# Communicating Mathematics in the Digital Era*

Eugénio M. Rocha† and José F. Rodrigues‡

## 1 Introduction

At the dawn of the digital age, the importance of mathematical documentation assumes a strategic and decisive character, both in current research as well as in the transmission of mathematical knowledge to future generations. Of course, this is something that happens in all branches of knowledge. Documentary resources such as databases, publication archives or electronic resources (digital books or simple web pages) as well as increasingly powerful search engines, are essential instruments not only for science but also for the economy and the information society. However, for Mathematics and all who use it, the importance of documentation is strengthened by the generality and permanence intrinsic to Mathematics, to which is due its unique situation amid the other branches of human knowledge.

Historically, the transmission of Mathematics is indissociably connected to the technology specific to each period, traversing the last the twenty-five centuries of civilization. From manuscripts and old copies to the typographical press, from classic libraries to digital repositories, from verbal or postal transmission to electronic communication, the technological revolution that we are living through presents mathematicians, mathematics professors and all professionals who use Mathematics with new challenges and new opportunities. Mathematical documents are currently created and transmitted in digital format, thus allowing inclusion of other elements besides the traditional text and static images. Computers and the Internet have radically altered our ways of communicating and sharing ideas and results, temporally and dimensionally enhancing the potentialities of the traditional forms of human thought and intelligence.

There are many players in this complex universe: authors, referees, editors, publishers, libraries, repositories, WWW servers, readers, financing foundations and agencies, universities, institutes and scientific societies. At the turn of the century, an ambitious idea that gave rise to the World-wide Digital Library of Mathematics (WDML) [46] started to mobilize efforts and to generate projects of national and international collaboration for the purpose of creating a gigantic virtual library to be used by the global mathematical and scientific community. However, this set of changes and adaptations that constitute the revolution in information technology has not yet settled down new ways of sharing and developing the Mathematical sciences. Given that Mathematics is particularly dependent on its literature, its credibility and accessibility, and that these might easily be jeopardized by the increasing ease of electronic

*This article is a translation of the Portuguese original one published in Boletim da SPM 53, Outubro 2005, pp. 121.

†CEOC and Univ. of Aveiro *(eugenio@mat.ua.pt)*

‡CMUC and Univ. of Lisbon *(rodrigue@ptmat.fc.ul.pt)*

publication and communication, the Sociedade Portuguesa de Matemática aims to contribute to the necessary awareness and exploration of the new techniques.

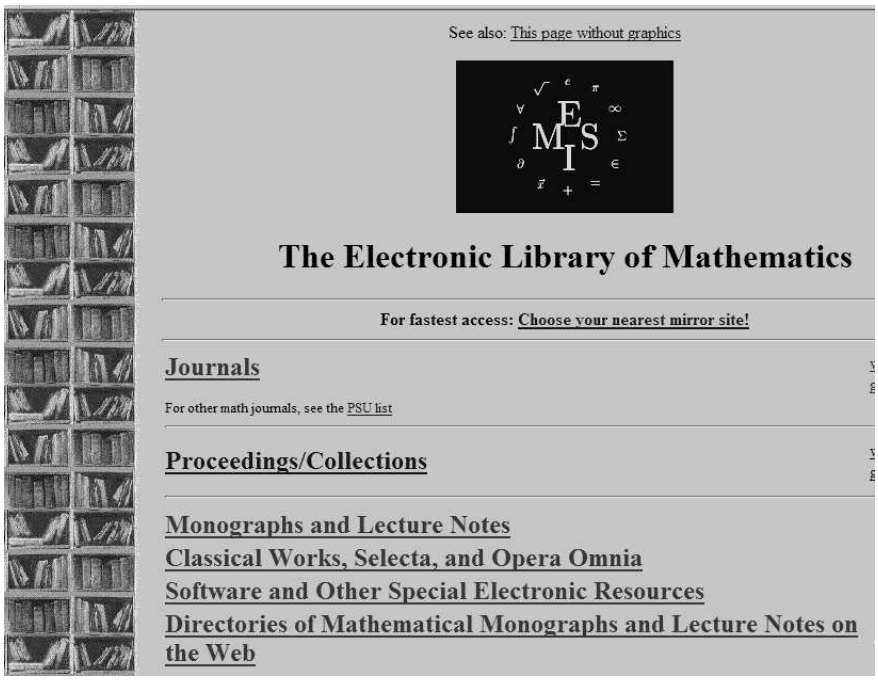
## 2 Electronic Publications in Mathematics

Five centuries separate the first typographical edition of the Elements of Euclid, printed in Venice in 1482 by Erhard Ratdolt, from the announcement in 1978 of the TEX language by its creator, the mathematician and computer scientist Donald Knuth [10]. This should be contrasted with the five years that elapsed between this event and the adoption by the Transactions and the Proceedings of the American Mathematical Society (AMS) of this digital system of mathematical composition. Although it was neither the first nor the only electronic typesetting system, TEX became stable and universally accepted by the mathematical community, despite the existence of certain dialects like LATEX and AMSTEX. For mathematicians today it constitutes almost a reinvention of the alphabet and has become an essential instrument of communication. Nowadays, practically all submissions of mathematical works, be they articles or monographs, are based on TEX and on its diverse forms of visualization (PostScript, PDF, etc.), which not only immensely facilitates their publication but also cuts the publication costs substantially, "democratizing" mathematical writing and typography. Moreover, the new versions of TEX permit the inclusion of hyperlinks in the electronic text, which allow new searches associated to hypertext and electronic publication in cyberspace.

While the 1980s marked the beginning of the systematic use of TEX in mathematical typography, the following decade saw the widespread use of electronic mail and, above all, the invention of the WWW. These new electronic resources have brought new methods of communication such as the servers and repositories of pre-publications and electronic journals. Scientific journals are the privileged vehicles for the publication of mathematical research, being an established form that goes back two centuries. The Journal für die Reine und Angewandte Mathematik was established in 1826 in Berlin by A.L. Crelle, and the Journal de Mathématiques Pures et Appliquées in 1836, by J. Liouville in Paris and both are still published today, being periodicals of unshakeable scientific reputation. The journal tradition is based on a publishing system of critical pre-reading (refereeing), decision (by the editorial commission) and scientific publication (by the publisher). This enriches the value of each author's work in a way that is difficult to quantify; the mathematical community values this system and considers it irreplaceable, while looking for new forms of guaranteeing it in the digital age.

