Philosophia Scientiæ
Travaux d'histoire et de philosophie des sciences

25-3 | 2021
L'analyse dans les mathématiques grecques

Rigorous Purposes of Analysis in Greek Geometry

Viktor Blåsjö

Electronic version
URL: https://journals.openedition.org/philosophiascientiae/3109
DOI: 10.4000/philosophiascientiae.3109
ISSN: 1775-4283

Publisher
Éditions Kimé

Printed version
Date of publication: 25 October 2021
Number of pages: 55-80
ISBN: 978-2-38072-041-9
ISSN: 1281-2463

Electronic reference
Viktor Blåsjö, "Rigorous Purposes of Analysis in Greek Geometry", *Philosophia Scientiæ* [Online], 25-3 | 2021, Online since 25 October 2021, connection on 08 November 2021. URL: http://journals.openedition.org/philosophiascientiae/3109 ; DOI: https://doi.org/10.4000/philosophiascientiae.3109

Rigorous Purposes of Analysis in Greek Geometry

Viktor Blåsjö
Mathematical Institute, Utrecht University (Netherlands)

Résumé : Dans la géométrie grecque, l'analyse est traditionnellement vue comme une technique heuristique. Cependant, de nombreuses analyses dans les traités mathématiques anciens sont difficiles à expliquer en ces termes. Cet article vise à montrer que l'analyse géométrique grecque peut aussi servir des enjeux mathématiques déterminés, qui, sans son apport, rendraient la synthèse incomplète. D'abord, lorsque la solution d'un problème est précédée par une analyse, celle-ci démontre rigoureusement qu'il n'y a pas d'autres solutions au problème que celles offertes dans la synthèse. Ensuite, chaque fois qu'on est en présence d'une hypothèse de construction qui dépasse l'utilisation de la règle et du compas, l'analyse démontre que le problème n'est pas seulement résolvable à partir de cette assomption, mais qu'il est en fait rigoureusement équivalent à elle.

Abstract: Analyses in Greek geometry are traditionally seen as heuristic devices. However, many occurrences of analysis in formal treatises are difficult to justify in such terms. I show that Greek analysies of geometrics can also serve formal mathematical purposes, which are arguably incomplete without which their associated syntheses are arguably incomplete. Firstly, when the solution of a problem is preceded by an analysis, the analysis latter proves rigorously that there are no other solutions to the problem than those offered in the synthesis. Secondly, whenever some construction assumption beyond ruler and compass is made, the problem is not only solvable by that assumption but is in fact equivalent to that assumption in a rigorous sense.

1 Why analyse?

Arguments called "analysis" occur a number of times in ancient geometry, but the precise meaning, nature, and purpose of analysis remains elusive.

Philosophia Scientiæ, 25(3), 2021, 55–80.

Analysis is traditionally viewed as a kind of method of discovery.[1] Faced with a geometrical problem, the method of analysis consists in assuming that it has been solved and then "reasoning backwards" from there in some sense. This is then followed by the "synthesis" in which the problem is solved in the "forward" direction, mirroring the steps of the analysis. Since the synthesis on its own solves the problem, it seems natural to interpret the analysis as a mere preliminary aid that is not logically necessary but heuristically helpful: in other words, a free-wheeling exploration that has no formal value until it is "confirmed by subsequent synthesis" [Heath 1925, I.140]. Thus, according to this view, analysis is "a heuristic method, essentially"; an "investigative tool that does not have proof character" [Sefrin-Weis 2010, 305].

Yet there are many indications that there must be more to Greek analysis, for if discovery and heuristic is all there is to it, then many aspects of the classical corpus of analysis are inexplicable and perplexing. It is not rare to find cases where the alleged heuristic advantages of analysis are not apparent, yet the analyses are carefully spelled out. For example, there are not a few cases in the Greek corpus where an analysis is spelled out even though the synthesis is obvious and clearly just as easy to come up with as the analysis.[2] Indeed, one can legitimately question or problematise the common assumption that analysis is more heuristic or easier to come up with than synthesis [Behboud 1994, 75], [Netz 2000, 143], and whether the analysis, once accomplished, is necessarily much help in finding the synthesis [Netz 2004, 191, 207, 217–218]. There are many examples where analysis does serve these purposes, but there are also not a few instances where it does not, yet is still included by Greek geometers. This suggests that the heuristic role of analysis is not the whole story. The Greeks seem to have felt that it served some other purpose as well.

By far the most substantial ancient account of what analysis is that has come down to us is that given by Pappus in his introduction to *Collection* VII [Jones 1986, 82–85]. But there are reasons to think that this classic description of analysis is misleading in some respects. In particular, it is perhaps deceptively susceptible to being interpreted in terms of modern conceptions of mathematics, which tend to take an axiomatic-deductive paradigm as primary and hence downplay the prominence of construction problems in the Greek tradition. Pappus seems to be speaking of relating propositional statements to one another as consequences and antecedents very much in the mould of deductive logic. Some older scholarship on analysis took this conception as its starting point and hence looked at the difference between analysis and synthesis as being fundamentally about direction of logical

1. This view is expressed in late ancient sources by Pappus (beginning of *Collection* VII, [Jones 1986, 82]), Proclus [Proclus 1970, 165], and Marinus [Taisbak 2003, 248]. It is still generally endorsed, for instance by [Robinson 1936, 464], [Knorr 1975, 9], [Jones 1986, 66], [Menn 2002, 195, 198, 204, 215, 223], [Fournarakis & Christianidis 2006, 47–48], [Sefrin-Weis 2010, xxii, 184].

2. E.g., Archimedes *Sphere and Cylinder* II.1 ("in this case, there is nothing 'heuristic' about analysis", [Netz 2004, 191]) II.3, Pappus *Collection* IV.39.

deduction [Heath 1925, I.139]. This has led to extensive debates as to whether analysis should be seen as consisting in deductions from the assumed result, or in looking for things that would imply the assumed result, or in a chain of reversible two-way implications from the assumed result down to something known. These debates have been inconclusive. More recently, many have felt that a focus on logical direction is not the best way to understand analysis [Behboud 1994, 52–57], [Hintikka & Remes 1974, 11, 31–32], [Berggren & Van Brummelen 2000, 2–8], [Sefrin-Weis 2010, 185–188], [Saito & Sidoli 2010, 583]. In fact, one very striking aspect of the Greek use of analysis is that it is almost always used for problems (that is, requests to construct a geometrical object with given properties) and rarely for theorems. This is contrary to what one would expect both from a modern point of view (where starting from Q when proving $P \implies Q$ is a standard trick often covered in "Intro to Proofs" books) and from the account given by Pappus (who mentions analysis as applied to theorems and to problems on seemingly equal footing). But "a consensus in the current literature", based on the actual Greek analytical corpus, is that "whatever analyses do, they do it better for problems than for theorems" [Netz 2000, 147].[3] Some have gone so far as to argue that "the instances of theoretic analysis [that is, analysis applied to theorems rather than to problems], rare on any account, all seem to be due to later commentators" and are often "artificial and unnecessary" [Knorr 1975, 360, 358].[4] This strongly suggests that the Greek conception of analysis cannot naturally be fitted into a modern axiomatic-deductive picture of mathematics. Instead it appears to have served a purpose that was somehow closely tied to problems and constructions.

The assumption that the synthesis formally speaking makes the analysis superfluous is commonly accepted.[5] According to such a view, the purpose of analysis would be external to technical mathematics itself. Instead it must serve something like heuristic, pedagogical, or rhetorical purposes.[6] But any such interpretation gives rise to the problem: "Why did the Greeks

3. For example, Pappus *Collection* VII, which consists of hundreds of propositions as a commentary following very closely key works in the analytical canon, uses the analysis-synthesis form for virtually all problems and only a tiny fraction of the theorems [cf. Behboud 1994, 57]. In Archimedes there are six analyses for problems (*Sphere and Cylinder* II: 1, 3–7) and none for theorems.

