The "Jahrbuch über die Fortschritte der Mathematik" as a part of zbMATH Open

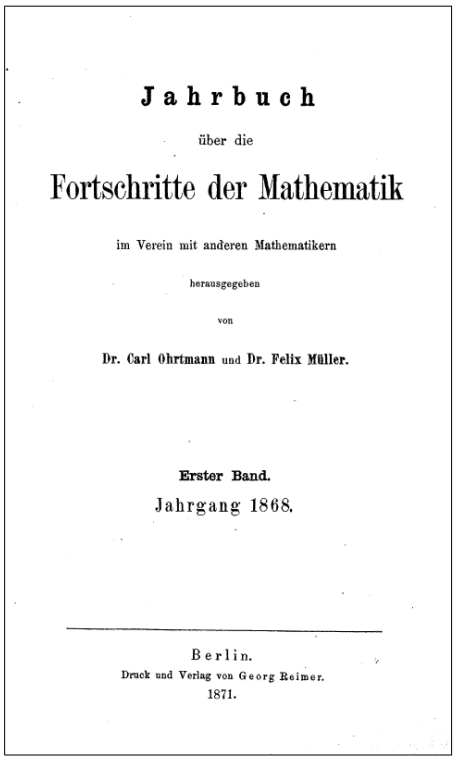
Olaf Teschke

2021 marks the 150th anniversary of the publication of the first volume of the "Jahrbuch über die Fortschritte der Mathematik" (JFM), as well as the transition of zbMATH to an open service 90 years after its founding. Initially digitised in the ERAM Project 1998–2004, JFM data have benefitted significantly from the subsequent integration into zbMATH, and are now available in a much enhanced form. We describe the improvements of the digital JFM version during the last decade, which are now available both as Open Access Database and Open Data.

1 Pieces of JFM history

According to most sources, Carl Ohrtmann and Felix Müller founded JFM at the end of 1868 with the aim of collecting, indexing, and reviewing the global mathematical literature in annual volumes. It followed the ideas of already established review journals like *Pharmaceutisches Central-Blatt* (later *Chemisches Zentralblatt*, founded in 1830), or *Fortschritte der Physik* (founded in 1847). The first JFM volume, covering 838 publications from 1868, was published in February 1871 – even in the beginning, the ideal of a complete and classified collection could only be achieved by a considerable delay.

Similar initiatives were at this time Boncompagni's *Bullettino di bibliografia e di storia delle scienze matematiche e fisiche* (1868), Darboux's *Bulletin des sciences mathématiques et astronomique* (1870) and the Dutch *Revue semestrielle des publications mathématiques* (1896), but JFM prevailed due to its broad community of expert reviewers, supported by the leading role of German mathematicians at that time. Until WWI, JFM defined the standard to judge contemporaneous research, and shaped both scope and classification of mathematics. The setback of the war, however, could never be regained: the admired comprehensiveness and the detailed knowledge of the organisation became a liability, leading to growing delays of up to seven years. Moreover, German was no longer the primary language of math publications, calling into question the existence of German-language-only reviews. In this critical situation, additional resources for JFM provided by the Prussian Academy since 1930 turned out to be double-edged: though they helped with the catching up, they forced JFM to follow its rather conservative politics, pursuing the ultimately infeasible aim of restoring JFM's pre-WWI status in a changed environment. This inflexibility intensified the decline of JFM, which ultimately deteriorated due to the outcomes of Nazi politics (the details of decline and fall of JFM are investigated in [7]). The *Zentralblatt für Mathematik und ihre Grenzgebiete* (ZfM, now zbMATH Open), founded in 1931 by Otto Neugebauer with the focus on timeliness instead of systematic comprehensive annual volumes, which assembled a new generation of a global community providing multilingual reviews, soon replaced JFM as the primary source of current research information, with the average review becoming available only after an impressive 0.58 years. After Neugebauer's emigration to the



Jahrbuch

über die

Fortschritte der Mathematik

im Verein mit anderen Mathematikern

herausgegeben

von

Dr. Carl Ohrtmann und Dr. Felix Müller.