The first appearance of solely electronic Mathematics journals took place as far back as 1992 with the ephemeral Ulam Quarterly that ceased publication with the third volume four years later, and with the Electronic Journal of Differential Equations and the Electronic Transactions on Numerical Analysis that have been published regularly since 1993. Today, there are over ten titles directly accessible from the EMIS/ ELibM [45] catalog of the European Mathematics Society (EMS). However, if the acceptance of this new type of journal took some time, with the gradual digitization of all Mathematics journals and their availability, at least for subscribers, in the WWW, the distinction starts to disappear. Beyond the advantage of accessibility, the solely electronic journals show that the cost of mathematical publications can be minimized. It must be said that a great number of journals, as for example, the ones edited by the recent EMS-Publishing House, already include hyper-links in their electronic editions, for formulas and for bibliography, including references to the MathSciNet [49] or Zentralblatt

MATH [50].




On the other hand, the last decade has witnessed an enormous reorganization of the publishing houses, with acquisitions and mergers, and simultaneously, a constant and paradoxical increase of subscription prices for those journals that belong to the main scientific publishing companies, which between them possess about 60% of the main journals. For example, D. Knuth [9], in a significant October 2003 letter to the Publishing Committee of the Journal of Algorithms, that he had helped found in 1979 for Academic Press, taking into account the subscription price and the number of pages published per year, noted with indignation typical of the scientific community that after many years of a steady 25 to 30 cents cost per page, the price had risen to 50 cents per page in 2002, after the acquisition of this journal by Elsevier.

In another interesting reaction to the evolution of subscription costs for Mathematics journals belonging to commercial publishing companies, the mathematician Joan Birman [1], of Columbia University, New York, observed that in the face of the commercial offensive of these costs, the mathematicians/publishers of the main Mathematics journals can and must react, looking for other publishing companies that offer to the same quality of publication for more reasonable prices. Birman gave as an example the collective resignation of 50 members of the editorial committee of the Journal of Logic Programming, after drawn out and fruitless negotiations on its price with Elsevier, followed by the foundation of the new journal Theory and Practice of Logic Programming, published by Cambridge University Press with a 55% price reduction. In this article, Birman classified the existing mathematical research journals into four categories, estimating that among the best, percentages would be 17%, 13%, 10% and 60%, respectively:

I) University journals, for example, the Annals of Mathematics (University of Princeton and the Institute for Advanced Studies) or the Annali della Scuola Normale Superiore di Pisa (Italy), that accept exchanges;

II) journals belonging to scientific societies, as the Transactions of the American Mathe-

matical Society or the Portugaliæ Mathematica, of the SPM;

III) journals belonging to University Presses, as for example, the European Journal of Applied Mathematics that belongs to Cambridge University Press - almost indistinguishable from type I), except that they do not accept exchanges;

IV) journals belonging to commercial publishing companies.

More precise data has recently been supplied by John Ewing [6], Executive and Publishing Director of the AMS, in a cautious article which nevertheless supports experimentation with new technologies to find better ways to communicate Mathematics. Ewing relates that, in 2001, Mathematical Reviews indexed or analyzed 51,721 articles from 1172 distinct journals. Of these journals, 591 (about 50%) were "cover to cover" (journals considered to be only about Mathematics) containing 30,924 (60%) articles, while the remaining 581 classified as "other", i.e., multidisciplinary or out of the mathematical "mainstream", contained the remaining 40% of Mathematics articles. However, that year, only 46 (4%) of the journals, containing only 1272 articles (2.5%), were considered "primarily-electronic", even though approximately 2/3 of all the articles, i.e., about 34,000, possessed hyper-links, meaning that at least they were available in electronic version. Also of interest is the information Ewing relays about the raised costs of subscriptions to commercial journals that, owned by great publishing companies, continue to generate high profits and to dominate the market. In 1991, only 24% of the journals were commercial and published 38% of the articles, but in 2001, some 30% (349) of the journals were commercial and published 48% (25,008) of the Mathematics articles! Ewing thus concluded that alternative electronic editions in Mathematics, not only the electronic journals but also the pre-publication repositories like arXiv [14] and MPRESS [34], still represent under 10% of all recent mathematical publication and do not present immediate danger to the great publishing companies, which have adapted to electronic publication by offering new services with the subscriptions.

On the other hand, Ewing [6] relates an extraordinarily interesting statistic about one of the new characteristics of the Mathematical Reviews, which consists of referring to the list of citations in articles of certain journals, which included, in 2002, about 340,000 citations made since 1998. Of these, 53% referred to articles published before 1990, and more than 28% were articles previous to 1980, which is particularly significant given the increase of Mathematics articles. This fact highlights the singularity of the importance and permanence of mathematical literature for the advance of this science, in contrast with the other sciences, including physics, whose literature has a more ephemeral character.

In Portugal, the mathematical press is five centuries old, going back to 1496 with the publication, in Leiria, of Abraham Zacutus' astronomical tables - the Almanach Perpetuum - and with the publication of the Tratado da Pratica d' Arismetyca, by Gaspar Nicolás, in Lisbon in 1516. In spite of its reduced dimension and impact in the recent development of Mathematics, the first periodic Portuguese mathematical journal dates from 1877 when the Journal das Ciencias Mathematicas e Astronomicas was founded in Coimbra by F. Gomes Teixeira. This journal was published for a quarter of a century. Nowadays, only one mathematics research journal is published in Portugal, namely Portugaliæ Mathematica which was founded in 1937 and is currently published by the SPM. Portugaliæ has used computer typesetting in TEX since Volume 47 in 1987, and has had an electronic version on the Internet, still limited, since 2001 in the EMIS ELibM [12].