4. This interpretation may well be a bit on the extreme side. For example, Sidoli & Saito note that there are theoretic analyses in Apollonius *On Cutting Off a Ratio* [Apollonius 1987] (although these are lemmas to construction problems) [Sidoli & Saito 2012, 27]. In any case, the general picture that analyses are especially associated with problems remains undeniable.

5. "The analysis [...] contains the creative mathematical work [but] the synthesis nevertheless is the part that carries the proof" [Sefrin-Weis 2010, xxiii]. "The synthesis will often be easy after a successful analysis, but it is the part that carries the formal proof" [Sefrin-Weis 2010, 188].

6. Knorr argues that analyses make the steps of the synthesis appear more naturally motivated [Knorr 1975, 9]. Similarly, Netz construes analysis as a "rhetorical tool" that mathematicians use to make their solutions appear more

publish their analyses?" This would seem to be "something strange, a puzzle: because, after all, the synthesis is sufficient. There seems to be no obvious function for the analysis, once it is directly followed by a synthesis. It is a duplication, doubly strange inside the Greek, highly economical, mathematical presentation" [Netz 2000, 146]. All the more so when the synthesis is simple, as happens regularly. Consider for example Archimedes's *Sphere and Cylinder*. This is a very advanced and authoritative masterpiece, in which, arguably, "intricacy and surprise govern the arrangement of the text" [Netz 2004, 21] in its theoretical development. Yet when treating problems Archimedes suddenly provides even the most elementary solutions with a meticulous analysis—such as in II.3, for example, where the synthesis is so straightforward and obvious that it make little sense to speak of any psychological need for the analysis as a tool of discovery or exposition or teaching. It seems more plausible that Archimedes considered the analysis somehow essential from a logical or mathematical point of view.

Furthermore, attempts to justify the purpose of analyses in extra-mathematical terms do not explain why analyses should be confined to problems only: why would the Greeks not seek such pedagogical and rhetorical advantages for theorems as well?[7] This again suggests that analyses were included for internal mathematical purposes.

Taken together, these considerations strongly suggest that the conception of analysis as a heuristic technique cannot be the whole picture. There must be something else to it, and this seems to be something that serves a formal (not psychological) role for problems (but not theorems). I shall highlight two purposes of analysis that fit precisely this explanatory gap.

2 Analysis as exhaustiveness

The first formal function of analysis I wish to highlight may be called *analysis as exhaustiveness*. I claim that the analysis not only uncovers how a problem can be solved; it also proves that this solution strategy covers all possible solutions.[8] By "solution" I shall mean, throughout this section, a solution configuration—that is, a geometrical object having the property demanded in the enunciation of the problem. For instance, the "solution" of the problem illustrated in Figure 2 consists of the lines DC, EA, AC, which are mentioned in the problem enunciation. It does not include the point F and associated lines

inevitable and definitive [Netz 2000, 153]; "a sort of apology for the synthesis" to make it "appear a bit more 'natural' " [Netz 2004, 207].

7. There is no consensus answer to this question. For some attempts, see [Jones 1986, 67], [Netz 2000, 153].

8. My thesis is exactly that expressed very clearly, but very briefly and in a footnote, by [Zeuthen 1886, 17] (also somewhat less directly in [Zeuthen 1896, 95]). Some modern scholarship has touched on similar connections between analysis and uniqueness [Saito & Sidoli 2010, 591, 600], [Acerbi 2011, 146–148].

which are merely stepping stones needed to construct and verify the solution, rather than being part of the solution object itself. By exhaustiveness I mean finding all solution objects in this sense. This is not to be confused with all possible solution strategies or methods. Of course one could always come up with other ways of producing the same final configuration through other intermediary steps, but this is not relevant to my interpretation of analysis.

The synthesis alone is *not* logically self-sufficient for a complete and exhaustive solution of the problem, contrary to what is commonly assumed. For the synthesis alone only proves that the particular construction offered in the synthesis in fact solves the problem. It typically says nothing about whether other solutions are possible. It is easy to imagine situations where a construction recipe could produce *a* solution to a given problem, yet fail to solve it generally. This can easily happen in underdetermined problem situations such as constructing a triangle with two given side lengths and a given angle other than the one enclosed by these sides, which often has two solutions. A synthesis might give one solution, and indeed prove that it is a true solution. But it would be a grave error to think that this exhausted the problem.

I say that, in many cases, the purpose of the analysis is precisely to ensure that the synthesis has not missed possible alternative solutions. Indeed, the analysis rigorously proves that no such omission is committed. The analysis starts by supposing the problem to have been solved in order to show that *any* solution to the problem leads to the setup of the synthesis. In this way, the analysis proves rigorously that there are no *other* solutions to the problem than those constructed in the synthesis.

A simple example is Apollonius *Conics* II.44: "Given [a parabola], to find a diameter" [Apollonius 2013, 154].[9] The definition of a diameter (Figure 1) makes the synthesis obvious: simply draw any two parallel lines crossing the parabola, bisect them, and join the two midpoints.[10] Clearly we can produce "an indefinite number of diameters" this way, as Apollonius observes. But can we produce *any* diameter this way? The synthesis does not answer this question, but the analysis does. The analysis starts with an arbitrary diameter, cuts across it with two parallels, and infers that by definition the diameter bisects them. This shows that any diameter is one of those generated in the synthesis, because it is producible from the bisection of a pair of parallels, which is exactly what the synthesis covers.

More formally, my general thesis can be stated as follows. Let S be the set of all solutions of a given problem, and let C be the set of all solutions constructed in the synthesis. The synthesis (including its demonstration) alone

9. Sidoli & Isahaya discuss this proposition in detail and argue that it is likely typical of geometrical analyses in the time of Euclid and Apollonius [Sidoli & Isahaya 2018, 17–20].

10. An earlier proposition (II.28) showed that bisecting a pair of parallel lines contained in a parabola implies bisecting all other lines parallel to them as well.

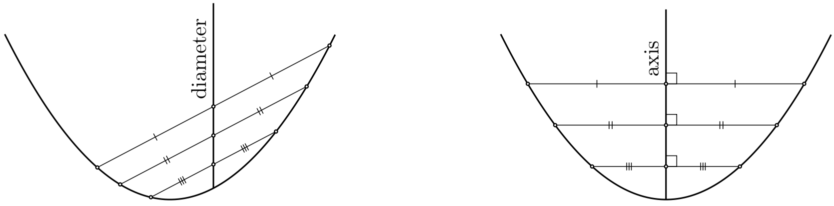


Figure 1: Left: A diameter of a parabola is a straight line that bisects each line in a family of parallel lines contained in the parabola. Right: An axis of a parabola is a diameter that cuts the parallel lines at right angles.

proves that $c \in C \implies c \in S$. The analysis, on the other hand, proves that $S \subseteq C$. It does this in the logically obvious way, namely by considering an arbitrary element $s \in S$ (that is, by starting with any solution of the problem) and proving that it is also an element of C (that is, that it is obtainable by the construction of the synthesis). From $c \in C \implies c \in S$ and $S \subseteq C$ it follows that $C = S$. That is, the analysis and synthesis together prove rigorously not only that the construction is correct but also that it covers all possible solutions.

Exhaustiveness is not the same thing as uniqueness. Many construction problems in the Greek corpus have one unique solution ($|S| = 1$). In such a case the subtleties of the above logical schema collapses. But there are good reasons to approach geometrical problem solving in terms of sets of solutions rather than individual solutions. One indication that Greek geometers were aware of this is that they sometimes explicitly prove that the solution they have given is unique, directly after a full analysis and synthesis have been successfully completed, even in cases where it is intuitively quite obvious that there is only one solution.[11] This suggests that in the preceding analysis and synthesis they allowed the possibility of sets of solutions and intended their arguments to apply to these sets, just as outlined above.

An example of this is Apollonius's *Conics* II.46. This is the same proposition as II.44, which we discussed above, but for an axis instead of a diameter (cf. Figure 1). The analysis and synthesis proceed in a manner analogous to II.44: the synthesis produces an axis and verifies that it is an axis, while the analysis shows that any axis could be obtained in that manner. This shows $C \subseteq S$ and $S \subseteq C$ respectively, which already proves that the synthesis has not missed any solutions. But this still does not say how many solutions there are, that is, the cardinality of S. Hence Apollonius adds a proof (by contradiction) that there could not be more than one axis of a parabola.

