The Greek Mathematical Corpus: A Quantitative Appraisal

Fabio Acerbi, Ramon Masià

Le corpus mathématique grec : une approche quantitative
The Greek Mathematical Corpus: A Quantitative Appraisal

FABIO ACERBI, CNRS, UMR8167 Orient et Méditerranée, équipe "Monde Byzantin", Paris
fabacerbi@gmail.com
RAMON MASIÀ, Magna Societas Catalanorum, passatge Sant Josep 10, 08340 Vilassar de Mar (Barcelona)
rmasia@gmail.com

résumé. 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.

abstract. The paper assesses the Greek mathematical corpus as a whole by using quantitative methods, whose methodological import is also discussed. A number of dynamics within the corpus are also outlined; they corroborate the view that, in Greek antiquity, demonstrative and non-demonstrative mathematics made one and the same universe of discourse.

mots-clés. mathématiques grecques, méthodes quantitatives, manuscrits, distribution de Gauss, distribution de Pareto
keywords. Greek mathematics, quantitative methods, manuscripts, Gauss distribution, Pareto distribution

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 as to content and style[1]. This universe of discourse is addressed to a readership that is made of disjoint continents: we cannot assume that a reader of Apollonius' celebrated treatise on conic sections could have belonged to the same social and cultural milieu as any user of Theon of Alexandria's

* We are grateful to Bernard Vitrac for a critical reading and for some suggestions.

[1] This universe of discourse is optimized: mathematical contents that are expressed within a stylistic code cannot be satisfactorily expressed by using a different code. The stylistic codes adopted in Greek mathematics are the demonstrative code, the procedural code, and the algorithmic code; See F. Acerbi, 2021a, sect. 1.1–3, for a detailed description of each of them.

"little" commentary on Ptolemy's astronomical tables[2]. However, different readerships do not necessarily mean a different "level" of the involved pieces of mathematics. These value-laden categories are not suitable 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 by 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 made one and the same universe of discourse. Of course, this perspective is not new[3]. We may, however, claim originality in employing 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, a research theme we shall pursue in the future[4]: it would be rash to decide a priori that some segments of Greek mathematics display a lexicon more suitable to a 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 the use of quantitative methods. These problems originate in an obvious feature of the subject-matter: the GMC is temporally, materially, and even definitionally coarse-grained. First, we are often unable to locate a Greek mathematician in time better than over a span of some centuries[5]. Second, some mathematical treatises are lost, or partly lost, in Greek[6]. Third, some mathematical works are not authorial undertakings but collections of disparate texts that got stabilized before the date of copying of their earliest extant manuscript witnesses[7]; for this reason, these works have been handed down to us as a unitary whole—worse than this, some of these collections are just philological artefacts assembled in modern times[8]. Finally, and in fact primarily, one has to decide what counts for a work to be a mathematical work. In other words, it is not immediately obvious what is a "(mathematical) work" to be included in the GMC and how can it be categorized, what is an "author" of any such "work", what temporal location must be assigned to any such "author". Apart from problems of coarse-graining, applying quantitative methods to studying the GMC entails introducing methods and tools that have to

[2] Of course, an 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 and S. Cuomo, 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, p. 59–73 (Hero of Alexandria), and in F. Acerbi, 2021a, p. 28–36 and 157–159 (Euclid). Corpus- and computational linguistics tools are applied to Medieval scientific texts written in Latin in Ph. Roelli, 2021; see also the synthesis in Ph. Roelli, 2020.
[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, p. 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 (*Cutting off of a ratio*), only four Books of his *Conics* are extant in Greek, three more we read only in Arabic translation, whereas Book VIII is lost. As we shall see in Sect. 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.

fit uniformly a database that contains fairly inhomogeneous items; for this reason, we shall adhere to a terminology uncommitted to considering these items as simply "works" written by a single "author".

These issues will be addressed in Sects. 2–6. We first discuss the several ways the GMC can be defined; our definition assigns a central role to works, not to authors (Sects. 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 the GMC (Sect. 4). A series of relevant pieces of information—like authorship (if any), temporal location, and size in words—are linked to these itemizers (Sects. 4–6). It is not said that a GMC-token is a work composed 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 in our perspective, is addressed in Sect. 7: how to categorize the GMC-tokens. We shall use a tripartite scheme: contents, genre, and style. Each of these categories is subdivided into suitable subcategories. In this way, the natural dynamics of the GMC on the time-axis will acquire further dimensions. Sects. 7 and 8 will present the results of our investigation in the form of charts and plots; these are followed, in Sect. 9, by a short assessment of methods and of 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 of them is strictly author-centered and hinges on the consensus of modern scholarship: a Greek mathematical text is any work of any ancient Greek author recorded in such a reference work as the *Dictionary of Scientific Biography*, or, alternatively, any work of anyone identified as a mathematician in the Pauly-Wissowa encyclopedic 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 to which we might hesitate to assign any mathematical content[10]. Moreover, this criterion just shifts the definitional problem to a higher category, namely, whether we have to include texts that belong to kinds like "astronomy" (certainly we have to), "engineering" (almost certainly we have to), or "astrology" (it is open to debate whether we have to) in the GMC.

