HOW TO BUILD A PROTOTYPE FOR A DISTRIBUTED DIGITAL MATHEMATICS ARCHIVE LIBRARY

GERHARD O. MICHLER

Abstract. The "Electronic Mathematics Archiving Network Initiative" (EMANI) recently announced by Springer Verlag, Tsinghua University Library (China), Göttingen State and University Library and Cornell University aims to insure the preservation and dissemination of mathematical information for future generations. In order to use the resulting distributed digital mathematical archive the mathematicians (and other scientists) have to be equipped with freely available readers which enable them to retrieve review articles on cited papers in a downloaded digital article from "Zentralblatt für Mathematik" or "Mathematical Reviews". Using then the permanent links between the review articles and their digital full texts on remote servers the scientists can also retrieve the digital full text of the cited articles. Such readers can only work if the scientific documents in the digital libraries have been provided with links between them and the review articles.

In this article the different steps to solve these technical problems will be discussed. Methods will be explained which allow to build a prototype of a distributed searchable, linked mathematical digital archive library.

1. Introduction

Mathematical research is so demanding that the complete proofs of very classical theorems like the odd order theorem of Feit and Thompson are still not part of the textbook literature. Therefore modern authors have to quote recent and old mathematical articles at the same time. On the other hand the mathematical literature has become so vast that an individual mathematician has a very hard time to get a survey about the existing literature even in his or her own special area of research. It is therefore an important issue for the mathematical community to use the present computer and network technology for building a distributed searchable and linked mathematics archive library.

The main research problems to be solved for achieving that goal are:

a) Extend the existing programs described in the author's article [11] for the recognition of cited articles in a reference list of a recent electronically published or retrodigitized back issue of a mathematical journal so that they can be applied automatically to many different journals.
b) Develop new linking techniques which automatically establish links from a cited article in a digital journal to its review article in "Mathematical Reviews" or "Zentralblatt für Mathematik" using only the bibliographic data produced by the recognition programs a).

The solutions of these two major problems suffice because "Zentralblatt für Mathematik" and "Mathematical Reviews" have begun to produce permanent links between their review articles and the digital full texts of the original mathematical articles either published recently in electronic form or retrodigitized like the back issues of the journals contained in JSTOR [7]. In her recent article [9], the executive editor of "Mathematical Reviews" writes: " The project of gathering and adding reference lists (together with the links, generated by a sophisticated matching algorithm) to the MR Database began in early 2001. Initially, lists from sixty-five journals are being added, for published issues from 2000 on. New issues of these journals are treated on receipt at the MR office on an ongoing basis. Later, other journals and possibly earlier time periods also will be covered. Thus, the numbers of both the MR Database items with full reference lists and the corresponding forward citations will grow over time, making MathSciNet an increasingly powerful tool with which to explore the interconnections among the literature. The addition of reference lists and the associated forward citations represent the most significant change in the structure and content of the MR Database in the over sixty years since it was founded."

In view of the importance of the contents of the databases of "Zentralblatt für Mathematik" and "Mathematical Reviews" for all mathematicians in the world, it is very important to produce software and readers which help the scientists to get instant links from downloaded digital articles to the review articles on their cited papers contained in their references, and from there to the digital full text of the cited articles whenever available by means of a permanent link.

In section 2.1 some methods will be described which helped to solve problem a) in the author's retrodigitization project "Archiv der Mathematik". More detailed descriptions are given in the author's articles [10] and [11]. In the meantime these programs have been extended to deal with several mathematical journals, which may even have no systematic ordering of the reference lists of cited articles like the ones of "Commentarii Mathematici Helvetici". They are described in section 2.2.

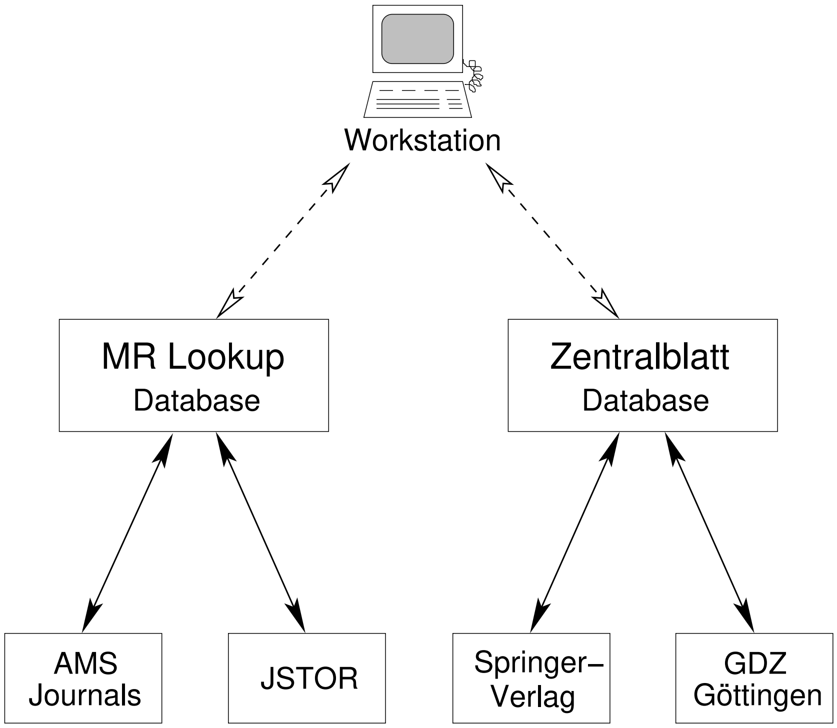
In both retrodigitization projects a printed article is scanned with 600 dots per inch. Then the commercial OCR system Fine Reader [3] is used to produce an ASCII file of the ordinary text part of the scanned page by separating the mathematical formulas and diagrams. Applying then new programs an Xdoc file of the ordinary text is produced, which also contains coordinates for each word. These coordinates are important for searching in the ordinary text and viewing the result on screen by means of the multivalent document system (MVD) of T. Phelps [16]. Coordinates are also necessary to add the URLs of a cited digital article at the correct place to an item of the list of references of a scanned back issue or an electronically published recent issue of a mathematical journal, see section 3, where a solution of problem b) is described.

Therefore multilayer document formats have to be used in a distributed, searchable and linked digital mathematical archive library. In the author's opinion the "DjVu"-reader is the best available technical solution for viewing both recent digital articles published in PDF format and also retrodigitized back issues of a mathematical journal, see the end of section 3.2.7. Furthermore, it is freely available in the internet for all users.

However, on the side of the document servers of a digital library all retrodigitized articles have to be stored in the multilayer DjVu-format described in 3.2.7. The advantage is that all fundamental programs have already been developed, see section 2.1 and 2.2. In particular, the most difficult problem of the automatic production of the bibliographic data in XML format of all the cited journal articles belonging to the list of references of a downloaded article has been solved, see sectin 3.2.3. The XML-records of the cited articles are then taken as input for the new programs producing the links to the two important review journals "Zentralblatt für Mathematik" and "Mathematical Reviews", see section 3.2.4. From there one obtains the URLs of the full digital texts of the cited articles provided there is a permanent link between the digital full text and its review articles. In order to put a link of a cited article to its review article one has to add the URLs of the review to the area of the corresponding citation in the list of references. This is done by means of the annotation layer of the DjVu format of the downloaded article, see section 3.2.5. Furthermore, with the help of the librarians the quality of the multilayer documents contained in a distributed digital library can be improved continuously by extensions or correcting errors found by users. If multilayer formats have been produced of all pages of the retrodigitized articles, then it is possible to search for words, names and mathematical expressions in the whole document. This has been achieved by converting the Xdoc format of the ordinary text into its DjVu format, see section 3.2.6.

If the multilayer DjVu document format will be used in the distributed digital archive library of the EMANI initiative then all articles contained in the different digital libraries of the initiative or other publishers which have permanent links to "Zentralblatt für Mathematik" or "Mathematical Reviews" will have the following properties:

a) They are linkable to each other,
b) their full text is searchable,
c) their cited articles can automatically be retrieved and photocopied,
d) all digital journal articles are also linked with their reviews in "Zentralblatt für Mathematik" and "Mathematical Reviews".