11. Apollonius *On Cutting Off a Ratio* I.1.1 [Apollonius 1987] and many times after that; Pappus *Collection* VII: 72, 85, 164, 218.

However, such cases do not validate the above schema completely, because in such cases there is no necessity for the logical function of the analysis that I have provided above. The synthesis (which proves $C \neq \emptyset$ and $c \in C \implies c \in S$) and the uniqueness proof ($|S| = 1$) alone prove that we have exhausted the problem ($C = S$). At least insofar as the uniqueness proof is independent of the analysis, which it sometimes is not (e.g., Pappus *Collection* VII.85). So this may appear to undermine the logical function I have assigned to analysis: in cases where a simple and brief uniqueness proof can be given, and this establishes exhaustiveness, why not do this instead of going through the sometimes rather laborious process of analysis just to accomplish the same goal? To this I have several replies. First of all, uniqueness is often not so straightforward. Already *Elements* I.1 arguably produces two equilateral triangles rather than one, and similar situations where solutions are unique only up to a mirror image or other trivial variations are commonplace in Greek geometry. I furthermore note that almost all of the uniqueness proofs just mentioned are little more than statements that uniqueness is obvious. They say something like: if I cut the thing on that side it would be too small and if on the other side it would be too big; therefore only the point I chose works. But this conclusion is not established by an application of a postulate or proposition, but rather seems to be considered intuitively evident. The analysis allows for a more formal argument. Finally, the analysis way of establishing exhaustiveness is a much more general method since it works regardless of how many solutions there are, whether none or one or many.

We can see the latter point in action in Apollonius's *On Cutting Off a Ratio* [Apollonius 1987].[12] Take for instance I.6.4, which is thoroughly discussed in [Saito & Sidoli 2010, 601–608]. Here the analysis proves that the problem reduces to what is in effect a quadratic equation. It has different numbers of solutions accordingly: 0, 1, or 2, depending on the parameters of the problem. The analysis is eminently useful: it proves that any solution must satisfy the quadratic equation in question. From this it follows rigorously that there are no solutions in certain cases (when the equation has no real roots). The synthesis only states and verifies solutions. The analysis, on the other hand, can prove the impossibility of solutions, by showing that any solutions would have to satisfy an impossible relation. Similarly in the cases where there are solutions. Again the synthesis states and verifies the various solutions. How do you know there are not further ones? You know this because any solution must correspond to the root of a quadratic equation, and this problem can be exhaustively treated. The uniqueness-proof strategy for proving exhaustiveness is not available when there are 0 or 2 solutions, but the analysis way of dealing with exhaustiveness handles it elegantly.

As this example illustrates, analysis is very useful for answering the questions: Under what conditions is the problem unsolvable? Or (more rarely and generally): How many solutions are there? The Greeks had a formal

12. Modern editions of this work are [Rashed & Bellosta 2010], [Apollonius 1987].

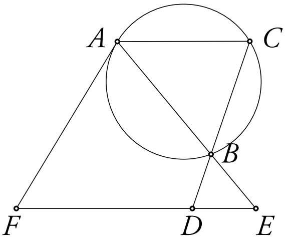


Figure 2: Pappus *Collection* VII.105.

name for arguments answering these questions: diorism. Unlike analysis itself, diorisms are generally recognised as formal arguments serving a rigorous role in showing that the problem has been completely solved. It is a mistake, in my view, to draw such a distinction between the two. Indeed, the diorism often immediately follows the analysis and relies on it.[13] On my reading, the analysis and diorism are both equally rigorous and indispensable arguments for establishing that the construction offered is the definitive one. When the problem is solvable, the analysis proves that the construction given in the synthesis covers *all* possible solutions ($S \subseteq C$), by an argument of the form $s \in S \implies s \in C$. When the problem is unsolvable, however, no set C is defined, so this argument cannot work. To ensure that we are not missing any solutions we must prove that the problem is unsolvable ($S = \emptyset$). The natural way to do this is to assume that there is a solution and prove that this leads to an impossibility ($s \in S \implies \perp$). The initial assumption of the analysis (let $s \in S$) is thus the natural starting point in both cases. In this respect, diorism and analysis are naturally related.

Let us spell out in more detail how exactly the analysis proves that $S \subseteq C$. To this end it is useful to employ the modern conventional division of an analysis-synthesis proof into four stages.[14] I shall illustrate these stages by a prototype example: Pappus *Collection* VII.105 (Figure 2). I shall give the full text of this problem and proof in the translation by [Jones 1986, 236] with minor adaptations and labelling of stages added in brackets. I also add footnotes explicating the steps of the proof. Importantly, the problem in fact has a second solution (Figure 3) which is not that of Pappus's figure. A second purpose of my notes on the proof is to verify that it applies equally well to this alternate configuration.

13. A canonical example is Archimedes *Sphere and Cylinder* II.7.
14. Introduced by [Hankel 1874, 144].

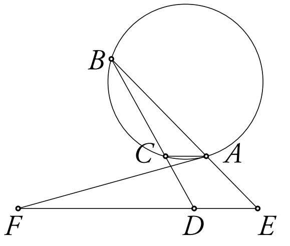


Figure 3: Alternate configuration for Pappus *Collection* VII.105.

Problem Given circle ABC in position, and two points D, E given,[15] to inflect a straight line DBE and, with it produced, to make AC parallel to DE.[16]

Analysis. Transformation. Let it be accomplished, and let FA be drawn tangent. Then since AC is parallel to DE, angle C equals angle CDE [*Elements* I.29]. But angle C equals angle FAE, because FA is tangent to, and AC cuts, the circle.[17] And hence angle FAE equals angle CDE. Thus points A, B, D, F are on a circle.[18] Hence the rectangle contained by AE, EB equals the rectangle contained by FE, ED.[19]

Resolution. But the rectangle contained by AE, EB is given, because it equals the square of the tangent from E to circle ABC.[20] Therefore also the rectangle contained by DE,

15. It is implicitly assumed that DE is outside the circle. This is evident for example in the synthesis where it is assumed that one can draw "the tangent [to the circle] from E".

16. In other words the problem is: Given a line segment DE and a circle, find a point B on the circle, such that the points C and A of intersection of DB and EB with the circle define a line AC parallel to DE.

17. *Elements* III.32, with our C = Euclid's D, our B = Euclid's C, given supplementary angle of FAC = angle CBA. Considering angle sum of triangle ABC, this means angle ACB = angle FAB = angle FAE. In the alternate case: *Elements* III.32, with our B = Euclid's D, our C = Euclid's C, gives supplementary angle of FAB = angle ACB. Hence angle ACB = angle FAE since angle FAE is the vertical angle of the supplementary angle of FAB.

18. By the converse of *Elements* III.22. angle FAE = supplementary angle of its opposite angle FDB. In the alternate case: Use instead the theorem that a quadrilateral is cyclic if the angle between a side and a diagonal (in our case angle BAF) is equal to the angle between the opposite side and the other diagonal (angle FDB).

19. By a common variant of *Elements* III.35–36, noted for instance in Heath's commentary thereon.

20. By the *Data* or "given" equivalent of *Elements* III.35.

EF is given.[21] And DE is given. Hence EF too is given. But it is also given in position; and E is given. Hence F too is given. But from a given point F a straight line FA has been drawn tangent to a circle ABC given in position. Hence FA is given in position and magnitude.[22] And F is given. Therefore A too is given. But E too is given. Therefore AE is given in position. But the circle too is given in position. Therefore point B is given. But each of D, E is given. Hence each of DB, BE is given in position.

Synthesis. The synthesis of the problem will be made as follows. **Construction.** Let the circle be ABC, and the given two points D, E, and let the rectangle contained by DE and some other line[23] EF be made equal to the square of the tangent from E, and from F let a straight line FA be drawn tangent to circle ABC,[24] and let AE be joined, and let DB be joined and produced to C,[25] and let AC be joined. I say that AC is parallel to DE.