A similar, author-centered, criterion points to the social role as recognized in ancient sources; let us quote the editor of the most important lexicon of Greek philosophers: "The main criterion [*scil. for assigning an entry of the lexicon 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

[9] These are Ch. C. Gillispie, 1970–81 and G. Wissowa *et al.*, 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 *Tetrabiblos*.

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 be eager to use for a mathematical activity: for instance, a "mathematician" may be someone engaged in doing astrology[12]. Second, and conversely, mathematical activities were practised by authors who were never called "mathematicians", simply because their main activities were categorized in a different way in antiquity: think of Hero of Alexandria, the "engineer"[13], who wrote a strictly mathematical commentary on Euclid's *Elements* and such an amazing piece of register-crossing pure mathematics as the *Metrics*[14]. Third, our evidence is sometimes so scanty as not to allow us any cross-checking. Fourth, sources attest to a sustained teaching activity of mathematics starting only from the beginning of the 4th century CE; attempts at locating Euclid, or any other mathematician prior to that period, in a teaching milieu are only grounded on an unwarranted back-projection of the modern organization of academic work[15]. Thus, it is clear that the social-role criterion for delimiting the GMC forces us to make a priori choices. For instance, consider the most natural ancient source for identifying Greek mathematicians: Pappus' *Collection*[16]. It is obvious that Apollonius—one of the greatest mathematicians of all times—must belong to the GMC, and in fact a lot of his mathematics is commented on by Pappus in the *Collection*; yet the only epithet Apollonius deserves in Pappus' work is "from Perge" (the town where he was born). Again, Pappus cites Plato in such a way that we should include Plato in the GMC[17], and on the other hand he does not cite Diophantus. Taking the set-theoretical union of such sources would not solve the problem, for Pappus' case just mentioned shows that we would be forced to include in the GMC authors whom no sensible scholar would rank as mathematicians.

---

11 R. Goulet, 2013, p. 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 to (arbitrary) choices in its third disjoint, and in particular thanks to the presence of the adverb "unmistakably". The lexicon is the *Dictionnaire des philosophes antiques* (R. Goulet, 1994–2018), and describes nearly 2500 philosophers, among which Ptolemy (for he wrote a short treatise on epistemology), Euclid (who is not reported to have written anything philosophical), and Hero of Alexandria (the same as Euclid).

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 that, more generally, designated anyone involved in investigations whose character was more or less markedly 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). Let us recall that the organization of scientific disciplines currently adopted, mainly for teaching purposes, in the Greco-Roman world was the so-called "quadrivium"; it comprises arithmetic, harmonic theory, geometry, astronomy (the ordering criterion is by decreasing abstraction). The generic name of each of these disciplines was μάθημα, so that a μαθηματικός was anyone engaged in any of these disciplines. See 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, sects. 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 take it for granted that, for instance, Aristotle's school worked exactly 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 should conclude that he was a top-rank mathematician who thoroughly studied the five regular polyhedra.

A third criterion holds that mathematics is what mathematicians do, and vice versa, where the meaning and the extension of the undefined term (namely, either "mathematics" or "mathematician") is settled 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 to know what it is that "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 what mathematics is by anchoring our choice to some piece of the historical record. Our criterion takes manuscript tradition as a suitable anchoring, delegates the choice to the scholars who have been responsible for handing Greek mathematics down to us, adds a ban, and is stated as follows: any structured macro-piece of discourse whose author figures (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 ban 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 producing new books that contain a representative segment of the Greek literary production; the new books were produced by copying from available manuscripts. This is the way most of ancient Greek literature has come to us[19]; the editions in which we read ancient Greek literature are based on these manuscripts.

In the Palaiologan period, huge mathematical encyclopaedias were composed by assembling primary sources. The encyclopaedias in question are made of a two-volume set, Par. gr. 2342 (*Diktyon* 51974) + Vat. gr. 198 (*Diktyon* 66829), and of two single, imposing manuscripts, Vat. gr. 191 (*Diktyon*

[18] R. Netz, 1997, p. 4, and R. Netz, 2002, p. 197, the latter explicitly reacting to the approach in S. Cuomo, 2000. See also the discussion in R. Netz, 1999, p. 277–278.
[19] As a (small) part of ancient Greek literature has only been transmitted by papyri, the predeterminer "most of" is necessary.

