Archive for History of Exact Sciences (2023) 77:451–499
https://doi.org/10.1007/s00407-023-00308-y

# Hero and the tradition of the circle segment

**Henry Mendell**[1]

Received: 27 September 2022 / Accepted: 22 February 2023 / Published online: 24 July 2023



## Abstract

In his *Metrica*, Hero provides four procedures for finding the area of a circular segment (with $b$ the base of the segment and $h$ its height): an Ancient method for when the segment is smaller than a semicircle, $(b+h)/2 \cdot h$; a Revision, $(b+h)/2 \cdot h+(b/2)^2/14$; a quasi-Archimedean method (said to be inspired by the quadrature of the parabola) for cases where $b$ is more than triple $h$, ${}^4/_3(h \cdot b/2)$; and a method of Subtraction using the Revised method, for when it is larger than a semicircle. He gives superficial arguments that the Ancient method presumes $\pi = 3$ and the Revision, $\pi = {}^{22}/_7$. We are left with many questions. How ancient is the Ancient? Why did anyone think it worked? Why would anyone revise it in just this way? In addition, why did Hero think the Revised method did not work when $b > 3\ h$? I show that a fifth century BCE Uruk tablet employs the Ancient method, but possibly with very strange consequences, and that a Ptolemaic Egyptian papyrus that checks this method by comparing the area of a circle calculated from the sum of a regular inscribed polygon and the areas of the segments on its sides as determined by the Ancient method with the area of the circle as calculated from its diameter correctly sees that the calculations do not quite gel in the case of a triangle but do in the case of a square. Both traditions probably could also calculate the area of a segment on an inscribed regular polygon by subtracting the area of the polygon from the area of the circle and dividing by the number of sides of the polygon. I then derive two theorems about pairs of segments, that the reviser of the Ancient method should have known, that explain each method, why they work when they do and do not when they do not, and which lead to a curious generalization of the Revised method. Hero's comment is right, but not for the reasons he gives. I conclude with an exploration of Hero's restrictions of the Revised method and Hero's two alternative methods.


Communicated by Alexander Jones.

✉ Henry Mendell
hmendell@calstatela.edu

1 Philosophy Dept, UCLA, Los Angeles, USA

## 1 Introduction

Hero of Alexandria in the *Metrica*,[1] followed by the Heronian *Geometrica* and preceded by Columella in his *De re rustica*,[2] preserves a method for calculating the area of a circular segment that is remarkable in its accuracy and is likewise peculiar in the very same matter. While we expect no justification of the method in the *Geometrica* or in a work on agriculture, it is notable that while Hero goes to some effort to justify an alternative method, he does little to establish this method nor a more ancient method on which it is based. It is also notable how modern commentators slide over the basis for both methods as well, usually accepting the little Hero says. Hero adds that the method only works well when the base of the segment is no more than three times its height and then offers the alternative method which he does establish, but which is much worse outside the limited case. This too is perplexing, as we shall see. He concludes his discussion with a fourth method, by subtraction, for the case where the height of the segment is greater than the radius. I shall start with a discussion of Hero's presentation in order to raise twelve questions about it. I do not pretend that I can answer all twelve, but they are all worth posing, and most will get plausible answers. I shall then use the appearance of the Ancient method, as I shall call it, in the third century BCE Demotic Egyptian treatise assembled by Parker from P. Cairo 89127-30, 89137-43, a text with Babylonian background, to establish some basis for why the rule might seem plausible even if it lacked a proof, although it was also known to be imperfect. While I do not know how the Babylonian mathematicians came to discover the ancient method nor how old it is, I will at least establish its use in Uruk, via the late fifth century BCE tablet W 23291-x, and will give some grounds for why it would be seen as plausible, although my analysis will also involve some startling coincidences. I shall also make some observations about an algorithm in the Old Babylonian tablet, BM 85194, and about the constants for polygons in the late Old

[1] There are now three editions of the *Metrica* available, Schoene (1903), with a German translation, Bruins (1964a), with an English translation (with only the Greek text in Bruins 1964b), and Acerbi and Vitrac (2014), with a Fnrech translation, based on the single manuscript, Cod. Seraglio G.I. I (formerly Const. Palat. Vet. 1), which Bruins (1964a) also reproduces in facsimile. I shall not attempt to diagnose the history of the text, whether certain parts are intrusions. For convenience, I will assume that Hero composed the *Metrica* sometime in the first or second century CE. See Acerbi and Vitrac (2014, pp. 16–26) for a discussion of the controversy and Masià (2015) for a more recent attack on one of the principal foundations for the dating of Hero, a lunar eclipse of 62 CE, wounded perhaps but not moribund (Masià, for example, considers it likely that Hero, *Dioptrics* 35, just happened to choose to display an instance, "let there have been found an eclipse …," chock full of very specific details, mostly completely unnecessary to his abstract argument). It is unimportant to my argument when Hero lived, but if my argument is right, it is unlikely that he lived in the second century or even the early first century BCE, at a time when the discovery of what I call the Revised method was made. For he would have presented it better.

[2] I put Columella before Hero for convenience and concede that this might be completely wrong. What is important for my argument is that Columella does not use the *Metrica*, although his procedures are very close to procedures presented in the *Metrica.* For example, in the case of the circular segment, he presents for the case of the segment less than a semicircle what I will call the Revised method, but uses a value for the height (5) and base (20) for which Hero recommends a different method. For a discussion of Columella in the Latin metrical tradition, see Bertoni (2017). Of treatises in the Heronian corpus, it is generally agreed that the *Geometrica*, *De mensuris*, and *Stereometrica* are largely composites and may be a good deal later than Hero, although it is debatable how much they reflect work of Hero and how much just the metrical tradition. See Acerbi and Vitrac (2014, pp. 429–507). A strong case could be made that all pre-medieval discussions of the areas of segments are uninfluenced by the *Metrica*.

Babylonian tablet, Susa I (Bruins and Rutten 1961, text III). All this will be prelude to the main part of our banquet, the explanation of Hero's peculiar remarks about the Ancient method and its revision, the Revised method, as I shall call it. I shall give a demonstration of a theorem about adjacent segments on a rectangle inscribed in a circle that explains very well why the Ancient method works as an approximation if $\pi = 3$. This proof, which involves adding and subtracting rectangles in a square, will only require material no more advanced than *Elements* I and will be designed to be accessible to Babylonian style manipulation. I shall then prove a general theorem, based on the first, for any value of $\pi$. This will show why the Revised method Hero cites is required, presumes that $\pi \approx 3\ {}^1/_7$, and works when it does work. This will also trivially allow for a generalization of the method for any approximation of $\pi$, should anyone care. The last step in my argument, albeit much less satisfactory, will be to explain why Hero incorrectly marks the limitations of the method in favor of the second method he uses, inspired by Archimedes' quadrature of the parabola, as well as to make some observations about his method of subtraction. Perhaps today, these methods have all become historical curiosities, but even curiosities can earn our respect and a careful explanation why they work when they work. In addition, these methods really do work. At the very least, we shall be able to trace an important and complex thread in the metrical tradition from Babylon to Hero, to show early mathematicians' awareness of the methods' limitations and of their validity. Hero's presentation has seemed a mess, enough so that some have questioned whether the text could originally have been presented as we have it. Hero cannot dispense with any of the three methods he advocates for measuring the area, unless he wishes to introduce a discussion of chord tables. He avoids doing this in the *Metrica*, perhaps for good reasons. But first, let me raise my puzzles by walking us through Hero's account.

## 2 Hero's account of four methods and the dozen questions they raise

We can, for practical purposes, distinguish three sorts of metrical practices in ancient texts. Some are simply geometrically based, that is, they are simply applications of a geometrical principle with numbers attached to lengths, areas, etc. Any error comes from outside the stipulated procedure.[3] Taking the area of a triangle as a multiple of half a side and the height from that side or as the square root of the multiple of half the perimeter and the difference between each side and half the perimeter only involves errors from the measurement of the respective lines or in chosen approximations of square roots. Even taking the area of a circle as a function of the square of the diameter,

[3] Acerbi (2021, pp. 2–22) distinguishes three "codes" of Greek mathematics, Demonstrative, Procedural, and Algorithmic. The Demonstrative Code is the familiar style of Euclid's *Elements*, the Procedural provides a general procedure verbally, without actual numbers (often in 1st person plural for actions) though often followed by an instance in the Algorithmic style, and the Algorithmic (often in the 2nd person imperative for actions) just provides the procedure as a paradigmatic instance with actual numbers. Although some late Babylonian texts are in Procedural form, the Egyptian and Babylonian texts that I shall consider are very close to the algorithmic style of Greek mathematics (indeed almost certainly its source). In my discussion, I shall speak of algorithms as paradigmatic procedures with actual numbers that the reader is to adapt, but will use 'procedure' as a general term for procedures given in Procedural form and for algorithms. When speaking of a text employing the Procedural Code, I shall speak, as here, of 'procedural form'.

in effect times a given parameter, $^3/_4$ or $(^8/_9)^2$, might be seen in this light, except that we have no idea how the respective procedures in Babylon or Egypt were arrived at. However, with the proof of Euclid, *Elements* xii 2, it is easily inferable that each circle has the same ratio in power to its diameter. With Archimedes, *Dimensio Circuli* 1, one can just take the area as the multiple of half the radius times the circumference, however, that may be determined.

The second group requires the procedure be based in an approximating series, i.e., a progressive approximation. This is typically the case for square roots in older texts, but is characteristic especially of Archimedean texts. Here, one has a procedure of approximation, typically by approaching from upper and/or lower bounds. An example might be the anthyphairetic procedure for getting at the ratio of diagonal to side of a square by building a series of pairs $(1,1) \rightarrow (3,2) \cdots \rightarrow (p_{\mathrm{m}},\ q_{\mathrm{m}}) \rightarrow (p_{\mathrm{m}} + 2q_{\mathrm{m}},\ p_{\mathrm{m}} + q_{\mathrm{m}})$. As Plato seems to hint in *Republic* VII, the square of the first number will differ from double the square of the second, alternatingly being larger or smaller by 1. On the other hand, the method that Hero uses for finding square roots[4] that seems to be based in Old Babylonian techniques,[5] $n \rightarrow p_1$ such that $p_1^2$ is near $n \rightarrow p_2 = {}^1/_2\left(p_1 + n/p_1\right) \rightarrow \cdots$ shoots in quickly on $\sqrt{n}$ from above. Other procedures involve closing in on a figure, typically curvilinear from the inside or outside. Choosing $3\ ^1/_7$ or $3\ ^{10}/_{71}$ for the parameter for multiplying the diameter to get the circumference of a circle is based in a progressive procedure of approximation, whether or not a Roman surveyor choosing the first number had read his Archimedes and whether or not the values for the square roots used by Archimedes would need to be recalculated to find a more precise value.

The third group involves an approximation that is, we might say, loose. Typically, one finds an average or close value, which may have some geometrical basis, but is not, at least as conceived, associated with a progressive procedure honing in on the actual figure. Typical are the Babylonian and Egyptian procedures for finding the volume of truncated cones, based on cylinders whose end circles are calculated from the averages of the lengths of the diameters or of the areas of the end circles of the truncated cones. Unlike methods based on Eudoxan successive inscriptions or Archimedean compression, there is no reason to expect that the approximations will sit on one side or the other of the actual value, regardless of the desires of a tax collector or a granary owner to maximize profit. In fact, unless the user has an independent simple geometrical or progressive procedure for determining a corrected value, she has no way of knowing whether the method sits on one side or oscillates on either side for different values, as, in fact, is the case with the volumes of truncated cones. All she can know is that it is near, in the way that an average might be near the deviations from the average. Obviously, even when systematic, rounding loosens all measure.

Furthermore, in the context of the Greek metrical tradition, even a loose approximation may come with a proof. For example, it is possible that one could show that a method works because it provides an area smaller than the actual area under the curve being studied. It might still be a loose approximation because it is neither a

4 E.g., *Metrica* I 8. See Acerbi and Vitrac (2014, pp. 121–4, 127–8).

5 See Fowler and Robson (1998), who state that Babylonians never iterated the approximation. However, P. Cairo (see footnote a in table under Fig. 5) and Greek mathematicians did so.

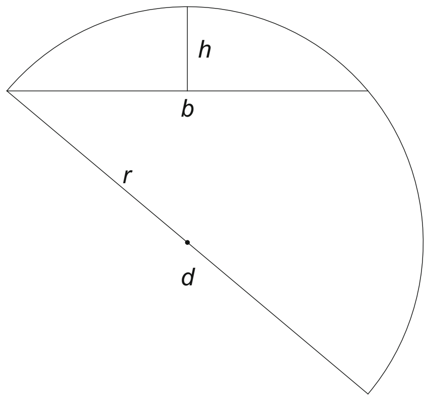



**Fig. 1** Segment in a circle with $h$ the height and $b$ the base of the segment, $r$ the radius and $d$ the diameter of the circle, and, not depicted, $h_2$ the height of the complementary segment

simple calculation of the area nor a member of a series that the mathematician or his contemporaries recognize as leading towards the area of the figure studied. For we might recognize that an approximation could be used in some series without it being historically part of an approximation series that someone actually used or even might have used. Hero provides an example of this in his presentation of the area of the segment.

One ground for suspicion that an approximation method is loosely geometrical is that the values leap on both sides of the true value, above and below. Of course, this will not always be the case. The anthyphairetic methods for square roots shift on both sides, but we also understand clearly why. In addition, it may well be the case that an approximation is always below or above a true value without its being within a series for making closer approximations. I shall argue that both the Ancient and the Revised ancient approximations mentioned by Hero are loose in this sense, that each will fall on both sides of a "true" value and that any mathematician who understood the way in which the two methods worked would have understood this very clearly, whether or not anyone actually did. Hero mentions four methods for measuring the area of a segment of a circle. Here and throughout this paper, $h$ is the height and $b$ the base of the segment, $r$ the radius and $d$ the diameter of the circle, and here, $h_2$ is the height of the complementary segment (Fig. 1).

Ancient method (in fact, at least late Babylonian and Demotic Egyptian, also found later in China[6]): $(b + h)/2 \cdot h$.
Revised Ancient method (post Archimedean): $(b + h)/2 \cdot h + (b/2)^2/14$.
Hero's quasi-Archimedean method (for case $b > 3\ h$), where one should note that only the geometrical theorem is presented and not a procedure (see footnote 3): $^4/_3\, b \cdot h/2$.

[6] *Nine Chapters* I probs 35, 36. See Chemla and Shuchun (2004, pp. 141, 190–3, 781–82), where the two examples are (with 1 mu = 240 bu): $b = 30$ bu, $h = 15$ bu (solution: 1 mu 97 $^1/_2$ bu, and $b = 78^1/_2$ bu, $h = 13^7/_9$ bu (solution: 2 mu 155 $^{56}/_{81}$ bu). The ancient commentary attempts to explain the procedure with a decomposition.

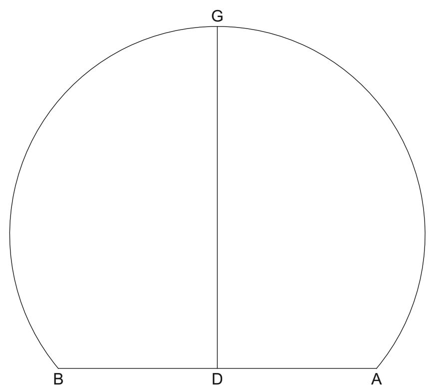
Hero's Subtraction method for $h > r$ or $b < 2\ h : h_2 = (b/2)^2/h \rightarrow$ $\text{area}_{\text{complement}} = (b + h_2) \cdot h_2/2 + (b/2)^2/14 \rightarrow d = h + h_2 \rightarrow \text{area}_{\text{circle}} =$ $(d^2 \cdot 11)/14 \rightarrow \text{area}_{\text{circle}} - \text{area}_{\text{complement}}$.

In addition, unstated by Hero, but which we might presume from his conditions for using the quasi-Archimedean method, though wrongly as it turns out (for the case where $b < {}^4\!/_3\ h$, i.e., $b > 3h_2$): $\text{area}_{\text{circle}} = (d^2 \cdot 11)/14 \rightarrow$ $\text{area}_{\text{circle}} - {}^4\!/_3\ b \cdot h_2/2$.

After providing three preliminary chapters of lemmata for establishing the quasi-Archimedean method, in *Metrica* I 30, Hero starts his main discussion with a "more careless," Ancient method for calculating the area of the segment, $(b+h)/2 \cdot h$, which he observes is equivalent to taking the circumference as three times the diameter (I use "circular-arc" for the circumference of a circle and any part of it) (Fig. 2):

> The ancients used to measure more carelessly the segment of a circle smaller than a semicircle. For by adding the base and altitude of it and taking half of these by the altitude they would declare the area of the segment of so much.
> In addition, these seem to follow on those who suppose the perimeter of the circle as triple the diameter.
> For if we measure the semicircle according to this hypothesis, the area of the semicircle will be in agreement with the mentioned method. For example, let there be a semicircle whose diameter is AB and altitude GD, and let the diameter be 12 units. Therefore, GD is 6 units. Accordingly, the circular-arc of the circle will be 36 units. Therefore, that of the semicircle will be 18 units. Therefore, since it was shown that the [rectangle enclosed] by the circular-arc and the [line] from the center is double the area, after multiplying 18 by 6 it is required to take the half. But they are 54 units. The same will be the case if you add 12 and 6, which become 18, taking half of which by those of the altitude. It becomes similarly 54.

**Fig. 2** *Metrica* I 30 Ancient method



Hero's argument is a little peculiar. We need to know that the example taken is general and not just an odd instance where half the area of the circle, taking the circumference as 3·diameter, equals $(b + h/2 \cdot h$, with $b$ equal to the diameter and $h$ the radius. In any case, one would quickly see that the value at the semicircle, is, with $d$ the diameter and $r$ the radius, $(d + d/2)/2 \cdot d/2 = {}^1/_2({}^3/_4 d^2) = {}^1/_2({}^1/_2 \cdot 3\ d \cdot r)$, half the Archimedean value for the area of a circle as $^1/_2$ the rectangle formed by the circumference of the circle rectified and the radius, with the circumference as 3 $d$, or, as we would say, taking $\pi$ as 3.

Why should one think that the method will work at all when the segment is not a semicircle? Why not use some other procedure, that is also equivalent to taking $\pi$ as 3, such as: $(b/2 + h)/2 \cdot b/2$ or $^3/_2\ h^2$? Or even as some student in the first or second century CE got the area of the circle from the semicircle $(d + h)^2/3$. Therefore, take half of this: $(b + h)^2/6$. The same student next (prob. 5) gives the Ancient rule for the semicircle as $(d + h)/2 \cdot h$.[7] Therefore, if his teacher intended him to understand that this otherwise overwrought procedure was to be generalized, we are back to the same question. Why? Also, how will one know that the procedure is any good?

Question 1: Why this procedure and not the others?
Question 2: Is there a way of making sense of Hero's remark that the method is tied to taking $\pi = 3$ that is more robust than the observation that it seems so in the case of the semicircle?
Question 3: Is there a way of justifying the procedure without having a 'correct' procedure with which to compare it that is historically plausible?
Question 4: How old is the method?

Hero next reports (I 31) a more precise and recent method:

> Others who investigated more precisely add to the mentioned area of the segment the 14th from the [square] of half the base. These, in fact, appear as following a different method according to which the circular-arc of the circle is triple the diameter of the circle and a 7th part larger. For if we similarly suppose diameter AB 14 units, altitude DG 7 units, the circular-arc of the semicircle will be 22 units. [Multiplied] by 7, this becomes 154, of which a half becomes 77. And declare the area of the semicircle as so much.
> The same [occurs] also if we do this: add 14 and 7, of which the half becomes 10 $^1/_2$. By 7 it becomes 73 $^1/_2$. In addition, the [square] from half of the base is 49 units. Universally [that is, always taking] the 14th of these becomes 3 $^1/_2$. Add these to 73 $^1/_2$. They become 77.