By putting all the programs described in this article together it is possible to design a complete system which will automatically transform a printed mathematical article into its multilayer DjVu format, which then can be stored automatically in a distributed digital library.

2. Retrodigitization projects

Since the last six years most of the important mathematical journals have been published electronically. The DFG "Sammelstelle" of the State and University Library Göttingen has permanent subcriptions to more than 790 journals in mathematical sciences, of which only 228 journals do not exist in electronic form.

In 1995 the JSTOR organisation [7] began to retrodigitize more than 100 scholarly journals of science, humanities and liberal arts. The 22 mathematical journals retrodigitized by JSTOR contain a moving wall of 5 years between the latest retrodigitized and the most recent electronically published volumes. In mathematics JSTOR offers only images of the journals. However all 22 journals are completely retrodigitized from their first volume onwards. The Metadata for the journal and its contents, such as article title, authors, etc. are hand-typed to ensure accuracy.

We now list the mathematics, and statistics journals in JSTOR. The URL from which one can get them is:
`http://www.jstor.org/cgi-bin/jstor/listjournal?frame=noframe&config=jstor`

Mathematics
American Journal of Mathematics 1878-1995
American Mathematical Monthly 1894-1995
Annals of Mathematics 1884-1995
Analyst 1874-1883
Journal of Symbolic Logic 1936-1996
Journal of the American Mathematical Society 1988-1995
Mathematics of Computation 1960-1995
Mathematical Tables and Other Aids to Computation 1943-1959
Proceedings of the American Mathematical Society 1950-1995
SIAM Journal on Applied Mathematics 1966-1995
Journal of the Society for Industrial and Applied Mathematics 1953-1965
SIAM Journal on Numerical Analysis 1966-1995
Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis 1964-1965
SIAM Review 1959-1995
Transactions of the American Mathematical Society 1900-1995

Statistics
Annals of Applied Probability 1991-1995
Annals of Mathematical Statistics 1930-1972
Annals of Probability 1973-1995
Annals of Statistics 1973-1995
Applied Statistics 1952-1996
Biometrics 1947-1995
Biometrics Bulletin 1945-1946
Biometrika 1901-1995
Journal of the American Statistical Association 1922-1995
Quarterly Publication of the American Statistical Association 1920-1921
Publications of the American Statistical Association 1888-1919
Journal of the Royal Statistical Society. Series A (Statistics in Society) 1988-1996
Journal of the Royal Statistical Society. Series A (General) 1948-1987
Journal of the Royal Statistical Society 1887-1947
Journal of the Statistical Society of London 1838-1886
Journal of the Royal Statistical Society. Series B (Methodological) 1948-1996
Supplement to the Journal of the Royal Statistical Society 1934-1947
Statistical Science 1986-1995
Statistician 1962-1996
Incorporated Statistician 1950-1961

The Retrodigitization Center at Göttingen University retrodigitizes a large number of back issues of many important mathematical journals, especially the ones published by Springer Verlag in Heidelberg,
see `http://134.76.163.65:80/advanced_search.html`

Nachrichten von der Königl. Gesellschaft der Wissenschaften und der Georg-August-Universität zu Göttingen 1864 - 1893

Mathematische Annalen **1**(1869) - **119** (1943/44); **120**(1947/49) - **122** (1950/51)

Mathematische Zeitschrift **1**(1918) - **50**(1944); from **51**(1947/49) - onwards.

Inventiones mathematicae 1966-1970.

Whereas in all these retrodigitization projects onyl TIFF files and bibliographical data of printed articles have been produced, the following three projects initiated by the author contribute constructive answers to the problems of:

the automatic recognition of the bibliographical data of a printed article and its cited references,

the searchability within its ordinary text, and
linking it to other cited digital articles.

2.1 "Archiv der Mathematik"

Since 1997 the author's study group has been involved in several retrodigitization projects. As the author is an editor of "Archiv der Mathematik" he was able to get the support of its publisher "Birkhäuser Verlag" for starting the retrodigitization of 6 volumes of the mathematical journal "Archiv der Mathematik" published in the years 1993-1996. The results of this research project were demonstrated at the international conference "Retrodigitization of mathematical journals and their integration into searchable digital libraries", which took place at Essen University in August 2000. Detailed surveys of the retrodigitization methods developed whithin this project have been published by the author in his articles [10] and [11].

The main steps of this mathematical text recognition system are:

1. Recognition of the ordinary text of a scanned page of a mathematical article by means of the commercial OCR product FineReader [3] and additional special programs and databases. These programs produce the ordinary text in Xdoc format, which assigns to each word of the obtained ASCII-file also its coordinates on the scanned page, see [11], p. 118, 119.

2. Separation of the mathematical formulas and symbols from the ordinary text by means of a separate program, see [11], p. 121.

3. Recognizing the mathematical symbols and formulas by means of Professor Okamoto's OCR-program EXP ([8] [14] [15]), which also produces a latex file of the recognized formulas, see [11], p. 120.

4. Reproducing the original coordinates of the beginning and the end of each mathematical formula box, and for attaching these coordinates to the corresponding latex files by another special program. A new program developed by Professor Waki during his stay at Essen University puts these mathematical data at the original place of the corresponding formula into the Xdoc file produced in step 1 by squeezing the scale of the latex files of the mathematical formulas, see [11], p.121.

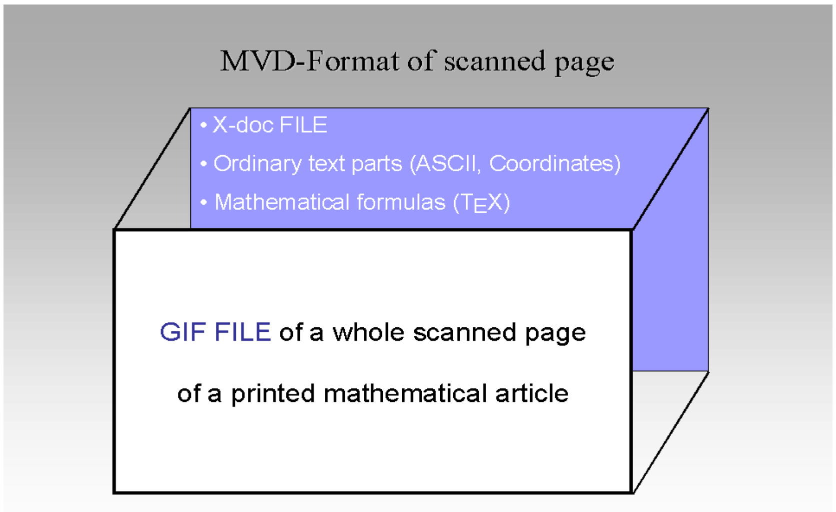
5. Recognition of the layout of the first page of an article to get its bibliographic data, see [11], p. 122.

6. Recognizing the references of a downloaded article to get the bibliographic data of each cited article of the list of references. With the obtained XML-data it is possible to retrieve the existing digital full texts of its references which are available in a digital library, see [10].

7. Using the multivalent document system (MVD) of T.A. Phelps [16], it is possible to combine the constructed two layers of a scanned page and to search for words and mathematical expressions of its contents, where the first layer consists of the GIF file of the scanned page, and the second layer consists of its Xdoc-file.

The multivalent document system designed by T.A. Phelps and R. Wilensky [17] has been produced by T.A. Phelps [16]. It is a general paradigm that regards complex documents as multivalent documents comprising multiple layers of distinct but intimately related content. So called behaviors bind together the disparate pieces of a multivalent document to present the user with a single unified conceptual document. Therefore this multivalent document system has been very useful for the author's retrospective digitization project.

8. The bibliographic metadata allow the integration of the MVD format of the scanned article into the digital library MILESS, see [12] and [6]. Thus it is possible to apply all search functions of MILESS to this MVD document see [4].



