Chapter 11

zbMATH Open as a tool for bibliographical studies

Klaus Hulek, Olaf Teschke

1 Introduction

Evaluations and rankings, be it of individuals or institutions, have become part of academic reality. These evaluations range from career-defining assessments of individuals to worldwide university rankings. Although the methodology of many of these evaluations has often been criticised, they remain ubiquitous with often extraordinary effects. The impact on individual careers, and hence lives, can be decisive. On a more global level these figures not only contribute significantly to the reputation of universities, but also affect the choices of perspective students and staff.

Various parameters are used to evaluate research performance, with bibliometric data playing an important role in (almost) all evaluations. Both, generating these data, as well as interpreting them, constitutes a major challenge. Therefore, it is important to understand the technical aspects, as well as the different parameters and perspectives, that go into bibliometric data.

The aim of this contribution is to show how zbMATH Open data provide insight into how mathematics is being published, with an emphasis to reveal how the genuine specifics of the mathematical literature render traditional subject-blind bibliometric approaches and measures inapplicable. Since most of zbMATH Open data – especially those relevant for bibliometric analysis – are openly available by a CC-BY-SA license [11] through the zbMATH Open REST API[1] [10], the following observations can not just be easily reproduced, but can serve as the basis for further, more sophisticated analysis.

2 Time line of mathematical references

It is a fundamental characteristic of mathematics that a theorem, once proved, remains valid forever. Nevertheless, scientific progress often leads to stronger and more general results which thus supersede earlier work. Hence the question about the relevance of older results, measured by the average time interval between publication and citation, is highly nontrivial. Other disciplines like biology, chemistry, physics, or medicine have recently seen a faster decline in citations [8] of a given paper, indicating that the half-life of publications might be decreasing.

[1] https://api.zbmath.org

With currently almost 50 million references available for a total of 2 million documents, the zbMATH Open citation database constitutes the largest curated citation database for mathematics.

To investigate the reference time line, it is not necessary to match the references to the database, which is only possibly for about 60% — the remaining 40% maybe unpublished work, or outside the scope of zbMATH Open. For this it is sufficient to extract the publication year from the string. This is the basis of the diagram in Figure 1, which shows the development of citation distances over time. It shows the average difference between the publication years of cited works and the publication year of the citing work, depending on the latter. An average was taken across all subject areas and all forms of publication.

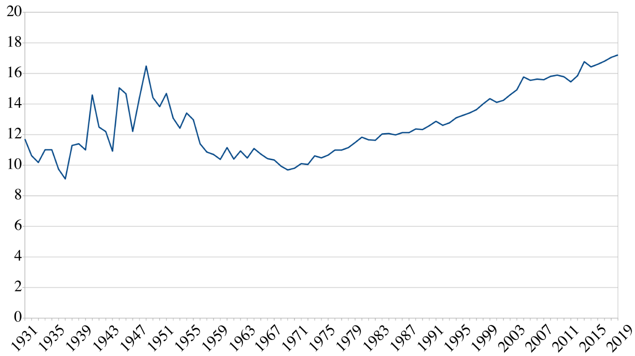

Figure 1. Average time interval between publication and references.

The striking result is that the average age of cited papers has actually grown constantly, and is now almost 18 years. To interpret the graph, various aspects must be taken into account. Reference data are currently available for only around 4% of publications before 1945. Therefore, this part is subject to increased uncertainty and shows a correspondingly erratic course. However, the effects of the two world wars are visible in the graphic, which led to a decrease in the interval between publication and citation in the following decade.

A plausible explanation for this is that publications from the war years, which would normally be cited more widely after a few years, are missing here and, on the other hand, reference is mainly made to more recent literature, especially since many working groups and networks had to be rebuilt. After WWII, this effect dominated until 1968. Since then, the diagram shows a continuous growth of this interval. This period of about 18 years after the end of WWII, before the citation distance starts to grow again, is a further indication that two decades represent a natural lower limit for average citation distances (excluding war effects), at least for the period in which extensive data are available.

3 Is there a half-life of mathematical results?

Vice versa, one can ask also how long a given work is cited. By investigating this with zbMATH Open data, we must keep in mind that, in contrast to the previous section, scope effects come into play – only the indexed citing documents contribute to the data.

The general concept of research impact suggests that research which is cited for a long time after its publication typically represents more significant contributions. However, this approach cannot be used to identify all outstanding publications. Important theorems often become so much part of common knowledge that a reference is no longer given. In other cases, more accessible version or survey articles are cited instead of the original work.

Instead, we consider mainly the question of longevity of references, i.e., the temporal distribution of references for documents published in a fixed year. One would expect that in general the number of citations increases sharply immediately after the publication year, but would show a steady decline. However, it turns out that the growth of the published literature has the strongest influence here, leading effectively to an unlimited growth of references to a set of documents with fixed publication year. Hence, in the diagram shown in Figure 2, the number of references to a given publication year is normalised by the overall number of references for the publication year of the citing documents. Moreover, for a better impression of the structure, the figures are also normalised with respect to the maximum of this figure.