If Hero's remark about the Ancient method is peculiar, his remark about the Revised method is just as peculiar. The fact that the method yields the same result in the case of the semicircle does not explain why the method is tied, except in this instance, to an Archimedean value for $\pi$. Yet, there are two other questions that are just as perplexing. Why add the 14th of the square of half the base? In addition, if this gives

[7] P. Vind. 26740, prob. 4. See Bruins et al. (1974). They do not propose a date for the writing, but the coauthors, Sijpesteijn and Worp (1974, p. 311), hazard first or second century CE for the Homeric quotation on the papyrus. Since these are school exercises, it does not seem likely that they would be separated by much time. Prob. 5 uses the Ancient method for a semicircle with a diameter of 10 schoenia.

a better approximation than adding the 14th of the square of the height, how does the mathematician know this? Which of these four procedures is better and how does anyone know this:

$$(b+h)/2 \cdot h + (b/2)^2/14$$

or

$$(b+h)/2 \cdot h + h^2/14?$$

or a simple adjustment by multiplying the new value of $\pi$ divided by the old value (see ps.-Hero, *De mensuris* 29, but indicated for segments larger than a semicircle),

$$^{22}/_{21} \cdot (b+h) \cdot h/2$$

or just a version of the circle based on the semicircle in terms of base and height (see ps.-Hero, *Geometrica* 20.4(S), but also indicated for segments larger than a semicircle)

$$b \cdot h \cdot {}^{11}/_{14}?$$

All of these are, after all, equivalent for the area of the semicircle. It is also not difficult to come up with other possible revisions of the Ancient method.

> Question 5: Without some way of comparing the values from some approximative algorithm to a true value, what is the basis for someone saying that the Revised method is better than the Ancient method?
> Question 6: Why does the Revised method add a part of the square of half the base and not something else?
> Question 7: Is there a way of making sense of Hero's remark that the method is tied to taking $\pi = 3\ {}^1/_7$ that is more robust than the observation that it seems so in the case of the semicircle?

Heuristically, my questions about Hero's methods began with the next, eighth question, which also will lead to the solution to the other questions. In Archimedes' *Measurement of the Circle*, $3\ {}^1/_7$ is, of course, an upper bound for the ratio of the circumference to the diameter, and hence for the ratio of the area of the circle to the square from the radius. It should be obvious to even a casual reader of the text that our version must be heavily redacted, rearranged, and abridged, but Knorr (1989, Part III) goes into much detail about how little we know about the original text. He argues that it is probable that Archimedes discussed the areas of sectors and segments in the book. The Revised method is unlikely to be from that book. The reason is that we expect from the way Archimedes presents the material that an upper bound will remain an upper bound in its application. That is, in principle, if Archimedes, or anyone, had used $3\ {}^1/_7$ or the lower bound $3\ {}^{10}/_{71}$ in building a theory of the area of segments, we would expect the rule to make it such that the area from the upper bound would be consistently above and that from the lower bound consistently below the true value for the area. This is not the case with the Revised formula, as is clear from Fig. 3, although a Greek might well have seen that the Ancient method is consistently below the true value. The reader will have to wait to see why this is the case, but let us just pose it as a big question.

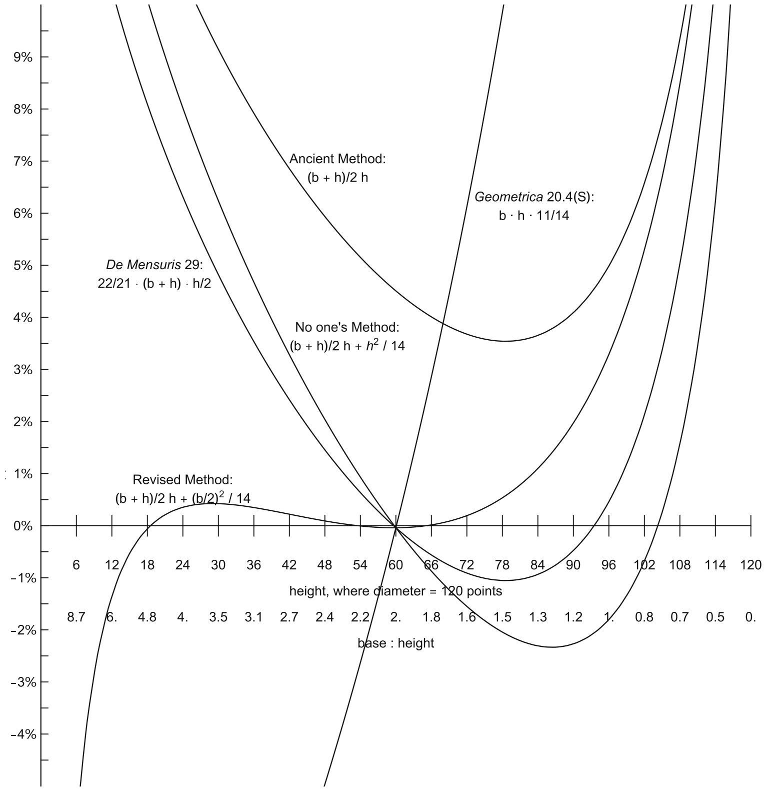



**Fig. 3** Comparison of Ancient and Revised methods to True area, along with three other algorithms (note that Fig. 3 does not depict what happens at the extremes. As the height gets smaller, the Ancient method provides values increasingly larger than the true. Indeed the inept method at the *Geometrica* 20.4(S) is equal to the Revised method when $h = b/25$ and the height is 3 $^{29}/_{200}$, and the error $-$ 17.38%. For lower values, it will be better than the Revised method. On the other side, when the segment is the circle ($b = 0$, $h = 120$), the Ancient method and the Revised method will give the same area for the circle, 7200, for an error of 36.3%. The method no one used will have error of 27.2%; that in *De mensuris* 29, an error of 33.3%; and that at *Geometrica* 20.4(S), 100%, as the area will be 0)

> Question 8: Why does the application of the value for an upper bound for $\pi$ in a procedure for finding the area of a segment fluctuate on both sides of the true value and what does this tell us, if anything, about the procedure?

Hero next (I 31) puts a limitation on the Revised method and introduces his quasi-Archimedean method for dealing with the bad cases.

> However, this method again also will not fit with every segment, but when the base is not larger than triple the altitude, since, note well, if the base is 60 units and the altitude 1, the figure enclosed [by 1 and 60 units] will be 60 units, which is, in fact, larger than the segment. But the 14th of the [square] from half of the base is larger than this. For it is 64 4 14′ [i.e. $(60/2)^2/14 = 64\ ^4/_{14}$]. Thus, the mentioned method will not fit with every segment, but, as was mentioned, when the base is not larger than triple the altitude.

But if it is larger than triple, we will use the following method:
32. Every segment of a circle is larger than a third again of a triangle having the same base as it and an equal height.

Here follows a careful proof that is straightforward if textually very problematic,[8] which we need not here go through, but which uses the lemmata proved in I 27–29, announced in 27 as being used for the measurement of the segment, but which are only used for this theorem. Indeed, this is the only claim that is actually proved about the area of the segment that can be applied to its measurement, although Hero gives no applications of the theorem. Rather he concludes:

This method works when the base is larger than triple the altitude.
If, however, a segment is enclosed by a straight-line and a parabola and the base of it is given and the altitude, that is the axis to the base, and we want to find the area of this, by measuring the triangle having the same base as it and an equal height and adding a third of it to it we will declare the area of the segment. For Archimedes showed in the *Method* every segment enclosed by a straight-line and section of a right-angled cone, that is of a parabola, is a third again a triangle having the same base as it and an equal height.

It is very nice to see Hero's views on the origin of the method, whether or not it is his own discovery or that of Archimedes,[9] and to know that the parabolic segment with the

[8] There are at least two lacunae in the proof of I 32, one in a part that seems garbled, at least to me. Let H be the area equal to triangle ABG inscribed in the semicircle and P the series of triangles on it approximating the semicircle and X areas equal to these, as set up in the proof. Hero seems to argue: (1) $^1/_3$ H $<$ X; (2) hence, H $<$ 3P; (3) hence, H + X $<$ 4X; (4) hence, by conversion (ἀναστρέψαντι), H + X $>$ [lacuna: $^4/_3$H]; (5) but H = ABG, and H + X = ABG + P (polygon inscribed in the circle); (6) hence, ABG + P $>$ a third-again H. Step (2) should be just: H $<$ 3X, as P is irrelevant to the next steps. As to the inference to step (5), if these were equalities, we would expect: H = 3X; H + X = 4X, so that H + X:H = 4:3. But the corresponding inference for the inequality from (3) and (4) is far from obvious! Why doesn't he argue: $^1/_3$ H $<$ X; hence, $^1/_3$ H+H $<$ X+H? The lemma that follows the text does not help, as it, in effect, divides H+X into $^3/_4$ (H+X) and $^4/_3$ (H + X) and infers that H + X = $^4/_3$($^3/_4$(H + X)). However, one still needs to infer from H $<$ 3X that H $<$ $^3/_4$ (H + X), which, of course, is tantamount to the claim in question. Therefore, it is not trivial to construct a coherent argument out of the text. We need a theorem, to be called 'by conversion', such as: $a+b < (n+1) \cdot b \rightarrow a+b : a > n+1 : n$, although the most natural proof would probably be something like: $a+b < (n+1) \cdot b \rightarrow a < n \cdot b \rightarrow a+n \cdot a < n \cdot a+n \cdot b \rightarrow (n+1) \cdot a < n \cdot (a+b) \rightarrow a+b : a > n+1 : n$; or more briefly: $a + b < (n + 1) \cdot b \rightarrow a < n \cdot b \rightarrow b : a > 1 : n \rightarrow a + b : a > n + 1 : n$—an inference of step (4) back to step (3) in order to infer step (5)! All of this might lead one to suspect that more is amiss than a small lacuna.

[9] See Knorr (1978, pp. 228–33 and 1989, pp. 497–502) for a comparison of Hero's proof of the quasi-Archimedean method *Metrica* I 28, 29, 27, 32 with Archimedes' geometric proof in *Quadrature of the Parabola* 19, 21, 23, 24 (Heath 1921, p. 330, and Bruins 1964a, part 3, p. 264, had already noted the close connection). Knorr proposes that the quasi-Archimedean method is due to Archimedes and that it is likely to have been part of the *Measurement of the Circle*, and that it formed the basis for the quadrature of the parabola, rather than the other way around (his most controversial claim). Thus, Hero's reference to the *Method* would then indicate Hero's lack of understanding of the origins of the quasi-Archimedean method. His principal argument is that there are much simpler proofs if one attends to the details of the proof in *Quad. of the Parab*. On his view, Hero could not be the discoverer of the method, but would have seen that it provides a useful lower bound. Although a scholion to *Geom*. 19 in Cod. Seraglio G.I. I f. $11^r$ from the twelfth century (according to Acerbi and Vitrac 2014, pp. 88, 98) attributes the theorem to Archimedes (see Hero, *Opera* v. 5, p. 228–9), Knorr (1989, p. 509 n. 34) dismisses this evidence for his thesis. For arguments against Knorr's theses, see Acerbi and Vitrac (2014, p. 217 n. 265). The issue requires a more detailed discussion.

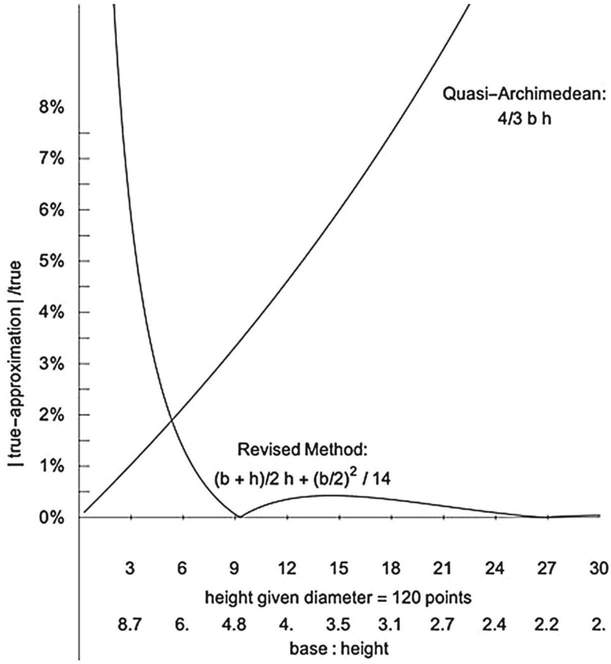



**Fig. 4** Comparison of the Revised and quasi-Archimedean methods

same base and height is a lower bound for the area of the circleand that a segment of a circle with a high ratio of base to height approximates a parabola, whose area he gives in procedural form, but nowhere here is what we need, an argument that the quasi-Archimedean method is better than the Revised method or that it is worse when the base is less than triple the altitude. A comparison of the two methods will bring this out. Hero chooses as an example for the weakness of the Revised method an instance where it is clear that the value is too high, larger than the rectangle enclosing the segment. In addition, this is all in the add-on from the revision: $(60 + 1)/2 \cdot 1 + 64\ {}^{2}/_{7} = 99\ {}^{9}/_{14} > 60 \cdot 1$, in fact more than half-again larger than the rectangle enclosing the segment. He does not need to appeal to a chord table to check the calculation. However, the ratio of base to height where the Revised method goes awry must have some foundation. In fact, the quasi-Archimedean method is better than the Revised method when the ratio is large, but only when it is larger than about 6.4 times the height (Fig. 4).[10]

> Question 9: What is the basis for Hero's claim that the quasi-Archimedean method is better when the ratio of base to height is larger than 3:1, and where does his error come from? Where is he getting this value to compare his two methods?

[10] Bruins (1964a, part 3, p. 266) suggests that the error is that Hero should have said that half the base is larger than three times the height. Even so, the Revised is still better here, while Bruins does not explain how Hero is supposed to have determined any boundary between the methods.

Finally (I 33), Hero presents a Subtraction method for measuring the segment larger than a semicircle. The method is what we expect. From the base and height, find the height of the complementary segment of the circle on the same base, then find its area by the Revised method (and we may suppose by the quasi-Archimedean method if the base is more than triple the height) and, having determined the diameter as the sum the heights of the two segments, find the area of the circle, and finally subtract the smaller segment from the area of the circle. That Hero is right that one needs to use some alternative method is clear from Fig. 3, whether or not the ancients only used the Ancient method for the case of segments smaller than a semicircle (it is mere irony that Fig. 2 above, based on the manuscript, is larger than a semicircle).

Question 10: Why does Hero think it necessary to use the separate method for the case where the segment is larger than a semicircle?

Furthermore, we expect that in the application of the Subtraction method, if the ratio of the base to the diameter minus the height is greater than triple, that it will use the quasi-Archimedean method. However, the example in the text involves a base and height of 14 units, so that the complementary segment has a height of 3 $^1/_2$ units, so that the base is 4 times the height.

Question 11: Why does his example for the Subtraction method use a case where the base is 4-times the height and yet uses the Revised method and not the quasi-Archimedean method?[11]

There is one final question, which probably cannot be resolved.

Question 12: Why does Hero present the three methods he recommends and the one he does not recommend in the messy order he does, that is, first lemmas for the quasi-Archimedean method, then the Ancient method, then the Revised method, then the quasi-Archimedean method with a proof, and finally the Subtraction method for the case where the height is greater than the radius?

In the course of this paper, I shall attempt to answer all twelve questions, and I think that we can arrive at satisfactory answers to most of them. All of these lie in dark shadows in small corners of history, and some will remain there. Without an explicit "And here's how I did it!" we will never know how the Ancient method was discovered. My goal here will be to give it some footing. I do believe that the Revised method will be completely explained, and that its explanation may shed light on the Ancient method. But that explanation, I admit, could well be a hologram, a fake image. Other questions, such as the source of Hero's error will be more speculative. With this, let us proceed.

[11] See Knorr (1989, p. 500), and also Høyrup (1997, pp. 241–2), who uses this anomaly to suggest that Hero originally placed the extensive discussion of the quasi-Archimedean method in the margin of the text, but that a later copyist absorbed it into the main text.

## 3 P. Cairo: the ancient method and the method of subtraction and division

Three problems in the third century Egyptian Demotic papyrus, P. Cairo J.E. 89140-43 (Parker 1972, probs. 36–8), hence just P. Cairo, present the Ancient method for the specialized cases of the areas of the segments on an equilateral triangle inscribed in a circle and for the areas of the segments on a square inscribed in a circle. The papyrus presumes already in problems 31, 32 that the parameter for $\pi$ is 3, which Hero supposed essential for the Ancient method. We shall see later that this is correct, but it will not be completely evident here. Problem 36 starts with the side of the triangle, finds its height and then the height of the segment (1/3 the height of the triangle). It next determines the area of a segment by the Ancient method and then calculates the area of the entire circle by adding up the areas of the three segments and the triangle. Next, it checks the result by determining the area of the circle from the diameter, the sum of the height of the segment and the height of the triangle. Since there is some rounding as the calculations proceed, numbers get fudged a little, but, even so, the calculated areas do not match. Therefore, it recalculates the height of the segment from a diameter determined by the circle's area as the sum of the triangle and three segments. Problem 37 does much the same for the case of the square, except that the check value is stipulated at the start, and the procedure seems to push the numbers to fit. Problem 38 repeats the work of 36, but in reverse. After determining the area of the triangle, it calculates the area of the circle from the diameter, again the height of the triangle plus the height of the segment, and then uses the Ancient method to determine the area of a segment, in order to find the area of the circle as the sum of three segments plus the area of the triangle and the circle. It then ends with the difference between the two calculated areas of the circle.

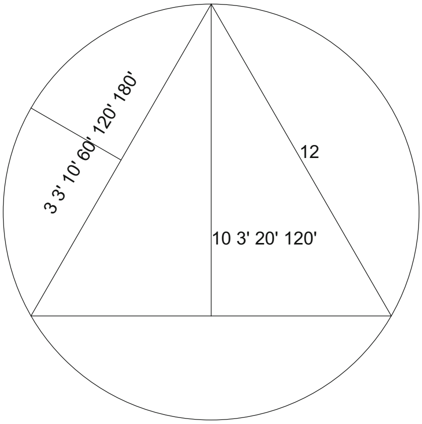
In each calculation, there are three possible sorts of error, miscalculation (the one clear instance possibly due to rounding issues), rounding (typically by dropping a fraction), and choice of the values for square roots (by the standard method from Babylon to Hero[12]). What is also important for us is that the author chose verifiable problems, as it is typical of problems in the papyrus to verify solutions; so this is probably an important part of its mathematical style. The checking is facilitated by the fact that the inscribed figures are both regular polygons, so that one has merely to multiply the area of the segment by the number of sides and add it to the area of the polygon.

In what follows, it is easiest to give the calculations in the form of a table. The numbers will be in quasi-Egyptian fractions, that is, a sequence, n′ m′, will mean $^1/_n + {}^1/_m$. The exceptions will be $^1/_2$, $^2/_3$, and $^5/_6$. Since the author of the papyrus has a penchant for base 60, I will also occasionally include sexagesimal values.[13] There is little point in giving full modern fractions, as they will not help in understanding

12 And found in a late papyrus P. BM 10520 prob. 62, unrelated to our text, but in Parker (1972).

13 However, I find it very difficult to endorse fully the view of Friberg (2005, p. 128), that all of the calculations but one were done originally in sexagesimal and then converted. Many of the calculations are simpler in sexagesimal, but they do not necessarily lead to the same rounding. Rather, as stated, the author obviously has a penchant for base 60 fractions, although some calculations may have been made as Friberg suggests.

**Fig. 5** P. Cairo, Problem 36. The segment on an equilateral triangle




either the relations between the numbers or the ways in which errors work. Where there is rounding or error, I will attempt to provide Egyptian numerals in the spirit of the author, but with the admission that though the spirit may be strong, the mind might not be (Fig. 5).