Despite the relative sluggishness in the process of substitution of paper by digital, and their continuing coexistence in the final version, paper has almost disappeared in the initial phases of submission, refereeing and production of scientific articles, where the use of email and TEX-

based files is almost universal. This confirms that electronic publication is a well-established and consolidated fact in 21st century Mathematics communication. Such a situation was already predictable in the 1980s: one of the authors of this article remembers how in 1988/1989 he published, in Lisbon, a collective book for Birkhäuser, for which more than half of the twenty contributions arrived electronically in TEX.

If the problems of production and distribution of mathematical literature are now accepted by the scientific community, both the long-term archive and the viability of access costs constitute serious problems, so far without stable solutions and thus a worry for both librarians and scientists.

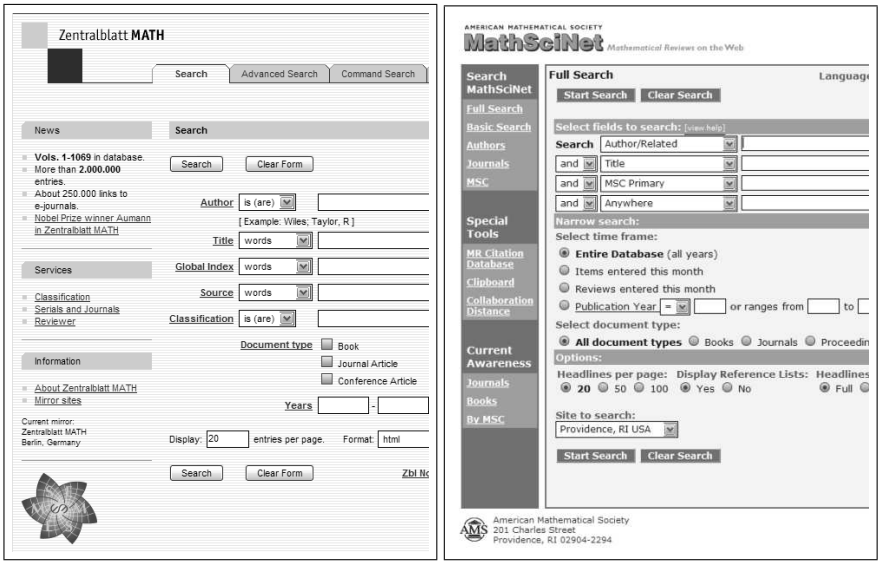
## 3 The Digital Library of Mathematics

Mathematical literature has a long tradition of organization and archiving, in part due to its accumulative nature, in part to the necessity felt by mathematicians to create databases and reviewing journals as Zentralblatt MATH (associated to the EMS) and the Mathematical Reviews (of the AMS), currently accessible on the Internet. This tradition goes back to the 19th century, from which the Jahrbuch über die Fortschritte der Mathematik (JFM) remains an outstanding example, having reviewed over 200,000 mathematical publications in the 68 volumes published between 1868 and 1942. This journal is currently the subject of an on-going project (Electronic Research Archive for Mathematics) of on-line digitization, having been distinguished recently with the 2005 award by the Special Libraries Association/Physics-Astronomy-Mathematics Division. In May 2005, the JFM site [32] already made use of 17,035 external hyper-links to facsimiles of original documents to be found in various libraries and repositories in Germany, France and U.S.A. With the gradual retro-digitization of mathematical documents and the placing online of global digital mathematical document repositories, all mathematical literature will one day be at the distance of a click for any person, anywhere at any time.

It is clear that this vision, susceptible of becoming reality more rapidly than might be imagined, still has many obstacles in front of it, including the technical problems of accessibility and long-term archive preservation. Among the current initiatives it is worth mentioning EMANI (Electronic Mathematics Archives Network Initiative) [22], a project that the German geometer Bernd Wegner, one of its main coordinators, describes in [13].

The current rate of mathematical publication is enormous; the Zentralblatt MATH databases, published since 1931, contain more than two million mathematical articles, growing at about 80,000 articles per year and covering about 2,300 journals, serial collections, conference proceedings, collected works and books. These numbers show why the traditional libraries struggle with problems of space, financing and other serious hindrances that in Portugal we know so well.

The European Mathematical Society (EMS), in cooperation with the Facinformationszentrum Karlsruhe, opened the EMIS portal, a European Service of Mathematical Information that has been available on the Internet since June 1995 [8]. With free access and with dozens of "mirrors" in the whole world, including one in Lisbon, this portal constitutes a most useful service to the international community with three main components: an electronic library, a collection of databases (some available through subscription) and projects. In 1995, the EMS made available in the EMIS an Electronic Mathematics Library (ELibM) [45], supervised by its Committee for Electronic Publication. From the 12 Mathematics journals that

