What More There Is in Early-Modern Algebra than its Literal Formalism*

Marco Panza
CNRS, IHPST
(CNRS and University of Paris 1, ENS, Paris)

August 27, 2010

1 Introduction

There are two views about early-modern algebra very often endorsed (either explicitly or implicitly). The former is that in early-modern age, algebra and geometry were different branches of mathematics and provided alternative solutions for many problems. The latter is that early-modern algebra essentially resulted from the adoption of a new literal formalism. My present purpose is to question the latter. In doing that, I shall also implicitly undermine the former[1].

I shall do it by considering a single example. This is the example of a classical problem. More in particular, I shall consider and compare different ways of understanding and solving this problem. Under the first understating I shall consider, this problem appears as that of cutting a given segment in extreme and mean ratio. This is what proposition VI.30 of the *Elements* requires. In section 2, I shall expound and discuss Euclid's solution of this proposition[2]. Then, in section 3, I shall consider other propositions of the same

*The present paper is based on a talk which, besides to the Ghent's conference *Philosophical Aspects of Symbolic Reasoning in Early Modern Science and Mathematics*, I also presented in some other occasion, including the Oberwolfach meeting *History and Philosophy of Mathematical Notations and Symbolism*, held in October 2009. I thank all the people in the attendances, for different questions and comments. I'm particularly grateful, for useful discussions and advises to Hourya Benis-Sinaceur, Karine Chemla, Massimo Galuzzi, Albrecht Heeffer, Antoni Malet, Jeff Oaks, and Roshdi Rashed.

[1]For a more articulated argumentation against this former view, cf. my [11]. The term 'early-modern algebra' is quite vague and open to many different understandings. What I mean by it is essentially the practice of dealing both with arithmetical and geometrical problems through a common approach mainly originated by Viète's and Descartes's achievements. Hence, early-modern algebra has to be conceived, in my parlance, as being essentially about geometrical concerns.

[2]According to its common use in geometry, the term 'solution' is equivocal: it can respectively denote the way a problem is solved, the argument that one relies on to solve a problem, the action of solving a problem, or even the object (or objects) whose construction is required by a problem . To avoid unfamiliar

Elements, and argue that they suggest another, quite different understanding of the same problem. Under this other understanding, this is viewed as the problem that of constructing a segment meeting a certain condition relative to another given segment. One way to express this condition is by stating the first of the three trinomial equations studied in al-Khwārizmī's *Algebra*, by supposing that this equation is geometrically understood and a particular case of it is considered[3]. In section 4, I shall consider this option, by focusing in particular, on Thābit ibn Qurra's interpretation and solution of this equation.

The transition from the former understanding to the latter hinges on the transition from a way of conceiving geometrical magnitudes and the relative problems and theorems, to another way of doing the same[4]. Whereas the former conception is proper to Euclid's geometry, the latter is proper to early-modern algebra. My main point will be that this latter conception is independent of the appeal to any literal formalism, as it is showed by the fact that it is already at work in Medieval Arabic geometry, in which there is no trace of such a formalism. Early modern algebra actually resulted from combining this conception with an appropriate use of a symbolic notation, which gave raise to the new literal formalism. In my view, this use was made possible by the adoption of this conception, but it was not merely a natural outcome of it. Hence, though early-modern algebra could have not development without the adoption of this conception, the latter cannot be reduced to the former. The purpose of section 5, the last one of my paper, will be that of discussing this connection.

---

phrases, in what follows I shall have no other option than using this term in these different senses, in different occasions. I hope the context will be enough to avoid any misunderstanding. In particular, when I shall claim that the problem I shall consider admits a unique solution, if it is conceived as Euclid does, I shall not mean that there is only one way to solve it, or only one argument allowing to solve it, within Euclid's geometry. I shall rather mean that, supposing that a certain segment be given, in order to solve this problem so conceived, one has to construct a certain unique object. In other terms, my point will be that, this segment being given, each way of solving this problem so conceived, or each argument through which it is solved, have to result in the construction of the same point on this segment.

[3] In my [11], I preferred avoiding to use the term 'equation' to refer to conditions like those that al-Khwārizmī's and, after it, al-Khayyām's *Algebra* ([1]; [14]), are about. I used instead the term 'equations-like problem' to refer to the problems associated to these conditions. My reason whose that, for al-Khwārizmī and al-Khayyām, stating such a condition was not the same as presenting a mathematical object, as it is the case when a polynomial equation is presented in the context of the formalism of early-modern algebra ([11], p. 124). I still think that this reason is good. But, for short and simplicity, I do not follow here my previous convention and I call al-Khwārizmī's conditions 'equations', as it is usually done. I hope the previous remark to be enough for avoiding any misunderstanding: when this term is used to refer to these conditions, it has not to be taken to have the same sense as when it is used in the context of early-modern algebra.

[4] A quite similar transition (perhaps the same one, also if conceptualized in a partial different way) is described by R. Netz in his [7]. About the convergences and divergences of Netz's and my views on this matter, cf. [11], p. 117 (footnote 46). Also Netz's discussion is focused on a single problem. This is significantly more complex that the one I shall consider here. Namely, differently from this latter one, it is not solvable within Euclid's geometry, that is, through a construction by ruler and compass (or elementary construction, according to the parlance I adopted in [12]). Far from undermining the effectiveness of my example, the relative straightforwardness of the problem I shall rather consider is intended to make clear that the shift I want to describe does not depend on the adoption of any mathematical tools exceeding the simple ones proper to Euclid's geometry, but was just a mere shift in conception.

2 Proposition VI.30 of the *Elements*

Proposition VI.30 of the *Elements* is a problem. It requires "to cut a given segment in extreme and mean ratio"[5]. Let AB be the given segment (fig. 1). The problem requires, in other terms, to construct a point E on it, such that

$$\mathsf{AB} : \mathsf{AE} = \mathsf{AE} : \mathsf{EB}. \tag{1}$$

Euclid's solution relies on that of proposition VI.29. This a problem in turn, and requires "to apply to a given segment a parallelogram equal to a given rectilineal figure with excess of [another] parallelogram similar to a [third] given one". If the given segment is CA, the given rectilineal figure is Γ, and the given parallelogram is Δ (fig. 2)[6], the problem consists in producing CA up to a point G such that the parallelograms CFDG and AEDG be respectively equal to Γ and similar to Δ.

To solve this last problem, Euclid begins by taking the middle point I of CA, and constructing on IA the parallelogram IALJ similar to Δ. Then, he constructs a parallelogram OPQK similar to IALJ and Δ, and equal to Γ and the same IALJ taken together. At this point, it is enough to produce IJ up to a point N such that NJ = OK, and then to complete the parallelogram NDMJ in such a way that its diagonal DJ be collineal with the diagonal AJ of IALJ, and to produce CA up to cut in G the side DM of this parallelogram. The point G will be the searched after one[7].

To construct the middle point of a given segment is easy. According to the solution of propositions I.1 and I.9-10, it is enough to describe two circles having this segment as a common radius and its extremities as the respective centres, and to join their intersection points (fig. 3).

It is also easy to construct a parallelogram similar to another given parallelogram on a given segment. Consider the case under examination. The segment IA being given (fig. 4), it is enough: to take on IA a point R such that RA be equal to the side ra of Δ; to construct on RA (according to the solution of proposition I.22) a triangle RAS having its sides equal to the sides of the triangle ras formed by tracing the diagonal rs of Δ (so that $\widehat{\mathsf{RAS}} = \widehat{\mathsf{ras}}$, according to proposition I.8 and the solution of proposition I.23); to trace the parallel IT to RS through I; to produce AS up to meet this parallel in L; and to complete the parallelogram IALJ.

[5]For my quotations from the *Elements*, I base on Heath's translation ([3]), though I slightly modify it somewhere.

[6]For the purpose of denoting with the same letters the points with play an analogues role in the following constructions related to the solution of proposition VI.30, I change some of the letters in Euclid's diagrams as they appear in [3].

[7]The proof is simple. By construction, the equalities

$$\mathsf{OPQK} = \Gamma + \mathsf{IALJ} = \mathsf{NDMJ}$$

hold. Hence Γ is equal to the gnomon NDMLAI. But, as I is the middle point of CA, and NDMJ is similar to IALJ, NEAI is equal both to FNIC and to AGML. Hence FNIC is equal to AGML, and then FDGC is equal to the gnomon NDMLAI, and consequently to Γ. On the other side, IALJ is similar to Δ by construction, to the effect that also EDGA is so.

Hence, the only critical step in the solution of proposition I.29 is the construction of the parallelogram OPQK (fig. 2). Euclid does not detail this construction, but it is easy to see that, in order to perform it, one has firstly to construct a rectilineal figure equal to Γ and IALJ taken together, and then apply the solution of proposition IV.25.

This can be done in different ways, depending on the nature of such a rectilineal figure. An obvious possibility is to take this figure to be the parallelogram VALU (fig. 5) which is got by constructing on IJ and in the angle $\widehat{\mathsf{CIJ}}$ the parallelogram VIJU equal to Γ.

This construction goes as follows, according to propositions I.42, and I.44-45. If Γ is not a triangle, divide it into several (non-overlapping) triangles by tracing appropriate segments joining two non-consecutive vertexes of it. Construct the middle point i, of a side fg of one of these triangles fgh. With i as vertex, construct the angle $\widehat{\mathsf{fij}}$ equal to $\widehat{\mathsf{CIJ}}$ (which can be done as explained above, according to the solution of propositions I.23) Through h, trace the parallel mk to fg meeting in m and k the parallel to ij through f and ij itself, respectively). The parallelogram fikm is equal to the triangle fgh, and its internal angle $\widehat{\mathsf{fik}}$ is, by construction, equal to the angle $\widehat{\mathsf{CIJ}}$. It is then easy to construct another parallelogram JWXY equal to fikm, and consequently to fgh, and having the sides JY and JW collineal to IJ and JL, respectively (it is enough to produce IJ up to a point Y such that JY = fm and take a point W on JL such that JW = fi). Produce then XW up to meet CA in Z and complete the parallelograms ZXnp and IJop in such a way that the diagonal of the former be collineal to the diagonal ZJ of the other parallelogram IZWJ. The parallelogram pIJo is equal to JWXY, and then to fgh. By repeating the same construction for all the triangles composing Γ, one gets then, step by step, the sought after parallelogram VIJU.

