Histoire & mesure

XXXVII-1 | 2022
Mesure, espace, sciences sociales (varia)

The Greek Mathematical Corpus: a Quantitative Appraisal

Le corpus mathématique grec : une approche quantitative

Fabio Acerbi and Ramon Masià

Electronic version
URL: https://journals.openedition.org/histoiremesure/15779
ISSN: 1957-7745

Publisher
Éditions de l'EHESS

Printed version
Date of publication: June 30, 2022
Number of pages: 15-36
ISBN: 978-2-7132-2943-5
ISSN: 0982-1783

Electronic reference
Fabio Acerbi and Ramon Masià, "The Greek Mathematical Corpus: a Quantitative Appraisal", *Histoire & mesure* [Online], XXXVII-1 | 2022, Online since 01 January 2024, connection on 11 January 2024. URL: http://journals.openedition.org/histoiremesure/15779 ; DOI: https://doi.org/10.4000/histoiremesure.15779

Histoire & Mesure, 2022, XXXVII-1, p. 15-36

The Greek Mathematical Corpus: a Quantitative Appraisal

Fabio Acerbi * & Ramon Masià **

Abstract. This paper assesses the Greek mathematical corpus as a whole using quantitative methods and discusses the methodological import of this approach. A number of dynamics within the corpus are also outlined. They corroborate the view that, in Greek antiquity, demonstrative and non-demonstrative mathematics belonged to one and the same universe of discourse.

Keywords. Greek mathematics, quantitative methods, manuscripts, Gaussian distribution, Pareto distribution

Résumé. Le corpus mathématique grec : une approche quantitative. Nous étudions le corpus mathématique grec au moyen de techniques quantitatives, dont nous discutons l'arrière-plan méthodologique. Cela nous permettra de mettre en valeur plusieurs dynamiques internes au corpus, et de donner des bases à la perspective critique selon laquelle, dans l'antiquité grecque, les mathématiques démonstratives et non démonstratives faisaient partie du même univers de discours.

Mots-clés. mathématiques grecques, méthodes quantitatives, manuscrits, distribution de Gauss, distribution de Pareto

* Centre national de la recherche scientifique (CNRS), UMR 8167 Orient & Méditerranée, équipe "Monde byzantin" (Paris, France). E-mail: fabacerbi@gmail.com

** Magna Societas Catalanorum (Barcelona, Spain). E-mail: rmasia@gmail.com

1. Introduction

Greek mathematics is something more than a handful of celebrated authors. It is a general universe of discourse canonized as a literary genre and comprising organized pieces of mathematics that may greatly differ in content and style.[1] This universe of discourse addresses a number of disjoined readerships: we cannot assume that a reader of Apollonius' celebrated treatise on conic sections could have belonged to the same social and cultural milieu as users of Theon of Alexandria's "Little Commentary" on Ptolemy's astronomical tables.[2] However, different readerships do not necessarily signify that the mathematics concerned can be classed on different "levels". Value-laden categories of this kind cannot be used to assess the extant Greek Mathematical Corpus (GMC henceforth) as a whole.

The specific aim of this study is to assess the GMC as a whole using quantitative methods. Such an approach will allow us to outline a number of dynamics within the GMC, and to corroborate the view that, in Greek antiquity, demonstrative and non-demonstrative mathematics belonged to one and the same universe of discourse. Of course, this perspective is not new.[3] We may, however, claim originality in the use of quantitative methods to assess the GMC *as a whole* and in the way such methods are implemented. The present study is also preliminary to the application of corpus linguistic techniques to segments of the GMC,[4] a research theme we shall pursue in the future: it would be rash to decide in advance that some segments of Greek mathematics comprise a lexicon more suitable for computational analysis than others without a preliminary assessment of the entire GMC.

The strategic aim of this study is to discuss the methodological problems raised by quantitative methods. These problems originate in an obvious feature of the subject matter: the GMC is temporally, materially, and even definitionally

We are grateful to Bernard Vitrac for his critical reading of the manuscript and his valuable suggestions.

1. This universe of discourse is optimized: mathematical contents expressed using one stylistic code cannot be satisfactorily expressed with another. The stylistic codes adopted in Greek mathematics are the demonstrative code, the procedural code, and the algorithmic code. See F. Acerbi, 2021a, Section 1.1–3, for a detailed description of each.

2. Of course, allowance must be made for the nearly five centuries that separate the two authors. Typical users of Theon's "Little Commentary" were professional astrologers of the middle and late Imperial Age: A. Jones, 1994.

3. Compare the approach adopted in S. Cuomo, 2000; ead., 2001.

4. The first application of computational linguistics to parts of the GMC is the detailed analysis of Archimedes' lexicon in R. Masià, 2012; very specific and less detailed discussions can also be found in F. Acerbi & B. Vitrac, 2014, pp. 59-73 (Hero of Alexandria), and in F. Acerbi, 2021a, pp. 28-36 and 157-159 (Euclid). Corpus and computational linguistics tools are applied to Mediaeval scientific texts written in Latin in P. Roelli, 2021; see also the synthesis in P. Roelli, 2020.

coarse-grained. First, we are often unable to locate a Greek mathematician in time more precisely than over a span of several centuries.[5] Second, some mathematical treatises have been lost, or partly lost, in Greek.[6] Third, some mathematical works are not authorial undertakings but collections of disparate texts that had been stabilized before their earliest extant manuscripts were copied;[7] for this reason, these works have been handed down as a unitary whole. Worse still, some of these collections are no more than philological artefacts assembled in modern times.[8] Finally, and in fact primarily, we must first decide what we mean by a mathematical work. In other words, it is not immediately obvious how a "(mathematical) work" for inclusion in the GMC should be defined and how it should be categorized, what corresponds to an "author" of any such "work", or what temporal location must be assigned to any such author. Apart from problems of coarse-graining, using quantitative methods to study the GMC entails introducing methods and tools applicable to a database filled with relatively heterogeneous items; for this reason, we shall adhere to a terminology that does not necessarily consider these items simply as works written by a single author.

These issues are addressed in Sections 2–6. We first discuss the several ways the GMC can be defined; our definition assigns a central role to works, not to authors (Sections 2–3). We then define the elements of our database. These elements are the "GMC tokens"; they are the core itemizers of our database and must be carefully distinguished from the works that make up the GMC (Section 4). A series of relevant pieces of information – like authorship (if any), temporal location, and size in words – are linked to these itemizers (Sections 4–6). It is not said that a GMC token is a work by a single author; conversely, parts of a treatise may be categorized as different GMC tokens. If a GMC token has an author, we explain how we locate the author's activity within the time span in which he can reasonably be assumed to have lived. A further issue central to our perspective, that of how to categorize the GMC tokens, is addressed in Section 7. We use a tripartite scheme: contents, genre, and style, with each of these categories being subdivided into suitable subcategories. In this way, the natural dynamics of the GMC on the time axis acquires further dimensions. Sections 7 and 8 present the results of our

5. Cases in point are Diophantus and Hero of Alexandria: see the most recent discussions in F. Acerbi, 2011b, p. 1, and F. Acerbi & B. Vitrac, 2014, pp. 15-22 and 103-115, both with bibliography, and R. Masià, 2015.