Prob. 36: with $b = 12$, the side of the equilateral triangle, to find the area of the triangle, the segments, and the circle.

| Index | Procedure | Calculated amount | Corrected | Sexagesimal |
|---|---|---|---|---|
| A | $b \cdot b$ | 144 | | |
| B | $^1/_2\, b \cdot {}^1/_2\, b$ | 36 | | |
| C | A − B | 108 | | |
| D | $\sqrt{\mathrm{C}} = h_{\text{triangle}}$ | 10 3′ 20′ 120′ | 1560′ truncated?[a] | 10;23,30 |
| E | $\mathrm{D} \cdot {}^1/_2\, b = \text{area}_{\text{triangle}}$ | 62 3′ 60′ | | 1,02;21 |
| F | $^1/_3\, \mathrm{D} = h_{\text{segment}}$ | 3 3′ 10′ 60′ 120′ 180′ | | 3;27,50 |
| G | $\mathrm{F} + b$ | 15 3′ 10′ 60′ 120′ 180′ | | 15;27,50 |
| H | G/2 | 7 $^2/_3$ 20′ 120′ 240′ 360′ | | 7;43,55 |
| I | $\mathrm{H} \cdot \mathrm{F} = \text{area}_{\text{segment}} = (h_{\text{segment}} + b)/2 \cdot h_{\text{segment}}$ | 26 $^5/_6$ 10′, which is too high[b] | 26 $^2/_3$ 10′ 90′ 240′ 1800′ 10368′ | Text: 26;56 Corrected: 26;46,57,20,50 |
| J | $3 \cdot \mathrm{I} = \text{area}_{3\ \text{segments}}$ | 80 $^2/_3$ 10′ 30′ | 80 3′ 90′ 300′ 86400′ | Text: 1,20;48 Corrected: 1,20;20,52,2,30 |

| Index | Procedure | Calculated amount | Corrected | Sexagesimal |
|---|---|---|---|---|
| K | J + E = $\text{area}_{\text{circle\_sum}}$ | 143 10′ 20′ | 142 $^{2}/_{3}$ 60′ 90′ 300′ 86400′ | Text: 2,31;11,30 Corrected: 2,22;41,52,02,30 |
| L | D + F = $h_{\text{triangle}}+ h_{\text{segment}} = d_{\text{circle}}$ | 13 $^{5}/_{6}$ 45′ | | 13;51,20 |
| M | 3 · L = circumference | 41 $^{1}/_{2}$ 15′ | | 41;34 |
| N | M/3 | 13 $^{5}/_{6}$ 45′ | | 13;51,20 |
| O | M/4 | 10 3′ 20′ 120′ | | 10;23,30 |
| P | N · O = $\text{area}_{\text{circle\_direct}}$ | 143 $^{5}/_{6}$ 10′ 30′ | 143 $^{5}/_{6}$ 10′ 30′ 90′ 240′ 2700′ | Text: 2,23;58 Corrected: 2,23;58,56,20 |
| Q | P − K $\text{area}_{\text{circle\_direct}}$ − $\text{area}_{\text{circle\_sum}}$ | $^{2}/_{3}$ 10′ 20′ | 1 5′ 12′ 900′ 14400′ 86400′ | Text: 0;49 Corrected: 1;17,4,17,30 |

[a]By the method for finding square roots (for $p^2$ near $n$, $\sqrt{n} \approx p + (n-p^2)/(2p)$), the amount should be 10 + (108 − 100)/(2 · 10) = 10 $^{2}/_{5}$ = 10 3′ 15′, which is less precise and 120′ larger than our value (and whose square is also inconveniently larger than 108), so that one needs a second go at the approximation. 10 $^{2}/_{5}$ − (108$^{4}/_{25}$ − 108)/(20 $^{4}/_{5}$) = 10 $^{2}/_{5}$ − 130′ = 10 3′ 20′ 60′ − (120′ − 1560′) = 10 3′ 20′ 120′ 1560′ = 1351/130. We may drop 1560′ as insignificant. A sexagesimal version of this will not be pretty, as $^{1}/_{13}$ is not a nice number: 10;23,32,18,27,41,32,18,27,42

[b]The multiplication of 6 terms by 6 terms will yield 36 terms, some of obviously small impact, which may be ignored. Nonetheless, the error left will be about 8′ 40′. We can imagine some of the error from rounding (still very high), but some must come from a calculating error. For example, the sum of 21, 2, $^{3}/_{20}$, $^{7}/_{3}$, $^{2}/_{9}$, $^{7}/_{10}$, $^{7}/_{60}$ reduces to 26 + 3′ + 5′ + $^{2}/_{9}$ + 60′+ extra values = 26 $^{2}/_{3}$ 10′ 180′+ extra values or even 26 $^{1}/_{2}$ 4′ 45′ (bumping $^{3}/_{4}$ to $^{5}/_{6}$?). The author could have achieved a more reasonable result by truncating or rounding from the exact value: 26 2/3 10′ 90′ 240′ 1800′ 10368′. A very different source of the error might result if we accept the thesis of Friberg (2005, op. cit.) that the original calculations were done in sexagesimal. For example, the author truncates 26;46,57,20,50, mis-writes the number as 26;56, and translates this into Egyptian fractions. Perhaps, but this is hardly a more satisfying story than someone simply writing $^{5}/_{6}$ for $^{2}/_{3}$

To review, given the sides of the equilateral triangle, the text finds its height, $h_{\text{triangle}}$, to find its area, $area_{\text{triangle}}$. It next finds the height of the segment, $h_{\text{segment}}$, as 1/3 $h_{\text{triangle}}$. It then uses the base of the segment (and side of the triangle) to find the area of the segment according to the procedure: $(h_{\text{segment}} + b)/2 \cdot h_{\text{segment}}$. Crucial to the example is that the inscribed figure is a regular polygon, so that all the segments are equal. Hence, it triples this and adds it to the area of the triangle to find the area of the circle as 143 10′ 20′. Next, it finds the area of the circle directly. The diameter $d$ is $h_{triangle} + h_{\text{segment}}$, and the circumference $c$ is 3-times this. The area by Problem 35 is $c/3 \cdot c/4$ (in effect, $(\pi \cdot d)/\pi \cdot (\pi \cdot d)/4$). Therefore, we have found the area of the circle in two different ways and they do not match by 49/60, according to the text. This is not a trivial difference, or so I shall argue. The discrepancy is reduced but not eliminated by the error in the calculation of the segment (row I).

I will not pretend to be able to explain the calculation that follows, which Parker reports to be in a fragmentary state[14]; however, the concluding line would appear to be a lowering of the height to accommodate the lower area calculated by taking the area from the sum (K 143 10′ 20′), as the height will be 1/4 the diameter, which is the square root of the area + 1/3 the area (Probs. 32–3).[15]

| | | | | |
|---|---|---|---|---|
| R | Revised $h_{\text{segment}}$ | 3 3′ 15′ 35′ 49′ | 3 3′ 15′ 35′ 49′ 196′ 11760′ | 3;26,56,19,35,31 Corrected: 3;27,15 |

This reduction of the height might then lead to a reduction in the area of the segment (now, 26 1/2 8′ 72′ 392′ 6174′ 172872′) and of the triangle (now, 62 14′ 98′) and who knows what else. My point is simply that the author knows that it is important to pursue the anomaly of the difference in value that the two methods of area measurement produce. Someone comparing the two methods of calculation will inevitably wind up with a problem, simply because, barring any other error, the direct calculation of the circle should be about 0.9% larger than the calculation by summing the triangle and segments. The author has got something like this right. The two values should not gel. Error must be noted!

Let us now look at problem 37. This problem also involves finding the area of a circle, but does not perform the check. Rather it starts with both the area and diameter of the circle as given. Presumably, the student is welcome to perform the calculation to see that the given value is correct, but it will be unnecessary. The problem employs a square inscribed in a circle, so that there will be 4 segments whose height will be half the difference of the side of the square and the common diameter of the square and the circle. The circle is given as 675 square cubits and the diameter $d = 30$, obviously the value we shall work with, with 675 the check at the end (Fig. 6).

[14] The text starts with 1 5/6 45, which is either $d - b$ or should be 13 $^{5}/_{6}$ 45′ $= d$. It then subtracts 80 $^{2}/_{3}$ 10′ 30′ 150′, which should be the area of the three segments with 150′ added (but why—it does make the fraction more normal to us, 23/30 vs. 4/5), from some number, 82… (presumably 82 1/2 + …, the result being 1 5/6…. But there is no calculated number that starts 82. It subtracts 1 $^{1}/_{2}$ from 4 $^{1}/_{2}$ (not in the extant text) to get 3, and gets 4 $^{5}/_{6}$ before ending (for us) with the final number.

[15] My analysis of the calculation is slightly different from Parker's, who thinks that the fraction at the end should be 39. If we convert to standard fractions, the area is: 143 $^{3}/_{20}$. $^{4}/_{3}$ of this will be: 190 $^{13}/_{15}$, whose square root is closest to 14 (an upper approximation), so that the approximation of the square root will be: $14 - (14^2 - (190 + 13/15)) \big/ (2 \cdot 14) = 13\ ^{49}/_{60}$, and $^{1}/_{4} \cdot 13\ ^{49}/_{60} = 3\ ^{109}/_{240}$, which can be expanded as above. This number is smaller than 3 3′ 15′ 35′ 39′; so I do not know how Parker came to his result, but using a lower bound such as 13 1/2 for the first approximation would lead to a result such as he suggests. However, none of this appears in the papyrus as Parker presents it.

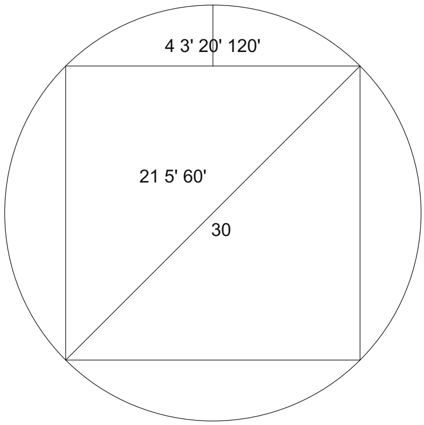



**Fig. 6** P. Cairo, Problem 37. Segment on a square

| Index | Procedure | Calculation | Corrected | Sexagesimal |
|---|---|---|---|---|
| A | $d \cdot d$ | 900 | | |
| B | A/2 | 450 | | |
| C | $\sqrt{B} = b_{\text{segment}} =$ side of the square | 21 5′ 60′ | 21 5′ 70′?[a] | 21;13 |
| D | $C^2$ | 450 | | |
| E | $d - C = 2h_{\text{segment}}$ | 8 $^2/_3$ 10′ 60′ | | 8;47 |
| F | $E/2 = h_{\text{segment}}$ | 4 3′ 20′ 120′ | | 4;23,30 |
| G | F + C | 25 $^1/_2$ 10′ 20′ 120′ | | 25;36,30 |
| H | G/2 | 12 $^2/_3$ 10′ 30′ 240′ | | 12;48,15 |
| I | $H \cdot F = (b_{\text{segment}} + h_{\text{segment}})/2 \cdot h_{\text{segment}} = \text{area}_{\text{segment}}$ | 56 4′ | The text fudges upwards from 56 5′ 60′ 90′ 300′ 1920′ | Text: 56;15 Corrected: 56;13,53,52,30 |
| J | $4 \cdot I = \text{area}_{4\ \text{segments}}$ | 225 | 224 $^5/_6$ 20′ 39′ 120′ 720′ 7200′ | Text: 225 Corrected: 3,44;55,35,30 |
| K | D + J | 675 | 674 $^5/_6$ 20′ 39′ 120′ 720′ 7200′ | 11,14;55,7,48,… |

[a]Suppose the estimate is 21, then the approximation will be $21+(450-441)/(2 \cdot 21) = 21+{}^9/_{42}$, which reduces to 21 5′ 70′. The author then makes the result base sixty friendly. 21 5′ 70′ is above the true value and is obviously, therefore, a closer approximation than 31 5′ 60′

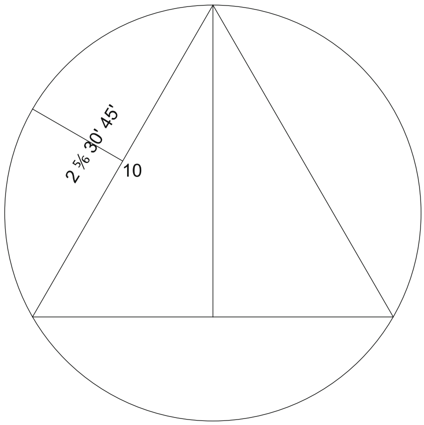



**Fig. 7** P. Cairo, Problem 38. The segment on an equilateral triangle

The value of the area of the circle should be 675 given the diameter of 30, that is $3 \cdot 30 = 90$. We then multiply $^{1}/_{3} \cdot 90$ by $^{1}/_{4} \cdot 90$. However, there is some fudging in K that might look minor, except that there was already fudging that increased the size of the side of the square, which decreased the height. Since, in the procedure $(b + h)/2 \cdot h$, a slight decrease in $h$ (caused by the increase in $b$) intuitively will have more effect than a small increase in $b$, the author has to do a little work to make the numbers work out. In fact, it will turn out that this is the one of two cases where, on the assumption that $\pi = 3$, the rule that the area of the segment $= (b + h)/2 \cdot h$ works perfectly, the other being the semicircle, that is, whether the author knows it or not, the method of determining the area of the circle as the sum of the inscribed, regular polygon and the segments, as determined by the Ancient method, should yield the same result as finding the area directly from the circumference only in the case of the square (the semicircle does not use the method). I suspect the author knows it. After all, there is no comment here on the issue.

Problem 38 starts with the calculation of the area of circle. Here, the side of the equilateral triangle and base of the segment $b = 10$ (Fig. 7).

| Index | Procedure | Calculation | Corrected | Sexagesimal |
|---|---|---|---|---|
| A | $b/2 \cdot b/2$ | 25 | | |
| B | $b \cdot b$ | 100 | | |
| C | B – A | 75 | | |
| D | $\sqrt{\text{C}} = h_{\text{triangle}}$ | 8 $^{2}/_{3}$ | $= 9 - (9^2 - 75)/(2 \cdot 9)$ | 8;40 |
| E | $\text{D} \cdot b/2 = \text{area}_{\text{triangle}}$ | 43 3′ | | 43;20 |
| F | $\text{D}/3 = h_{\text{segment}}$ | 2 $^{5}/_{6}$ 30′ 45′ | | 2;53,20 |
| G | $\text{D} + \text{F} = d_{\text{circle}}$ | 11 $^{1}/_{2}$ 30′ 45′ | | 11;33,20 |
| H | $3 \cdot \text{G} = c_{\text{circle}}$ | 34 $^{2}/_{3}$ | | 34;40 |

| Index | Procedure | Calculation | Corrected | Sexagesimal |
|---|---|---|---|---|
| I | H/3 | 11 $^1/_2$ 30′ 45′ | | 11;33,20 |
| J | H/4 | 8 $^2/_3$ | | 8;40 |
| K | I · J = $\text{area}_{\text{circle\_direct}}$ | 100 15′ 90′ | 100 9′ 27′ | Text: 1,40;04,40 Corrected: 1,40;08,53,20 |
| L | $F + b$ | 12 $^5/_6$ 30′ 45′ | | 12;53,20 |
| M | (F + $b$)/2 | 6 3′ 9′ | | 6;26,40 |
| N | M · F = $\text{area}_{\text{segment}}$ | 18 $^1/_2$ 10′ 60′ | 18 $^1/_2$ 10′ 60′ 1620′ | Text: 18;37 Corrected: 18;37,02,13,20 |
| O | 3 · N = $\text{area}_{3\text{ segments}}$ | 55 $^5/_6$ 60′ | 55 $^5/_6$ 60′ 540′ | Text: 55;51 Corrected: 55;51,06,40 |
| P | O + E = $\text{area}_{\text{circle\_sum}}$ | 99 6′ 60′ | 99 6′ 60′ 540′ | Text: 1,39;11 Corrected: 1,39;11,06,40 |
| Q | K – P | $^5/_6$ 18′ 180′ | $^5/_6$ 9′ 54′ | Text: 0;53,40 Corrected: 0;57,46,40 |

Again, as we saw in the earlier version, the direct calculation was slightly larger than the calculation by summation. Therefore, the author here too notes the discrepancy between the direct method and summation of the triangle and the segments as derived by the Ancient method.

It is an important part of the P. Cairo treatise to check results. Here, there are two methods of approach, one the direct method of taking $^1/_{12}$ of the circumference squared: $^1/_4$ (3$d$) · $^1/_3$ (3$d$). The other takes the area of a regular $n$-gon with sides of length $b$ and finds the height $h$ of the segments. The area of the circle is then $n \cdot ((b+h)/2 \cdot h)$ + the area of the $n$-gon. For one case, the results are the same, albeit with some noticeable fudging, while for the other there is a noticeable and noted error. It must have been part of the pedagogy of the method to know this, while at the same time, it is hard to know how many examples of such checking could be afforded to the student. One suspects that these are nigh they.[16] As to the Ancient method for finding the area of the segment, well, it works perfectly on a semicircle, perfectly on a side of a square, and fairly well on one figure in between, the equilateral triangle. Therefore, it works fairly well. The difficulty for the method is that one cannot usually check it out by subtracting a known figure from an enclosing circle. Should we assume, as is plausible, that the reader took it for granted that the procedures for the circle and inscribed polygon were accurate and the Ancient method an approximation, to be used

[16] A seventeenth century BCE text from Susa has values for the areas of pentagons, hexagons, and heptagons (see Addendum II to Sect. 3). Therefore, it is possible that Egyptians also had more.

with caution? Furthermore, does the author have any inkling why it works perfectly in the case of the square but not in the case of the equilateral triangle?

A final issue lies behind the method of checking. Problem 36 seems to continue the adjustment by setting up a new height of the segment. But it also suggests another technique, which we can call Subtraction and Division. Take the area of the circle as derived from the diameter and subtract the area of the triangle to get the three segments $\left(143\ {}^{5}\!/_{6}\ 45'\ 10'\ 30' - 62\ \ 3'\ \ 60' = 81\ \ {}^{1}\!/_{2}\ \ 10'\ \ 30'\ \ 180'\right)$, and taking a third, we get the area of the segment (27 6′ 30′ 90′ 540′), where the area of the segment $= 1/n \cdot$ (${}^{3}\!/_{4}\ d^2$ − the area of the inscribed, regular $n$-gon). It is perfectly possible, but not very likely that the author of P. Cairo did not see this as a trivial consequence of his method of checking. It may well have figured, however, in Old Babylonian methods for calculating segments on squares and other figures.

## 4 A late Babylonian example on a 3–4–5 rectangle, with amazing coincidences

The appearance of the Ancient method in P. Cairo establishes it outside the world of Greek mathematics, but it would also be nice to establish it within the Middle Eastern world and to establish it as a general method and not something for special cases of segments on equilateral triangles and squares where it is not needed in any case.[17] It is also important because it is in this way that we shall better come to understand Hero's strange remarks about the method. Friberg (2005, p. 133), states, "In Babylonian mathematics, the use of the rule is not documented." Therefore, it is nice to report there is a very nice example of the method on a tablet copied from a writing board in the late fifth century BCE, from Uruk, W 23291-x problem 1,[18] which employs the Ancient method. Since this is not evident in the original publication, which interprets the argument very differently, and because I make no claims to being a scholar of this material, I will reproduce the translation of Friberg et al. (1990, p. 487) and will give my interpretation on the right. The diagram of the tablet has a circle with four numbers, the three areas and the length of the circumference of the semicircle, imitated in Fig. 8a. My figure, Fig. 8b, produces the elements of my interpretation. Here, $b$ is the diameter of the circle and the cross-diameter of the inner figure; $p$ the circular-arc of the semicircle; $h$ will be the height of a segment on $b$, and $w$ the width of two segments together, so that $w = 2\ h$. $C$ will be a constant whose properties will be determined later. Since it will be important for my interpretation where the sexagesimal place is, I will freely use the semi-colon to mark the fractional part of a number. The inner figure, called a 'heart', will turn out to be formed from two equal circular segments.

17 To put this in perspective, the survey of evidence in Friberg (2005, 133–136) suggests that there is only indirect evidence for the Ancient method.

18 Published in Friberg et al. (1990, pp. 487–93). See also Friberg (2007, pp. 321–325). For the provenance of the tablet and the Šangû-Ninurta family that produced and owned it, see Robson (2008, pp. 227–37).