Multiple active semantic layers of the contents of a scanned page

Since 1997 all articles of "Archiv der Mathematik" have been published electronically in PDF format. Using the public domain program *Image Magick 4.2.7* to the PDF files of the digital pages of a recent article a complete TIFF file of each page of the volumes **68** (1997) - **73** (1999) has been produced. Applying then the above mentioned programs for the recognition of the layout of the first page of an article one obtains automatically all the bibliographical data of these articles in XML-format. Thus it has been possible to include all the data of the 6 recent volumes of Archiv der Mathematik into the digital library MILESS. Furthermore, applying the programs for the recognition of the references to the TIFF files of these articles it is possible to link the recent with the back issues of the journal contained in MILESS.

Using the metadata of the page MILESS provides links of the article to the full text of all the cited articles contained in the library server of MILESS. However, there are only a few digital articles in it. This is a great restriction which we want to overcome in this **new** project.

2.2 "Commentarii Mathematici Helvetici"

With permission of Birkhäuser Verlag and the editor in chief the author's study group has recently retrodigitized the volumes **69**(1994) - **71**(1996) of the journal "Commentarii Mathematici Helvetici". Furthermore, the recent electronically published volumes **72** (1997) - **74** (1999) have been linked to the retrodigitized volumes in the digital library system MILESS, which also contains the data produced in the above mentioned "Archiv der Mathematik"-project.

This second project required completely new methods for recognizing the references of recent or back issues of the "Commentarii Mathematici Helvetici". The editors of this journal did not pay any attention to have a uniform headline of the references of all articles published in this journal. The following examples have been found: References, Literatur, Bibliographie, Bibliographique, Bibliographiques, Références, Références Bibliographiques.

The authors used completely different abbreviations of a cited article in their lists of references, e. g. [1] or [2], 1. or 2., [An] for author Anderson, [GLP] for the 3 authors Gaboriau, Levitt and Paulin, [Bi 76] for Bieri's article published in 1976, [B1] and [B2] for 2 cited articles of S. Bochner, [Hartshorne], the complete name of the author, and no abbreviations at all.

Also the authors did not care about the orderings of the abbreviations of the first names. Sometimes they are before and sometimes they are behind the last names, even in the list of references of one article. In cited articles with several authors the different names are separated by commas or "and". Again this is not done uniformly.

The important separation of the names of the authors from the title of a cited article is also done very incoherently. So a comma, a dot and a colon are used arbitrarily.

In the meantime, Dr. R. Staszewski and Mrs. D. Mirrouch have developed a new recognition program for the references which can deal with all these obstacles. If the abbreviation of the cited journal which is used by the author of the downloaded article is listed in the constructed database containing the relevant bibliographic data of the mathematical journals then the recognition rate of the references is almost 90%, as is shown by the following table. It provides the recognition rates of all cited journal articles of the 8 mathematical papers published in issue 3 of volume **69** (1994) of the journal "Commentarii Mathematici Helvetici", pp. 335-499. For the linking programs described in section 3 only the bibliographical data ⟨ author ⟩, ⟨ title ⟩, ⟨ journal name ⟩ and ⟨ ISSN ⟩ of a cited article are required. The last column in the following table gives the total number of cited articles for which all these relevant bibliographic data have been produced in XML format.

Experimental Results

Pages of references	Number of all references	Number of journal articles	Recognized XML-data
350	17	11	11
374	13	5	2*
415/416	23	14	12*
430	12	8	7
446	6	5	4*
472/473	12	9	9
482	9	5	5
498/499	15	13	10 †

*Low recognition rate caused author's wrong abbreviations of journal names.
† Low recognition rate caused by mistakes of used OCR system.

In the future the XML-data for a cited article in a list of references of a recent electronically published or retrodigitized article of a mathematical journal will be given in the following order:

`ISSN|JOUR|AUTH|VID|IID|PID|YNO|TYP|KEY|MRID|TTL`

where:

ISSN	=	ISSN
JOUR	=	Journal title of abbreviation
AUTH	=	Author name(s)
VID	=	Volume number
IID	=	Issue ID number
PID	=	Initial Page number
YNO	=	Year
TYP	=	Resource type (for compatibility with CrossRef.; in Batch MR Lookup this value is ignored)
KEY	=	User supplied key
MRID	=	Identifier (Mathematical Reviews Number)
TTL	=	Item title

This is the same structure which has been used by the American Mathematical Society MR Lookup reference tool for linking [13]. It is compatible with matching done by the commercial CrossRef. initiative, see [11]. The above described format of the output of the XML-data of the Essen program for recognizing a cited article in a reference list will be independent of the particular mathematical journal. This will be important for solving the reference linking problems described lateron.

2.3 Separating Mathematical Formulas from Ordinary Text

In general, commercial OCR systems like FineReader cannot recognize mathematical formulas, even if they are only easy algebraic expressions like $p^{\alpha}\ q^{\beta}$. The occurance of such mathematical formulas may even destroy the recognition result of the remaining ordinary text.

Therefore it is necessary to separate the mathematical formulas from the ordinary text of an article for improving the recognition rate of the ordinary text. This is a special problem dealt with in the joint German-Japanese research project "Structure Analysis and retrodigitization of scientific documents" lead by the author, Professor M. Okamoto (Nagano) and Professor M. Suzuki (Fukuoka). The Deutsche Forschungsgemeinschaft (DFG) has granted financial support for 2 years for this project, which has started on Oct. 1, 2000. The Japanese partners have found financial support by the Ministry of Education of Japan and the Japanese Society for Promoting Sciences (JSPS). Due to this cooperation the Essen study group has access to the commercial OCR system ExpressReader [2] of Toshiba Ltd. It contains new algorithms for recognition of mathematical symbols and formulas.

In this German-Japanese cooperation new software has been developed for the separation of the mathematical formulas from the ordinary text, and the production of the Xdoc-files of the ordinary text of a page of a mathematical article. Thus, the recognition of the ordinary text of such a page has been improved considerably. Recently, Professor Suzuki's study group (Kyushu University) has incorporated these programs into an interface "Infty" of the Toshiba OCR system ExpressReader. Furthermore, the Xdoc format has been replaced by XML. Thus any good scanner together with ExpressReader and Infty will be able to produce automatically a "TIFF file" and a file consisting of the "Ordinary Text in XML format". The results of this cooperation have been demonstrated at the "Workshop on Mathematical Documents Recognition and Digitization", which took place at Kyushu University in October 2001. It was organized by Professor M. Suzuki and the author.

3. Programs producing a multilayer DjVu document from a printed or an electronically published article

3.1 Establishing links in a distributed digital library

Due to severe copyright problems it will not be possible to incorporate the full texts of all electronically published recent and retrodigitized back issues of all major mathematical journals into one digital library as in the "Archiv-project" 2.1. This cannot even be done at a "DFG-Sammelstelle für Mathematik" like the one of the State and University Library of Göttingen University. Also the CrossRef. initiative [1] of the commercial publishers will not be able to create an almost complete distributed mathematical archive library containing at least all recent electronically published articles. So far no Japanese mathematical journal is covered by CrossRef.

On the other hand "Mathematical Reviews" and "Zentralblatt für Mathematik" contain review articles of most of the mathematical literature appeared in the last 60 years. Therefore each mathematician whose university has a legal subscription to them has instant access to these two important mathematical databases. Both review journals have permanent links of some review articles to the digital full text of the corresponding mathematical articles. In the future it will be in the interest of the commercial publishers to have all digital issues of their journals linked to these review journals. Furthermore, "Mathematical Reviews" and "Zentralblatt für Mathematik" have links to the Tiff-files of the back issues of 22 important mathematical journals retrodigitized by JSTOR, see section 2. Any mathematician whose university has a subscription to JSTOR can at least link a downloaded article read on screen to the pictures of the scanned pages of a cited article contained in the JSTOR database.

Already this huge mathematical literature in digital form which has been linked via the two mathematical review journals shows that even without the solution of the copyright problems it is possible to build a prototype of a *distributed, linked* and *searchable* mathematical archive library to which the individual mathematician has instant access provided he or she works at a university or research institute which has a legal subscription to "Mathematical Reviews", "Zentralblatt für Mathematik", JSTOR and the mathematical digital journals of his or her interest.