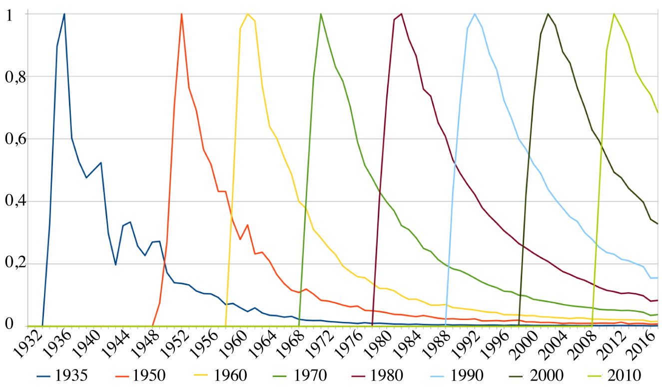

Figure 2. Timeline of relative citations to mathematical papers for different fixed publication years.

The x-axis shows publication years; the y-axis the number of citing articles relative to the overall references in the citing year and the year of maximum citations.

With these normalizations, the figures match more closely the expected shape. However, the relative maximum is usually obtained only after three years (or, taking the results of Section 5 into account, in average five years after it has been posted on the arXiv for core mathematics papers). Moreover, the decline remains relatively smooth, with more than half of the relative citations being generated more than eight years after publication. It should also be noted that the aggregations provide no information on the share of highly cited publications, which, based on the analysis of some samples, appears to grow over time (especially for books). Such an analysis would be beyond the scope of this note, but readers are invited to explore this effect by using zbMATH Open data.

Taking all this into account, it becomes clear that the usually applied short-term bibliometric measures (such as three- or five-year impact factors) miss the crucial part of the relevant citation information. Vice versa, assuming the usual timeframes in scientific careers, there seems no meaningful way to include into decision-making measures which only have a chance to become relevant about a decade after their underlying idea went public.

Another caveat would be that this diagram aggregates publications from all mathematical areas. However, both citation behaviour and publication growth depends heavily on the subject, so it seems natural to take subject specifics into account.

4 One step further: Subject specifics

One might wonder whether it is possible to differentiate this general picture further by taking mathematical subjects into account. Matching citations to zbMATH Open provides MSC information and raises the natural question what the interdependence between mathematical fields and citation networks is.

Figure 3 shows that there is indeed a strong concentration along the diagonal (which means that the bulk of references point to papers with the same MSC), although there obviously exist further cross-references which might be worth investigating in a more detailed analysis.

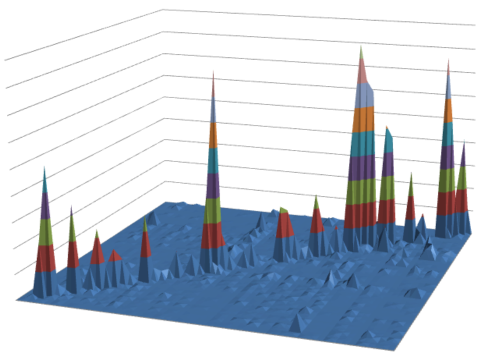

Figure 3. Cross-MSC citation map.

The strong concentration on the diagonal (which is, by the way, an indication that the MSC actually depicts clusters of related work well) can serve as a justification that restricting to area-preserving citations serves well as a first approximation.

We employ here the following distribution into mathematical subdomains, as employed in [15, 17]:

- Gen: General Mathematics; History; Foundations. This corresponds to sections 00, 01, 03, 06, 08, and 18 of the Mathematics Subject Classification MSC
- Disc: Discrete Mathematics. Convex Geometry; MSC sections 05, 52
- NTAG: Number Theory. Algebra. Algebraic Geometry. Group theory; MSC sections 11, 12, 13, 14, 15, 16, 17, 19, 20
- Ana: Real and Complex Analysis; MSC sections 26, 28, 30, 31, 32, 33, 40, 41.
- OpTh: Harmonic and Functional Analysis; Operator Theory; MSC sections 42, 43, 44, 46, 47.
- DIEq: Differential and Integral equations; MSC sections 34, 35, 37, 39, 45.
- OptCS: Optimization. Numerical Analysis. Computer Science. Algorithms; MSC sections 49, 65, 68, 90, 93, 94.
- ProbStat: Probability Theory and Statistics. Applications to Economics, Biology and Medicine; MSC sections 60, 62, 91, 92.
- TopGeom: Topology and Geometry; MSC sections 22, 51, 53, 54, 57, 58.
- MaPh: Mathematical Physics; MSC sections 70, 74, 76, 78, 80, 81, 82, 83, 85, 86.