Once this last parallelogram have been constructed, the construction of the parallelogram OPQK (fig. 6) goes as follows, according to the solution of proposition VI.25. Proceeding in the same way as in the construction of the parallelogram VIJU, construct, on the side rb of Δ and in the angle $\widehat{\mathsf{erb}}$ equal to $\widehat{\mathsf{ras}}$, the parallelogram drbc equal to VALU. According to the solution of proposition IV.13, construct a segment OP mean proportional between dr and ra (supposing that tO = dr and Oq = ra, and that u is the middle point of tq, OP is the perpendicular through O to tq up the the circle of centre u and diameter tq). Then construct the two angles $\widehat{\mathsf{KOP}}$ and $\widehat{\mathsf{OPK}}$ equal to $\widehat{\mathsf{bra}}$ and $\widehat{\mathsf{rab}}$, respectively, and complete the triangle OPK and the parallelogram OPQK.

At this point, the solution of proposition VI.30 is easy. It includes two steps. The former consists in constructing the square ABHC on the given segment AB (fig. 7). The latter consists in applying to the side AC of this square a rectangle equal to this same square with excess of a parallelogram similar to such a square, *i. e.* with excess of a square. This reduces to construct, according to the solution of proposition VI.29, the rectangle GDFC equal to the square ABHC and such that AG = GD. The sought after point is the intersection point E of the side DF of GDFC and the given segment AB[8].

When applied to the case considered in proposition VI.30, the solution of proposition

[8]The proof is easy. As GDFC = ABHC and AEFC is a common part of both, GDEA = EBHF. Hence, according to proposition I.14:

$$\mathsf{GA} : \mathsf{EF} = \mathsf{EB} : \mathsf{AE},$$

VI.29 is simplified with respect to the general case expounded above. This is because the role of both the rectilineal figure Γ and the parallelogram Δ is played by the same figure, and this figure is a square, namely the square ABHC. Let us see how this solution applies in this case.

Let the square ABHC be given (fig. 8). Construct the middle point I of AC, and, on AI, the square LAIJ. According to the general procedure expounded above, one should then construct the rectangle JIVU equal to ABHC. There would be a quite simple way to do it without following the general step-by-step procedure expounded above: insofar as ABHC can be divided in four squares equal to LAIJ, it would be enough to produce LJ up to a point U such that JU be equal to 4 times LJ, and then complete the rectangle JIVU (which would be the same as constructing four squares equal to LAIJ on each other). But this is useless, in fact. Insofar as the role of the parallelogram Δ is played by the square ABHC, one can take, indeed, the side CH of this square to coincide with the side rb of Δ (compare fig. 8 with fig. 6) to the effect that, after having constructed the rectangle JIVU, one should construct on CH $\equiv$ rb the rectangle rbcd equal to LAVU. Now, as this last rectangle would be, for construction, equal to LAIJ and ABHC taken together, this step can easily be performed directly, without relying on the construction of the rectangle JIVU. It is enough to construct on CH $\equiv$ rb the square rbwv equal to ABHC, then produce Cv up to the point d such that vd be equal to the half of AI, and finally complete the rectangle rbcd. The next step consists in constructing a segment OP mean proportional between rd and the homologous side of Δ. But, insofar as Δ is nothing but the square ABHC, a coincide with A and this side is then nothing but AC. Hence, what has to be constructed is a segment OP mean proportional between rd and AC $\equiv$ ar. The more natural way to do it, in order to get a geometrical configuration analogous to that which enters the solution of proposition VI.29, is to perform this construction on the segment qt collineal to LJ, such that qL = LJ and Jt = rd (to the effect that qt = Ad), by taking the point O to coincide with the point J. If u is the middle point of qt, to construct the point P, it is then enough to trace the semicircle of centre u and diameter qt, and the perpendicular OP to qt through J up to this semicircle. If one proceeds this way, the construction of the square OPQK equal to CHcd is useless, since the point P directly coincide with the point N (compare now fig. 8 with fig. 2 and fig. 7), the point D is got as the intersection point of the perpendicular to JN through N and the diagonal AJ of the square LAIJ produced, and the point G is got by completing the rectangle GDNI. The sough after point E is the intersection point of the side DN of this rectangle and the given segment AB (the construction of the point M which complete the rectangle MDNJ is then useless).

Insofar as ABHC is a square, AI is equal to the half of the given segment AB, and LAIJ is equal to a square having the half of such a given segment as side. Hence, the rectangle rbcd is equal both to a square having the given segment as side taken together with a square having the half of such a given segment as side, and to a rectangle having as sides

---

that is:

$$\mathsf{AB} : \mathsf{AE} = \mathsf{AE} : \mathsf{EB},$$

as required.

the given segment and a segment equal to this same given segment taken together with another segment equal to a fourth of it. This is also the case of the square OPKQ. Its side OP is then the side of a square equal to such a last rectangle. It follows that the segment AE results from cutting off the segment LA, equal to an half of the given segment, from the segment LE equal to a side of a square equal to this same rectangle.

This description is cumbersome. Still, the adoption of an appropriate notation allows rephrasing it in a quite simple way. Call the given segment 'a'. For any segment α, denote then with '$\frac{\alpha}{n}$' (where 'n' denotes, in turn, a natural number greater than 1) any segment equal to one n-th of α, and with '$S(\alpha)$' any square having as side a segment equal to α. For any another segment β, denote then with '$R(\alpha,\beta)$' any rectangle having as sides two segments equal to α and β, respectively. Use moreover the sign '$+$' and '$-$' to indicate the requirements of taking two geometric objects together and of cutting off one of them from the other, respectively

Using this notation, one can then write the following equalities:

$$\begin{aligned}
\mathsf{AB} &= a;\\
\mathsf{AI} &= \frac{a}{2};\\
\mathsf{LAIJ} &= S\left(\frac{a}{2}\right);\\
\mathsf{rbcd} &= S(a) + S\left(\frac{a}{2}\right) = R\left(a, a+\frac{a}{4}\right) = \mathsf{OPQK}.
\end{aligned}$$

Let now a^* be a segment such that

$$R\left(a, a+\frac{a}{4}\right) = S(a^*).$$

Then:

$$\mathsf{OPQK} = S(a^*) \quad ; \quad \mathsf{OP} = a^* \quad ; \quad \mathsf{AE} = a^* - \frac{a}{2}.$$

Euclid's construction as a whole can then be rephrased through three quite simple equalities:

$$\mathsf{AB} = a \quad ; \quad R\left(a, a+\frac{a}{4}\right) = S(a^*) \quad ; \quad \mathsf{AE} = a^* - \frac{a}{2}. \tag{2}$$

Mathematically, the possibility of this rephrasing leaves no doubt. Still, this is not the same as admitting that this rephrasing is appropriate for the purpose of describing Euclid's solution as it is, in fact. If this were so, it would be natural to wonder why Euclid expounds such a solution in a relatively so complex way, when he could have described it through such three simple equalities or, at least, through a vernacular reformulation of them, for example as follows: construct any segment equal to the given one and a fourth of it taken together; construct any square equal to the rectangle having segments equal to this last segment and to the given one as sides (which is the same as finding a mean proportional between two sides of this last rectangle); cut off a segment equal to an half of the given segment from a segment equal to a side of this square; this results in a segment meeting the condition of the problem.

This question is ill-stated, however. Since behind such a rephrasing of Euclid's solution—and, *a fortiori*, behind that consisting of equalities (2)—there is an understanding of this solution that is quite far from Euclid's.

I call 'purely quantitative' a geometric problem that asks for the construction of whatever segments (or whatever geometric objects of a certain sort, which are determined if appropriate segments are so), which are supposed to meet some conditions depending only on the relative size of these segments and of other given ones, regardless to their respective positions. In other terms, such a problem asks for the construction of whatever segments belonging to the equivalence classes of segments which are specified by stating these conditions (the relevant equivalence relation being equality). Let a be any given segment. Hence, one would be stating purely quantitative problems by requiring, for example, to construct a segment x such that

$$a : x = x : a - x, \tag{3}$$

or that

$$R\,(x, a + x) = S(a), \tag{4}$$

whatever the given segment a and the positional relations of x and a might be. But this is not what propositions VI.30 and VI.29 require. They are rather, I would way, positional problems, each of which have a unique possible solution (or a finite number of possible symmetric solutions)[9].

Proposition VI.30 is a problem of partition of a given segment. It requires constructing a certain point on such a given segment, and is then solved if and only if this very point is constructed. According to a classical terminology, proposition VI.29 is a problem of application of an area with excess of a parallelogram similar to a given one. Though classical, this denomination is potentially misleading, however. Strictly speaking, Euclid is not asking to construct any parallelogram having a certain area (whatever an area might be for him). He is rather requiring to construct a certain parallelogram having a base collineal with a given segment and a vertex coincident with an extremity of such a segment.

To be more precise, in proposition VI.30, the segment AB being taken as given (fig. 1), Euclid does not require to construct any segment x smaller then AB and equal to a mean proportional between a segment equal to AB itself and another segment equal to that which have to be joined to x itself for getting a segment equal to AB. He rather requires to construct a certain determinate point on AB, namely the point E (which is a much more natural requirement to be advanced in his language). Analogously, in proposition VI.29, the segment CA being taken as given (fig. 2), Euclid does not require to construct any parallelogram equal to Γ suitable for being divided into two parallelograms, one of which has a base equal to CA and the other is similar to Δ[10]. He rather requires to

[9] Cf. footnote (2), above.