66822) and Vat. gr. 192 (*Diktyon* 66823). The two-volume set was penned about 1360–70 by the scholar and copyist Malachias; it contains a fully-fledged *quadrivium*. The second manuscript was written by a number of unknown copyists about 1270–90 and was assembled about 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 whole manuscript tradition of Greek scientific works. These encyclopaedias do not count as decisive manuscripts witnesses of the works they contain: there usually are 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 felt themselves in strict continuity with ancient Greek thought, took to be the GMC.

Our criterion is a disjunction with three disjuncts, plus a ban. The first disjunct selects the contents of the four manuscripts. In particular, these manuscripts contain the entire corpus of harmonic theory; this is one of the four disciplines of the *quadrivium*, hence a perfectly legitimate mathematical discipline on ancient standards. The second and the third disjunct, along with the recursive qualification, enlarge this core of GMC-texts to "exegetic chains". In this way, we take account of the fact that Greek mathematics is first and foremost a literary genre, which constitutes itself as a tradition. Accordingly, works of authors like Menelaus, Hero of Alexandria, or Anthemius of Tralles, and the entire poliorcetic corpus, which do not figure in any of the above manuscripts[22], are GMC-texts because they elaborate on GMC-texts (second disjunct of the criterion, with the recursive qualification). Likewise, works of Archimedes are not included in any of the above manuscripts, but his works are GMC-texts because an author included in Vat. gr. 191 and Par. gr. 2342, Theodosius, 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); this entails including the entire set of mathematical papyri[23]. As for the ban, this makes the manuscript-based sufficient condition a criterion, by adding a necessary condition. The ban allows eliminating astrology, and, for instance, all works of an author that ancient categorizations would have regarded as philosophical, as most of Proclus' commentaries and some of Iamblichus' works[24].

Our criterion is partly author-centered, but this is the only approach that allows us to reckon the entire production of, for instance, Archimedes and Hero of Alexandria among the GMC-texts. Most

[20] On these four manuscripts, see F. Acerbi, 2016, *passim* but in particular p. 154–160, 189–190 (Par. gr. 2342 and Vat. gr. 198), and F. Acerbi & A. Gioffreda, 2019, p. 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 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 more or less invasive revision of it, usually carried out in coincidence with a new edition: in most manuscripts, we read Euclid's *Elements* in the light recension authored by the middle 4th-century scholar Theon of Alexandria; we have access to Apollonius' *Conics* only through the heavy recension of the early 6th-century Neoplatonic philosopher Eutocius. See F. Acerbi, 2016 for a complete survey of Byzantine recensions.

[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' *Introductio arithmetica* among the GMC-texts.

importantly, using authorship as a bridge between classes of GMC-texts allows us to reckon lost works among the GMC-texts. This is crucial because most of Apollonius' production is lost, but we are nevertheless able to estimate its size (see Sect. 5 below).

4. *Defining a GMC-Token*

A GMC-token that figures in our database is the form a (self-contained part of a)[25] GMC-text has in the modern critical edition that serves as a reference[26]. As for requiring that a GMC-token is the representative of a self-contained piece of mathematical discourse, the boundaries of self-containment are usually obvious: most of our GMC-tokens are canonical kinds of works, which we may categorize within the literary-theoretical genre "treatise"[27]. By definition, a treatise is written by an author, even if his or her name may be unknown to us, and even if we do not know much more than his or her name about this author. There are, however, GMC-tokens—and they may be significant and fairly-sized GMC-tokens—that are not authorial undertakings. These are the collections of problems that make 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 in the course of 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 have provided a shape, and boundaries, to these collections, making them such self-contained and structured macro-pieces of mathematical discourse as are found in 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 we shall embed GMC-tokens in the time-axis will neutralize this feature.