6. For example, most advanced treatises of Apollonius are totally lost (a discussion can be found in F. Acerbi, 2011a), one of them is lost in Greek but has survived in Arabic translation (*On the Cutting Off of a Ratio*), only four Books of his *Conics* are extant in Greek, three more exist only in Arabic translation, and Book VIII is lost. As we shall see in Section 5, the size of several lost works can be estimated.

7. Most works included in the geometric metrological corpus are of this kind; see the discussion in F. Acerbi & B. Vitrac, 2014, Étude complémentaire III.

8. See again the discussion in F. Acerbi & B. Vitrac, 2014, Étude complémentaire III.

investigation in the form of figures and plots; these are followed, in Section 9, by a short assessment of our methods and results.

2. Defining the GMC: false leads

The hardest task is to delimit the GMC. This requires jointly defining what is a Greek mathematical text. Several definitions can be envisaged. The simplest is strictly author-centred and hinges on the consensus of modern scholarship: a Greek mathematical text is any work by any ancient Greek author recorded in a reference work such as the *Dictionary of Scientific Biography*, or, alternatively, any work of anyone identified as a mathematician in the Pauly-Wissowa encyclopaedic lexicon.[9] This criterion is perfectly legitimate, but it has two major drawbacks: on the one hand, it neglects most anonymous mathematical works; on the other, it includes works which may have little or no mathematical content.[10] Moreover, it simply shifts the definitional problem to a higher category, namely, whether we should include texts in fields such as astronomy (certainly), engineering (almost certainly), or astrology (perhaps) in the GMC.

A similar author-centred criterion points to the social role as recognized in ancient sources; as stated by the editor of the most important lexicon of Greek philosophers: "The main criterion [*scil.* for assigning a lexicon entry to a name] was for a person to have been described as a philosopher or a philosopher of some philosophical school in ancient sources, to have produced or have been said to have produced philosophical treatises, to have expressed unmistakably philosophical ideas, or to have taught philosophy to some disciple(s)".[11] This criterion can hardly be applied to mathematics. First, several epithets were used in antiquity to qualify someone as a mathematician, but their meanings overlap and, more importantly, overlap with meanings we might not wish to use for a mathematical activity: for instance, a mathematician may be someone who practices astrology.[12] Second, and conversely, authors

9. These are C. C. Gillispie, 1970–81 and G. Wissowa, 1894–1972, respectively. The latter is the standard lexicon for classical antiquity.

10. Cases in point are Ptolemy's epistemological work about the criterion of truth and his astrological treatise *Apotelesmatika*.

11. R. Goulet, 2013, pp. 12-13. Note the careful wording in disjunctive form: any of these conditions suffices for identifying a writer as a philosopher. Of course, this definition leaves room for (arbitrary) choices in its third disjunct, in particular due to the presence of the adverb "unmistakably". The lexicon is the *Dictionnaire des philosophes antiques* (R. Goulet, 1994–2018), and describes nearly 2,500 philosophers, among whom Ptolemy (he wrote a short treatise on epistemology), Euclid and Hero of Alexandria (who are not reported to have written anything philosophical).

12. A widely used Greek epithet for astronomer, and even for astrologer, is μαθηματικός (see the terminological point in Sextus Empiricus, *Adv. math*. V.1–2), a term more generally used to describe anyone involved in investigations whose character was more or less markedly

who practiced mathematics but were never called mathematicians, simply because their main activities were categorized differently in antiquity: think of Hero of Alexandria, the "engineer",[13] who wrote a strictly mathematical commentary on Euclid's *Elements*, and an amazing example of register-crossing pure mathematics such as the *Metrics*.[14] Third, our evidence is sometimes so scant that cross-checking is impossible. Fourth, sources indicate that teaching of mathematics did not become a sustained activity until the beginning of the 4th century CE; attempts at locating Euclid or any other mathematician prior to that period in an educational context are grounded on an unwarranted back-projection of the modern organization of academic work.[15] The social-role criterion for delimiting the GMC thus clearly obliges us to make a priori choices. For instance, consider Pappus' *Collection*, the most natural ancient source for identifying Greek mathematicians.[16] It is obvious that Apollonius – one of the greatest mathematicians of all time – must be included in the GMC, and in fact a lot of his mathematics is commented on by Pappus in the *Collection*; yet Apollonius is merely described as "from Perge" (the town where he was born) in Pappus' work. Likewise, Pappus' references to Plato fit the criteria for inclusion in the GMC,[17] yet he does not cite Diophantus. Taking the set-theoretical union of such sources would not solve the problem, for the case of Pappus shows that we would be obliged to include in the GMC authors whom no sensible scholar would define as mathematicians.

A third criterion holds that mathematics is what mathematicians do, and vice versa, where the meaning and the extension of the undefined term

---

theoretical (see the tripartition of theoretical philosophy discussed by Ptolemy, following Aristotle, *Metaph*. E.1, at the beginning of *Alm*. I.1 and of *Harm*. III.6). Note that the organization of scientific disciplines adopted, mainly for teaching purposes, in the Greco-Roman world was the so-called *quadrivium*; it comprises arithmetic, harmonic theory, geometry and astronomy (ordered by decreasing level of abstraction). The generic name of each of these disciplines was μάθημα, so that a μαθηματικός was anyone engaged in any of these disciplines. See B. Vitrac, 2005 for the classifications of sciences in Greek antiquity.

13. The Greek term is normally μηχανικός.

14. See again the discussion in F. Acerbi & B. Vitrac, 2014, Sections 1–2.

15. This anachronistic perspective is not typical of the modern assessment of the social role of ancient mathematicians: even if it is clear that ancient philosophers were actively and constitutively engaged in teaching, scholars since the end of the 19th century have assumed, for instance, that Aristotle's school worked in exactly the same way as the German universities of the period did, or, more precisely, as the great *Akademien* projects did and still do: a project leader distributes the research tasks and a crowd of slaves and sub-slaves carries them out, but the leader publishes the results.

16. For a presentation of this huge mathematical compilation, in which so many mathematicians are mentioned, see A. Jones, 1986, Introduction. Pappus lived in the first half of the 4th century CE.

17. At *Coll*. V.34, Pappus says that the well-ordered solids are "not only the five figures by the most divine Plato", but also the thirteen semi-regular solids discovered by Archimedes. If this were the only piece of information on "divine" Plato, we might conclude that he was a top-ranking mathematician who made a detailed study of the five regular polyhedra.

(namely, either "mathematics" or "mathematician") are established a priori. A definition of this kind is used for instance by Reviel Netz: "[w]hoever has written (or perhaps merely produced orally) an argument showing the validity of some claim, using the techniques we identify with Greek mathematics (a lettered diagram, a specific mode of language use) is […] a mathematician", and again: "anyone who has written down an original mathematical demonstration, no matter in what context".[18] This definition excludes all mathematics that is not demonstrative. Moreover, as we have no way of determining what "we identify with Greek mathematics" other than by reading it in the work of a Greek mathematician, this definition is circular.

3. Delimiting the GMC

The point of the discussion presented in the previous section is clear: a choice must be made. Granted, but who makes the choice? Of course, *we* make the choice, but we may reduce its dependence on a modern notion of mathematics by anchoring our choice to some piece of the historical record. Our criterion takes manuscript tradition as a suitable anchoring condition, delegates the choice to the scholars responsible for handing Greek mathematics down to us, adds an exclusion, and is stated as follows: any structured macro-piece of discourse whose author is included (false ascriptions included) in the major scientific encyclopaedias of Palaiologan Byzantium, or which elaborates on any of these pieces, or which is elaborated on by any of these pieces, is a GMC text, and so on recursively. We exclude pieces of discourse of the said authors that no ancient authority would have included in a *quadrivium*.