[10] Notice that if Euclid's proposal, in stating proposition VI.29, had been that of advancing a purely quantitative problem, it would have been pointless to take Γ and Δ to be any polygon and any parallelogram, respectively. This would have involved no gain of generality with respect to taking Γ and Δ to be a square and a rectangle, respectively.

construct a certain determinate parallelogram, having a vertex in C and a base collineal to CA, namely the parallelogram FDGC (or at least one of the four symmetric parallelograms FDGC, CGD′F′, D″EAG′, and G′AE′D‴: fig. 9).

Now, the previous rephrasing of the solution of proposition VI.30 respects its being concerned with the construction of a certain determinate point on the given segment, namely with the point E on AB (fig. 1) But, within the context of this rephrasing, the fact that the object to be constructed be a point on the given segment, rather than any segment equal to $a^* - \frac{a}{2}$, appears as an immaterial and extrinsic constraint. Such a rephrasing is based, indeed, on the understanding of the successive steps of the solution as responses to purely quantitative sub-problems, which is clearly not Euclid's understanding.

It is true that the segment OP (fig. 6) entering the solution of proposition VI.29 is supposed to be any mean proportional between any two segments equal to dr and ra, respectively Still, this appears to be more a trick for avoiding a too intricate diagram than an intrinsic feature of this solution. Once the segment OP and the associate parallelogram OPQK are constructed, Euclid immediately constructs the segment JN, equal to the side OK of this parallelogram in its appropriate position (fig. 2). In other terms, he is certainly concerned with a purely quantitative condition, but he understands it as an ingredient of the positional problems he is interested to.

The previous rephrasing of Euclid's own solution of proposition VI.30 is then unfaithful to such a solution. But is it also, merely, a pointless modernization of Euclid's argument? In the following part of my paper I shall show that this is not so, in fact.

3 Comparing Proposition VI.30 with Other Proposition of the *Elements*

If proposition VI.30 is compared with proposition I.3 of the same *Elements*, it becomes natural to understand the former as requiring to divide the given segments into two segments one of which results from cutting off the other from the given one. Using, as above, the sign '−' to indicate the requirement of cutting off a segment from another, one could then rephrase the condition (1) as follows

$$\mathsf{AB} : \mathsf{AE} = \mathsf{AE} : \mathsf{AB} - \mathsf{AE}. \qquad (5)$$

This rephrasing is not innocent. Since, it displaces the focus of the problem from the construction of the point E into the construction of a segment AE having a certain relation with the given segment AB. Once this change of focus is admitted, the requirement that the sought after segment be taken on AB, that is, that it a segment AE, its extremity E being on AB appears as an extrinsic constraint: as a pointless specification of a purely quantitive condition like

$$\mathsf{AB} : x = x : \mathsf{AB} - x, \qquad (6)$$

which only differ from condition (3) for the way as the given segment is denoted. The passage from (5) to (6) or (3) is not anodyne, however, since it results in transforming the positional problem stated by proposition VI.30 into a purely quantitative one.

This is a radical change in my mind. But it does not depend on the acquisition of some new mathematical resources with respect to Euclid's (the crucial identification of the segment EB with the result of cutting off AE from AB is openly licensed and also suggested by proposition I.3, as said). It merely depends on a change of focus, or, more generally, a shift in conception.

One could object that also the use of a literal notation and of an operational arithmetical sign like '$-$'—*i. e.* of terms like 'a', 'x', and '$a-x$'—is relevant. In other terms, one could object that condition (3) and the problem connected to it are significantly different from condition (1) and proposition VI.30 for a reason that hinges essentially on the fact that the former are stated by using a literal notation and an operational arithmetical sign like '$-$'. This is wrong, however, and depends on a confusion (which is not less misleading for its being recurrent). This is the confusion between the semantic function of these terms and their syntactical function[11].

The apparent plausibility on the objection depends on the fact that the terms 'a','x' and '$a-x$' are implicitly supposed to be involved in the whole formalism of early-modern algebra. If it were so, these terms would be also used for their syntactical function, to the effect, for example, that the condition (3) should be taken to be *ipso facto* equivalent to the condition

$$x^2 + ax = a^2. \tag{7}$$

But nothing like this is implied in the passage from (1) to (3), as I have described it above. In (3), the terms 'a','x' and '$a-x$' are only intended to denote some segments (either given or sought after), that is, they are merely used for their semantic function[12].

[11]My use of the two adjectives 'semantic' and 'syntactical' in order to account for the distinction I'm willing to point out is far from mandatory. I could have used another terminology, instead. The advantage of my choice is that these adjectives immediately evoke some opposite features of language that this distinction is concerned with. Still, there are many other aspects of the meaning (or meanings) usually ascribed to these adjectives which are not relevant for my present purpose. By using these adjectives, I do not intended to evoke them. In section 5 below,.I shall come back on this distinction, by trying to clarify it further

[12]Notice however that there is an essential difference between the symbol 'a' on one side, and the symbols 'x', and '$a-x$', as well as the symbols '$\frac{a}{2}$', '$\frac{a}{4}$', '$a+\frac{a}{2}$', '$a+\frac{a}{4}$', 'a^*', '$a^*-\frac{a}{2}$', '$S(a)$', '$S\left(\frac{a}{2}\right)$','$S\left(a+\frac{a}{2}\right)$','$R\left(a,a+\frac{a}{4}\right)$','$S(a^*)$', '$R(x,a+x)$' used above, on the other side. The symbol 'a' is intended to be a proper name of a certain segment, and is then, on this respect, analogous to terms like 'AB', 'AE', 'EB', or even 'ABHC' or 'CGDF' that are just used by Euclid as proper names of segments and polygons, respectively. The symbols 'x', '$a-x$', '$\frac{a}{2}$', '$\frac{a}{4}$', etc. are intended to be, instead, non-definite descriptions denoting any element of a certain class of equivalence of geometrical objects of the appropriate sort (the relevant equivalence relation being equality). As a consequence, there is an essential difference between the meaning to be assigned to the sign '$=$' in formulas like '$\mathsf{AB}=a$' where this sign stands between two proper names and in formulas like '$\mathsf{AI}=\frac{a}{2}$', '$\mathsf{IJLA}=S\left(\frac{a}{2}\right)$' or '$S(a)+S\left(a+\frac{a}{2}\right)=R\left(a,a+\frac{a}{4}\right)$' where it stands between a proper name and a non-definite description, or between two non-definite descriptions. In the former case, this symbol indicates identity; in the latter it merely indicates equality. These differences reflect a distinctive feature of purely quantitative problems: the fact that they involve one or more principal or independent segments, in terms of which the segments to be constructed are characterized, and that these segments are taken to be given as such at the beginning of the construction that is intended to solve the problem, whereas the segments constructed during this construction are (for the very nature

Hence, this last condition is not significantly distinct from the requirement that the segment to be constructed be a mean proportional between a segment equal to the given one and another segment equal to that resulting from cutting a segment equal to it off from a segment equal to the given one. On the other hand, this condition is not equivalent to the condition (7), since the passage from (3) to (7) depends on a formalism that goes not necessarily together with the use of the terms 'a','x' and '$a - x$' according to their semantic function.

It follows that, according to their intended function (which is here only the semantic one), the use of terms like 'a','x' and '$a-x$' is nothing more but a convenient linguistic trick for stating a purely quantitative problem. This use allows to avoid cumbersome vernacular expressions in favor of shorter, nimbler and unambiguous ones, and could thus help for going ahead in complex reasoning, or for expounding complex arguments in a simplified and clearer way. But it is in no way essential for the purpose of stating such a purely quantitative problem. Hence, the passage from proposition VI.30 to a purely quantitative problem (which is the passage I focus my attention on) does not depend on the use of such terms, that is, on the use of a literal notation and of an operational arithmetical sign like '$-$'.

This being said, consider now proposition II.2 of the same *Elements*. It is a theorem: "if a segment is cut at random, the rectangle contained by the whole and each of the [two] parts is equal to the square on the whole". The expression 'the rectangle contained by the whole and each of the two parts' refers to the rectangle contained by the whole and one of the two parts and the rectangle contained by the whole and the other part taken together. Supposing that the segment AB is cut at random in C[13] (fig. 10), Euclid takes his theorem to state, indeed, that "the rectangle contained by AB, BC together with the rectangle contained by BA, AC is equal to the square on AB". To prove that this is so, he constructs the square DEBA, traces the perpendicular FC to AB through C, and takes the rectangles DFCA and FEBC to be "the rectangle contained by BA, AC" and "that contained by AB, BC", respectively, since AD = AB = BE. In other terms, supposing that α, β, and γ be three distinct segments, Euclid takes the rectangle "contained by" α, β to be the same as (and not just equal to) the rectangle "contained by" α, γ, if $\beta = \gamma$ (which is what allows him giving sense to the notion of rectangle "contained by" two collineal segments). This suggests that he takes expressions like 'the rectangle contained by α, β' and 'the square on β' to refer to any rectangle having as sides two segments equal to α and β, and to any square having as side a segment equal to α, respectively. These are purely quantitative notions: they do not depend on the mutual positions of the relevant segments, but only on the fact that some of them are equal to others.

Despite this, proposition II.2 appears as a positional theorem: a theorem whose content depends on the mutual position of the objects it is concerned with. Since, it asserts something about whatever segment split up into two other segments by any point on it,

of the problem) whatever elements of certain classes of equivalence of segments defined in terms of the given ones.

[13] I come back now to using the same letters as in Euclid's diagrams as they appear in [3].

namely that any pair of rectangles one side of which is equal to the former segment, while the other side is respectively equal to the two parts of it, are, if taken together, equal to any square having this former segment as side.

