what he calls a 'general law' (or an 'axiom') inherent in the nature of the mind, that all questions 'which it [the mind] asks must be answerable'. This is an important remark, and I will return to it. For the moment, suffice it to say that this is why Hilbert seeks 'projection onto the conceptual level', a term we will elucidate in Section 8.3.1 (see especially p. 217).

These remarks suggest why Hilbert says that seeking insight into apparent cases of unprovability is the 'most fruitful and deepest principle in mathematics', and partly explain therefore his interest in purity questions. But it would be misleading to suggest that Hilbert regards 'purity' questions as *merely* heuristic in this way, at least judging by his foundational work on geometry, which Hilbert took to be the paradigm of foundational analysis. Purity of method investigations, in so far as they concern not just open problems but also the analysis of the sources of mathematical knowledge, *was* fundamentally important to Hilbert. But to understand exactly how, it is necessary to recognize the novelty of Hilbert's approach to foundational issues, and to see the considerable effect this had on the type of mathematical knowledge obtained through a purity investigation. There is a sense in which mixture of mathematical domains is intrinsic to Hilbert's investigations; on the one hand, higher-level mathematics is essential to foundational investigation of theories, even relatively low-level ones, and this higher level also instructs the fundamental source of geometrical knowledge. On the other hand, the examples given in Sections 8.4.1–8.4.3 make it clear that, while 'purity' results *do* throw light on on the question of the appropriate sources of knowledge for geometry, parallel to this are abstract mathematical/logical results at what Hilbert calls the 'conceptual level'. Not only does this level largely emancipate mathematics from the epistemological constraints of the 'appropriate', but it is an essential part of what is attained by a 'purity' result. But there is a twist, which we will consider at the end in Section 8.5.

8.2 The foundational project

Before we go further, it is worth pointing out a number of things about Hilbert's approach to geometry.

In the first place, Hilbert's axiomatic presentation as it appeared in the *Grundlagen* of 1899 (and, to some extent, the preceding lectures) builds on the tradition of synthetic geometry, a tradition which saw a strong revival through the work of Monge and von Staudt in the middle of the 19th century. In particular, Hilbert attempted to avoid where possible the direct intrusion of numerical elements. Underlying this, at least in part, was a view that geometry is of empirical and intuitive origin, and concerns 'the properties of things in space'.[6] This is concisely summarized in Hilbert's introduction to his 1891 lectures on projective geometry:

[6] This view is what underlies the *Vorlesungen* of Pasch from 1882. It is, of course, of much older provenance.

> I require *intuition and experimentation*, just as with the founding of physical laws, where also the *subject matter* [Materie] *is given through the senses*.
>
> In fact, therefore, the *oldest geometry* arises from *contemplation of things* in space, as they are given *in daily life*, and like all science at the beginning had posed *problems of practical importance*. It also rests on the *simplest kind of experimentation* that one can perform, namely on *drawing*. (Hilbert (*1891, Introduction, p. 7), p. 23 in Hallett and Majer (2004))

But he goes on to say that Euclidean geometry had '... an *essential defect*: it had no *general method*, without which a *fruitful further development* of the science is impossible' (*ibid.*, p. 8). This defect was rectified through the invention of analytic (Cartesian) geometry, which indeed provided a powerful, unified method. Nevertheless, this brought its own disadvantages:

> As important as this *step forward* was, and as *wonderful as the successes were*, nevertheless *geometry* as such *in the end suffered* under the *one-sided development of this* method. One *calculated* exclusively, *without having any intuition of what was calculated*. One lost the *sense for the geometrical figure*, and for *the geometrical construction*. (Hibert (*1891, p. 10), p. 24 in Hallett and Majer (2004))

In what follows, Hilbert makes it clear that he sees the movement in the 19th century to promote synthetic geometry as at least in part a reaction to this. This movement concentrated on projective geometry (what was often called '*Geometrie der Lage*'), but Hilbert's aim was to reformulate and restructure full Euclidean geometry itself as far as possible in an essentially synthetic way. In doing so, he develops geometry in a modern axiom system building up from the simplest possible projective framework (an incidence and order geometry), and arriving at full Euclidean geometry with, for example, the standard results from a Euclidean theory of congruence, of proportions, area (or surface measure), and parallels, all of which is sometimes called by Hilbert 'school geometry', and developed (where possible) before continuity is broached. This synthetic restructuring is much clearer in the 1898/1899 lectures which preceded the *Grundlagen* than it is in the *Grundlagen* itself.

The tone of the remarks from 1891 leaves the impression that, according to Hilbert, Euclidean geometry represents knowledge of a certain kind, an impression which is strengthened by Hilbert's repeated declaration that an axiomatisation in any area of science always begins with a certain domain of 'facts [*Tatsachen*]'. By this, Hilbert does not mean just facts in the sense of empirical facts or even established truths, though such things might be included, but simply what over time has come to be accepted, for example, from an accumulation of proofs or observations. Geometry, of course, is the central example: there are empirical investigations, over 2,000 years of mathematical development, including 300 years' experience with analytic geometry, and the 19th-century growth of this into function-theoretic investigations, the discovery of various 'impossibilities', and above all the independence of Euclid's Axiom of Parallels. Within this domain, certain 'basic facts' are isolated (different sets of 'basic facts' could be chosen, depending on the purpose), and an axiomatization based on these is designed, central assumptions being identified and grouped.

This grouping, fundamental to Hilbert's mature presentation of geometry, is based on a natural conceptual division (e.g. all the congruence assumptions are grouped together), but this conceptual division also in its turn corresponds to different levels of empirical/intuitive justification. For instance, in his short holiday lecture course from 1898, Hilbert says, when presenting Euclidean geometry:

> 4.) 5.) do not have the same *empirical*, constructible character as 1.) 2.) 3.), these being established through a finite number of experiments [*Versuchen*]. (Hilbert (*1898, p. 18), p. 171 in Hallet and Majer (2004))

Axiom groups 1.)–3.) here are the incidence, order, and congruence axioms, while 4.) and 5.) are the Parallel and Archimedean axioms respectively. In the 1898/1899 lecture notes, he writes:

> A *general remark on the character of our axioms I–V* might be pertinent here. The axioms I–III [incidence, order, congruence] state very simple, one could even say, original facts; their validity in nature can easily be demonstrated through experiment. Against this, however, the validity of IV and V [parallels and continuity in the form of the Archimedean Axiom] is not so immediately clear. The experimental confirmation of these demands a greater number of experiments. (Hilbert (*1899, pp. 145–146), p. 380 in Hallett and Majer (2004))

For Hilbert, central independence results concerning the axioms underline the conceptual division, for they show that different levels of empirical/intuitive support are essential. For example, Hilbert regarded the Archimedean Axiom, interpreted as about actual space, as something like an empirical principle (or 'highly intuitive'), since it encapsulates the assumption that all distances, the cosmic and the sub-atomic, can be measured along the same scale. But its independence shows that it has to be supported on grounds separate from the arguments for the more elementary axioms, e.g. congruence, since it is not a logical consequence of them, and 'logic does not demand it'.[7]

[7] See Hilbert (*1917/1918, 18; 1918, 408–409; 1919, 14–15, book p. 11) and also Hilbert (*1905, 97–98).

Similar points hold about the Parallel Axiom. In his 1891 lectures on projective geometry, Hilbert remarks in a note that:

> This *Parallel Axiom* is furnished by intuition. *Whether this latter is innate or nurtured*, whether *the axiom corresponds to the truth*, whether it *must be confirmed by experience*, or *whether this is not necessary*, none of this concerns us here. We treat of intuition, and this demands the axiom. (Hilbert (*1891, p. 18), p. 27 in Hallett and Majer (2004). See also the 1894 lectures, pp. 88–89, pp. 120–121 of (Hallett and Majer, 2004).

In the *Ausarbeitung* of the 1898/1899 lecture notes, Hilbert is somewhat more circumspect:

> ... the question as to whether our intuition of space has an *a priori* or empirical origin remains unelucidated. (Hilbert, *1899, p. 2) p. 303 in Hallett and Majer (2004).

But the proof of the independence of the Axiom shows again that a distinct justification has to be given:

> Even the philosophical value of the investigation should not be underestimated. If we wish to apply geometry to reality [*Wirklichkeit*], then intuition and observation must first be called on. It emerges as advantageous to take certain small bodies as points, very long things with a small cross-section, like for instance taught threads etc., as straight lines, and so on. Then one makes the observation that a straight line is determined by two points, and in this way one observes [as correct] the other facts expressed in the Axioms I–III of the schema of concepts. Non-Euclidean geometry, i.e. the axiomatic investigation of the Parallel Proposition, states then that to know that the angle sum in a triangle is 2 right angles a new observation is necessary, that this is no way follows from the earlier observations (respectively from their idealised and more precise contents). (Hilbert (*1905, pp. 97–98))

For Hilbert, this shows that Euclid's instincts had been quite right, that the axiom is *required* as a new assumption to prove certain central, intuitively correct 'facts' such as the angle-sum theorem or the existence of a rectangle:

> Even if Euclid did not state these axioms [incidence, order, and congruence] completely explicitly, nevertheless they correspond to what was intended by him and his successors down to recent times. However, when Euclid wanted to prove further propositions immediately furnished by intuition, propositions such as the presence of a quadrilateral with four right-angles, so he recognised that these axioms do not suffice, and therefore erected his famous Parallel Axiom. ... The brilliance it demanded to adopt this proposition as an axiom can best be seen in the short historical sketch: Stäckel–Engel Parallellinien, Teubner, 1895 (Stäckel and Engel, 1895. (Hilbert's lecture notes for 1898/1899, p. 70; p. 261 in Hallett and Majer (2004))

Two pages later, Hilbert remarks:

> One can indeed say that Gauss was the *first one* in about *2100 years* who first *understood and completely grasped* why Euclid adopted the Axiom of Parallels as one of the axioms. (*Ibid.*, p. 74; p. 263 in Hallett and Majer (2004))[8]

The empirical/intuitive admixture and the conceptual division and its corresponding demands for different levels of empirical justification is a large part of what lies behind the statement in the published *Grundlagen* (p. 4) that each of the Axiom Groups represents 'certain connected facts of our intuition' (Hallett and Majer, 2004, p. 437).

Just as important as the acknowledgement of the empirical/intuitive roots of geometry is a precise assessment of the relation of Hilbert's work to Euclid's. An important component of Euclid's work in the opening books of the *Elements* was the desire to show that certain central propositions can be established using only 'restricted methods', and thus embodies a certain 'purity of method' concern. By this is not merely meant that Euclid's system was an axiomatization; it is illustrated rather by the way in which Euclid withholds deployment of the Parallel Postulate until rather late on in Book I, and by the way he restricts use of congruence arguments using the translation and flipping of figures in the plane, even when proofs would be easier when these techniques are employed. Some of Hilbert's work in geometry is a direct continuation of this kind of investigation. For one thing, Hilbert's lectures of 1898/1899 which immediately precede the *Grundlagen*, are on Euclidean geometry, concentrating on, and analyzing, the theoretical results in the early part of the *Elements*, in particular on congruence, proportion and area, the involvement (or exclusion) of the Archimedean Axiom, and the Parallel Axiom itself. Hilbert shows, for example, that we can dispense with what is initially adopted in his lectures as Axiom 9, added to the Euclidean system by many of Hilbert's predecessors in the 19th century, stating the assumption that all 'straight angles' (angles on a straight-line) are congruent. More significantly, Hilbert also shows that a Euclidean theory of linear proportion and of surface content (roughly, polygonal area) can be developed without the implicit assumption that the content measures or the 'lengths' are themselves magnitudes, an assumption he directly criticizes Euclid for.[9] Connected to this is Hilbert's proof of what

[8] The examination of the Parallel Axiom is an excellent example of what Hilbert calls 'analyzing intuition': independence confirms Euclid's 'intuition' to use the axiom. We will comment on 'analyzing intuition' below; see the remarks on p. 292.

[9] Hilbert's student Dehn showed that the same does *not* hold for tetrahedral/polyhedral volume; the Archimedean Axiom is required. Hilbert posed the problem in the late 1890s; see the 1898 *Ferienkurs* (p. 26), and the notes for the 1898/1899 lectures (pp. 106 and 169 respectively), all in Hallett and Majer

he calls the Pascal Theorem (usually called Pappus's Theorem) using just the plane part of the elementary axioms together with congruence, but without any involvement of continuity, a result new to Hilbert. More generally, the whole of 'school geometry' can be developed without continuity assumptions, showing that Euclid was right not to include explicit continuity principles among his axioms.

Of course, there are ways in which Hilbert's treatment is not 'purely' or historically Euclidean, for example in developing the core of projective geometry on the basis of the order and incidence axioms alone (i.e. before the Congruence and Parallel Axioms have been introduced), and in its further concentration on the Desargues and Pascal theorems. An underlying concern here is the relationship between Euclidean geometry and the coordinate structure of analytic geometry, for part of what Hilbert investigates are the hidden field properties in segment structure showing that the basic magnitude principles (represented by the core of the ordered field axioms) are true of linear segments once addition and multiplication operations have been defined for them in a reasonable way. Part of the point is surely to defend Euclid, in that Hilbert shows that the 'theory of magnitudes' arises intrinsically, and does not have to be imposed from without by some extra assumptions, but another motive is clearly to show that the central guiding assumption of analytic geometry, coordinatization by real numbers, is not *ad hoc*, a central concern of Hilbert's since at least the 1893/1894 lectures. The guiding insight is clearly that, since analytic geometry was, at the time Hilbert was writing, the pre-eminent way to pursue Euclidean geometry, careful analysis of any suitable synthetic replacement should reveal some central conceptual parallels with analytic structures.[10] Indeed, despite the desire to keep synthetic Euclidean geometry as far as possible independent of analytic geometry, Hilbert did impose a strong adequacy condition, namely 'completeness' with respect to analytic geometry, i.e. the demand that a satisfactory synthetic axiomatization should be able to prove all the geometrical results that analytic geometry could.[11]

Nevertheless, despite these modern flourishes, it is clear (above all from the 1898/1899 lectures) that Hilbert's own investigations were profoundly influenced by Euclid's. In the Introduction to his 1891 lectures on projective geometry, Hilbert gives a short but highly illustrative survey of geometry. He

(2004, pp. 177, 284, 392 respectively). The problem reappears as Problem 3 in Hilbert's famous list of mathematical problems set out in 1900; see Hilbert (1900*c*, 266–267).

[10] Desargues's Theorem is essential to this: see Section 8.4.1.

[11] For a discussion of what Hilbert meant by completeness in 1899 when calling for a 'complete' axiomatization of geometry, see Section 5 of my Introduction to Chapter 5 in Hallett and Majer (2004, 426–435).

divides geometry into three domains: (1) intuitive geometry, which includes 'school geometry', projective geometry, and what he calls 'analysis situs'; (2) axiomatic geometry; and (3) analytic geometry. Axiomatic geometry, according to Hilbert, 'investigates which axioms are used in the garnering of the facts in intuitive geometry, and sets up systematically for comparison those geometries in which various of these axioms are omitted', and its main importance is 'epistemological'.[12] The description of axiomatic geometry is a reasonably good, rough description of much of what Hilbert actually carries out systematically in the period from 1898 to 1903. He does indeed investigate the 'facts' obtained from intuition; in the *Grundlagen* (p. 3) he actually describes his project as being fundamentally a 'logical analysis of our spatial intuition' (Hallett and Majer, 2004, p. 436), a description which also appears in the *Ausarbeitung* of the 1898/1899 lecture notes (p. 2), where he says 'we can outline our task as constituting a *logical analysis of our faculty of intuition* [*Anschauungsvermögens*]', and where 'the question of whether spatial intuition has an apriori or empirical character is not hereby elucidated' (see Hallett and Majer 2004, p. 303). Hilbert *does* consider the geometries one obtains when various axioms are 'set aside', non-Euclidean geometry, non-Archimedean geometry, non-Pythagorean geometry, etc., and we will see some examples later on. Such investigations necessarily embrace questions of what can be, and cannot be, proved on the basis of certain central propositions, either axioms or central theorems. That the main benefit of doing this is said by Hilbert to be 'epistemological' is also understandable. If empirical investigation and geometrical intuition are the first sources of geometrical knowledge, then Hilbert's dissection of Euclidean geometry is indeed an 'analysis' of this source, revealing the propositions responsible for various central parts of our intuitive geometrical knowledge.

In short, surely some of Hilbert's work can be seen as stemming from a rather straighforward epistemological concern with the purity of method, namely showing that P (or some theoretical development $\mathcal{T}$) can be deduced solely using some specified axioms Σ (or more generally $\Sigma - \Gamma$), and this is designed at least to explore the epistemological underpinnings of the axioms.[13] This kind of investigation fits with the *traditional* conception of axiomatic

[12] See Hilbert (*1891, 3–5), pp. 21–22 in Hallett and Majer (2004).