Before discussing the import of our sufficient condition, let us clarify a couple of non-obvious points. First, the Palaiologan dynasty ruled the Byzantine Empire from 1261 to 1453, namely, from the end of the Latin rule (1204–1261) to the year in which the Ottomans finally seized Constantinople, and in which, consequently, Greek manuscripts started to drift massively to the West. This was the period of Byzantine history in which mathematics was mostly praised and "practised"; imperial patronage was decisive. Second, as is the case for so many other intellectual activities in Byzantium, practising mathematics mainly consisted in transmitting the Greek heritage by copying from available manuscripts to produce new books that contained a representative segment of the Greek literary production. This is how most

18. R. Netz, 1997, p. 4; id., 2002, p. 197, the latter explicitly reacting to the approach in S. Cuomo, 2000. See also the discussion in R. Netz, 1999, pp. 277-278.

ancient Greek literature has come to us;[19] the editions we read are based on these manuscripts.

In the Palaiologan period, huge mathematical encyclopaedias were composed by assembling primary sources. The encyclopaedias in question comprise a two-volume set, Par. gr. 2342 (*Diktyon* 51974) + Vat. gr. 198 (*Diktyon* 66829), and two single, imposing manuscripts, Vat. gr. 191 (*Diktyon* 66822) and Vat. gr. 192 (*Diktyon* 66823). The two-volume set was penned in around 1360–70 by the scholar and copyist Malachias; it contains a complete *quadrivium*. The second manuscript was written by a number of unknown copyists about 1270–90 and was assembled in around 1296 by the scholar John Pediasimos.[20] Vat. gr. 192 was written by a team of professional copyists in the same period as Vat. gr. 191 and likewise assembled for encyclopaedic purposes. There are no comparable collections in the entire manuscript tradition of Greek scientific works. These encyclopaedias do not count as decisive manuscript witnesses of the works they contain: there are usually more ancient and more authoritative witnesses, and Malachias' great *quadrivium* only carries recensions.[21] Yet, these encyclopaedias offer a clear picture of what Byzantine scholars, who saw themselves as direct inheritors of ancient Greek thought, took to be the GMC.

Our criterion is a disjunction with three disjuncts, plus an exclusion. The first disjunct selects the contents of these four manuscripts. In particular, they contain the entire corpus of harmonic theory; this is one of the four disciplines of the *quadrivium*, hence a perfectly legitimate mathematical discipline under ancient standards. The second and the third disjunct, along with the recursive qualification, extend this core of GMC texts to include "exegetic chains". We thus take account of the fact that Greek mathematics is first and foremost a literary genre, representing a tradition in itself. Accordingly, works of authors like Menelaus, Hero of Alexandria, or Anthemius of Tralles, and the entire poliorcetic corpus, which do not feature

19. As a (small) part of ancient Greek literature has been transmitted by papyri only, the predeterminer "most of" is necessary.

20. On these four manuscripts, see F. Acerbi, 2016, *passim* but in particular pp. 154-160, 189-190 (Par. gr. 2342 and Vat. gr. 198), and F. Acerbi & A. Gioffreda, 2019, pp. 30-34 and 41-44 (Vat. gr. 191), and 19, 30, and 44-46 (Vat. gr. 192), with a detailed description of their content (featuring small corrections to the descriptions found in the standard catalogues) and a complete bibliography. Further bibliographic information can be found by searching the reference website Pinakes (URL: https://pinakes.irht.cnrs.fr/) by using the *Diktyon* numbers given in the text. For Byzantine mathematics, see F. Acerbi, 2020.

21. A recension of a mathematical work is a revision of its content, usually carried out in conjunction with a new edition and varying in extent: for example, most manuscripts of Euclid's *Elements* correspond to the light recension authored by the mid-4th-century scholar Theon of Alexandria; the only extant version of Apollonius' *Conics* is the heavy recension of the early 6th-century Neoplatonic philosopher Eutocius. See F. Acerbi, 2016 for a complete survey of Byzantine recensions.

in any of the above manuscripts,[22] are GMC texts because they elaborate on other GMC texts (second disjunct of the criterion, with the recursive qualification). Likewise, the works of Archimedes are not included in any of the above manuscripts, but they are nonetheless GMC texts because Theodosius, an author included in Vat. gr. 191 and Par. gr. 2342, wrote a (now lost) commentary on Archimedes' *Method* (third disjunct of the criterion). The entire metrological corpus is included because ancient sources ascribe parts of it to Hero (parenthetic qualification in the first disjunct); the entire set of mathematical papyri are included accordingly.[23] As for the exclusion, this makes the manuscript-based sufficient condition a criterion, by adding a necessary condition. The exclusion allows us to eliminate astrology, and, for instance, the works of any author that ancient categorizations would have regarded as philosophical, as is the case for most of Proclus' commentaries and some of Iamblichus' works.[24]

Our criterion is partly author-centred, but this is the only approach that allows us to count the entire production of authors such as Archimedes and Hero of Alexandria among the GMC texts. Most importantly, using authorship as a bridge between classes of GMC texts allows us to estimate the share of lost works among the GMC texts. This is crucial because most of Apollonius' production is lost, but we can nevertheless estimate its size (see Section 5).

4. Defining a GMC token

In our database, a GMC token corresponds to a (self-contained part of a)[25] GMC text as it appears in the modern critical edition that serves as a reference.[26] As for the requirement that a GMC token be representative of a self-contained piece of mathematical discourse, the boundaries of self-containment are usually obvious: most of our GMC tokens are canonical works which we can categorize within the literary-theoretical genre of the "treatise".[27] By definition, a treatise is written by an author, even if his or her

22. No works of Menelaus have survived in Greek.

23. An overview of mathematical papyri is contained in R. Bagnall & A. Jones, 2019.

24. The principle of exegetic chains allows us to include Proclus' commentary on *Elements* I and Iamblichus' paraphrase of Nicomachus' *Introduction to Arithmetic* among the GMC texts.

25. See also Section 7 for this qualification.

26. This requirement is motivated by the fact that almost any occurrence of a GMC token in a manuscript carries a specific text, which differs to a certain extent from the text of any other occurrence of the "same" GMC token. Philology uses standard methods to take these differences into account; the result is a "stabilized" text, printed in a "critical edition". Quantitative methods can only be applied to the text presented in an edition.

27. These boundaries are only "usually" obvious: in many manuscripts, the *Elements*, which comprises 13 Books, has been transmitted with an adjunct made of two spurious Books,

name may be unknown to us, or even if we know little more than his or her name. However, some GMC tokens – that may be significant and large – are not authorial undertakings. These are the collections of problems that make up the geometric metrological corpus and the collections of scholia (namely, annotations to other GMC tokens carried in the margins of the relevant manuscripts).[28] The non-authorial GMC tokens have usually taken shape across many centuries by means of accretions to some original core. They are the result of two editorial activities: first, the activity of late antique and Byzantine copyists and scholars who shaped and delimited these collections, making them into the self-contained and structured macro-pieces of mathematical discourse typical of Byzantine manuscripts; second, the activity of modern editors, who have sometimes combined Byzantine collections to make GMC tokens. These GMC tokens are nothing more than philological artefacts. However, the way the GMC tokens are embedded in the time axis will neutralize this feature.