This understanding is reinforced by Euclid's proof of this theorem, which is entirely diagrammatic. The segment AB being considered, the crucial step in in this proof is the claim that the rectangles DFCA and FEBC taken together are equal to the square DEBA. As this claim is not further justified, it is obvious to reconstruct Euclid's argument as follows: the rectangles DFCA and FEBC taken together coincide with the square DEBA; but, according to common notion I.4, "things which coincide with one another are equal to one another"; hence, the rectangles DFCA and FEBC taken together are equal to the square DEBA; still, these rectangles are those contained by BA, AC and by AB, BC, respectively; then these last rectangles taken together are equal to the square DEBA. This argument crucially depends on the mutual position of the rectangles DFCA and FEBC and the square DEBA. And this is just what reinforces the understanding of proposition II.2 as a positional theorem.

Still, there is no doubt that for Euclid, any rectangle is equal to any other rectangle having equal sides (this is a particular case of proposition I.35), and any pair of segments respectively equal to the two parts into which a third segment is split up are, if taken together, equal to this last segment and to any other segment equal to it. Hence, the same argument also proves the following implication: if two segments taken together are equal to a third one, then any pair of rectangles having as sides two segments equal to this third one and to the two others, respectively, are, if taken together, equal to any square having as side a segment equal to the third one.

This is no more a positional theorem, however. Using the same notation as before, it can be rephrased as follows:

$$\text{if} \quad a = b + c \quad \text{then} \quad R(a, b) + R(a, c) = S(a). \tag{8}$$

Moreover, if three segments a, b and c are such that $b+c = a$, then if a segment equal to b is cut off from a segment equal to a, a segment equal to c results. Hence, from implication (8) it follows that

$$R(a, b) + R(a, a - b) = S(a). \tag{9}$$

Both implication (8) and equality (9) are purely quantitative theorems: they are about some geometric objects which are supposed to meet some conditions depending only on the relative size of these same objects, regardless to their respective positions (*i. e.*, they are about the equivalence classes of these geometric objects that are specified by stating these conditions, the relevant equivalence relation being equality).

One can resist, of course, in understanding proposition II.2 as being equivalent to these theorems, and, then, in ascribing these theorems to Euclid. Still, without admitting that, for Euclid, the rectangle contained by two segments α and β is the same as the rectangle contained by α and γ, if $\beta = \gamma$, the very statement and proof of proposition II.2 are hard to understand. And, once this is admitted, the understanding of this proposition as being equivalent to implication (8) and equality (9) becomes very natural. Nevertheless, I do not want to argue that this understanding is Euclid's. For my present purpose, it is enough to have fixed such an understanding and to have shown that it is at least suggested by the very way Euclid expresses himself and reasons.

Consider now proposition VI.16 of the *Elements*, again, in particular its first part: "if four segments are proportional, the rectangle contained by the extremes is equal to the rectangle contained by the means". The terms 'the rectangle contained by the extremes' and 'rectangle contained by the means' are subject to the same considerations relative to the analogous terms occurring in proposition II.2 and its proof. In proving proposition VI.16, Euclid considers four given segments, AB, CD, E, and F (fig. 11) such

$$\mathsf{AB} : \mathsf{CD} = \mathsf{E} : \mathsf{F}$$

and construct, on AB and CD, respectively, two rectangles ABKG and CDLH, such that $\mathsf{AG} = \mathsf{F}$ and $\mathsf{CH} = \mathsf{E}$ then he reasons on these last rectangles. It is then clear, again, he takes the rectangle contained by two segments α and β to be the same the same as the rectangle contained by α and γ, if $\beta = \gamma$. It is thus natural to understand proposition VI.16 as stating the following theorem: if four segments are proportional, then any rectangle having as sides two segments equal to the extremes is equal to any rectangle having as sides two segments equal to the means. In the same notation as before:

$$\text{if} \quad a : b = c : d \quad \text{then} \quad R(a, d) = R(b, c).$$

Now, by comparing this implication with (3), one gets

$$R(a, a - x) = S(x).$$

Hence, from (9), it follows:

$$S(x) + R(a, x) = S(a). \tag{10}$$

If the previous rephrasing of propositions II.2, VI.16 and VI.30 is admitted, this last proposition can be reduced, then, to the problem of constructing a segment x such that this last equality holds, supposing that the segment a is given.

Compare now this equality with the construction that provides Euclid's solution of this proposition. If AB (fig. 7) is the given segment a, then AE satisfies the former: if one takes AE to be a value of x, then the rectangle AEFC and the square GDEA are values of $R(a, x)$ and $S(x)$, respectively, and, if they are taken together, they are equal to the square ABHC which is a value of $S(a)$.

My terminology I use requires an explanation. As I use them, the terms 'x', '$R(a, x)$', '$S(x)$', and '$S(a)$' are not singular, that is, they do not respectively refer to a unique object, but rather to whatever element of a certain equivalence class of objects. By saying that a certain segment is a value of x, I mean that this segment is an element of the equivalence class to whatever element of which 'x' refers (as the relevant equivalence relation is equality, it follows that, if α and β are two values of x, then $\alpha = \beta$). The same convention apply to '$R(a, x)$', '$S(x)$', and '$S(a)$' and to any other terms like those: by saying, for example that a certain square is a value of $S(a)$, I mean that it is a square having a segment equal to a as side.

This explains the relation between the equality (10) and the geometrical configuration that Euclid's solution of proposition VI.30 depends on: this configuration provides a positional model for this equality. Other models are possible, however. And this is not only because one can construct the point E—providing the unique solution of this proposition, according to Euclid's understanding—in many other ways. But also and overall because, even if the given segment is always the same—that is, a—there are infinite many segments other than AE which satisfy this equality as values of x[14]. All of them provide a solution of the purely quantitative problem of constructing a segment x such that equality (10) holds.

All these remarks are not enough for solving such a problem. And, it is neither enough for this purpose to observe that, according to implication (8), from equality (10), equality (4) follows, and that this last equality is such that constructing a segment x that satisfies it is the same as solving a problem of application of an area, understood as a purely quantitative problem (namely the problem of application of an area whose solution is involved in Euclid's solution of proposition VI.30). What is still lacking is constructing a segment suitable for providing a value of x. An appropriate formulation of the problem can help, however, in identifying a positional model suggesting a simple construction of such a segment. To say the same in another way: in order to look for a solution of the problem, one can try to reduce it to some other simpler problem, which can be done, in turn, by appropriately transforming the condition that this problem is concerned with.

4 Thābit ibn Qurra's interpretation of al-Khwārizmī's first trinomial equation

One such transformation is implied by Thābit ibn Qurra's solution of al-Khwārizmī's first trinomial equation. Expounding this treatment is a way of showing that the rephrasing of problems like proposition VI.30 as purely quantitative problems, and the appeal to appropriate other propositions of the *Elements*, understood in turn as purely quantitative theorems, for justifying such a rephrasing are not merely pleasant possibilities suggested to a modern reader by certain aspects of Euclid's text and arguments, but were already part of mathematical practice several centuries before early-modern algebra developed. This is the purpose of the present section.

Al-Khwārizmī's *Book of Algebra and al-Muqābala*[15] deals with a combinatorial system including three basic elements, or modes: "Numbers", "Roots", and "Squares". These

[14] Cf. footnote (12), above.

[15] Al-Khwārizmī's treatise was probably written in the first part of 9th century. I base my analysis on Rashed's edition and French translation of it ([5]). This is accompanied by a rich introduction which, among other things, stresses the connection between Al-Khwārizmī's algebra, Thābit's treatment of his trinomial equations, and the tradition of the *Elements*. Tough my views are not always the same as Rashed's, much of what I shall say on these matters has been inspired by what he argues for in this introduction. For a much earlier edition of al-Khwārizmī's treatise, together with an English translation, cf. [4].

elements are combined so as to get six equations: three are binomial equations, the three other trinomial equations[16]. For short (using a notation which is not al-Khwārizmī's), one could write them as follows:

$$S = R \quad ; \quad S = N \quad ; \quad R = N \quad ;$$
$$S + R = N \quad ; \quad S + N = R \quad ; \quad R + N = S$$

(where 'N' refers to a Number, 'R' to some Roots, and 'S' to some Squares; the meaning to assign to the sign '+' will be clarified later).

These equations can be firstly understood as (shortenings of) statements of general canonical problems about Numbers, Roots, and Squares. As these equations exhaust the possible combinations admitted by the system (supposing that the commutativity of '+' is taken for granted, and repetitions are not allowed), these are all the problems of this sort that can be stated. Still, these same equations can also be understood as possible forms of different problems about what Numbers, Roots, and Squares are taken to stay for. In this case, the fact that these equations exhaust the possible combinations admitted by the system merely entails that these problems can have six different forms, that is, that al-Khwārizmī's algebra is concerned with all the problems that can take one of these forms, under an appropriate interpretation of Numbers, Roots, and Squares.

But what Numbers, Roots, and Squares can stay for? In other terms, which are their possible interpretations? These are two: Numbers, Roots, and Squares can both stay for numbers, in the usual sense of this term, or for geometrical magnitudes, namely for rectangles whose sides are taken to be equal to certain segments which are either given or to be constructed.

This being said, let us stop with generalities about al-Khwārizmī's algebra. On one side, I do not want to say too much that I shall not have room to justify (through a detailed analysis of al-Khwārizmī's arguments). On the other side, the consideration of a single example is enough for my present purpose and will also allow to clarify some of the previous generalities.

This is the example of the general canonical problem corresponding to the first trinomial equation, that is, $S + R = N$, or, in al-Khwārizmī's own parlance: "The Squares and the Roots equal to a Number" ([5], pp. 100-101).

Al-Khwārizmī describes its solution by expounding his well-known arithmetical algorithm in the particular case where one Square and ten Roots are equal to thirty-nine (dirhams). The consideration of this particular case is enough for making the arithmetic interpretation manifest. In this particular case, the problem consists in determining a number (which al-Khwārizmī takes to be a number of dirhams) such that by adding its square to the number resulting by multiplying it for ten one gets thirty-nine. In modern terms, one has then to solve the quadratic equation:

$$x^2 + nx = m$$

[16]Cf. footnote (3), above.

where $n = 10$, $m = 39$ and x is taken to be a number[17]. (which implies that the Square is nothing but a product of numbers).