In view of the programs and software tools described in the previous section one only needs to write new programs which automatically establish the links from the cited articles of a mathematical article stored in a digital library to their corresponding review articles in "Zentralblatt für Mathematik" and "Mathematical Reviews". Whenever there are permanent links between the review and the digital full text of the cited article on a server of a commercial publisher or a university library, the user can automatically retrieve this full text from the remote server.

However, to establish such automatic links from a TIFF file one needs a multilayer document format for the retrodigitized articles. Furthermore, each user of such a distributed linked digital library has to be equipped with a reader enabling him or her to view such a document on screen. Of course, such a reader has to be freely available in the internet. The user also wants to search in the ordinary text of a retrodigitized article. At present there seems to be only the DjVu reader which offers all these abilities freely. It can be downloaded from
http://www.lizardtech.com

On its first page there is the bottom "DjVu plug in" .

The price has to be paid on the sides of the document servers of the digital libraries, because all retrodigitized articles have to be stored in the multilayer DjVu format.

DjVu is an image compression technique specifically designed for the compression of scanned documents at a high resolution, see [5]. Like T. Phelps' MVD system [16] it works with different layers and has a software to link the corresponding points of the two layers. Therefore the DjVu technique is an alternative to the use of MVD. However, for large scale applications it is necessary to use a commercial product of LizardTech, Inc., Amsterdam, Holland.

Equipped with the DjVu technology the problem of producing automaticl links can be solved as is described in the following section.

3.2 Description of the recognition, conversion and linkage softwares

The transformation of a printed article into a multilayer DjVu document can be achieved by the production of the following softwares:

3.2.1 Software to recognize the layout of the first page of a scanned article to get its bibiographic data

At present several publishers like Academic Press and universities like the Digitization Center of Göttingen University scan many mathematical journals. Thus it is worthwile to apply the recognition programs described in section 2.1 for the automatic production of the bibliographic data of a scanned article, at least for the journal volumes published between 1960 and 1995. Of course, the original programs developed in the "Archiv der Mathematik" project have to be adjusted to the different journals.

However, it should be mentioned that the automatically produced XML-data of the scanned articles can be checked by linking them to "Mathematical Reviews" or "Zentralblatt für Mathematik" using the new software to be developed in project 3.2.4, see below.

3.2.2 Software to separate the mathematical formulas from the ordinary text of an article in order to produce its Xdoc format

Since mathematical papers contain mathematical formulas the methods described in section 2.3 have to be applied in order to improve also the recognition rate of the ordinary text of these digital articles. This will be necessary for the recognition of the references and linking them to their digital full texts.

Using a good scanner and the OCR system ExpressReader with interface Infty described in 2.3 from each page of a scanned mathematical article one can produce automatically its "TIFF file" and a file consisting of the "ordinary text in XML format". The only program of which not even an experimental version has been built so far is the one which converts the "ordinary text in XML" into the "ordinary text in DjVu format".

Therefore the originally developed programs described in section 2.3 are used to produce the Xdoc file of the ordinary text of a scanned page from the output of the OCR system ExpressReader.

3.2.3 Software to recognize cited articles of recent electronically published or retrodigitized mathematical journals

Although the programs described in sections 2.1 and 2.2 work very well for producing the XML-records of cited articles of a downloaded digital mathematical article the following example shows that each mathematical journal not yet dealt with creates new recognition problems.

BIBLIOGRAPHY

1. N. Blackburn, *On a special class of p-groups*, Acta Math., **100** (1958), 45-92.
2. ———, *Generalizations of certain elementary theorems on p-groups*, Proc. London Math. Soc. (3), **11** (1961), 1-22.
3. R. Brauer, *On the connection between the ordinary and the modular characters of finite groups*, Ann. of Math. (2), **42** (1941), 926-935.
4. W. Burnside, *Theory of Groups of Finite Order*, Cambridge, 1911.
5. ———, *On groups of order $p^\alpha q^\beta$*, Proc. London Math. Soc., (2), **2** (1904), 432-437.
6. L. E. Dickson, *Linear Groups*, New York, 1958.

W. Feit, J. G. Thompson, Solvability of groups of odd order. Pacific Journal of Mathematics, Vol. 13 (1963), 775-1029

Here the use of bars instead of repeating the names of the authors of previous articles creates a new problem. Again, it has been solved in the meantime. But in this cooperation the programs developed in Essen will be applied to the Tiff files of the lists of references of the retrodigitized journals mentioned in section 9.2 and 9.3. Thus the database containing all used abbreviations of cited mathematical journals will be enlarged. This helps to improve the recognition rate.

3.2.4 Software to link the XML-data of these cited articles to their reviews in "Mathematical Reviews" or "Zentralblatt für Mathematik"

With the above mentioned reference recognition procedures one obtains for each cited digital article a corresponding XML-record, like in the following example:

```
<REFERENCE>
 <COORDS>   8 266 2949 145                          </COORDS>
 <JOURNAL>  Annals of Mathematics                   </JOURNAL>
 <ISSN>     0003-486X                               </ISSN>
 <VOLUME>   121                                     </VOLUME>
 <INITPAGE> 215                                     </INITPAGE>
 <LASTPAGE> 249                                     </LASTPAGE>
```

```
 <YEAR>     1985                                              </YEAR>
 <AUTHORS>
  <AUTHOR>
   <LASTNAME> HARER                                          </LASTNAME>
   <FIRSTNAME>                                               </FIRSTNAME>
   <FIRSTINIT> J.                                            </FIRSTINIT>
   <SECNDINIT> L.                                            </SECNDINIT>
  </AUTHOR>
 </AUTHORS>
 <TITLE>    Stability of the homology of the mapping class
            groups of orientable surfaces
 </TITLE>
</REFERENCE>
```

This XML-record is the essential input for the software to be developed for automatic linkage to "Zentralblatt für Mathematik" and "Mathematical Reviews". The coordinates in DjVu format of the area of this citation in the list of references are also produced automatically. They are piped through all the following procedures applied to this XML-record until the URL of the corresponding review article has been produced. This URL and the coordinates allow the automatic production of the link from the cited article to its review article. The author's collaborator Dr. M. Kratzer has started some experiments for developing such a software for links to

<u>Mathematical Reviews</u>

In this subproject a suitable parser `xml2mrq` has to be developed which converts each XML-record contained in a file <*name*>`.xml` into a Batch MR Lookup Query API. Precisely the parser sorts the data between the various XML-tags into the corresponding API field. The syntax of the Batch MR Lookup Query API is as follows:

`ISSN|JOUR|AUTH|VID|IID|PID|YNO|TYP|KEY|MRID|TTL`

For the example given above `xml2mrq` would generate the line

```
0003-486X|Annals of Mathematics|Harer, J.|121||215|
1985||||  Stability of the homology of the mapping class
groups of orientable surfaces
```

in the output file <*name*>`.mrq` produced by `xml2mrq`.

The non-interactive download utility `wget` under GNU General Public License interprets the list of Query APIs in <*name*>`.mrq` as a list of HTTP requests and performs the actual query process at the MR Lookup database saving the results in a third file <*name*>`.res`.

A further advantage of the use of Batch MR Look up is its error correcting side effect:

Due to the "triple query" strategy used by Batch MR Lookup —

1st: query: try to match `AUTH` and two other fields,
2nd: query: try to match `JOUR` and `VID` and `PID` and one other field,
3rd: query: try to match `TTL` and two other fields,

— there's always a good chance that also erroneous items can be identified uniquely in which case Batch MR Lookup returns the complete and correct(!) API of the cited article.

For the above example Batch MR Lookup returns the information

`0003-486X|Ann. of Math. (2)|Harer|121|2|215|1985|||`*87f:57009*
`|Stability of the homology of the mapping class groups of orientable surfaces`

Thus, one obtains the "Mathematical Reviews" number `MRID` (typed in italics) of the article's review.