Demonstration. For since the rectangle contained by FE, ED equals the square of the tangent from E,[26] while the rectangle contained by AE, EB too equals the square of the tangent,[27] therefore the rectangle contained by AE, EB equals the rectangle contained by FE, ED. Hence points A, B, D, F are on a circle.[28] Therefore angle FAE equals angle BDE.[29] But angle FAE also equals angle ACB in the alternate segment.[30] Hence angle ACB equals angle BDE. And they are alternate angles. Thus AC is parallel to DE.

---

21. As in note 19.
22. It is obvious that FA is not uniquely determined and that there are in fact two possible tangents.
23. It is implied that F is taken on the extension of ED.
24. Again, it is obvious that FA is not uniquely determined and that there are in fact two possible tangents.
25. Of course, in the alternate case, it is AE that needs to be produced (that is, extended) rather than DB, but this is a cosmetic matter. In these kinds of situations, the logic of the proof only concerns points defined as intersections of lines (that may be thought of as infinite). Whether the lines need to be produced or not is not relevant to the proof, only to the practical drawing. As is well known, there are many instances already in the *Elements* where the proof should be taken to apply also to variants of the figure that "would require a slight modification in the wording", not least with regard to whether certain segments need to be produced or not [Heath 1925, I.245]. The same goes for the phrasing in terms of "produced" in Pappus's statement of the problem itself.
26. By definition of the point F.
27. *Elements* III.36.
28. By the converse of the theorem used in note 19.
29. By the converse of the reasoning in note 18.
30. As in note 17.

We can formulate the principles of this and other analysis-synthesis arguments in general terms as follows. In the transformation stage, we assume a solution to be available to us, and draw constructions and inferences from this starting point. In this way we prove that solving the problem amounts to producing some auxiliary configuration (in our example, to finding an auxiliary point F that stands in a certain relation to given entities).

Typically, the transformation uses only reversible inferences and constructions [Behboud 1994, 65], so that solving the problem is strictly equivalent to finding the auxiliary configuration. This is certainly very natural and useful since the steps will be reversed later in the proof. Nevertheless, on my reading, it is not necessary to stipulate this as a logical requirement of the transformation per se. The important thing, rather, established in the transformation is that *any* solution maps to the auxiliary configuration in question.

The resolution confirms that the scaffolding built up in the transformation is indeed sufficient to bridge the gap between the givens and the sought. Unlike the transformation, the resolution is not reasoning "backwards" from the solution but rather forwards *from* the givens of the problem *to* the solution via the auxiliary configuration. This confirms that the solution is recoverable from the auxiliary configuration.

The transformation and resolution together prove that any solution to the problem can be obtained via the auxiliary configuration.

The construction gives a recipe for producing the solution configuration from scratch, going through the path laid out in the analysis. The important thing for our logical purposes is that the synthesis clearly covers all possible auxiliary configurations of the type established in the analysis. In fact, in our example as in many others, the first step of the construction is virtually *defined* as the act of accomplishing the auxiliary configuration in question. While it is not evident that there could not be other solutions of some completely different kind, it is evident that no auxiliary configuration of this type is missed by the construction. But since the analysis proved that all solutions are recoverable from such an auxiliary configuration, it follows that the construction produces all possible solutions. The subsequent demonstration of course verifies that the construction indeed fulfils the requirements of the problem. In other words, the construction defines C, and the demonstration proves that $c \in C \implies c \in S$.

It is not necessary for this argument that the point F should be unique. Indeed, one could argue that the analysis and synthesis also cover the cases in Figure 4, insofar as the point labelling A, C and D, E can be considered pairwise interchangeable.[31] It is not important to my purposes whether Pappus or other Greek mathematicians intended the scope of their proofs to be

31. It is evident that the same proof applies to these configurations. They produce the same segment AC as in the previous figures, but as far as my account of analysis is concerned this is not important. The analysis shows that the synthesis covers all solutions without saying anything about how many solutions there are.

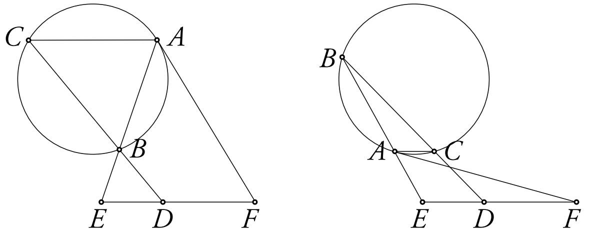


Figure 4: Other alternate configurations for Pappus *Collection* VII.105.

interpreted in this kind of way. My reading works whether this interpretation is granted or not. But let us suppose for the sake of argument that Pappus's proof should be read in this more flexible sense that includes the configurations of Figure 4, because this will allow us to highlight some general points about how far uniqueness requirements can be relaxed with the analysis still being able to perform the task I ascribe to it.

In general terms, we can state the logical schema of the analysis in my sense as follows. We are interested in linking the set of solution configurations S to the given initial configuration (which we may think of as a set G consisting of a single configuration). We do this by means of a set of auxiliary or intermediate configurations I. In our example we can define these sets as follows (cf. Table 1). The elements of the set of solutions S are geometrical configurations, so essentially the figures shown above. Except the solutions only concern what is mentioned in the problem, so we must delete the point F and its associated lines. The set G contains the given configuration, meaning the circle and the segment DE only. The set I can be defined as the set of configurations that contain the givens and an additional point F on the extension of the given line DE on the side of D such that $DE \cdot EF = ET^2$, where ET is a tangent to the circle. Allowing for Figure 4, we should consider all of these elements as concerning only the geometric entities, not their lettering. Thus, in our example, S and I have two elements (corresponding to the two possible ACs and the two possible Fs, respectively), while G has one.

The constructions involved in the analysis and synthesis are effectively mappings between these sets. A construction can be thought of as a function that takes one geometrical configuration as input and produces one or more geometrical configurations as output. In simple cases where there is a single unique solution, we typically have simple bijections $S \longleftrightarrow I \ \longleftrightarrow G$. But more generally we can think of the relations between these sets in terms of multivalued functions. Thus, for example, the transformation stage of the analysis shows that any solution $s \in S$ is associated with some auxiliary configuration $i \in I$, but the analysis does not care whether the construction that produces this i is uniquely determined or not. It could just as well involve ambiguous or case-splitting steps, such as taking the tangent to a circle from a point. So really the transformation is about a multivalued function $S \longrightarrow I$. The transformation shows that any solution $s \in S$ maps to some auxiliary configuration(s), that is, some subset of I.

The resolution then looks at a multivalued function in the other direction, $I \longrightarrow S$. Again, multivaluedness does not matter. The goal of the resolution is to prove that any $s \in S$ can be obtained from some $i \in I$. It does this by confirming that the s we started with—which was an arbitrary element of S—can indeed be recovered from I, that is, that this s is among the outputs of the multivalued function $I \longrightarrow S$ defined by the construction (the construction in the resolution is implied by the chain of givens; later in the synthesis the same construction is spelled out more explicitly). The analysis transformation already outlined a path connecting any $s \in S$ to some $i \in I$. It remains for the resolution to verify that this path is constructively traversable in the reverse direction, and that the construction covers all such paths. See Section 3 for more detail on how the machinery of "givens" contributes to this.

In the same way the resolution also confirms that the multivalued function $G \longrightarrow I$ defined by the construction in the synthesis does indeed cover all of I. Note indeed in our example that the first step of the construction is in effect the definition of I, as is very typical.[32]

These things taken together prove that the composite multivalued function $G \longrightarrow I \longrightarrow S$ is onto. Since the image of this composite multivalued function is C, by definition, this proves that all of S is contained in C, that is, that the synthesis produces all possible solutions. (Of course, these explications in terms of sets and functions must not be taken too literally. They are not meant as a rigorous formalisation of analysis-synthesis proofs altogether, but only as a schematic representation of how analysis-synthesis proofs can deal with sets of solutions and branching constructions and still serve the purpose I have attributed to them.)