[13] Another geometrical example outside the framework of Euclidean geometry might be Hilbert's analysis of Lie's work, where Lie is criticized for the way he uses the full theory of differentiable functions to analyze the concept of motion. Hilbert showed that much weaker assumptions than Lie's (assumptions approximating more to the 'ancient Euclid', as he puts it) will suffice, Lie's assumptions being 'foreign to the subject matter, and because of this superfluous'. For further details, see Hallett and Majer (2004, 9).

investigation, and might be used to describe, say, Pasch's work on projective geometry in his (Pasch, 1882), an important influence on Hilbert's work, and also Frege's work on arithmetic, for Frege attempted to show that natural number arithmetic can be derived from purely logical principles alone, with the help only of appropriate definitions. Both Frege's and Pasch's projects attempt to show that, when properly reduced or reconstructed, the respective mathematical theories represent knowledge of a certain, definite kind, logical and empirical respectively. The projects thus have conscious epistemological aims, and are very broadly 'Euclidean' in the general sense that they attempt to demonstrate that certain bodies of knowledge can be *deduced* from a stock of principles circumscribed in advance, a demonstration which can only be effected by the actual construction of the deductions.

While these *are* very important elements in Hilbert's treatment of Euclidean geometry, they are by no means exhaustive.

For one thing, the description of Hilbert's geometrical work as an extension of the Euclidean project does nothing to explain its entirely novel contribution to foundational study. The novelty comes in treating the *converse* of the Euclidean foundational investigation, asking in addition the question: Can we show that a given proposition P (or some theoretical development $\mathcal{T}$) *cannot* be deduced (carried out) solely using the precisely specified 'restricted methods' Σ?[14] In other words, as Hilbert puts it in the *Grundlagen* (pp. 89–90), we seek to 'cast light on the question of which axioms, assumptions or auxilliary means are necessary in the proof of certain elementary geometrical truths'. It is important to see that Hilbert's investigations are not just *complementary* to the Euclidean ones; indeed, to describe them so would be to underemphasize their novelty. The very pursuit of this new line of investigation involved Hilbert in a radical transformation of the axiomatic pursuit of mathematics. As Hilbert expressed it in some lectures from 1921/1922 (Hilbert, *1921/1922, pp. 1–3):

> The further development of the exact sciences brought with it an essential transformation of the axiomatic method. On the one hand, one found that the propositions laid down as axioms in no sense could be held sublimely free from doubt, where no difference of opinion is possible. In particular, in geometry the

[14] There are two important variations to this:

1. Show that P (or some theoretical development $\mathcal{T}$) *can* be deduced (carried out) using Σ, but *not* using Σ^- (slight weakening)
2. Show that P (or some theoretical development $\mathcal{T}$) *cannot* be deduced (carried out) using Σ, but *can* with Σ^+ (slight strengthening).

> evidence in favour of the Parallel Axiom was called into question. ... In this way, there developed the view that the essential thing in the axiomatic method does not consist in the securing of absolute certainty, which is transmitted to the theorems by means of logic, but in this, that the investigation of the logical interconnections is separated off from the question of the actual truth of the axioms.[15]

This Hilbert emphasized as being the 'main service' of the axiomatic method. The interest is therefore much more logical than epistemological.

The remark of Hilbert's just quoted leads to a second sense in which Hilbert's work is new, and which has not yet become clear: part of the point of Hilbert's investigation is to effect an *emancipation* from the sources of knowledge provided by the 'facts' or by intuition. As I hope will be shown, this plays a significant part in understanding the role of purity of method investigations in Hilbert's investigation of geometry.

The radicalness of Hilbert's approach has three important features, to be dealt with briefly in the next section.

8.3 Independence and metamathematical investigation

8.3.1 Interpretation

It is clear that in his study of geometry, Hilbert's focus is less on questions of provability from a circumscribed stock of principles than on *unprovability*, in other words, on independence. The basic technique which Hilbert adopted for this investigation is that of modelling, more strictly, of translating the theory to be investigated into another mathematical theory. For this, it is essential (and Hilbert is very clear about this) that the primitive concepts employed are not tied to their usual fixed meanings; they must be free for *reinterpretation*. Indeed, in his 1898/1899 lectures Hilbert stresses that the most difficult part in carrying out the investigation will be separating the basic terms from their usual, intuitive associations.[16] It follows that the axioms cease to have fixed meaning, and thus cease to be, for someone like Frege, genuine axioms at all.

This was not just a matter of expediency for Hilbert, done for the sake of the independence proofs; along with it goes a new picture of mature mathematics,

[15] See also Bernays (1922, 95). Bernays's remarks are very reminiscent of the long passage from Hilbert (*1921/1922) just quoted. Bernays was the *Ausarbeiter* for the lecture notes.

[16] See p. 7 of the 1898/1899 lectures (Hallett and Majer, 2004, 223).

intrinsic to which is the view that in general no *one* interpretation of an axiom system is privileged above others, despite what might seem like the overwhelming weight of the interpretation underlying the 'facts' as originally given: thus in the case of geometry the weight of the 'intuitive' or 'empirical' origins. Thus, we see a radical departure from the kind of enterprise Frege and Pasch (and even Euclid) were engaged in, enterprises part of whose very point was to *explain* (and thereby *delimit*) the primitives. Hilbert's axiomatic method abandons the direct concern with the kind of knowledge the individual propositions represent because they are about the primitives they are, and concentrates instead on what he calls 'the logical relationships' between the propositions in a theory.

Hilbert's fundamental supposition of foundational investigation, going back to 1894 and stated repeatedly thereafter, is that a theory is only 'a schema of concepts' which can be variously 'filled with material'. He says:

> In general one must say: Our theory furnishes only the schema of concepts, which are connected to one another through the unalterable laws of logic. It is left to the human understanding how it applies this to appearances, how it fills it with material [*Stoff*]. This can happen in a great many ways. (Hilbert (*1893/1894, p. 60), or Hallett and Majer (2004, p. 104))

In the 1921/1922 lectures already referred to, Hilbert calls axiomatization of this kind a 'projection into the conceptual sphere':

> According to this point of view, the method of the axiomatic construction of a theory presents itself as the procedure of the mapping [*Abbildung*] of a domain of knowledge onto a framework of concepts, which is carried out in such a way that to the objects of the domain of knowledge there now correspond the concepts, and to statements about the objects there correspond the logical relations between the concepts.
>
> Through this mapping, the investigation is completely severed from concrete reality [*Wirklichkeit*]. The theory has now absolutely nothing more to do with the real subject matter or with the intuitive content of knowledge; it is a pure *Gedankengebilde* [construct of thought] about which one cannot say that it is true or it is false. (Hilbert (*1921/1922, p. 3), forthcoming in Ewald and Sieg (2008))

It is not that a mathematical theory in this sense has nothing to do with reality; indeed it may have *more* to do with it, for the connections might be established 'in a great many ways', to use Hilbert's phrase from 1894.

It is important to see how this view fits with the insistence in the 1890s, outlined above, that the root of geometry is empirical, and that geometry is, as Hilbert frequently put it, the 'most perfect natural science'. For Hilbert, geometry is a natural science primarily because it can be *applied* to nature to furnish a more or less accurate *description* of it. The term 'description' is important. In the *Vorrede* to his lectures on mechanics published in 1877 (Kirchhoff, 1877), Kirchhoff insists that physics should only set itself the restricted aim of describing the phenomena, and not that of trying to get at the underlying 'causes', since these frequently suffer from deep-seated 'conceptual unclarity'. In his 1894 lectures on the foundations of geometry, Hilbert refers directly to Kirchhoff's aim of 'describing' only, and states that correspondingly the aim of geometry is to 'describe' the 'geometrical facts'. See Hilbert (*1893/1894, p. 7) or Hallett and Majer (2004, p. 72). And the basis of application in this sense is *interpretation* (or, as Hilbert says, '*Deutung*'), and this is essentially inexact. In these 1894 lectures, Hilbert says:

> With the axioms given hitherto, the existence [incidence] and position [order] axioms, we can already describe a large collection of geometrical facts and phenomena. We require only to take bodies for points, straight lines and planes, for the relation of passing through, touching, for being definite, immovable or fixed (perhaps, in the unrefined sense, when nudged by the hand). The bodies we should think of as finite in number, and such that the axioms are satisfied under this interpretation [*Deutung*] (for which, as one recognises, it will be necessary to have for the bodies taken in place of points, straight lines, planes, something like grains, rods or stretched threads or wire, cardboard) and indeed precisely. Then we know that all the propositions set up so far are also satisfied, and indeed precisely satisfied.

The direct continuation of this passage also makes it clear that applications only hold approximately, and this is stated in one breath with the claim that theories are only schemata of concepts, represented here by the ellipsis:

> If one finds that, with an application, the propositions are not satisfied (or not precisely satisfied), this arises because an inappropriate application has been taken, i.e. the bodies, movement, touching do not agree with our scheme of axioms. In this case it will be necessary to replace the things: bodies, movable, touching, by others, perhaps by smaller grains, blots [*Klexe* (sic)], tips, thinner wires, thinner cardboard, touching with firmer contact, movability [of the bodies] even when we blow on them [*Anpusten*], in such a way that the axioms are satisfied. Then we know that the propositions also hold (precisely). ... But always when the axioms are satisfied, the propositions also hold. The easier and more far reaching the application, then so much better*) the theory.
>
> ---
>
> *) All systems of units and axioms which describe the phenomena are equally justified. Show nevertheless that the system given here is in a certain respect the uniquely possible one. (Hilbert (*1893/1894, pp. 60–60A), pp. 104–105 in Hallett and Majer (2004))

Hilbert repeatedly stresses the inexactness of application, for example on p. 92 of the 1893/1894 lectures (Hallett and Majer, 2004, p. 122), or p. 106 of the

1898/1899 lectures (Hallett and Majer, 2004, p. 283). And he even states the approximate nature of the physical interpretations of geometry as a *reason* why we require the logical development of geometry separately:

> In physics and nature generally, and even in practical geometry, the axioms all hold only approximately (perhaps even the Archimedean Axiom). One must however take the axioms precisely, and then draw the precise logical consequences, because otherwise one would obtain absolutely no logical overview. Necessarily finite number of axioms, because of the finitude of our thought. (Note on the front cover of Hilbert's copy of the *Ausarbeitung*. Date unknown, but after 1899, and very probably before 1902; see Appendix to Chapter 4 in Hallett and Majer (2004, p. 401), remark [15].)

Two things are immediately clear from these passages. (1) Many interpretations are possible, and even desirable, even where we are concerned apparently with only one general area of application (the application of the *same* theory to the *same* spatial world). (2) There is an implicit assumption that the (unspecified) logical apparatus is *sound*, i.e. if the axioms hold under an interpretation, then so do the theorems. (See also Hilbert's letter to Frege of 29.xii.1899.) Most importantly, this latter means that the internal (logical) workings of the theory, in short ('pure' or 'free') mathematics, can proceed independently of any particular application, and this is the case even when one might firmly believe that application is the primary purpose. This is stressed by Hilbert in the continuation of the passage from the 1921/1922 lectures quoted on page 296:

> Nevertheless this framework of concepts has a meaning for knowledge of the actual world, because it represents a 'possible form in which things are actually connected'. It is the task of mathematics to develop such conceptual frameworks in a logical way, be it that one is led to them by experience or by systematic speculation. (Hilbert (*1921/1922, p. 3) in Ewald and Sieg (2008))

In short, interpretation (quite possibly manifold interpretations) can establish *various* connections to the world, and the more the better.[17] Thus, the mathematical theory is not *determined* by *Wirklichkeit*, it does not necessarily extend our knowledge of it (it might or might not), and is in the end not responsible to it. Mathematics can learn from intuition, observation, and empirical investigation more generally, but is not to be their slave, even when they have played a major part in the establishment of the domain of 'facts', and therefore in the axiomatization itself. A prime example is the

[17] See again Hilbert's letter to Frege of 29.xii.1899. In his own letter of 27.xii.1899, to which Hilbert's is a reply, Frege had objected to the very idea of considering different interpretations for geometry.

formulation of the congruence axioms. Linear congruence in geometry, both the idea of congruence, and the central propositions governing it, was originally motivated by simple observations about movement of rigid bodies in space, but Hilbert's axioms are no longer to do with movement itself, though the connection between them and the movement of rigid bodies is not hard to discern. Rather, Hilbert's view (see Hilbert, *1899, pp. 59–60) is that a proper mathematical analysis of spatial movement *requires* an independently established and neutral notion of congruence. Thus the abstract notions are to be applied in the analysis of movement, but geometry with Hilbert's congruence axioms is not dependent upon the purely empirical matter of whether or not there are in fact rigid bodies, and thus whether bodies can indeed ever be congruent in the intuitive sense. In pursuing mathematical investigations of geometry, one is investigating 'possible forms of connection' and not necessarily 'actual connections', an emancipation which reflects the ascent to what Hilbert calls the conceptual.

It follows from this account that even *if* the underlying notion of intuition were strong enough to guarantee the 'apodeictic certainty' of the axioms, it would still be possible to drop axioms, or modify them, or replace them, and it would still be part of the task of mathematics to investigate the consequences of so doing. In other words, a strong notion of intuition would not restrain Hilbert's axiomatic programme.

There is another important consequence of this, namely that theories cannot be straightforwardly true or false through correctly representing some fixed subject matter, or failing to represent it. This is stated clearly in the 1922/1923 lectures, but it is already clear in the 1898/1899 lectures and the 1899 correspondence with Frege. Theories, variously interpretable, are either 'possible' or not, and what shows their 'possibility' is a demonstration of consistency, in which case the mathematical theory 'exists'. Thus, for Hilbert, the correct account of truth and falsity with respect to mathematical theories is that of consistency/inconsistency. Derivatively a mathematical *object* exists (relative to the theory) if an appropriate existence statement can be derived within the consistent theory. As Hilbert puts it in his 1919 lectures:

> What however is meant here by existence? If one looks more closely, one finds that when one speaks of existence, it is always meant with respect to a *definite system taken as given*, and indeed this system is different, according to the theory which we are dealing with. (Hilbert, 1919, p. 147, book p. 90).

Hilbert sums this up in his 1902 lectures on the foundations of geometry:

> We must now show the *freedom from contradiction of these axioms taken together*; ... .
> In order to *facilitate* the *understanding* of this, we begin with a remark:

> The things with which mathematics is concerned are defined through axioms, *brought into life*.
>
> The axioms can be taken quite arbitrarily. However, if these axioms contradict each other, then no logical consequences can be drawn from them; the system defined then does not exist for the mathematician. (Hilbert (*1902, p. 47) or Hallett and Majer (2004, p. 563))

This complex of views is a very important foundation for Hilbert's theory of what he calls 'ideal objects', where the concepts, freed from the constraints of the actual, are 'completed'. In his 1898/1899 lectures, Hilbert notes that while the Archimedean Axiom as sole continuity axiom shows that to every point, there corresponds a real number, the converse will not generally hold:

> That to every real number there corresponds a point of the straight line does not follow from our axioms. We can achieve this, however, by the *introduction of ideal (irrational) points (Cantor's Axiom)*. It can be shown that these ideal points satisfy all the axioms I–V; it is therefore a matter of indifference whether we introduce them first here or at an earlier place. The question whether these ideal points actually 'exist' is for the reason specified completely idle [*völlig müssig*]. As far as our knowledge of the spatial properties of things based on experience is concerned the irrational points are not necessary. Their use is purely a matter of method: *first with their help is it possible to develop analytic geometry to its fullest extent.* (Hilbert (*1899, pp. 166–167); Hallett and Majer (2004, p. 391))

Thus, while one might start from the idea that the points in geometry correspond to points whose existence in actual space can be shown, perhaps through construction, this idea is left behind; as Hilbert says, it is 'idle' to consider the question of whether the new points 'actually exist'.[18] It is precisely in this context that Hilbert first introduces his new notion of mathematical existence; thus, full Euclidean geometry 'exists' because we can furnish a model for the theory by using analysis. Put more abstractly, the procedure is this. We begin with a Euclidean geometrical system without continuity axioms, and where 'real' points are instantiated by geometrical constructions. These constructions correspond to various algebraic fields over the rationals (depending on what is permissible in construction). These number fields are seen to have a maximal extension in the reals; therefore, we postulate that there are points corresponding to *all* real numbers. These new objects are then 'ideal' with respect to the original, real points. The geometrical system corresponding

[18] This view, that we can decide to extend the field of objects without being constrained by what 'really exists', is very much in evidence in Dedekind's 1872 memoir on continuity and irrational numbers; see Dedekind (1872, 11). The most notable difference compared with Hilbert is in the idea that objects are 'created' to fill the 'gaps'; Hilbert's reliance on consistency avoids this notion of 'creation'.

to the completion is axiomatized; this new system is consistent, since it has a model in the reals. As Hilbert says:

> *Euclidean geometry exists, so long as we take over from arithmetic the proposition that the laws of the ordinary real numbers lead to no contradiction.* With this we have shown the existence of all those other geometries which we have considered in the course of this investigation. (Hilbert (*1899, p. 167); Hallett and Majer (2004, p. 391). See also Hilbert's own lecture notes, p. 104; in Hallett and Majer (2004, p. 280).)

