Green, Gold, Platinum, Nickel: On the Status of Open Access in Mathematics

Olaf Teschke (FIZ Karlsruhe, Berlin, Germany)

The various facets of Open Access (OA) have been a topic in the EMS Newsletter for many years, including articles on such diverse topics as a global description of the publishing ecosystem with its political ramifications [1] and the concrete circumstances and policies leading to the flip of a single journal [2]. Recently, the announcement of the plan of cOAlition S [3] and the choice of Elsevier as a subcontractor to monitor the progress of OA by the European Commission triggered many reactions, including a statement by the EMS [4]. The aim of this note is to give a status report on the various OA facets in the mathematical corpus, which may also provide some quantitative information helpful to estimate the impact of envisioned changes.

The ongoing progress of green OA in mathematics

Almost three years ago, we discussed the progress of green OA in mathematics [5]. While much more data are now available, it is still true that the arXiv accounts for most of the green OA deposits in mathematics. In addition, institutional repositories seem to contribute a growing share. One caveat here is that an arXiv deposit may not count as full green OA since frequently only preliminary versions are available; funders' policies may help to encourage mathematicians to include updates with equal math content to the published version (even though the layout may differ).

Figure 1 shows the share of arXiv versions matched to publications indexed in zbMATH for the main MSC subjects.[1] We can see that for several core areas in mathematics, almost half (or more) of recent articles are available. Quite remarkably, the share of arXiv submissions is sometimes now even higher than in areas of mathematical physics (MSC 81–83), which traditionally formed arXiv's nucleus. Given the relatively stable share of around 40% in the latter areas over the last few years, propelled by large journals like JHEP, which have a 100% arXiv overlap, one might wonder which figures would form a possible limit of saturation; but, so far, the overall trend still points upward.

This strong growth is less obvious from the total figures, which are still slightly below 10% of recent publications; this is, however, due to the effect of large applied areas being less present on the arXiv. However, with the achieved share and the ample amount of reference data available, we are in a position to test the effectiveness of the arXiv with respect to early dissemination.

To do so, we employ citations as a proxy for awareness. Naturally, there is the obstacle that citation numbers are influenced by many factors (not related to the quality of a paper) [6] and that the presence on the arXiv may come along with an inherent bias. Therefore, we select several larger, mostly general, mathematics journals (for our approach to this term, see a previous newsletter article [7]), which have reached a share of around 50% of arXiv overlap. Altogether, they account for about 12% of the published maths arXiv. This aggregation aims at minimising selection or subject bias. The following table and picture show publication vs. citation year for nine jour-