This account shows the viability of conceiving analyses and syntheses as describing sets of solutions rather than individual solutions. As I have shown, in our example, there are two solutions that are both covered by the same analysis and synthesis, so the figure can be considered a particular case whereas the proof applies generally to the full set of solutions. By thinking in terms of sets of solutions in the manner I have outlined, this is handled just as easily as if there had been only one solution. Indeed, the existence of situations like this, where there are multiple solutions, is all the more reason to make sure that the set of solutions provided in the synthesis is exhaustive, which is exactly what the analysis does.

In conclusion, I have outlined a general account of how analysis can serve a particular formal function. I believe analyses in the Greek corpus are typically

32. An analogous example is Pappus *Collection* VII.72, where H is the point defining I. Also analogous are cases where I consists in a point defined as the intersection of conic sections, as in for example Pappus *Collection* IV.31.

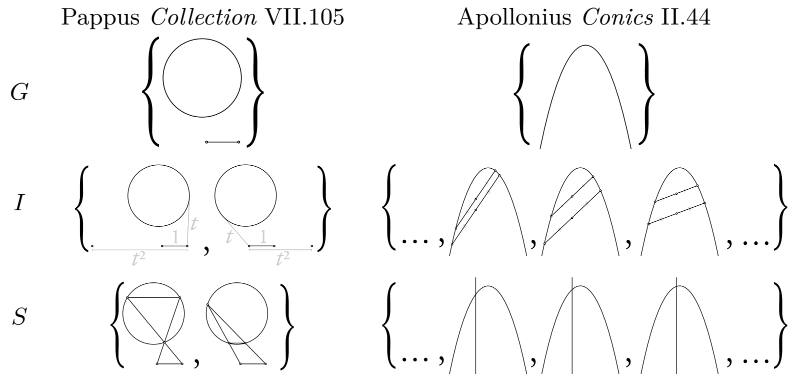


Table 1: The sets G, I, S in the case of two illustrative problems. Light-grey elements are explanatory notes, not themselves part of the set elements.

susceptible to such a reading. At the very least, I have shown that my reading works for Pappus *Collection* VII.105, which is a prototype example of an analysis-synthesis argument often expressly used as a generic illustration of the entire genre,[33] and a complex one in that there are non-trivially distinct solutions.

3 Role of the *Data* and meaning of "given"

The resolutions of analyses are characterised by inferences of the form "X is given, so Y is given". The precise meaning of "given" in Greek geometry has been the subject of much scholarly attention but has eluded simple characterisation. To understand the meaning of "given", the key source is Euclid's *Data*, which proves theorems of the form "if X is given then Y is given", and is hence the canonical source for establishing the validity of the inferences made in analysis resolutions. Just as the resolution of a problem analysis matches up step by step with the construction in the synthesis, so also the *Data* provides direct companions for many constructions in the *Elements*. For example, *Data* 39 ("if each of the sides of a triangle be given in magnitude, the triangle is given in form") corresponds to *Elements* I.22 ("to construct a triangle from three straight lines which are equal to three given [straight lines]"), and *Data* 90 ("if from a given point a straight line be drawn tangent to a circle given in position, the straight line drawn will be given in position

33. [Hankel 1874, 143], [Heath 1925, I.141], [Ito 1980, 10], [Fournarakis & Christianidis 2006, 48] all use VII.105 to illustrate analysis generally. [Behboud 1994, 57] uses the very similar variant proposition VII.107 in place of VII.105 as his prototype example.

and in magnitude") corresponds to *Elements* III.17 ("from a given point to draw a straight line touching a given circle").

It is evident that being "given" is very closely related to being "determined" in some sense. However, for the purposes of my account of analysis, it is essential that this not be confused with "uniquely determined". Being "given" does not entail uniqueness. Indeed, the proof of *Data* 39 follows *Elements* I.22 (which shows the possibility of constructing a triangle with three given sides), not I.7–8 (which proves that there is only one such triangle). *Data* 90 too is evidently not bothered by the fact that the object shown to be "given" is not uniquely determined by the conditions imposed on it (this is the ambiguity which corresponds precisely to the step at note 22 above).[34] This is necessary for my account, because, as I argued above, the logic of the analysis is best phrased in terms of sets of solutions. If "given" meant "uniquely determined", then this point of view would collapse.

Nevertheless, I propose that "given" does indeed mean "determined", but in the following sense. The key purpose of the concept of "given" for the purposes of my account of analysis is this: the *Data* version of a construction from the *Elements* confirms that the construction is done *without loss of generality*. On its own, a solution to a problem in the *Elements* produces an object that satisfies the problem posed, and verifies that it does so. As such, it leaves the question open whether this is the most general solution possible, or whether there could be other solutions missed by this construction. A key concern of the *Data*, I believe, is to alleviate this worry. The *Data* does this by verifying, anytime a construction involves some apparent specificity (such as placing a given length along a particular line, or constructing a tangent of a particular length), that this level of specificity is not gratuitous but rather forced upon us by the specifications given in the problem statement. The proofs of the *Data* check that the specificity assumed in each step of a construction is truly "given"—that is to say, implied by the conditions imposed by the statement of the problem itself—rather than a substantive restriction of the generality of the problem.

It is because the *Data* has done this work that a problem analysis such as that of the Pappus problem above can verify that the synthesis construction does not miss any solutions. Inferences of the form "X is given, so Y is given" means not only that Y is constructible from X, but also that this construction operates without loss of generality, so that it will not miss any such Y. Recall that, on my account, the goal of the resolution is to prove that $I \longrightarrow S$ is onto. It is for this purpose that it is important to know that each step of the

34. Another example of a similar nature is *Data* 31: "If from a given point a straight line given in magnitude be drawn to meet a straight line given in position, the line drawn is also given in position". Thus "either the proposition is false or 'given in position' does not mean 'unique'—unless we extend the property 'given in position' to cover *symmetrical* positions" [Taisbak 2003, 105]. I note that even if this very liberal notion of "uniqueness" is granted the argument against uniqueness based on *Data* 39 still stands.

construction does not introduce overspecificity, but rather solves its subtask with full generality. This ensures that the mapping $s \mapsto i$ is reversible since it rules out the possibility that the construction ran back up again starting from i “branches off” into a restricted subset of S.

4 Analysis as justification of means

I shall now argue that in some cases analysis serves a second rigorous function, in addition to proving exhaustiveness as above. This point of view comes into play when the solution assumes a construction method other than ruler and compass, or the Euclidean postulates.

Consider for example Archimedes’s *Sphere and Cylinder* II.1. I will summarise the basic structure of it in modern terms.[35] This is the problem: given a cylinder (say with diameter d and height h), to find a sphere (say with diameter x) with the same volume. Archimedes has already proved the basic volume relation, $\frac{3}{2}d^2h = x^3$, as a theorem earlier in the treatise (I.34). The problem therefore amounts to extracting a cube root (or, in more classical language, taking two mean proportionals). The structure of the analysis is in effect: suppose such a sphere has been found; then $x = \sqrt[3]{\frac{3}{2}d^2h}$. The structure of the synthesis is in effect: form $\frac{3}{2}d^2h$ and extract its cube root and make a sphere with this diameter; then this is the sought sphere.

In the synthesis, Archimedes simply stipulates that the cube root can be extracted (“let two mean proportionals be taken”). The proof of the synthesis shows that, if this is granted, then the problem can be solved. In the analysis, the implication goes the other way, so that together the analysis and synthesis show that not only can the problem be solved by taking mean proportionals, but the problem is in fact *equivalent* to doing so. This two-way implication aspect is known in the literature.[36]

But I shall go further and argue that this can be read as a justification of the tools assumed in the solution in that it serves to answer the possible objection that the synthesis has assumed too advanced a method. In a sense the analysis shows that we have not assumed too advanced means for the problem, and hence justifies the use of this additional construction rule. This may be called *analysis as justification of means.*

35. Thus I am using algebraic formulations whereas Archimedes of course writes purely geometrically. These differences are immaterial to my argument. A detailed exposition using both geometric and algebraic explications of Archimedes’s proofs is [Dijksterhuis 1987, chap. V].