In general, it will be noted that our criterion allows including in the GMC a good deal of mathematics that is considered "peripheral" by current scholarship. On the other hand, it is obvious that our criterion is worded so as to allow including in the GMC all authors and works we reasonably regard as "mathematics". The main adjustment lies in allowing "exegetic chains", so as to include in the GMC

---

[25] See also Sect. 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, 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 entails 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 fully-fledged commentaries, but we shall treat all scholia as if they were independent pieces of writing. Outside the GMC but still in a Greek-reading area, the category of non-authorial works includes the so-called Byzantine *Rechenbücher* (see F. Acerbi, 2019) and the Easter Computi (see F. Acerbi, 2021b).

works that do not figure in the selected manuscripts. We might have disposed of "exegetic chains" in our criterion by enlarging 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. Yet, it is a welcome fact that only a handful of manuscripts should be added to the above-mentioned encyclopaedias in order to get the entire GMC corpus as it is identified according to the sole first disjunct of our criterion[29]. Finally, the fact that the formulation of our criterion is not *so* contrived shows that the Byzantines have had a clear perception that Greek mathematics constitutes a tradition: we must not forget that they have already contributed in a decisive way to deciding what kind of Greek mathematics we had to read.

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 shall take the dates of birth and death (or, faute de mieux, the *post quem* and *ante quem* dates) currently assumed by modern historiography for granted, 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 shall restrict the lifespan of a given author to 80 or 100 years centered on the middle point of the currently assumed lifespan, even if the possible lifespan is wider. For example, we assign Archimedes the lifespan [–287,–212] because we know that he died in 212 BCE, when the Romans seized Syracuse, and a late source claims that he lived 75 years[31]. We assign the lifespan [–240,–160] to Apollonius, [–320,–240] to Euclid, [90,170] to Ptolemy. All these dating we deem "exact", because our documents allow us to determine the lifespan quite accurately; 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 datings "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 grounds 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 the interval [200,700]; the scholia to the *Elements* are located in the interval [500,700], even if most of them can be argued to be later.

---

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

[30] A survey of the several ways Greek mathematicians are dated is in F. Acerbi, 2010, p. 79–86. Astronomers can usually be dated accurately because they may report observations of celestial phenomena that can be identified with certainty. It goes without saying that we are unable to assign a "publication date" to any GMC-token; even the relative chronology of the works of such authors as Archimedes is uncertain. Moreover, there are GMC-tokens (a well-known example is Apollonius' *Conics*) whose parts were composed in different times and "published" by sending these parts to different addressees (in ancient terminology, the "parts" are the "books" of which a treatise is made)

[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.

Our estimates entail that the segment of the time-axis in which the GMC is embedded is the interval [–380,700], that is, more than one thousand years.

5. *Estimating the Size of a GMC-Token*

Since a GMC-token is a self-contained work, the natural size we may attach to it is the number of words it contains. This number is calculated in a uniform way, as follows. If a GMC-token is extant in its entirety, we use the number of word-tokens indicated in the *Thesaurus Linguae Graecae* online as the base estimate of its size[32]; to simplify matters, this number includes the segments of the text 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 constitute a very small percentage of a work's 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 the works that are totally lost or that are 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 made of several treatises and now almost entirely lost. For some works of the analytical corpus, in which what is solved is a single problem that branches off in a multitude of cases, Pappus adopts the counting unit "diagram". For instance, the two books of Apollonius' *Plane loci* comprise 15 loci (each 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

[32] This has the consequence that we must exclude astronomical tables from the GMC. The website of the *Thesaurus Linguae Graecae* is 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 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' *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] That is, number of lost books : number of extant books = number of words contained in the lost books : number of words contained in the extant books.

[35] A case in point is Theon of Alexandria's commentary on the *Almagest*, whose 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 only be read in a Byzantine recension, Books VIII–X and XII–XIII can only be read—certainly mutilated and possibly in a recension again—in the 13th century manuscript Vat. gr. 1087. Despite these problems, it is clear that the size of the first books was larger than the size of the last books: basics such as how to divide two sexagesimal numbers have to be explained at the beginning of the commentary.

[36] Another matter is to estimate how many Greek mathematical texts are radically lost, in the sense that we do not have any piece of information about them; for instance, we do have a piece of information on Demetrius' of Alexandria work of higher-order curves: it is just the title "Linear investigations", mentioned, along with its author, by Pappus in *Coll*. IV.58. See Sect. 9 for a discussion of this issue.

words vs. number of propositions of the works for which these two quantities are known (Chart 1. below is also found at https://bit.ly/3t7HLMq, attached to the database)[37].