The aggregation over all publication years aims to eliminate the growth effects mentioned earlier. Figure 4 shows the relative distribution of references for these ten MSC clusters in relation to the interval between publication and citation (from 0 to 24 years). It is evident that a long-term decay for relative citation frequencies of

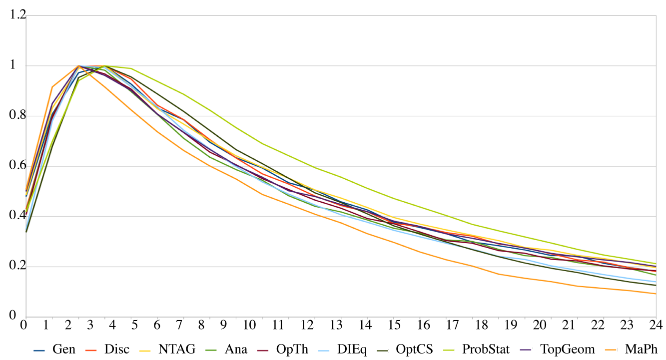

Figure 4. Relative time intervals for subject-preserving citations.

subject-preserving citations exists, but there is also a significant long tail. A notable exception is mathematical physics, where the initial relative citation rate is much higher before descending much more quickly. For the remaining areas, the diagram confirms that citation metrics that only cover a short time interval can hardly have any significance for mathematics. With the observed distribution, it becomes obvious that any measure that will not omit the most relevant information must cover a span of at least five years.[2,3]

5 Effects of publication delay

Yet another prevalent effect which provides a strong argument against the use of short-term bibliometric measures in mathematics is the exceptionally long publication delay due to the rigorous, and hence often extensive, peer-review process. zbMATH Open data can be used surprisingly easily to determine its magnitude. This is done in the following way. For many years, the arXiv has established itself as the standard preprint repository for many areas in mathematics, often preceding the actual publication by several years. Since 2016, zbMATH matches mathematical publications to their arXiv versions. As shown in [16], the arXiv is rarely used for retrospective self-archiving, hence the difference between arXiv submission and publication date can serve as a proxy for publication delay.

The diagram in Figure 5 shows the distribution of articles with respect to publication year for various arXiv submission years. As it can be seen from zbMATH Open data, the average publication delay accounts for about 18 months, but may vary significantly depending on the journal, subject, or individual paper. The effect of the subject could again be explored further by an MSC-based analysis.

---

[2] In fact, the temporal development in Figure 4 does not seem to be consistent with the results in the previous section and Figure 2. However, the decline is due to two effects: On the one hand, the citations are summarised across all years, so that the effect of publication growth is leveled out. On the other hand, citations with a large time interval are more often cross-area and therefore not included in Figure 4.

[3] Another methodological artifact should be noted that could also influence the results of other statistical studies (such as [8]): studies are often limited to the so-called top 10 % papers (this refers to papers with high short-term citation numbers, whether justified or not). With such a selection, some areas would be over-represented, even based on the zbMATH Open data, and would suggest a faster relative decline in citations than justified.

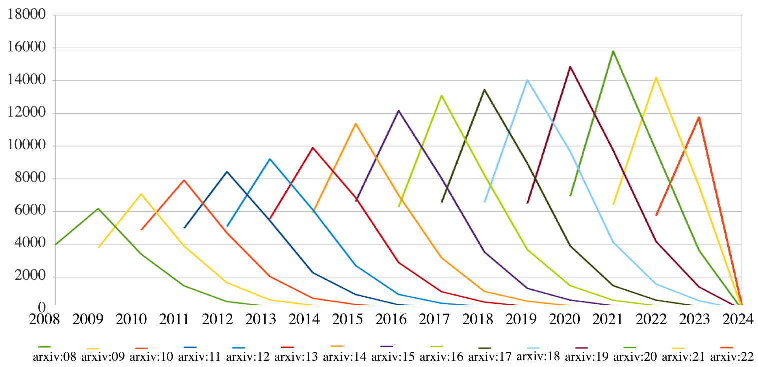

Figure 5. Publication timeline for several arXiv submission years.

With the same categories as in the previous section, the diagram in Figure 6 shows the distribution of the number of arXiv submission with respect to the average difference to the submission and publication year.

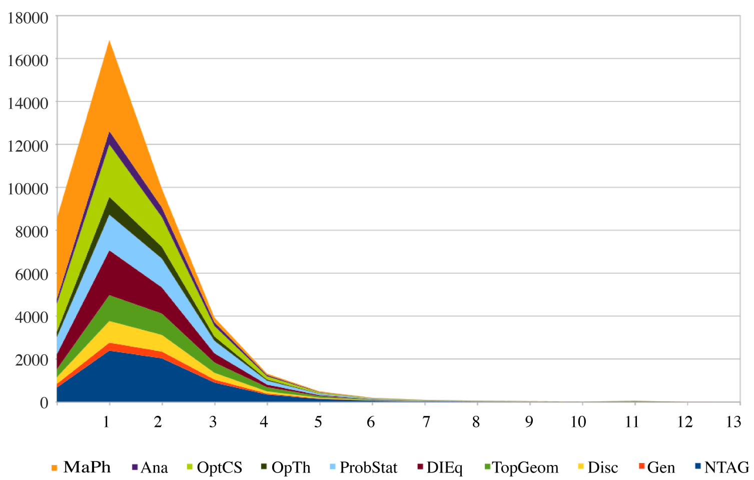

Figure 6. Publication delay based on arXiv submission dates for several mathematical areas.