36. [Berggren & Van Brummelen 2000, 8, 16]. Closely related to this is the idea that and analysis shows of what kind or degree a problem is, in the sense of something like Pappus’s classification of problems into plane, solids, and linear. It is evident that analyses of non-plane problems is linked to such questions, as has often been noted [e.g., Knorr 1975, 344], [Sefrin-Weis 2010, xxx, 188, 274, 293, 301, 305].

In Archimedes's case, nobody knew how to take two mean proportionals with ruler and compass,[37] so if the problem could in fact be solved by ruler and compass, then Archimedes could be criticised for assuming unnecessarily advanced methods where simple ones would do. But the analysis proves that Archimedes is completely safe from such a critique. For while the synthesis proves that the ability to extract cube roots implies the ability to solve the problem, the analysis shows that the ability to solve the problem implies the ability to extract cube roots. So if in fact someone would come along and solve the problem by ruler and compass, then the analysis shows that the extraction of cube roots too can be done by ruler and compass.[38] Hence Archimedes has proved rigorously that it is impossible to find a simpler construction than the one he has given, for any possible "simpler" solution must in fact entail the ability to extract cube roots and hence be logically on par with Archimedes's solution rather than simpler than it.

My point can be stated more explicitly as follows. Geometrical problems have the form: "given X, construct Y". To construct something means to reach it by applying a small set of approved construction rules. These rules can be thought of as problems of the same form as above—such as "given two points, construct a straight line joining them"—but whose solution is assumed rather than proved. Let E be the set of construction rules allowed in the *Elements*,[39] and let $P(E)$ be the set of all problems that can be solved by these means. Higher Greek geometry deals with problems that are not in $P(E)$. Their solution therefore involves some additional construction step, that introduces a point determined by means of neusis, conic sections, or higher curves such as the conchoid or quadratrix. Let p denote such a problem and let f denote the additional construction rule assumed for its solution.[40]

37. This is of course impossible, as we now know and as the Greeks in all likelihood strongly suspected.

38. In effect, by tacking the steps of the analysis on at the end of the ruler-and-compass solution of the problem. To turn this composite construction into a recipe for finding the cube root of any given x one only needs to restrict it to a suitable choice of cylinder, such as $d = 1$, $h = \frac{2}{3}x$.

39. This means essentially Postulates 1–3 and 5 of Book I, with some minor accommodations for informal assumptions that Euclid does not articulate explicitly. See, e.g., [Sidoli 2018], where, in particular, Euclid's postulates (419–424) and problems (426–431) are both interpreted as functions introducing certain geometrical entities.

40. We must understand f here as being restricted to the precise point-introduction rule used in the particular construction in question. For example, the Greeks knew how to duplicate a cube by conic sections. In modern terms, the intersection of the hyperbola $xy = 2$ and the parabola $y = x^2$ has x-coordinate $\sqrt[3]{2}$, which is the side of a cube with twice the volume of a unit cube. But we must not think of the assumption of this solution as being the ability to draw conic sections and taking their intersections. That would be much too general. All the solution really needs is the ability to produce a point that stands in a specific relation to the existing geometrical configuration already in place at that stage of the proof. Characterising this point as the intersection of two conics is a natural way of formulating and conceptualising this

Then solving p really means showing that $p \in P(E, f)$, that is, that p is in the set of all problems that can be solved by Euclidean means together with the additional rule f. Now, as we observed, problems and construction rules formally have the same structure and the only thing that separates one from the other is whether we have decided to accept it as a "primitive" given or not. Hence it makes just as much sense to think of f as a problem and p as a construction rule. I claim that this is in effect what the analysis does. In fact, the analysis shows that $f \in P(E, p)$.

One important purpose of doing this, I claim, is to justify the use of the additional construction rule f. Ideally, to justify the use of f, you would want to show that the problem cannot be solved by Euclidean means alone ($p \notin P(E)$). As is well known, results of this type for the main classical construction problems were proved only in the 19th century. Nevertheless, analysis can provide the next best thing. By proving $f \in P(E, p)$, the analysis shows that *any* solution of the problem implies the ability to do f. Thus, in a rigorous sense, we did not assume "too much" when we allowed ourselves to take f as given.

Another way of putting it is that the analysis shows that it is impossible to solve the problem by "simpler" means than f in a certain rigorous sense. For suppose the problem could be solved by ruler and compass, $p \in P(E)$. Then the analysis shows that f itself can be done by ruler and compass, $f \in P(E, p) = P(E)$. Hence when we assumed f we did in fact *not* assume anything beyond the reach of ruler and compass, and hence the supposedly "simpler" solution by ruler and compass is as a matter of fact not simpler at all (in the sense of being based on a more restrictive set of assumptions). In other words, although the analysis itself does not answer the question of whether or not $P(E) = P(E, f)$, the analysis *does* show that there is no set of assumptions $A \supseteq E$ such that $p \in P(A) \subsetneq P(E, f)$: there is no "simpler"—meaning more restrictive—set of assumptions that accomplishes the solution. This fits with a documented concern in the Greek tradition of hierarchy of methods issues. The methodological dictum that it is "not a small error for geometers" to solve a problem "from a non-kindred kind", that is, using more advanced assumptions or methods than necessary, is found in Pappus and seems to have been a prominent concern for Apollonius as well [Knorr 1975, 346], [Sefrin-Weis 2010, 145, 271–275], [Jones 1986, 530], [Knorr 1982, 272], [Acerbi 2018, 273, 276].

Consider for example Pappus *Collection* IV.31 (Figure 5). Everything in the analysis and synthesis of this problem proceeds by ruler and compass, except the introduction of the hyperbola. One might therefore object: yes, this solves the problem, but by advanced means—perhaps a simpler solution is possible, that does not require the ability to draw hyperbolas but only ruler

in the Greek framework, but this does not mean that the proof relies in an essential way on the ability to draw these conics, let alone any two conics. We shall see an example of this below.

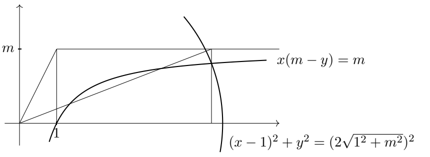


Figure 5: Angle trisection by conics.

and compass? Ideally, we would like to be able to show that such a thing is not impossible. Unfortunately the Greeks did not have the tools to do this. Nevertheless, with the help of the analysis we could still give at least some reply to the sceptical question. The first step in defending ourselves against such an objection could be to say: yes, the synthesis is best formulated in terms of drawing a conic section, but really all it actually requires is the ability to find one single point of intersection (the intersection of the hyperbola and the circle), not the ability to draw the entire curve. The analysis proves that *any* solution to the problem implies the ability to find that intersection point (in other words: if you can trisect the angle you can also recreate the intersection point by ruler and compass from there). Thus we have rigorously proved that it is impossible to solve the problem by more restrictive means than the ones we employ in the synthesis, for if any purportedly simpler solution was provided, it would entail our assumption and hence not be based on more restrictive assumptions after all.

In the notation introduced above, the additional construction rule f assumed in this case would be: given line segments 1 and m, to construct the point of intersection of $x(m-y)=m$ and $(x-1)^2+y^2=(2\sqrt{1^2+m^2})^2$. This formulation is agnostic as to how this point of intersection is found. The logical function of analysis that I have outlined is based on this f only, and hence does not capture the idea that the problem is solvable *by conics*, but rather focusses only on one very specific, technical construction principle. In this way my account of analysis misses what is arguably a fundamental point of the proof and others like it. This is indeed a drawback. The function of analysis that I have outlined in this section is not the greatest thing a geometer might wish for. It would be more interesting indeed to have systematic results characterising the scope of ruler and compass constructions, the scope of constructions by conics, and so on, in general terms. For example, Viète later showed that

$$\text{any cubic or quartic equation } \in P(E, p_1, p_2) \subseteq P(E, F)$$

where

$$p_1 = \text{cube duplication}$$
$$p_2 = \text{angle trisection}$$
$$F = \text{neusis}$$

But such results were not forthcoming in Greek times. The function of analysis I have outlined doesn't say anything about such "general scope" (of conics, neusis, etc.) questions. This is indeed another respect in which this function of analysis doesn't give the most complete answer to means-justification-questions. Nevertheless, the form of analysis I have highlighted is at any rate better than merely assuming the method f without justification (as would have been the case if the problem was solved by the synthesis alone).