In general, our criterion enables us to include in the GMC a good deal of mathematics that is considered peripheral by current scholarship; at the same time, our criterion is clearly worded to allow the inclusion of all authors and works we reasonably regard as mathematics. The main adjustment lies in allowing exegetic chains, so that works that do not figure in the selected manuscripts can also be included. We might have eliminated exegetic chains in our criterion by broadening the set of selected manuscripts, but this amounts to including non-encyclopaedic compilations, thereby making our definition of the GMC as arbitrary as any other choice discussed above. Fortunately, only a handful of manuscripts must be added to the above-mentioned encyclopaedias in order to get the entire GMC corpus as identified under the first disjunct of our criterion.[29] Finally, the fact that the formulation of our criterion is not *so* contrived shows that the Byzantines clearly understood that Greek mathematics constitutes a tradition: we must not forget that they had already contributed decisively to deciding what kind of Greek mathematics we had to read.

---

one of which was authored by Hypsicles (2nd century BCE), the other by an unknown pupil of the architect Isidorus of Miletus (early 6th century). That these Books are spurious is obvious, but their mere presence means that the boundaries of the *Elements* are less "obviously" defined than, for instance, the boundaries of the *Almagest*.

28. The former is edited by J. L. Heiberg in volumes IV and V of the *Opera omnia* of Hero of Alexandria; a selection of the latter is edited in the volumes of Euclid's *Opera omnia*. Scholia may be extracts from commentaries, but we shall treat all scholia as if they were independent pieces of writing. Outside the GMC but still in a the context of ancient Greek writings, the category of non-authorial works includes the so-called Byzantine *Rechenbücher* (see F. Acerbi, 2019) and the Easter Computi (see F. Acerbi, 2021b).

29. This follows from the fact that Greek mathematics is transmitted in corpora, namely, thematically-based sets of works collected in a single manuscript.

Once a GMC token is individuated, a major problem is to endow it with a diachrony. This means locating each GMC token in time. In the case of authorial GMC tokens, this is equivalent to dating the Greek mathematician who authored the intended GMC token.[30] We apply the dates of birth and death (or, *faute de mieux*, the *post quem* and *ante quem* dates) currently accepted by modern historiographers, mark them on the time axis, and then elaborate an output (the GMC token distribution) on this time segment as indicated in Section 6. For reasons of definiteness – and this is a non-trivial assumption – we restrict the lifespan of a given author to 80 or 100 years centred on the midpoint of the currently assumed lifespan, even if the possible lifespan is wider. For example, we assign to Archimedes the lifespan [–287,–212] because we know that he died in 212 BCE, when the Romans captured Syracuse, and a late source claims that he lived for 75 years.[31] We assign the lifespan [–240,–160] to Apollonius, [–320,–240] to Euclid, and [90,170] to Ptolemy. We deem that all these date ranges are exact because our documents allow us to determine the lifespan quite accurately, and all these lifespans are set to 80 years. Authors whose biographical data are less clear-cut are assigned a lifespan of 100 years; we deem these date ranges to be approximate. For instance, we assign the interval [0,100] to Hero of Alexandria, and the interval [200,300] to Diophantus.

If a GMC token is non-authorial, it is located on the time axis on the basis of historical evidence; we are generous in assigning the size of the operative time interval. On account of the mechanism of formation outlined above, the geometric metrological corpus is assigned to the interval [200,700], and the scholia to the *Elements* are located in the interval [500,700], even if most of them may well be later.

Based on our estimates, the GMC is embedded in the segment of the time axis corresponding to the interval [–380,700], a period of more than one thousand years.

30. A survey of the several ways Greek mathematicians are dated is given in F. Acerbi, 2010, pp. 79-86. Astronomers can usually be dated accurately because they may report observations of celestial phenomena that can be identified with certainty. We are, of course, unable to assign a "publication date" to any GMC token; even the relative chronology of the works of authors such as Archimedes is uncertain. Moreover, some GMC tokens (a well-known example is Apollonius' *Conics*) comprise parts composed at different times and "published" by sending these parts to different addressees (in ancient terminology, the "parts" are the "books" making up a treatise).

31. We disregard the fact that this source is notoriously unreliable. As there is no year 0 CE, 212 BCE corresponds to –211 and not to –212. This discrepancy is irrelevant in our perspective.

5. Estimating the size of a GMC token

Since a GMC token is a self-contained work, we can define its natural size as the number of words it contains. This number is calculated uniformly as follows. If a GMC token is extant in its entirety, we use the number of word tokens indicated in the online *Thesaurus Linguae Graecae* as the base estimate of its size.[32] To simplify matters, this number includes the text segments that the editor of the intended work has marked for deletion because they are regarded as non-authentic (with an important exception,[33] these segments always represent a very small percentage of a work's total size). The size of the lost part of partly lost works is estimated by bookwise proportion with respect to the extant part,[34] even if, in some cases, it is obvious that the size of what has been lost was smaller than the size estimated by proportion.[35] The size of several works that are totally lost or transmitted only in translation can be estimated thanks to a remarkable resource:[36] their description in Pappus' *Collection*. In particular, Pappus gives the number of propositions contained in any work included in the so-called "analytical corpus", a very advanced mathematical resource comprising several treatises and now almost entirely lost. For some works of the analytical corpus, in which a single problem branches off into a multitude of cases, Pappus adopts the "diagram" counting unit. For instance, the two books of Apollonius' *Plane Loci* comprise 15 loci (each

32. This means that we must exclude astronomical tables from the GMC. See the *Thesaurus Linguae Graecae* website, URL: http://stephanus.tlg.uci.edu/.

33. The exception is the Euclidean corpus. For instance, the material that Heiberg "bracketed" or placed in the Appendices of his 4-volume edition of the *Elements* amounts to about 8% of the text. In his forthcoming Budé edition of the *Elements*, B. Vitrac (*per litteras*) will bracket about 12% of the transmitted text of the Euclidean treatise as spurious. For this reason, in our database, the texts that Heiberg placed in the Appendices (these are mainly lemmas and alternative proofs) are assigned to Anonymous, under the title "Elements alternative proofs", and similarly for the alternative proofs contained in Euclid's *Data* and for one of the two redactions in which we read Euclid's *Optics* and *Phaenomena* (and, separately, the latter's alternative proofs).

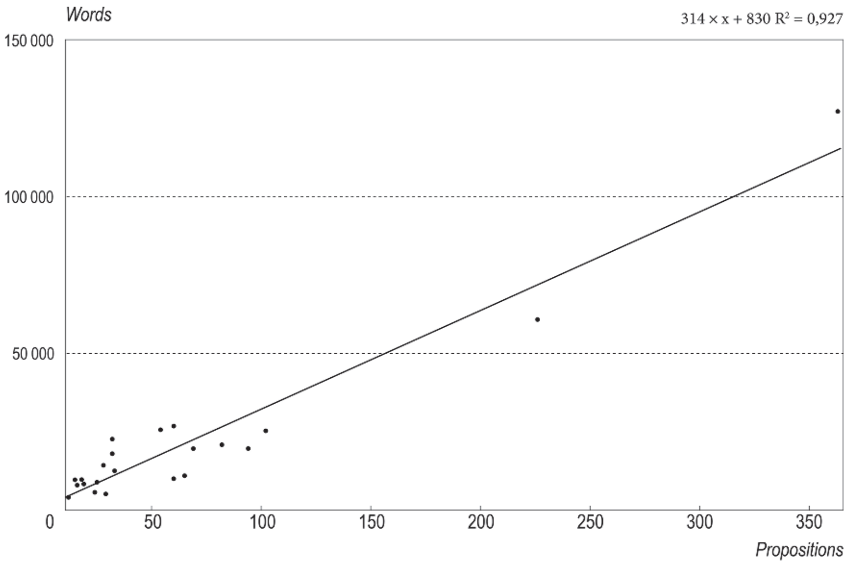
34. The expression "by bookwise proportion" means that the ratio of number of lost books to number of extant books is equal to the ratio of number of words contained in lost books to number of words contained in extant books.