From the point of view of this theory, any dispute then over the 'real' existence of some points as opposed to others is 'idle'; the *theory* exists, and it has the existence of all these points as consequences; the real/ideal distinction was always only ever relative to the original domain. As Hilbert puts it later:

> The terminology of ideal elements thus properly speaking only has its justification from the point of view of the system we start out from. In the new system we do not at all distinguish between actual and ideal elements. (Hilbert, 1919, p. 149, book p. 91)[19]

Furthermore, the core system of elementary geometry possesses many different and incompatible 'ideal extensions', some of them giving rise to the 'other geometries' Hilbert mentions. All 'exist' in so far as they are consistent.

In short, advanced 'pure' mathematical knowledge is now knowledge about various conceptual schemes, and a significant part of this is logical knowledge, in the first instance, knowledge of the existence or *non*-existence of derivations from certain starting points. It is the abandonment of the idea of fixed content and the shift to the 'conceptual sphere' which allows Hilbert to introduce a sophisticated, and *relative*, conception of the 'real/ideal' distinction, one which is of great importance for his later foundational work.

The arrival at this 'pure thought form' is, I think, the key to the correct explanation for the progression from intuition to concept to ideal stated at the beginning of Hilbert's *Grundlagen* in an epigram from Kant's *Kritik der reinen Verknunft*.[20] And it is as a 'pure thought form', thus through the projection into the 'conceptual sphere', that mathematics falls under the 'general law' on the nature of the mind which Hilbert stated in his lecture on mathematical problems, namely that all questions 'which it [the mind] asks must be answerable'.

Important in all this is the relationship *between* mathematical theories. For one thing, the approximate (not to say coarse) nature of interpretation in

[19] See also p. 153 of the 1919 lecture notes (Hilbert, 1919, book p. 94).

[20] The epigram is the following passage: 'So fängt denn alle menschliche Erkenntnis mit Anschauungen an, geht von da zu Begriffen, und endigt mit Ideen.' (See *Kritik der reinen Vernunft*, A702/B730. A very similar remark is also to be found at A298/B355.)

the observable world means that these interpretations are less than ideal as ways of demonstrating ‘possibility’, and the same holds for investigations of independence. Secondly, the process of ideal extension as described above appears to rely heavily on the notion of embedding one mathematical theory in another, and indeed Hilbert himself describes the procedure in just this way in his 1919 lectures:

> Precisely stated, the method [of ideal elements] consists in this. One takes a system which behaves in a complicated way with respect to certain questions which we want to answer, and one transforms to a new system in which these problems take on a simpler form, and which in addition has the property that it contains a sub-system isomorphic to the original system. ... The theorems concerning the objects of the original system are now special cases of theorems of the new system, in so far as one takes account of the conditions which characterise the sub-system in question. And the advantage of this is that the execution of the proofs takes place in the new system where everything becomes much clearer and easier to take in [*übersichtlich*]. (Hilbert, 1919, pp. 148–149, pp. 90–91 of the book)

This procedure can be brought out, too, by looking more closely at the general form of argument in the independence and relative consistency proofs, a form first articulated by Poincaré in one of his explanations of the proof of the independence of the Parallel Postulate.[21] The general procedure is this. Proving the independence of Q from $P_1, P_2, \ldots, P_n$ with respect to a given conceptual scheme $\mathcal{T}_1$, that is, showing that there cannot be a derivation of Q from $P_1, P_2, \ldots, P_n$, involves showing that, if there were, there would then also be a derivation of $\tau(Q)$ from $\tau(P_1), \tau(P_2), \ldots, \tau(P_n)$ with respect to another conceptual scheme $\mathcal{T}_2$, where: (*a*) we know that each of the $\tau(P_i)$ are *theorems* of $\mathcal{T}_2$; (*b*) we know that the conceptual scheme $\mathcal{T}_2$ can *already* derive $\neg\tau(Q)$; (*c*) we assume (or know) that the scheme $\mathcal{T}_2$ is not in fact inconsistent. Here, crucially, $\tau(P)$ is some map from formulas to formulas which preserves logical structure. (This point was first articulated in Frege (1906).) An obvious variant of this procedure can be used to prove the relative consistency of theory $\mathcal{T}_1$ with respect to $\mathcal{T}_2$. Hilbert puts it this way: After specifying a minimal Pythagorean (countable) field Ω of real numbers which determines an analytic model of the axioms for Euclidean geometry, he continues:

> We conclude from this that any contradiction in the consequences drawn from our axioms would also have to be recognizable in the domain Ω. (Hilbert (1899, p. 21); p. 455 in Hallett and Majer (2004))

Thus, crucial to the investigation are what we might call *base theories*, those through which the propositions of the ‘home theory’ are interpreted.

[21] See Poincaré (1891).

8.3.2 Non-elementary metamathematics

The second very important thing to recognize about Hilbert's foundational enterprise as developed in his study of geometry follows on from this. When looking for base theories, the natural thing is to choose those over which one has a fine degree of control, perhaps because they have only countably many elements which can be individually described.

I mentioned above that Hilbert tried to separate synthetic geometry from analytic geometry, but the separation is only at the mathematical level; at the metamathematical level, the link is retained. For one thing, as we have remarked, analytic geometry is taken as the measure of synthetic geometry (for example, via the demand for completeness). But more importantly here, real and complex analysis in its broadest sense is taken as the fundamental tool in the logical analysis of synthetic geometry. Since synthetic geometry has as one of its goals the aim of matching analytic geometry, then it should be clear that ordinary analytic geometry based on the complete ordered field of the real numbers provides a model for (disinterpreted) synthetic geometry. It is not a large leap from this to expect that one can find (or fabricate) substructures of the reals (or wider analysis) which will correspond to the variations of axioms and central propositions of synthetic geometry, not least because analytic geometry gives point-by-point and line-by-line control over the geometrical structure. For instance, in the standard arrangement, lines are given by simple linear functions of the number pairs giving the coordinates of points. But in principle, a vast range of other functions could be chosen, showing one kind of behaviour within a certain region, and a quite different behaviours outside that region. Indeed, this is the lesson taught by the various models of non-Euclidean geometry. A fundamental presupposition of Hilbert's investigation, therefore, is the presence of the full panoply of analytic techniques. Thus, while Hilbert's axiomatization of geometry distances itself from the analytic developments of the 19th century, the full range of analytic geometry is made available, not to prove results in the theory itself, but to prove results *about* the theory, and in particular to throw light upon the underlying source of knowledge. I will attempt to draw out the points made here in the examples given in Section 8.4.

8.3.3 Foundations

There is another sense in which Hilbert's project is radically different from other foundational projects at the time, projects with a Euclidean flavour, and this is that Hilbert does not automatically seek a more primitive conceptual level. Of course, this might be done for specific reasons in certain circumstances, and of course important mathematical information might be gained from doing so, whether one succeeds or fails. The point is that it is not the central goal of foundational investigation. This means that, to a large extent, Frege-style definitions can be dispensed with. They appear now rather as assignments for the purpose of modelling the principles being investigated. The prime example is perhaps Hilbert's use of the theory of real numbers (or rather a minimal Pythagorean field drawn from them) to model the axioms of elementary geometry by giving 'temporary' definitions of point, straight line, congruence, and so on; this is in place of the standard procedure behind analytic geometry, which effects a reduction of geometry to analysis by making these assignments as Frege-style definitions. Similarly, one could use the Dedekind-cut construction or the Cauchy-sequence construction to model the axioms for the theory of real numbers, rather than to define them, and likewise with the 'definitions' of the integers as equivalence classes of ordered pairs of natural numbers, complex numbers as ordered pairs of reals, and so on.[22]

There is a further important point to be made. Even if we seek a conceptual reduction, it is important to have available 'local' axioms, for instance for the natural numbers or for the reals. To take an example, in seeking to show in his 1898/1899 lectures (and later the *Grundlagen*) that line segments themselves exhibit a field structure like that of the real numbers, it was necessary for Hilbert to have a detailed axiomatization of field structure, in order to say that the fields are alike in certain respects but differ in others. This is the origin of Hilbert's axiomatisation of the real numbers in the lecture notes, published as axioms for what Hilbert calls 'complex number systems' in the *Grundlagen* of 1899, and then completed in (Hilbert, 1900*a*). The point is simple: to show that a conceptual 'reduction' (like Frege's) has worked, one has to be able to derive theorems which say that the defined objects (e.g. the numbers) have the right properties.[23] Thus, one has to have in effect an axiom system, and this is *overriding*.

Frege sought a reduction to fewer and fewer principles, as in a sense did Euclid; Hilbert's work on the other hand shows that, in some key cases if

[22] Hilbert's later paper 'Axiomatisches Denken' makes the parallel point that the uncovering of ever 'deeper' primitives is a constant theme in the development of mathematics and physics. The paper even seems to suggest that the use of the Axiomatic Method incorporates such a search; he uses the phrase '*Tieferlegung der Fundamente*'. The point to stress, though, is that there is never thought to be an ultimate conceptual layer, and the 'foundations' for a theory might be given, and productively so, in different, even incompatible, ways. As Boolos observes with respect to the Frege analysis of number and the attempt to reduce arithmetic to logic:

> Neither Frege nor Dedekind showed arithmetic to be part of logic. Nor did Russell. Nor did Zermelo or von Neumann. Nor did the author of *Tractatus* 6.02 or his follower Church. They merely shed light on it. (Boolos, 1990, 216–218).

[23] Frege seeks to establish just this in his (1884, §§78–83).

we have more axioms, then our logical analysis can be more refined. A nice example is provided by the maintenance of the Archimedean Axiom as an independent axiom while seeking an additional axiom which leads to point-completeness. A natural axiom would have been some version of the limit axiom which Hilbert had used in earlier axiomatizations (e.g. in his 1893/1894 lectures, in Hilbert (1895), and in his 1898 *Ferienkurs*), and which is even presented as a possibility in the 1898/1899 lectures. But this axiom would 'hide' the Archimedean Axiom because it would imply it. Instead, Hilbert retains the Archimedean Axiom and seeks an axiom which will complement it and boost its power to full completeness. The axiom he chooses is his own *'Vollständigkeitsaxiom'*.[24] Many of Hilbert's metamathematical investigations concern the precise role of the Archimedean Axiom in the proofs of central theorems, typically whether, when used, its use is necessary or not. One example is given by Dehn's work on the Legendre Theorems and the Angle-Sum Theorem (Dehn, 1900), another concerns the involvement of the axiom in the theories of plane area and three-dimensional volume. The axiom and the requisite analyses are also clearly important in the development of non-Archimedean geometries; and its presence *without* the Completeness Axiom is very important in the search for countable models. Yet another example is provided by Hilbert's own work on the Isoceles Triangle Theorem to be considered below in Section 8.4.2.

The point about the definitions, and this point about not reducing the conceptual level, are, of course, closely related. The creative purpose behind defining, as Frege recognized, is that with properly chosen definitions, unprovable propositions become provable, and therefore there is no need to take them (and attendant propositions) as axioms. Examples abound, but a striking one is furnished, of course, by Frege's own work. With the Frege definition of number, we can *prove* what is now known as *Hume's Principle* (*HP*),[25] the principle in effect from which Frege arithmetic is derived. Without such a proof, we would have to take the principle as a primitive axiom, and the terms '*NxFx*' (i.e. the numbers) as primitive terms, contrary to Frege's intention. However, it might be argued that the really revealing thing about Frege's work is not that, with the Frege definition, *HP* is provable, but rather that the central arithmetic principles can be proved from *HP* alone, given Frege's other definitions.

[24] See my *Introduction* to Chapter 5 of Hallett and Majer (2004, §5, 426–435).

[25] That is, the second-order proposition: '$\forall F, G[NxF = NxG \leftrightarrow F \approx G]$', where '$F$', '$G$' are variables for concepts, 'NxF' is a singular term standing for 'the number of things falling under the concept F', and where '$\approx$' stands for the (definable) relation of equinumerosity between concepts. See Frege (1884, §63).

So, to sum up the points in these last two sub-sections: for Hilbert, we do *not* generally seek 'more primitive' conceptual levels from which a theory can be finally deduced; moreover, often conceptually 'higher' mathematics is intrinsically necessary for the investigation of conceptual schemes, even elementary ones.

8.4 The 'purity of method' reconsidered

Let us now return to the way that 'purity of method' is dealt with in Hilbert's geometrical work. I will consider three examples, concerning respectively Desargues's Theorem, the Isoceles Triangle Theorem and the Three Chord Theorem, all of them extremely elementary geometrical results, and all three of which touch centrally on the intuitive 'facts' behind geometry. This is no accident. We have seen that for Hilbert the main source of knowledge behind traditional geometry is a mixture of intuition and empirical investigation (experiment), a mixture ultimately behind the successful axiomatization. But, starting with informal 'purity' questions, Hilbert's metamathematical analysis of the 'facts' uses higher mathematics, which in turn informs elementary geometrical knowledge. None of the examples treated is fully represented in the original 1899 version of the *Grundlagen*. The central result on Desargues's Theorem (Section 8.4.1) *is* in the *Grundlagen*, but what leads up to this result, namely, the philosophical reflection and analysis undertaken in the 1898/1899 notes, is suppressed; the analysis of the Three Chord Theorem (Section 8.4.3) is an important part of the 1898/1899 lectures, but only the abstract mathematical result, and not the analysis itself, appears in the *Grundlagen*; and the analysis of the Isoceles Triangle Theorem (Section 8.4.2) makes no appearance, being first dealt with in the 1902 lectures.

8.4.1 Desargues's Theorem

The first example I want to consider where purity of method and the analysis of intuition play a significant role is in Hilbert's treatment of Desargues's Theorem in elementary projective geometry. Suppose given two triangles ΔABC and $\Delta A'B'C'$, not in the same plane, which are so arranged that the lines AA', BB', CC' meet at a point. Desargues's Theorem then says that the three points of intersection generated by the three pairs of straight lines AB and $A'B'$, BC and $B'C'$, AC and $A'C'$ themselves lie on a straight line. Intuition might be said to play a role from the beginning, since it is very easy to 'see' the correctness of the theorem; the intersection points must all lie in the planes of both triangles, and these planes intersect in a straight line. Cast in Hilbert's system, it is easily proved using just the full (i.e. planar and spatial) incidence and order axioms (Groups I and II). However, the theorem has a restricted version, where the triangles in question both lie in the *same* plane.[26] This version is not so easily visualizable, neither is it especially easy to prove. More importantly, the standard proof goes via the *unrestricted* version, using a point outside the plane of the triangles to reconstruct a three-dimensional Desargues's arrangement whereby the three-dimensional version of the theorem can be applied. In short, this proof, too, calls on all the axioms of I and II, even though the theorem itself appears to involve only *planar* concepts, i.e. is concerned only with the intersection of lines in the same plane. An indication of the standard way of proving the planar version of Desargues's Theorem is given by Fig. 8.1. Intuition might seem to play a role here, too, since it relies on projection of the original triangles up from the plane they lie in; in short, one can 'see' that planar Desarguesian configurations are reflected in spatial Desarguesian configurations and conversely.

This situation is remarked upon by Hilbert in his 1898/1899 lectures as raising a purity problem. Hilbert writes:

> I have said that the *content* of Desargues's Theorem is important. For now however what's important is its *proof*, since we want to connect to this a very important *consideration*, or rather *line of enquiry*. The theorem is one of plane geometry; the proof nevertheless makes use of space. The question arises whether there is a proof which uses just the *linear and planar axioms*, thus I 1–2, II 1–5. Thus here

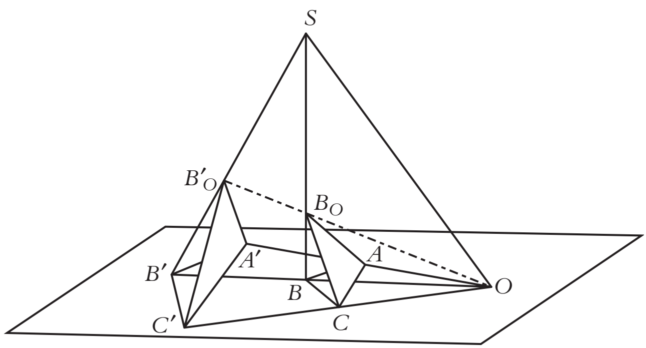

Fig. 8.1. Diagram for the usual proof of the Planar Desargues's Theorem, taken from: Hilbert and Cohn-Vossen (1932), p.108, Fig. 134. ΔABC and $\Delta A'B'C'$ are the two triangles in the same plane, and S is an auxilliary point chosen outside this plane.