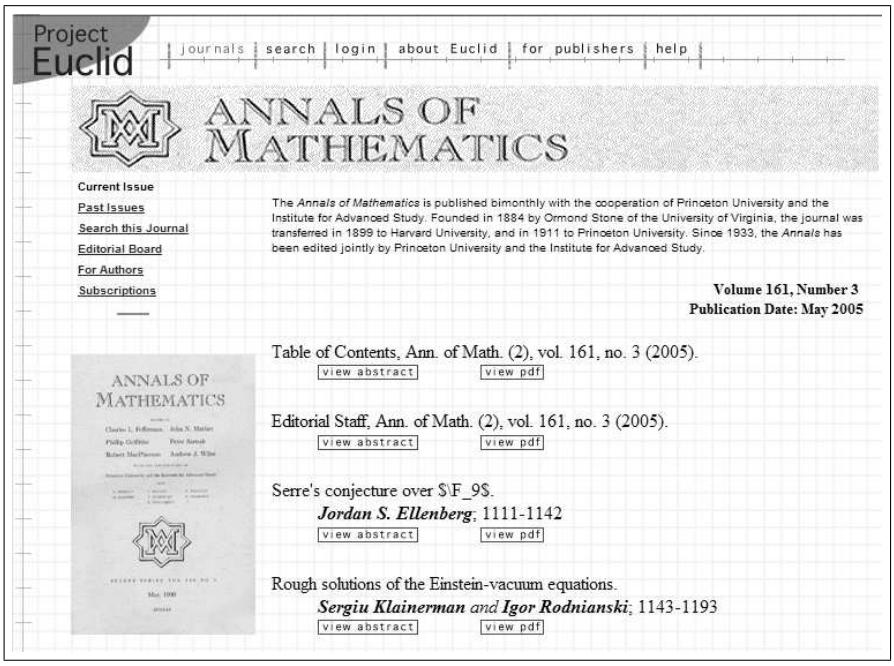



were available in November 1995, it progressed to the current catalog of 69 journals (October 2005), including the Portugaliæ Mathematica, as well as a long list of Proceedings/Collections (including the Proceedings of the International Congress of Mathematicians, 1998 (Berlin)), Monographs and Lecture Notes, Classic Works (including the collected works of Hamilton and Riemann), electronic and computational resources and, recently, a free access directory of books and monographs on the Internet. All the material is of free access, except for the availability of journals that satisfy the so-called "moving wall", i.e., a period of restricted access, usually of three to five years, after which the articles become available free. This period allows the publishing companies to recoup investments made in the publication process. In the follow-up to the International Congress of 1998 in Berlin, the International Mathematical Union (IMU) created a Committee for Electronic Information and Communication (CEIC). In the context of the initiatives planned for the World Mathematical Year, in 2000, mathematicians had started to speak of the necessity of digitalizing the historical Mathematical literature in order to make it available and to relate it with the current mathematical production, which tends to be all digital, in one way or another. Philippe Tondeur, professor of the University of Illinois and director of the NSF division of Mathematics (National Science Foundation, USA) between 1999 and 2002, supported the idea of establishing an international project for massive digitization of articles with the aim of creating the Digital Mathematics Library (DML). To this end, the NSF financed the Library of the University of Cornell for a year, supporting a discussion group on the subject, currently known as the DML Planning Group. To encourage the idea, John H. Ewing (executive director of the AMS) published an article [7] in 2002 in which he considers the possibility, using current technology, of digitizing all the old mathematical documents, estimated by Keith Dennis of Cornell to be about 50 million pages, and to make them available on the Internet. One of the key points made by Ewing that provoked a positive impact in the scientific community was the observation that Mathematics is, in fact, the ideal branch of learning to carry out an enterprise like the DML. It should be pointed out that the average life of a Mathematics article is approximately 10 years, whereas in other sciences it could be as little as 6 months [44].

The great aim of the DML is, therefore, to congregate all mathematical literature and to make it available through a centralized portal, accessible via the Internet, that coordinates

the various sources. To this effect, it shall be necessary to digitize all paper-based literature (a gigantic task that will only be possible if it is decentralized and performed with common or compatible standards), to create a system that integrates all literature in digital format and to establish connections with databases such as those of MathReviews (AMS) and Zentralblatt (EMS); in short, to adopt technical norms that adequately guarantee the correct dissemination, permanent update, long-term preservation and, of course, the integrity of copyright.

As recognized by the IMU, in particular by its current president John Ball, of the University of Oxford, the DML "is a vital effort for the mathematical community" and presently there exists a world-wide movement of coordination, but also of competition, in order to execute the multiple projects necessary to this collective enterprise. This process, necessarily complex and distributed by several partners, involves the production of digital documents, presuming that the standards are established for their technical specifications, for the metadata, for interoperability and interconnection, for the archive, update and maintenance, besides the problems of cost and economic models of sustainability, the question of intellectual property and copyright and, finally, the international management and coordination.




## 4 (Retro)Digitization

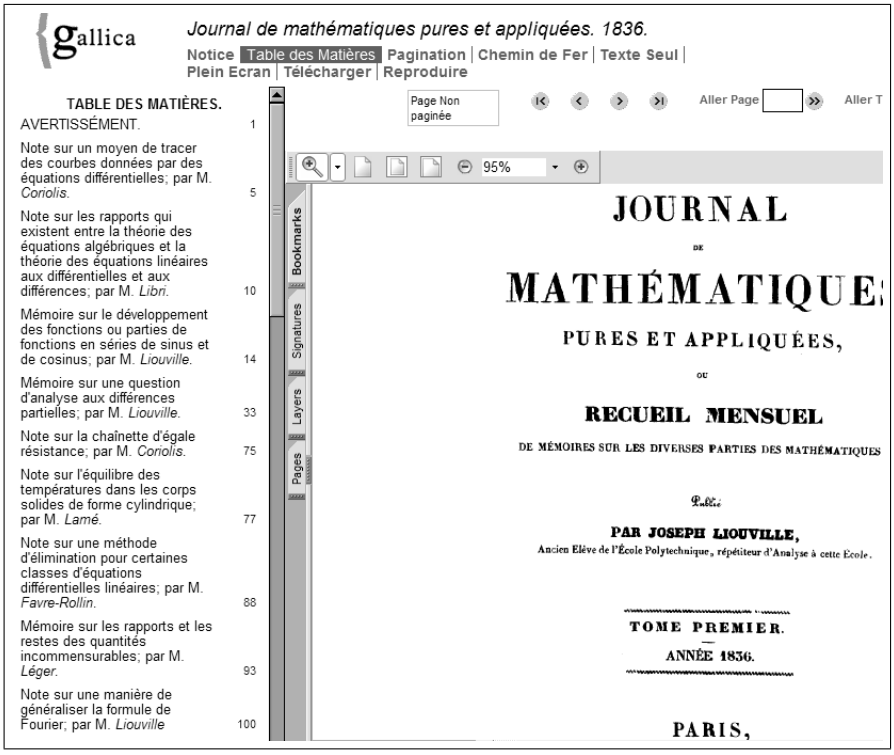
Retrodigitization is the name of the process that consists of creating digital copies of documents that exist only in paper format. One of the aims of the Digital Mathematics Library (DML) is to retrodigitize all the existing mathematical legacy that is not yet in digital format. The retrodigitization process is basically composed of two phases: obtaining a digital version of the paper document and structuring the acquired information in a useful and accessible database. In a first phase, an image is usually created for each page of the document and kept in PDF or DjVu format. After that, the metadata are created and text recognition software (OCR) is used to construct indexes that allow for electronic searching of the content. Financially the process is not expensive, considering that many projects of massive retrodigitization have

sent their documents to be digitized in countries where wages are very low, so the majority of costs correspond to the introduction of bibliographical data and equipment. The usual cost per page is on average 2 euros, 10% of which is the cost of the actual digitization.