35. A case in point is Theon of Alexandria's commentary on the *Almagest*. Books I–IV and VI are transmitted by the early 9th century manuscript Laur. Plut. 28.18, Book V can only be read in the margins of Vat. gr. 198, Book XI is irremediably lost, Book VII can be read in a Byzantine recension only, and Books VIII–X and XII–XIII can be read – certainly mutilated and possibly in a recension again – only in the 13th century manuscript Vat. gr. 1087. Despite these problems, it is clear that the first books were larger than the last books: basics such as how to divide two sexagesimal numbers have to be explained at the beginning of the commentary.

36. Another problem is to estimate how many Greek mathematical texts are radically lost, in the sense that we have no information about them at all. For instance, the sole reference to Demetrius of Alexandria's work on higher-order curves is the title "Linear investigations", mentioned, along with its author, by Pappus in *Coll*. IV.58. See Section 9 for a discussion of this issue.

of which is a “main problem” of sorts), but Pappus states that they “contain 147 theorems or diagrams” (*Coll.* VII.26). Relying on Pappus’ descriptions, we estimate the size in words of these works by linear regression from the size in words vs. number of propositions in the works for which these two quantities are known (Figure 1).[37]

Figure 1. *Linear regression words vs. propositions*



Note. This figure is also accessible online, attached to the database.
URL: https://nakala.fr/10.34847/nkl.22517tj9

6. Mapping the GMC tokens on the time axis

Once the GMC tokens are defined and temporally located, we elaborate an output (the GMC token distribution) on the relevant time segments as follows.

If nothing is known about the author(s) of a GMC token or if this author does not exist, we stipulate that the intended GMC token was composed at a uniform rate during the relevant time segment. Thus, the output of an

37. Using linear regression is more robust than using the average number of words in a proposition as the scaling factor. This is because most Greek mathematical treatises carry preliminaries, like prefaces or lists of principles, that do not contain propositions. The size of such preliminaries is normally uncorrelated with the number of propositions that follow.

anonymous author or the activity that has produced a collection of mathematical items (as scholia and metrological works) is allocated to the relevant time segment by means of a continuous uniform distribution; the area under the distribution is equal to the size of the output as measured in words.

If an item of the GMC is authorial and the author's name is known, this item is allocated to the relevant time segment by means of a suitable Gaussian distribution centred on the midpoint of the interval (in ancient terms, this coincides with the *akmē* of the author). Here too, the area under the distribution is equal to the size of the output as measured in words.[38] If an author can be dated exactly in the sense specified above, the standard deviation σ of the Gaussian distribution is 1/12 of his lifespan; if an author cannot be dated exactly, the standard deviation is 1/6 of his lifespan. This is a simple numerical encoding of the loss of information with respect to the case where an author can be dated exactly.

For example, we assume that Archimedes' production is centred on year 37.5 of his currently assumed lifespan of 75 years, and that he wrote about 68% of his production ± 6.25 years therefrom (that is, when he was 31 to 44), about 95% of his production ± 12.5 years therefrom (that is, when he was 25 to 50), and almost nothing when he was very young and very old.[39] Our choice is dictated by requirements of uniformity, even if this choice is arbitrary whenever an author cannot be dated exactly; for this reason, we have usually estimated an author's lifespan by a time segment of no more than a century.

7. Categorizing GMC tokens

An overall assessment of the GMC cannot simply consist in putting the entire GMC on the time axis and estimating the temporal evolution of its size. Several literary-theoretical parameters that characterize the GMC have also evolved over the timespan of nearly 1,000 years that supports the GMC, and one of our aims is to reveal such partial evolutions, showing at the same time that they are related to each other.

---

38. The two distributions are chosen according to standard criteria of information theory: the uniform distribution is the maximum entropy distribution among all continuous distributions that are supported in the intended interval; the normal distribution is the maximum entropy distribution among all real-valued distributions supported in an interval and with assigned mean and standard deviation. For basics of information theory see T. M. Cover & J. A. Thomas, 2006.

39. The example of Archimedes also shows that some historical data must be neglected for uniformity's sake, for Archimedes was apparently active until the very end of his life.

The GMC is appraised by introducing the following literary-theoretical categorizations:

– *Contents*. We modify a standard *quadrivium* and set the possible contents as pertaining to arithmetic (as, for instance, in Diophantus' *Arithmetica*; in modern terms, this is number theory; this includes harmonic theory, represented for instance by Ptolemy's *Harmonica*, and "logistic", a discipline whose primary aim is computation, represented, for instance, by the geometric metrological corpus), geometry (a part of Euclid's *Elements* and the whole of Apollonius' *Conics*), astronomy (Ptolemy's *Almagest*); or applied mathematics, under the heading "technical" (Euclid's *Optica*; these are mainly geometric models of natural phenomena other than astronomical phenomena). Different parts of a work may belong to different subcategories; for instance, *Elem*. VII–IX belong to arithmetic, the rest of the *Elements* to geometry; for this reason, the *Elements* is split into two GMC tokens.

– *Literary genre*, which is subdivided as follows: formal treatises (such as Apollonius' *Conica*; a well-defined subject is treated, and the work is intended for specialists), popularization (Nicomachus' *Introduction to Arithmetic*), commentaries (their explicit goal being to elucidate formal treatises, as in Pappus' and Theon's commentaries on the *Almagest*), and compilations (the scholia and the geometric metrological corpus).

– *Style*, which can be canonical, that is, demonstrative (Euclid's *Elements*, which is a sequence of propositions not embedded in a discursive frame) or algorithmic (the geometric metrological corpus), mixed (Ptolemy's *Almagest*; the discursive frame is present but the mathematical content is the focus of the treatise), informal (Nicomachus' *Introduction to Arithmetic*, where it can be argued that the discursive frame carries the focus).

Assigning some of these categories is easy and uncontroversial; assigning others may not be. We decided to split some treatises into several GMC tokens in order to fine-tune the above categorizations. Some treatises may be written in different styles, such as Hero's *Metrica*, for instance, which mixes demonstrative and algorithmic codes. The reader will find all our choices in the database uploaded online.[40] As shown in Figure 2, the database sets out the following items in tabular form: extremes of the time range; whether the time range is exact or approximate; author (including Anonymous); title of the GMC token; title of the work; fraction of a whole work in word size; number of "books"; number of words; number of characters; ratio of characters to words; whether the last two figures are exact or approximate; content; genre; style; number of "diagrams" for lost works; number of propositions for extant works; comment; parameters and nature of the GMC token distribution.