In order to justify this algorithm, he shifts to geometry, however ([5], pp. 108-113). Namely, he offers two different positional models for the problem. In both of them, the Square is interpreted as being the square GDEA (fig. 12 and 13a)[18] and its sides are interpreted as being equal to one Root. In the former, the four small squares Γ, Δ, Θ, Λ (fig. 12) are then taken to have sides equal to a fourth of the number of Roots; the square GDEA and the four rectangles Π, Σ, Φ, Ψ taken together are taken to be equal to the Number; and the problem is conceived as that of constructing a side of the square GDEA, supposing that the sides of the squares Γ, Δ, Θ, Λ are equal to a fourth of a given segment, and the square GDEA and the four rectangles Π, Σ, Φ, Ψ taken together are equal to a given rectangle or square. In the latter, instead, the square Γ (fig. 13a) is taken to have sides equal to an half of the number of Roots; the square GDEA and the two rectangles Π, Ψ taken together are taken to be equal to the Number; and the problem is conceived as that of constructing a side of the square GDEA, supposing that the sides of the square Γ are equal to an half of a given segment, and the square GDEA and the two rectangles Π, Ψ taken together are equal to a given rectangle or square.

Let a and b be to a given segments. Using the previous notation, the condition to be met, as interpreted in the former and in the latter model respectively, could be written as follows:

$$S(x) + 4R(\frac{a}{4}, x) = S(b) \quad \text{and} \quad S(x) + 2R(\frac{a}{2}, x) = S(b). \qquad (11)$$

Insofar as it is natural to conceive the four rectangles Π, Σ, Φ, Ψ in the former model (fig. 12), and the two rectangles Π, Ψ in the latter (fig. 13a) as being equal respectively to the four and the two equal parts in which it is divided a rectangle having a side equal to one Root and another side equal to the number of Roots, these two conditions appear as appropriate rephrasings of the same condition:

$$S(x) + R(a, x) = S(b). \qquad (12)$$

Under the geometric interpretation, the problem appears then to consist in constructing a segment x such that this condition obtains, provided that a and b are given segments: in the particular example considered, a is supposed to measure 10 (unities of length), and b is supposed to be such that any square having a side equal to it measures 39 (unities of surface).

I do not want to argue that al-Khwārizmī actually understands his problem this way (I would have not have space for justifying this claim, and it is not essential for my purpose). I

[17] I cannot enter here the quite delicate question relative to the suppositions that one should make about the nature of this number, which is of course a crucial question to be addressed for clarifying the arithmetical interpretation of al-Khwārizmī's general canonical problems (but is not relevant for my present purpose). For a discussion of the notion of number in Arabic algebra, cf., for example, [8], § 2, and [9], § 5.

[18] For reasons of uniformity with respect to the letters used to denote points in the diagrams relative to Euclid's proposition VI.30 , I change the letters in al-Khwārizmī's diagrams as they appear in the French translation included in [5], pp. 110, 112.

limit myself to remark that this understanding is compatible with his geometric arguments. These arguments goes as follows.

In both models, the largest square has sides equal to one Root plus the half of the number of Roots. In the former model, it is also equal to the Number (which is equal to the square GDEA and the four rectangles Π, Σ, Φ, Ψ taken together: fig. 12) plus four squares whose sides are equal to a fourth of the number of Roots (the squares Γ, Δ, Θ, Λ). In the latter model, it is also equal to the Number (which is equal to the square GDEA and the two rectangles Π, Ψ taken together: fig. 13a) plus a square whose sides are equal to an half of the number of Roots. Since, four squares whose sides are equal to a fourth of the number of Roots are equal, if taken together, to a square whose sides are equal to an half of the number of Roots, in both cases, the Root is found by retracing an half of the number of Roots from the side of a square equal to the Number plus a square whose sides are equal to an half of the number of Roots. In the example considered, it is then equal to the side of a square equal to $39 + 25 = 64$ (unities of surface), *i. e.* 8 (unities of length), minus the side of a square equal to 25 (unities of surface), *i. e.* 5 (unities of length), which is the same as that which is prescribed by the arithmetic algorithm.

Under the previous understanding, this is the same as observing that condition (12) is equivalent to the conditions

$$S(b) + 4S\left(\frac{a}{4}\right) = S(x + \frac{a}{2}) \quad \text{and} \quad S(b) + S\left(\frac{a}{2}\right) = S(x + \frac{a}{2}). \tag{13}$$

The problem is thus reduced to two other problems whose solution is quite simple: a and b being given, construct a square equal to a square whose sides are equal to b taken together with a square whose sides are equal to an half of a; then cut off a segment also equal to an half of a from a side of this square.

Now, it is enough to suppose that a is the same segment as b, to reduce condition (12) to condition (10). Hence, al-Khwārizmī's handling of his first trinomial equation suggest reducing the problem stated in Euclid's proposition VI.30, understood as a purely quantitative problem, to the problem of constructing a segment x such that

$$S(a) + 4S\left(\frac{a}{4}\right) = S(x + \frac{a}{2}) \quad \text{or} \quad S(a) + S\left(\frac{a}{2}\right) = S(x + \frac{a}{2}). \tag{14}$$

This reduction is different in nature from that through which I passed above from proportion (3) to equality (10). Whereas this last reduction is justified by a mere inspection of the purely quantitative condition that the problem requires to meet, the former is justified by the inspection of two positional models of the original problem. In al-Khwārizmī's argument, condition (10) is firstly rephrased under the forms of conditions (11)[19]; these two

[19] This passage depends in fact on the replacement of $R(a, x)$ with $4R(\frac{a}{4}, x)$ and $2R(\frac{a}{2}, x)$, respectively, which is not justified by an inspection of the two positional models, but rather suggests them. A justification of this replacement could then depend on an appeal to proposition II.1 of the *Elements* understood as a purely quantitative theorem, in analogy with the understanding suggested above for proposition II.2. Al-Khwārizmī does not make, however, any explicit reference to the *Elements*, and it is far from sure that this reference is implicit for him. Proposition II.1 is the following: "If there be two segments and one

conditions suggests then two positional models; the inspection of these models suggests two new purely quantitative conditions, namely conditions (13); finally conditions (14) are got by supposing that segment a be the same as segment b[20].

In Thābit's solution of al-Khwārizmī's equation, this difference disappears. The occasion for expounding this solution is a short treatise, composed around the middle of 9th century: *Restoration of algebraic problems through geometrical demonstrations*[21]. The aim of this treatise is just that of making manifest the geometric interpretation of al-Khwārizmī's trinomial equations: this is the restoration that its title refers to. I consider of course only Thābit's handling of the first of these equations.

Thābit opens his argument by referring to a quite simple diagram (fig. 14). In it, the Square is interpreted as being the square DBAC[22]. The segment EB and the rectangle GEBD are then taken to be equal, respectively, to the number of Roots and to the Roots themselves. Hence, the whole rectangle GEAC is equal to the Square and the Roots taken together, and then to the Number.

In order to fix the interpretation of the segment EB and the rectangle GEBD as being respectively equal to the number of Roots and to the Roots themselves, Thābit requires that EB be "how many times the unity by which lines are measured as the supposed number of Roots", and remarks that "the product of AB and the unity by which lines are measured is the Root", to the effect that "the product of AB and BE is equal to the Roots" ([13], p. 160-161). It seems then that Thābit intends his diagram as a geometric illustration of the problem arising from the arithmetical interpretation of al-Khwārizmī's equation. Once having offered this illustration, he remarks, however, that the problem associated to this equation reduces to "a known geometric problem": supposing the segment EB to be

of them be cut into any number of parts whatever, the rectangle contained by the two [whole] segments is equal to the rectangles contained by the uncut segment and each of the parts [in which that other segment is cut]." Understood as a purely quantitative theorem, this proposition can be rephrased through the following implication

$$\text{If} \quad b = c + d + \ldots + e \quad \text{then} \quad R(a,c) + R(a,d) + \ldots + R(a,e) = R(a,b),$$

where a and b are any two given segments. Implication (8) can clearly be taken as a particular case of this last one.

[20]Notice that, in the case where segment a be the same as segment b, it is enough to take a configuration symmetric to that corresponding to the latter model (fig. 13*b*) to get back to the configuration constituted by the square MDNJ (fig. 8) entering Euclid solution of proposition VI.30. The argument leading to this configuration is however quite different in the two cases.

[21]I base my analysis on Rashed's edition and French translation ([13], pp. 159-169), which is accompanied by an introduction stressing the connection between Thābit's treatise and al-Khwārizmī's (*ibid.* pp. 153-158). The same translation is also integrally quoted (with the exception of the two first introductory lines) in [4], pp. 33-41, where the same connection is also stressed, as I have already said in footnote (15). For an earlier edition of Thābit's treatise, together with a German translation, cf. [6]. A quite similar argument is also advanced in Abū Kāmil's *Algebra* dating back to the end of 9th century. An English translation of the relevant passage (together with the appropriate references to the original text and his critical editions) in included in Appendix B of [9].

[22]Apart for the addition of 'G', the letters are now those of Thābit's diagram, as it appears in the French translation included in [13], p. 162.

"know", one "joins" the segment BA to it, and one supposes that the "product of EA and BA is known"; the problem consists then in determining the segment BA.

Supposing that α and β are two collineal segments, like BA and EB, or BA and EA, Thābit's expression 'the product of α and β' seems to be intended to have the same reference as Euclid's expression 'the rectangle contained by α, β', namely any rectangle having having as sides two segments equal to α and β, respectively. This suggests two things. The former is that Thābit conceives the "known geometric problem" (which is, of course, perfectly independent of the arithmetic interpretation of al-Khwārizmī's equation) as a purely quantitative problem: supposing that a is a given segment (to which BE is supposed to be equal), construct a segment x such that $R(x, a + x)$ is equal to a given rectangle or square. The latter thing is that Thābit is willing to solve this problem by appealing to appropriate propositions of book II of the *Elements*, understood as a purely quantitative theorems, in turn. Both things are confirmed by the way Thābit continues.