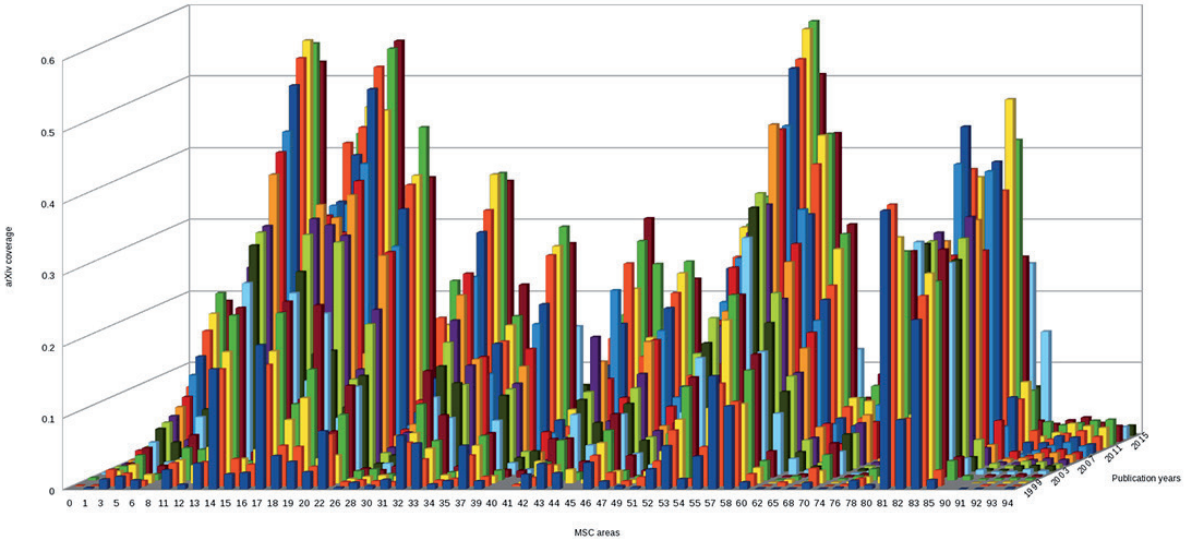

Figure 1: Progress of the arXiv share for the main MSC classes for publication years 1999–2017.

[1] Recent figures can be retrieved via zbMATH by adding "en:arXiv" to queries; in particular, the filters for MSC, years, authors or journals provide detailed information.

Table 1: Citations per year for nonarXiv/arXiv articles published.

Publication year	2011		2012		2013		2014		2015	
Type	nonarXiv	arXiv	nonarXiv	arXiv	nonarXiv	arXiv	nonarXiv	arXiv	nonarXiv	arXiv
# of publications	2006	1137	1877	1302	1616	1238	1752	1509	1885	1768
Citations in 2011	448	519	106	192	47	96	7	24	5	15
Citations in 2012	719	1098	462	452	119	186	40	86	15	22
Citations in 2013	740	1202	834	660	289	383	92	161	37	75
Citations in 2014	902	1520	1361	1031	900	767	434	630	109	258
Citations in 2015	987	1511	1483	1083	1234	1034	1060	1079	493	725
Citations in 2016	927	1498	1363	1098	1245	990	1260	1188	1057	1161
Citations in 2017	905	1408	1363	1031	1151	988	1300	1253	1321	1301
Citations in 2018	764	1055	1020	846	864	815	994	1008	1064	1144

nals (*Adv. Math, J. Algebra*, *J. Funct. Anal.*, *J. Math. Anal. Appl.*, *Int. Math. Res. Not.*, *Math. Z.*, *Pac. J. Math.*, *Proc. AMS* and *Trans. AMS*), which have the advantage of a rough balance recently of nonarXiv/arXiv articles, as well as presenting a reasonably broad variety of subjects.

Figure 2 illustrates the ratio of citations adjusted by article numbers in relation to the timeframe.

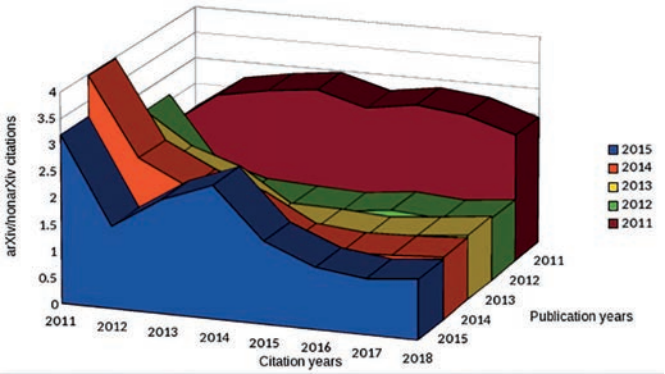

Figure 2: Adjusted citation ratio ((arXiv cit)*(nonarXiv pub)/(non-arXiv cit)*(arXiv pub))

As a first observation, we see that the ratio is always greater than one. That it approaches one in the long term seems to indicate that there is indeed no inherent selection bias. However, prior to publication, this ratio is usually larger than two, confirming (unsurprisingly) that green OA via the arXiv is indeed an effective way of dissemination (and not just preservation, as emphasised by plan S).