40. URL: https://nakala.fr/10.34847/nkl.22517tj9. See F. Acerbi & R. Masià, 2022.

Figure 2. *A part of the GMC token database*

						prop/words regression			…
						314*x + 830			…
date b	date e	date	author	chunk	work	#work	#bo	Words	…
100	700	approximate	Anonymous	arithmetical papyri	arithmetical papyri	1,00	NA	10000	…
100	700	approximate	Anonymous	astronomical papyri	astronomical papyri	1,00	NA	10000	…
200	700	approximate	Anonymous metrology	Geodesy	Geodesy	1,00	1	4831	…
200	700	approximate	Anonymous metrology	Geometrica	Geometrica	1,00	NA	32923	…
200	700	approximate	Anonymous metrology	Liber geeponicus	Liber geeponicus	1,00	1	6534	…
100	300	approximate	Didymus	Measures of marbles and timber	Measures of marbles and timber	1,00	1	1344	…
100	300	approximate	Diophanes	metrological compilation	metrological compilation	1,00	1	3211	…
100	700	approximate	Anonymous	metrological papyri	metrological papyri	1,00	NA	10000	…
200	700	approximate	Anonymous metrology	On measures	On measures	1,00	1	5027	…
200	700	approximate	Anonymous metrology	Stereometrica	Stereometrica	1,00	NA	17529	…
-200	-120	exact	Diocles	On burning mirrors	On burning mirrors	1,00	1	3970	…
90	170	exact	Ptolemy	Planisphere	Planisphere	1,00	1	7110	…
-287	-212	exact	Archimedes	Stomachion	Stomachion	1,00	1	7110	…
-320	-240	exact	Euclid	Division of figures	Division of figures	1,00	1	10250	…
0	100	approximate	Hero of Alexandria	On mirrors	On mirrors	1,00	1	10250	…
-240	-160	exact	Apollonius	On contacts	On contacts	1,00	2	19670	…

…	prop/words regression					prop/chars regression							Graphs		
…	314*x + 830			314	830	1510*x + -3199	1510	-3199							
…	#work	#bo	Words	Chars		data	content	genre	style	numb	number	Con	μ	σ	gauss?
…	1,00	NA	10000	45000	4,5	approximate	arithmetical	compilation	canonical				400	173,21	NO
…	1,00	NA	10000	45000	4,5	approximate	astronomical	compilation	canonical				400	173,21	NO
…	1,00	1	4831	23930	4,95	exact	geometrical	compilation	canonical				450	144,34	NO
…	1,00	NA	32923	155171	4,71	exact	geometrical	compilation	canonical				450	144,34	NO
…	1,00	1	6534	31824	4,87	exact	geometrical	compilation	canonical				450	144,34	NO
…	1,00	1	1344	6388	4,75	exact	geometrical	compilation	canonical				200	33,33	SI
…	1,00	1	3211	15629	4,87	exact	geometrical	compilation	canonical				200	33,33	SI
…	1,00	NA	10000	45000	4,5	approximate	geometrical	compilation	canonical				400	173,21	NO
…	1,00	1	5027	23847	4,74	exact	geometrical	compilation	canonical				450	144,34	NO
…	1,00	NA	17529	80573	4,60	exact	geometrical	compilation	canonical				450	144,34	NO
…	1,00	1	3970	11901	3,00	approximate	geometrical	treatise	canonical	10			-160	6,67	SI
…	1,00	1	7110	27001	3,80	approximate	astronomical	treatise	canonical	20			130	6,67	SI
…	1,00	1	7110	27001	3,80	approximate	geometrical	treatise	canonical	20			-250	6,25	SI
…	1,00	1	10250	42101	4,11	approximate	geometrical	treatise	canonical	30			-280	6,67	SI
…	1,00	1	10250	42101	4,11	approximate	technical	treatise	canonical	30			50	16,67	SI
…	1,00	2	19670	87401	4,44	approximate	geometrical	treatise	canonical	60			-200	6,67	SI

8. Figures and charts

Figure 3 shows a Sunburst interactive chart that represents the GMC. It shows the prominent role of astronomy, the extent of Ptolemy's legacy, and the productivity of the two late antique authors Pappus and Theon.[41]

We also present an interactive timeline graph that represents the GMC token distributions in terms of works, words, and characters on the interval [–450,750] (Figure 4). This graph also displays all the categories established in Section 7, and the integral of the distribution curves. These integral lines show the cumulative evolution of the indicated parameters.

Both interactive graphs, which are automatically generated from our database by means of non-trivial dedicated software, are uploaded at our GitHub pages;[42] detailed instructions for use are also provided there.

41. This is based on the pie charts found in R. GOULET, 2013 and in B. VITRAC, 2021, p. 40.
42. URL: https://ramonmf.github.io/GraphAGM/.

Figure 3. *Interactive Sunburst chart representing the GMC*

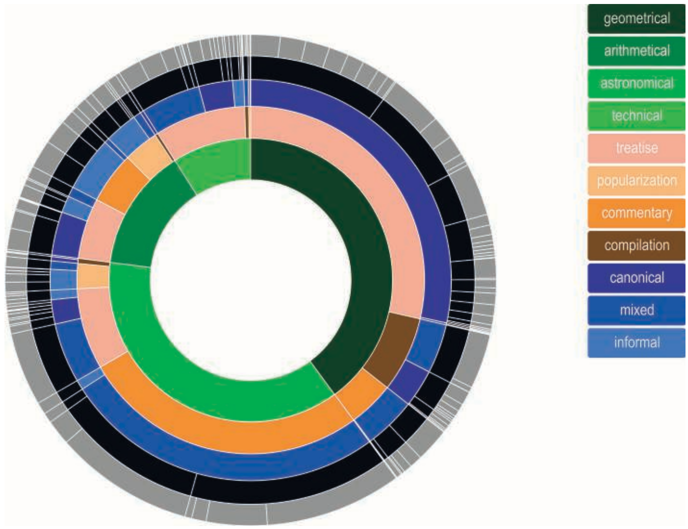
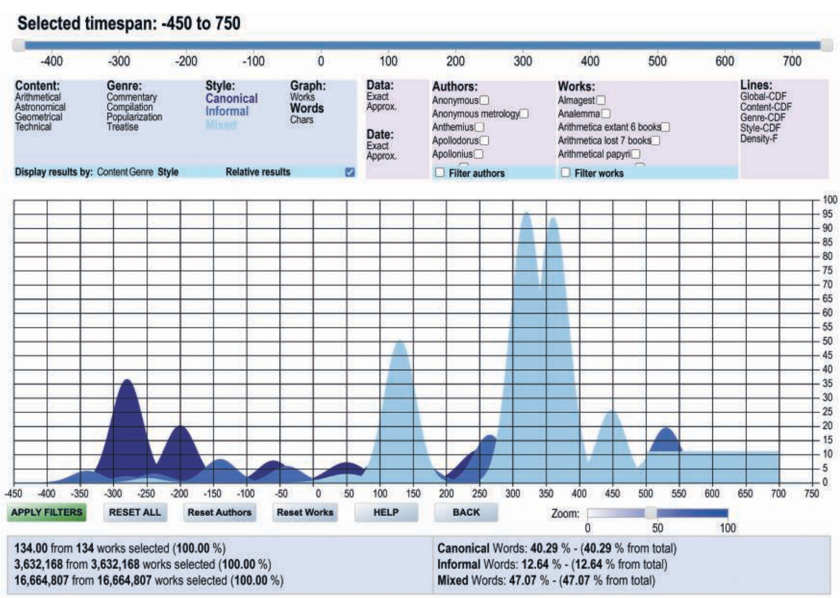