[26] Both versions have fully equivalent converses, where it is *assumed* that the three points of intersection lie on a line, and it is then *shown* that the lines AA', BB', CC' meet in a point.

> for the first time we subject *the means of carrying out a proof* to a critical *analysis*. (Hilbert (*1898/1899, p. 30), p. 236 in Hallett and Majer (2004))

Then follow the general remarks about 'purity of method' which were quoted earlier (see p. 282). What this passage raises is the question of the *appropriateness* of a given proof, a 'purity' question *par excellence*, for here we have to do with a theorem, the Planar Desargues's Theorem, which is proved by means which at first sight might be thought inappropriate, owing to the proof's appeal to *spatial* axioms.

What we are faced with first is an independence question, and Hilbert sets out to show that the proof of the *Planar* Desargues's Theorem is *not* possible without appeal to spatial incidence axioms. As he puts it:

> We will rather show that Desargues's Theorem is *unprovable* by means of I 1–2, II 1–5. One will thus be *spared the trouble* of looking for a proof in the plane. For us, this is the first, simplest example for the proof of *unprovability*. Indeed: To satisfy us, it is necessary either *to find* a proof which operates just in the plane, or *to show that there is no such proof*. Prove so, that we specify a system of things = points and things = planes for which axioms I 1–2, II 1–5 hold, but the Desargues Theorem does not, i.e., that a plane geometry with the axioms I, II is possible without the Desargues Theorem. (Hilbert (*1898/1899, p. 30), pp. 236–237 in Hallett and Majer (2004))

To show this, Hilbert constructs an analytic plane of real numbers, with the closed interval $[0, +\infty]$ removed. The key thing is that some of the new straight lines are gerrymandered compositions: *below* the x-axis, the line is an ordinary straight (half-)line; *above* the x-axis, the line is in fact an arc of a circle uniquely determined by the conditions (*a*) that it goes 'through' the origin, O (it is open-ended here, since O is not itself in the model), and (*b*) that the given straight line below the x-axis is a tangent to it. Hilbert's figure (Fig. 8.2) illustrates the essentials of the model.[27]

The first thing to note about this is that we are here operating at one remove from the intuition ordinarily thought to underlie projective geometry. For one thing, the model produces highly unintuitive 'straight lines', even though they are pieced together using *intuitable* objects. The gerrymandered 'straight lines' are again 'intuitable' in the sense that one can easily visualize them (witness Fig. 8.2). Moreover, there is a sense in which lines made up of two distinct pieces are extremely familiar to us; think of a straight stick (a line segment) half immersed in still, clear water and then viewed from above looking down into the water. (Indeed, if it is viewed when the water is not still, then at any instant

[27] For full details, see Hilbert (*1898/1899, 31), or the *Ausarbeitung*, p. 28; these are pp. 237 and 316 respectively in Hallett and Majer (2004).

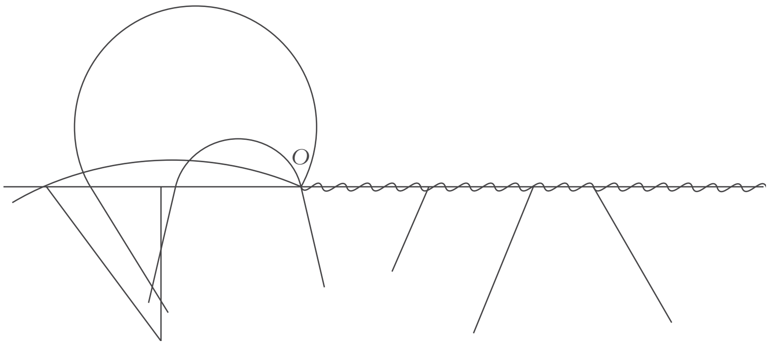

Fig. 8.2. Model for the failure of Desargues's Theorem; diagram taken from p.31 of Hilbert's 1898/1899 lecture notes on the foundations of Euclidean geometry.

the upper part will appear as an ordinary straight line, while the lower part will be a non-straight curved line. It might be said that what Hilbert has done here is to take a simple example like this and to choose a particularly simple curve.) However, while the direct perception ('intuition') of such objects is perhaps familiar, the analytic treatment is essential, and not merely a convenient way of proceeding. For one thing, the use of the algebraic manipulation is indispensable; it has to be shown that the model satisfies the plane axioms of I and II, and this is by no means trivial. For instance, in considering Axiom I1, it has to be shown that any two points determine a straight line in the new sense, including the case where one point lies above the x-axis (has positive y-coordinate), and one point lies below the x-axis (has negative y-coordinate); in other words, it has to be shown that there is always a circle passing through 0 and the upper point and which cuts the interval $[-\infty, 0)$ in a point below 0 such that the tangent to the circle at that point is a straight line which passes through the given point in the lower half-plane. This, however, is correct, as careful calculation shows.[28] It is hard to see how this could be accomplished without the use of calculation. The point is that the use of the analytic plane and the accompanying algebra gives Hilbert extremely fine control over the pieces and how to glue them together in the right way, even though the result (when transferred back to the intuitive level) is fairly easily understood visually.

Hilbert's model specifies the first example of a non-Desarguesian geometry. Hilbert's treatment of the Desargues's Theorem in the 1898/1899 notes is

[28] Hilbert does not explicitly adddress this case, and I am grateful to Helmut Karzel for bringing it to my attention. It is more fully explained in my n. 46 to the text of Hilbert (*1898/1899) in Hallett and Majer (2004, 237).

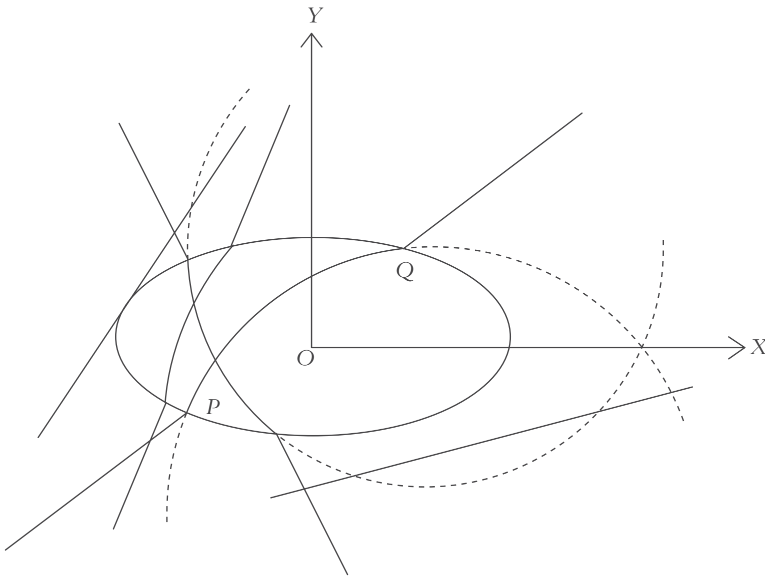

Fig. 8.3. Model for the failure of Desargues's Theorem used in the *Grundlagen der Geometrie*; diagram adapted from p.53.

fully projective, and takes place before the congruence axioms are even stated. In the *Grundlagen*, all the axioms are set out before there is any significant development. The statement of Desargues's Theorem given there involves parallels, and the proof from the full spatial axioms (and the Parallel Axiom) is called upon, though not given (p. 49). Hilbert then remarks that one can give a simple *planar* proof provided a central result from the theory of proportions is used. Since congruence is involved, this means in effect that congruence can replace the spatial assumptions involved in the usual proof. Hilbert then gives another planar model (see Fig. 8.3), thus creating a non-Desarguean affine geometry, in which Desargues's Theorem and the appropriate congruence assumption (the Triangle Congruence Axiom, IV 6 of the *Grundlagen*, III 6 of the lectures) both fail. The investigation is thus different, and certainly adds important information to that of the lectures.[29] But the same points as were made above hold. Hilbert again takes an ordinary analytic plane, but this time one with a certain 'distorting' ellipse around the origin. Here, straight lines which would normally pass through the origin are 'distorted' by the

[29] As against this, though, note that the whole fascinating discussion of the import of Desargues's Theorem, of '*Reinheit der Methode*', etc. is quite lacking in the first and subsequent editions of the *Grundlagen*.

elliptical 'lens' such that in the interior of the ellipse they describe arcs of circles, and the new 'straight lines' are these straight line/circle arc/straight line composites. (See Fig. 8.3 and then Hilbert's *Grundlagen*, pp. 52–54, in Hallett and Majer (2004, pp. 488–491).) Once again the idea has connections with perceptual experience; once more, we see a case where the result issues from interplay between geometrical intuition and abstract, non-intuitive calculation. But again note that it is this calculation which shows that the model works, for once again, precision is the key. We also have a case which is similar to the Isoceles Triangle Theorem result in Section 8.4.2, for what is shown here is that spatial assumptions can be avoided if (here) we adopt the full set of axioms governing plane congruence. It is also worth bearing in mind that highly refined mathematical/numerical models like these are what show for Hilbert that certain 'intuitive' geometries 'exist'.

One of the results of Hilbert's investigation is that one need not look any further for a purely plane proof (from I 1–2, II) of the Planar Desargues's Theorem; one can, as Hilbert puts it, 'sich die Mühe sparen'. But this is by no means the end of the matter. As Hilbert states, our 'drive' for mathematical knowledge is only satisfied when one can establish *why* matters cannot go in the initially expected way. (See the passage from the *Grundlagen* cited on p. 201.) Since Desargues's Theorem *can* be proved from the axiom groups I and II, which among other things put conditions on the 'orderly' incidence of lines and planes, one might say that the theorem is a *necesssary* condition for such incidence. Hilbert now conjectures the following:

> Is the Desargues Theorem also a *sufficient* condition for this? i.e. can a system of things (planes) be added in such a way that all Axioms I, II are satisfied, and the system before can be interpreted as a sub-system of the whole system? Then the Desargues Theorem would be the very condition which guarantees that the plane itself is distinguished in space, and we could say that everything which is provable in space is already provable in the plane from Desargues. (Hilbert (*1898/1899, p. 33), p. 240 in Hallett and Majer (2004))

Hilbert shows that this conjecture is indeed corrrect, and the result is achieved by profound investigations of the relationship between the geometrical situation and the analytic one, the overall result being a re-education of our geometrical intuition, for what it reveals is that the Planar Desargues's Theorem in effect actually has spatial content. This provides an explanation of why it is (in the absence of congruence and the Parallel Axiom) that the Planar Desargues's Theorem cannot be *proved* without the use of spatial assumptions, and it provides a beautiful example of Hilbert establishing in the fullest way possible *why* 'impure' elements are required in the proof of Desargues's Theorem, the grounds for the 'impossibility of success' in trying to prove Desargues's

Theorem in the plane. When Hilbert first articulated the conjecture in the passage just quoted, he clearly had no proof of it, the same being true at the time the corresponding place in the *Ausarbeitung* was composed.[30] However, towards the very end of the *Ausarbeitung* (pp. 159–160; Hallett and Majer (2004, p. 387)), Hilbert presents a proof showing the conjecture to be correct.[31] The situation is a little complicated, but worth sketching.

(*i*) Hilbert's first step is to set up a *segment calculus* using the full axioms I–II together with IV (i.e. the axioms of incidence, order, and parallels, following the enumeration of the lectures; in the *Grundlagen*, these axioms are I–III). Segments are in effect to be treated as magnitudes, though the relevant properties are not assumed, but have to be demonstrated: addition and multiplication are defined for segments, the basis being an arbitrarily chosen system of axes; the zero and unit segments are defined; negative segments and fractional segments are then also defined, as well as an ordering relation. The Planar Desargues's Theorem is then called on in a fairly natural way to show that these operations satisfy all the axioms for a non-commutative ordered field. Hilbert also shows that the usual equation for a straight line holds, where segment 'magnitudes' are taken as what the variables vary over.

(*ii*) The Planar Desargues's Theorem is available because of the presence of the spatial incidence axioms in I. Not only are these the *only* spatial assumptions, their sole purpose here is to guarantee the presence of the Planar Desargues's Theorem. If we drop these axioms, but adopt the Planar Desargues's Theorem instead, in effect as a new axiom, then we have apparently *just* planar axioms, but importantly the results about the segment field still go through.

(*iii*) Hilbert's third step is to point out that the segment magnitudes can be used in pairs to coordinatize the plane; this is done in a very natural way, given the axes which lie at the base of the definitions of addition and multiplication. Hilbert argues now that we can use *triples* of these magnitudes as the basis of a coordinatization of *space*. The space thus determined will satisfy all the original axioms I–II, IV (I–III in the enumeration of the *Grundlagen*), and will be such that the original plane (now characterized by coordinate pairs) will be a plane *in the new space*; moreover, the fact that the right incidence

[30] In the *Ausarbeitung* of the lectures notes, Hilbert introduces the conjecture with the words 'The question is … '. He then writes:

> This question is probably to be answered in the affirmative; one could then say in this case: The Desargues Theorem is the resultant of the elimination of the spatial axioms from I and II. (p. 32; Hallett and Majer, 2004, 318)

[31] For the evidence that Hilbert did not have the proof until late on in the course, see Hallett and Majer (2004, 189–190).

axioms hold for planes means that this plane will be 'properly' embedded in the space.

What this shows is that that the Planar Desargues's Theorem is a *sufficient* condition for the orderly incidence of lines and planes, in the sense that it can be used to *generate* a space. We thus have an explanation for why the Planar Desargues's Theorem cannot be proved from planar axioms alone: the Planar Desargues's Theorem appears to have spatial content. Moreover, as is clear in Hilbert's statement of his conjecture and from his construction, Desargues's Theorem could now be taken as an axiom to act *in place* of the spatial incidence axioms. The use of the theorem as an alternative axiom is quite standard in modern treatments of projective geometry, especially in treatments of affine planes; see, for example, Coxeter (1974) or Samuel (1988).

In this example, the investigation begins with a straightforward 'purity' question, which involves the analysis of an easy result underwritten, if not generated by, intuition; the analysis itself consists of an interplay between intuitive geometrical configurations and analytic geometry, though the precise results are obtained via the analytic models. The final analysis produces results which inform or educate (perhaps even challenge) our intuition.

8.4.2 The Isoceles Triangle Theorem

The Isoceles Triangle Theorem (*ITT*) of elementary geometry says that the base angles in an isoceles triangle are equal or, equivalently, that in a triangle whose base angles are equal, the sides opposite the equal angles are equal. The theorem figures early in Book I of Euclid's *Elements* as Proposition 5. Euclid's proof relies on the congruence of certain triangles using the side-angle-side criterion, the justification for this criterion being given in I, 4. The proof of I, 4 relies on a superposition argument, something which after this point Euclid seems keen to avoid. In Hilbert's axiomatization, the theorem is proved from incidence, order, and congruence axioms, the latter being designed to avoid physico-spatial assumptions, such as those involving rigid body movement or superposition. Hilbert's Triangle Congruence Axiom, bringing together assumptions about linear congruence and angle congruence, itself directly legitimizes the side-angle-side criterion for triangle congruence; the other usual criteria for triangle congruence are then given in the subsequent triangle congruence theorems, all of which rest ultimately on the Triangle Congruence Axiom. Hilbert's proof of the *ITT*, as given in his 1898/1899 lectures (the theorem is not mentioned in the *Grundlagen*) rests on a simple observation, namely that in the isoceles triangle ΔABC (where AB and BC are taken to be the sides which are equal) we can describe the triangle in two distinct ways, namely as ΔABC and ΔACB, and the triangles so described must be congruent by the side-angle-side criterion; consequently, the base angles must be the same. The proof appears to be, thus, little more than trivial.

Hilbert's proof can be traced to Pappus; Euclid's is different and a bit more involved. In his edition of the *Elements*, Heath notes of Pappus's proof:

> This will no doubt be recognised as the foundation of the alternative proof frequently given by modern editors [of the *Elements*], though they do not refer to Pappus. But they state the proof in a different form, the common method being to suppose the triangle to be taken up, turned over, and placed again upon *itself*, after which the same considerations of congruence as those used by Euclid in I, 4 are used over again. (Heath (1925, Volume 1, p. 254))

But Heath points out that Pappus himself avoids the assumptions about 'lifting' and 'turning', simply describing the same triangle in two ways. Hilbert thus follows Pappus, not those 'modern editors' Heath mentions.