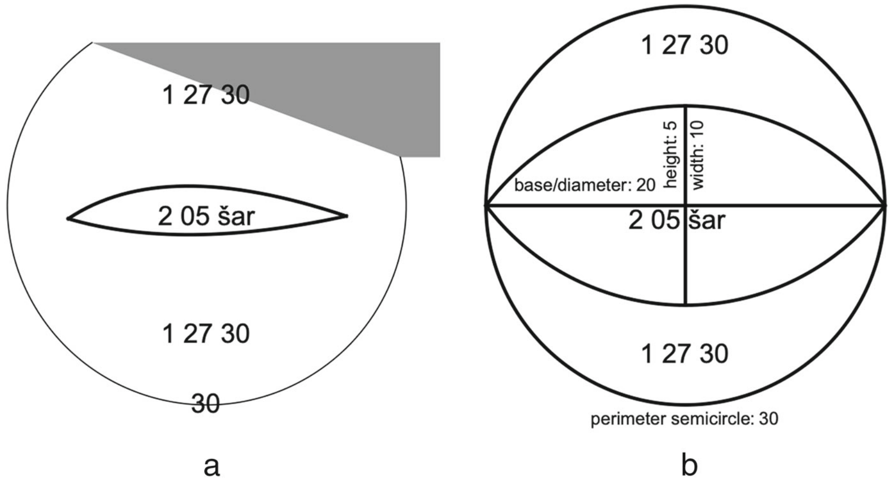



**Fig. 8** W 23291-x, problem 1: **a** (as on tablet); **b** (precise version, with values added)

| | |
|---|---|
| 10 is my [??] | |
| 10$^?$ of the extension$^?$ of the heart is what? | It will turn out that $10 = w$ |
| 20 steps of 10 is 3,20 | $b = 20$ (if $30 = p$, then $(30/3) \cdot 2 = b$, cf. Friberg et al. 1990) |
| | $b \cdot w = 3{,}20$ |
| Since 10 is $^1/_2$ of [20], | $w/b = 10/20 = 1/2$ |
| (0;0)7 30, $^1/_2$ of (0;)15 | $(w/b) \cdot {}^1/_4 = {}^1/_2 \cdot 0{;}15 = 0{;}07{,}30$ |
| to (0;)30 pair, then (0;)37 30 | $0{;}30 + {}^1/_4\, w/b = 0{;}37{,}30$ |
| 3$^!$ 20 steps of (0;)37 30 is 2 05, | $(w \cdot b) \cdot (0;30 + {}^1/_4 w/b) = 3,20; \cdot 0;37,30 = 2,05; = \text{area}_{\text{heart}}$ |
| 1 iku 25 šar, this is the area | 1 iku 25 šar, this is the area$_{\text{heart}}$, where 1 iku = 100 šar |
| (0;)30 of the crescent-field, | $p = 30;$ |
| the area is what? | |
| (0;)30 steps of (0;)30 is (0;)15, | $p^2 = 30; \cdot 30; = 15{,}00;$ |
| <steps of> 5 50 go is 1 27;30 | $C = 0{;}05{,}50$, so that $15{,}00; \cdot 0{;}05{,}50 = 1{,}27{;}30$ šar |
| 1 ubu 37 $^1/_2$ šar | = 1 ubu 37 $^1/_2$ šar, where 1 ubu = 50 šar |
| this is 1 crescent-field | |
| Steps of 2 I have gone | |
| 1,27;30 steps of 2 go, then 2,55, | two fields = $2 \cdot 1{,}27{;}30 = 2{,}55$ šar |
| 1 iku 1 ubu 25 šar, | |
| these arc 2 crescent-fields | |

| | |
|---|---|
| Heap them, then | |
| all of them, then 3 iku | total area = 2,55 + 2,05 šar = 5,00 šar = 3 iku |

Obviously, the interpretation derives from its coherence. First, it is important to see that this is equivalent to the Ancient method. The algorithm on this interpretation exemplifies the following procedure:

$$(w \cdot b) \cdot (1/2 + (w/b) \cdot 1/4).$$

Since the heart is equal to 2 segments and $w = 2h$:

$$w \cdot b \cdot (1/2 + 1/4\ w/b) = 2h \cdot b \cdot (1/2 + 1/4 \cdot 2h/b) = 2 \cdot (b + h)/2 \cdot h$$

Conceptually, the method takes half $w \cdot b$ and adds on an extra amount, something that we might expect in a Babylonian context, in effect the square on half the width. But this leads us back to our initial questions about the method itself.

In their publication of the tablet, Friberg et al. (1990) provide a very different analysis of the problem. Although I think it is wrong as a reconstruction, the mathematics underlying their reconstruction will prove to have some truly startling and weird consequences which may prove very important for the history of the Ancient method. They start with Old Babylonian constants for two figures: the grain, basically taking two segments on a square inscribed in a circle and forming an oval from them; and the ox-eye, taking two segments on an equilateral triangle inscribed in a circle and, again, forming an oval (Fig. 9).[19] The constants are numbers for the base, width (double the height of the segment), and the area (presumably formed by a version of Subtraction and Division), based on taking the arc of the segment as a unit, with $\pi = 3$. The user can then multiply these constants by the size of the similar arcs on other figures to determine the lengths or by their squares to determine the areas of the larger or smaller similar figures. These constants appear in lists of many sorts of constants or coefficients on several, mostly Old Babylonian tablets.[20] The calculations assume, again, the length of the arc, $a$, on the segment as 1;0 (keeping in mind that 1 is indeterminate in its value and can be any $1 \cdot 60^{\pm n}$), with $\pi = 3$, $\sqrt{2} = 1;25$ (i.e., 17/12), $\sqrt{3} = 1;45$ (i.e., 7/4), with $c$ the circumference, $d$ the diameter, $A$ the area. Therefore, since $a_{\text{grain}} = c_g/4$, the three constants for the grain will be (with $c_g = 4a_{\text{grain}}$, $d_g = c_g/\pi = 4a_{\text{grain}}/\pi = 1;20 \cdot a_{\text{grain}}$):[21]

$$b_g = 4 \cdot a_{\text{grain}}/\pi \cdot \sqrt{2}/2 = 1;20 \cdot 0;42,30 \cdot a_{\text{grain}} = 0;56,40 \cdot a_{\text{grain}} = 0;56,40$$
$$w_g = 4 \cdot a_{\text{grain}}/\pi - b_g = (1;20 - 0;56,40) \cdot a_{\text{grain}} = 0;23,30 \cdot a_{\text{grain}} = 0;23,30$$

[19] The first successful analysis of these figures is Vaiman (1963, pp. 75–76, 79–80). See also Robson (1999, pp. 45–48).

[20] See Robson (1999) for a comprehensive discussion of how constants work throughout Old Babylonian mathematics and administration, her preferred term being 'coefficients'. She discusses eight Old Babylonian lists and four later lists.

[21] See Friberg et al. (1990, p. 490) for an elegant geometrical derivation of the constants for the grain and ox-eye.

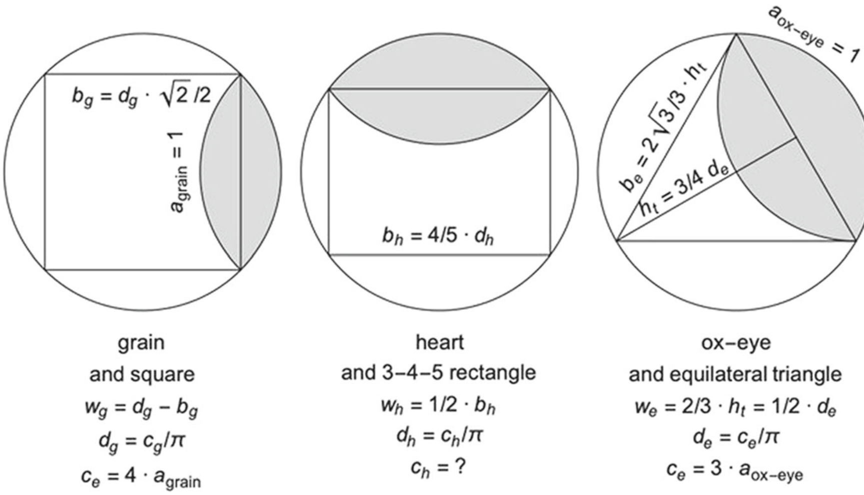



**Fig. 9** The grain, the heart, and the ox-eye on their respective polygons and known lineal dimensions for the base figures

$$\begin{aligned} A_g &= \tfrac{1}{2} \cdot (\text{circle} - \text{square}) = \tfrac{1}{2}\,(c_g^2/(4\,\pi) - b_g^2) \\ &= \tfrac{1}{2} \cdot \left((4 \cdot a_{\text{grain}})^2/(4\,\pi) - (4 \cdot a_{\text{grain}}/\pi \cdot \sqrt{2}/2)^2\right) \\ &= \left(2 \cdot a_{\text{grain}}^2/\pi - 4 \cdot a_{\text{grain}}^2/\pi^2\right) = (0;40 - 0;26,40) \cdot a_{\text{grain}}^2 \\ &= 0;13,20 \cdot a_{\text{grain}}^2 = 0;13,20 \end{aligned}$$

Again, since $a_{\text{ox-eye}} = c_e/3$, the three constants for the ox-eye will be (with $c_e = 3a_{\text{ox-eye}}$, $d_e = c_e/\pi = 3/\pi \cdot a_{\text{ox-eye}} = a_{\text{ox-eye}}$, and the height of the inscribed triangle, $h_t = \tfrac{3}{4}d_e$):

$$\begin{aligned} b_e &= h_t \cdot 2\sqrt{3}/3 \cdot a_{\text{ox-eye}} = \tfrac{3}{4} \cdot 1;10 \cdot a_{\text{ox-eye}} = 0;52,30 \cdot a_{\text{ox-eye}} = 0;52,30 \\ w_e &= \tfrac{1}{2}\, d_e = 0;30 \cdot a_{\text{ox-eye}} = 0;30 \\ A_e &= \tfrac{2}{3}\,(\text{circle} - \text{triangle}) = \tfrac{2}{3}\,(c_e^2/(4\pi) - \tfrac{1}{2}\, b_e \cdot h_t) \\ &= \tfrac{2}{3}\left((3 \cdot a_{\text{ox-eye}})^2/(4\pi) - \tfrac{1}{2} \cdot (\tfrac{3}{4}d_e)^2 \cdot 2\sqrt{3}/3\right) \\ &= \tfrac{3}{2} \cdot a_{\text{ox-eye}}^2/\pi - 9/(8\pi^2) \cdot \sqrt{3} \cdot a_{\text{ox-eye}}^2 = (\tfrac{1}{2} - \tfrac{7}{32}) \cdot a_{\text{ox-eye}}^2 \\ &= 0;16,52,30 \cdot a_{\text{ox-eye}}^2 = 0;16,52,30 \end{aligned}$$

Friberg et al. then normalize these two sets of figures to a common value with the base of the heart, $b_{\text{joint}} = 20$. This means that we need to multiply the remaining values for the grain by $b_{\text{joint}}/b_g$ or its square for the area, and for the heart by $b_{\text{joint}}/b_e$ or its square for the area. Hence (I will give a general analysis later, but note that they calculate these directly from the values for the circumferences by the algorithms above

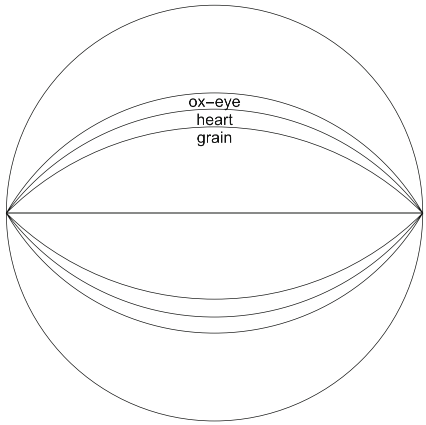



**Fig. 10** Comparison of the grain, the heart, and the ox-eye

and not from the constants in the tables that result from the algorithms) (Fig. 10):

$$w_{\text{g-joint}} = 8;20$$
$$A_{\text{g-joint}} = 1{,}40$$
$$w_{\text{e-joint}} = 11;40$$
$$A_{\text{e-joint}} = 2{,}30$$

From these two areas, we can compute the area of the crescents composed from the semicircle on their common base (the long diameter of the grain and ox-eye), $b_{\text{joint}} = 20$, less half the figure (i.e., one circular segment):

$$\text{Semicircle} = {}^1\!/_2\, b^2_{\text{joint}}\, \pi/4 = {}^3\!/_8\, 20^2 = 2,30$$
$$\text{Crescent}_{\text{grain-joint}} = \text{semicircle} - {}^1\!/_2 A_{\text{g-joint}} = 2{,}30 - {}^1\!/_2\; 1{,}40 = 1{,}40$$
$$\text{Crescent}_{\text{ox-eye-joint}} = \text{semicircle} - {}^1\!/_2 A_{\text{e-joint}} = 2{,}30 - {}^1\!/_2\;\; 2{,}30 = 1{,}15$$

Now, the average of these, ${}^1\!/_2 \cdot \left(\text{Crescent}_{\text{grain-joint}} + \text{Crescent}_{\text{ox-eye-joint}}\right) = {}^1\!/_2\,(1{,}40 + 1{,}15) = 1{,}27;30$. Amazingly, this just happens to be the area of the crescent outside the heart. Therefore, they hypothesize that the heart (double the segment that is the semicircle less the crescent) is calculated as the mean between the grain and the heart. If we normalize the constant for the grain and the heart, taking the arc of the semicircle on the joint base, $a_{\text{semicircle}} = 1$ (they take 1,0, but it really does not matter), the factor for conversion will be $(a_{\text{crescent}}/p)^2 = (1/30)^2 = 0;02^2 = 0;00{,}04$, since the semi-circumference $p = 30$ and the base semi-circumference $a_{\text{crescent}} = 1$:

$$\text{Constant}_{\text{crescent-grain-joint}} = 0;00{,}04 \cdot 1{,}40 = 0;06{,}40$$
$$\text{Constant}_{\text{crescent-ox-eye-joint}} = 0;00{,}04 \cdot 1{,}15 = 0;05.$$

However, the average of these is 0;05,50, the very constant that we find in the text. On their view, the Babylonians started with the crescent as cut off by a segment with height half the radius, a perfect crescent, and then construct the heart as a complementary figure. Using these as a basis for the area of the heart, they then reconstruct the calculation of W23291-x prob. 1 as taking the mean of adjusted constants for the grain and the ox-eye as a multiplier of the rectangle $b \cdot r = 20 \cdot 10 = 3{,}20 = b \cdot w_{\text{heart}}$, where $r$ is the height of the semicircle (i.e., the radius):

$$A_{\text{g-joint}} = 1{,}40 = C_{\text{g-joint}} \cdot b \cdot r = C_{\text{g-joint}} \cdot 3;20, \quad \text{so that} \quad C_{\text{g-joint}} = 0;30$$
$$A_{\text{e-joint}} = 2{,}30 = C_{\text{e-joint}} \cdot b \cdot r = C_{\text{e-joint}} \cdot 3;20, \quad \text{so that} \quad C_{\text{e-joint}} = 0;45.$$

Hence, the area of the heart will be calculated from the mean of $C_{\text{g-joint}}$ and $C_{\text{e-joint}}$.

On their interpretation, the calculation proceeds as follows (with $\textit{diff} = C_{\text{e-joint}} - C_{\text{g-joint}}$):

$$b \cdot r \text{ (since } r \text{ is half of } b, \text{ the diameter of the circle)} \rightarrow {}^1\!/_2\, \textit{diff} \rightarrow C_{\text{g-joint}} + {}^1\!/_2\, \textit{diff}$$
$$\rightarrow b \cdot r \cdot (C_{\text{g-joint}} + {}^1\!/_2\, \textit{diff})$$
$$20 \cdot 10 = 3{,}20 \rightarrow {}^1\!/_2\; 0;15 = 0;07;30 \rightarrow 0;30 + 0;07{,}30 = 0;37{,}30$$
$$\rightarrow 3{,}20 \cdot 0;37{,}30 = 2{,}05$$

It is utterly amazing that these numbers should work out so well, which commends the interpretation, even if Friberg et al. have to done much hunting to find them; for many variants on how the calculations are accomplished (such as taking ${}^1\!/\!\sqrt{3}$ instead of ${}^1\!/_3 \cdot \sqrt{3}$, or taking $(\sqrt{3})^2$ as ${}^{49}\!/_{16}$, or working directly from the given constants for the grain and ox-eye) would have led to very different results. But there are several problems with it. In effect, the calculation amounts to finding $b \cdot r \cdot C_{\text{h-joint}}$, where the heart constant $C_{\text{h-joint}} = C_{\text{g-joint}} + {}^1\!/_2\ diff$. First, do we expect $C_{\text{h-joint}}$ to be calculated out, rather than simply given? Nor are we warned what these constants are nor that they come from constants for other figures. Where do these numbers come from? Why is *diff* given and not $C_{\text{e-joint}}$? They are numbers truly out of the blue. In fact, they are not standard constants at all, since the constants are supposed to be based on the length of the arc of the heart taken as 1, and not, as here, the rectangle enclosing the heart, namely $b_{\text{joint}} \cdot r = b_{\text{joint}} \cdot w_{\text{heart}}$, or some such thing. In fact, Friberg et al. rightly set the basis of the constant for the crescent to be the semi-circumference, where the arc is 30 times (or 0;30) the base value, 1;00, and the area of the $\text{crescent}_{\text{heart}} = 0;05{,}50$, but if this were also the basis for grain and ox-eye constants used to calculate the area of the heart, they are different from the actual constants we find in Old Babylonian texts and would need themselves to be calculated from the procedures used for calculating the grain and ox-eye constants, just as Friberg et al. do. And yet, if the basis were the semi-circumference, we would expect the calculation to be something like $30^2 \cdot 0;08{,}20 = 2{,}05$, where 0;08,20 is the average of 0;06,20 (grain) and 0;10 (ox-eye) for the semi-circumference = 1;00. Finally, line 4 in their translation, "Since 10 is ${}^1\!/_2$ of [20]," seems not to be part of their algorithm. Something is deeply amiss in all this. Yet, they have found a tantalizing coincidence. In fact, as amazing as this is, something even more amazing and zany lurks in the background, as this coincidence can be explained, but only, so far as I can tell, by two other, much weirder coincidences.

Let us look more carefully at the geometry behind the heart. Let us suppose that the base/diameter of the heart is a side of a figure inscribed in a circle, as are the grain and ox-eye. We can imagine a segment on the base of an isosceles triangle with base equal to height, or a segment on the side of an isosceles triangle with side to base to height of 5:6:4,[22] but most plausibly the heart is the double of the segment on the long side of the 3–4–5 (width, length, diameter) rectangle inscribed in a circle (or just the 3–4–5 triangle), the "favorite rational rectangle" of Babylonian mathematicians.[23] Although the arc on the width is very close to the side of a regular, inscribed decagon (see Sect. 5), and the length is close to the side of a regular, inscribed heptagon, neither seems to be relevant to Babylonian mathematics.[24] Therefore, there really is no arc known to our Uruk mathematician that he can use as the basis for the heart, as there is in the case of the grain and the ox-eye. Indeed, although the value for the crescent is based on a known arc, the semicircle, we expect, on the analogy of the grain and ox-eye, the constants for the heart to be based on the arc of the segment on the long side of the 3–4–5 rectangle, and a value for either arc may not be part of Babylonian mathematics.[25]

[22] At *Metrica* I 18, Hero will use 5 of these triangles (double 3:4:5 right triangles) to approximate the inscribed, regular pentagon, so that the segment on the base would be checkable by Subtraction and Division (see Sect. 5), but this triangle is from the center and is not inscribed in the circle, so that there is no trivial relation between the segments here and those on the approximate pentagon. The approximation may well be Old Babylonian, for which see Addendum II at the end of this section.

[23] See Robson (1999, p. 44) quoting Friberg. To see that this could be a 3–4–5 rectangle, the height of the segment is $h = {}^1\!/\!_2\ w = 5$, with $b = 20$. Therefore, if the segment is on a rectangle, the diameter will be $(b/2)^2/h + h = (20/2)^2/5 + 5 = 25$, and, by the Pythagorean rule, the width of the rectangle will then be 15. BM 85194 Rs. I 33–43, probs. 20, 21, consists of two problems, one finding the base of the segment from the height and circumference and the other the height from the circumference and base. See Neugebauer (1935, pp. 169–160), also Høyrup (2002, pp. 272–275) and Friberg (2007, pp. 43–46). One calculation in each problem is wrong, but with correct numbers. If we think of an inscribed rectangle with sides $b_1, b_2$, and heights of the segments on each as $h_1$ and $h_2$, with the diameter $d$ implicitly a third of the circumference $c$, then the procedure of prob. 20 may very naturally be read in this way, to find $b_1$: $c = 60$, $h_1 = 2 \to d(= c/3) = 20 \to h_1^2 = 4$(corrected: $2 \cdot h_1 = 4$) $\to b_2 = d - 2 \cdot h_1 = 20 - 4 = 16 \to d^2 = 6{,}40 \to b_2^2 = 4{,}16 \to d^2 - b_2^2 = 2{,}24 \to b_1 = \sqrt{d^2 - b_2^2} = 12$. Prob. 21 is in reverse, to find $h_1$: $c = 60, b_1 = 12 \to d(= c/3) = 20 \to d^2 = 6{,}40 \to b_1^2 = 2{,}24 \to d^2 - b_1^2 = 6{,}40 - 2{,}24 = 4{,}16 \to b_2 = \sqrt{d^2 - b_1^2} = 16 \to h_1^2 = d - b_1 = 20 - 16 = 4 \to h_1 = \sqrt{4} = 2$ (last 2 steps corrected: $2 \cdot h_1 = d - b_1 = 20 - 16 = 4 \to h_1 = {}^1\!/\!_2\ 2 \cdot h_1 = 2$). The base of the segment or chord could be seen as the short side 12 of a 12–16–20 rectangle, with the height of the segment equal to 2.