Figure 4. *Interactive timeline graph representing the GMC*



9. Final assessment; aims and limitations of our approach

We comment on some features of the GMC distribution on the time axis. The distribution of the number of works is remarkably uniform. Predictably, the end of the time interval is less productive. The more refined GMC token distribution in terms of words highlights a number of dynamics within the GMC. Some features are obvious: the seven "bumps" are, in succession, Euclid, Apollonius, Ptolemy, Pappus, Theon, Proclus, and the Alexandrian Neoplatonic school as represented by Eutocius and John Philoponus. Archimedes, the most celebrated mathematician of antiquity, is overshadowed by Euclid and Apollonius. Long gaps[43] characterize two eras of transition: from the Hellenistic period to Roman rule, and the crisis of the third century. The size of single works increases in late antiquity: while the outputs of Euclid and Ptolemy are comparable, the commentators Pappus and Theon make the largest contributions to the GMC.[44] There is very little time overlap between authors: the GMC as it is handed down to us is a tradition comprising separate contributions spread over a millennium; Greek science was more a literary tradition – which could sleep for centuries before being revived – than a body of knowledge with any socio-economic function. The GMC is quantitatively dominated by three sub-traditions: Euclid and his legacy, Apollonius and his legacy, and, almost overwhelmingly, Ptolemy and his legacy.

As for genre and style, formal treatises were gradually replaced by secondary outputs such as commentaries and scholia; use of the demonstrative code declined, insofar as partly supplanted by the algorithmic code, and a mixed style characterized by register-crossing became widespread and systematic.

In terms of content, the GMC turns out to be less proof-and-geometry centred than usually believed. Greek geometry virtually disappeared with Apollonius (only 20% of all geometric treatises were written after him, a time interval representing 80% of the entire timeline); the revival of interest in late antiquity was triggered by scholastic activity. Geometry was progressively replaced by arithmetic and mathematical astronomy, which attracted talented people at a time of massive interest in astrology during the Greco-Roman period.[45] Astronomical commentaries were written to expound Ptolemy's system and to explain how to use his astronomical tables (less than 20% of all astronomical commentaries were written before 300 CE, which amounts

---

43. The first gap is widened by a selection effect. For instance, most of Hipparchus' output was replaced by the more advanced models elaborated in Ptolemy's *Almagest*.

44. This is not surprising: the commentary on Aristotle's *Physics* authored by Simplicius (a schoolmate of Eutocius and Philoponus) is the largest work of the entire Greek literature.

45. See A. Jones, 1994.

to 65% of the entire timeline). Astronomy spawned the new research field of logistic.

The feature just outlined, already mentioned at the beginning of our study, is the historical point we wish to make. We have identified a fairly comprehensive set of Greek mathematical texts based on a factually-rooted definition, for inclusion in the GMC. The GMC so identified mainly comprises texts that are not strictly demonstrative and do not have the structure of a mathematical treatise. However, thanks to the mechanism of exegetic chains, such texts are closely interconnected with works that adopt the demonstrative code and that are set out as formal treatises. This gives quantitative substance to the claim that, in Greek antiquity, demonstrative and non-demonstrative mathematics formed one and the same universe of discourse.

Our study also aims to make a methodological point. Our approach to the tagging, normalization, and processing of data in order to reveal its main quantitative features involved the creation of dedicated software to produce a specific mode of display. As we are interested in temporal evolution – that is, a process indexed by a continuous argument – no static table, histogram, pie or line chart is able to faithfully represent the data. As the GMC is categorized by several parameters, a specific display mode is needed to represent them all at once and in an economical way. More important, however, is the way we have tagged the database in order to embed historical information so as to yield an output on the time segment; this is done by means of the GMC token distribution, which encodes information about when and how a GMC token has come into existence.

This leads us to the limitations of our approach. The main question is whether the GMC is representative of the "real thing" – namely, the entire Greek mathematical output as it has actually existed – or not. In a sense, this is an idle question: one cannot process an empty database, so the contribution of lost works to a quantitative analysis of any well-defined literary corpus is nil. After all, any overall assessment of the Greek mathematical output – take, for instance, Heath's well-known synthesis, which *nolens volens* is still the reference textbook – is grounded on extant works. What is interesting, however, is the means by which we can estimate what has been lost. Broadly speaking, there are three ways in which a work can be lost. A work can be lost but we may have enough information to estimate its size; this scarcely represented, yet non-empty category – which crucially includes most treatises of Apollonius' – was the focus of our attention in Section 5. A work can be lost but we know that it existed: we may have its author and its title, or its author and a definite description that may not coincide with the title, or its author and something that we may take to be its description (references to authorless works are extremely rare), the latter subcategory raising the problem of whether similar descriptions do refer to the same work or not. Finally, a work can be totally lost, in the sense that we are unaware of its existence. This last category can only be the subject of an informed guess; the second category may undergo a quantitative analysis, which also provides us with the grounds for the just-mentioned guess. Both employ citations in other works as their database. Let us see how.

The only attempt at estimating the overall number of Greek mathematicians (from an author-centred perspective) is that of R. Netz. His argument runs as follows.[46] Let us take a set of N elements and make random and independent choices from it. In the first choice, we select n elements; in the second, we select k elements. It turns out that the two selections share r elements. As the choices are assumed to be independent and random, the ratio of the number of shared elements r to the number of elements in either one of the selected subsets, say n, is the same as the ratio of the number of elements in the other subset, namely, k, to N. This gives $N = (nk)/r$. In our case, N is the "absolute number of Greek mathematicians active in antiquity", n and k are the number of such mathematicians cited in specific sources. Netz eliminates seven names of mathematicians that "must" be cited, such as Euclid. He then takes Pappus, Proclus, Eutocius, and the manuscript tradition as specific sources. There are six possible pairings, and they give the following values for N:[47] 130, 104, 135, 101, 303, 89. "As a safer guess" – and, we may add, for the excellent reason that Netz had just stated that there are "144 individuals about whom we can make a *guess* that they may have been mathematicians"[48] – Netz takes 300 to be the absolute number of Greek mathematicians active in antiquity yielded by this estimate, which he subsequently increases to 1,000 to be on the safe side.

B. Vitrac has made a careful analysis of the citations of scientific works in scientific and non-scientific texts.[49] His analysis shows that extant works correspond to about 30% of the works referred to, the latter amounting to 380–440 works distributed among about 150 authors.[50] If we look closely at his data, we see that Thales, Democritus, Plato's pupil Philip of Medma,

46. See R. Netz, 1999, n. 68 on pp. 282-283.

47. The numbers refer to the natural sequence of pairings of the four-item list provided: 130 refers to the Pappus-Proclus pairing; 104 to the Pappus-Eutocius pairing, etc.

48. Their list is given in R. Netz, 1997, pp. 6-9. Note that in this publication, the list is simply presented as a "catalogue of mathematicians", with no double modal attenuation, no reference to any "guess", no italics. As citations appear to follow Pareto's rank/frequency law (see M. E. J. Newman, 2006), it is quite obvious that Netz' random model is incorrect; moreover, eliminating the most cited mathematicians introduces a bias instead of eliminating it. Thanks to Pareto's law, an estimate of the overall number of ancient mathematicians can be provided starting from the number of cited ones, as we shall show in a forthcoming study.