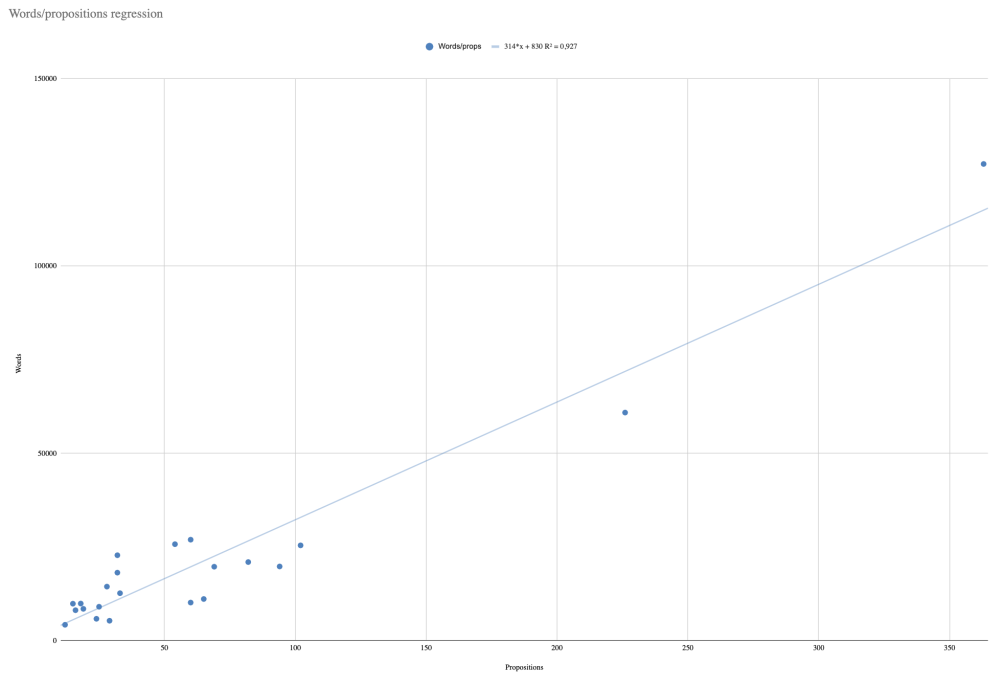

Chart 1. Linear regression words vs. propositions

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 of the author(s) of a GMC-token or if this author does not exist, we stipulate that the intended GMC-token has been composed uniformly during the relevant time segment. Thus, the output of an anonymous author or the activity that has originated a collection of mathematical items (as scholia and metrological works) is adapted 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 name of the author is known, this item is adapted to the relevant time segment by means of a suitable Gaussian distribution centered on the midpoint of the interval (in ancient terms, this coincides with the *akmē* of the author); the area under the distribution is

[37] Using linear regression is more robust than using the average number of words in a proposition as the scaling factor. The reason is that 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 any of them.

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 $\frac{1}{12}$ of his lifespan; if an author cannot be dated exactly, the standard deviation σ of the Gaussian distribution is $\frac{1}{6}$ of the author's lifespan. This is a simple numerical encoding of the fact that there is a loss of information with respect to the case in which an author can be dated exactly.

For example, we assume that Archimedes' production is centered 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 during the timespan of nearly 1000 years that supports the GMC, and one of our aims is to exhibit such partial evolutions, showing at the same time that they are related to each other.

The GMC has been 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", which is the 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 (Apollonius' *Conica*; this means that a well-defined subject is treated, and the work is intended for specialists), popularization

---

[38] The choice of the two distributions is made 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 Th. 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 as far as the very end of his life.

(Nicomachus' *Introduction to arithmetics*), commentaries (they expressly set elucidating formal treatises as their goal, 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 that is 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 arithmetics*, where the discursive frame can be argued to carry the focus).

Assigning some of these categories is easy and uncontroversial; assigning others may not be. A fine-tuning of the above categorizations motivates our decision of splitting some treatises into several GMC-tokens. Some treatises may be written in different styles, as for instance Hero's *Metrica*, which mixes the demonstrative and algorithmic codes. The reader will find all our choices in the database uploaded at https://bit.ly/3t7HLMq. As shown in the image below (Chart 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; the ratio characters/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.