[24] See Addendum II at the end of this section. If the reconstruction is correct, we find a very different value for the heptagon.

[25] We can derive constants based on the arc on the heart, $a_{\text{heart}} = 1$, and the ratio of the circumference to this arc, here $f$, from $A_{\text{heart}} = 125$, and $b = 20$, along with the known equivalences for the basic figure, $d_{\text{h-basic}} = c_{\text{h-basic}}/3$, $b_{\text{h-basic}} = {}^4\!/\!_5\ d_{\text{h-basic}} = {}^4\!/\!_{15}\ c_{\text{h-basic}}$, the small side of the rectangle $s_{\text{h-basic}} = {}^3\!/\!_5\ d_{\text{h-basic}} = {}^3\!/\!_{15}\ c_{\text{h-basic}}$. Let $f \cdot a_{\text{heart}} = c_{\text{h-basic}}$, so that $d_{\text{h-basic}} = {}^1\!/\!_3 f \cdot a_{\text{heart}}$, and $m$ be the amount that is multiplied by $b_{\text{h-basic}}$ to get $b = 20 \cdot a_{\text{heart}}$. We then have two equations: $A_{\text{h}} = m^2 \cdot A_{\text{h-basic}}$, and $b = m \cdot b_{\text{h-basic}}$. The first may be expanded as: $125\ a_{\text{heart}}^2 = m^2 \cdot \left(2/f \cdot \text{circle} - {}^1\!/\!_2\ 3\text{–}4\text{–}5 \text{ rectangle}\right) = m^2 \cdot \left(2/f \cdot (f \cdot a_{\text{heart}})^2/12 - {}^1\!/\!_2 \cdot {}^3\!/\!_{15} \cdot {}^4\!/\!_{15} \cdot (f \cdot a_{\text{heart}})^2\right) = m^2 \cdot (f/6 - {}^2\!/\!_{75}\ f^2) \cdot a_{\text{heart}}^2$. The second equation is: $20\ a_{\text{heart}} = m \cdot {}^4\!/\!_{15}\ c_{\text{h-basic}} = {}^4\!/\!_{15}\ m \cdot f \cdot a_{\text{heart}}$. It follows that $m = 22$, and $f = {}^{75}\!/\!_{22}$. I could find no intuitive way of reconstructing a derivation of $f$ from the grain and ox-eye or in any other way; moreover, the denominator of $f$ is unfriendly so that a Babylonian would find $f$ very distasteful, 3;24,32,43,38,10,54,32,…, although it would be needed to determine all the constants for the heart. I do not see a plausible analysis along these lines.

It is possible, however, that some sort of calculation such as Friberg et al. propose, amazing as it is, justified a different procedure. To see this, I shall, for brevity, provide algebraicized reductions. Following Friberg et al., I shall use the base $b$ to construct a factor by which to multiply the constant, first setting $\pi = 3$, and then $\sqrt{2} = 17/12$ and $\sqrt{3} = 7/4$. The four basic values for the grain and ox-eye, may be calculated, the width, the base, and the area by the Subtraction and Division method (*Asd*), also, but not part of the known texts, by the Ancient method (*Aam*), where $Aam = (b + w/2) \cdot w/2$. The diameter of the circle for the grain, $d_{\text{g-basic}} = 4/\pi$, and for the ox-eye, $d_{\text{e-basic}} = 3/\pi$. The value, $b = 20$, proves irrelevant to our concerns:

$$b_{\text{g-basic}} = d_{\text{g-basic}}/\sqrt{2} = 2\sqrt{2}/\pi$$
$$w_{\text{g-basic}} = d_{\text{g-basic}} - b_{\text{g-basic}} = 2 \cdot (2 - \sqrt{2})/\pi$$
$$Asd_{\text{g-basic}} = 1/2\ (\text{circle} - \text{square}) = 1/2\ (d^2_{\text{g-basic}} \cdot \pi/4 - b^2_{\text{g-basic}}) = 2(\pi - 2)/\pi^2$$
$$Aam_{\text{g-basic}} = (b_{\text{g-basic}} + w_{\text{g-basic}}/2) \cdot w_{\text{g-basic}}/2$$
$$= 1/2\ 2 \cdot (2 - \sqrt{2})/\pi \cdot (2\sqrt{2}/\pi + 1/2\ 2 \cdot (2 - \sqrt{2})/\pi)$$
$$= 2/\pi^2$$

$$b_{\text{e-basic}} = h_t \cdot 2/\sqrt{3} = (3/4\ d_{\text{e-basic}}) \cdot 2/\sqrt{3} = 3\sqrt{3}/(2\pi)$$
$$w_{\text{e-basic}} = 2 \cdot (1/4\ d_{\text{e-basic}}) = 3/(2\pi)$$
$$Asd_{\text{e-basic}} = 2/3\ (\text{circle} - \text{triangle}) = 2/3\ (d^2_{\text{e-basic}} \cdot \pi/4 - 1/2 \cdot 3/4\ d_{\text{e-basic}} \cdot b_{\text{e-basic}})$$
$$= 2/3\ (9/\pi^2 \cdot \pi/4 - 3/8 \cdot 3/\pi \cdot 3\sqrt{3}/(2\pi))$$
$$= 3\ (4\pi - 3\sqrt{3})/(8\pi^2)$$
$$Aam_{\text{e-basic}} = (b_{\text{e-basic}} + w_{\text{e-basic}}/2) \cdot w_{\text{e-basic}}/2 = 9\ (1 + 2\sqrt{3})/(16\ \pi^2).$$

For the widths, one will multiply the basic value by these factors, where $b_{\text{joint}} = b$:

$$b_{\text{joint}}/b_g = b \cdot \pi/(2\sqrt{2}) = (b \cdot \pi \cdot \sqrt{2})/4$$
$$b_{\text{joint}}/b_e = b \cdot 2\pi/(3\sqrt{3}) = (2b \cdot \pi \cdot \sqrt{3})/9.$$

For the areas, one will use the squares of these values:

$$(b_{\text{joint}}/b_g)^2 = (b^2 \cdot \pi^2)/8$$
$$(b_{\text{joint}}/b_e)^2 = (4b^2 \cdot \pi^2)/27$$

From these, the widths and areas by both Subtraction and Division (*Asd*) and by the Ancient method (*Aam*) may be calculated:

$$w_{\text{g-joint}} = 2 \cdot (2 - \sqrt{2}) \cdot (\text{b} \cdot \pi \cdot \sqrt{2})/4 = b \cdot (\sqrt{2} - 1) = 5/12\ b = 0;25 \cdot b$$
$$Asd_{\text{g-joint}} = 2 \cdot (\pi - 2)/\pi^2 \cdot (b^2 \cdot \pi^2)/8 = 1/4\ b^2(\pi - 2) = 1/4\ b^2 = 0;15 \cdot b^2$$
$$Aam_{\text{g-joint}} = 2/\pi^2 \cdot (b^2 \cdot \pi^2)/8 = 0;15 \cdot b^2$$

$w_{\text{e-joint}} = 3/(2\pi) \cdot (2b \cdot \pi\sqrt{3}/9) = \sqrt{3}/3 \cdot b = {}^{7}/_{12} \cdot b = 0;35 \cdot b$
$Asd_{\text{e-joint}} = 3\ (4\pi - 3\sqrt{3})/(8\pi^2) \cdot (4b^2 \cdot \pi^2)/27$
$= b^2 \cdot (4\pi - 3\sqrt{3})/18 = 1/18\ b^2(12 - 3\sqrt{3}) = {}^{3}/_{8}\ b^2 = 0;22,30 \cdot b^2$
$Aam_{\text{e-joint}} = 9\ (1 + 2\sqrt{3})/(16\pi^2) \cdot (4b^2 \cdot \pi^2)/27$
$= {}^{1}/_{2}\ w_{\text{e-joint}}(b + {}^{1}/_{2}\ w_{\text{e-joint}})$
$= {}^{1}/_{2}\ b^2 \cdot (1 + 2\sqrt{3}) = {}^{3}/_{8}\ b^2 = 0;22,30 \cdot b^2$
$w_{\text{heart}} = b/2 = 0;30 \cdot b$
$w_{\text{heart-average}} = (w_{\text{g-joint}} + w_{\text{g-joint}})/2 = ({}^{5}/_{12}\ b + b \cdot 7/12) = b/2 = 0;30 \cdot b$
$Aave_{\text{heart}} = (Asd_{\text{g-joint}} + Asd_{\text{e-joint}})/2 = {}^{5}/_{16} \cdot b^2 = 0;18,45 \cdot b^2$
$Aam_{\text{heart}} = {}^{1}/_{2}\ w_{\text{h}}\ (b + 1/2\ w_{\text{h}}) = {}^{5}/_{16} \cdot b^2 = 0;18,45 \cdot b^2.$

There are three remarkable coincidences here, all a function of taking $\pi = 3, \sqrt{2} = {}^{17}/_{12}$, and $\sqrt{3} = {}^{7}/_{4}$. We have these five equivalences:

$$Asd_{\text{g-joint}} = Aam_{\text{g-joint}}.$$

From our Egyptian text, we should not be surprised that Subtraction and Divide and the Ancient method turn out equal if $\pi = 3$.

$$w_{\text{heart}} = w_{\text{heart-average}}.$$

This is just peculiar. There is nothing in the area of the heart being the mean of the areas of the grain and ox-eye that would suggest that the heart's width should also be the mean of their widths.

And now for something amazing,

$$Asd_{\text{e-joint}} = Aam_{\text{e-joint}} = 0;22,30 \cdot b^2.$$

Surprise! Well, I was surprised. This is precisely what we do not expect. The important observation of the Egyptian text implied that Subtraction and Division and the Ancient method here yield different results. Given $\pi = 3$, these two area calculations are equal iff $\sqrt{3} = 7/4$!

Next, we see that

$$Aam_{\text{heart}} = {}^{1}/_{2}\left(Aam_{\text{g-joint}} + Aam_{\text{e-joint}}\right).$$

This is just as astonishing and results solely from the taking $\sqrt{3} = {}^{7}/_{4}$!

Once we see this, we should also be astonished that the width of the heart is the average of widths of the ox-eye and the grain and that its area is also the mean between the areas of the ox-eye and the grain. Again, since the Ancient method is a parabolic curve, with $b$ fixed, it is not possible for values given $w_h + e$ and $w_h - e$ to have as their average the value for $w_h$:

$$\begin{aligned} &1/2\,(1/2\;(w_h + e)\,(b + 1/2\,(w_h + e)\\ &\quad + 1/2\,(w_h - e)\,(b + 1/2(w_h - e)) - 1/2 w_h\,(b + 1/2 w_h) = e^2/4. \end{aligned}$$

The weird effect can only be a result of the approximation taken for $\sqrt{2}$ and $\sqrt{3}$ and never their inverses or squares. Once we have these four coincidences, the fifth will trivially follow, that the mean of the areas of the ox-eye and grain will be the same as the area of the procedure, that is, as I have suggested, by the Ancient method, the oddity with which we began:

$$Aave_{\mathrm{heart}} = Aam_{\mathrm{heart}}.$$

I have no deeper explanation for these flabbergasting coincidences. But, for every case where a Babylonian mathematician is known to have found the area of a segment on a circle, the accurate Subtraction and Division method or an educated guess based on it turns out to be equivalent to the Ancient method. IF the Babylonian mathematicians delayed evaluating $\sqrt{3}$ and $\sqrt{2}$, that would be a very good confirmation, indeed, of both methods. Yet, since they are based on $(\sqrt{2})^2 = 2$ and $(\sqrt{3})^2 = 3$, the Babylonian values of the constants for the areas of the grain and the ox-eye suggest that they actually did. Of course, we do not know whether anyone actually followed this line of thought, but if they had, the Ancient method would have been confirmed in this weird way.

One might now ask why the Egyptian did not get this result as well. He starts with the side of the triangle as 12 and, using the procedure for square roots, takes the height of the triangle as $\sqrt{108} = 10\ 3'\ 20'\ 120' = 10\ {}^{47}/_{120}$. Since $\sqrt{108} = 6\sqrt{3}$, he might have chosen instead the less accurate $6 \cdot {}^{7}/_{4} = 10\ {}^{1}/_{2}$. Even so, he would not have gotten the same result for the two methods (the Ancient method would yield $27\ {}^{1}/_{8}$, and the Subtraction and Division method 28). He would have had likewise to have delayed taking the approximation for $\sqrt{3}$ to the very end, a process that is hard to see in his practice.

In any case, whether or not the analysis of Friberg et al. is a plausible account of Babylonian practice (I confess that I am unfit to judge), it seems much more likely that W23291-X prob. 1 employs the Ancient method and that what Friberg et al. attribute to Babylonian mathematicians would be at most background to their use of the Ancient method. We have also established in passing that there are three inscribed figures on which they built their accounts of segments, the equilateral triangle, the square, and the 3–4–5 rectangle, and that the Ancient method oddly 'works' for all three. Our next consideration will concern segments on an inscribed rectangle.

### 4.1 Addenda

I. As an addendum to this discussion, it might be useful to look at the garbled calculation of the segment on an equilateral triangle in the old Babylonian 'recombination' tablet,[26] BM 85194, Rs. III 1–8, prob. 29, which stumped Neugebauer (1935, pp. 188–190) and Friberg (2005, p. 133), although I will only slightly improve on Neugebauer's analysis. The text assumes the arc of the segment, $p = 1{,}0$, and the length of the base of the segment, $b = 50$. From this, we can infer that the angle is about 120°, so that it is reasonable to presume that the example is the segment on an equilateral triangle, although, taking $\sqrt{3} = 7/4$, the base should be 52,50, but with Neugebauer's $\sqrt{3} = 5/3$, $b = 50$.[27] Now the procedure, which on the surface makes no sense, is: $A = (p - b) \cdot b - (p - b)^2$, where the subtraction is miscalculated as $8{,}20 - 1{,}40 = 7{,}30$ (we expect 6,40). The first thing to notice is that for diameter $d$, $p = d$, and that the copyist missed this. Therefore, the procedure should probably be $d = p \rightarrow A = (d - b) \cdot b - (d - b)^2$. But, although, $d - b$ seems to make no geometrical sense, it can, in fact, be taken (with the height of the segment $h_s = {}^1\!/_4\, d$):

$$d - b = {}^1\!/_6\, d = {}^2\!/_3 \cdot h_s,$$

so that the procedure is now:

$$A = ({}^2\!/_3 \cdot h_s) \cdot b - ({}^2\!/_3 \cdot h_s)^2 = 10 \cdot 50 - 1{,}40 = 6{,}40,$$

which is equivalent to:

$$({}^2\!/_3\, h_s)\,(b - ({}^2\!/_3\, h_s)).$$

However, if, as happens elsewhere in BM 85192 (see footnote 23), the values are correct, but not the calculation, then the procedure might be:

$$A = ({}^2\!/_3 \cdot h_s) \cdot b - {}^1\!/_2\, ({}^2\!/_3 \cdot h_s)^2 = 10 \cdot 50 - 50 = 7{,}30,$$

and equivalent to:

$$({}^2\!/_3\, h_s) \cdot (b - {}^1\!/_3\, h_s).$$

By Subtraction and Division, $Asd = {}^1\!/_3\, ({}^3\!/_4\, d^2 - {}^1\!/_2\, b({}^3\!/_4\, d)) = {}^1\!/_4\, d^2 - 1/8\, (b \cdot d) = 15{,}00 - 6{,}15 = 8{,}45$, and by the Ancient method, $Aam = (b + h_s)/2 \cdot h_s = 65 \cdot 15/2 = 8{,}07;30$.[28] The value in BM 85194 is low but not way off. If this analysis is at all right, the mathematician who invented the procedure, whatever it was, seems to have been playing with some combination of the height and base of the segment. In any case, the constants for the grain and ox-eye show that the algorithm in BM 85194 was not among the best available in old Babylonian mathematics.

II. Bruins and Rutten (1961), Text III (Tablet I), from Susa, late seventeenth century BCE, includes a long list of constants, including constants for the area of the pentagon

[26] For the classification, see Friberg (2005, pp. 92–94).

[27] Neugebauer (1935, pp. 189) takes $\sqrt{3} = 5/3$, so that $b = 2 \cdot \sqrt{3}/3 \cdot ({}^3\!/_4\, d) = 50$. How the base comes to be 50 is not part of my analysis.

[28] The difference between the Ancient method and that in the problem, for what it is worth, is obviously: $h_s \cdot (13h_s - 3b)/18$.

(line 26): 1,40, for the area of the hexagon (line 27), 2,37,30, and for the area of the heptagon (line 28), 3,41.[29] These are unique among the many lists of constants, so that it is very difficult to know how well disseminated they were. Bruins and Rutten plausibly reconstruct the calculations underlying the constants, but it is important to keep in mind that these are reconstructions (as is so much in my paper). The basic idea is that the values are based in taking these figures as inscribed in a circle and taking the arc on a side as 1, in conformity with other figures, such as the grain and ox-eye. Text II consists of two diagrams decomposing a hexagon and heptagon into triangles formed by the radii and sides for the purpose of finding their areas, so that it is natural to assume that this is how the constants were determined. Then, most of the rest follows (with $c_n$ the circumference enclosing the $n$-gon, $r_n$ its radius, $b_n$ its side, $s_n$ its half side, $h_n$ the height of the triangle from the center):

$$
\begin{aligned}
& c_5 = 5\,a_5 \rightarrow r_5 = c_5/(2\pi)\cdot {}^5\!/\!_6 = 0;50 \\
& \quad \rightarrow s_5, h_5, r_5 \text{ are sides of a } 3-4-5 \text{ triangle} \\
& \quad \rightarrow s_5 = 0;30, h_5 = 0;40 \rightarrow \text{Area}_{\text{triangle}-5} = s_5 \cdot h_5 0;20 \\
& \quad \rightarrow \text{Area}_5 = 5 \cdot 0;20 = 1;40 \\
& c_6 = 6\,a_6 \rightarrow r_6 = c_6/(2\pi) = {}^6\!/\!_6 = 1 \\
& \quad \rightarrow b_6 = 1 \text{ (equilateral triangle)} \rightarrow s_6 = 0;30 \rightarrow h_6 = {}^1\!/\!_2 \quad \sqrt{3} \\
& \quad \rightarrow h_6 = {}^7\!/\!_8 \approx 0;52,30 \rightarrow \text{Area}_{\text{triangle}-6} s_6 \cdot h_6 = 0;26,15 \\
& \quad \rightarrow \text{Area}_6 = 6 \cdot 0;26,15 = 2;37,30 \\
& c_7 = 7\,a_7 \rightarrow r_7 = c_7/(2\pi) = {}^7\!/\!_6 \cdot 1;10 \\
& \quad \rightarrow b_7 = 1 \text{ (the arc and side almost co-inside?)} \rightarrow s_7 = 0;30 \\
& \quad \rightarrow h_7^2 = r_7^2 - s_7^2 = 1;21,40 - 0;15 = 1;06,40 = 1 + 0;20^2 \\
& \quad \rightarrow h_7 \approx 1 + 0;20^2/(2 \cdot 1) = 1;03,20 \text{ or } h_7 \approx 1 + 0;20^2/(2 \cdot 1;10) = 1 + {}^1\!/\!_{21} \\
& \quad \rightarrow \text{Area}_{\text{triangle-7}} \approx h_7 \cdot s_7 = 0;31,40 \text{ or Area}_{\text{triangle-7}} \approx 0;30 + {}^1\!/\!_{42} = {}^{11}\!/\!_{21} \\
& \quad \rightarrow \text{Area}_7 \approx 7 \cdot 0;31,40 = 3;41,40 \text{ or Area}_7 \approx 7 \cdot \left({}^{11}\!/\!_{21}\right) = {}^{11}\!/\!_3 = 3;40 \\
& \quad \rightarrow \text{Area}_7 \approx 3;41.
\end{aligned}
$$

IF this analysis of Bruins and Rutten is correct, then there is nothing to prevent a mathematician in Susa from determining the area of the segment or of a double segment by Subtraction and Division on any of these, as in the case of the grain and ox-eye:

$$
\begin{aligned}
\text{Area}_{\text{circle-5}} &= c_5^2/(4\pi) = 5^2/12 = 2;05 \\
&\rightarrow \text{Area of 5 segments}_5 = 2;05 - 1;40 = 0;2 \\
&\rightarrow \text{Area}_{\text{segment-5}} = 0;25/5 = 0;05 \rightarrow \text{Area}_{\text{oval-5}} = 0;10
\end{aligned}
$$

[29] The first discussion of this is Bruins (1951, 18–20), followed Bruins and Rutten (1961, pp. 23–30, 32–3), who publish an old Babylonian, Susa tablet (text II) with a diagram for the area of a hexagon on one side, whose algorithm is easy to reconstruct from the numbers given, and a heptagon on the other, which is not re-constructible, as well as this tablet (Text III, Tablet I). See also Vaiman (1963, pp. 76, 82–83) and Robson (1999, pp. 48–50). For the provenance of Tablet I (text III), see also Robson (1999, pp. 19–21).

$$\text{Area}_{\text{circle-6}} = c_6^2/(4\pi) = 6^2/12 = 3$$
$$\rightarrow \text{Area of 6 segments}_6 = 3 - 2;37,30 = 0;22,30$$
$$\rightarrow \text{Area}_{\text{segment-6}} = 0;22,30/6 = 0;03,45 \rightarrow \text{Area}_{\text{oval-6}} = 0;07,30$$

$$\text{Area}_{\text{circle-7}} = c_7^2/(4\pi) = 7^2/12 = 4;05$$
$$\rightarrow \text{Area of 7 segments}_7 = 4;05 - 3;41 = 0;24$$
$$\rightarrow \text{Area}_{\text{segment-7}} = 0;24/7 = 0;03,25,43 \rightarrow \text{Area}_{\text{oval-7}} = 0;06,51,26$$

Perhaps these numbers will show up one day on a Babylonian tablet.