Heath notes, though, that even Pappus's proof depends indirectly on a superposition argument, since it rests on Euclid's I, 4. Hilbert's proof does not, simply because the side-angle-side congruence criterion depends on the Triangle Congruence Axiom, not on any special proof method. It seems that matters have been shifted to the 'conceptual sphere', and no longer directly concern intuitions of space and movement in space.

Nevertheless, a 'purity of method' question can still be posed at the intuitive level: is 'flipping' and turning of the triangles, and thus a spatial dimension, *essential* to the argument for *ITT*? In some work first presented in his 1902 lecture course on the foundations of geometry, and then more fully in a paper published later in that year, Hilbert transposed this question at the intuitive level to a correlative question at the level of his axiomatization.[32] At the centre of concern is the Triangle Congruence Axiom. In its usual form, this involves no reference to what intuitively might be called the 'orientation' of the triangles involved; two triangles with matching side-angle-side combinations will be determined as congruent regardless of whether their orientation is the same or not. To mix idioms, accepting the quasi-physical interpretation of the concepts and axioms, and with the notion of congruence formulated

[32] The lecture notes are Hilbert (*1902), in Chapter 6 of Hallett and Majer (2004); Hilbert's paper is Hilbert (1902/1903). This paper was reprinted as Anhang II to the editions of the *Grundlagen der Geometrie* from the Second on. An Appendix (mentioned below) was made to the first reprinting in Hilbert (1903), and then a note (also mentioned below) was added to the Sixth Edition published in 1923, i.e. Hilbert (1923). The Appendix was radically revised by Arnold Schmidt for the Seventh Edition (Hilbert 1930) of the *Grundlagen*, as Hilbert makes clear in the Preface, this being the last edition of the monograph to be published in Hilbert's lifetime. We will actually concentrate here on Hilbert's presentation in his 1905 lectures on the 'Logische Principien der Mathematik', since of all the versions, this is the one where Hilbert is most philosophically expansive about the results and what they show. These lectures will appear in Ewald, Hallett and Sieg (Forthcoming).

with physical manipulation and superposition in view, triangles of different orientation can only be shown congruent by superposition after *lifting* and *turning* them in space. Thus, under such a physical interpretation, Hilbert's axiom licenses 'flipping' as a kind of displacement, and not just 'sliding' within the same plane. Seen in this way, the proofs of the *ITT* all exploit such flipping, either directly or indirectly, even, it might be said, Hilbert's in his 'neutral' system. The original 'purity of method' question, namely, 'Is flipping essential to the proof?', now has a direct correlate with respect to Hilbert's system: Is the Triangle Congruence Axiom (which can be seen to license 'flipping') essential to the proof of *ITT* in the axiomatic system, divorced as it is from the intuitive/empirical perspective? In order to pose the question more precisely, Hilbert considers a *weakened* version of the Triangle Congruence Axiom, one which insists in effect that, before the side-angle-side criterion can be used to license triangle congruence, the triangles be of the same orientation in the plane. This weakened axiom no longer underwrites the usual proofs of the *ITT*, as a quick look at the Pappus/Hilbert argument indicates; can the *ITT* nevertheless be proved? As Hilbert puts it in his presentation of the work in his 1905 lectures, the original 'purity of method' question is now transformed into a question primarily about logical relations:

> There now arises the question of whether or not the [original] broader version of the [Triangle Congruence] Axiom contains a superfluous part, whether or not it can be replaced by the restricted version, i.e. whether it is a logical consequence of the restricted version. This investigation comes to the same thing as showing whether or not the *equality of the base angles in an isoceles triangle* is provable on the basis of the restricted version of the congruence axiom. The question has a close connection with that of the validity of the theorem that the sum of two sides of a triangle is always greater than the third. (Hilbert, *1905, pp. 86–87)

This question is the beginning of Hilbert's investigations, and the consequences are fascinating, both for the abstract mathematical structure, for what the investigations tell us about 'geometrical intuition', and because of the connection Hilbert draws with the property of triangles he states, which we will refer to here as the Euclidean Triangle Property.

Hilbert first shows that the restricted congruence axiom together with the *ITT* itself implies the normal Triangle Congruence Axiom (see Hilbert, *1902, p. 32).[33] The key question, in effect first raised in Hilbert (*1902), is then

[33] In a note added to Appendix in the Sixth Edition of the *Grundlagen*, Hilbert points out that this is in fact only the case if one adds a further congruence axiom guaranteeing the commutativity of angle addition. See Hilbert (1923, 259–262); a suitable further axiom due to Zabel is stated on p. 259. This note does not appear in subsequent editions, but Schmidt's revised version of Appendix II adopts a similar additional congruence axiom, a weaker one attributed to Bernays. See Hilbert (1930, 134).

this: What has to be added to the usual system, with both the Parallel Axiom and the Archimedean Axiom, but only the restricted Triangle Congruence Axiom, to enable the Isoceles Triangle Theorem (or indeed the usual Triangle Congruence Axiom) to be proved?

Hilbert's first answer to this question is given in the 1902 lectures, and then in the 1902/1903 paper, namely: one *can* prove the *ITT* if one adopts alongside the Archimedean Axiom a second continuity axiom which Hilbert calls the *Axiom der Nachbarschaft*. This latter axiom (see Hilbert, *1902, p. 84) states that, given any segment *AB*, there exists a triangle ('oder Quadrat etc.') in whose interior there is no segment congruent to *AB*. Hilbert later gave other answers, focusing, in place of *Nachbarschaft*, on an axiom he calls the *Axiom der Einlagerung* (Axiom of Embedding), which says that if one polygon is embedded in another (i.e. its boundary contains interior but no exterior points of the first), then it is not possible to split the two polygons into the same number of pairwise congruent triangles. (Bernays later gave an essential simplification.)[34]

The bulk of Hilbert's 'purity' investigation is now devoted to showing that this axiom (or respectively *Einlagerung*) *and* the Archimedean Axiom are both essential if the *ITT* (respectively the broader version of the Triangle Congruence Axiom) is to be proved in this way. What Hilbert shows is that geometries can be constructed in which all the plane axioms hold (with the weaker version of triangle congruence), and where *Nachbarschaft* (respectively *Einlagerung*) holds, but where the Archimedean Axiom and the *ITT* both fail, or similarly where the Archimedean Axiom holds, but where *Nachbarschaft* (respectively *Einlagerung*) and the *ITT* fail.

In his 1902 lectures, Hilbert is very careful to set out some of the important things which can be reconstructed on the basis of the weaker congruence axiom[35], and among them is Hilbert's equivalent of Euclid's theory of linear proportion. On the other hand, Hilbert's full theory of triangular area ('surface content'), a conscious reconstruction of Euclid's, does not apparently go through. Euclid's fundamental theorem concerning this (*Elements*, I, 39) is that triangles on the same base and with the same area must have the same height. In establishing his version of this theorem (that two triangles which are *inhaltsgleich* and on the same base have the same height), Hilbert defines the

[34] This new 'very intuitive' requirement was first introduced by Hilbert in a section added to the first reprinting of his paper in the Second Edition of the *Grundlagen*, i.e. Hilbert (1903, 88–107). Schmidt's revision of Appendix II for the Seventh Edition (1930) omitted the consideration of the *Einlagerungsaxiom*. It was revived by Bernays in Supplements to later editions, where he points out the simplification.

[35] See §A of the 1902 lectures, which occupies pp. 26–32 of Hilbert (*1902), to be found in Hallett and Majer (2004, 553–556).

notion of the surface measure, or *Flächenmass*, of a triangle as the product of half the base and the height.[36] But in order for this definition to be a good one, it has to be shown that this quantity is independent of the choice of which side is to be the base. However, this apparently simple fact depends on the recognition that two right-angled triangles constructed within the given triangle are similar, and thus it depends on the theory of triangle congruence. But the triangles in question are not in the same orientation, and consequently it seems that the restricted version of the Triangle Congruence Axiom is not sufficiently strong for the purpose at hand, and that the demonstration required must ultimately depend on the unrestricted Triangle Congruence Axiom.

The construction of the models showing the various independence results is not so clear in the 1902 notes, but is given much more explicitly in the 1902/1903 paper, which was already clearly in preparation while the 1902 lectures were being held, and then in the 1905 lectures. We will also concentrate on the first part of Hilbert's independence investigation, where the Archimedean Axiom fails, but *Nachbarschaft* holds, for this is very revealing of Hilbert's approach to foundational questions.

To show the necessity of the Archimedean Axiom, Hilbert constructs a model in which all the axioms up to the Congruence Axioms (with, of course, the restricted Triangle Congruence Axiom), hold, *Nachbarshaft* holds, but the Archimedean Axom fails. In this model: (*i*) the Isoceles Triangle Theorem fails; (*ii*) the Euclidean triangle property fails; and (*iii*) the theory of triangle area fails to go through. The basis of this model is a non-Archimedean field, which he calls T, using as elements the power series expansions

$$a_0 t^n + a_1 t^{n+1} + \cdots$$

around a certain fixed parameter t, with real number coefficients, a_0 non-zero, and n (which can be zero) a rational number; arithmetic calculation using these expansions takes place in the standard way. The ordering on the elements of T is defined as follows: for any elements α, β in the field, $\alpha > 0$ if $a_0 > 0$, a_0 being the first coefficient of the expansion of α, and < 0 if $a_0 < 0$; $\alpha > \beta$ if $\alpha - \beta > 0$. These stipulations are enough to show that t is infinitesimal, since $1 - mt > 0$ whatever natural number m is; any α whose expansion begins with t or indeed any t^n is infinitesimal. If $\alpha = a_0 t^n + a_1 t^{n+1} + \cdots$ is such an infinitesimal element, then e^α itself will be in the field, and will be representable by a power series expansion whose first element is $1 \cdot t^0$.

[36] See the 1899 *Grundlagen*, p. 43, or the *Ausarbeitung* of the 1898/1899 lectures, p. 131, or pp. 372–373 and 478–479 respectively in Hallett and Majer (2004).

Suppose now that α and β are in the field; we can then consider the complex number $\alpha + i\beta$, and these elements give rise to the complex extension of the field. In fact, such an element will be represented by an appropriate power series built on t where the coefficients are imaginaries. If $\alpha + i\beta$ is in the complex extension, then so will $e^{\alpha+i\beta}$ be.

Hilbert now defines an analytic geometry as follows. Points are defined by pairs of coordinates (x, y), where x and y are elements of T. Straight lines will be given by linear equations in the usual way, and so the usual incidence and order axioms will then hold, as well as the Parallel Axiom; the Archimedean Axiom, of course, will not hold. This leaves just the congruence axioms to consider, and, as one might expect, this is where the art comes. First, two segments are said to be congruent if parallel transport can shift the line segments so that they start at the coordinate origin, and then (keeping one of the segments fixed) one of them can be rotated into the other by a positive rotation through some angle θ. The two 'movements', parallel transport and rotation, are of course, properly speaking, transformations of the plane onto itself. What will be congruent to what will then depend on the definition of the transformation functions; parallel transport is trivial, so the key matter is the transformation function corresponding to rotation.

This function is based on the following trick. The coordinate (x, y) can be coded as a single complex number $\alpha = x + iy$ in the complex extension of T. The function e^{α} can then be used to form another complex number, which can then be decoded to form a new coordinate (x', y'). Suppose (x, y) is the coordinate of one end of a segment whose other end-point (perhaps after parallel transport) is at $(0, 0)$. Suppose the segment is positively rotated through an angle $\theta + \tau$, where θ is real (and could be 0) and τ is infinitesimal. Then the new coordinates are given by the formula

$$x' + iy' = e^{i\theta+(1+i)\tau} \cdot (x + iy)$$

Clearly $(0, 0)$ transforms into the coordinate $(0, 0)$.[37]

Segments can be assigned length acccording to the following procedure. A segment on the x-axis is said to have length l when one of its end-points lies at the origin and the other has x-coordinate $\pm l$. Any other segment has length l if one end can be shifted to the origin by parallel transport and the other end can be rotated into $(\pm l, 0)$ by the rotation function or where $(\pm l, 0)$ can be

[37] As Hilbert makes clear, it can be shown that given any line segment and any point in the plane, a rotation can be found so that the rotated line passes through that point.

rotated into it. Parallel transport does not alter length, but crucially rotation sometimes does.[38]

Hilbert now considers the point (1, 0) on the x-axis, and rotates this positively through the infinitesimal angle t. This will give a segment OA, where O is the origin and A has coordinates (x', y') determined by

$$x' + \mathrm{i}y' = \mathrm{e}^{(1+\mathrm{i})t} \cdot (x + \mathrm{i}y) = \mathrm{e}^{(1+\mathrm{i})t} \cdot x = \mathrm{e}^{(1+\mathrm{i})t}$$

since $y = 0$ and $x = 1$. This is then:

$$\mathrm{e}^{t+\mathrm{i}t} = \mathrm{e}^t \cdot \mathrm{e}^{\mathrm{i}t} = \mathrm{e}^t \cdot (\cos t + \mathrm{i} \sin t) = \mathrm{e}^t \cos t + \mathrm{i}\mathrm{e}^t \sin t$$

giving the coordinates

$$x' = \mathrm{e}^t \cos t, \quad y' = \mathrm{e}^t \sin t$$

According to the method of determining length, segment OA necessarily has length 1.

Now we construct the reflection of A in the x-axis, giving A' with coordinates $(x', -y')$, and A and A' are then connected by a line perpendicular to the x-axis; denote by B the point where this line cuts the axis (Fig. 8.4).

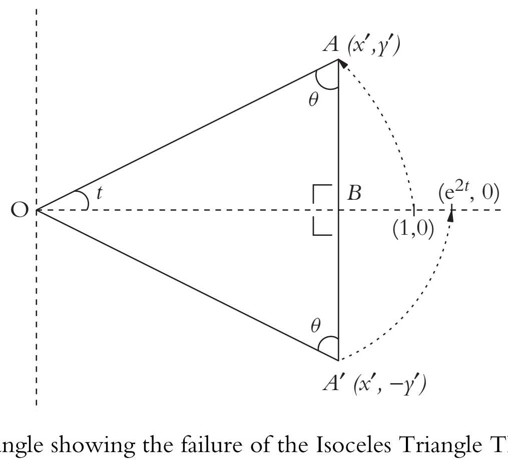

Fig. 8.4. The triangle showing the failure of the Isoceles Triangle Theorem.

[38] A key part of the proof is showing, of course, that the Congruence Axioms hold. Hilbert's work is not explicit about this. In his new version of Appendix II to the Seventh Edition of the *Grundlagen*, Schmidt uses one 'congruence mapping' in place of the distinct notions of parallel transport and rotation. This is essentially the rotation function described above, with 'transport' parameters added, i.e. $x' + \mathrm{i}y' = \mathrm{e}^{\mathrm{i}\theta+(1+\mathrm{i})\tau} \cdot (x + \mathrm{i}y) + \lambda + i\mu$, where λ, μ are taken to be elements of T. Schmidt then proceeds to give careful verifications of the Congruence Axioms.

Clearly, $AB = BA'$ and $OB = OB$; both angles $\angle OBA$ and $\angle OBA'$ are $\frac{\pi}{2}$. The question is now to compare OA and OA'. What Hilbert does is to rotate OA' positively through the angle t. This results in new coordinates (x'', y'') for the end-point, where

$$\begin{aligned} x'' + \mathrm{i}y'' &= \mathrm{e}^{(1+\mathrm{i})t} \cdot (\mathrm{e}^t \cos t - \mathrm{i}\mathrm{e}^t \sin t) \\ &= \mathrm{e}^{(1+\mathrm{i})t} \cdot \mathrm{e}^t(\cos t - \mathrm{i} \sin t) \\ &= \mathrm{e}^{(1+\mathrm{i})t+t} \cdot (\cos t - \mathrm{i} \sin t) \\ &= \mathrm{e}^{(1+\mathrm{i})t+t-\mathrm{i}t} \\ &= \mathrm{e}^{2t} \end{aligned}$$

Therefore, $x'' = \mathrm{e}^{2t}$, and $y'' = 0$. (See Fig. 8.4.) Thus, the rotation takes the point A' to the x-axis, namely to the point $(\mathrm{e}^{2t}, 0)$. But this point cannot coincide with $(1, 0)$, since $\mathrm{e}^{2t} > 1$. (This is because the power series expansion of e^{2t} is $1.t^0 + 2t + \frac{4t^2}{2!} + \cdots$; since the expansion of 1 is $1.t^0 + 0 + 0 + \cdots$, the series for $\mathrm{e}^{2t} - 1$ is $2t + \frac{4t^2}{2!} + \cdots$. The first coefficient of this is 2, hence $\mathrm{e}^{2t} - 1 > 0$, and so $\mathrm{e}^{2t} > 1$.) The length of OA' is therefore greater than 1, and so different from that of OA. The base angles in the extended triangle $\Delta OAA'$ are the same; the Angle-Sum Theorem holds, since the Parallel Axiom does. In consequence, we have an example of a triangle for which the base angles are equal, but in which the sides opposite these angles are not, violating the *ITT* (in its converse version).