OA in mathematics journals: moving walls, (non-)APCs and hybrids

While this indicates that green OA serves well in supporting the dissemination and preservation role, the journal system is still (and perhaps, due to the growth of publications, more than ever) essential for ensuring scientific quality. Indeed, mathematicians consider the quality of peer review as by far the most important asset of a journal, much more than other aspects, such as access [8]. This seems completely sensible given the potentially enormous waste of resources connected to research based on false assumptions or incentives for duplicated or not genuinely new work. Hence, it is in the interest of the scientific community to maintain a reliable system of quality control and reputation management, without diverting too many resources in this process. In the past, this has been provided by the subscription model, with the revenues of society journals frequently used for grants, prizes, etc. Inherently tied to the print costs, subscriptions also came along with an incentive to create quality content. This aspect has faded in the age of digitisation and large bundles; indeed, the subscription model came under just criticism in connection with cases of monopolistic profits, frequently generated by the sale of bundles bloated by dubious content and lopsided incentives due to the problematic use of bibliometric measures by funders. The push for OA models was also a means to counteract these developments.

Due to the extreme longevity of mathematical research [9], by far the gravest issue with respect to making the maths corpus openly available is still to open up the knowledge assembled in the past. This is especially true since green OA simply doesn't work retroactively on a large scale [5]. The obvious solution is provided by wall ("embargo") policies, which require making the content available after a defined period of time. This approach has been pursued in various initiatives; notably, EuDML [10] has been very successful in assembling free resources, especially from European society journals. ProjectEuclid [11] and MathNet.Ru [12] have pursued a similar approach. The push from mathematicians has also motivated other publishers to adopt moving wall policies. Experience from societies (e.g. the AMS and the EMS) indicates that this hasn't harmed subscriptions; in addition, it perfectly complemented the arXiv corpus. Another important side effect of establishing sustainable moving walls is that they effectively undermine monopolistic approaches.

On the other hand, even with voluntary self-archiving and moving walls broadly available, a fraction of the corpus would remain at least temporarily behind paywalls. More importantly, both the mature status of green OA and the importance of historical content is quite specific for mathematics; other areas, along with funders, have therefore more eagerly ventured to establish business models allowing for immediate OA at the journal source. So far, the main distinction is the existence of article processing charges (APCs) (usually referred to as "gold OA") or the lack thereof ("platinum" or "diamond

OA"), in which case the infrastructure costs are usually covered by a third party (very often by an individual or local efforts but sometimes by larger infrastructure funds or libraries realigning subscription resources). APCs have been successfully established in several amply funded disciplines (e.g. medicine and biology) to an extent that both existing large commercial publishers consider them as the major tool for future revenue growth beyond the saturated subscription market (see the 2018 Springer Nature IPO prospect) and various new players have been attracted by this market. This development has been accompanied by intense lobbying both at European and national level, supported by the establishment of policies strongly in favour of APCs (e.g. the UK Finch report [13]). However, mathematicians overwhelmingly reject APCs for different reasons, perhaps the most prominent being that they install incentives to eradicate quality control (see N. Taubert [14] for a very detailed sociological study about the attitude of mathematicians toward the various aspects of OA). Below are some indications that this expectation has been mostly confirmed so far. A related development has been the installation of hybrid OA models, i.e. APCs for single articles in subscription journals. While some publishers adopted it as a means to allow submissions under constraints like those imposed by the Finch report, they are now widely considered to be problematic due to the option of double dipping when connected with the sale of large bundles of these journals, and limited benefits for science. While several measures have been proposed to adjust pricing with respect to APC-covered articles, the inherent lack of transparency of the bundle sales, and of a penalties mechanism in case of double charges, makes it hard to imagine them as a viable long-term option, although they may have at least the advantage that there are less incentives for lower quality, since this may destroy the journal's reputation and hence affect the demand for subscriptions. This would, however, only be true in the rare case of non-bundled sales. In any case, hybrid OA has failed to meet the demand for broad OA (in mathematics, its existence is extremely scarce), such that even APC advocates seem now to be in favour of its abolishment.

Finally, non-APC journals exist in a large variety. Some of them were established at the beginning of the digitisation era and flourished basically as one-person projects, with all the relevant advantages (reduced overhead) and disadvantages (lack of sustainability and scalability). Others are supported by institutions, foundations or donations. Some projects have established a scalable technical infrastructure, e.g. by creating arXiv overlay journals and employing open source software like OJS, but they must rely on third-party resources for the remainder of work (see, for example, M. C. Wilson et al. [2]). While the list of quality platinum journals in mathematics is considerable (see below), they are far from being already sufficient to ensure a full transformation to OA. That would require the allocation of considerable funds from libraries (or funders currently willing to cover APCs) to their maintenance. If the implementation of plan S would lead to such a support, this could lead to immense progress. However, experience advises that even in the case of existing funds, the transition of a journal to such a model requires at least a year; hence the process could be obstructed by too close timeframes.