## 5 How to derive the ancient method and the revised method: why they work

The oddity of the Ancient method is that it adds on to $1/2\ b \cdot h$ half the square of $h$ (Fig. 11). Now we can imagine a dirty way of thinking about this, which is probably how most people try to explain it.[30] We look at the segment with an inscribed triangle. Its area will be $1/2\ b \cdot h$. Clearly, it is too small. Therefore, we need to approximate with a figure that is larger. Obviously, the rectangle $b \cdot h$ is way too large. Therefore, let us take an average, the trapezoid with height $h$, base $b$, and upper side also equal to $h$, clearly $(b+h)/2 \cdot h$. We could consider some other approximation, 3/2 $b^2$, or even 3 $h^2$, but these would not come out of thinking about the figure in this simple way. Such mean taking is typically Babylonian. This might even be right. If this is how the method was conceived, one might well be wary of it, as the author of P. Cairo shows, or espouse it because one has been seduced by the coincidences associated with W 23291-x.

However, there is nothing in this method that bears any obvious relation to taking $\pi = 3$. The feature that Hero observes would just be a fortuitous accident, and he would be wrong. That too is possible. What would he know of intuitions in another culture, centuries earlier?

I would like to suggest a more precise mean taking. While I cannot be certain that anyone in the Babylonian world ever thought like this, I think we can be fairly certain that the reviser did and that this is what Hero did not quite understand or at least did not quite represent accurately. Let us return to the figure of the rectangle inscribed in the circle. Every segment is on such a rectangle and is adjacent to two equal segments. Let us call the equal segments 'complements' and the unequal segments 'adjacent segments' (with $segment_i$ being the area of segment$_i$). My first observation is something that is so obvious that it is hard to imagine others not noticing it. Indeed, if the scribe of BM 85194, probs. 20-1, did not get it, the author of the problem he

[30] For example, Friberg (2005, p. 133) thinks it has to do with a trapezoid rule. Had its Babylonian provenance not been so tentative, I suspect that more people would have used this sort of diagram to explain the algorithm.

**Fig. 11** A common way to understand the Ancient method

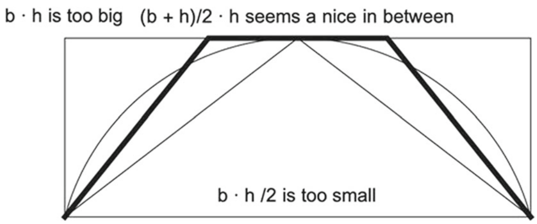
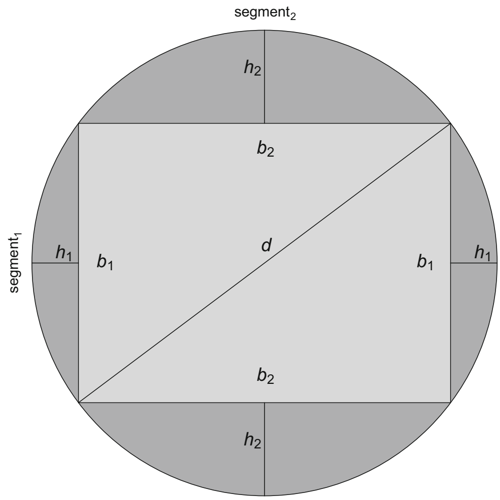



**Fig. 12** Adjacent segments on a rectangle inscribed in a circle




copied certainly did (see footnote 23).[31] Let $h_1$, $b_1$ and $h_2$, $b_2$ be the heights and bases of adjacent segments on a rectangle inscribed in a circle of diameter $d$ (Fig. 12).

$$\text{Observation 1}: \; h_1 = {}^1\!/\!_2\,(d - b_2) \;\text{ and }\; h_2 = {}^1\!/\!_2\,(d - b_1)$$
$$\text{Or } d = 2h_1 + b_2 \text{ and } d = 2h_2 + b_1$$
$$\text{Or } b_1 = d - 2h_2 \text{ and } b_2 = d - 2h_1.$$

The second observation is something that is basic to Old Babylonian treatments of Subtraction methods for obtaining constants for the calculations of the areas of irregular figures such as the grain and ox-eye. For this observation, however, it is necessary to assume that $\pi = 3$.

Observation 2: the area of the circle is equal to the area of the rectangle and twice the areas of a pair of adjacent segments. That is,

$$ {}^3\!/\!_4\, d^2 = b_1 \cdot b_2 + 2\,(segment_1 + segment_2)$$
$$\text{Or:}\quad 2\,(segment_1 + segment_2) = {}^3\!/\!_4\, d^2 - b_1 \cdot b_2.$$

[31] Friberg (2007, pp. 43–45) provides an example, MS 3049, prob. 1, of finding the height from the chord by taking the base/chord of the segment from the height and the diameter, by taking $2 \cdot \sqrt{(d/2)^2 - (d/2 - h)^2}$. This does not require Observation 1 below.

In what follows, I shall not use standard labels in the diagrams. The lengths of lines will be indicated, and rectangles will be labeled with capital letters.

**Theorem 1** *On the assumption that* $\pi = 3$, *that is that the area of a circle is* $^3\!/_4$ *the square of the diameter, if a rectangle is inscribed in the circle, the sum of the areas of any two adjacent segments on the rectangle is half the sum of the base and height times the height of each* (Fig. 13).

*Let there be a circle with a rectangle inscribed in it, with segment*$_1$ *having base of length* $b_1$ *and height of length* $h_1$ *and segment*$_2$ *having base of length* $b_2$ *and height of length* $h_2$. *Then,*

$$segment_1 + segment_2 = {}^1\!/_2\,(b_1 + h_1) \cdot h_1 + {}^1\!/_2\,(b_2 + h_2) \cdot h_2.$$

Derivation: Let there be a circle with a rectangle inscribed in it, with the horizontal length longer than the vertical width, and let the rectangle be divided into quadrants, and let the circle be inscribed in a square, where the lines dividing the rectangle into quadrants are extended so as also to divide the square into quadrants. For convenience, let a common diameter of the circle and of the rectangle be drawn. Mark the rectangle in gray.

Since, assuming that $\pi = 3$, the area of the circle $= {}^3\!/_4\, d^2$, color in the three quadrants of the square in dark gray, with the area shared with the rectangle being medium gray, leaving the remaining quadrant of the rectangle in light gray. This will encompass three of the four quadrants of the circumscribing rectangle. Extend all lines in these three quadrants to the circumscribing rectangle. We can thus say that $^3\!/_4$ of the circumscribing square is equal to the circle, so that $^3\!/_4$ of the square less the inscribed rectangle is also equal to the four *segments* on the rectangle. In other words, the dark/medium gray area less the medium/light gray area is equal to the four *segments*. Having removed the medium gray area, the dark gray area less the light gray area is equal to the four *segments*.

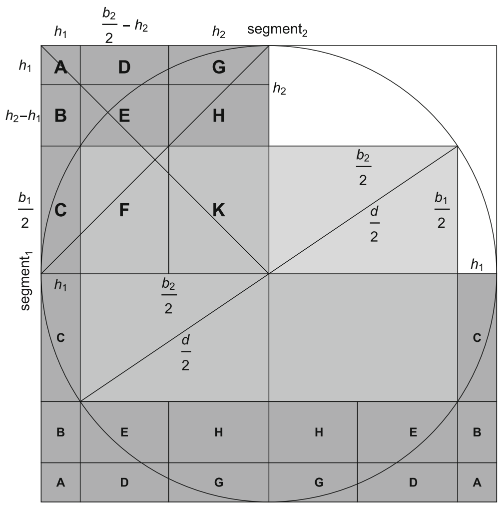



**Fig. 13** Theorem 1

Of course, all of the quadrants of the rectangle are equal. Therefore, I have divided up the upper left quadrant of the square into nine sections. We need to show that the light gray area is equal to no more than triple some selection of portions from the dark gray region of the upper left quadrant. Sections FK are equal to 1/4 the rectangle and so to the light gray area. Hence,

$$3\ (\mathrm{AB\ C\ DE\ GH}) - \mathrm{FK} = 4\ segments \text{ on the rectangle.}$$

Since the rectangle is oblong, the heights of the adjacent segments are unequal. Let us examine how I have divided up the square. First, there are the lines of the bases of the segments, so that the horizontal lengths of DG, EH, FK are $b_2/2$ and those of A, B, $\mathrm{C} = h_1$, while the heights of AB, DE, GH are $h_2$ and those of C, F, K are $b_1/2$. However, I have also added two lines dividing AB, DE, GH, so that the heights of A, D, G are $h_1$, and those of B, E, H are $h_2-h_1$. Similarly, I have divided DG, EH, FG, so that the horizontal lengths of G, H, K are $h_2$ and those of D, E, F, $b_2-h_2$. In all this, what is important is the following:

$$\begin{aligned}&\text{A is a square with side } h_1,\\&\text{GH is a square with side } h_2.\\&\text{Hence, EHFK is a complement square of A}\\&\text{and CF is a complement square of GH.}\end{aligned}$$

Since the diameter will cut the squares, we can also apply Euclid, *El*. I 43, that the complement rectangles will be equal. K = ABED, and DG = BC.

Additionally, it follows that

$$\begin{aligned}&\text{CF is the square of } b_1/2\\&\text{and EHFK is the square of } b_2/2.\\&\text{Hence, by the Pythagorean rule, EHFK + CF = the square on } d/2\\&\quad = 1/4 \text{ the square on the diameter = the square of the quadrant.}\\&\text{Thus, ABCDEFGHK = EHFK + CF}\\&\text{It clearly follows that F = ABDG}\\&\text{We also saw that K = ABED}\\&\text{Hence FK = 2 ABD + E + G.}\end{aligned}$$

Now, I divide up the other two quadrants outside the rectangle in the same way. Hence, the areas of the three square quadrants after the inscribed rectangle is taken out will equal: 3 ABCDEGH.

Hence,

$$\begin{aligned}4 segments &= 2\ (segment_1 + segment_2)\\&= 3\ \mathrm{ABCDEGH} - (2\ \mathrm{ABD} + \mathrm{E} + \mathrm{G})\\&= \mathrm{ABD} + 3\mathrm{C} + 2\mathrm{E} + 2\mathrm{G} + 3\mathrm{H}\\&= \mathrm{A} + \mathrm{GH} + 2\mathrm{C} + \mathrm{B} + \mathrm{C} + \mathrm{D} + 2\mathrm{E} + \mathrm{G} + 2\mathrm{H}\end{aligned}$$

$$= \text{A} + \text{GH} + 2\text{C} + \text{DEGH} + \text{BC} + \text{E} + \text{H} \text{ (but we saw BC} = \text{DG)}$$
$$= \text{A} + \text{GH} + 2\text{C} + \text{DEGH} + \text{DEGH}$$
$$= \text{A} + \text{GH} + 2\text{C} + 2\text{DEGH}$$
$$= h_1^2 + h_2^2 + 2h_1 \cdot b_1/2 + 2h_2 \cdot b_2/2$$
$$= h_1^2 + h_2^2 + h_1 \cdot b_1 + h_2 \cdot b_2$$

Hence, since each of $segment_1$ and $segment_2$ equals the area of its opposite segment,

$$segment_1 + segment_2 = {}^1\!/_2\,(h_1^2 + h_1 \cdot b_1) + {}^1\!/_2\,(h_2^2 + h_2 \cdot b_2)$$
$$= {}^1\!/_2\,(h_1 + b_1) \cdot h_1 + {}^1\!/_2\,(h_2 + b_2) \cdot h_2$$

If the heights are equal, that is, the rectangle is a square, then we do not need the horizontal line, but the proof is otherwise the same, with B, E, H eliminated. Q.E.D.

The proof for $\pi = 22\big/7$ or for any value of $\pi$ (including a true value) now treats this as a lemma, namely:

**Lemma** Given a rectangle inscribed in a circle, with sides $b_1$ and $b_2$ and $h_1$ and $h_2$ the heights of the segments on them, with $d$ the diameter of the circle,

$${}^3\!/_4\,d^2 - b_1 \cdot b_2 = (h_1 + b_1) \cdot h_1 + (h_2 + b_2) \cdot h_2$$

It then follows straightforwardly that:

**Theorem 2** *For whatever* $\pi$ *is taken to be*:

$$4\ segments = \pi \cdot d^2/4 - b_1 \cdot b_2$$
$$= \pi \cdot d^2/4 - {}^3\!/_4\,d^2 + (h_1 + b_1) \cdot h_1 + (h_2 + b_2) \cdot h_2$$
$$= (\pi - 3) \cdot (d/2)^2 + (h_1 + b_1) \cdot h_1 + (h_2 + b_2) \cdot h_2$$
$$= (\pi - 3) \cdot ((b_1/2)^2 + (b_1/2)^2) + (h_1 + b_1) \cdot h_1 + (h_2 + b_2) \cdot h_2$$
$$\text{(since } (d/2)^2 = (b_1/2)^2 + (b_1/2)^2\text{).}$$

Hence, $segment_1 + segment_2$
$= (h_1 + b_1)/2 \cdot h_1 + (b_1/2)^2 \cdot (\pi - 3)/2 + (h_2 + b_2)/2 \cdot h_2 + (b_2/2)^2 \cdot (\pi - 3)/2$, Q.E.D.

Obviously, if the reviser of the Ancient method understood something of this and took $\pi$ to be approximately 3 1/7 (or something equivalent), then

**Corollary to Theorem 2**

$$segment_1 + segment_2$$
$$\approx (h_1 + b_1)/2 \cdot h_1 + (b_1/2)^2 \cdot (22/7 - 3)/2 + (h_2 + b_2)/2 \cdot h_2 + (b_2/2)^2 \cdot (22/7 - 3)/2$$
$$\approx (h_1 + b_1)/2 \cdot h_1 + (b_1/2)^2/14 + (h_2 + b_2)/2 \cdot h_2 + (b_2/2)^2/14$$

The approximation of the segment will then be quite reasonably the average, namely, what we in fact find:

$$segment \approx (h + b)/2 \cdot h + (b/2)^2/14.$$

Indeed, just as it would be surprising if no users of the Ancient method were completely unaware of Theorem 1, it is even harder to imagine that the reviser did not grasp something of this corollary, even without Theorem 2. For, again, how otherwise would he have realized that he needed to take half the square of the base, instead of the height?

Most important, Theorem 1 captures why the rule is essentially tied to taking $\pi = 3$ and why it is exact for only two cases: where the segment is on a semicircle, and there is no inscribed rectangle; and where the figure is a square, and all four segments are equal. It also shows that in every pair of adjacent segments, $segment_1$ and $segment_2$, that are not are on a square, if we could check the area for each segment, again on the assumption that $\pi = 3$:

$$\text{if } (b_1 + h_1) \cdot h_1 > segment_1, \quad \text{then } (b_2 + h_2) \cdot h_2 < segment_2$$

Once one realizes this, one sees immediately that Fig. 3 is deeply flawed. Of course, that graph does not represent the fluctuation of the Ancient method around the true, as it is always less than the true value. However, what it should show is the fluctuation about the true value on the assumption that $\pi = 3$. One may well suspect that what is going on there in the fluctuation of the Ancient value is a fluctuation, not with respect to a true value, but about a norm for $\pi = 3$. This is far enough from a true value for $\pi$ that the Ancient method does not ever result in values above the True area of the segment. Therefore, in the case of the Revised method too, 22/7 is close enough to the actual value of $\pi$ that pairs of adjacent segments can fluctuate on both sides of the actual, respective values.

Each method works for both segments when the deviation is not large, which normally it is not. We can see why, when the ratio of the base to the height is large, the error of one of the segments should also be large. For the adjacent segments will diverge in size, so that they will also diverge from the mean. All this is clear in Fig. 14, which compares the values for the Ancient and Revised methods with 'true' values for $\pi = 3$ and $\pi = 3\ ^1/_7$, respectively, and Fig. 15, which shows the values for the adjacent segments based on the Revised method and their mean area. We see clearly the intersection with the true values at the semicircles (radius = 60 points) and at the square ($h \approx 17.57$ points, $b \approx 84.85$, and $b : h \approx 4.83$).[32] It happens that the inaccuracy comes about when the segment is the smaller, since there are two points of accuracy, the limiting case of the segment on the semicircle and the segment on the square. However, without a deeper means of checking, it is hard to see how one could know. The author of P. Cairo has a means, but not the means to exploit it extensively.

[32] For the 'true' value of the area of the segment, I used, in Mathematica™: *pival* $r$^2 ArcCos[$(r - h)/r$]/Pi − Sqrt[2 $h\ r - h^2$]$(r - h)$, where *pival* is whatever value of π was appropriate and Pi the built-in value of π. However, when the value π is distant from the true value, as in taking $\pi = 3$, and the angle of the sector is small, the value of the arc/circumference becomes distorted, so that it does not seem meaningful to construct a curve. In Fig. 14, I do not include small values of *h,* where $h : r < 1{:}12$).