						prop/words regression					prop/chars regression							Graphs		
						314*x + 830			314	830	1510*x + -3199	1510	-3199							
da	**da**	**date**	**author**	**chunk**	**work**	**#w**	**#**	**Word**	**Chars**		**data**	**content**	**genre**	**style**	**nu**	**nun**	**C**	**μ**	**σ**	**gauss?**
100	700	approximate	Anonymous	arithmetical papyri	arithmetical pap	1,00	NA	10000	45000	4,5	approximate	arithmetical	compilation	canonical				400	173,21	NO
100	700	approximate	Anonymous	astronomical papyri	astronomical pa	1,00	NA	10000	45000	4,5	approximate	astronomical	compilation	canonical				400	173,21	NO
200	700	approximate	Anonymous met	Geodesy	Geodesy	1,00	1	4831	23930	4,95	exact	geometrical	compilation	canonical				450	144,34	NO
200	700	approximate	Anonymous met	Geometrica	Geometrica	1,00	NA	32923	155171	4,71	exact	geometrical	compilation	canonical				450	144,34	NO
200	700	approximate	Anonymous met	Liber geeponicus	Liber geeponicu	1,00	1	6534	31824	4,87	exact	geometrical	compilation	canonical				450	144,34	NO
100	300	approximate	Didymus	Measures of marble	Measures of ma	1,00	1	1344	6388	4,75	exact	geometrical	compilation	canonical				200	33,33	SI
100	300	approximate	Diophanes	metrological compil	metrological con	1,00	1	3211	15629	4,87	exact	geometrical	compilation	canonical				200	33,33	SI
100	700	approximate	Anonymous	metrological papyri	metrological pap	1,00	NA	10000	45000	4,5	approximate	geometrical	compilation	canonical				400	173,21	NO
200	700	approximate	Anonymous met	On measures	On measures	1,00	1	5027	23847	4,74	exact	geometrical	compilation	canonical				450	144,34	NO
200	700	approximate	Anonymous met	Stereometrica	Stereometrica	1,00	NA	17529	80573	4,60	exact	geometrical	compilation	canonical				450	144,34	NO
-200	-120	exact	Diocles	On burning mirrors	On burning mirr	1,00	1	3970	11901	3,00	approximate	geometrical	treatise	canonical	10			-160	6,67	SI
90	170	exact	Ptolemy	Planisphere	Planisphere	1,00	1	7110	27001	3,80	approximate	astronomical	treatise	canonical	20			130	6,67	SI
-287	-212	exact	Archimedes	Stomachion	Stomachion	1,00	1	7110	27001	3,80	approximate	geometrical	treatise	canonical	20			-249,	6,25	SI
-320	-240	exact	Euclid	Division of figures	Division of figure	1,00	1	10250	42101	4,11	approximate	geometrical	treatise	canonical	30			-280	6,67	SI
0	100	approximate	Hero of Alexand	On mirrors	On mirrors	1,00	1	10250	42101	4,11	approximate	technical	treatise	canonical	30			50	16,67	SI
-240	-160	exact	Apollonius	On contacts	On contacts	1,00	2	19670	87401	4,44	approximate	geometrical	treatise	canonical	60			-200	6,67	SI

Chart 2. A part of the GMC-token database

8. *Charts*

We present a Sunburst interactive chart that represents the GMC (Chart 3). It exhibits the prominent role of astronomy, the extent of Ptolemy's legacy, and the productivity of the two late antique authors Pappus and Theon[40].

[40] This is an evolution of the pie charts found in R. Goulet, 2013 and of the pie chart found in B. Vitrac, 2021, p. 40.

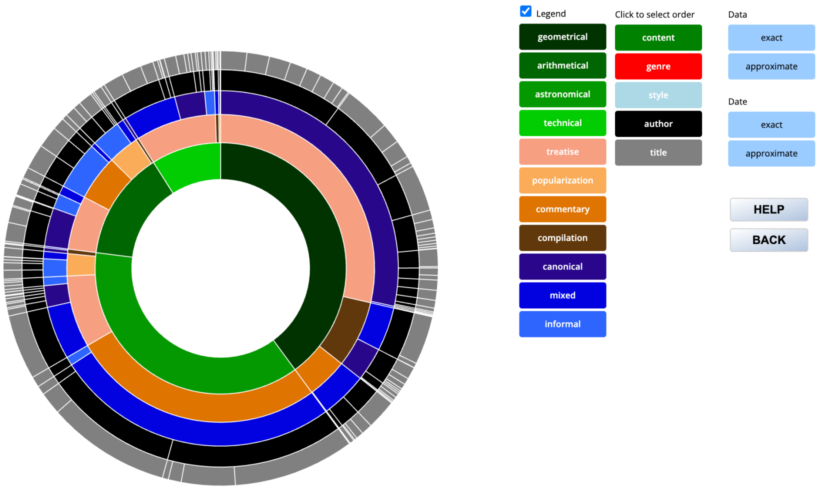

Chart 3. Interactive Sunburst representing the GMC

We also present a Timeline interactive chart that represents the GMC-token distributions in terms of works, words, and characters on the interval [–450,750] (Chart 4). This chart also exhibits all categories established in Sect. 7, and displays the integral of the distribution curves: these integral lines show the cumulative evolution of the indicated parameters.