One notices, e.g., that for mathematical physics (MaPh) the difference is much smaller than for core mathematics areas – e.g., for NTAG the average difference exceeds two years.

This adds further evidence that short-term bibliometric measures are inadequate for mathematics – indeed, the widely varying publication delay is a strong argument in itself that the two-year impact factor, which is often used in bibliometrics, is highly unreliable for mathematics journals [9].

6 Aggregated journal information

Citation data is often used in aggregated form, in particular summarised for journals, individuals or institutions. In this section we discuss the case of journals in more detail. Mathematical journals are characterised to varying degrees by the areas represented. In simple terms, one can differentiate between those for special topics from cross-field to general mathematics journals, although the definition of a general journal is not trivial and even in such cases the regional representation can vary widely [15].

In addition, the focus changes over time, there are changes in editors, and sometimes journals are renamed or produce spin-offs. The variance of citation measures is even greater between specialist journals to which the subject specifics considered in the previous section can be directly transferred.

The diagram in Figure 7 shows the total publication and citation numbers for four classes of journals. This is based on the zbMATH Open internal categorisation of journals. This classification is done less with the aim of a ranking than with a quick decision on priorities in the workflow and, ideally, a fair balance of specialist areas. It therefore differs in detail from other approaches (such as the Scandinavian or Australian ranking), but of course all highly relevant general mathematical journals (Acta Mathematica, Annals, Duke, Inventiones, JAMS, JEMS, Publ. IHES, . . .) are represented in the 164 journals in the FAST TRACK category, as are the leading journals in the respective specialist areas. The other three categories distinguish further workflow priorities, with category 3 journals containing usually only a small fraction of research mathematics.

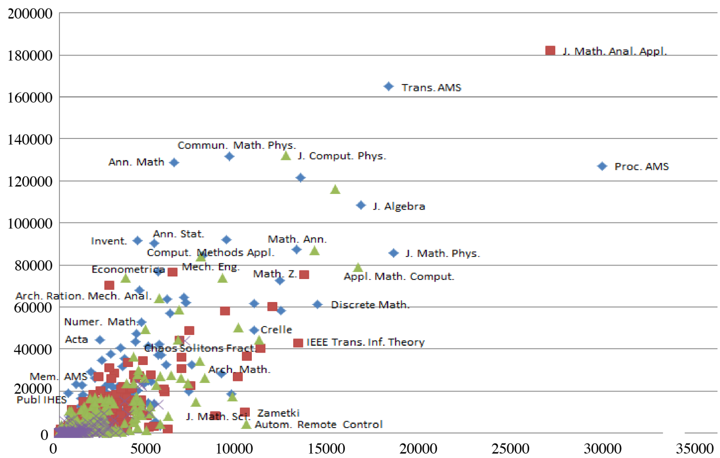

Figure 7. Publication- (x-axis) und citation (y-axis) figures of mathematical journals from four zbMATH Open categories: FAST TRACK (diamond), 1 (square), 2 (triangle) und 3 (cross).

In this diagram, the total number of all publications in the journal (x-axis) is related to the total number of all citations in this journal (y-axis). Accordingly, the slope of the origin line determined by the entry of a journal can be seen as a proxy for the average *impact factor*.

It is obvious that the spread is very wide within all categories. Indeed, the slopes vary very much with the mathematical specialties; in fact, they are strongly influenced by them. Although the average gradient in the FAST TRACK category is above category 1, it is apparently not significant, given the individual deviations. What is also striking is the often high increase in the next category 2, which is due to the fact that here a particularly large number of journals from mathematical physics or engineering are represented, so that the citation patterns of these areas dominate. In addition, there is an increased presence of journals in this category from countries (such as Iran), in which the evaluation of scientists is often very strictly linked to bibliometric values and thus a correspondingly adapted publication behaviour is enforced.

A possible conclusion is that aggregated citation information is primarily shaped by factors such as the area profile or the scientific environment – it is only after taking these dominant parameters into account that a noticeable correlation of a numerical citation indicator with the assessment made by experts can be observed. In order to analyse this in more detail, it is just as necessary to have this granular profile information available as well as to be aware of the influences of time delay and data availability and accuracy mentioned above.

7 Aggregated author information

While the previous analysis was mostly document-based, it is also worthwhile taking a more author-centred point of view when analysing publication behaviour. Such an analysis, however, requires extremely precise authorship data, since otherwise error propagation would disturb any derived quantities, making meaningful conclusions impossible. In this section, we take advantage of the significant progress of the zbMATH Open author disambiguation during the past years. Methods and progress on this matter have been amply described in [14, 18]. Nevertheless, we would like to mention that currently only roughly 3.5% of authorships are ambiguous (compared to 5% in 2018), despite the growing ratio of authorships involving Chinese names, which cause the most complicated disambiguation tasks. Most large clusters of Chinese names have now been successfully analysed (e.g, more than 1,500 documents involving the most frequent single name Wang, Wei have been distributed to currently 344 identities). The by now highly efficient author disambiguation will help to eliminate distortions in the subsequent analysis (which will take into account only the 96.5% of unambigious assignments).