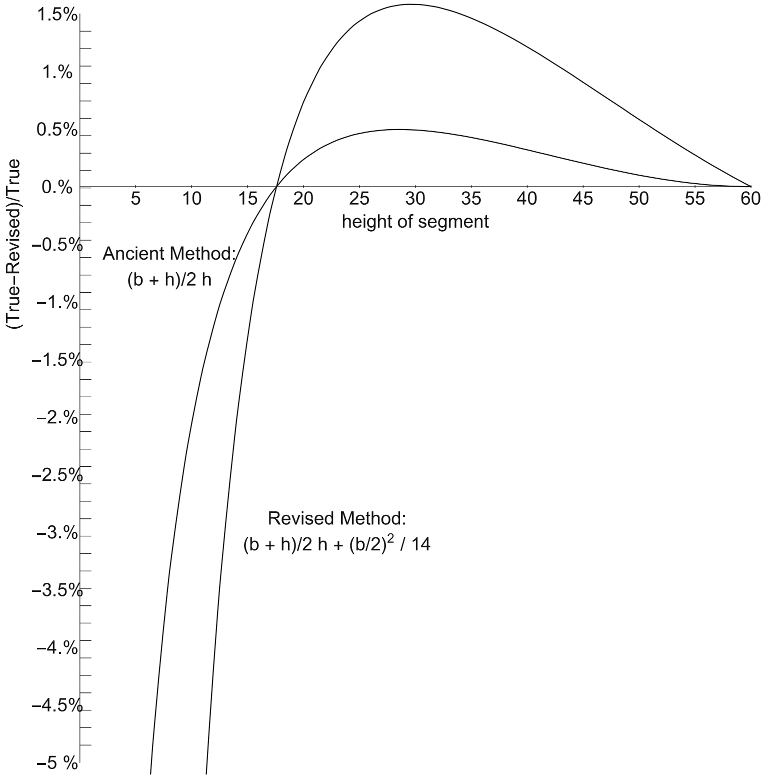



**Fig. 14** Comparison of the Ancient and Revised methods to True area based on $r = 60$ and $\pi = 3$ or $\pi = 3\ 1/7$

The Revised method obviously has a terminus post quem, Archimedes' publication of the *Dimensio Circuli*, but I have only a guess at who and when. Did the discoverer know of a Babylonian antecedent, or was he relying on Egyptian treatises, or had the Ancient method already become a part of the Greek metrical tradition? And if so, was the vehicle of transmission treatises like P. Cairo? These are nice matters for speculation, but not presently answerable.[33]

It is evident from Theorem 2 that we may now state the Revision of the Ancient method as a generalized approximation of a segment:

$$segment \approx (h + s) \cdot h + (b/2)^2 \cdot (\pi - 3)/2$$

The same issues will apply. When the segment is that on an inscribed square the rule will be accurate for that value of $\pi$, but in any case,

$$\text{if } (b_1 + h_1) \cdot h_1 + (b_1/2)^2 \cdot (\pi - 3)/2 > segment_1, \text{ then } (b_2 + h_2) \cdot h_2 + (b_2/2)^2 \cdot (\pi - 3)/2 < segment_2$$

[33] See Knorr (1982) for an argument that transmission into Greek mathematics from Babylon was via Egypt.

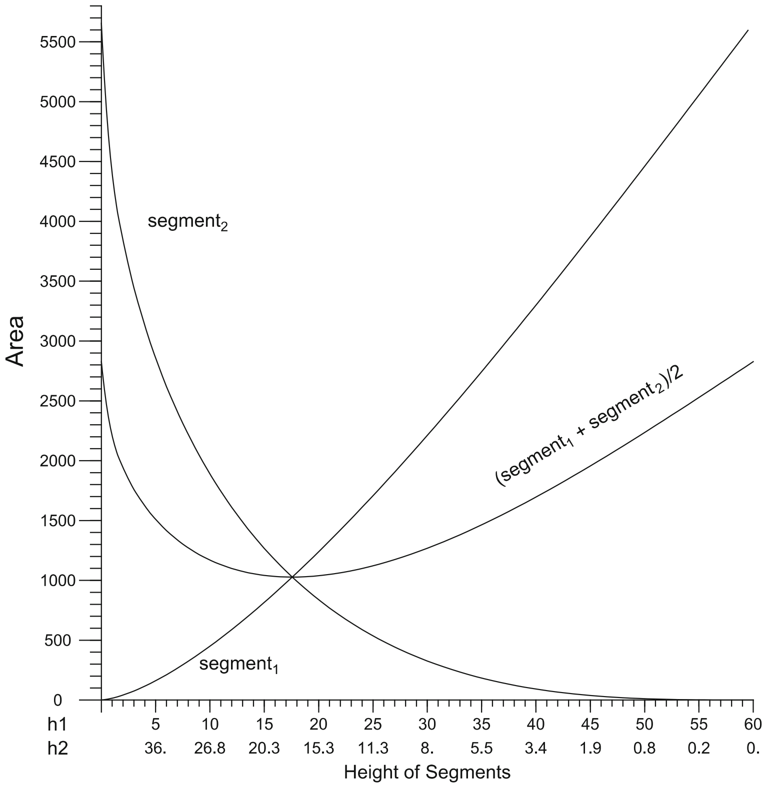



**Fig. 15** Comparison of the areas of adjacent segments for $r = 60$ points and $\pi = 3\ ^1/_7$

I have no idea whether anyone would be interested in this modern approximation of the area of a segment, but, if $\pi \approx 3\ ^1/_7$ is good enough, then the Revised method is probably good enough.

It should also be clear what is strange and unfortunate about Hero's description of both methods. Of course, he is correct in explaining that in the limiting case of the semicircle, as he could have added for the case of the square, each method is essentially tied to its presumed value for the ratio of circumference to diameter. But this is because there is a more fundamental theorem that Hero does not prove, that establishes both methods, but only accurately establishes the areas of adjacent segments together and not individually. His comments, therefore, that the application of the methods to the semicircle justifies and explains each are misleading and peculiar, but comprehensible once we understand the underlying theorem. On the other hand, it would seem that Hero does not know the underlying theorem, and we shall see that there is some further evidence for this.

## 6 Hero's method, the quasi-Archimedean area as $4/3 \cdot b \cdot h/2$: a puzzle remains

Unless one has an independent way of checking the areas, e.g., by regular polygons, there really is no way of knowing that the Revised Version is ever inadequate, although one can guess that it must be, simply because $3^1/_7$ is an upper bound. This is partly why I think it unlikely that the Revised method was part of the *Dimensio Circuli* and was probably introduced somewhat later than Archimedes. Does it become irrelevant with the introduction of trigonometry, when one really can use a Subtraction method to get the area of a segment? Obviously not! Indeed, it is perfectly consistent with the utility of the method that it be introduced by someone like Hipparchus. But the only requirement is that it is well established by the time that Columella is giving advice to farmers in the mid first century CE.

We saw that Hero's assertion that the Revised method does not work well when the base is more than triple the height is problematic. It would have been natural to presume that Hero uses trigonometry and subtraction to check the Revised method, except that it would require that he made an error of calculation. And this is a problem. Why does he think the Revised method is worse than it actually is? In addition, why, as Knorr and Høyrup observe (see footnote 11), does he then use the Revised method to illustrate the Subtraction method, in an example where the base is 14 units and the height is calculated to be 3 $^1/_2$ units, that is where the base is four times the height? In fact, the situation appears even worse than they suggest. The ratio of the base of the segment on a square to the height is $1{:}^1/_2\,(\sqrt{2}-1) \approx 1{:}^1/_2\,(^{17}/_{12}\ -1) \approx 4\ ^4/_5\ :1$, which is much larger than either the example Hero uses or 3:1. Yet, this is where the Ancient and Revised methods actually are most accurate, in fact geometrically based, and, as we have seen, demonstrably so. For this, one would not even have to rely on a trigonometric chord table. One can see this simply by the Subtraction and Division method suggested by P. Cairo. Hero's remarks are indeed mysterious.

Nor will it be an adequate response that Hero did not like the fact that the Revised method gives values higher than the true for ratios even where the height is larger than the height of the segment on the square. First, the ratio is much larger than 3:1, and secondly $3^1/_7$, which he knows gives a high value, is the value that he is using for circle calculation.

In my discussion, I have avoided the thorny issues of the unity of the 7 chapters I 27–33, on the segment in the *Metrica*. I am not good at judgments of taste and I do not see why, for example, what others see as a messy organization might not be exactly what the author wanted, why the author of the discussion might not set up the preliminaries for a theorem he is keen to present before the full discussion of the segment and then, with these out of the way, proceed to present two traditional ways of finding the area, the Ancient and the Revision, proudly give the theorem (whatever may be his contribution to it), which is far from technically trivial, and then conclude with the Subtraction method. He might well have thought the lemmata a distraction from the tight discussion of I 30–33.

The only other oddity is that he gives no metrical example of the quasi-Archimedean method, but this is hardly uncommon in the *Metrica*.[34] The only real oddity, then, is the issue that Knorr and Høyrup raise, and which we see is much, much worse. The Revised method is better than the quasi-Archimedean method for ratios where the base is less than about 6.4 times the height, he ignores his own rule, and the best application of the method is ruled out by his restriction. Is there a better case for suggesting that there is an error in the text? Either Hero really blew it, or he meant to or did say something else, such as that the half base must not be greater than three times the height (see footnote 10). That, at least, would be in the ball park of correct and would not go against any examples. Given that he repeats the error four times in the course of I 31–32, emendation is out of the question. But faulty memory might not be. Or perhaps there is some other reason entirely for what he says.

It is natural to want to save an author from a blunder. The quasi-Archimedean method appears nowhere else in Hero's corpus.[35] Why not think of it as an intrusion? This is not so easy, given that it really does work better than the Revised method for a high ratio of the base to height and given the difficulty of the proof, as well as the pride with which the author announces it. It cannot be a general substitute for the Revised method, but something is needed for the high ratios. Therefore, the only problem is the ratio, three times, rather than slightly more than six times.

More to the point, however, assuming Hero understood that the correct ratio is six times, as we can see from Fig. 14, the drop in error is very fast after the ratio for the square, but not severe. We can make up a table of simple calculations of the sort Hero might have presented, or used, comparing the Revised method and the quasi-Archimedean with ratios from twice to nine times, along with the extreme example from *Metrica* I 31.

| (* in Hero) | Height | Base | Revised method | | Quasi-Archimedean method | | Method with chords |
|---|---|---|---|---|---|---|---|
| Ratio | | | Calculation | Result | Calculation | Result | Result[a] |
| | | | $(h+b)/2 \cdot h + (b/2)^2/14$ | | $4/3\ b \cdot h/2$ | | |
| 2x* | 7 | 14 | $(7+14)/2 \cdot 7 + 7^2/14$ | 77 | $2/3 \cdot 7 \cdot 14$ | 65 1/3 | 77;00 |
| 3x | 7 | 21 | $(7+21)/2 \cdot 7 + (10^1/_2)^2/14$ | 105 7/8 | $2/3 \cdot 7 \cdot 21$ | 98 | 106;18 |
| 4x* | 3 1/2 | 14 | $(3\ ^1/_2 + 14)/2 \cdot 3\ ^1/_2 + 7^2/14$ | 34 1/8 | $2/3 \cdot 3^1/_2 \cdot 14$ | 32 2/3 | 34;17 |
| 4x*[b] | 4 | 16 | $(4+16)/2 \cdot 4 + 8^2/14$ | 44 4/7 | $2/3 \cdot 4 \cdot 16$ | 42 2/3 | 44;47 |

[34] See Acerbi and Vitrac (2014, pp. 58–59, 411–427). Cf. *Metrica* II 10, III 11–18, 23, II 10, III 11–18, 23, which do not give metrical examples.

[35] See Høyrup (1997, p. 242). So far as I can tell, none of the four methods for finding the area of the segment are used in the *Metrica* outside I 30–33, although III 18, to divide a circle into thirds with two straight lines, could easily have used something of these methods. The scholion to *Geom*. 19 in Cod. Seraglio G.I. I f. 11[r] (Hero, *Opera* v. 5, p. 228–9, see footnote 9) illustrates the quasi-Archimedean method with $b = 40$ and $h = 10$: triangle inscribed in the segment $= 200 \rightarrow 1/3 \cdot 200 = 66\ 2/3 \rightarrow$ segment $\approx 200 + 66\ 2/3 = 266\ 2/3$. Of course, the scholiast is reading and commenting on the entire manuscript, which includes the *Metrica*.

| (* in Hero) | Height | Base | Revised method | | Quasi-Archimedean method | | Method with chords |
|---|---|---|---|---|---|---|---|
| Ratio | | | Calculation | Result | Calculation | Result | Result[a] |
| 5x | 7 | 35 | $(7+35)/2 \cdot 7 + (17\,{}^{1}/_{2})^2/14$ | $168\,{}^{7}/_{8}$ | ${}^{2}/_{3} \cdot 7 \cdot 35$ | $163\,{}^{1}/_{3}$ | 168;38 |
| 6x*[c] | $3\,{}^{1}/_{3}$ | 20 | $(3\,{}^{1}/_{3}+20)/2 \cdot 3\,{}^{1}/_{3} + 10^2/14$ | $46\,{}^{2}/_{63}$ | ${}^{2}/_{3} \cdot 7 \cdot 42$ | $44\,{}^{4}/_{9}$ | 45;29 |
| 7x | 4 | 28 | $(4+28)/2 \cdot 4 + 14^2/14$ | 78 | ${}^{2}/_{3} \cdot 4 \cdot 28$ | $74\,{}^{2}/_{3}$ | 76;03 |
| 8x | $3\,{}^{1}/_{2}$ | 28 | $(3\,{}^{1}/_{2}+28)/2 \cdot 3\,{}^{1}/_{2} + 14^2/14$ | $69\,{}^{1}/_{8}$ | ${}^{2}/_{3} \cdot 3\,{}^{1}/_{2} \cdot 28$ | $65\,{}^{1}/_{3}$ | 66;22 |
| 9x | 7 | 63 | $(7+63)/2 \cdot 7 + (31\,{}^{1}/_{2})^2/14$ | $315\,{}^{7}/_{8}$ | ${}^{2}/_{3} \cdot 7 \cdot 63$ | 294 | 298;00 |
| 10x | 7 | 70 | $(7+70)/2 \cdot 7 + 35^2/14$ | 357 | ${}^{2}/_{3} \cdot 3\,{}^{1}/_{2} \cdot 24$ | $326\,{}^{2}/_{3}$ | 331;13 |
| 60x* | 1 | 60 | $(1+60)/2 \cdot 1 + 30^2/14$ | $94\,{}^{11}/_{14}$ | ${}^{2}/_{3} \cdot 1 \cdot 60$ | 40 | 40;03 |

[a]Using a sexagesimal, unlimited precision calculator that I have written in Visual Basic™ for Excel™, that, inter alia, interpolates from Ptolemy's chord tables, after finding the radius, I took the arc-chord ($b \cdot 60/r$), divided the angle by 360, and multiplied this by 3;8,34,37 (an approximation of $3\,{}^{1}/_{7}$) and the square of the radius to get the area of the section. I then subtracted $(r-h) \cdot b/2$ to get the area of the segment, all calculated with the precision, except for $\pi$, kept to 1 or 2 places. This paltry attempt at authenticity is probably unnecessary, since we do not know what Hero's chord table would have looked like anyway, while the difference between the values obtained and values obtained in more conventional ways is trivial

[b]*Geometrica* 24.51 (ms. S), without a given procedure as part of the calculation of the area of a segment larger than a semicircle, $b = 16$, $h = 16$ by subtraction. This is one of two examples (three passages) outside the *Metrica* that use ratios of $b : h > 3{:}1$, both within a subtraction argument. Columella, *De re rustica* V 2.9 uses these values for his sole example but rounds down, taking 64/14 as "ea efficit IV paulo amplius," whence, he actually gets 44 as the area

[c]*Geometrica* 20.8-11 (ms. S and mss. AC differ in 8–9, the area of the smaller segment is in 11), This is a calculation of a larger segment by subtraction, with $b = 20$ and $h = 30$. See also *De mens*. 32, which merely instructs us, in procedural form, to follow the procedure and does not provide values. These are two of three instances outside the *Metrica* of values $b : h > 3{:}1$. See Acerbi and Vitrac (2014, p. 225 n. 289) for a survey of occurrences of the Revised method, values used and variations

I have no idea whether Hero would have gone through the labor of testing the two loose methods against a geometrical test in a random way. The books on *Straight Lines in a Circle* referred to in I 22, 24 should have been a substantial treatise that discussed the lengths of chords. It is hard to imagine that an Alexandrian mathematician would not be in a position to make such a test. But that does not mean he actually did.[36] It would certainly be an onerous task.

Given that the goal of handbooks such as the *Geometrica* and the *Metrica* is to provide useful and handy methods for calculation, it is easy to see why certain complex methods are not going to be recommended, even if they were available. Nonetheless, it is natural that Hero will recommend methods that are easy. His quasi-Archimedean method is easier to use than the Revised method and is reasonably accurate when the ratio of $b{:}h$ is high. He also knows that the Revised method gets very bad eventually. All this is clear. "What is good enough?" is a different question. But how he knows

[36] See Acerbi and Vitrac (2014, pp. 29–30, 205 n. 227), who think it likely that the treatise was a development of a chord table.

when it goes bad or if its going bad is even the reason for the recommendation to switch at the 3:1 ratio or at any ratio is the unanswered question.

With this, let me make one last attempt a good story. But it is no more than a good story.

There is an obvious way in which Hero might have checked the Revised method, the one we have already inferred from P. Cairo, by Subtraction and Division. In *Metrica* I 17–25, he establishes procedures for deriving the areas of regular polygons from their sides, from triangle up to the dodecagon, except for the trivial square. All he needed to do was to do was something that would have taken a few extra moments, to have derived the height of a segment and the area of the segment by the Revised method, and then to have found the area of the circumscribing circle and to subtract the area that he has just found, and to divide by the number of sides of the polygon.Then, if things did not turn out right, he might also have checked the ratio of the side of the segment to the height. Here is a problem. The ratio of the side of every regular polygon inscribed in a circle to the height of the segment on it is larger than 3:1. Almost every one of these calculations will involve discrepancies due to the values for square roots chosen and for the value of $\pi$, that are built into the nature of the Revised method itself, except for the segment on the square.

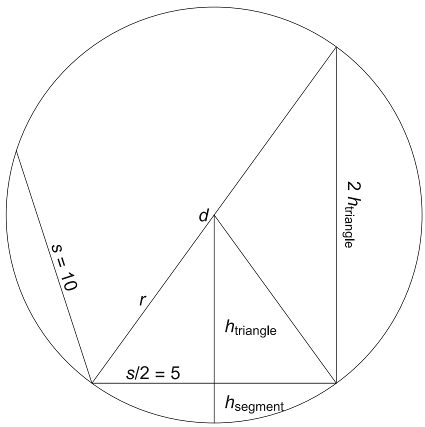
Except for the case of the triangle, Hero provides all the numbers and ratios for every odd-numbered polygon adequate to do both calculations. That is, the calculations provide enough to get ratios between the side, the diameter, and the altitude of the triangle on the side whose vertex is the center of the triangle. That is not the case for the even numbered polygons, where the length of the diameter and radius are not used in the calculation. Of course, it would be easy to apply the Heronian procedure to derive square roots for these, even when they involve two applications of the procedure (e.g., $\sqrt{(5/2+\sqrt{5}/2)} \approx \sqrt{(5/2+9/8)} \approx 61/32$). For our purposes, however, the odd-numbered polygons will be enough, except that we will also need some construction of the triangle and square. For the equilateral triangle (I 17), Hero actually avoids finding the height by determining the area as the square root of $3/16$ of the square of the square (δυναμοδύναμις) of the side, but 7/4 is well enough established as a value for $\sqrt{3}$ and $17/12$ for $\sqrt{2}$. In each case, the problem is to find the area of the figure with a side of 10. Therefore, I will give the ratios, what they represent, and then the results of calculation (Fig. 16).

Equilateral triangle: we saw the same example in P. Cairo and may ignore the oddities of the Babylonian ox-heart, so that we expect a discrepancy. Let the side, $s$, be 10. Then, $h_{\text{triangle}} = \sqrt{(s^2-(s/2)^2)} \approx \sqrt{(10^2-5^2)} = 5\sqrt{3} \approx 5 \cdot 7/4 \approx 8\ 3/4$. As before, $h_{\text{segment}} = 1/3\ h_{\text{triangle}} \approx 2\ 11/12$. Finally, $d = 20\ \sqrt{3}/3 = 11\ 2/3$. The area of the triangle from *Metrica* I 17 is: 43 1/3.

Square: we expect the values to be close. Let the side, $s$, be 10. Hence, $d = s\sqrt{2} = 10\sqrt{2} \approx 10 \cdot 17/12 = 14\ 1/6$, and $h_{\text{segment}} = 1/2\ (d-s) \approx 2\ 1/12$. The area of the square obviously is: 100.