If Batch MR Lookup wasn't able to find a unique match it returns the Query API as it was submitted. So in order to decide if a query was successful one simply has to browse through the lines of `<`*name*`>.res`, search each of them for the ninth delimiter "`|`" and analyze the characters following it. Exactly this will be the job of a converting tool to be named **`res2html`**. Starting from the `MRID` it takes a concatenation of strings to establish a hyperlink to the review of the cited article:

```
<a href="http://www.ams.org/mathscinet-getitem?mr=87f:57009">MR Review of [H]</a>
```

For lines which contain an empty `MRID` field **`res2html`** has to generate an appropriate error message: E.g. unsuccessful queries may result from citations of too old articles (earlier than 1940; not contained in MR Lookup database) or from the fact that a cited paper was just to appear or even a preprint.

Whenever Mr. Kratzer's program was able to get a MR number of a cited article having a link to its digital full text, then MathSciNet automatically provides the information and the link to the full text. In the example above MathSciNet shows:

87f:57009 57M99 (20F34)
Harer, John L. (1-MD)
Stability of the homology of the mapping class groups of orientable surfaces
Ann. of Math. (2) **121** (1985), no.2, 215–249.

Article

Now clicking on Article JSTOR replies: "We are sorry. You do not have access to JSTOR from current location. The article you request is:

Stability of the homology of the mapping class groups of orientable surfaces, pp. 215-249
Annals of Mathematics Vol.121, Iss. 2
John L. Harer

Nevertheless, even this information proves that in theory one could retrieve the full digital text of this quoted article.

Zentralblatt für Mathematik

The author's study group will develop a similar software which automatically will produce links from downloaded digital mathematical journals to their reviews in the "Zentralblatt für Mathematik".

Although this software will be slightly different from the one described above, this part of the project does not cause any technical difficulties. Most likely, after January 1, 2002 all electronically published articles published in the mathematical Springer journals will be linked to "Zentralblatt für Mathematik". Furthermore, the retrodigitized volumes of "Archiv der Mathematik" and "Commentarii Mathematici Helvetici" will be stored at the SUB Göttingen and linked to the Zentralblatt für Mathematik.

3.2.5 Software to build html-links of the references into the DjVu format of a retrodigitized article

The commercial "DjVu" converter transforms the TIFF file of a scanned page into both:

1) the DjVu graphic layer, and
2) DjVu text layer.

Furthermore, DjVu contains an "annotation layer" which can be used to incorporate html-links into the graphic layer as follows. Using the software described in 3.2.3 one obtains the XML records of a cited journal article from its TIFF file. With them the programs described in 3.2.4 produce the URL of MathSciNet and the Mathematical Review number of the cited article, if available. Dr. Nörenberg has written an experimental program which automatically puts such a URL and the coordinates of the rectangular area of the text of the corresponding reference in pixels of the TIFF file of the scanned page at the correct spot of the "annotation layer". A similar program will be written for producing links to "Zentralblatt für Mathematik".

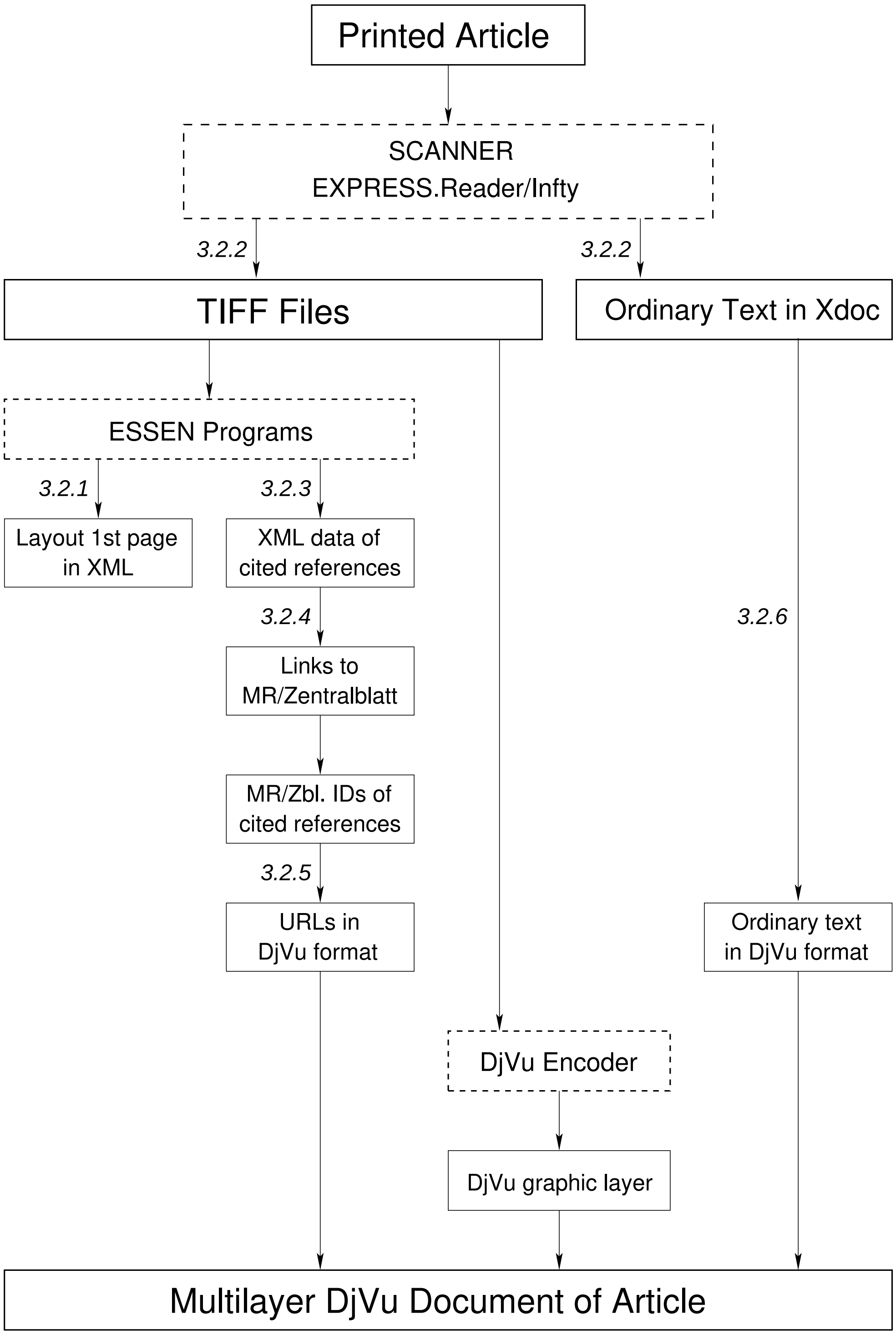
3.2.6 Software to convert the Xdoc format of the ordinary text into its DjVu format

The Xdoc format and the DjVu format of the ordinary text of a page of a mathematical article present both this text in ASCII and assign to each word of it coordinates. Since the ASCII text is the same one only has to convert the corresponding coordinates of a word from Xdoc format into DjVu format. For this transformation the author's collaborator Dr. Nörenberg has written a computer program.

3.2.7 Production of the multilayer DjVu document of a mathematical article

Using a good scanner and the OCR system ExpressReader together with all the programs 3.2.1 till 3.2.6 one obtains for each printed article its TIFF files, its bibliographic data in XML format, the URLs of its cited articles in the DjVu annotation layer, and its ordinary text in the DjVu textlayer. From the TIFF files of the pages of the article their DjVu graphic layers can be produced by means of the DjVu encoder, which also helps to put all three DjVu layers together.

The following picture describes the production of the resulting multilayer DjVu document of a mathematical article in detail


Printed Article
SCANNER
EXPRESS.Reader/Infty
3.2.2
3.2.2
TIFF Files
Ordinary Text in Xdoc
ESSEN Programs
3.2.1
3.2.3
Layout 1st page in XML
XML data of cited references
3.2.4
3.2.6
Links to MR/Zentralblatt
MR/Zbl. IDs of cited references
3.2.5
URLs in DjVu format
Ordinary text in DjVu format
DjVu Encoder
DjVu graphic layer
Multilayer DjVu Document of Article

Although DjVu is an image compression technique designed for the compression of scanned documents at a high resolution, the DjVu format of an electronically published article can also be produced. All that is necessary is to produce TIFF files of its pages by means of a public domain program like *Image Magick 4.2.7*, see step 8 of section 2.1. Then apply the above described programs 3.2.1 till 3.4.6.