Simple calculation in Hilbert's model also shows that $OB + BA' < OA'$, violating the Euclidean Triangle Property. As for the theory of triangle area, Hilbert shows that the Pythagorean Theorem for right-angled triangles holds in the form 'The sum of the square surface areas over the two legs of the right angle equals the square over the hypothenuse', since this depends only on the weaker Triangle Congruence Axiom. But the squares over BA' and BA are the same; since the square over OB equals itself, those over OA and OA' must be the same, too. But $OA < OA'$; hence the usual analytic conclusion from the Pythagorean Theorem ('The length of the hypothenuse is the square root of the sum of the squares of the lengths on the other two sides') fails, because we will have here sums which are the same, but squared hypothenuse lengths which are different.[39] For this reason, the geometry constructed is

[39] As Hilbert says in Hilbert (*1902, 125a): '... one can *no longer conclude that the sides are equal from the fact that the squares are*'. See Hallett and Majer (2004, 597).

called by Hilbert *non-Pythagorean geometry*. Moreover, since $OA < OA'$, the square over OA will fit inside the square over OA' with a non-zero area left over, so we have a square which is of equal area-content to a proper part of itself, violating one of the central pillars of the Euclidean/Hilbertian theory of triangle area-content.[40]

It is important to see that these are not just technical results, for they give us a great deal of information about what axioms are necessary for the reconstruction of classical Euclidean geometry. Hilbert sums it up in his 1905 lectures as follows:

> The result is particularly interesting again because of the way continuity is involved. In short:
>
> 1.) The theorem about the isoceles triangle and with it the Congruence Axiom in the broader sense is *not provable* from the Congruence Axiom in the narrower sense when taken with the other plane axioms I–IV [thus excluding continuity assumptions].
>
> 2.) Nevertheless, it becomes provable when one adds continuity assumptions, in particular the Archimedean Axiom.
>
> From this we see therefore that in the Euclidean system properly construed, and which allows us to dispense with continuity assumptions, the broader Congruence Axiom is a necessary component. The investigations which I have here set in train throw new light on the connections between the theorem on the isoceles triangle and many other propositions of elementary geometry, and give rise to many interesting observations [*Bemerkungen*]. Only the axiomatic method could lead to such things. (Hilbert, *1905, pp. 86–87)

This last remark is amplified by Hilbert in the direct continuation of this passage, which also ties the nature of the investigation directly to the view of geometry which we have seen emerging, namely *emancipation* from interpretation and intuition without losing contact with what underlies the axioms:

> When one enquires as to the status within the whole system of an old familiar theorem like that of the equality of the base angles in a triangle, then naturally one must liberate oneself completely from intuition and the origin of the theorem, and apply only logically arrived at conclusions from the axioms being assumed. In order to be certain of this, the proposal has often been made to avoid the usual names for things, because they can lead one astray through the numerous associations with the facts of intuition. Thus, it was proposed to introduce into the axiom system new names for point, straight line and plane etc., names which will recall only what has been set down in the axioms. It has even been proposed

[40] Hilbert states the conclusion as follows: '*The theory of surface content depends essentially on the theorem concerning the base angles of an isoceles triangle; it is thus not a consequence of the theory of proportions on its own*' (Hilbert (*1902, 125*a*), or Hallett and Majer (2004, 597)).

> that words like equal, greater, smaller be replaced by arbitrary word formations, like *a*-ish, *b*-ish, *a*-ing, *b*-ing. That is indeed a good pedagogical means for showing that an axiom system only concerns itself with the properties laid down in the axioms and with nothing else. However, from a practical point of view this procedure is not advantageous, and also not even really justified. In fact, one should always be guided by intuition when laying things down axiomatically, and one always has intuition before oneself as a goal [*Zielpunkt*]. Therefore, it is no defect if the names always recall, and even make it easier to recall, the content of the axioms, the more so as one can avoid very easily any involvement of intuition in the logical investigations, at least with some care and practice. (Hilbert, *1905, pp. 87–88)

The emancipation from intuition and the 'origin' of the theorem concerned is clearly indicated in this passage, but Hilbert's remark that 'one always has intuition before oneself as a goal' is also very important. Although the *achievement* of the results necessarily requires the deliberate suspension of intuition, crucially they yield important information *about* intuition. One thing which Hilbert stresses in the passage quoted on p. 237 is that adopting the full Triangle Congruence Axiom allows us to avoid any continuity assumption in demonstrating the *ITT*. That is certainly correct, and this fact belongs alongside others concerning congruence and continuity, a prime example being Hilbert's reconstruction of the Euclidean theories of proportion and surface area without invoking continuity. But there are other subtle conclusions to be drawn. For example, as explained above, the (apparently planar) full Triangle Congruence Axiom appears to contain some hidden spatial assumption, since it licenses 'flipping' arguments. But now Hilbert's independence results seem to show that we *can* get the result without the spatial assumption, and thus with genuinely planar congruence axioms, *if* we accept some modest continuity assumptions about the plane. So one has a choice between spatial assumptions in the planar part of geometry, or continuity assumptions. However, what makes this a rather more complicated matter is that the main continuity assumption involved, the Archimedean Axiom, is quasi-numerical, something one might think should be avoided as far as possible in a purely 'geometrical' axiomatization.

But whatever the right conclusion to be drawn about our geometric intuitions, two very important things seem to follow from Hilbert's analysis. On the one hand, the results seem to strengthen Hilbert's holism about geometry, at least in the sense that the axioms are very intricately involved with each other, and that there might be more than one way to achieve many important results. Secondly, the adumbration of the intuitive picture here, and perhaps its correction or adjustment, follows from the high-level logico-mathematico investigation which Hilbert engages in. The information is obtained only by using very sophisticated mathematical analysis: complex analysis over a non-Archimedean field. In other words, the higher mathematical/logical analysis based on complex numerical structures instructs and informs lower-level geometrical intuition. This, note, is just such a case where 'function theory' is used to show something about elementary geometry. (See the quotation from Hilbert on p. 201 above.)

In short, this example illustrates Hilbert's view that one is *guided* by geometrical intuition, one asks questions *suggested* by intuition, but in the end it is higher mathematics which *instructs* intuition, not the other way around. Thus, while Hilbert does carry out a kind of 'purity of method' investigation, it is much more focused, as he puts it, on the 'analysis of intuition'. One of the reasons why Hilbert thinks that intuition *requires* analysis is that it is not, for him, a *certain* source of geometrical knowledge, and certainly not a *final* source. Thus the analysis, which is designed to throw light on the question: what is one committed to exactly when one adopts certain principles, among them principles suggested by intuition?

8.4.3 The Three Chord Theorem

The third example considered here concerns another fairly elementary geometrical theorem, which says that the three chords generated by three mutually intersecting circles (lying in the same plane) always meet at a common point. Call this the Three Chord Theorem (*TCT*). (See Fig. 8.5.)

The theorem is not an ancient one, but was apparently first discovered by Monge in the middle of the 19th century. It has an interesting generalization, much studied by 19th century geometers, concerning the lines of *equal power*

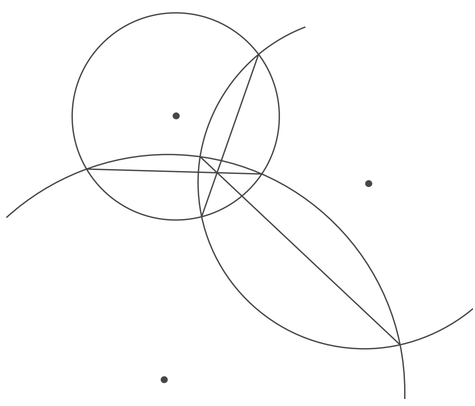

Fig. 8.5. Diagram of the Three Chord Theorem, adapted from the *Ausarbeitung* of Hilbert's 1898/1899 lectures, p.61, (Hallet and Majer (2004, 335)).

between two circles. The power of a point O with respect to a circle $\mathcal{C}$ in the same plane can be defined thus: Consider any straight line OPQ through O which cuts $\mathcal{C}$ in the two (not necessarily distinct) points P and Q. The power of O with respect to $\mathcal{C}$ is now the product $OP \cdot OQ$. This is a constant for O and $\mathcal{C}$, since this ratio is the same wherever P (respectively Q) happens to lie on the circle. (See Euclid's *Elements*, Book III, Proposition 35.) If O is inside the circle, the power is negative; if it is outside, and P and Q coincide, then OPQ is a tangent, and the power is OP^2. Given two circles, we can consider the points which have the same power with respect to both circles; these lie on a straight line perpendicular to the line joining the centres of the circles, a line which is called the *radical axis* or the *line of equal power* (in German, a '*Chordal*') between the two circles. (See Coolidge, 1916, Theorem 167.) What was often called the *Hauptsatz* of the theory of these lines is: Given three circles in a plane, intersecting or not, the three lines of equal power that the three pairs of circles generate must intersect at a single point. (*Ibid.*, Theorem 168.) Clearly, the TCT is just the case where each pair of circles intersect each other. There is another interesting special case. Consider a triangle, and consider three circles in the plane of the triangle centred respectively on the three vertices. As the radii of the three circles tend to zero, the three lines of equal power will tend to the perpendicular bisectors of the three bases. It follows that these three bisectors have a common point of intersection.

Hilbert gives a direct proof of the TCT in his 1898/1899 notes. The proof, which is clearer in the *Ausarbeitung*, pp. 61–64 (Hallett and Majer, 2004, pp. 335–337), assumes that the three circles in question, all in the plane α, are the equators of three spheres which have just two points P and Q of mutual intersection; it is then shown fairly easily (*loc. cit.*, p. 61) that the three chords of the circles, which lie, like those circles, in α, must all intersect the line PQ. But since neither P nor Q lies in α, PQ has only one point in common with α, and this point is therefore common to all three chords. Thus the simple proof depends on the assumption that three mutually intersecting spheres intersect in exactly two points, in other words, that these points *exist*. Hilbert's focus is then on this assumption, no longer on the TCT itself.

Hilbert's next step is to connect this with the theorem that a triangle can be constructed from any three line segments which are such that any two of the segments taken together are greater than the third. Call this the Triangle Inequality Property. In Euclid, this is proved in I, 22 (see Heath, 1925, Volume 1, pp. 292–293); call this the Triangle Inequality Theorem. The question Hilbert asks is: On what assumption is the proof of *this* proposition based? In his lecture notes (p. 64), Hilbert states this triangle property in the same breath with another, namely that, given a straight line and a circle in the same plane, if the line has both a point in the interior of the circle and a point outside it, then it must intersect it, and in exactly two points.[41] Call this the *line-circle property*. Similar to this is what can be called the *circle-circle property*, namely a circle with points both inside and outside another circle has exactly two points of intersection with it.[42] The connection to the Triangle Inequality Theorem is not surprising, for the circle-circle property is precisely what Euclid implicitly relies on in the proof of I, 22.

It is important to see Hilbert's question as one concerning 'purity of method'. It is well known that similar questions arise with the very first proposition of Euclid's *Elements* (I, 1) which shows how to construct an equilateral triangle on a given base *AB*. Euclid's construction takes the two circles whose centres are the endpoints of the base and whose radii are equal to the base; either of the two points of intersection of the circles can be taken as the third apex, *C*. But *do* the circles actually intersect? A standard objection is that there is no guarantee of this. Heath notes:

> It is a commonplace that Euclid has no right to assume, without premissing some postulate, that the two circles *will* meet in a point *C*. To supply what is wanted we must invoke the Principle of Continuity. (Heath, 1925, Volume 1, p. 242)

And by a 'Principle of Continuity', Heath means something like Dedekind continuity (*op. cit.*, pp. 237–238). Heath also cites one of Hilbert's contemporaries, Killing, as invoking continuity to show the line-circle property (Killing, 1898, p. 43). And many commentators on Euclid (for example, Simson in the 1700s) raised this point with respect to the proof of the Triangle Inequality Property, while at the same time stating that it is 'obvious' that the circles intersect, and that Euclid was right not to make any explicit assumption. What Hilbert investigates is what formal property of space corresponds to the implicit underlying existence/construction assumption.

Hilbert constructs a model of his geometry (i.e. Axiom Groups I–III) in which the existence assumption *fails*, and where the Triangle Inequality Theorem (Euclid's I, 22) also fails. Thus, the necessary conditions for Hilbert's proof of the Three Chord Theorem are not present in a geometry based solely on I–III. Moreover, since the Euclidean proof of I, 22 is based on a simple straightedge and compass construction, Hilbert's result is tantamount to saying that his axiom system does not have enough existential 'weight' to match this particular construction. This result is interesting, because, beginning with an

[41] Note that this has to be assumed for the power of point to be defined for all points and all circles.

[42] For similar properties, see pp. 62, 65 of Hilbert's lecture notes, and pp. 63–64 of the *Ausarbeitung*, i.e. Hallett and Majer (2004, 335–337).

intuitively inspired 'purity of method' question, it issues in a result on the abstract, conceptual level; not just this, but the result also shows why the assumption behind Euclid's I, 1 *is* justified in an elementary way.

Hilbert begins to consider this metamathematical problem on p. 65 of his lecture notes, pp. 64ff. of the *Ausarbeitung*. He constructs in effect the smallest Pythagorean sub-field of the reals which contains 1 and π, which yields a countable model of the axioms I–III (indeed of I–V, the whole system of the 1898/1899 lectures and the first edition of the *Grundlagen*) when the usual analytic geometry is constructed from pairs of its elements taken as coordinates. However, the number $\sqrt{1-(\pi/4)^2}$ does not exist in this field.[43] Since 1, 1, and $\frac{\pi}{2}$ are in the field, the model will possess three lines which satisfy the triangle inequality, but from which no triangle can be constructed. This is well illustrated by the diagrams in the *Ausarbeitung*, pp. 66 and 68, here Figs. 8.6 and 8.7.

One can see from Fig. 8.7 that the upper apex of the triangle depicted in Fig. 8.6 ought to have coordinates $(\pi/4, \sqrt{1-(\pi/4)^2})$; but this is not a point in the model. The same example shows that there can be a line partially within and partially without a circle but which intersects the circle nowhere: take the vertical line $x = \frac{\pi}{4}$ in the diagrams; this ought to meet the circle $x^2 + y^2 = 1$ at points with y-coordinates $\pm\sqrt{1-(\pi/4)^2}$, but again these points are missing.

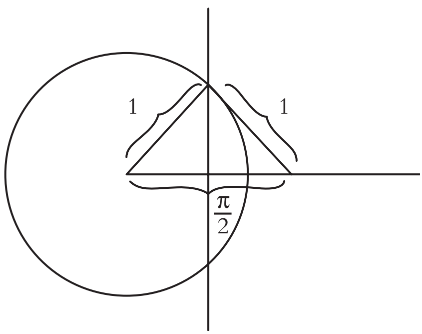

Fig. 8.6. Model of the Failure of the Triangle Inequality Property.

[43] Hilbert's quick sketch of the argument is as follows. Suppose $\sqrt{1-(\pi/4)^2}$ were in the Pythagorean field constructed, then, since this field is minimal, it would be represented by an expression formed from π and 1 by the five operations allowed; Hilbert denotes this expresssion by $A(1,\pi)$. But then, as he points out, $A(1, t)$ must represent the corresponding element $\sqrt{1-(t/4)^2}$ of the corresponding minimal field constructed from 1 and the real number t, whatever t is. However, while $A(1, t)$ is always real, it is obvious that $\sqrt{1-(t/4)^2}$ will be imaginary for t sufficiently large t. Hence, $A(1, t)$ will not always represent it.

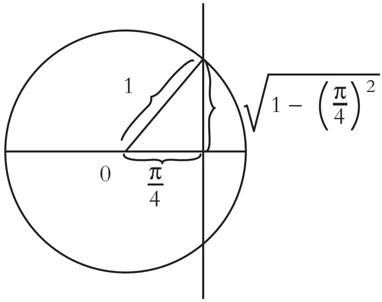

Fig. 8.7. Model of the Failure of the Triangle Inequality Property, continued.