In recent years, the idea of the DML has been consolidated through various projects and initiatives, so that the number of retrodigitized journals available on-line is already significant. In the electronic addresses collected by Ulf Rehmann [20], 1,874 books are indexed (at least 385,236 pages) and 145 journals (at least 2,942,143 pages); in some cases, complete access is paid for. This work is the result of regional projects financed by national agencies or Mathematics societies from countries such as Germany (EMANI, GDZ), France (GALLICA, NUMDAM), United States (JSTOR), Poland (BWM) or Portugal (PM-SPM in contribution with the National Library), as for example:

- BWM - Biblioteka Wirtualna Matematyki [17] with 3 journals (1888-1993), including Fundamenta Mathematicae (1920-1993) and the Studia Mathematica (1929 - 1964);
- GALLICA [24] with 230 Mathematics entries, including Liouville's Journal (1836-1932), the 4 volumes of the Histoire de Mathématiques by J.F. Montucla (1799-1802) and several works by Huygens, Euler, Fourier, Cauchy, Darboux and Jordan, among others;
- GDZ-Göttingen - Göttinger Digitalisierungszentrum [25] with 28 journals and several monographs (1777-1997), including the complete works of Gauss, Klein and Hilbert, the editions of 1898 and 1939 of the Encyklopädie der Mathematischen Wissenschaften mit Einschlussiher Anwendunge, and the Zentralblatt für Mathematik und ihre Grenzgebiete (1931-1978);
- JSTOR - Journal STORage [33] with 28 journals (1800-2002), including the Annals of Mathematics (1884-1997), the Journal of the AMS (1988-1997), the Journal of the Royal Statistical Society (1988-1998), the Mathematische Annalen (18691996), SIAM Review (1959-1997) and the Proceedings of the AMS (1950-1997);
- NUMDAM - NUMérisation de Documents Anciens Mathématiques [35] with 6 journals (1864-2000), including the complete text of the work Eléments de Géometrie Algébrique by A. Grothendieck and J. Dieudonné, the journals Annales de l'Institut Fourier (1949-1997) and Annales Scientifiques de l'École Normale Supérieure (1864-1997);
- PM-SPM - Portuguese project [38] of retrodigitization of the Portugaliæ Mathematica (1937-1993).

Despite the relevance of the amount of information on offer, the current implementation of the DML does not yet satisfy the premise of centralized access. The various projects operate in a relatively isolated and independent way, with the data being made available in very distinct forms. As observed by Ewing in [7], it would make sense to create a coordinating entity to define guidelines and technical practices to follow for all such projects. However, the ones that already exist continue to work without input from, or rule-setting by, any higher authority. Consequently, there are visible disparities between services and functionalities available to the user of each project. For example, GDZ is of free access but it does not perform OCR on the images, and so the search for keywords in the content of its documents is impossible. JSTOR, that allows an integrated search and makes available various functionalities such as obtaining documents in tiff, pdf and PostScript formats, is a commercial venture, charging




thousands of dollars for access. NUMDAM, which contains links to the MathSciNet and Zentralblatt MATH, is a free service based in the University of Grenoble and supported by the French CNRS, but, like the remaining projects, it was in the beginning limited to national journals. In the technical context, despite the inexistence of a definite and globally accepted norm, several entities publish recommendations on good practice to adopt in the digitization process, as for example Minerva eEurope [26] and the Committee of Electronic Information and Comunications of the International Mathematics Union (CEIC/IMU) [41]. On the other hand, Mathematics presents concrete difficulties resulting from the dispersion and diversity of documents. As mathematical journals were (and are) published variously by commercial publishing houses, university publishing companies, scientific societies, departments of mathematics, and even by groups of mathematicians, it thus becomes difficult to acquire/transport the original copies or to negotiate copyrights. Currently, digitization and the subsequent creation of digital libraries is so important for the preservation of the cultural and scientific heritage that in a document dated 28 April 2005 to the Presidency of the European Council, six Heads of State and Government defended the creation of a European virtual library. The Commission congratulated itself for this plan and will contribute to its accomplishment through the emblematic initiative called i2010 Digital Libraries [31].

## 5 Search and Metadata

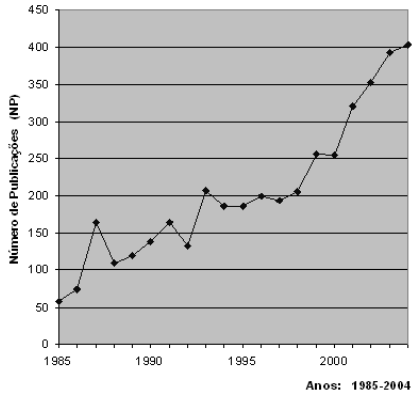
State of the art scientific activity depends heavily on the search and consultation of bibliographical material. However, the gigantic amount and the permanent growth of scientific literature constitute a difficulty for authors and appraisers in identifying and using all the information relevant to their work. According to the article [11], the most-cited articles are

those of easiest access (for example on-line). In particular, in Computer sciences, in a sample of 119,924 conference articles, the average number of off-line articles cited is 2.74, while the average number of on-line articles cited is 7.03. Note that, in this scientific area, conference articles carry more weight than journal articles, having an average acceptance of 10%. In Mathematics, no similar studies have been published, but they would probably follow this standard. Forms of global and integrated search are ever more important and necessary for the development and progress of scientific research.