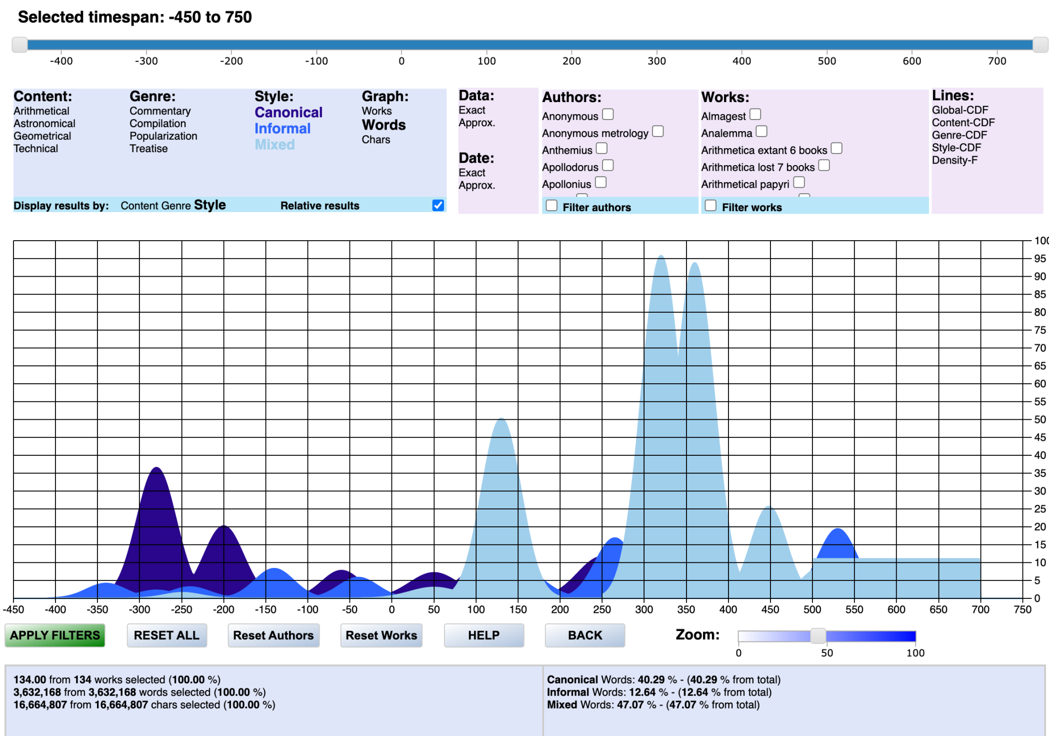

Chart 4. Interactive Timeline representing the GMC

Both interactive charts, which are automatically generated from our database by means of non-trivial dedicated software, are uploaded at https://ramonmf.github.io/GraphAGM/; detailed instructions of use are also provided there.

9. *Final Assessment; Aims and Limitations of our Approach*

We outline some features of the distribution of the GMC on the time axis. The distribution of the number of works exhibits a remarkable uniformity. As is to be expected, 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[41] characterize two eras of transition: the transition from the Hellenistic period to Roman rule and the "crisis" of the third century. The size of single works increases in Late Antiquity: if the productions of Euclid and of Ptolemy are comparable, the commentators Pappus and Theon contribute the largest outputs of the GMC[42]. Temporal coincidence of authors is very infrequent: the GMC as it is handed down to us is a tradition made of punctual contributions that spread over a millennium; Greek science was more a literary tradition—which can sleep for centuries before being revived—than a body of knowledge of any socio-economic use. 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 second-order activities like commentaries and scholia, the demonstrative code was less practised, insofar as partly supplanted by the algorithmic code, and the use of a mixed style characterized by register-crossing became widespread and systematic.

As for contents, we are presented a picture of a less proof-and-geometry centered GMC than usually believed. Greek geometry virtually died with Apollonius (only 20% of all geometric treatises were written after him, a time interval that is 80% of the entire Timeline); the regain of interest in Late Antiquity was triggered by scholastic activity. Geometry was progressively replaced by arithmetic and mathematical astronomy, which attracted talented people on the background of the explosion of astrology in the Greco-Roman period[43]. Astronomical commentaries were written whose goal was to expound Ptolemy's system and the way of use of his astronomical tables (less than 20% of all

[41] The first gap is widened by a selection effect. For instance, most of the production of Hipparchus was replaced by the more advanced models elaborated in Ptolemy's *Almagest*.
[42] 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.
[43] See A. Jones, 1994.

astronomical commentaries were written before 300 CE, which amounts to 65% of the entire Timeline). Astronomy triggered the new research field of logistic.