49. See the discussion in B. Vitrac, 2021, pp. 13-21, 37-42, 58 and Appendices IV and V. Note that Vitrac includes citations found in the *Fihrist*, a late tenth-century catalogue of literature in Arabic.

50. The range 380–440 comes from the above-mentioned uncertainties inherent in the definition of the third subcategory of lost works by known authors.

Eudoxus, Hipparchus, and Ptolemy account for the lion's share of lost works mentioned in non-technical sources; technical sources offer a less spiky distribution, giving prominence to Archytas, Archimedes (mostly mechanical works), Eratosthenes, Apollonius, and Hipparchus. The overall list of lost works (some of which are included in the GMC, however) highlights the complete loss of pre-Euclidean mathematics, which mostly lies before the lower bound of our timeline, but confirms that the distribution of the GMC is not seriously affected by gaps that were originally filled by "bumps" – the only obvious loss of this kind is represented by Hipparchus' writings, which were superseded by those of Ptolemy.

Bibliography

Data

Acerbi, Fabio & Masià, Ramon, "The Greek Mathematical Corpus: a Quantitative Appraisal (database)", dataset (CC-BY-NC-4.0), Nakala, 2022. DOI: https://doi.org/10.34847/nkl.22517tj9

Printed sources

Acerbi, Fabio (ed.), *Diofanto*, *De polygonis numeris*, Pisa and Roma, Fabrizio Serra Editore, 2011b.

Acerbi, Fabio & Vitrac, Bernard (eds.), *Héron*, *Metrica*, Pisa and Roma, Fabrizio Serra Editore, 2014.

Bagnall, Roger Shaler & Jones, Alexander (eds.), *Mathematics, Metrology, and Model Contracts. A Codex from Late Antique Business Education (*P.Math.*)*, New York, New York University Press, 2019.

Heiberg, Johan Ludvig & Menge, Heinrich (eds.), *Euclidis opera omnia*, 8 vol., Leipzig, B. G. Teubner, 1883–1916.

Jones, Alexander (ed.), *Pappus of Alexandria, Book 7 of the* Collection, 2 vol., New York, Springer, 1986.

Nix, Ludwig, Schmidt, Wilhelm, Schöne, Hermann & Heiberg, Johan Ludvig (eds.), *Heronis Alexandrini opera quae supersunt omnia*, 5 vol., Leipzig, B. G. Teubner, 1899–1914.

Secondary sources

Acerbi, Fabio, *Il silenzio delle sirene. La matematica greca antica*, Roma, Carocci, 2010.

—, "Pappus, Aristote et le τόπος ἀναλυόμενος", *Revue des études grecques*, vol. 124, 2011a, pp. 93-113.

—, "Types, Function, and Organization of the Collections of Scholia to the Greek Mathematical Treatises", *Trends in Classics*, vol. 6, 2014, pp. 115-169.

—, "Byzantine Recensions of Greek Mathematical and Astronomical Texts: A Survey", *Estudios bizantinos*, vol. 4, 2016, pp. 133-213.

Acerbi, Fabio, "Byzantine *Rechenbücher*: An Overview with an Edition of *Anonymi* L and J", *Jahrbuch der Österreichischen Byzantinistik*, vol. 69, 2019, pp. 1-57.

—, "Arithmetic and Logistic, Geometry and Metrology, Harmonic Theory, Optics and Mechanics", in Stavros Lazaris (ed.), *A Companion to Byzantine Science*, Leiden, Brill, 2020, pp. 105-159.

—, *The Logical Syntax of Greek Mathematics*, New York and Heidelberg, Springer, 2021a.

—, "Byzantine Easter Computi: An Overview with an Edition of *Anonymus* 892", *Jahrbuch der Österreichischen Byzantinistik*, vol. 71, 2021b, pp. 1-62.

Acerbi, Fabio & Gioffreda, Anna, "Manoscritti scientifici della prima età paleologa in scrittura arcaizzante", *Scripta*, vol. 12, 2019, pp. 9-52.

Cover, Thomas M. & Thomas, Joy A., *Elements of Information Theory*, 2nd ed., London, Wiley, 2006.

Cuomo, Serafina, *Pappus of Alexandria and the Mathematics of Late Antiquity*, Cambridge, Cambridge University Press, 2000.

—, *Ancient Mathematics*, London and New York, Routledge, 2001.

Gillispie, Charles Coulston (ed.), *Dictionary of Scientific Biography*, 16 vol., New York, Ch. Scribner's Sons, 1970–1981.

Goulet, Richard, "Ancient Philosophers: A First Statistical Survey", in Michael Chase, Stephen R. L. Clark & Michael McGhee (eds.), *Philosophy as a Way of Life*, London, Wiley, 2013, pp. 10-39.

— (ed.), *Dictionnaire des philosophes antiques*, 7 vol., Paris, CNRS Éditions, 1994–2018.

Jones, Alexander, "The Place of Astronomy in Roman Egypt", in Timothy D. Barnes (ed.), *The Sciences in Greco-Roman Society*, Edmonton, Academic Printing and Publishing, 1994, pp. 25-51.

Masià, Ramon, "La llengua d'Arquimedes en De Sphaera et Cylindro", PhD thesis in mathematics, Universitat de Barcelona, 2012.

—, "On Dating Hero of Alexandria", *Archive for History of Exact Sciences*, vol. 69, 2015, pp. 231-255.

Netz, Reviel, "Classical Mathematics in the Classical Mediterranean", *Mediterranean Historical Review*, vol. 12, 1997, pp. 1-24.

—, *The Shaping of Deduction in Greek Mathematics*, Cambridge, Cambridge University Press, 1999.

—, "Greek Mathematicians: A Group Picture", in Christopher J. Tuplin & Tracey E. Rihll (eds.), *Science and Mathematics in Ancient Greek Culture*, Oxford, Oxford University Press, 2002, pp. 196-216.

Newman, Mark E. J., "Power Laws, Pareto Distributions and Zipf's Law", *Contemporary Physics*, vol. 46, 2006, pp. 323-351.

Roelli, Philipp, "What Kind of Latin Was Used in Scientific Communication in the Early Middle Ages?", in *La conoscenza scientifica nell'Alto Medioevo. Atti della LXVII Settimana di Studio, Spoleto, 25 aprile – 1 maggio 2019*, Spoleto, Centro Italiano di Studi sull'Alto Medioevo, 2020, pp. 1021-1058.

—, *Latin as the Language of Science and Learning*, Berlin, De Gruyter, 2021.

Vitrac, Bernard, "Les classifications des sciences mathématiques en Grèce ancienne", *Archives de philosophie*, vol. 68, 2005, pp. 269-301.

—, "Quand ? Comment ? Pourquoi les textes mathématiques grecs sont-ils parvenus en Occident ?", online document, 2021.
URL: https://hal.archives-ouvertes.fr/hal-03328383

Wissowa, Georg (ed.), *Paulys Real-Encyclopädie der classischen Altertumswissenschaft*, 1st series, 24 vol. in 43 tomes, Stuttgart 1894–1963, 2nd series, 10 vol. in 19 tomes, Stuttgart, 1914–1972.