### 5.1 Mathematical databases

Besides historical catalogs, such as the Jahrbuch über die Fortschritte der Mathematik (JFM) (1868-1944) and the Répertoire Bibliographique DES Sciences Mathématiques (1894-1912) [40], the current MATHDI (MATHematics DIdactics) [48], accessible through the EMIS, constitutes an on-line database of bibliography of mathematical education that is associated with the German journal ZDM (Zentralblatt für Didaktik DES Mathematik) and contains over 112,500 citations (1976-Jun/2005).

But the main support tools for current Mathematical research are the databases of Zentralblatt MATH and the MathSciNet, each one with around 2 million entries and with their own search engines in the respective on-line versions. These not only allow document searches, but also supply some statistics on publications. For example, to justify the passage of Portugal from Group I to Group II of the IMU, the Portuguese National Commission of Mathematics conducted a quantitative survey of publications referenced in MathReviews (Books, Journals and Proceedings), counting per year those publications with at least one Portuguese institution based author in the period from 1985 to 2004 (search "Inst Code=p- *" on the 26 of October, 2005). This is illustrated in the graph and table.




| Ano | NP | Ano | NP |
|---|---|---|---|
| 1985 | 58 | 1996 | 200 |
| 1986 | 75 | 1997 | 193 |
| 1987 | 164 | 1998 | 205 |
| 1988 | 109 | 1999 | 256 |
| 1989 | 120 | 2000 | 255 |
| 1990 | 138 | 2001 | 320 |
| 1991 | 165 | 2002 | 353 |
| 1992 | 132 | 2003 | 393 |
| 1993 | 207 | 2004 | 404 |
| 1994 | 187 | | |
| 1995 | 186 | **Total** | **4.120** |

### 5.2 Portuguese online library of knowledge (*b-on*)

The Online Library of Knowledge (b-on) [16] allows full access (complete text and unlimited access) for all members of the national academic scientific community, centres of research and development, public administration institutions, non-profit institutions and hospitals, to some

of the main sources of international scientific knowledge. Moreover, it allows a rationalization of costs, through a global negotiation with the scientific publishing companies. A year and a half after its creation, b-on, with reference to more than 16 thousand publications and 4 million visualizations of integral texts, has altered its portal to improve its usability, and has increased the available functionalities. These include interfaces adapted to the user's profile (beginner/regular/expert), parameterization of alert, and context help, among others. b-on is a federated search engine (based on the Metalib), used by the universities of Harvard, Stanford and several Finnish universities. The current number of adhering institutions is 69, having been 48 at start-up. Contents are distributed by areas: Art and Humanities (11.58%), Science and Technology (20.69%), Health Sciences (17.99%), Social Sciences (38. 52%) and Physics, Chemistry and Mathematics (11.22%); according to data supplied by the Agency for the Society of Knowledge (UMIC). The costs of its functioning have been estimated as 7.5 million Euros in 2004 and 10.8 million Euros in 2005, 56.6% of this being supported by the State and the remainder by the adhering institutions.

The importance of b-on is recognized internationally. Besides the Economist Intelligence Unit having cited it as a model in a report on cases of success in the use of structural funds for the countries of the enlargement, this initiative has been presented in several international forums. The interim evaluation report for the Lisbon Strategy [23] by the High Level Group, presided over by Wim Kok, recognizes the importance of the Society of Knowledge in the design to transform Europe into the most competitive space in the world.

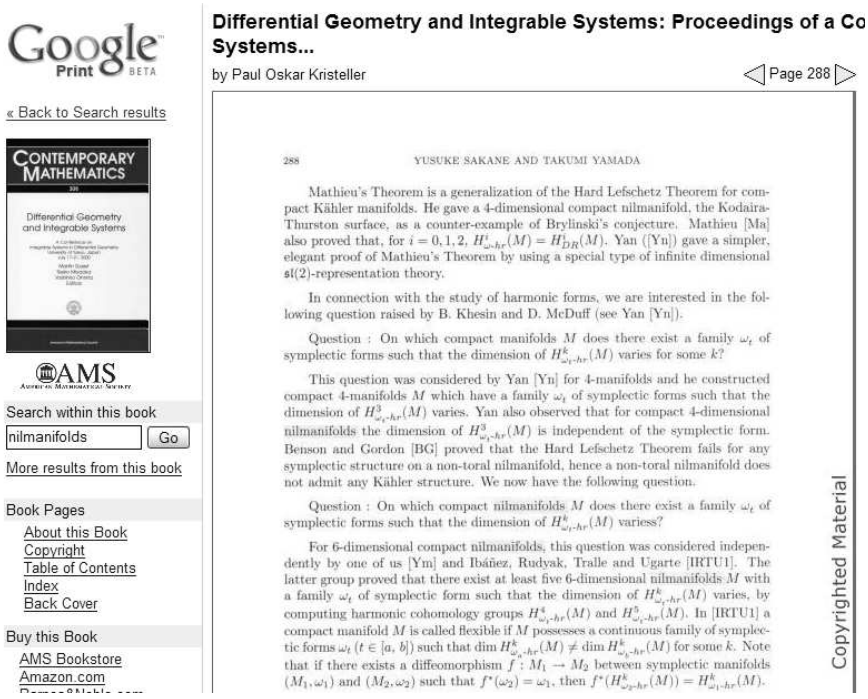
However, despite the enormous progress that recourse to b-on represents for mathematicians in Portugal, important gaps still exist in the area of Mathematics, such as access to JSTOR, or to the journals of the EMS and the AMS and, in particular, the MathReviews. On the other hand, the impact of b-on in the classical libraries has yet to be evaluated, as well as the possible future consequences of the lack of paper versions of the main journals in case of insufficient financing or discontinuation of some on-line services.