There are a few rare cases in the Greek geometric corpus where neusis or conics are used to solve a certain problem even though it would have been possible to solve it purely with ruler and compass.[41] These problems do not come with analyses, but, if we tried to supply analyses for such cases, how would my account of analysis as justification of means play out? There are two possibilities: either f can be effected by ruler and compass, or it cannot. My analysis applies in either case.

Consider the case where f can be constructed by ruler and compass. Then an analysis is perfectly possible. The interpretation of analysis as a justification of means is applicable, but what it justifies is the use of f, not the use of neusis or conics. Those advanced methods may give a more compact way of expressing the step f, but that step in and of itself does not actually depend on those means. So in this case $P(E, f) = P(E) = P(E, p)$. As always, the analysis proves that $f \in P(E, p)$ and that there is no more restrictive set of assumptions that solves the problem, which is correct. One may feel that this analysis "misses the point", because it doesn't take into account that f was described in terms of conics or neusis. The analysis is indeed blind to this fact and would have said the exact same thing if f had been done by ruler and compass. This could be perceived as a drawback or limitation, since analysis in this sense is somewhat divorced from certain elements of the mathematical practice. On the other hand, one could argue that the analysis is doing the right thing in that it focusses on what the proof actually needed formally speaking, namely f, and not the manner in which f was described, which is arguably a cosmetic or psychological question.

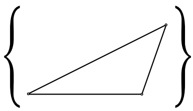
Next consider the case where f cannot be constructed by ruler and compass. Then the solution to the problem via f depends in an essential way on the extra means (conics, neusis, etc.), although there would be another path, not via f, that led to the solution without such means. So the means truly are not justified, and indeed the analysis cannot be used to prove that they are. For in this case $f \notin P(E, p) = P(E)$, so the analysis transformation can never arrive at f, and hence the analysis for this case can never be completed. (It is understood, of course, that in any analysis-synthesis pair,

41. [Dijksterhuis 1987, 138–139], [Heath 2002, c–ciii], [Heath 1921, 183, 196], [Heath 1896, cxxix]. All of these examples are from works where the overall purpose of the treatise means that a reduction to ruler and compass in those isolated cases would arguably have been pointless, since those works as a whole are concerned precisely with things that cannot be reduced to ruler and compass.

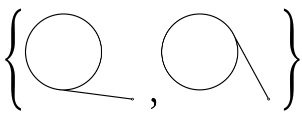
f must be the only non-ruler-and-compass step. Any analysis transformation uses only ruler-and-compass steps until it reaches the point corresponding to f, or, what is the same thing, to the intermediate configuration(s) I.) Hence it is impossible for the analysis to give a "false positive" justification of the means f in any such case.

5 Summary of the two purposes of analysis

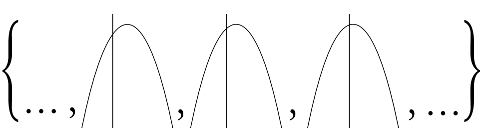
Any problem has a set S of solutions. For example, there is one triangle with three given sides (Euclid *Elements* I.22),[42]

$S =$  $|S| = 1$

two tangents to a given circle from a given point (Euclid *Elements* III.17),

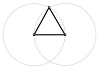
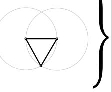
$S =$  $|S| = 2$

and infinitely many diameters for a given parabola (Apollonius *Conics* II.44),

$S =$  $|S| = \infty$

A solution to a problem is a construction recipe that produces some set of purported solution objects C. Constructions typically produce not just one object but a set of objects, due to case-splitting steps. For example, Euclid's construction of an equilateral triangle (*Elements* I.1) produces two of them:

42. Strictly speaking, whenever we speak of enumerations of elements such as these, we are assuming implicitly some notion of equivalence under which two figures are regarded as "the same". The *Data* has the necessary technical language for this: we speak of geometrical entities being "given in form" if we only care about the shape and not where it is located, but "given in position" if we distinguish between identical shapes differently located. I shall leave this technical issue implicit. But I note that my account of analysis, by being based on cardinalities of sets, naturally raises such issues. So this squares well with the technical attention to exactly such issues in the *Data*.

$C = \Bigg\{$  ,  $\Bigg\}$

The synthesis proves that constructed objects are solutions: $C \subseteq S$. But it leaves open the question: are they *all* the solutions? $C = S$? The analysis proves that all possible solutions are among those constructed: $C \supseteq S$. Hence the analysis-synthesis pair altogether proves that the construction is exhaustive: $C = S$.

The synthesis defines a construction $G \longrightarrow I \longrightarrow S$ from the set of the given configuration G, via some set of intermediate configurations I, to the set of solutions S. The analysis transformation $I \longleftarrow S$ shows that any solution is obtainable by "the kind of construction" that the synthesis uses. It remains (for the analysis resolution, i.e. chain of givens) to verify that the synthesis includes *all constructions of that kind*, i.e., that *no step limits the generality of the construction.* This is purpose of the machinery of "givens". "If X is given, then Y is given" means: We can construct Y from X in such a way that any specificity that Y has is unavoidable (i.e., a consequence of the the specificity of X). In other words, the construction does not introduce specificity or limit the generality of the problem.

The purpose of analysis is to show that the synthesis construction $G \longrightarrow S$ is *onto* (i.e., hits all of S). The analysis transformation verifies the onto-ness of the synthesis construction *globally* by showing that any solution is obtainable by the type of construction given in the synthesis, while the analysis resolution verifies the onto-ness of the individual steps of this construction *locally* by showing that no step introduces overspecificity.

A second purpose of analysis applies for non-ruler-and-compass problems. If a problem p is not solvable by ruler and compass, it is solved by some additional construction rule f (such as an application of conics, neusis, etc.). Problems and construction rules have the same form: "given X, construct Y". So it makes just as much sense to think of f as the problem and p as the construction rule. This is precisely what the analysis does.

Let $P(X)$ be the set of all problems solvable by means X, and let E be the set of Euclidean means (ruler and compass). The synthesis proves $p \in P(E, f)$. The analysis proves $f \in P(E, p)$.

Do we "assume too much" when we take f as given? The analysis enables us to answer: No. The analysis result $f \in P(E, p)$ means that *any* solution of the problem implies the ability to do f. Hence there is no set of assumptions $A \supseteq E$ such that $p \in P(A) \subsetneq P(E, f)$: there is no "simpler" (= more restrictive) set of assumptions that accomplishes the solution. In other words, since *any* set of assumptions sufficient for p are also sufficient for f, it is impossible to solve the problem by simpler methods (= methods whose scope includes p but not f).

The analysis does not show that the problem cannot be solved by Euclidean means alone ($p \notin P(E)$). But it still justifies the use of the means f in that:

either	the problem *is not* solvable by ruler and compass	$p \notin P(E)$
$\Longrightarrow$	additional means necessary	
$\Longrightarrow$	f is justified	
or	the problem *is* solvable by ruler and compass	$p \in P(E)$
$\Longrightarrow$	f can be done by ruler and compass	$f \in P(E,p) = P(E)$
$\Longrightarrow$	f doesn't assume anything beyond ruler and compass	$P(E,f) = P(E)$
$\Longrightarrow$	f is justified	

6 Conclusion

The two functions of analysis I have outlined have much in common. In terms of logic, they both start from the supposition that the problem has been solved because this is the proper logical starting point for a rigorous derivation of *properties that any solution must have*. In terms of purpose, they are both essential parts of a completely satisfying solution, since they rigorously dismiss certain critical objections, to which the solution would have been susceptible if the synthesis alone had been given.

