

Some thoughts
on the near-future
Digital Mathematics Library

Thierry Bouche

Cellule MathDoc & institut Fourier, Grenoble

Launching DML-CZ
Czech Digital Mathematics Library
Charles university, Prague, 11th June 2007

NUMDAM



cedram

Outline

- 1 The mathematical literature
- 2 The electronic mathematical literature
- 3 The Digital Mathematical Library
- 4 Implementation
- 5 Conclusions

The mathematical literature

Specificities

- Mathematical literature never becomes obsolete
- It's valid only as a *whole*, building a wide network of references
- It's useful to other sciences in an *asynchronous* fashion
- It must be carefully archived, indexed and preserved
- It must be accessible over the long term

The mathematical literature

Milestones

- 1665** Birth of scientific journals (*Journal des sçavans*, *Philosophical transactions*)
- 1800** About 200 journals where math articles are published
- 1810** First math-only journal (*Annales de mathématiques pures et appliquées*, aka *Annales de Gergonne*)
- 1850** About 1000 mathematical research articles published each year
- 1950** About 6000 mathematical research articles published each year

The mathematical literature

Milestones

1978-1986 T_EX

1992 arXiv, math preprints (Physics: 1991)

1994 First non specialized math-only electronic journal
it's free
(*New York Journal of Mathematics*)

1995 JSTOR digitises 6 English speaking math journals
(400 000 pages)

2000 100 000 articles considered, 75 000 reviewed *Math. Reviews*
or *Zentralblatt MATH*
600 cover-to-cover journals, 1500-2000 serials scanned

2008 4.5 million pages digitised, 65% of core journals available
digitally?

The mathematical literature

The impossible catalogue

- 1868 *Jahrbuch über die Fortschritte der Mathematik*
- 1894 *Répertoire bibliographique des sciences mathématiques*
("valuable" references from 19th century)
- 1931 *Zentralblatt für Mathematik und ihre Grenzgebiete*
- 1940 *Mathematical reviews*
- 1970 Classification AMS
- 1990 Electronic versions (MathSci Disc, CompactMath)
and online access (telnet. . .) MSC
- 1995 Web access (MathSciNet, ZMATH)
- 2000 Links to original texts
- 2002 Bibliographies, backward links
- 2004 μ -DML :-)
- 2009 EuDML?

The mathematical literature

Time scale

- Instant preprint circulation (labs, arXiv, email, home pages)
- Actual publication delayed 1-2 years
- About 50% citations in today's bibliographies are more than 10 years old
- About 25% citations in today's bibliographies are more than 20 years old
- Among the 100 most cited items in MR biblios, 96 are books (88 in NUMDAM biblios)

The mathematical E-literature

Interrogations

- How open archives and formal journals fit together?
- Reliability of digital libraries?
- Who pays for electronic-only free access journals?
- Does the author-pay model help improve the quality?
- Does the author-pay model help lower global costs?
- Will independent publishers survive?
- Who will maintain the files after their commercial life?

The mathematical E-literature

Wishes (IMU 2002, 2005)

- Free metadata and navigation
- Eventual open access (moving wall)
- No long-term economic, legal, technical barriers
- No dependance upon viability of any economic agent
- A universal reference digital mathematical library (DML)

The mathematical E-literature

Content providers

- **Gallica** retrodigitised, public domain (old), free, French speaking
- **GDZ** retrodigitised, free, not only German
- **NUMDAM/CEDRAM** retrodigitised/publishing platform, moving wall, not only French
- **JSTOR** retrodigitised, not-for-profit, English only, (expensive) subscription library service
- **project Euclid** retrodigitised and publishing platform, not-for-profit
- **Oxford University Press** retrodigitised/publishing platform, no moving wall, English only
- **Elsevier** publishing platform, retrodigitised content as one optionnal package
- **Springer** publishing platform, retrodigitised content as one optionnal package (English only)

The mathematical E-literature

Journal accessibility report

- Acta math. Mittag-Leffler[†] (1882-2005) ; Springer (1882-1997), Springer (1997-)
- Ann. Math. JSTOR (1884-2001), Euclid (2001-)
- Bull. LMS OUP (1865-)
- CRAS Gallica (1835-1965) ; Elsevier (1997-)
- Crelle GDZ (1826-1997) ; Walter de Gruyter (1999-)
- Duke Math. J. Euclid (1935-1999), Euclid (2000-)
- Liouville Gallica (1836-1935) ; Elsevier (1997-)
- Math. Ann. GDZ (1869-1996) ; Springer (1869-1997), Springer (1997-)
- Pacific J. Math. Euclid (1951-)
- Théor. nombres Bordeaux Séminaire : GDZ (1972-1988) ; Journal :
NUMDAM (1989-2003) ; ELibM (1994-2005) ; CEDRAM
(1989-2006)

Digital Mathematical Library

The vision

“In light of mathematicians’ reliance on their discipline’s rich published heritage and the key role of mathematics in enabling other scientific disciplines, the Digital Mathematics Library strives to make the entirety of past mathematics scholarship available online, at reasonable cost, in the form of an authoritative and enduring digital collection, developed and curated by a network of institutions.”

(Cornell NSF project 2002,
endorsed by IMU 2006)

A digital library?

Context

- Libraries buy and store publisher's (paper) production
- They preserve it and provide access to their patrons
- Why should this change because of a format move?
- Will the mathematical knowledge remain part of our common, freely accessible heritage?
- Or is it going to be confiscated by private interests?

A digital library?

Traditional components—digital counterparts

Selection Selecting collections by subject, document type. . .

Acquisition Retrospective digitisation *and*
ingesting current production

Cataloguing Capture, produce, import, enhance metadata

Archiving Collections, file names, identifiers

Preservation Hardware maintenance, emulation, management. . .

Access Easy access, file conversions, interfaces maintenance

Implementation

“The Digital Mathematics Library strives to make

- ~~the entirety of past mathematics scholarship~~
 - **available online, and preserved offline**
 - **at reasonable cost,**
 - in the form of an **authoritative** and **enduring** digital collection,
 - developed and curated by a network of **institutions.**”
- + updated continuously with publisher supplied new content
- + with sophisticated search interfaces and interoperability services

Implementation

“Available online”

Collections should be

- Cared for and accessed locally
(digital files preserved physically at each participating institution:
not a virtual library)
- Usable, accessible globally
(though a virtual union catalogue, and metadata sharing
with cooperating services like reviewing databases
or more general search engines, portals, etc.)

Implementation

“Reasonable cost”

Business model should be modelled on the current library system:

- Free to patrons
- Free to anyone would be appreciated,
but not at the risk of losing the sustainability
or reliability of the system

Implementation Institutions

Should be

- Scientifically reliable (authoritative)
- Long lasting (enduring)
- Not-for-profit
- Committed to the effort
(digital legal deposit for mathematical content?)

Conclusions

- Mathematicians are waiting for a reference digital library
- It should be a distributed collection of physical archives
- It has to be a public service (at least not-for-profit)
- lasting for ever
- But it should keep current!
- Immediate free access is *not* mandatory
- Eventual open access *is* mandatory
- The French NUMDAM+CEDRAM was a preview
- DML-CZ is a promising new one!
- The foreseen EuDML could be the first large-scale implementation!