In his review of Hilbert's *Grundlagen*, Sommer remarks that one cannot prove the line-circle property in Hilbert's system (Sommer, 1899/1900, p. 291). What Sommer observes already follows from what Hilbert shows in his lectures, and what he says there explicitly (e.g. pp. 64–65). Sommer, of course, certainly knew that Hilbert had shown this. For one thing, this is the same Julius Sommer, the 'friend', who, along with Minkowski, is thanked by Hilbert at the end of the first edition of the *Grundlagen* for help with proofreading. For another, Sommer refers to Hilbert's lectures course for 1898/1899 directly, in a different context (*loc. cit.*, p. 292). Furthermore, the mathematical core of the result is in fact *present* in the *Grundlagen*, if only implicitly and in a more abstract form. Let us turn to this now, and then return to Sommer's remark.

The model described above really presents a result about the abstract conceptual structures which mature axiom systems represent, for what Hilbert has in effect shown, in modern algebraic terms, is that not every Pythagorean field is Euclidean. An ordered field is said to be *Pythagorean* when the Pythagorean operation holds, that is, when $\sqrt{x^2 + y^2}$ is in the field whenever x and y are. (Note that, for each real r, there is a minimal, and countable, Pythagorean sub-field of the reals containing the rationals and r, a fact which Hilbert frequently employs in his independence proofs.) An ordered field K is said to be *Euclidean* when for any non-negative element $x \in K$, we also have $\sqrt{x} \in K$. It is obvious that every ordered Euclidean field is Pythagorean, but Hilbert shows here that the converse fails, for his model is formed from a field which is Pythagorean; $\frac{\pi}{4}$ is an element, given that π is, and so are $(\frac{\pi}{4})^2$ and $1 - (\frac{\pi}{4})^2$; but, as we have seen, $\sqrt{1-(\pi/4)^2}$ is not, and so the field cannot be Euclidean.[44] The key point is summed up in the following result: In an analytic geometry whose coordinates are given by an

[44] That Hilbert had in fact shown this was first pointed out to me by Helmut Karzel.

ordered Pythagorean field, one can always construct a triangle from three sides satisfying the triangle inequality if and only if the underlying coordinate field is also Euclidean. Indeed, for an analytic geometry based on an ordered field, the Euclidean field property is equivalent to the line-circle property, and this is in turn equivalent to the circle-circle property, the property directly relevant to Hilbert's proof of the Three Chord Theorem (Hartshorne, 2000, pp. 144–146). Given this, the connection between the Euclidean field property and the formation of a triangle from any three lines satisfying the triangle inequality is obvious.

To repeat, the result is a thoroughly abstract one, a result about fields. The inspiration is again intuitive, but this time the major fruit is a new theorem in abstract mathematics. (The result can be found in the *Grundlagen der Geometrie*, though the background work concerns solely the algebraic equivalents to elementary constructions.)

This algebraic result is strongly hinted at in the *Ferienkurs* Hilbert gave in 1898. (See Hallett and Majer, 2004, Chapter 4; see pp. 22–23 of Hilbert's course.) Hilbert poses the question of whether, given a segment product $c \cdot d$, there is a segment x such that $x^2 = c \cdot d$, i.e. a square root for $c \cdot d$. His comment suggests that he thinks this is not always the case, and this is precisely what the counterexample outlined above confirms. For example, in the second diagram given, consider the horizontal and vertical products formed by the four segments arising from the intersection of the horizontal and vertical chords. The horizontal segment product (our $c \cdot d$) is $(1 + \pi/4) \cdot (1 - \pi/4)$, which equals the vertical segment product $\sqrt{(1 - (\pi/4)^2)} \cdot \sqrt{(1 - (\pi/4)^2)}$; thus, the x sought is $\sqrt{(1 - (\pi/4)^2)}$, which does not exist in the model, as we have seen. Thus, a question from the 1898 *Ferienkurs* is answered.

Hilbert notes that the problem exhibited here does not just arise because of the involvement of the transcendental number π; he gives an example of an elementary number which will be in any Euclidean field over the rationals, but which is not in the minimal Pythagorean field, namely $\sqrt{1 + \sqrt{2}}$. (See Hilbert's own lecture notes, p. 67, and also p. 67 of the *Ausarbeitung*. Hartshorne (2000, p. 147) gives further details of the counterexample.) Another example is given in the *Grundlagen* itself, to which we will come in a moment.

What is now interesting is how this *abstract* result is used to yield more information at the intuitive level, at the level of synthetic, Euclidean geometry rooted in elementary constructions. In the 1898/1899 lectures, Hilbert himself seems to suggest that the problem might have to do with a continuity assumption. On p. 64 of the *Ausarbeitung*, he says when assuming either the line-circle or circle-circle properties, one is actually assuming that 'the circle is a closed figure'.[45] Moreover, it is precisely in the context of the failure of continuity in Hilbert's system that Sommer makes his remark about the line-circle property, adding that 'it remains undecided whether or not the circle is a closed figure' (Sommer, 1899/1900, p. 291), thus adopting Hilbert's terminology from the lectures.

But continuity is not necessary to close this particular gap; just assuming the Euclideanness of the underlying field will do, and results in Hilbert's *Grundlagen* make this quite clear. In his lectures, having pointed out the problem with constructing the triangle described above, Hilbert says:

> We will return to these considerations later, after we have built up geometry completely, and when we investigate the means which can serve in construction. We will then become acquainted with the fine distinction which arises when one is allowed to use a pair of compasses [*der Zirkel*] in an unrestricted way, or whether it can only be used for measuring off segments and angles (the right-angle suffices). (Hilbert (*1898/1899, p. 67), p. 260 in Hallett and Majer (2004))[46]

Hilbert does not return to these matters in his own lecture notes, although there is a section in the *Ausarbeitung*, pp. 170–173 which takes up directly the question of which geometrical constructions are performable in his axiom system.[47] This discussion is generalized in Chapter VII of the first edition of the *Grundlagen* (Hallett and Majer, 2004, Chapter 5), and it is this which we will consider here, although none of the discussion is motivated, as it is in the preceding lecture notes, by the original consideration of the *TCT*. Hilbert proves two results: (1) Any constructions carried out and justified on the basis of Axioms I–V are necessarily constructions using just a straightedge and a what he calls a 'segment mover [*Streckenübertrager*]' (for which a pair of dividers would serve), the first for drawing straight lines, and the second for measuring off segments. (2) The algebraic equivalent to these constructions is the Pythagorean field. The use of the '*Streckenübertrager*' corresponds to the *restricted* use of the pair of compasses in constructions, i.e. marking off given

[45] A related remark is made in the original lecture notes, p. 64, and here Hilbert adds that Euclid has 'a similar sounding axiom'. There is, however, no such assumption in the *Elements*, either in the Postulates or under the Common Notions. Hilbert may have been referring indirectly to the Euclidean Definitions. For Euclid, a circle is a certain kind of figure, and a figure is 'that which is contained by any boundary or boundaries' (Definition 14); see Heath (1925, Volume 1, p. 153). Perhaps the somewhat vague 'contained by' and 'boundary' suggest 'closed', and perhaps that the circle has no 'gaps'.

[46] On p. 68 of the *Ausarbeitung* of these notes (Hallett and Majer, 2004, 339–340), Hilbert writes:

> We will discover among other things that it makes an essential difference, whether one is allowed the unrestricted use of a pair of compasses or only allowed to use it for the measuring off of segments and angles.

[47] See also Hilbert's 1898 *Ferienkurs*, pp. 12–14; (Hallett and Majer, 2004, Chapter 3).

radii using the compass as, say, a pair of dividers. What this means is that the constructions licensed by Hilbert's geometrical system of the 1890s can be carried out (i.e. the existence of the points constructed justified) in any analytic geometry whose coordinates form a Pythagorean field, even the minimal Pythagorean field built over the rationals. (These are Theorems 40 and 41 of the first edition of the *Grundlagen*, pp. 79–81.)[48] Hilbert then remarks that:

> From this Theorem [41], we can see immediately that not every problem solvable by use of a pair of compasses can be solved by means of a ruler and segment mover [*Streckenübertrager*] only. (Hilbert, *Grundlagen*, p. 81, i.e. p. 516 in Hallett and Majer (2004))

In other words, the *unrestricted* use of the pair of compasses in constructions is *not* justified in his system.

To show this, Hilbert first gives an example of a real number which cannot be in the minimal Pythagorean field built over the rationals, namely $\sqrt{2|\sqrt{2}| - 2}$, despite the fact 1 and $|\sqrt{2}| - 1$ are both in the field. (The example is thus slightly different from that given in the lectures.) It follows that we cannot construct by means of 'Lineal und Streckenübertrager' a right-angled triangle with sides of length 1 (hypothenuse), $|\sqrt{2}| - 1$ and $\sqrt{2|\sqrt{2}| - 2}$, since the latter length cannot correspond to an element in the minimal Pythagorean field; hence the construction problem is not soluble in Hilbert's geometry. But, as Hilbert remarks (*Grundlagen*, p. 82), the problem *is* immediately soluble by a compass construction; the number Hilbert specifies ($\sqrt{2|\sqrt{2}| - 2}$) is, of course, in any Euclidean ordered field built over the rationals.

The central point is now this: If K is the set of all real numbers obtained from the rationals by the operations of addition, multiplication, subtraction, and division, and such that K contains square roots for all of its positive elements, then K is Euclidean and is the smallest field over which straightedge and compass constructions can be carried out. (See Hartshorne, 2000, p. 147.) The salient point is even clearer in Hilbert's Theorem 44 (p. 86), which deals with the problem of characterizing which straightedge and compass constructions can be carried out in his geometry (i.e. reduced to constructions by straightedge and *Streckenübertrager*). In the statement of the condition (see *Grundlagen*, Theorem 44, p. 86), Hilbert quite clearly expresses the fact that the algebraic condition corresponding to the compass construction is that each number in

[48] Kürschák showed that the *Streckenübertrager* can be dispensed with in favour of an *Eichmaß*, i.e. a device which measures off a single fixed segment. (See Kürschák 1902.) Hilbert makes a corresponding adjustment, with acknowledgement to Kürschák's work, in the Second Edition of the *Grundlagen* (Hilbert, 1903, 74, 77).

the field of coordinates has a square root in the field, i.e. is Euclidean.[49] Thus, the problem is not to do with a failure of continuity, as Sommer suggests, very possibly leaning on Hilbert's original remark, but rather with the failure of a very much weaker field property, and it seems that Hilbert would certainly have been fully aware of this, by the time of the composition of the *Grundlagen* if not earlier. However, having said this, it is not entirely clear what principle should be added to the axiom system to guarantee Euclideanness; adding the circle-circle property itself as an axiom might appear somewhat *ad hoc*.

Finally, to come back to the constructions involved in Euclid's proofs of I, 1 and I, 22, although both apparently involve ruler and circle constructions, an adequate construction for the *first* case *can* be given using Pythagorean operations alone, thus, uses the compass only in the 'restricted' sense. An equilateral triangle can be constructed by Pythagorean operations just in case $\sqrt{3}$ is in the underlying coordinate field. But $\sqrt{3} = \sqrt{1 + (\sqrt{2})^2}$, and $\sqrt{2} = \sqrt{1 + 1^2}$. Hilbert shows this: see the *Ausarbeitung*, p. 173. Hence, the equilateral triangle can be constructed in Hilbert's axiom system. (The actual construction is given in the 1898 *Ferienkurs*, p. 15: see Hallett and Majer (2004, p. 169).) Thus, Euclid's construction here does not in the least assume continuity, or even 'Euclideanness'.

In sum, what we have here is another investigation which begins with a 'purity of method' question, which then employs higher mathematics in its pursuit, achieves an abstract result, and also uses the knowledge gained to inform us and instruct us about elementary geometry, the geometry closest to intuition.

8.5 Conclusion

Let us draw some general conclusions from these examples.

The concern with 'purity of method' usually focuses on some general consideration of 'appropriateness'; this at least is the way that Hilbert casts

[49] In the 1902 lectures, for the purposes of showing the independence of the *Vollständigkeitsaxiom*, Hilbert contructs a (countable) model based on a minimal Pythagorean field. He adds that this geometry (i.e. model) is particularly interesting, for

> ... it contains only points and lines which can be found solely by measuring off segments and angles.
>
> As we have shown in our *Grundlagen*, page 81ff., not every segment can be constructed by means of measuring off segments alone. Take as an example the segment $\sqrt{2\sqrt{2} - 2}$. (Hilbert (*1902, 89–90); Hallett and Majer (2004, 581)).

it in the passages quoted in the Introduction. And at *some* level, this was a concern of Hilbert's. There are many examples in Hilbert's 1898/1899 lectures on Euclidean geometry where Hilbert is concerned to show that some implicit assumption made by Euclid or his successors is in fact dispensable; some of these were cited in Section 8.2. But the examples we have concentrated on in Section 8.4 show that the focus on eliminating 'inappropriate' assumptions was only one facet of Hilbert's work on geometry; the cases examined are all ones where the appropriateness of an assumption is initially questioned, but where it is shown that it is indeed *required*. Among other things, this forces a revision of what is taken to be 'appropriate'. It is part of the lesson taught by the examples that in general this cannot be left to intuitive or informal assessment, for instance, the intuitive assessment of the complexity of the concepts used in the assumption in question.

Furthermore, full investigation of geometry requires its axiomatization, and proper examination of this requires that it be cut loose from its natural epistemological roots, or, at the very least, no longer immovably tied to them. According to Hilbert's new conception of mathematics, an important part of geometrical knowledge is knowledge which is quite independent of interpretation, knowledge of the logical relationships between the various parts of the theory, the way the axioms combine to prove theorems, the reverse relationships between the theorems and the axioms, and so on, all components we have seen in the examples. And in garnering this sort of geometrical knowledge, there is *not* the restriction to the 'appropriate' which we see in the 'Euclidean' part of Hilbert's concerns. What is invoked in pursuing this knowledge might be some highly elaborate theory, as it is in the analysis of the Isoceles Triangle Theorem, a theory far removed from the 'appropriate' intuitive roots of geometry. Even in the cases of the fairly simple models of the analytic plane used to demonstrate the failure of Desargues's Planar Theorem, the models though visualizable are far from straightforwardly 'intuitive'.

One might be tempted to say that the knowledge so achieved is not geometrical knowledge, but rather purely formal logical knowledge or (as it would be usually put now) *meta*-geometrical knowledge. But although this designation is convenient in some respects, it is undoubtedly misleading. As we have seen, the 'meta'-geometrical results have a direct bearing on what is taken to be geometrical knowledge of the most basic intuitive kind; in particular it can reveal a great deal about the *content* of intuitive geometrical knowledge. In short, it effects an alteration in geometrical knowledge, and must therefore be considered to be a *source* of geometrical knowledge. To repeat: for Hilbert, meta-mathematical investigation of a theory is as much a part of the study of a theory as is working out its consequences, or examining its foundations in the way that Frege, for instance, does.

Thus, for Hilbert's investigations in geometry, 'purity of method' analysis in the standard sense is elaborated into the 'analysis of intuition'. This resolves into two separate investigations, one at the intuitive level, and one at the abstract level, levels which frequently interact and instruct each other. Furthermore, extracting this information often itself involves a detour into the abstract. One example is given by the investigation of Desargues's Planar Theorem, where the use of the structure of segment fields requires first an abstract axiomatization of ordered fields. In particular, as the examples treated here make abundantly clear, higher mathematics is used to instruct or adumbrate intuition, or at the very least to instruct us about it and what it entails.

The second conclusion concerns the notion of the 'elementary' or 'primitive' with respect to a domain of knowledge. The examples we have considered show that often we have to adopt a *non-elementary* point of view in order to achieve results about apparently elementary theorems. Hilbert often stressed, just as Klein did, that elementary mathematics must be studied from an advanced standpoint. Certainly as far as Hilbert's work is concerned, this is much more than a pedagogical point, although Hilbert often stressed this, too.[50] For one thing, the examples considered show that genuine knowledge concerning the elementary domain can flow from such investigations, and they also show that apparently elementary propositions contain within themselves non-elementary consequences, often in a coded form.[51] Furthermore, consideration of the intuitive and elementary is used to generate results at the *abstract* level; one example of this was the result about the abstract theory of fields sketched in Section 8.4.3.

But there is a further question about appropriateness. Investigation of independence inevitably involves mathematics as broadly construed as possible, since it involves the construction of models, indeed, requires the precise description which is only afforded by mathematical models. Given this, one might ask whether there is an appropriate limit on the mathematics which can be used for the analysis of the intuitive. There is an obvious practical limitation: in constructing models, one naturally uses those branches of mathematics which are most familiar, and which will afford the finest control over the models we construct. In Hilbert's case, resort to higher analysis is especially natural, given the extensive theoretical development of analytic geometry in the 19th

[50] See e.g. the introductory remarks in Hilbert's *Ferienkurs* for 1896 (Hilbert, *1896), in Hallett and Majer (2004, Chapter 3).