However, there exists a long tail of less prominent journals without APCs (see the table below). Frequently, their number is cited as an argument that the majority of OA journals don't require APCs, making a swift and broad transition likely. However, a closer look reveals that they are of quite a diverse nature. For instance, a considerable number have been founded in countries that rely heavily on quantitative measures of research evaluation. Since the introduction of the Hirsch index, which takes both publication and citation figures into account, the incentives have grown there to create large clusters of publications that are often only small improvements or variations of each other. There are a number of OA journals aimed at bringing them to account. Notably, they have a thematic bias toward certain areas, e.g. fixed point theorems, fractional DE, elementary numerical methods, inequalities, fuzzy structures, classes of functions or sequence spaces, and combinatorial identities. In other cases, groups of researchers are engaged in a relatively small subject that finds obstacles to being published in traditional journals so, again, founding a specific OA journal is a welcome solution. In both examples, frequently (but not necessarily), there is a considerable overlap of the authors with the editorial board, making it harder to judge the quality of peer review. On the other hand, such overlap may also occur initially for very good OA journals, just to get the project started. Often, only time will tell in what direction the journal will evolve.[2]

Frequently, journals (or whole publishers) with dubious policies are referred to as predatory OA. This name is somewhat imprecise: similar bad practices can also occur in subscription journals (as in the case of Chaos, Solitons & Fractals, where the accumulation of bogus science also served commercial interests by inflating citation and publication figures). On the other hand, inappropriate quality control may not always derive from bad intentions. Perhaps the most clear-cut cases are APC publishers who limit the rejection of papers out of commercial interest, and even get rid of boards they consider to be too selective. Often, this is accompanied by practices like spam invitations (both for editors and authors).

Table 2 gives a rough breakdown of the various OA journal types in zbMATH[3] according to five categories. In addition, the number of total and recent (2016–2018) articles in these is given, indicating that sheer journal numbers may be misleading. Note that the categories are not a fully adequate proxy for quality – they are primarily internally employed for workflow priorities at zbMATH.

[2] The author would suggest naming the less convincing efforts of this type 'nickel OA' from the element of the platinum group famously named after a mountain spirit responsible for letting precious metals vanish from the ore. But one has to keep in mind that a large spectrum of alloys is possible.

[3] In the zbMATH serial search, results can be restricted to OA by adding the st:o option. Additional options are available under the recently introduced structured serial search.

Table 2: Categorisation of various journal types in zbMATH, with the number of journals, and their corresponding total and recent (since 2016) articles.

Category	non-APC OA			APC OA			Moving Wall (Eventually OA)			Subscription		
	Journal	Total#	Recent#	Journal	Total#	Recent#	Journal	Total#	Recent#	Journal	Total#	Recent#
FAST TRACK	13	9053	885	4	226	128	26	191721	16419	128	350425	28384
CAT 1	113	99279	13394	8	6786	922	36	116641	12386	336	512591	53734
CAT 2	201	73563	9316	39	21607	2513	23	39553	2878	435	545825	55961
CAT 3	114	14332	1785	52	10300	2359	2	1252	125	227	159908	17365
Under scrutiny[1]	85	20719	1370	31	7011	7720	1	2440	194	86	87452	6295

One caveat here is that these distinctions have been made partially heuristic and might be subject to discussion individually. As an example, looking at one of the smallest categories (the four FAST TRACK APC journals), both the Forum of Mathematics journals seem to very frequently waive APCs, even though they are much lower than average journal APCs, so they might almost be considered to be non-APC journals. The other two are also peculiar since they are the APC spinoffs of the Proceedings and Transactions of the AMS (titled "B"), which puts them close to the function of a hybrid journal (and therefore the subscription category). Both also offer a good measure for the acceptance of APCs in the mathematical community: since 2015, the subscription variants had almost a hundred times more articles than their "B" alternatives. The vast majority of APC journals indexed in zbMATH are currently published by Hindawi, SpringerOpen or de Gruyter Open.