Prominent aspects of the classical corpus that are not easily explained by other means make perfect sense on my reading. The functions of analysis I have highlighted have nothing to do with the notion of analysis as a method of discovery, so they explain why analyses occur in sparsely worded formal treatises and even where heuristic or pedagogic motives seem inapplicable. They also explain why analyses are particularly relevant to problems, and hardly at all to theorems. Indeed, my interpretation fits well with the documented concern for constructions and the scope of particular construction methods in Greek geometry.

If my account is correct, the famous description of analysis by Pappus—the only surviving methodological description of analysis in ancient sources of any detail—is inadequate and incomplete. Pappus states, correctly, that analysis starts by assuming that a solution has been affected and proceeds to reason from this assumption onwards. But he fails to highlight what is in my view the reason for doing so, and instead overemphasises, in my view, the exploratory, heuristic purpose of analysis.

But there are good grounds for not according Pappus's passage too much weight. "Pappus is often careless and sometimes demonstrably in error [...] about earlier Greek geometry" [Toomer 1976, 14], and his mathematical credentials are far from impeccable [Knorr 1975, 339], [Knorr 1989, 31, 226–228], [Knorr 1982, 253], [Jones 1986, 1], [Neugebauer 1975, 968]. His description of analysis is vague and seemingly incoherent, to the extent that a number of scholars have felt that interpretations assuming textual corruption is the only way to make coherent sense of it [Berggren & Van Brummelen 2000, 8]. And, in what may be a product of Platonic bias [Knorr 1975, 360], he seemingly construes analysis as applicable to problems and theorems in equal measure, which is at odds not only with my view but with the Greek geometrical record generally. Furthermore, Pappus was far removed from the leading geometers who had actually done the real work, writing many hundreds of years after the composition of all the works that by his own reckoning constitute the "treasury of analysis". Finally, Pappus most likely did not intend his brief characterisation of analysis to be complete, but rather as "remarks to introduce a broad field expounded in a number of treatises that he expected his reader to be able to read" [Saito & Sidoli 2010, 583]. For these reasons I do not see it as a problem that my account of analysis does not align perfectly with Pappus's description.

Bibliography

Acerbi, Fabio [2011], The language of the Givens: Its forms and its use as a deductive tool in Greek mathematics, *Archive for History of Exact Sciences*, 65(2), 119–153, doi: 10.1007/s00407-010-0072-z.

—— [2018], Hellenistic mathematics, in: *The Oxford Handbook of Science and Medicine in the Classical World*, edited by P. T. Keyser & J. Scarborough, New York: Oxford University Press, chap. 3, 269–298, doi: 10.1093/oxfordhb/9780199734146.013.69.

Apollonius [1987], *On Cutting Off a Ratio*, Fairfield: Golden Hind Press, ed. by E. M. Macierowski & R. H. Schmidt.

—— [2013], *Conics Books I–IV*, Santa Fe: Green Lion Press, ed. by C. R. Taliaferro & M. N. Fried.

Behboud, Ali [1994], Greek geometrical analysis, *Centaurus*, 37(1), 52–86, doi: 10.1111/j.1600-0498.1994.tb00002.x.

Berggren, Lennart J. & Van Brummelen, Glen [2000], The role and development of geometric analysis and synthesis in Ancient Greece and Medieval Islam, in: *Ancient & Medieval Traditions in the Exact Sciences: Essays in Memory of Wilbur Knorr*, edited by P. Suppes, J. M. Moravcsik, & H. Mendell, Stanford, CA: CSLI Publications, 1–31.

Dijksterhuis, Eduard Jan [1987], *Archimedes*, Princeton: Princeton University Press, with a new bibliographic essay by W. R. Knorr. (Original Dutch edition in 1938).

Fournarakis, Philippos & Christianidis, Jean [2006], Greek geometrical analysis: A new interpretation through the "givens"-terminology, *Bollettino di Storia delle Scienze Matematiche*, 26(1), 33–56.

Hankel, Hermann [1874], *Zur Geschichte der Mathematik in Alterthum und Mittelalter*, Leipzig: Teubner.

Heath, Thomas L. [1896], *Apollonius of Perga: Treatise on Conic Sections*, Cambridge: Cambridge University Press.

—— [1921], *A History of Greek Mathematics*, vol. I: From Thales to Euclid, Oxford: Clarendon Press, 2nd edn.

—— [1925], *Euclid: The Thirteen Books of the Elements*, Cambridge: Cambridge University Press.

—— [2002], *The Works of Archimedes*, Mineola: Dover Publications.

Hintikka, Jaakko & Remes, Unto (eds.) [1974], *The Method of Analysis: Its Geometrical Origins and Its General Significance*, *Boston Studies in the Philosophy of Science*, vol. XXV, Dordrecht; Boston: Reidel.

Ito, Shuntaro [1980], *The Medieval Latin Translation of the Data of Euclid*, Tokyo: University of Tokyo Press.

Jones, Alexander [1986], *Pappus of Alexandria: Book 7 of the Collection*, New York: Springer.

Knorr, Wilbur R. [1975], *The Ancient Tradition of Geometric Problems*, Boston: Birkhäuser.

—— [1982], The hyperbola-construction in the *Conics*, Book II: ancient variations on a theorem of Apollonius, *Centaurus*, 25(4), 253–291, doi: 10.1111/j.1600-0498.1981.tb00647.x.

—— [1989], *Textual Studies in Ancient and Medieval Geometry*, Boston: Birkhäuser.

Menn, Stephen [2002], Plato and the method of analysis, *Phronesis*, 47(3), 193–223, doi: 10.2307/4182698.

Netz, Reviel [2000], Why did Greek mathematicians publish their analyses?, in: *Ancient & Medieval Traditions in the Exact Sciences: Essays in Memory of Wilbur Knorr*, edited by P. Suppes, J. M. Moravcsik, & H. Mendell, Stanford, CA: CSLI Publications, 139–157.

—— [2004], *The Works of Archimedes*, vol. 1, Cambridge: Cambridge University Press.

NEUGEBAUER, Otto [1975], *A History of Ancient Mathematical Astronomy*, New York: Springer.

PROCLUS [1970], *A Commentary of the First Book of Euclid's Elements*, Princeton: Princeton University Press.

RASHED, Roshdi & BELLOSTA, Hélène [2010], *Apollonius de Perge, La section des droites selon des rapports*, Berlin: Walter de Gruyter.

ROBINSON, Richard [1936], Analysis in Greek Geometry, *Mind*, 45(180), 464–473, doi: 10.1093/mind/XLV.180.464.

SAITO, Ken & SIDOLI, Nathan [2010], The function of diorism in ancient Greek analysis, *Historia Mathematica*, 37(4), 579–614, doi: 10.1016/j.hm.2009.12.003.

SEFRIN-WEIS, Heike [2010], *Pappus of Alexandria: Book 4 of the Collection*, London: Springer.

SIDOLI, Nathan [2018], Uses of construction in problems and theorems in Euclid's *Elements* I–VI, *Archive for History of Exact Sciences*, 72(4), 403–452, doi: 10.1007/s00407-018-0212-4.

SIDOLI, Nathan & ISAHAYA, Yoichi [2018], *Thābit ibn Qurra's Restoration of Euclid's Data*, Cham: Springer.

SIDOLI, Nathan & SAITO, Ken [2012], Comparative analysis in Greek geometry, *Historia Mathematica*, 39(1), 1–33, doi: 10.1016/j.hm.2011.09.002.

TAISBAK, Christian Marinus [2003], *Euclid's Data*, Copenhagen: Museum Tusculanum Press.

TOOMER, Gerald J. [1976], *Diocles on Burning Mirrors*, Berlin: Springer.

ZEUTHEN, Hieronymus Georg [1886], *Die Lehre von den Kegelschnitten im Altertum*, Copenhagen: A. F. Höst & Sohn.

—— [1896], *Geschichte der Mathematik im Altertum und Mittelalter*, Copenhagen: A. F. Höst & Sohn.