The feature just outlined is the historical point we wish to make, and we have already mentioned at the beginning of our study. A factually-rooted definition of "Greek mathematical text" has led us to identify a fairly comprehensive set—the GMC—of such items. The GMC so identified is mainly made of texts that are not strictly demonstrative nor have the structure of a mathematical treatise. However, thanks to the mechanism of "exegetic chains", such texts are strictly intertwined 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 made one and the same universe of discourse.

Our study, however, also aims at making a methodological point. This point consists in the way we have tagged, normalized, and processed our data in order to exhibit its main quantitative features. This approach requires conceiving a specific mode of display and writing a dedicated software that implements it. As a temporal evolution is at issue—that is, a process indexed by a continuous argument—no static table, histogram, pie or line chart is suitable faithfully to 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 suitable 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 to 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 that the contribution of lost works to any 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 the extant works. What is interesting, however, is the way we may estimate what has been lost. Broadly speaking, there are three three ways in which a work can be lost. A work can be lost but we may happen to have enough information to estimate its size; this scarcely-represented yet non-empty category—which crucially includes most treatises of Apollonius—has been the focus of our attention in Sect. 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 (authorless works referred to 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. The 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 (thus, the perspective is author-centered) has been put forward by R. Netz. His argument runs as follows[44]. Let us have a set of *N* elements and let us make random and independent choices from it. In the first choice, we select *n* elements; in the second choice, 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 of the selected subset, 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*[45]: 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"[46]—Netz takes 300 to be the absolute number of Greek mathematicians active in antiquity yielded by this estimate, which he subsequently enlarges to 1000 in order to rest on even safer grounds.

A careful analysis of the citations of scientific works in scientific and non-scientific texts has been carried out by B. Vitrac[47]. His analysis shows that the percentage of extant works is about 30% of the works referred to, the latter amounting to 380–440 works distributed among about 150 authors[48]. If we look closely at his data, we see that the lion's share among lost works mentioned in non-technical sources is held by those of Thales, Democritus, Plato's pupil Philip of Medma, Eudoxus, Hipparchus, and Ptolemy; technical sources offer a less peaked distribution, giving anyway 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 screened away by Ptolemy's.

Bibliography

[44] See R. Netz, 1999, n. 68 on p. 282–283.

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

[46] Their list is given in R. Netz, 1997, p. 6–9. We may note that, in this publication, the list is simply presented as a "catalogue of mathematicians": no double modal attenuation, no guess around, 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.

[47] See the discussion in B. Vitrac, 2021, p. 13–21, 37–42, 58 and the Annexes IV and V. Note that Vitrac includes the citations in the *Fihrist*, a late tenth century catalogue of literature in Arabic.

[48] The range 380–440 comes from the above-mentioned uncertainties inherent in the definition of the third subcategory of lost works of known author.

PRINTED SOURCES

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

Acerbi, Fabio & Vitrac, Bernard (eds.), *Héron*, *Metrica*, Pisa – 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 LITERATURE

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

Acerbi, Fabio, "Pappus, Aristote et le τόπος ἀναλυόμενος", *Revue des Études Grecques*, vol. 124, 2011a, p. 93–113.

Acerbi, Fabio, "Types, Function, and Organization of the Collections of Scholia to the Greek Mathematical Treatises", *Trends in Classics*, vol. 6, 2014, p. 115–169.

Acerbi, Fabio, "Byzantine Recensions of Greek Mathematical and Astronomical Texts: A Survey", *Estudios bizantinos*, vol. 4, 2016, p. 133–213.

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

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

Acerbi, Fabio, *The Logical Syntax of Greek Mathematics*, New York – Heidelberg, Springer, 2021a.

Acerbi, Fabio, " Byzantine Easter Computi: An Overview with an Edition of *Anonymus* 892", *Jahrbuch der Österreichischen Byzantinistik*, vol. 71, 2021b.

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

Cuomo, Serafina, *Ancient Mathematics*, London – 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, p. 10–39.

Goulet, Richard (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, p. 25–51.

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

Masià, Ramon, "On Dating Hero of Alexandria", *Archive for History of Exact Sciences*, vol. 69, 2015, p. 231–255.

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

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

Netz, Reviel, "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, p. 196–216.

Newman, Mark E. J., "Power laws, Pareto distributions and Zipf's law", *Contemporary Physics*, vol. 46, 2006, p. 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, p. 1021–1058.

Roelli, Philipp, *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, p. 269–301.

Vitrac, Bernard. "Quand ? Comment ? Pourquoi les textes mathématiques grecs sont-ils parvenus en Occident ?", 2021, hal-03328383.

Wissowa, Georg *et al*. (eds.), *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.