Before coming to that, another remark is appropriate. The geometrical configuration represented by Thābit's diagram is the same as that which proposition II.3 of the *Elements* is concerned with. This is a theorem, again: "if a segment is cut at random, the rectangle contained by the whole and one of the [two] parts is equal to the rectangle contained by the [two] parts and the square on the aforesaid part". The same considerations made above about proposition II.2, also apply in this case[23], and suggest an analogous understanding of the former proposition as a purely quantitative theorem:

$$\text{if} \quad a = b + c \quad \text{then} \quad R(a, b) = R(b, c) + S(b),$$

from which it immediately follows that

$$S(b) + R(b, c) = R(b, b + c),$$

or, by appropriate replacements,

$$S(x) + R(a, x) = R(x, a + x). \tag{15}$$

Hence, it is enough to apply proposition II.3 so understood, in order to rephrase condition (12) as:

$$R(x, a + x) = S(b), \tag{16}$$

which just corresponds to the "known geometric problem" to which Thābit reduces al-Khwārizmī's equation, this problem being understood as a purely quantitative one, and is nothing but a generalization of condition (4).

There is thus room for advancing that Thābit's diagram is intended as an implicit reference to proposition II.3 used in order to reduce al-Khwārizmī's equation, interpreted under the form of condition (12), to such a known geometric problem. This is all the more plausible that this diagram does not play any further role in the continuation of Thābit's argument, which is now time to consider.

[23] Euclid's proof of proposition II.3 is perfectly analogous to his proof of proposition II.2.

He explicitly appeals to proposition II.6 of the *Elements* by claiming that, according to it, if F is the middle point of EB, then the "the product of EA and AB" and the "square of BF" taken together are equal to the "square of AF" ([13], p. 162-163).

Proposition II.6 is also a theorem: "if a segment is cut in half and a segment is joined to it in a straight line, the rectangle contained by the whole [segment] with the joined segment and the joined segment, taken together with the square on the half, is equal to the square on the [segment] composed by the half and the joined segment". To prove it, Euclid considers a segment AB (fig. 15) cut in half at C and produced up to a point D. Then he takes: the "rectangle contained by the whole [segment] with the joined segment and the joined segment" to be the "rectangle contained by AD, DB", which he identifies, in turn, with the rectangle KMDA (supposing that MD = BD); the "square on the half" to be the "square on CB", which he identifies, in turn, with the square EGHL (supposing that EL = CB); the "square on the [segment] composed by the half and the joined segment" to be the "square on CD", which he identifies, in turn, with the square EFDC. What the proposition asserts is thus that the "rectangle contained by AD, DB", *i. e.* the rectangle KMDA, taken together with the "square on CB", *i. e.* the square EGHL, is equal to "square on CD", *i. e* the square EFDC. The proof is then very easy, since to get it, it is enough to remark that the rectangle KLCA is equal to the rectangle GFMH.

Thābit limits then himself to apply Euclid's proposition to the configuration constituted by the segment EA (fig. 14) and the points B and F on it, and to re-states it as such, by replacing Euclid's expressions 'the rectangle contained by −, − ' and 'the square on −' with 'the product of − and − ' and 'the square of −'. What is important, however, is the use he does of this proposition. As in the "known geometric problem" to which al-Khwārizmī's equation has been reduced, the product of EA and BA (or rectangle by EA, BA) is supposed to be known, and this is also the case of the square of (or on) FB, since this last segment is equal to an half of a given segment, from proposition II.6, it follows that also the square of (or on) FA is known, which entails that FA itself is known. To solve the problem is then enough to cut off FB, from FA, since the result of this is just BA, which comes then to be known.

This argument goes together with no explicit construction, and, its only correlate on the diagram is the display of point F. It seems then that proposition II.6 is used as a sort of rule of inference, independent of any construction or positional model, and allowing a further reduction of the problem[24].

For Euclid's proof of this proposition to work, the square EGH (fig. 15) has to be taken to be the square on CB. This is a further clue for understanding Euclid's expressions of the form 'the square of α' and 'the rectangle by α, β', and the corresponding Thābit's one of the form 'the square on α' and 'the product of α and β' as referring to any square having as side a segment equal to α, and to any rectangle having as sides two segments equal to α and β, respectively. The fact that Thābit feels no need to represent the squares of (or on)FB and FA, together with his shift from the former expressions to the latter ones make evident that he just understands these expressions this way.

[24]The exact sense in which I speak here of rule of inference will be clarified in section 5, below, p. 22.

In my notation, his understanding of proposition II.6 is then expressed by rephrasing it under the form of the following equality

$$R(a+b,b)+S\left(\frac{a}{2}\right)=S\left(\frac{a}{2}+b\right),$$

or, by appropriate replacements,

$$R(x,a+x)+S\left(\frac{a}{2}\right)=S\left(\frac{a}{2}+x\right). \tag{17}$$

His argument reduces then to a comparison of this last equality with the equality (16), so as to get the second of the two equalities (13), to which the problem is finally reduced.

The conclusion of this argument clearly parallels that which al-Khwārizmī reaches through the consideration of the second of his geometrical models. Still, it is justified by the inspection of no geometrical model, and merely depends on an application of proposition II.6, understood as a purely quantitative theorem used as a rule of inference.

5 Early-Modern Algebra and Purely Quantitative Theorems and Problems

As observed before, if a is taken to be the same segment as b, condition (12), and then al-Khwārizmī's equation understood as a geometric problem, reduces to condition (10), and the second of the two equalities (13) is transformed in the second of the two equalities (14). Hence, when applied to the problem stated by proposition VI.30, Thābit's argument suggests a quite easy way to solve this problem: according to proposition I.47 (the Pythagorean theorem), it is enough to cut the given segment AB in its middle point C (fig. 16), to construct the right angled rectangle ABD, whose sides BD is equal to CB, or an half of AB, and to cut off a segment equal to such a side from the hypotenuse of this triangle. The segment AF so constructed satisfies condition (10) for $a =$ AB: if IA = AF, the square AFHG and the rectangle IJBA taken together are equal to the square KLBA. Hence, if the point E is taken on AB so that AE = AF, AB is cut in extreme and mean ratio, as required.

This provides, in the same time, the (unique) solution of proposition VI.30, and a quite simple positional model for the purely quantitative problem associated to it, whose solutions include, beside AE, also, AF, AG, AI, GH, FH, BJ, and any other segment equal to them. The relevant point here is that this easy way of solving proposition VI.30 only appears if this proposition is converted into such a purely quantitative problem, and this last problem is then reduced to an easier one by relying on proposition II.6, understood in turn as a purely quantitative theorem used as a rule of inference. Furthermore, also to prove that the solution that is got this way is appropriate, one has to rely either on this last proposition so understood, or to some analogous theorem.

In an extreme synthesis, my point is that the passage from Euclid's understanding of proposition VI.30 and his way to solve it, to its conversion into a purely quantitative problem, which is solved in turn through an appeal to a purely quantitative theorem used as a rule of inference, manifest a structural feature of early-modern algebra, which is both essential to it (if this is used in geometry) and independent of any literal formalism.

Of course, there is much more in early-modern algebra. Still, this much more is, in my sense, dependent on such an understanding of geometrical problems and theorem, and on this way of solving them through appropriate reductions which are not suggested by the consideration of appropriate positional models of them, but rather suggest these models. In other terms, this is a necessary condition that made possible the development of early-modern algebra (as it is used in geometry) and the adoption of its literal formalism.

So, a natural question arises: what more has been needed for the shaping of early-modern algebra? Or better: what there is more in it than in an argument like Thābit's? Suppose to transcribe this argument using a skillful notation, like that I have used above. This would be certainly not enough to reproduce the early-modern algebraic setting. But, why not? What would be still lacking?

Let me to conclude my paper by sketching an answer to these questions.

According H. Bos, early modern geometry presents a peculiar form of problematic analysis that is made possible by the introduction of the algebraic formalism, especially thank to Viète, and that he calls 'analysis by algebra' ([2], pp. 97-98). Its crucial difference from Pappusian analysis is, for Bos, that it requires rephrasing the problems to be solved through a system of equations written in such a formalism. There is no doubt that the development of early-modern algebra goes together with the adoption of a non-Pappusian kind of analysis, which is clearly described by Viète in the *Isagoge* ([15]). Still, it seems to me that Vietian analysis is a particular form of a more general kind of analysis, crucially different from Pappus', which is, as such, perfectly independent of Viète's formalism, and depends, instead, on the conceptual outcome I have described above.

Elsewhere ([11]), I have called 'trans-configurational' this kind of analysis, since it consists in transforming the configuration of data and unknowns that comes with the problem to be solved, into another configuration. The starting configuration is fixed through a system of purely quantitative conditions that known and unknown magnitudes are supposed to meet, then analysis operates on this system of conditions through appropriate rules of inference, so as to transform it in a new equivalent, but essential different system of conditions, expressing a new configuration of data and unknowns. Typically, this final system of conditions suggests a simple positional model for the given problem (which is not suggested by the original one). The synthesis (which is nothing but the solution of the given problem) consists then in the construction of the sought after elements of this model starting from the given ones.

Thābit's previous arguments is an example of trans-configurationnal analysis. Consider the particular case corresponding to Euclid's proposition VI.30. Once this proposition is converted into a purely quantitative problem, it requires constructing a mean proportional x between any segment equal to a given segment a and any other segment segment $a - x$ resulting from taking away a segment equal to x from a segment equal to a, that is, a segment x meeting the following condition

$$a : x = x : a - x.$$

Trans-configurational analysis transforms this problem as follows:

- For proposition VI.16, it reduces it to the problem of constructing a segment x such that
$$S(x) = R(a, a - x).$$
- For proposition II.2 (*i. e.*, according to equality (9)), this last condition reduces to the other one:
$$S(x) + R(a, x) = S(a)$$
- For proposition II.3 (*i. e.*, according to equality (15)), this reduces in turn to:
$$R(x, a + x) = S(a)$$
- Finally, for proposition II.6 (*i. e.*, according to equality (17)), this reduces to:
$$S(a) + S\left(\frac{a}{2}\right) = S\left(x + \frac{a}{2}\right),$$
which immediately suggests a simple positional model, and then a solution.