Erster Band.

Jahrgang 1868.

Berlin.

Druck und Verlag von Georg Reimer.

1871.

Figure 1. The first page of the first JFM volume

EMS MAGAZINE 122 (2021) — DOI 10.4171/MAG-58

U.S., the second journal he founded, *Mathematical Reviews* (MR) took the lead following similar principles, leaving JFM only the role of a well-organised collection of historical interest. Consequently, attempts to revive JFM after WWII failed, resulting in volume 67 of 1942 being the last one ever published – though former JFM staff members played a crucial role in the successful resurrection of *Zentralblatt* after the war.

2 JFM digitisation

Still after its end, JFM was used as a standard reference by MR and ZfM for earlier literature. Hence, at the dawn of math digitisation, it was natural to include information from JFM covering publications from 1868 through 1942. Initially suggested by then-EiCs of MR and ZfM Keith Dennis and Bernd Wegner as a joint National Science Foundation (NSF) and Deutsche Forschungsgemeinschaft (DFG) project in 1997, JFM was digitised as a part of the Electronic Research Archive for Mathematics (ERAM) project (1998–2002) funded by DFG and conducted by TU Berlin, Göttingen State and University Library (SUB) Göttingen and FIZ Karlsruhe. Along with the digitisation of important historical mathematical sources at SUB, JFM data were seen as a building block of the World Digital Mathematics Library initiative championed by the International Mathematics Union (IMU). Although clearly German funds alone could not be sufficient to achieve all these desirable objectives, significant results were obtained: all volumes of JFM were transformed into LaTeX and made freely available in a database, allowing for a search distinguishing author and reviewer names, titles, and review texts. At that time, it was likely the largest free LaTeX transcription project to have ever existed, and beyond its actual output, it provided some insights into feasible procedures. The combination of OCR techniques and manual transcription of formulae proved to be a manageable approach, although there was a relatively significant spread in the error rates of the various companies given the initial samples. Fortunately, DFG funds made it possible to choose the companies with the lowest error rates for the remaining parts; though some of the numerous errors coming from the initial low-cost alternatives can still be found in the data, and represent challenges both to the reader and for derived information like author disambiguation. As a collection, the LaTeX transcription of the Jahrbuch also provides a good gold data for next-level digitisation approaches like those outlined in [2].

3 Integration of JFM into zbMATH

The results of the digitisation project were made freely available after its ending in the JFM database, but a lot of work remained to be done (apart from a good compilation of earlier publications on JFM, one can find in [4] an account of the missed objectives within the ERAM project). Among the desirable features which were goals of the project that were not attained were: the standardisation of journals, author disambiguation, and interlinking with full-texts. Since these projects were being undertaken at the same time on a broader scale in the zbMATH database, JFM stakeholders allowed for the integration of the project data into zbMATH under the condition of providing resources for their enhancement. About a decade ago, this column [3] described the status at the start of the integration process. Since then, vast improvements have been achieved. The bibliographical sources, initially just a string, which could vary for a single journal from, say, Clebsch Annalen to Klein Annalen to Mathematische Annalen, with different abbreviations and mixed Arabic and Roman volume numbers. They have now not only been standardised, but have evolved into a full-scale journal database facilitating faceted searches including granular information such as titles, publishers, ISSN, main subjects, time periods, countries, languages or Open Access information and issue-level browsing. Thanks to these assignments, automated generation of full-text links is now possible. While the first JFM database did not contain a single DOI, these are now available for more than 20 % of the 223,276 JFM documents, along with 19,015 links to free EuDML entries, 8,587 to Gallica, and many to a number of other free digital libraries.