We will first employ the zbMATH Open author database to derive figures on the number of actively publishing mathematicians in a given year. Some effects showing changing publication frequency and collaboration behaviour will become visible. With the assignment of MSC (Mathematical Subject Classification) classes since the 1970s, it is possible to analyse and compare these figures for different mathematical areas. For convenience (and to achieve some historical coherence by avoiding effects from the evolution of MSC) this is done for the set of ten clusters of main MSC classes which we already introduced above.

When one focuses on author counts, instead of publication numbers, one has to keep in mind that the distribution of papers is extremely biased. The median author has 2 publications, while the average publication number is about 7.9, with the maximal number of publication for a single author being 1769 (further data can easily derived from the zbMATh Open API).

There are many reasons why many authors are only connected with one paper. The obvious one is a short career in academia, often just a PhD thesis and one paper derived from this. Other people may have longer careers in research, but may switch to application areas where they drop out of the scope of zbMATH Open. In any case, this large percentage is the main reason for a large coincidence of the author and document count, as shown in Figure 8.

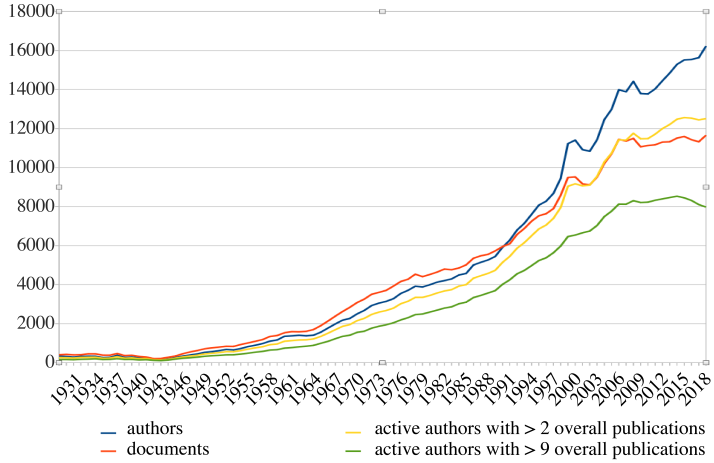

Figure 8. Actively publishing authors per calendar year, in relation to documents.

In spite of the possible methodological issues discussed above, two trends are clearly visible: (1) the number of active authors grows much quicker than that of the overall publications, and (2) the figure of established researchers with a larger number of papers grows much slower. Two main effects can conceivably play a role here

– the publication frequency and the collaborative behaviour. Due to the large number of authors with very few papers, a detailed analysis of the publication frequency is highly complicated, especially since it then seems appropriate to also involve an analysis of the length of the publications in such a study.

The overall length of publications has actually been decreasing. But this phenomenon is due to the shrinking role of books. Papers in journals have in fact become longer, at least in some areas [6]. Further effects here come from the replacement of printed by fully electronic versions and different journal policies. Again, this makes a more detailed analysis, which would also need to involve the journal status, as well as the area, quite demanding and is thus beyond the scope of this contribution. In other sciences a tendency to split results into least publishable units has been reported. At this stage our data do not allow us to draw substantiated conclusions on this for mathematics.

We will, however, see that the changing collaboration behaviour is likely to be a major factor in the increased growth of authors.

7.1 Collaboration behaviour and subject-based figures

Historically, mathematical publications were predominantly single-authored. Recently, this has changed significantly, following similar trends in other sciences. Though the overall effect is strongly driven by application areas, the phenomena are visible throughout mathematics. We employ the same categories as for the analysis of publication delay and obtain in Figure 9 a diagram of average authorships per publication for the calendar years.

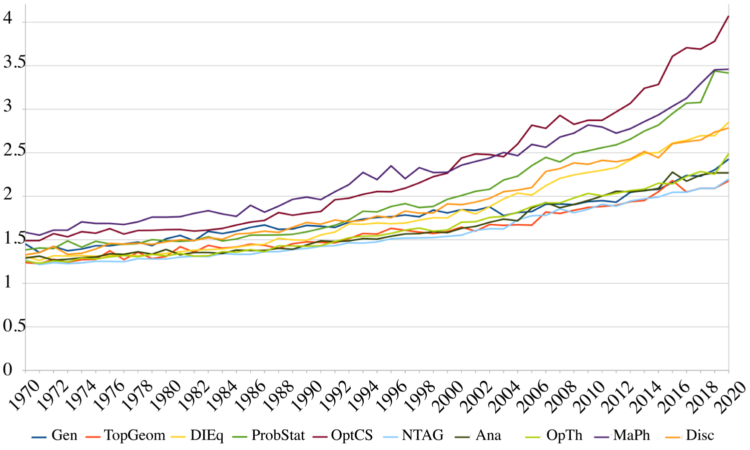

Figure 9. Average number of authors for a paper in clusters of ten mathematical areas.