### 5.3 Google scholar and Google print

Not long ago, Google made available the services "Google Scholar" (GScholar) [28] and "Google Print" (GPrint) [27], still in an experimental phase. GScholar is a bibliographic search mechanism, counting, currently, with more than 500 million references to scientific documents. Among other interesting characteristics, it allows for a search that includes articles, summaries and citations, locates a document in a library (when not available on-line), and allows the grouping of several versions of a work (for example, pre-print, journal article). GPrint, in turn, permits the search and visualization of book pages with the search keywords appearing in color, transits to a shopping portal on the web, or verifies its existence in libraries belonging to the WorldCat [47]. The number of pages acessible for visualization depends on the status of the book: in free access, the whole book is obtainable; conditioned access permits visualization of two pages either side; unavailable access offers only bibliographical references. Google encourages publishers of all disciplines to make available references to their material [43]. Total control is delegated to the publishers concerning publication policies and compels them to make available at least a summary for each article. GScholar promises to deal with all violations of copyright that are communicated to it, as determined in the Digital Millennium Copyright Act [19].

Google also encourages libraries to use GScholar [42]. This service allows members of a University to find through GScholar a link to the catalogue entry of a document in its library,

whenever this is available. The system also connects to WorldCat, permitting for example, a search for the closest library where the document is available. In GScholar it is possible to personalize some parameters, for example, to configure a default localization; allowing a user outside the University Campus to accomplish searches as though on Campus. The authentication of users and control of access to documents are the entire responsibility of the libraries.




These services are a sample of what will be possible to access on the Internet when the millions of books from five great libraries (of the Universities of Oxford, Harvard, Standford, Michigan and of the New York Public Library), being digitalized by Google, start to be available in its database.

## 5.4 Metadata

Metadata refers to the set of additional data that allows one to identify, to classify, to search and to transform concrete data (documents), being elements of a database or catalog. According to one of the collaborators of the NUMDAM in [3], metadata exist in three distinct categories: external, which are those visible to the users (author, title, etc.); internal, being those necessary for correct system operation (identification, placement and date of the file, etc.); and exportable, which serve essentially to make our data recognizable by others. People interested in using metadata are generally researchers, who wish to find information relevant to their work; publishing companies, to divulge their product; digitization projects/centres, that need to organize their material; reference journals/catalogs, to catalog digitized material; libraries, to make their content available; archives, concerned with preservation. In short, it may be said that the metadata serve the purposes of research, compatibility, identification, management and preservation.

While the choice of internal and external metadata depends entirely on each project, in the

interests of compatibility some standards should be adopted, such as the OAI Protocol and Dublin Core Metadata. In fact, institutions dedicated to the development of metadata norms already exist: PREMIS [39] (durable preservation) or OCLC [36] (digital catalog register). It should be noted that the problems to be solved are innumerable. These include, among others, variations in the name of an author (for example, John Miller or J. Miller); the order of presentation (for example, John Miller or Miller, John) that directly affects the process of ordering (for example, should Andrew Lloyd Webber be placed in the A, L or W); and problems of transliteration (for example: Cebycev or Tschebyscheff). On the purely technical side, solutions do exist (for example, Authority files, PURL, DOI, URN, etc) for such problems, however, more coordination is needed between their technical availability and their uses.

# 6 Final Considerations

The current publication process is in crisis and is undergoing profound transformation. Its economic models are in question and there is controversy between the traditional reader-payer model and the author-payer model; this latter is on the rise as it is the support for Open Access (OA).

The main idea of OA is to break the trade barriers to literature, on the one hand allowing free access to the information and, on the other, minimizing the costs of dissemination by delegating them to the responsibility of the authors or of public or private institutions. The declarations of Budapest [18] and Berlin [15] are well known, and as can be seen in the Directory of Open Access Journals [21] more than 1800 journals currently exist only in OA, of which more than 60 are in the area of Mathematics. Guidelines for the sustainability of journals in the OA regime [30, 29, 37] can be found on the Internet.

One of the strongest arguments for OA is based on the observation that research is financed by governmental institutions and, therefore, society should have the right of access to the scientific content paid for by their taxes. Another argument points out that public money should be used for the common good, not for the benefit of private publishing companies that acquire the copyright and profit from it, without paying the authors. This idea underpins the Declaration on Access to Research Data from Public Funding, an agreement signed under the auspices of the OECD (Organization for Economic Co-operation and Development), to which Portugal has belonged since the 4th of August 1961.

This model is, however, far from finding consensus not only between enterprise specialists, like J.J. Esposito [4], who anticipates a general increase in the cost of publications for research in the OA model, but also between members of the mathematical community. John Ewing [5], aware that 75% of the annual income of the AMS comes from its publications, considers that the solution is not in free access to publications, but in creating a demand to lower the prices of journal subscriptions for librarians and academics. J. Ball and J. Borwein [2] affirm that the OA model is ill- adapted to Mathematics, for it might place its publications at the mercy of the administrators of the universities and other institutions, as well as at the mercy of other, richer, disciplines. This may cause discrimination between Mathematicians according as to whether they can fund their publications in cheaper or more expensive journals.

In the current context, it seems to us to be of the greatest importance that the mathematical community participates in the discussion with the various players in the process in order to guarantee the quality, access to and the preservation of Mathematics. To this end, the Portuguese Society of Mathematics supports the holding of the Conference "Communi-

cating Mathematics in the Digital Era", between the 15th and the 18th of August 2006 in the University of Aveiro (http://www.cmde2006.org). This event aims at fostering the debate between the interested parties on issues underlying Digital Libraries, Electronic Mathematics Resources, Electronic Publication, Digitization, OA Initiatives, Copyrights, Metadata and Rules in the specific case of Mathematics.

## References


[1] Birman, Joan S. *Scientific Publishing: A Mathematician's Viewpoint*, Notices of the AMS 47 n.7 (2000), 770–774.

[2] Ball, J. and Borwein, J. *ACCESS: Who gets what access, when and how?*, MSRI workshop, Berkeley (2005).

[3] Bouche, T. *Digitisation and metadata overview*, MSRI workshop, Berkeley (2005).

[4] Esposito, J. J. *The devil you don't know: The unexpected future of Open Access publishing*, First Monday 9 n.8 (August 2004).

[5] Ewing, John H. *Misdirection*, in: Proceedings of the ECM 4 satellite conference publishing at KTH, Stockholm, Sweden (June 2004).

