G. Törner
Gerhard-Mercator-University of Duisburg
Faculty of Science - Department of Mathematics

Math*Diss* International

On the one hand, dissertations are the most important documents of personal qualification for young scientists. On the other hand, dissertations also represent a significant part of the current status of scientific research in the literature. They include recent research findings and innovative methodical approaches and are therefore especially important for the scientific public in the field. Here immediate, current accessibility is critical.
In view of the different demands posed on dissertations by societies and scientists in any field, global access to the documents from a single source remains impossible. A foundation for the handling of electronic dissertations in Germany has been laid in the project *Dissertations Online*[2]. However, in the field of mathematics, additional points stemming from the relative frequent inclusion of formulas in mathematic dissertations must be considered.
Within the scope of the project **Math*Diss* International**, a permanent international online full-text document server for mathematical dissertations will be established. In this connection, questions concerning an online presentation of the documents and the problems of long-term archiving (from TeX resp. LaTeX documents) will be considered. They include the question of how to homogenize such files in order to enable their later conversion into programming languages following XML.
Furthermore, the expansion of research possibilities using online documents is being planned. Providing access to the tables of contents, lists of tables and illustrations and bibliographies on the LaTeX level is of top priority. Because of the structure of mathematical documents written in LaTeX we have a lot of high quality information which gathers dust in the archives without being used for the retrieval of scientific documents. This situation should be changed and it could be changed because LaTeX has become a widely accepted tool in mathematical literature.
In the course of this lecture we should like to demonstrate the idea and the subject specific background of Math*Diss* International.
In the later presentation *MathDiss Service* by Th. Fischer the tools and the practical part of our project will be represented.

[2] http://www.dissonline.de

Math*Diss* International

sponsored dy

Prof. Dr. Günter Törner

Gerhard-Mercator-Universität Duisburg

Content

- **The ETD-Scene**
- **Features and characteristics of mathematical literature**
- **Objectives of Math*Diss* International**

Some observations on ETDs

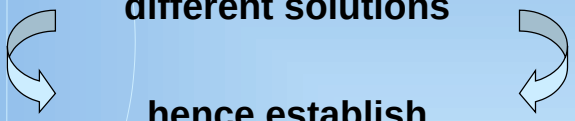
Problems of

- **accessing and viewing**
  - Full text often not available or only available by paying charges
- **formats**
  - The unsolved problem of the formats
    - Presentation format
      PDF does it, but not in all cases
    - Archiving format
      PDF is not adequate.
- **retrieval**
  - often restricted search in only a few basic fields (Author, Title, Keywords)
  - often lack of content-information
  - often no standard language for the information
  - often no subject specific information

The large spectrum of EDTs... (DDC)

- Philosophy
- Religion
- Social Sciences
- Languages
- Pure Sciences
  - Math
  - Astronomy
  - Physics
  - Chemistry
  - Earth Science
  - Paleontology
  - Life Sciences
  - Botany
  - Zoology
- Applied Sciences & Technology
- Arts
- Literature
- History & Geography
- General

ETDs in different disciplines require different solutions



hence establish the Subject Specific Services (SSS)

ETDs in Mathematics

General Requirements
and
Subject Specific Features

Mathematics ... as a well-organized bibliographical subject

There exist three independent databases covering a major part of mathematical literature
(more than 30.000 new articles per year)

- Mathematical Reviews - MR
  (run by American Mathematical Society - AMS)
  Online database
- Zentralblatt für Mathematik - Zbl
  (run by European Mathematical Society - EMS)
  Online database
- Russian Reviewing Journal (Izvestija Mathematics)

Mathematics ... as a research discipline

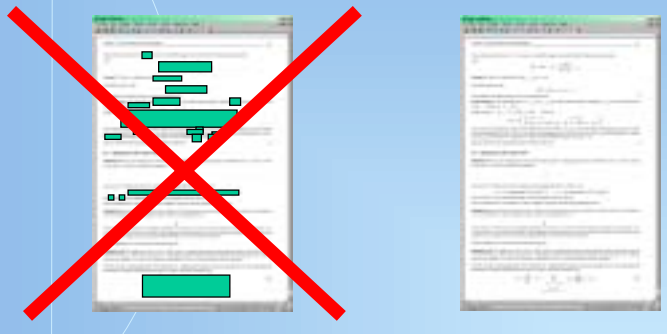
- There exists a world-wide accepted classification system:
  MSC – *Mathematical Subject Classification*
- MSC is continuously updated by the International Mathematical Union (IMU), the international panel for all learned societies of mathematics; actual version MSC 2000.
- MSC identifies nearly 80 subdisciplines within math; each subdiscipline itself is then covered by a further three-steps-classification system using digits and letters.
- MSC provides the scientists with a large list of keywords.

Mathematics ... as the language of its documents

A mathematical document…

- is more than only standard-text (English, Spanish, German,...) with graphics and layout elements.