This piece of analysis is structurally akin Vietian one, but it does not relies on the literal formalism of early-modern geometry. Since, as I have observed above, the notation used to expound this argument enters it only for its semantic function, that is, it is only employed for the purpose of avoiding too long and boring expressions and phrases: any step in this argument could be rephrased in a longest and cumbersome but essentially unmodified way without using this notation, and replacing it by a codified but non symbolic language as Euclid's or Thābit's.

The answer I suggest for the previous questions depends on a further clarification of this point. The symbols that my notation consists of do not only denote objects that one could also denote in a different way, but they overall compose complex expressions that transform into each other not because of some syntactical rules, but rather because of theorems that concern these objects and are proved by reasoning on them in a way that is independent of the use if this notation. Consider, as an example, the last step of the previous argument. It is not licensed by any syntactical rule relative to such a notation, but it rather depends on the admission of equality (17), which is proved by reasoning on an appropriate configuration of rectangles and squares. By saying that proposition II.6 enters the previous argument as a rule of inference, I do not mean that this proposition provides any such syntactical rule, but simply that it allows to pass from a certain equality to another.

The literal formalism of early-modern algebra works in an essentially different way. The signs composing it are also used for their syntactical function. This means that they are part of a syntactical system providing appropriate rules for transforming the expressions composed by these symbols into each other. These rules express the properties of a number of operations, and these properties are independent of the objects these operations are supposed to apply (that is, they are the same whatever these objects might be: numbers, segments, or any other sort of quantities), to the effect that they can be so expressed also without having care to the semantic functions of the relevant signs, that is, to their power to denote some objects. The use of literal signs for their syntactical function is, I argue, the essential novelty coming together with the adoption of the formalism of early-modern algebra.

Of course, this means neither that the same signs were previously adopted for their semantic function, nor that they did not also comply with this last function within such a formalism. As a matter of fact, the former thing is only very partially true. The latter is plainly false, instead. What is relevant is rather that the two functions of signs, the semantic and the syntactical one, are integrated within the formalism of early modern algebra: these signs both denote objects (according to some appropriate interpretation of them), and compose expressions that transform into each other according to some syntactical rules, *i. e.* of rules that directly apply to these signs independently of their power to denote any object.

But, if this so, one could object, there is no room to argue, I as I just did, that the shift in conception that I accounted for was a necessary condition for the adoption of the literal formalism of early modern algebra. The emergence of such a formalism, one could argue, only depended on the fixation of appropriate syntactical rules applying to appropriate signs; and just insofar as these rule are syntactical in the previous sense, this could be but independent on any way of conceiving geometrical objects and their relation.

This objection is clearly flawed, however, and this is so for a quite trivial reason. Since the features of the syntactical system involved in the literal formalism of early modern algebra were just motivated (as it happens for any syntactical system to be used in mathematics) by the purpose of making some interpretations of this formalism possible. This system was just conceived in order to provide a tools for dealing with arithmetical and/or geometrical objects. The fact that the rules involved in it are syntactical, that is, directly apply to the relevant signs independently of their power to denote any object, does not entail that they were not conceived for the purpose of providing such a tool. Now, it is easy to understand that it was only insofar as geometrical objects came to be conceived as pure quantities—that is, as possible *relata* of a system of purely quantitative conditions—that a geometrical interpretations of a literal formalism as that of early-modern algebra (intrinsically unsuitable for expressing the positional relations of geometrical objects) could be licensed. This is just the reason for I argue that the adoption of such a formalism was made possible by the shift in conception that I have tried to describe[25].

[25]One could perhaps argue that things should be seen in the other way around, that is, that it was the rephrasing of some geometrical problems (like that stating by proposition like VI.30) using a literal formalism like that of early modern algebra that forced the understanding of these problems as a purely quantitative ones. But, for such a line of argumentation to be plausible, one should also explain what motivated this rephrasing. A possible response could be that this rephrasing was motivated by the desire

At this point, another story should be told: a story accounting for the way as such a formalism was conceived, so as to be appropriate for integrating the semantic and the syntactical functions of the signs involved in it. This is a quite complex story, that has been told many times in different ways. Of course, I'd have my own way to tell it[26]. But this is not something I can do here.

References

[1] U. al Khayyām. *L'algèbre d'Omar al-Khayyāmī*. Duprat, Paris, 1851. Publiée, traduite et accompagnée d'extraits de manuscrits inédits par F. Woepcke.

[2] H. J. M. Bos. *Redefining Geometrical Exactness. Descartes' Transformation of the Early Modern Concept of Construction.* Springer Verlag, New York, Berlin, etc., 2001.

[3] Euclid. *The Thirteen Books of the Elements.* Cambridge Univ. Press, Cambridge, 2nd edition, 1926. Translated with introduction and commentary by Sir Thomas L. Heath; 3 vols.

[4] M. ibn Musa al Khwārizmī. *The algebra of Mohammed ben Musa.* Printed for the Oriental Translation Fund. and sold by J. Murray, London, 1831. Edited and translated in English by F. Rosen.

[5] M. ibn Mūsā al Khwārizmī. *Le commencement de l'algébre.* Blanchard, Paris, 2007. Critical edition of al-Khwārizmī's *Book of Algebra and al-Muqābala*, with French Translation, Introduction and Commentaries by R. Rashed.

[6] P. Luckey. Ṯābit b. Qurra über den geometrischen Richtigkeitsnachweis der Auflösung der quadratischen Gleichungen. *Berichte über die Verhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig. Mathematisch-physikalische Klasse*, 93:93–114, 1941.

[7] R. Netz. *The Transformation of Mathematics in Early Mediterranean World: From Problems to Equations.* Cambridge University Press, Cambridge, New York, etc., 2004.

of stating (and solving) geometrical problems by using a language analogous to that of arithmetic. The choice between my hypothesis and this alternative one should be ultimately justified by detailed textual considerations that, here, I can only suggest to undertake. Still, it seems to me that the evidence I offered in my paper, relative both to Euclid's *Elements* (in section 3) and to al-Khwārizmī's and Thābit's treatises (in section 4) is at least enough to show that the understanding of geometrical problems as a purely quantitative ones is not only independent of the use of a literal formalism like that of early modern algebra, but is also suggested (independently of any concern for such a literal formalism) by the way Euclid expresses himself and reasons, and explicitly at work many century before early-modern age, in the context of a manifest effort of unification of arithmetic and geometry.

[26] Concerning this matter, I can here only refer the refer to the Introduction of my [10], pp. 1-44.

[8] J. Oaks. Al-Khayyam's scientific revision of algebra. *Preprint, University of Indianapolis*, 2010.

[9] J. Oaks. Geometry and proof in Abū Kāmil's algebra. In M. Abdeljaouad, editor, *Actes du dixieme colloque Maghrebim sur l'histoire des mathématiques arabes: Tunis, 29-31 Mai, 2010*. To appear in 2011.

[10] M. Panza. *Newton et les origines de l'analyse: 1664-1666*. Blanchard, Paris, 2005.

[11] M. Panza. What is new and what is old in Viète's analysis restituita and algebra nova, and where do they come from? some reflections on the relations between algebra and analysis before Viète's. *Revue d'Histoire des Mathématiques*, 13:83–153, 2007.

[12] M. Panza. Rethinking geometrical exactness. *Historia Mathematica*, 37:???–???, 2010.

[13] R. Rashed, editor. *Thābit ibn Qurra: Science and Philosophy in Ninth-Century Baghdad*. Walter de Gruyter GmbH & Co, Berlin, 2009.

[14] R. Rashed and B. Vahabzadeh. *Al-Khayyām mathématicien*. Blanchard, Paris, 1999.

[15] F. Viète. *In artem analiticem isagoge: seorsim excussa ab Opere restitutae mathematicae analyseos, seu Algebra nova*. apud I. Mettayer, Turonis, 1591.