As will be shown by an example in the next section the whole *DjVu document of a page* of a mathematical article looks like a rectangular trisected sheet

graphic layer DjVu coordinates for compressed picture of a page
text layer Recognized ASCII text .... DjVu coordinates of ASCII symbols
annotation layer Links of cited articles to "Mathematical Reviews" with coordinates of the rectangular areas of the text of each reference in pixels of the TIFF file of the page

Using the DjVu-reader this format allows the user to search for words, names and mathematical expressions in the whole ordinary text of the paper, and to link the cited references automatically to their reviews in "Mathematical Reviews" or "Zentralblatt für Mathematik", and to their digital full texts, provided they are contained in the distributed digital library.

4. Example of a multilayer DjVu document

The following picture 4.1 shows the contents of page 61 of the article:
Allan Hatcher, Homological stability for automorphism groups of free groups. Commentarii Mathematici Helvetici **70** (1995), 39-62.

Its mathematical formulas are framed by black boxes produced by the OCR system, see picture 4.2. The truncated Xdoc file of its ordinary text is shown in picture 4.3. Its transformation into DjVu format is given in picture 4.4. The annotation layer with the URLs of the links ot the cited articles is found in 4.5. Finally pictures 4.6 and 4.7 show the review article of the second cited article [BH] and its bibliographic data in JSTOR.

4.1 Printed page

the punctures. The quotient $\mathbb{S}_0/\mathrm{Aut}(F_n)$ is finite since there are only finitely many isotopy classes of 2-spheres in a punctured S^3, such a 2-sphere being determined up to isotopy by how it separates the punctures.

At this point we need the fact that $\mathrm{Aut}(F_n)$ has a torsionfree subgroup of finite index. Namely, the composition $\mathrm{Aut}(F_n) \to GL_n(\mathbb{Z}) \to GL_n(\mathbb{Z}_p)$ has torsionfree kernel for $p \geq 3$. For the second of these two maps this is explained on p. 40 of [B], and for the first map it can be deduced as a pleasant exercise from a result of Culler [Cu] that any periodic automorphism of F_n is induced by a periodic homeomorphism of a finite graph having fundamental group F_n. If G denotes the kernel of $\mathrm{Aut}(F_n) \to GL_n(\mathbb{Z}_p)$ then G acts freely on $\mathbb{S}_0$ and so the quotient $\mathbb{S}_0/G$ is a finite CW complex $K(G, 1)$. Applying the Lyndon–Hochschild–Serre spectral sequence to the fibration $K(G, 1) \to K(\mathrm{Aut}(F_n), 1) \to K(\mathrm{Aut}(F_n)/G, 1)$ we can deduce that the homology groups of $\mathrm{Aut}(F_n)$ are finitely generated: The E^2 terms are $H_p(\mathrm{Aut}(F_n)/G; H_q(G))$, $H_q(G)$ is finitely generated since we have a finite $K(G, 1)$, and the homology groups of $\mathrm{Aut}(F_n)/G$ are finitely generated with any finitely generated twisted coefficient system since finite groups have CW $K(\pi, 1)$'s with finite skeleta.

Since $\mathbb{S}_0/G$ has dimension $2n-2$, $\mathrm{Aut}(F_n)$ had v.c.d. at most $2n-2$. On the other hand, the v.c.d. is at least $2n-2$ since $\mathrm{Aut}(F_n)$ contains a free abelian subgroup of rank $2n-2$ generated by the $2n-2$ automorphisms which fix all generators x_i of F_n except for one x_k with $k > 1$, which is sent to $x_1 x_k$ or $x_k x_1$.

REFERENCES

[B] BROWN, K. S., *Cohomology of Groups*, Springer-Verlag, 1982.

[BH] BESTVINA, M. and HANDEL, M., *Train tracks and automorphisms of free groups*, Annals of Math. *135* (1992), 1–51.

[Ch_1] CHARNEY, R., *Homology stability for GL_n of a Dedekind domain*, Invent. Math. *56* (1980), 1–17.

[Ch_2] CHARNEY, R., *A generalization of a theorem of Vogtmann*, J. Pure Appl. Alg. *44* (1987), 107–125.

[C] COHEN, F. R., *Artin's braid groups, classical homotopy theory, and sundry other curiosities*, Contemp. Math. *78* (1988), 167–206.

[Cu] CULLER, M., *Finite groups of outer automorphisms of a free group*, Contemp. Math. *33* (1984), 197–207.

[CV] CULLER, M. and VOGTMANN, K., *Moduli of graphs and automorphisms of free groups*, Invent. Math. *84* (1986), 91–119.

[G] GERSTEN, S. M., *A presentation for the special automorphism group of a free group*, J. Pure Appl. Alg. *33* (1984), 269–279.

[GT_1] GOLDSTEIN, R. Z. and TURNER, E. C., *Automorphisms of free groups and their fixed points*, Invent. Math. *78* (1984), 1–12.

[GT_2] GOLDSTEIN, R. Z. and TURNER, E. C., *Monomorphisms of finitely generated free groups have finitely generated equalizers*, Invent. Math. *82* (1985), 283–289.

[H] HARER, J. L., *Stability of the homology of the mapping class groups of orientable surfaces*, Annals of Math. *121* (1985), 215–249.

[Hat] HATCHER, A., *On triangulations of surfaces*, Top. and its Appl. *40* (1991), 189–194.

[HM] HATCHER, A. and MCCULLOUGH, D., *Finite presentation of 3-manifold mapping class groups*, Springer Lecture Notes *1425* (1990), 48–57.

4.2 Separation of Mathematics

the punctures. The quotient $\mathbb{S}_0/\mathrm{Aut}(F_n)$ is finite since there are only finitely many isotopy classes of 2-spheres in a punctured S^3, such a 2-sphere being determined up to isotopy by how it separates the punctures.

At this point we need the fact that $\mathrm{Aut}(F_n)$ has a torsionfree subgroup of finite index. Namely, the composition $\mathrm{Aut}(F_n) \to GL_n(\mathbb{Z}) \to GL_n(\mathbb{Z}_p)$ has torsionfree kernel for $p \geq 3$. For the second of these two maps this is explained on p. 40 of [B], and for the first map it can be deduced as a pleasant exercise from a result of Culler [Cu] that any periodic automorphism of F_n is induced by a periodic homeomorphism of a finite graph having fundamental group F_n. If G denotes the kernel of $\mathrm{Aut}(F_n) \to GL_n(\mathbb{Z}_p)$ then G acts freely on $\mathbb{S}_0$ and so the quotient $\mathbb{S}_0/G$ is a finite CW complex $K(G, 1)$. Applying the Lyndon–Hochschild–Serre spectral sequence to the fibration $K(G, 1) \to K(\mathrm{Aut}(F_n), 1) \to K(\mathrm{Aut}(F_n)/G, 1)$ we can deduce that the homology groups of $\mathrm{Aut}(F_n)$ are finitely generated: The E^2 terms are $H_p(\mathrm{Aut}(F_n)/G; H_q(G))$, $H_q(G)$ is finitely generated since we have a finite $K(G, 1)$, and the homology groups of $\mathrm{Aut}(F_n)/G$ are finitely generated with any finitely generated twisted coefficient system since finite groups have CW $K(\pi, 1)$'s with finite skeleta.

Since $\mathbb{S}_0/G$ has dimension $2n-2$, $\mathrm{Aut}(F_n)$ had v.c.d. at most $2n-2$. On the other hand, the v.c.d. is at least $2n-2$ since $\mathrm{Aut}(F_n)$ contains a free abelian subgroup of rank $2n-2$ generated by the $2n-2$ automorphisms which fix all generators x_i of F_n except for one x_k with $k > 1$, which is sent to $x_1 x_k$ or $x_k x_1$.