There are significant differences between different clusters. Examples are given by OptCS (where the average now exceeds 4), MaPh, or Probstat (almost 3.5) and TopGeom or NTAG (about 2.2). In spite of this, however, the overall tendency is clear – collaboration has significantly increased in all fields. With mathematics being a very international enterprise, this seems to hold true globally, although samples indicate that figures may differ geographically, which may be explained both by area correlation or national science policies. However, such an analysis would again exceed the space of this article, and will be left to subsequent studies (again, the reader is encouraged to employ data available from the zbMATH Open API for a more sophisticated analysis).

Analogously, a breakdown can be made of the actively publishing mathematicians in each field; see Figure 10.

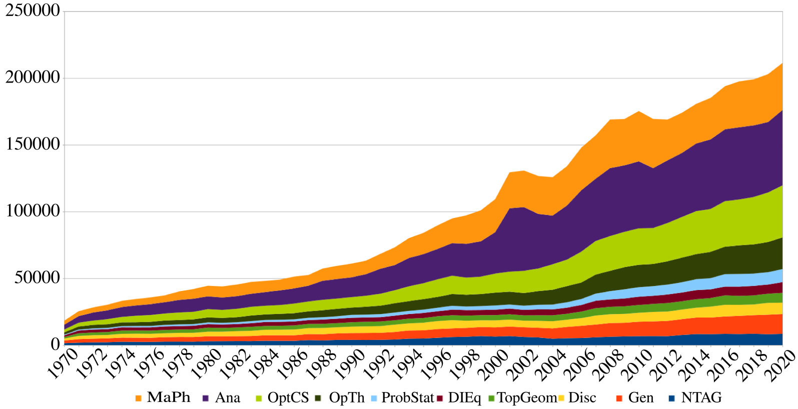

Figure 10. Actively publishing persons in ten clusters of math subjects.

There is a small caveat here – actively publishing mathematicians are evaluated separately for each area, so in the cumulative display, people active in several clusters may appear several times (the comparison with Figure 8 shows that this effect amounts to an about 20% increased height).

Summarizing, we can say that the publication behaviour has clearly changed throughout mathematics towards a more collaborative attitude, but the intensity with which this happens is somewhat different in different areas.

7.2 Citation and coauthor networks

Another aspect, which is relevant in connection with the observed increased collaboration, is the question as to how citations are distributed within the coauthor network. Although it is for many reasons clear that mathematical achievements cannot be compared on the basis of simple (especially, short-term) citation counts (cf. [1,2,4]), there is still a prevailing notion that some (possibly vaguely defined) impact is correlated with aggregated citations. For a better understanding of what citations reflect, we would here suggest a first step into an empirical analysis of their distribution in the collaboration network. Although there have been suggestions of a bibliometric index involving collaboration distances [3], it appears that such approaches have never been applied to real-world databases. One reason might be that such an analysis requires very precise authorship data, since otherwise the error propagation would lead to ever more unreliable results as the coauthor distance grows. In bibliometrics, the discussion is mostly restricted to the zero level (i.e., a possible exclusion of self-citations). This is unlikely to provide a comprehensive understanding.

The mathematics collaboration graph has been investigated frequently, especially in [13], based on zbMATH data. While the median distance in its large connected component is 5, the situation is different when one looks at the collaboration distance for citing authors.

Here one would naturally expect shorter collaboration distances. Since higher collaboration distances are linked to a higher error probability, we restrict our discussion to the ranges from 0 (self-citations), 1 (coauthor citations), 2, 3 and more than 3. The distribution shown in the diagram in Figure 11 indicates that these seem indeed the most significant categories.

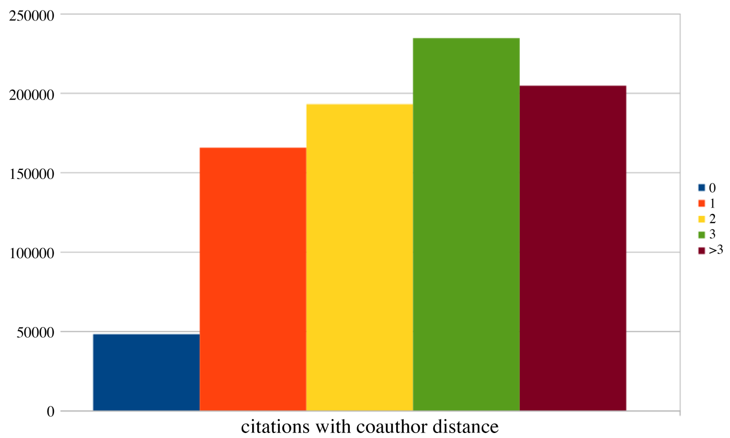

Figure 11. Minimal collaboration distance for citations of zbMATH Open authorships.

More precisely, we computed for each authorship in a paper cited in zbMATH Open the minimal collaboration distance to the citing paper (note that due to multiple authorships, the total number is larger than the overall number of matched references in the database). The figures show that both, the average and the median collaboration distance, is 3. The aggregation for authors, however, seems to indicate that the distribution is somewhat uneven; see Figure 12.