Figures

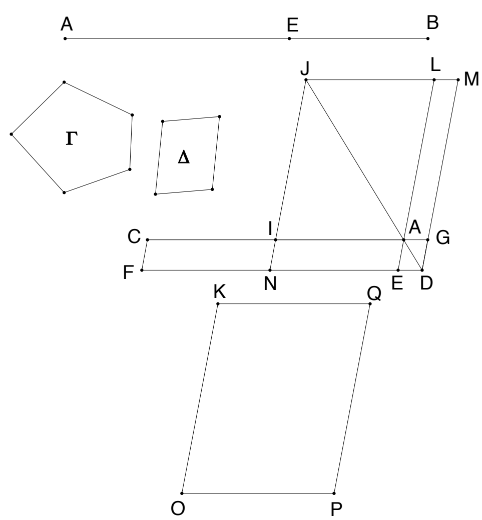

Figure 1

Figure 2

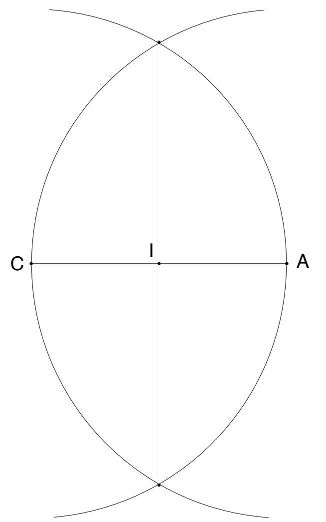

Figure 3

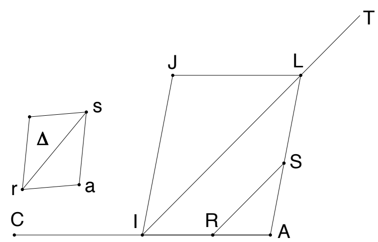

Figure 4

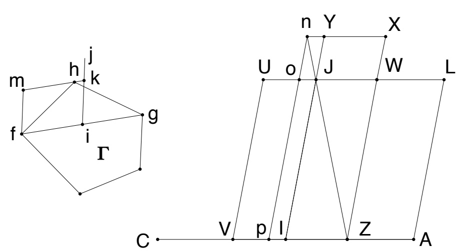

Figure 5

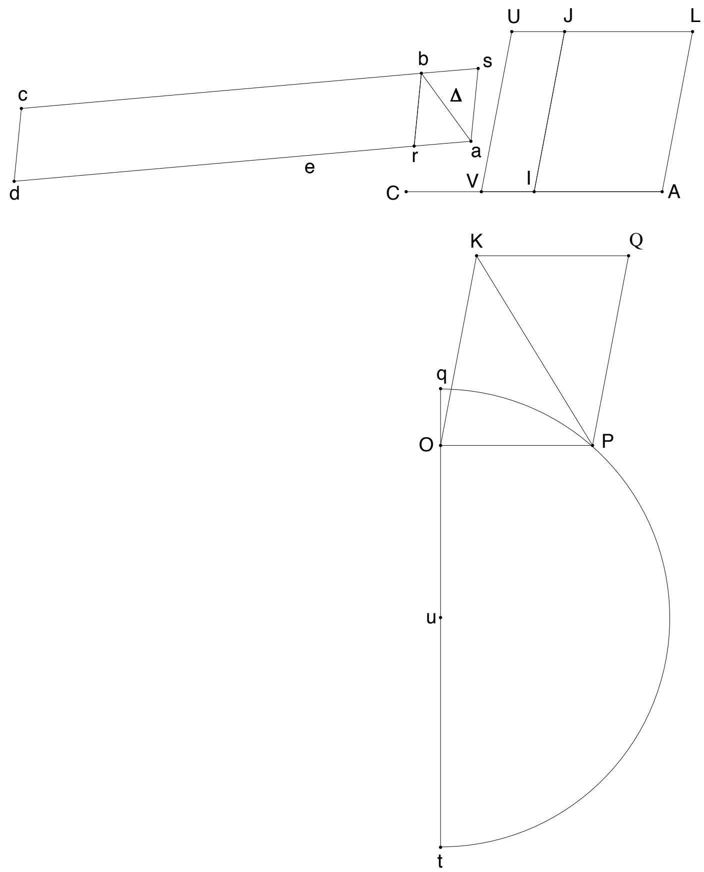

Figure 6

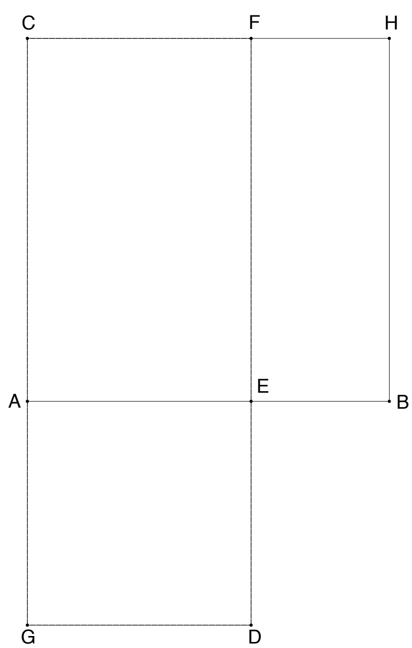

Figure 7

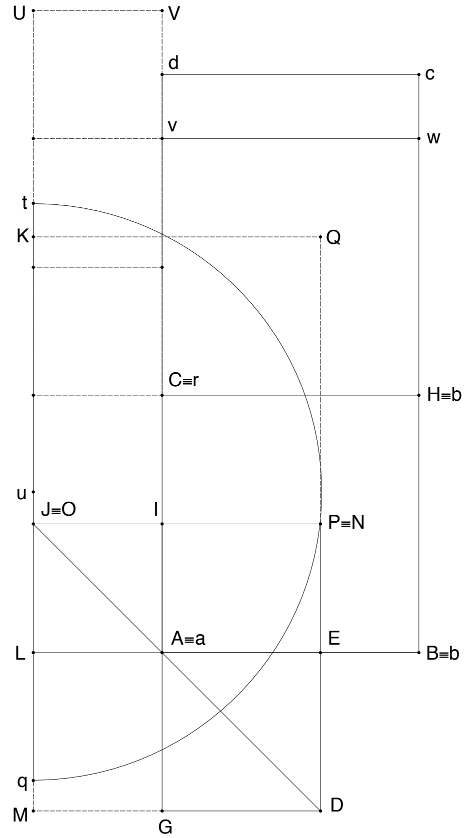

Figure 8

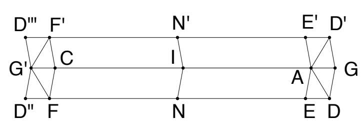

Figure 9

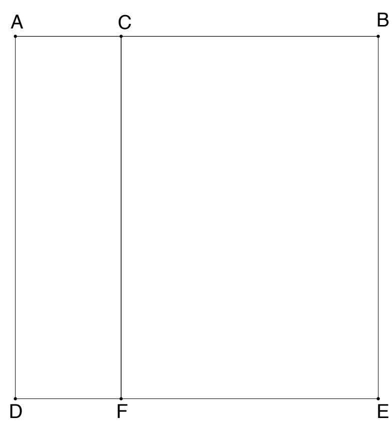

Figure 10

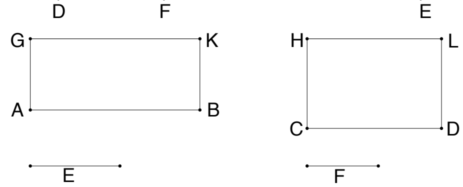

Figure 11

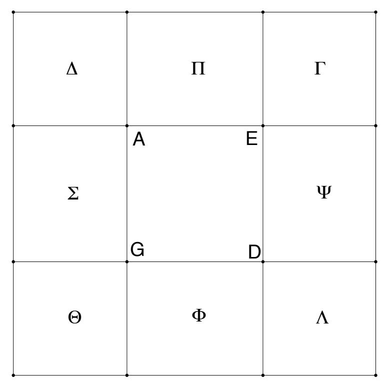

Figure 12

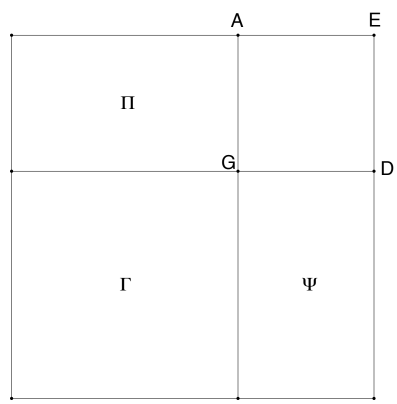

Figure 13a

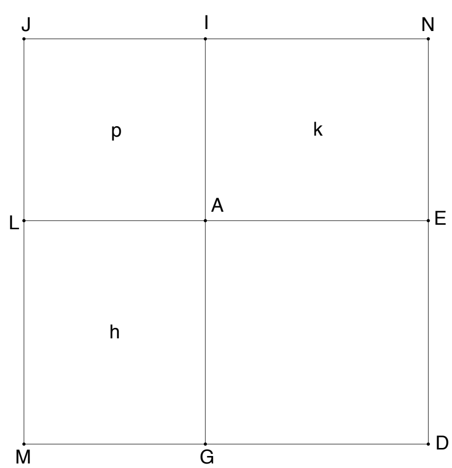

Figure 13b

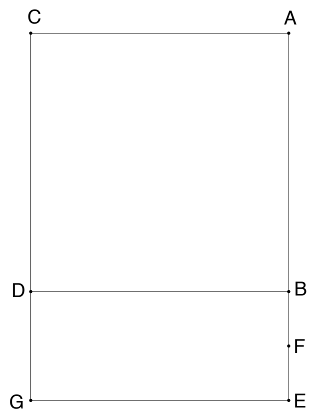

Figure 14

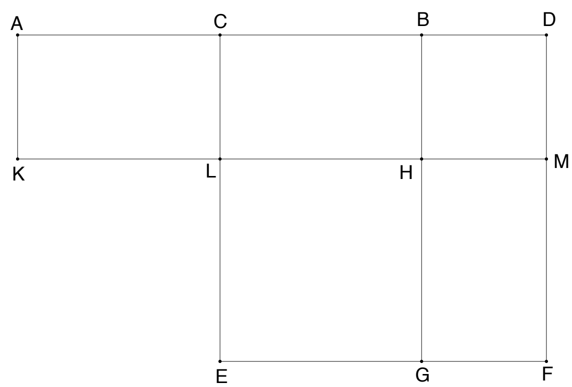

Figure 15

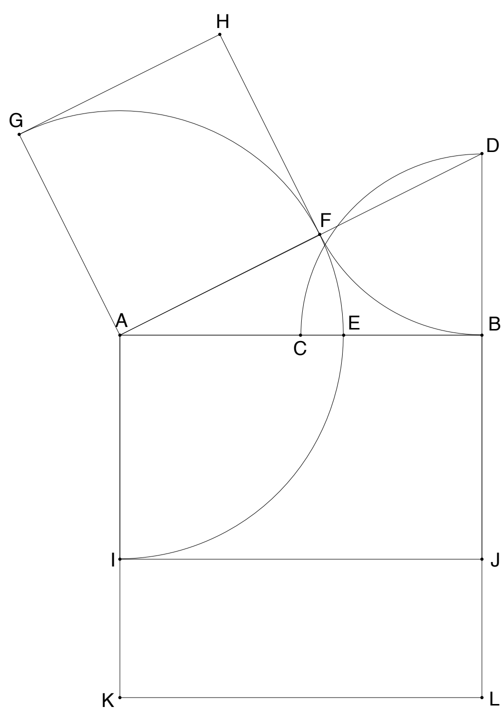

Figure 16