[51] There is surely here more than an analogy with the 'hidden higher-order content' stressed by Isaacson in connection with the Gödel incompleteness phenomena for arithmetic. See Isaacson (1987).

century, which among other things produced the intricate analytic descriptions of non-Euclidean geometry (models), and also involved the treatment and extension of intuitively based geometrical ideas by highly unintuitive means. It is also worth noting that the reason behind the very development of analytic geometry was to be able to solve problems posed by synthetic geometry in an analytic way, and to construe the solutions synthetically. There is a clear sense that this is also what much of Hilbert's work with analytic models does.

But there is a philosophical reason which goes along with this. Part of the point of Hilbert's axiomatization of geometry is to remove it to an abstract sphere 'at the conceptual level' where it is indeed divorced from its 'natural' interpretation in the imperfectly understood structure of space, becoming in the process a self-standing theory. The ideal for Hilbert in this respect was the theory of numbers, and, by extension, analysis. The important philosophical point about the theory of numbers is that for Hilbert it was entirely 'a product of the mind'; it is not in origin an empirical, or empirically inspired, theory like geometry, and can be considered in some sense as already on the 'conceptual level'. Thus, Hilbert shared Gauss' view of the difference in status of geometry and number theory. In short, these geometrical investigations and the consequent extension of geometrical knowledge presuppose arithmetic at some level.

That elementary arithmetic is *explicitly* presupposed is made clear at the beginning of Hilbert's 1898/1899 lectures:

> It is of importance to fix precisely the starting point of our investigation: *We consider as given the laws of pure logic and full arithmetic.* (On the relationship between logic and arithmetic, cf. *Dedekind*, 'Was sind und was sollen die Zahlen?' [i.e. Dedekind (1888)].) Our question will then be: *What propositions must we 'adjoin' to the domain defined in order to obtain Euclidean geometry?* (*Ausarbeitung*, p. 2, p. 303 in Hallett and Majer (2004).)[52]

This use of number theory and analysis reflects two important philosophical positions which Hilbert held at that time concerning arithmetic and analysis. First, just prior to this (e.g. *circa* 1896; see the *Ferienkurs* mentioned in n.50), he seems to have held a version of the 'Dirichlet thesis' that all of higher analysis will in some sense 'reduce' to the theory of natural numbers, a thesis which is stated without challenge in the *Vorwort* to Dedekind's 1888 monograph. Secondly, there is clear indication that he thought of arithmetic as conceptually prior to geometry. This is illustrated in his 1905 lectures.

[52] There is no corresponding passage in Hilbert's own lecture notes, suggesting perhaps that Hilbert became aware a little later that something ought to be said about the foundation for the mathematics presupposed in the investigation of geometry.

Hilbert remarks that there are in principle three different ways in which one might provide the basis of the theory of number: 'genetically' as was common in the 19th century; through axiomatization, Hilbert's preferred method whenever possible; or *geometrically*. With respect to the latter, after sketching how such a reduction might in principle proceed, Hilbert says the following:

> The objectionable and troublesome [*mißliche*] thing in this can be seen immediately: it consists in the essential use of geometrical intuitions and geometrical propositions, while geometry and its foundation are nevertheless less simple than arithmetic and its foundations. One must also note that to lay out a foundation for geometry, we already frequently use the numbers. Thus, here the simpler would be reduced to the more complicated, or in any case to more than is necessary for the foundation. (Hilbert, *1905, p. 9)

But while the foundational investigation of geometry presupposes arithmetic (and analysis), there was at this time no similar foundational investigation of arithmetic, and no investigation of the conceptual connection between more elementary parts of arithmetic and higher arithmetic and analysis.[53] In particular, this complex of theory was not subject to axiomatic analysis and indeed not even axiomatized. In the course of Hilbert's work on geometry, he does axiomatize an important part of it, namely the theory of ordered fields, mainly for the purpose of revealing certain analytic structure in the geometry of segments in the analysis of Desargues's Theorem, giving rise to the system of complete, ordered fields published in Hilbert (1900*a*). Nevertheless, the theory of natural numbers was not treated axiomatically by Hilbert until very much later. And the important extensions of Archimedean and non-Archimedean analysis involving (say) complex function theory were never treated by Hilbert as axiomatic theories. Of course, what is in question in the work examined here are certain aspects of the foundations of geometry. Nevertheless, Hilbert was well aware that the results garnered are in a certain strong sense relative, and that the foundational investigation of geometry must be part of a wider foundational programme. Indeed, Hilbert's famous lecture on mathematical problems from 1900 sets out (as Problem 2) precisely the problem of investigating the axiom system for the real numbers, in particular showing the mathematical existence of the real numbers, where there is no recourse to a natural theoretical companion such as is possessed by synthetic geometry. Thus, a limited foundational investigation gives birth to another more general one.

[53] In the 1920s, Hilbert stated decisively his rejection of the Dirichlet thesis, though it is not clear when he abandoned it.

Bibliography

Bernays, Paul (1922), 'Über Hilberts Gedanken zur Grundlegung der Arithmetik', *Jahresbericht der deutschen Mathematiker-Vereinigung*, 31, 10–19.

Boolos, George (1990), 'The Standard of Equality of Numbers', in George Boolos (ed.), *Meaning and Method: Essays in Honor of Hilary Putnam* (Cambridge: Cambridge University Press) 261–277. Reprinted in Boolos (1998, 202–219). Page numbers refer to this reprinting.

—— (1998), *Logic, Logic and Logic* (Cambridge, MA: Harvard University Press).

Browder, Felix (1976), 'Mathematical Developments Arising from the Hilbert Problems', *American Mathematical Society, Proceedings of Symposia in Pure Mathematics*, 28 (parts 1 and 2).

Coolidge, Julian Lowell (1916), *A Treatise on the Circle and the Sphere* (Oxford: Clarendon Press). Reprinted by Chelsea Publishing Company, New York, 1971.

Coxeter, H. M. S. (1974), *Projective Geometry*, 2nd edn (Toronto: University of Toronto Press). Reprinted in 1987 by Springer-Verlag, Berlin, Heidelberg, New York.

Dedekind, Richard (1872), *Stetigkeit und irrationale Zahlen* (Braunschweig: Vieweg und Sohn). Latest reprint, 1969. Also reprinted in Dedekind (1932, 315–332); English trans. in Ewald (1996, 765–779).

—— (1888), *Was sind und was sollen die Zahlen?* (Braunschweig: Vieweg und Sohn). Latest reprint, 1969. Also reprinted in Dedekind (1932, 335–391); English trans. in Ewald (1996, 787–833).

—— (1932), *Gesammelte mathematische Werke, Band 3. Herausgegeben von Robert Fricke, Emmy Noether and Öystein Ore* (Braunschweig: Friedrich Vieweg und Sohn). Reprinted with some omissions by Chelsea Publishing Co., New York, 1969.

Dehn, Max (1900), 'Die Legendre'sche Sätze über die Winkelsumme im Dreieck', *Mathematische Annalen*, 53, 404–439.

Ewald, William (ed.) (1996), *From Kant to Hilbert*. 2 vols. (Oxford: Oxford University Press).

Ewald, William and Sieg, Wilfried (eds.) (2008), *David Hilbert's Lectures on the Foundations of Logic and Arithmetic, 1917–1933* (Heidelberg, Berlin, New York: Springer). *David Hilbert's Foundational Lectures, Vol. 3*.

Ewald, William, Hallett, Michael, and Sieg, Wilfried (eds.) (Forthcoming), *David Hilbert's Lectures on the Foundations of Logic and Arithmetic, 1894–1917* (Heidelberg, Berlin, New York: Springer). *David Hilbert's Foundational Lectures, Vol. 2*.

Frege, Gottlob (1884), *Die Grundlagen der Arithmetik* (Wilhelm Koebner). Reprinted by Felix Meiner Verlag, Hamburg, 1986, 1988. English trans. by J. L. Austin as *The Foundations of Arithmetic*, 2nd edn (Oxford: Basil Blackwell, 1953). The Austin edition is a bilingual one, and the page numbers are the same for both the German and English texts.

——(1906), 'Über die Grundlagen der Geometrie', *Jahresbericht der deutschen Mathematiker-Vereinigung*, 15, 293–309, 377–403, 423–430. Reprinted in Frege (1967, 262–272).

——(1967), *Kleine Schriften*, ed. I. Angelelli (Hildesheim: Georg Olms Verlag).

HALLETT, Michael (1979*a*), 'Towards a Theory of Mathematical Research Programmes (I)', *British Journal for the Philosophy of Science*, 30, 1–25.

——(1979*b*), 'Towards a Theory of Mathematical Research Programmes (II)', *British Journal for the Philosophy of Science*, 30, 135–159.

HALLETT, Michael and MAJER, Ulrich (eds.) (2004), *David Hilbert's Lectures on the Foundations of Geometry, 1891–1902* (Heidelberg, Berlin, New York: Springer). *David Hilbert's Foundational Lectures, Vol. 1.*

HART, W. D. (1996), *The Philosophy of Mathematics* (Oxford: Oxford University Press).

HARTSHORNE, Robin (2000), *Geometry: Euclid and Beyond* (New York, Berlin, Heidelberg: Springer).

HEATH, Thomas L. (1925), *The Thirteen Books of Euclid's Elements*. 3 vols. 2nd edn (Cambridge: Cambridge University Press).

HILBERT, David (*1891), '*Projektive Geometrie*', Lecture notes for a course held in the Wintersemester of 1891 at the University of Königsberg. Niedersächsische Staats- und Universitätsbiblithek, Göttingen. First published in partial form in Hallett and Majer (2004, 21–64).

——(*1893/1894), '*Grundlagen der Geometrie*', Lecture notes for a course to have been held in the Wintersemester of 1893/1894 at the University of Königsberg. Niedersächsische Staats- und Universitätsbiblithek, Göttingen. First published in Hallett and Majer (2004, 72–144).

——(1895), 'Über die gerade Linie als kürzeste Verbindung zweier Punkte', *Mathematische Annalen*, 46, 91–96. Reprinted as Anhang I in Hilbert (1903).

——(*1896), '*Feriencurs*', Notes for a short course for *Oberlehre*, given at the University of Göttingen during the Easter holidays, 1896; Niedersächsische Staats- und Universitätsbiblithek, Göttingen. First published in Hallett and Majer (2004, 152–158).

——(*1898), '*Feriencurs*', Notes for a short course for *Oberlehre*, given at the University of Göttingen during the Easter holidays, 1898; Niedersächsische Staats- und Universitätsbiblithek, Göttingen. First published in Hallett and Majer (2004, 160–184).

——(*1898/1899), '*Grundlagen der Euklidischen Geometrie*', Lecture notes for a course held in the Wintersemester of 1898/1899 at the Georg-August Universität, Göttingen. Niedersächsische Staats- und Universitätsbiblithek, Göttingen. First published in Hallett and Majer (2004, 221–301).

——(*1899), '*Elemente der Euklidischen Geometrie*', *Ausarbeitung* by Hans von Schaper of the lecture notes Hilbert (*1898/1899). Niedersächsische Staats- und Universitätsbiblithek, Göttingen, and the Mathemtisches Institut of the Georg-August Universität, Göttingen. First published in Hallett and Majer (2004, 302–406).

—— (1899), 'Grundlagen der Geometrie', in *Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmals in Göttingen* (Leipzig: B. G. Teubner). Republished in chapter 5 of Hallett and Majer (2004 436–525).

—— (1900*a*), 'Über den Zahlbegriff', *Jahresbericht der deutschen Mathematiker-Vereinigung*, 8, 180–185. Reprinted (with small modifications) in 2nd to 7th editions of Hilbert (1899).

—— (1900*b*), 'Les principes fondamentaux de la géométrie', *Annales scientifiques de la École Normale Supérieur*, 103–209. French trans. by L. Laugel of Hilbert (1899), with additions by Hilbert. Additions republished in Hallett and Majer (2004, 526–529).

—— (1900*c*), 'Mathematische Probleme', *Nachrichten von der königlichen Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse*, 253–296. English trans. by Mary Winston Newson in *Bulletin of the American Mathematical Society*, (2)**8** (1902), 437–479, and in Browder (1976), 1–34. Partial English trans. in Ewald (1996), Vol. 2, 1096–1105.

—— (*1902), '*Grundlagen der Geometrie*', *Ausarbeitung* by August Adler for lectures in the Sommersemester of 1902 at the Georg-August Universität, Göttingen. Library of the Mathematisches Institut. First published in chapter 6 of Hallett and Majer (2004, 540–606).

—— (1902/1903), 'Über den Satz von der Gleichheit der Basiswinkel im gleichschenklichen Dreieck', *Proceedings of the London Mathematical Society*, 35, 50–67. Republished as Anhang II in Hilbert (1903), and in revised form in subsequent editions.

—— (1903), *Grundlagen der Geometrie. Zweite Auflage* (Leipzig and Berlin: B. G. Teubner). Revised edition, also contains the sections added to Hilbert (1900*b*).

—— (*1905), '*Logische Principien des mathematischen Denkens*', Lecture notes for a course held in the Sommersemester of 1905 at the Georg-August Universität, Göttingen. Library of the Mathematisches Institut. To appear in Ewald, Hallett and Sieg (Forthcoming).

—— (*1917/1918), '*Prinzipien der Mathematik*', *Ausarbeitung* by Paul Bernays of notes for lectures in the Wintersemester 1917/1918, Mathematisches Institut of the Georg-August Universität, Göttingen. Published in Ewald and Sieg (2008).

—— (1918), 'Axiomatisches Denken', *Mathematische Annalen*, 78, 405–415. Reprinted in Hilbert (1935, 146–156); English trans. in Ewald (1996), Vol. 2, 1105–1115.

—— (1919), *Natur und mathematisches Erkennen* (Basil, Berlin, Boston: Birkhäuser Verlag). Lecture notes for a course held in the Wintersemester of 1921/1922 at the Georg-August Universität, Göttingen, and prepared by Paul Bernays. Library of the Mathematisches Institut. First published in 1992 in an edition edited by David Rowe.

—— (*1921/1922), '*Grundlagen der Mathematik*', Lecture notes for a course held in the Wintersemester of 1921/1922 at the Georg-August Universität, Göttingen, and prepared by Paul Bernays. Library of the Mathematisches Institut. First published in Ewald and Sieg (2008).

——— (1923), *Grundlagen der Geometrie. Sechste Auflage* (Leipzig and Berlin: B. G. Teubner). Revised edn, with extra notes.

——— (1930), *Grundlagen der Geometrie. Siebente umgearbeitete und vermehrte Auflage* (Leipzig and Berlin: B. G. Teubner). New edn, extensively revised by Arnold Schmidt.

——— (1935), *Gesammelte Abhandlungen, Band 3* (Berlin: Julius Springer).

Hilbert, David and Cohn-Vossen, Stefan (1932), *Anschauliche Geometrie* (Berlin: Julius Springer). 2nd edn published by Springer-Verlag in 1996. English trans. *Geometry and the Imagination* by P. Nemenyi (Chelsea Publishing Company, 1952).

Isaacson, Daniel (1987), 'Arithmetical Truth and Hidden Higher-order Concepts', in *Logic Colloquium '85* (Amsterdam: North-Holland Publishing Co.). Reprinted in Hart (1996, 203–224).

Killing, Wilhelm (1898), *Einführung in die Grundlagen der Geometrie. Band II* (Paderborn: Druck und Verlag von Ferdinand Schöningh).

Kirchhoff, Gustav (1877), *Vorlesungen über mathematische Physik. Band I: Mechanik* (Leipzig: J. Barth).

Kürschák, Josef (1902), 'Das Streckenabtragen', *Mathematische Annalen*, 55, 597–598.

Pasch, Moritz (1882), *Vorlesungen über neuere Geometrie* (Leipzig: B. G. Teubner).

Poincaré, Henri (1891), 'Les géometries non-euclidiennes', *Revue générale des sciences pures et appliquées*, 2, 769–776. Reprinted with alterations in Poincaré (1902).

——— (1902), *La Science et l'Hypothése* (Paris: Ernst Flammarion). English translation as *Science and Hypothesis*, Walter Scott Publishing Co., London, reprinted by Dover Publications, New York, 1952.

Samuel, Pierre (1988), *Projective Geometry* (Berlin, Heidelberg, New York: Springer-Verlag). English trans. by Silvio Levy of *Géometrie Projective* (Paris: Presses Universitaires de France, 1986).

Sommer, Julius (1899/1900), 'Hilbert's Foundations of Geometry; Review of David Hilbert: *Grundlagen der Geometrie*', *Bulletin of the American Mathematical Society*, 6 (2nd ser.), 287–289.

Stäckel, Paul and Engel, Friedrich (eds.) (1895), *Theorie der Parallellinien von Euklid bis auf Gauss, eine Urkundensammlung zur vorgeschichte der nichteuklidischen Geometrie* (Leipzig: B. G. Teubner).