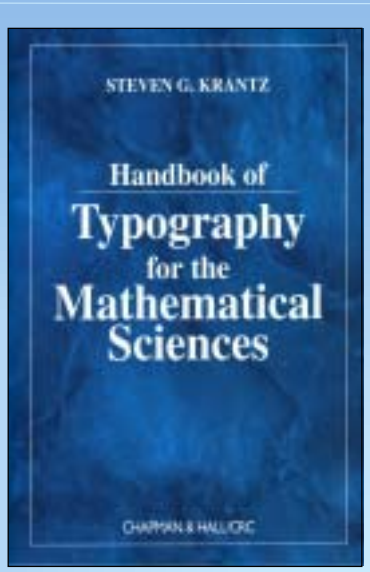
Paradigm of a mathematical text



Mathematics ... as the language of its documents

A mathematical document…

- is more than only standard-text (English, Spanish, German,...) with graphics and layout elements.
- is using an internationally standardized notation and well-defined typography



Mathematics ... as the language of its documents

A mathematical document…

- is more than only standard-text (English, Spanish, German,...) with graphics and layout elements.
- is using an internationally standardized notation.
- is a “multi-lingual” text; each mathematical discipline has its own “letters” additionally to some core alphabet of mathematics.

Mathematics possesses an adequate, high professional, sophisticated and world-wide accepted type-writing system (TeX, LaTeX, designed by D. Knuth), also used in other disciplines like physics, chemistry and engineering.

Mathematics ... as the subject field

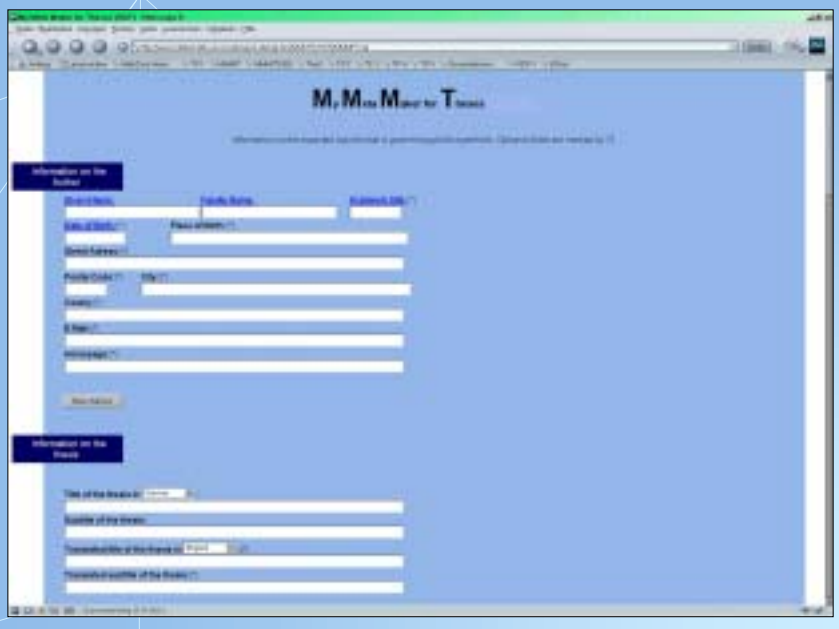
- Full-text search?
- Search for or within formulas?
  - Examples
    - exchangeable $\Leftrightarrow$ interchangeable $\Leftrightarrow$ permutable
    - $(X_n)_{n\in I} \Leftrightarrow (Y_k)_{k\in J}$

    Makes no sense!
- Conclusia:
  - To gain information it is sufficient to have
    - MSC-coordinates
    - table of contents
    - bibliography
    - index
    - table of figures
    - …

Objectives of Math*Diss* International

- Combine MetaData and Meta-Information
  in accordance to
  - Dissertation Online (German National Library)
  - NDLTD
- Integrate TeX / LaTeX-functions for gaining retrieval data
- Use adequate, high professional and sophisticated type-writing system, which accommodate the needs of typography and the technical possibilities.

Math*Diss* Metamaker



Math*Diss* - Retrieval

- Math*Diss* retrieval interface in Göttingen
- Integration into international systems (NDLTD, etc.)
- Integration in national systems (TheO, etc.)
- Development of an Allegro-Database containing
  - Metadata
  - Structural data
  - Textual data
- Short Information
  - Abstracts
  - Table of contents and bibliography

Math*Diss* - Files

- LaTeX file(s) (ASCII text)
- Normed input-files referred to Dante-Server (LaTeX-User group), *otherwise*
- Author-designed input-files (optional)
- Metadata provided by the
  - Author
  - Library
- Meta-information as automatically extracted from the LaTeX-files

International Cooperation

- Official Member of NDLTD
- Partner of OAI by the University Library of Duisburg
- Cooperation with Zentralblatt Math:
  every Math*Diss* Dissertation will be refereed in Zentralblatt
- Cooperation with the
  “Österreichische Dissertationsdatenbank”
- Cooperation with
  Koordinationsstelle Dissertation Online (DDB)
- …

Further Information

- **Further Presentation**

  Math*Diss* Service
  Th. Fischer (4 p.m.; this afternoon)

- **Project Homepage**

  http://www.ub.uni-duisburg.de/mathdiss/

- **Email**

  Günter Törner
  toerner@math.uni-duisburg.de

  Thorsten Bahne
  bahne@math.uni-duisburg.de