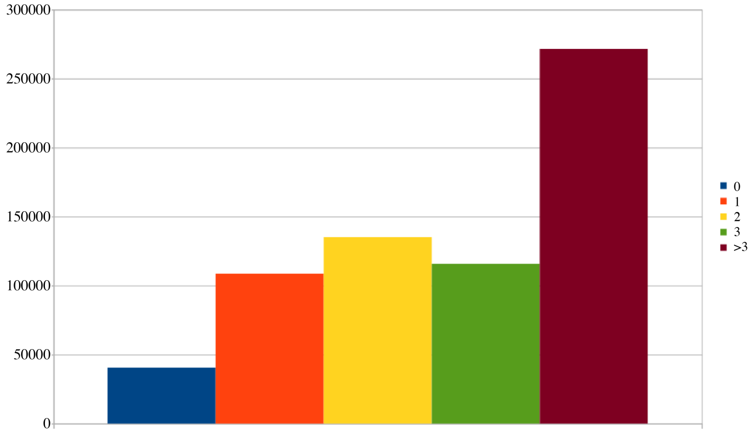

Figure 12. Number of authors in zbMATH Open with median collaboration distance n for their citations.

Of the 671,513 cited authors evaluated, most (271,435) have median collaboration > 3 distance for their citations, with a second maximum at distance 2. When we restrict this analysis to the top 15,000 cited authors in zbMATH Open (which account for more than half of all citations), the picture is, however, different; see Figure 13.

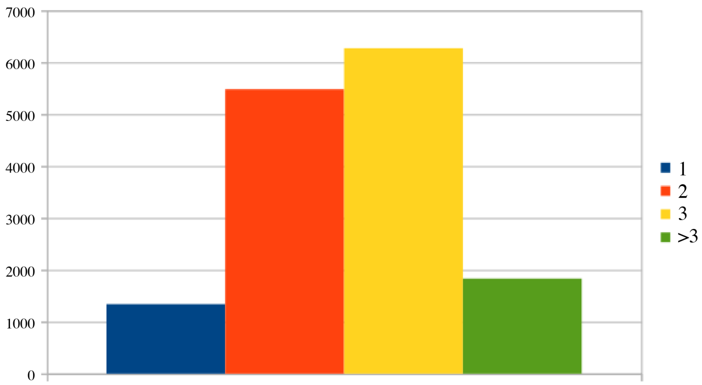

Figure 13. Number of top 15,000 cited authors in zbMATH Open with median collaboration distance n for their citations.

One sees that the distribution in Figure 12 derives from the large number of rarely cited (and thus presumably also rarely collaborating) authors, which therefore necessarily also have larger collaboration distances. For the 100 authors with most citations in zbMATH Open, the picture is even clearer; see Figure 14.

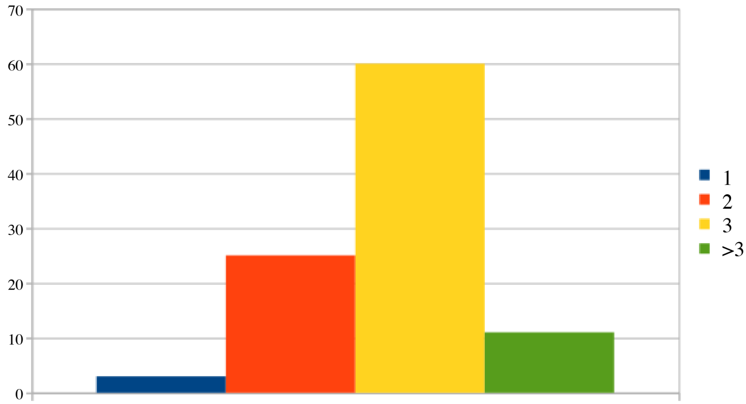

Figure 14. Number of top 100 cited authors in zbMATH Open with median collaboration distance n for their citations.

In the presence of a high number of citations, a median of three for the collaboration distance of citations seems indeed to be the default value, which is very much the standard for today's mathematical community. The larger value of four occurs almost exclusively for older mathematicians with fewer collaborations (e.g., Kolmogorov, Mac Lane, or Pólya), or in bordering areas for which collaboration paths may exist only outside the database (e.g, Barabási or Hawking). On the other hand, Erdős, who is obviously at a disadvantage due to his huge collaboration network, is almost the only elder famous mathematician with median 2; else, median 2 occurs mostly for younger mathematicians where the citations are more likely to derive from a narrower community. Especially, the rare cases of median 1 (i.e., most citations are self-citations or come from immediate coauthors) indicate almost invariably a very particular citation network.

Finally, we compare the collaboration distance (CD) distribution of zbMATH Open citations for the Fields Medalists (FM) and the highly cited researchers (HCR) in mathematics 2022 [5] of the Clarivate database:

CD	0	1	2	3	> 3
FM	7,129	37,576	117,667	193,372	130,562
HCR	29,893	139,980	164,290	175,220	81,515

The huge difference between the distribution in both series is obvious. Although the Clarivate HCR gather a much larger total citation number, only a relative small fraction affects collaboration distances ≥ 2, which usually accounts for most of the citations. By far most of HCR citations derive from the close coauthor network, and the median of two differs significantly from the corresponding figure of the most cited authors in zbMATH Open. Even as much as 10% of Clarivate HCR turn out to have an extreme collaboration median of two for their zbMATH Open citations, i.e. most of their citations are self- or coauthor citations. The difference of median citation distance for Clarivate HCR in comparison to highest cited zbMATH Open authors may indicate that the Clarivate database contains many more sources that involve large numbers of self- and coauthor citations. This adds evidence to the observation in [7] that citations for Clarivate HCR contain a significantly higher number of self-citations. Indeed, the difference exists not just at level zero, but becomes even more significant in the full distribution of citations with respect to the collaboration distance.