[6] Ewing, John H.*Predicting the Future of Scholarly Publishing*, in Electronic Information and Communication in Mathematics, F. Bai and B. Wegner (Eds) LNCS#27300, Springer (2003), 52–58.

[7] Ewing, John H. *Twenty centuries of mathematics: Digitizing and disseminating the past mathematical literature*, Notices (2002), 771–777.

[8] Jost, M. and Wegner, B. *EMIS 2001 - A World-Wide Cooperation for Communicating Mathematics Online*, in: Electronic Information and Communication in Mathematics, F. Bai and B. Wegner (Eds) LNCS#2730, Springer (2003), 87–94.

[9] Knuth, Donald E. *Letter to the Editorial Board*, Journal of Algorihms (October 2003).

[10] Knuth, Donald E. *Mathematical tyupography*, Bull. Amer. Math. Soc. (N.S) 1 n.2 (1979), 337–272.

[11] Lawrence, Steve. *Free online availability substancially increases a paper's impact*, Nature - webdebates (2005).

[12] Rodrigues, J. F. *Revistas Matemáticas Portuguesas*, Boletim da SPM 50 (2004), 19–36.

[13] Wegner, B. *EMANI - A Project for the Long-Term Preservation of Electronic Publications in Mathematics Online*, in: Electronic Information and Communication in Mathematics, F. Bai and B. Wegner (Eds) LNCS#2730, Springer (2003), 178–188.

[14] *arXiv.org e-Print archive*, [webpage] URL: http://arxiv.org/.

[15] *Berlim Open Acces Declaration*, [webpage] (2003) URL: http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html.

[16] *Biblioteca de Conhecimento Online (b-on)*, [webpage] URL: http://www.b-on.pt/.

[17] *Biblioteka Wirtualna Matematyki*, [webpage] URL: http://matwbn.icm.edu.pl/.

[18] *Budapest Open Acces Initiative*, [webpage] (2001) URL: http://www.soros.org/openaccess/.

[19] *Digital Millennium Copyright Act*, [doc. web] URL: http://www.copyright.gov/legislation-/dmca.pdf.

[20] *DML: Ulf Rehmann's extensive and detailed listing*, [webpage] URL: http://www.mathematik-.uni-bielefeld.de/řehmann/DML/dml_links.html.

[21] *DOAJ - Directory of Open Access Journals*, [webpage] URL: http://www.doaj.org/.

[22] *EMANI - Electronic Mathematical Archiving Network Initiative*, EMIS/Springer/SUB Göttingen-/Cornell Library/Tsinghua Library, [webpage] URL: http://www.emani.org/.

[23] *Facing the challenge: The Lisbon strategy for growth and Employment*, High Level Group, [webpage] URL: http://www.eiro.eurofound.eu.int/2004/12/feature/eu0412205f.html.

[24] *Gallica - Bibliothèque numérique de la Bibliothèque Nationale de France*, [webpage] URL: http://gallica.bnf.fr/.

[25] *GDZ-Göttingen (Göttinger Digitalisierungszentrum)*, [webpage] URL: http://www.sub.uni-goettingen.de/gdz/.

[26] *Good practices in digitisation*, Minerva eEurope, [webpage] URL: http://www.minervaeurope.org-/listgoodpract.htm.

[27] *Google Print (beta)*, Google, [webpage] URL: http://print.google.com/.

[28] *Google Scholar (beta)*, Google, [webpage] URL: http://scholar.google.com/.

[29] *Guide to business planning for converting a subscription-based journal to OA*, BOAI, [webpage] URL: http://www.soros.org/openaccess/oajguides/html/business_converting.htm.

[30] *Guide to business planning for launching a new OA journal*, BOAI, [webpage] URL: http://www-.soros.org/openaccess/oajguides/html/business_planning.htm.

[31] *i2010 Digital Libraries*, Europe's Information Society, [doc. web] URL: http://europa.eu.int-/information_society/activities/digital_libraries/doc/pt_comm_digital_libraries.pdf.

[32] *Jahrbuch Search*, [webpage] URL: http://www.emis.de/MATH/JFM/.

[33] *JSTOR - Journal STORage*, [webpage] URL: http://www.jstor.org/.

[34] *MPRESS – The Mathematics Preprint Search System*, EMS, [webpage] URL: http://mathnet.preprints.org/.

[35] *NUMDAM - NUMérisation de Documents Anciens Mathématiques*, [webpage] URL: http://www-.numdam.org/.

[36] *OCLC: a worldwide library cooperative project*, [webpage] URL: http://www.oclc.org-/digitalpreservation/why/digitalregistry/.

[37] *Plublishing open-access journals*, PLoS, [doc. web] URL: http://www.plos.org/downloads-/oa_whitepaper.pdf.

[38] *Portugaliæ Mathematica Retrodigitalizada pela Biblioteca Nacional*, [webpage] URL: http://purl.pt/404/.

[39] *PREMIS - PREservation Metadata: Implementation Strategies Working Group*, [webpage] URL: http://www.oclc.org/research/projects/pmwg/.

[40] *Répertoire Bibliographique des Sciences Mathématiques*, [webpage] URL: http://math-doc.ujf-grenoble.fr/RBSM/.

[41] *Some Best Practices for Retrodigitization*, Committee for Electronic Information and Communication of the International Mathematical Union, [doc. web] (June 2005) URL: http://www.ceic.math.ca/Publications/retro_bestpractices.pdf.

[42] *Support for Libraries*, Google, [webpage] URL: http://scholar.google.com/scholar/libraries.html.

[43] *Support for Scholarly Publishers*, Google, [webpage] URL: http://scholar.google.com/scholar-/publishers.html.

[44] *Science Citation Index*, [webpage] URL: http://www.isinet.com/.

[45] *The Electronic Library of Mathematics*, EMIS, [webpage] URL: http://www.emis.de/ELibM.html.

[46] *WDML - World Digital Mathematics Library*, [webpage] URL: http://wdml.org/.

[47] *WorldCat*, [webpage] URL: http://www.oclc.org/worldcat/.