REFERENCES

[B] BROWN, K. S., *Cohomology of Groups*, Springer-Verlag, 1982.

[BH] BESTVINA, M. and HANDEL, M., *Train tracks and automorphisms of free groups*, Annals of Math. *135* (1992), 1–51.

[Ch$_1$] CHARNEY, R., *Homology stability for GL_n of a Dedekind domain*, Invent. Math. *56* (1980), 1–17.

[Ch$_2$] CHARNEY, R., *A generalization of a theorem of Vogtmann*, J. Pure Appl. Alg. *44* (1987), 107–125.

[C] COHEN, F. R., *Artin's braid groups, classical homotopy theory, and sundry other curiosities*, Contemp. Math. *78* (1988), 167–206.

[Cu] CULLER, M., *Finite groups of outer automorphisms of a free group*, Contemp. Math. *33* (1984), 197–207.

[CV] CULLER, M. and VOGTMANN, K., *Moduli of graphs and automorphisms of free groups*, Invent. Math. *84* (1986), 91–119.

[G] GERSTEN, S. M., *A presentation for the special automorphism group of a free group*, J. Pure Appl. Alg. *33* (1984), 269–279.

[GT$_1$] GOLDSTEIN, R. Z. and TURNER, E. C., *Automorphisms of free groups and their fixed points*, Invent. Math. *78* (1984), 1–12.

[GT$_2$] GOLDSTEIN, R. Z. and TURNER, E. C., *Monomorphisms of finitely generated free groups have finitely generated equalizers*, Invent. Math. *82* (1985), 283–289.

[H] HARER, J. L., *Stability of the homology of the mapping class groups of orientable surfaces*, Annals of Math. *121* (1985), 215–249.

[Hat] HATCHER, A., *On triangulations of surfaces*, Top. and its Appl. *40* (1991), 189–194.

[HM] HATCHER, A. and MCCULLOUGH, D., *Finite presentation of 3-manifold mapping class groups*, Springer Lecture Notes *1425* (1990), 48–57.

4.3 Xdoc file


```
[a;"XDOC.10.0";E;"Infty"]
[d;"00701061.xdc"]
[p;1;P;83;S;0;1666;0;0;2984;4496]
[t;1;1;0;0;A;"";"";"";0;0;0;0;1]
[f;0;"<DEFAULT>";R;q;10;V;0;0;0;33;100]
[f;1;"Courier";R;q;12;F;60;50;10;50;100]
[f;2;"F";R;q;48;V;0;0;0;26;100]
[s;1;584;0;80;p;2]Homological[h;942;31]stability[h;1192;30]for[
h;1301;29]automorphism[h;1738;29]groups[h;1955;31]of[h;2046;22]
free[h;2168;30]groups[h;2386;508]61[y;2950;0;77;0;H]
[s;1;11;0;296;p;2]the[h;115;36]punctures.[h;508;39]The[h;679;36
]quotient[h;1004;39][h;1434;34]is[h;1518;36]finite[h;1725;38]si
nce[h;1928;36]there[h;2136;36]are[h;2276;36]only[h;2460;36]fini
tely[h;2730;35]many[y;2956;0;277;0;H]
[s;1;10;0;404;p;2]isotopy[h;263;29]classes[h;518;32]of[h;625;21
]2-spheres[h;967;29]in[h;1059;31]a[h;1127;27]punctured[h;1506;3
3][h;1587;85]such[h;1824;31]a2-sphere[h;2211;29]being[h;2425;29
]determined[h;2841;30]up[y;2956;0;401;0;H]
[s;1;10;0;510;p;2]to[h;76;38]isotopy[h;368;38]by[h;486;38]how[h
;668;36]it[h;749;37]separates[h;1102;38]the[h;1243;36]punctures
.[y;1636;0;492;0;H]
[s;1;133;0;620;p;2]At[h;219;33]this[h;373;33]point[h;588;32]we[
h;710;33]need[h;900;35]the[h;1038;34]fact[h;1198;34]that[h;1368
;32][h;1671;33]has[h;1817;33]a[h;1888;32]torsionfree[h;2294;34]
subgroup[h;2652;34]of[h;2761;24]finite[y;2956;0;599;0;H]
[s;1;10;0;736;p;2]index.[h;213;27]Namely,[h;530;28]the[h;661;24
]composition[h;1114;24][h;2187;24]has[h;2324;24]torsionfree[h;2
721;24]kernel[y;2956;0;708;0;H]
[s;1;12;0;834;p;2]for[h;110;28][h;325;61]For[h;512;35]the[h;649
;37]second[h;918;37]of[h;1030;26]these[h;1226;36]two[h;1387;35]
maps[h;1603;35]this[h;1758;36]is[h;1843;34]explained[h;2206;36]
on[h;2326;35]p.[h;2421;37]40[h;2536;35]of[h;2646;28][h;2771;59]
and[y;2956;0;816;0;H]
[s;1;10;0;942;p;2]for[h;110;28]the[h;240;28]first[h;399;27]map[
h;575;27]it[h;647;27]can[h;792;30]be[h;896;28]deduced[h;1207;29
]as[h;1303;28]a[h;1367;27]pleasant[h;1679;27]exercise[h;1971;29
]from[h;2166;26]a[h;2229;28]result[h;2445;28]of[h;2549;19]Culle
r[h;2779;28][[Cu][y;2955;0;923;0;H]
[s;1;10;0;1049;p;2]that[h;146;32]any[h;300;33]periodic[h;612;32
]automorphism[h;1154;33]of[h;1262;24][h;1341;50]is[h;1440;32]in
duced[h;1742;33]by[h;1856;33]a[h;1925;32]periodic[h;2236;32]hom
eomorphism[h;2856;32]of[y;2964;0;1030;0;H]
[s;1;10;0;1156;p;2]a[h;46;34]finite[h;251;33]graph[h;483;35]hav
ing[h;750;34]fundamental[h;1220;34]group[h;1458;34][h;1547;81]I
f[h;1690;28][h;1772;30]denotes[h;2064;34]the[h;2202;33]kernel[h
;2445;35]of[h;2556;24][y;2943;0;1138;0;H]
[s;1;12;0;1272;p;2][h;306;32]then[h;486;36][h;575;31]acts[h;736
;34]freely[h;958;32]on[h;1074;36][h;1160;63]and[h;1348;34]so[h;
1452;34]the[h;1589;32]quotient[h;1910;34][h;2119;31]is[h;2198;3
3]a[h;2266;32]finite[h;2470;36][h;2646;21]complex[y;2956;0;1261
(Xdoc file truncated)
```