Pentagon (*Metrica* I 18): Hero treats the square root of 5 as 9:4, which conveniently results from dividendo on the ratios of a 3:4:5 right triangle, here the right triangle

**Fig. 16** General diagram for segments on regular *n*-gons, where *n* is odd




formed by the altitude from the center,[37] so that hypotenuse + altitude:altitude ≈ $\sqrt{5}$ : 1 ≈ 9:4 → hypotenuse:altitude ≈ 5:4 (see also the procedure for the dodecagon, *Metrica* I 23). Hence, a 3–4–5 triangle is involved in a metrical approximation of a pentagon. The ratios in the text then are radius:height:side ≈ 5:4:6. Therefore, if the side $s$ is 10, $h_{\text{triangle}} \approx 10/6 \cdot 4 = 6\ ^2/_3$ and $d = 10/6 \cdot 5 \cdot 2 = 16\ ^2/_3$, i.e., $r \approx 8\ ^1/_3$. Hence, $h_{\text{segment}} = r - h_{\text{triangle}} \approx 1\ ^2/_3$. The area of this pentagon is calculated in *Metrica* I 18 as 166 2/3, although he points out that one can get a more precise result with more precise values for numbers, $n^2$ and $m^2$, such that $n^2$ is closer to $5m^2$ than 81 to 5 · 16, his way of describing the approximation.

Hexagon (*Metrica* I 19): It is probably unnecessary to include a hexagon, especially, as Hero builds the calculation of the area on the area of the equilateral triangle. Nonetheless, as it involves no more than $\sqrt{3} \approx 7/4$, it is not egregious to include it. For, if we let $s = 10$, $d = 20$, and $h_{\text{triangle}} \approx 8\ ^3/_4$. Hence, $h_{\text{segment}} \approx 1\ ^1/_4$. The area calculated will be: 259.

Heptagon (*Metrica* I 20): the ratio of the radius: side ≈ 8 : 7 and of the side: $\text{height}_{\text{triangle}} \approx 42 : 43$, so that diameter:radius:side:$\text{height}_{\text{triangle}}$ will be 96:48:42:43. Let the side be 10, so the adjustment will be $^5/_{21}$. Hence, $d \approx 22\ ^6/_7$, while the radius ≈ $11\ ^3/_7$ and $\text{height}_{\text{triangle}} \approx 10\ ^5/_{21}$, so that $h_{\text{segment}} \approx 1\ ^4/_{21}$. Hero calculates the area as $358\ ^1/_3$.

Nonagon and hendecagon (*Metrica* I 22 and 24): both passages refer to books *On straight-lines in a circle*, which may have constructed a table of chords, but certainly explored metrical values. The diagram takes a diameter and a line from the other end point of the side/chord to the diameter. This triangle will be similar to and double a triangle from the center where one side is the radius and the other the height.

Nonagon: $d : s = 3{:}1$ and $h_{\text{triangle}} : s \approx {}^1/_2 \cdot 17{:}6 \approx 17{:}12$. Let $s = 10$. Then, $d \approx 30$; $r \approx 15$; $h_{\text{triangle}} \approx 14\ ^1/_6$. Therefore, $h_{\text{segment}} \approx {}^5/_6$. The area is: $637\ ^1/_2$.

[37] Hero shows in a lemma to I 18 that (with ρ a right angle) if a right triangle has an angle $^2/_5 \cdot \rho$, with $a$ the adjacent leg to the angle (and opposite the $^3/_5 \cdot \rho$ angle) and $h$ the hypotenuse, the square on $h + a$ is 5-times the square on $a$. Since the angle of the right triangle from the center, i.e., half the triangle from the center, is $^1/_2 \cdot (^4/_5 \cdot \rho)$, the square on the radius/hypotenuse of the circle/triangle + the altitude is 5 times the square on the altitude.

Hendecagon: $d : s = 25{:}7$ and $2h_{\text{triangle}} : s \approx 24{:}7$, so that $h_{\text{triangle}} : s \approx 12{:}7$. Let $s = 10$. Then, $d \approx 35\ ^5/_7$; $r = 17\ ^6/_7$; $h_{\text{triangle}} \approx 17\ ^1/_7$. Therefore, $h_{\text{segment}} \approx {}^5/_7$. The area is: 942 $^6/_7$.

With this, we may now construct a table with the different values for the segments (the second numbers for the circle area and difference/$n$ take $d^2$ without using the approximated square root):

| $n$-gon | Ratio $b{:}h$ | Circle area | $n$-gon area | Difference/$n$ | Revised method | Quasi-Arch | Modern |
|---|---|---|---|---|---|---|---|
| 3-gon | 3 $^3/_7$ | 106 $^{17}/_{18}$, 104 $^{16}/_{21}$ | 43 $^1/_3$ | 21 $^{11}/_{54}$, 20 $^{10}/_{21}$ | 20 $^{1255}/_{2016}$ | 19 $^4/_9$ | 20.47 |
| 4-gon | 4 $^4/_5$ | 157 $^{347}/_{504}$, 157 $^1/_7$ | 100 | 14 $^{851}/_{2016}$, 14 $^2/_7$ | 14 $^{751}/_{2016}$ | 13 $^8/_9$ | 14.27 |
| 5-gon | 6 | 218 $^{16}/_{63}$ | 166 $^2/_3$ | 10 $^{20}/_{63}$ | 11 $^{32}/_{63}$ | 11 $^1/_9$ | 11.06 |
| 6-gon | 8 | 314 $^2/_7$ | 259 | 9 $^3/_{14}$ | 8 $^{183}/_{224}$ | 8 $^1/_3$ | 9.06 |
| 7-gon | 8 $^2/_5$ | 410 $^{170}/_{343}$ | 358 $^1/_3$ | 7 $^{3254}/_{7203}$ | 8 $^{197}/_{441}$ | 7 $^{59}/_{63}$ | 7.69 |
| 9-gon | 12 | 707 $^1/_7$ | 637 $^1/_2$ | 7 $^{31}/_{42}$ | 6 $^{151}/_{504}$ | 5 $^5/_9$ | 5.91 |
| 11-gon | 14 | 1002 $^{64}/_{343}$ | 942 $^6/_7$ | 5 $^{135}/_{343}$ | 5 $^{30}/_{49}$ | 4 $^{16}/_{21}$ | 4.81 |

If you have a penchant for a good story, I am not sure that the one I am about to give will satisfy you. I certainly do not put much store in its likelihood, but it has, at least for me, a feel of verisimilitude, and that may be what is wrong with it. It is curious that every regular polygon has a base/side to height ratio greater than three times, while the pentagon and up have ratios greater than or equal to six times. These are also all the segments that can be tested without resorting to Hipparchus' theory of chords, other than, of course the limit case of the semicircle. There is no getting around the fact that the square works fairly well, as we expect. But we also expect some error in the case of the equilateral triangle, and it is there on either calculation. Do not look at the column on the right—I put it there for completeness. In fact, do not look yet at the column next to it with Hero's own method. I do not know why the values for the hendecagon's are so in sync, but all the polygons are all a little problematic when compared with each other. In our table, the quasi-Archimedean, against Hero's theorem, even comes out larger than that by Subtraction and Division, which should lead one to think the approximations very inadequate, as Hero noted in his account of the pentagon. Let us suppose that he got results something like this. Having assumed that the method of Subtraction and Division is completely reliable if the approximations are good, Hero might have thought that the only really reliable values in the application of the Revised method are those near the case that he could see was proven, namely the semicircle. If this is right, by not having our theorem on adjacent segments, Hero has missed understanding why the Revised method works—and why it does not. He then condemns all non-trivial applications which he can check, namely where the ratio is that of the height on a regular polygon, or where the ratio of base to height is larger than 3:1. Finally, he found his own alternative that seemed to work even better (well, not quite if you now look at the second column from the right). But it was his method. Therefore, this is a good story. Is it true? Maybe! Maybe not?

## 7 Very brief remarks on the Subtraction method: another puzzle remains

I do not know what would count for Hero as inadequate values for the area, nor how he would check. The obvious way for someone in his time would be to do the tedious work of using a geometrically based method, such as Subtraction and Division. It is evident that the Ancient method and the Revision also work well when the segment is larger than a semicircle, just not when it is much larger, speaking very roughly, when the height is $^3/_2$ the radius or about $^7/_8$ the base, that is the complement of the segment on the side of the equilateral triangle, the error is only about 2% for the Revised method. A graph (Fig. 17) might, then, be useful comparing the difference in values between the Revised method and the Subtraction method. As before, it might be interesting to have them compared to an actual value as well.

The chart reveals two things that are unexpected, and perhaps amusing. The Revised method is fairly effective so long as the height is less than about $^3/_2$ the radius. Perhaps

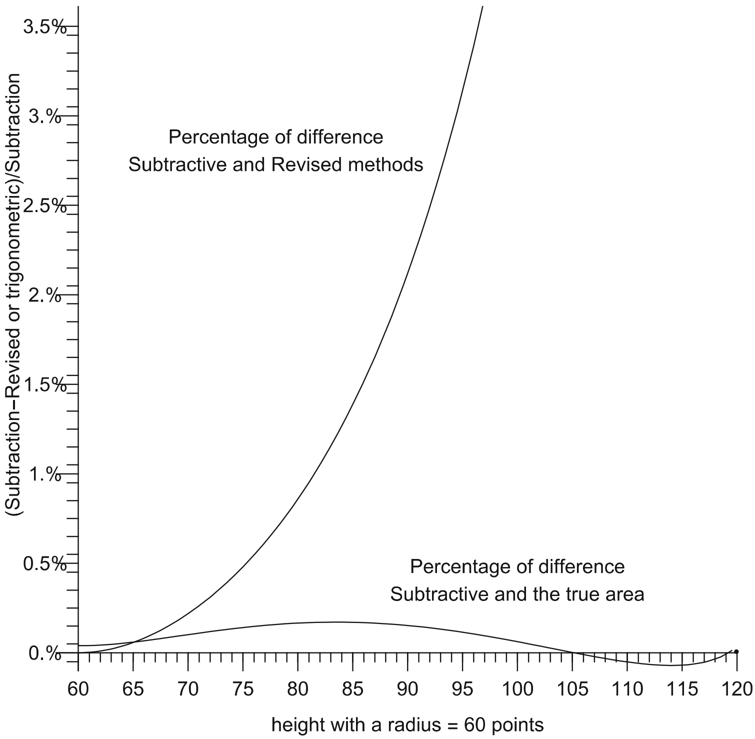



**Fig. 17** Comparison of the Subtraction method with the Revised method and the True area

this is too inaccurate for some people. But it is certainly comparable, for example, to taking $\pi = 3$, or some other calculations to be found in the Heronian tradition, or even some other methods for finding the area of a segment larger than semicircle.[38] What is truly surprising, however, is how good the Subtraction method is, regardless of the inaccuracies of the calculation of the complement segment. Quite the contrary, when the height is more than 5/6 the diameter, and you are making the worst estimate of the area of the complementary segment, you need not worry at all. Your error is still going to be less than 0.1%. Therefore, Hero was right to ignore his restriction in applying the Revised method in using it for the Subtraction method. If the base is 60 units and the height 900 units, go right ahead and derive the complement's height as 1 unit. Take the area as 125 $^2/_7$ and calculate the area of the circle ($901^2 \cdot {}^{11}/_{14}$) as 637844 $^9/_{14}$ and tell us that the area is 637719 $^5/_{14}$, because your handy sexagesimal calculator and chord table will come up with 637811;16. Then, you will see why. The complement is so small that any bad but reasonable calculation will make for an accurate estimate of the segment.

But is it at all likely that anyone in antiquity investigated these issues of accuracy with any depth and beyond looking at extreme cases, such as Hero does in the *Metrica*? It is more likely that these issues of accuracy are guided more by intuition and the knowledge that the Subtraction method should work exactly as well as the Revised method (and apparently even better) while these other methods fall apart quite visibly at extremes. We thus come back to the same issue. When someone speaks of error, sometimes they are being methodical, but not always. Often they are just waving hands.

## 8 Conclusions: some questions answered

I began with twelve questions, and some more came up along the way. To some extent, my puzzles are a small part of the history of mathematics. They constitute a corner in the world of mensuration, not the loftiest corner in the pantheon of past historical interest. Yet, there is an important path that I have traced from Babylon to some mathematician in the late Hellenistic age who found an interesting property, which was then lost, as people applied the results, a much better way of measuring something.

I have shown why the Ancient method works and in a deeper way that it really is essentially connected to taking $\pi = 3$. This also led to a better way of understanding the procedure, I think, than has been done before. Problems remain. Did anyone in the Babylonian world see the property that serves as a solid foundation for the procedure? At least, it is plausible that the author of P. Cairo knew that the segment of the square was unproblematic, while the segment on the equilateral triangle was,

[38] Besides occurrences of the Ancient method in the *Stereometrica* (see Acerbi and Vitrac 2014, p. 223, n. 286), I noted in Sect. 1 *Geometrica* 20.4(S): $(b, h) \rightarrow b \cdot h \rightarrow b \cdot h \cdot 11 \rightarrow b \cdot h \cdot {}^{11}/_{14} \rightarrow$ A, with an error from 0 to 9.1% for the small range $2{:}1 > b{:}h \geq 1{:}1$, but any lower is noticeably bad; *De mensuris* 29 : $(b, h) \rightarrow (b{+}h) \rightarrow (b{+}h) \cdot h \rightarrow (b{+}h) \cdot h/2 \rightarrow 1/21(b{+}h) \cdot h/2 \rightarrow (b{+}h) \cdot h/2{+}1/21(b{+}h) \cdot h/2 \rightarrow$ A, an error of about 0–5.0% for $2{:}1 > b{:}h > 2{:}3$. The error for the Revised method with $b : h = 2{:}3$ is 8.5%. The angle in the circle for this segment is about 286°. See Fig. 3. Both are much worse than the method of Subtraction. My point here is not that anyone is paying attention, but rather that preferences one way or another are probably based on intuitions about extreme cases and suspicions. But could anyone doing any serious checking miss how bad the method of *Geometrica* 10.4(S) is?

a crucial consequence of the property. In addition, I have established the procedure as Babylonian and showed that a weird coincidence of numbers could have led a Babylonian mathematician to miss completely what was problematic in P. Cairo and so to think the method confirmed, although I leave it to others to determine more precisely the method's provenance in that world. With much more confidence, I think I have established the basis for the Revised method. Indeed, the Revision allowed us to find a nice generalization: for any preferred value of $\pi$, the sum of the areas of two adjacent segments is $(b_1+h_1)/2 \cdot h_1+(b_1/2)^2 \cdot (\pi-3)/2+(b_2+h_2)/2 \cdot h_2+(b_2/2)^2 \cdot (\pi-3)/2$, so that the area of a segment is approximately $(b+h)/2 \cdot h+(b/2)^2 \cdot (\pi-3)/2$. Since the values for the areas of two segments will sum up to a common value that is correct, each will fluctuate around a 'true' value but will be completely accurate for two cases, the trivial semicircle and the segment on a square. This is why the Revised version uses a value of $\pi$ that is above the true value but fluctuates on both sides of the True area.

The question of Hero's quasi-Archimedean method and his claim that it needs to be used when $b{:}h > 3:1$ remains the mystery it was when I began. Perhaps there is yet more mystery. He uses illicit values for the bases and heights on the way to calculating the complementary segments, but lo and behold, it turns out to be completely justified, even in the crazy case where, it turns out, the errors in the calculation are superfluous. The real mystery is why he rules out one of the two cases where the Revised rule is perfectly accurate, the segment on the square. Here, we can only speculate that there is a major textual error, that Hero made a big blunder, had faulty memory, thought his quasi-Archimedean method was adequate and easier, or the always to be considered, some other good story. I suggested one, but is it merely a good story? Life is messy; texts are messy; history is messier.[39]

## Declarations

**Conflict of interest** The author states that there is no conflict of interest.



## References

Acerbi, Fabio. 2021. *The Logical Syntax of Greek Mathematics*. Switzerland: Springer.

Acerbi, Fabio, and Bernard Vitrac. 2014. *Héron d'Alexandrie: Metrica: introduction, texte critique, traduction française et notes de commentaire*. Pisa: Fabrizio Serra Editore.

39 I should like to thank Jens Høyrup for his guidance, especially on some Babylonian issues, and Fabio Acerbi for some early advice.

Bertoni, Daniel. 2017. Geometry and Genre in Columella. *American Journal of Philology* 138: 527–554.
Bruins, E.M. 1951. Nouvelles Découvertes sur les Mathématiques Babyloniens. *Les Conférences du Palais de la Découverte* série D, no. 11.
Bruins, E.M. 1964a. *Codex Constantinopolitanus*. Leiden: Brill.
Bruins, E.M. 1964b. *Heronis Alexandini Metrica*. Leiden: Brill.
Bruins, E.M., and M. Rutten. 1961. *Mémoires de la Mission Archéologique en Iran, Tome XXXIV: Textes Mathématiques de Suse*. Paris: Paul Geuthner.
Bruins, E.M., P.J. Sijpesteijn, and K.A. Worp. 1974. A Greek mathematical papyrus. *Janus* 61: 297–312.
Chemla, Karine, and Guo Shuchun. 2004. *Les Neuf Chapitres: Le Classique mathématique de la Chine ancienne et ses commentaire*. Paris: Dunod.
Fowler, David, and Eleanor Robson. 1998. Square Root Approximations in Old Babylonian Mathematics: YBC 7289 in Context. *Historia Mathematica* 25: 366–378.
Friberg, Jören. 2005. *Unexpected Links between Egyptian and Babylonian Mathematics*. Singapore: World Scientific Publishing.
Friberg, Jören. 2007. *Amazing Traces of a Babylonian Origin in Greek Mathematics*. Singapore: World Scientific Publishing.
Friberg, Jören., Hermann Hunger, and Farouk al-Rawi. 1990. 'Seed and Reeds' a Metro-Mathematical Topic Text from Late Babylonian Uruk. *Baghdader Mitteilungen* 21: 483–557 (**Tables 46–48**).
Heath, T.L. 1921. *A History of Greek Mathematics*, vol. 2. Oxford: Oxford University Press.
Høyrup, Jens. 1997. Hero, ps-Hero, and Near Eastern Practical Geometry: An investigation of *Metrica*, *Geometrica*, and other treatises. In Klaus Döring, Bernhard Herzhoff & Georg Wöhrle (eds.), *Antike Naturwissenschaft und ihre Rezeption*, Band 7. Trier: Wissenschaftlicher Verlag Trier, pp. 67–93. Reprinted in Høyrup, *Selected Essays on Pre- and Early Modern Mathematical Practice*. Berlin: Springer, Ch. 10, 229–255 **(references to this edition)**.
Høyrup, Jens. 2002. *Lengths, Widths, Surfaces: A Portrait of Old Babylonian Algebra and Its Kin*. New York: Springer.
Knorr, Wilbur Richard. 1978. Archimedes and the Elements: Proposal for a Revised Chronological Ordering of the Archimedean Corpus. *Archive for History of Exact Sciences* 19: 211–290.
Knorr, Wilbur Richard. 1982. Techniques of Fractions in Ancient Egypt and Greece. *Historia Mathematica* 9: 133–171.
Knorr, Wilbur Richard. 1989. *Textual Studies in Ancient and Medieval Geometry*. Boston: Birkhäuser.
Masià, Ramon. 2015. On Dating Hero of Alexandria. *Archive for History of Exact Sciences* 69: 231–255.
Neugebauer, Otto. 1935. *Mathematische Keilschrift-Texte*. Berlin: Springer.
Parker, Richard A. 1972. *Demotic Mathematical Papyri. Brown Egyptological Studies vii*. Providence: Brown University Press.
Robson, Eleanor. 1999. *Mesopotamian Mathematics, 2100–1600 BC: Technical Constants in Bureaucracy and Education*. Oxford: Clarendon Press.
Robson, Eleanor. 2008. *Mathematics in Ancient Iraq: A Social History*. Princeton: Princeton University Press.
Schoene, Hermann. 1903. *Heronis Opera quae supersunt: vol. iii: Rationes Dimetiendi et Commentatio Dioptrica*. Leipzig: Teubner.
Sijpesteijn, P.J., and K.A. Worp. 1974. Literary and Semi-Literary Papyri from the Vienna Papyrus Collection. *Chronique D'égypte* 49: 342–350.
Vaiman, A.A. 1963. Istolkvanie geometričeskich postoyannich iz suzskogo klinopisnogo spiska I (Suzi) (Interpretation of the geometric constants in the cuneiform tablet I from Susa). *Vestnik Drevnei Istorii* 83: 75–86.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.