This indicates that the distribution of citations with respect to the collaboration distance provides a more meaningful impression of the "impact" reflected by citations. However, since it obviously depends heavily on both the age of the author and the size of the subject areas, it appears not advisable to derive yet another bibliometric measure from it. Rather, the distribution should be taken into account along with other information (such as age or subject specifics), to better understand what is usually hidden in total citation figures.

8 Conclusions

We have outlined how data available from zbMATH Open can be employed for a transparent investigation of publication and citation structures in mathematics. Even these few figures make it clear that common bibliometric measures appear to be ill-suited to reflect just only the formal bibliometric structure in mathematics publications, let alone can serve as proxies for scientific excellence. Throughout the note, we indicated several further questions which may deserve a more thorough investigation, for which data are available from the zbMATH Open API. The interested reader is encouraged to pursue a deeper analysis!

References

[1] T. Bouche and O. Teschke, An update on time lag in mathematical references, preprint relevance, and subject specifics. *Eur. Math. Soc. Newsl.* **106** (2017), 37–39

[2] A. Bannister, K. Hulek, O. Teschke, Das Zitationsverhalten in mathematischen Arbeiten. Einige Anmerkungen. *Mitt. Dtsch. Math.-Ver.* **25** (2017), no. 4, 208–214

[3] M. Bras-Amorós, J. Domingo-Ferrer, V. Torra, A bibliometric index based on the collaboration distance between cited and citing authors. *J. Informetrics* **5** (2011), no. 2, 248–264

[4] T. Bouche, O. Teschke, and K. Wojciechowski, Time lag in mathematical references. *Eur. Math. Soc. Newsl.* **86** (2012), 54–55

[5] Clarivate Highly Cited Researchers in mathematics 2023. https://clarivate.com/highly-cited-researchers visited on 14 March 2024

[6] E. Dunne, Are math papers getting longer? Blog article https://blogs.ams.org/beyondreviews/2021/10/14/are-math-papers-getting-longer/, Oct 14 (2021), visited on 14 March 2024

[7] E. Dunne, Don't count on it. *Notices Amer. Math. Soc.* **68** (2021), no. 1, 114–118

[8] P. Della Briotta Parolo, R. Kumar Pan, R. Ghosh, B. A. Huberman, K. Kaski, and S. Fortunato, Attention decay in science. *Journal of Informetrics* **9** (2015), no. 4, 734–745

[9] A. Ferrer-Sapena, E. A. Sánchez-Pérez, F. Peset, L.-M. González, and R. Aleixandre-Benavent, The lack of stability of the impact factor of the mathematical journals. In *Proceedings of ISSI 2015 Istanbul: 15th International Society of Scientometrics and Informetrics Conference, Istanbul, June 29 – July 3, 2015*, edited by A. A. Salah, Y. Tonta, A. A. Akdag Salah, C. Sugimoto, U. Al, pp. 415–416, Bogaziçi University Printhouse, Istanbul, 2015

[10] M. Fuhrmann and F. Müller, A REST API for zbMATH Open access. *Eur. Math. Soc. Mag.* **130** (2023), 63–65

[11] K. Hulek and O. Teschke, The transition of zbMATH towards an open information platform for mathematics. *Eur. Math. Soc. Newsl.* **116** (2020), 44–47

[12] K. Hulek and O. Teschke, How do mathematicians publish? – Some trends. *Eur. Math. Soc. Newsl.* **129** (2023), 36–41

[13] M. Jost, N. D. Roy, and O. Teschke, Another update on the collaboration graph. *Eur. Math. Soc. Newsl.* **100** (2016), 58–60

[14] H. Mihaljević-Brandt and N. Roy, zbMATH author profiles: open up for user participation. *Eur. Math. Soc. Newsl.* **93** (2014), 53–55

[15] H. Mihaljević-Brandt and O. Teschke, Journal profiles and beyond: what makes a mathematics journal "general"? *Eur. Math. Soc. Newsl.* **91** (2014), 55–56

[16] F. Müller and O. Teschke, Progress of self-archiving within the DML corpus, with a view toward community dynamics. In *Intelligent Computer Mathematics. CICM 2016. Lecture Notes in Computer Science* Vol. **9791**, edited by M. Kohlhase, M. Johansson, B. Miller, L. de Moura, and F. Tompa, pp. 63–74, Springer, Cham, 2016

[17] N. Schappacher, *Framing global mathematics—the International Mathematical Union between theorems and politics*. Springer, Cham, 2022

[18] O. Teschke and B. Wegner, On authors and entities. *Eur. Math. Soc. Newsl.* **71** (2011), 43–44