Author disambiguation has been particularly challenging for JFM entries: first names were usually abbreviated or completely missing, and typos from the chunks digitised with lower quality complicated the situation further. Approaches which work well for modern publications such as analysis of coauthor and reference networks fail due to the lack of reference data and the fact that at that time most publications were single-authored. Thus, progress depended on purely human checking of authorship assignments; fortunately this has now been done for a large part of the JFM data. As a result, JFM authorship data now contribute to comprehensive author profiles for mathematicians of three centuries. The situation is, however, less ideal for reviewer information: reviewer signatures mostly lack first names, which makes their identification very complicated[1]. A precise disambiguation and integration into person profiles remains to be done.

Likewise, the integration of JFM into zbMATH has led to its extension to a citation database. Not only could citation data be added for more than 6,000 JFM documents, but references from later publications to JFM could be matched within the integrated corpus. The resulting dataset provided the opportunity for unique analysis of long-term citation behaviour in mathematics, part of which was reported in this column [1].

[1] This is also likely the reason why the number of reviews for prominent mathematicians in earlier publications turn out to be frequently incorrect.

4 JFM as part of zbMATH Open

One major drawback, however, was that these improvements – facilitated by zbMATH resources – were only available within the commercial zbMATH database, and hence only partially accessible. Although limited to subscribers (except for the reduced free results), the improved functions led to a gradual shift in usage from the free project version of JFM to the zbMATH subset. In 2020, JFM documents were >20 times more often accessed in their zbMATH version compared to the old JFM database. The transition of zbMATH to the zbMATH Open service at the beginning of 2021 [5] resolved this dilemma: the enhanced JFM data within zbMATH now became completely free, and also provides all the information of the old JFM database as a subset[2]. In fact, this transition achieved even more: the EMS, named by the project partners as the holder of the JFM data, agreed to make it available under a CC-BY-SA 4.0 data. In particular, this dataset is completely available via the zbMATH Open API introduced earlier in this column [6]. However, there still remains a great deal to do. As mentioned above, reviewer disambiguation is lacking. There is a small overlap for the years of 1931–1945 when JFM and zbMATH were published in parallel; corresponding items should ideally be merged. Furthermore, many comments from the mathematics history community have been collected during the project; they are not in publishable condition, however, due to their heterogeneous nature. Perhaps most importantly, the digitisation of historical fulltexts may benefit greatly from recent technological developments, and their integration would facilitate additional functions like fulltext or formula search for mathematical content over a period of 150 years.

[2] This also allowed to discontinue the technically outdated old JFM interface. Note that search results in zbMATH Open can be filtered to JFM documents by adding "dt:JFM" to a query.

References

[1] 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)

[2] M. Beck, I. Beckenbach, T. Hartmann, M. Schubotz and O. Teschke, Transforming scanned zbMATH volumes to LaTeX: planning the next level digitisation. *Eur. Math. Soc. Newsl.* **117**, 49–52 (2020)

[3] S. Göbel, Jahrbuch über die Fortschritte der Mathematik and Zentralblatt MATH – reporting on more than 140 years of mathematics. *Eur. Math. Soc. Newsl.* **73**, 45–46 (2009)

[4] S. Göbel, W. Sperber and B. Wegner, 150 Jahre—Ein Rückblick auf das Jahrbuch über die Fortschritte der Mathematik. *Mitt. Dtsch. Math.-Ver.* **28**, 95–107 (2020)

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

[6] M. Schubotz and O. Teschke, zbMATH Open: Towards standardized machine interfaces to expose bibliographic metadata. *EMS Magazine* **119**, 50–53 (2021)

[7] R. Siegmund-Schultze, *Mathematische Berichterstattung in Hitlerdeutschland*. Studien zur Wissenschafts-, Sozial- und Bildungsgeschichte der Mathematik [Studies on the Scientific, Social and Educational History of Mathematics] 9, Vandenhoeck & Ruprecht, Göttingen (1993)

Olaf Teschke is managing editor of zbMATH Open and vice-chair of the EMS Committee on publications and electronic dissemination.

olaf.teschke@fiz-karlsruhe.de