On the other hand, a large majority of journals are still not immediately OA. In the top categories, this especially concerns journals by society publishers, while this column is overall dominated by Springer and Elsevier.

Some conclusions

Mathematics has made considerable progress toward an OA corpus, especially in the direction of green OA via the arXiv. There is certainly room for improvement, both concerning the coverage, the deposited versions and the licences (older arXiv content usually comes only with a distribution licence, while there are now diverse compliant licenses like CC BY 4.0, CC BY-SA 4.0 and CC BY-NC-SA 4.0). Diversity is important here; e.g., CC BY (as recommended by plan S) would, e.g., allow for rebundling genuinely OA content in commercial platforms even behind paywalls, in contrast with the NC (non-commercial) option. The SA (share alike) option would require making also derived work equally openly available. It should be noted that the arXiv also provides, as quite a unique feature, the distribution of LaTeX sources, which already facilitates formula search [15] and might be essential in the future to develop semantic features. This is complemented by a large ecosystem of journals that are platinum or moving wall OA, accounting for a considerable proportion of core mathematics. However, a swift change to immediate OA, like that envisioned by plan S, would pose considerable challenges in order to maintain the benefits of the ecosystem of quality control. Especially, APCs appear to be an inadequate alternative. While large commercial publishers have shown a lot of flexibility in generating revenues, society publishers would face imminent problems. An appropriate means to enable their OA transition could consist of funding models supporting platinum OA models, ideally via platforms that also allow for support in sustaining the diverse existing platinum projects, which too often rely on local efforts with rather scarce resources.

[4] The "under scrutiny" category comprises journals that are currently (re)considered, both from category and indexing viewpoints. Especially in the area of low-category APC journals, indexing has been discontinued in several cases due to questionable peer review. It indeed became quite a pattern that already suspicious journals eventually started publishing article of elementary proofs of the Riemann hypothesis or the Goldbach conjecture.

References

1 M. Farge and F. Hélein, Is the system of scientific publications on the eve of a revolution? And if so, toward what? *Eur. Math. Soc. Newsl.* 108, 35–40 (2018; Zbl 1392.68436).

2 M. C. Wilson et al., Flipping JACO. *Eur. Math. Soc. Newsl.* 109, 38–41 (2018; Zbl 06980714).

3 https://www.scieur.org/coalition-s.

4 http://euro-math-soc.eu/news/18/10/1/ems-new-developments-concerning-open-access.

5 F. Müller and O. Teschke, Will all mathematics be on the arXiv (soon)? *Eur. Math. Soc. Newsl.* 99, 55–57 (2016; Zbl 1345.68267).

6 A. Bannister and O. Teschke, An update on time lag in mathematical references, preprint relevance, and subject specifics, *Eur. Math. Soc. Newsl.* 106, 41–43 (2017; Zbl 1384.01099).

7 H. Mihaljević-Brandt and O. Teschke, Journal profiles and beyond: what makes a mathematics journal "general"? *Eur. Math. Soc. Newsl.* 91, 55–56 (2014; Zbl 1296.01048).

8 C. Neylon et al., Results of a worldwide survey of mathematicians on journal reform. *Eur. Math. Soc. Newsl.* 103, 46–49 (2017; Zbl 1366.01071).

9 T. Bouche et al., Time lag in mathematical references, *Eur. Math. Soc. Newsl*. 86, 54–55 (2012; Zbl 1263.01012)].

10 https://eudml.org.

11 https://projecteuclid.org.

12 https://www.mathnet.ru.

13 https://www.acu.ac.uk/research-information-network/finch-report-final.

14 N. Taubert, *Fremde Galaxien und abstrakte Welten – Open Access in Astronomie und Mathematik: Eine soziologische Analyse*. 373 p., urn:nbn:de:0070-pub-29150353 (2018).

15 F. Müller and O. Teschke, Full text formula search in zbMATH, *Eur. Math. Soc. Newsl*. 102, 51 (2016; Zbl 1366.68355).

Photo and CV of the author can be found in previous Newsletter issues.