4.4 DjVu Text layer

```
TXTa.......Homological stability for automorphism groups of free groups
61 the punctures. The quotient  is finite since there are only finitely
many isotopy classes of 2-spheres in a punctured such a2-sphere being
determined up to isotopy by how it separates the punctures. At this
point we need the fact that  has a torsionfree subgroup of finite index.
Namely, the composition  has torsionfree kernel for  For the second of
these two maps this is explained on p. 40 of  and for the first map it
can be deduced as a pleasant exercise from a result of Culler [Cu] that
any periodic automorphism of is induced by a periodic homeomorphism of
a finite graph having fundamental group  If  denotes the kernel of
then acts freely on  and so the quotient  is a finite  complex  Applying
the Lyndon-Hochschild -Serre spectral sequence to the fibration we can
deduce that the homology groups of  are finitely generated: The  terms
are  is finitely generated since we have a finite  and the homology groups
of are finitely generated with any finitely generated twisted coefficient
system since finite groups have  with finite skeleta. Since  has dimension
had v. c. d. at most  On the other hand, the v. c. d. is at least  since
contains a free abelian subgroup of rank  generated by the  automorphisms
which fix all generators  of  except for one  with  which is sent to  or
REFERENCES
B] BROWN, K. S. Cohomology of Groups, Springer-verlag, 1982.
[BH] BESTVINA,  and HANDEL,  Train tracks and automorphisms offree groups,
Annals of Math.

[Ch|] CHARNEY,  Homology stability for  of a Dedekind domain, lnvent. Math.
56(1980), l-17.
 CHARNEY,  A generalization of a theorem of Vogtmann,  Pure
 COHEN, F. R. Artin's braid groups, classical homotopy theory, and sundry
 other curiosities,Contemp. Math.
[CU]CULLER,  Finite groups of outer automorphisms of  free group, Contemp.
Math. 33(1984), 197-207.
 CULLER,  and VOGTMANN,  Moduli of graphs and automorphisms of free groups,
 lnvent. Math.
 GERSTEN, S. M. A presentation for the special automorphism group of  free
 group,  Pure Appl. Alg. 33(1984), 269-279.
 GoLDsTEIN, R. Z. and TURNER, E. C. Automorphisms of free groups and their
 fiXed points, lnvent. Math.
 GOLDSTEIN, R. Z. and TURNER,  generatedfree groups have finitely generated
 equalizers, Invent. Math.
 HARER, J. L. Stability of the homology of the mapping class groups of
 orientable surfaces, Annals of Math.
Hat]HATCHER, A. On triangulations of surfaces, Top. and its Appl. 40(1991),
189-194.
[HM]HATCHER, A. and MCCULLOUGH,  Finite presentation of 3-manifold mapping
class groups, Springer Lecture Notes
..............b..,..H.#.>.-....A.........f.-................-...
...........O.-................-...............-.............<.
-..............d.-...............-.............8.-.........}..
....-....K.........h.-..........$...e.-..........'.....-........
..$...!.-..........'.....-.........."...2.-..........$.....-....
......&.....-..........$.....-..........$...h.-..........$.....-
..........$.....-..........#.....-...........?...-....Q........
...-...............-............K.-.............A.-.........
.....?.-........    (coordinate section truncated)
```

4.5 DjVu Annotation layer


```
ANTa....(maparea (url "http://www.ams.org/mathscinet-getitem?m
r=92m:20017" "\_ref" ) "BH" (rect 8 1588 2949 145 ) (xor ) ) (
maparea (url "http://www.ams.org/mathscinet-getitem?mr=81h:180
10" "\_ref" ) "Ch\_1" (rect 8 1504 2949 73 ) (xor ) ) (maparea
 (url "http://www.ams.org/mathscinet-getitem?mr=89h:18010" "\_
ref") "Ch\_2" (rect 8 1420 2949 73) (xor ) ) (maparea (url "ht
tp://www.ams.org/mathscinet-getitem?mr=90a:20077" "\_ref") "C"
 (rect 8 1256 2949 145) (xor ) ) (maparea (url "http://www.ams
.org/mathscinet-getitem?mr=86g:20027" "\_ref") "Cu" (rect 8 11
04 2949 145) (xor ) ) (maparea (url "http://www.ams.org/mathsc
inet-getitem?mr=87f:20048" "\_ref") "CV" (rect 8 928 2949 145)
 (xor ) ) (maparea (url "http://www.ams.org/mathscinet-getitem
?mr=86f:20041" "\_ref") "G" (rect 8 774 2949 145) (xor ) ) (ma
parea (url "http://www.ams.org/mathscinet-getitem?mr=86h:20031
" "\_ref") "GT\_1" (rect 8 597 2949 145) (xor ) ) (maparea (ur
l "http://www.ams.org/mathscinet-getitem?mr=88b:57007" "\_ref"
) "GT\_2" (rect 8 430 2949 145) (xor ) ) (maparea (url "http:/
/www.ams.org/mathscinet-getitem?mr=87f:57009" "\_ref") "H" (re
ct 8 266 2949 145) (xor ) ) (maparea (url "http://www.ams.org/
mathscinet-getitem?mr=92f:57020" "\_ref") "Hat" (rect 8 181 29
49 73) (xor ) ) (maparea (url "http://www.ams.org/mathscinet-g
etitem?mr=91g:57013" "\_ref") "HM" (rect 8 17 2949 145) (xor )
 )
```


Remark: This is the end of the multilayer DjVu document of p. 61 of the article: Allan Hatcher, Homological stability for automorphism groups of free groups. Commentarii Math. Helv. **70** (1995), 39-62.

Finally it has to be mentioned that there are permanent links from "Mathematical Reviews" to the digital full texts in JSTOR of the following articles cited in Hatcher's paper:

[BH] Bestvina, M., and Handel, M., Train tracks and automorphisms of free groups, Annals of Math. **135** (1992), 1-51.

[H] Harer, J. L., Stability of the homology of the mapping class groups of orientable surfaces, Annals of Math. **121** (1985), 215-249.

4.6 MR Review

AMERICAN MATHEMATICAL SOCIETY
MathSciNet *Mathematical Reviews on the Web*

Full Search	Search Author Database	Browse Current Books
Basic Search	Search Journals Database	Browse Current Journals
Search CML	Search MSC	Browse by MSC

Matches for: MR Number=92m:20017 **MSN-Support** **Help**

Return to List Item: **1** of **1** Go To Item #:

0 items in clipboard; (500 item maximum; items will be lost after 2 hours of MathSciNet inactivity)

Retrieve in PDF format. Add citation to clipboard

92m:20017 20E05 (20F28 20F32 20F34 57M07 57M15)
Bestvina, Mladen(1-UCLA); Handel, Michael(1-CUNY7)
Train tracks and automorphisms of free groups.
Ann. of Math. (2) **135** (1992), no. 1, 1--51.

Article

References: 0 **Reference Citations: 6** **Review Citations: 15**

The outer automorphism group $\hbox{Out}(F\sb n)$ of a free group of rank n is in many ways analogous to the mapping class group of a 2-manifold (the group of homeomorphisms modulo isotopy). Thurston, completing a program begun many years ago by Nielsen, classified mapping classes into three types: finite order, pseudo-Anosov, and reducible (a piecing-together of simpler types). Replacing the 2-manifold by a finite 1-dimensional CW-complex, or graph, and the homeomorphisms by homotopy equivalences, the authors here carry out an analogous program that gives an algorithm to produce, for each outer automorphism, an "efficient" homotopy equivalence that induces it on the fundamental group of a graph. This has many consequences, notably a proof of a conjecture of P. Scott that the subgroup of elements of $F\sb n$ fixed by a given automorphism has rank at most n.

An automorphism is called reducible if there are proper free factors $G\sb 1,\cdots,G\sb k$ of $F\sb n$ such that the automorphism permutes the conjugacy classes of the $G\sb i$ and $G\sb 1*\cdots*G\sb k$ is a free factor of $F\sb n$; otherwise it is irreducible. The authors prove that each irreducible automorphism Φ is induced by a "train track map", which means a homotopy equivalence $f\colon \Gamma\to \Gamma$ of some graph Γ, such that the restriction of $f\sp k$ to each edge is locally injective for all $k>0$. The transition matrix has (i,j)-entry equal to the number of times that the image under f of the ith edge crosses the jth edge in either direction. It has a unique positive eigenvector, with eigenvalue $\lambda\geq1$. The train track maps are exactly those whose transition matrices achieve the minimal possible value of λ for their induced automorphism. When $\lambda=1$, Φ has finite order; otherwise λ measures the growth rate of words in the group as powers of Φ are applied. The dynamics of the lift of f to the universal cover of Γ can be studied, yielding information about the subgroup of elements fixed by Φ. In particular, its rank is at most 1 when Φ is

4.7 Full text of Cited Article [BH]

We're sorry. You do not have access to JSTOR from your current location.

The article you requested is:

Train Tracks and Automorphisms of Free Groups
Mladen Bestvina, Michael Handel
Annals of Mathematics, 2nd Ser., Vol. 135, No. 1, pp. 1-51

Options for Access to JSTOR

Even though you do not have access to JSTOR from your current location, there may be another option for you. Please read the following options:

1. **US Participating Institution**

   If you are trying to access JSTOR from an off campus location or are using a campus dial-in line that is not set up to provide access to restricted resources, you are not recognized as being affiliated with a participating site. You may wish to contact a librarian at your institution about local options for remote access.

2. **UK or Ireland Participating Institution**

   If you are a member of a participating Higher Education institution in the UK or Ireland, please use JSTOR's Athens logon page.

3. **Individual Account Holders**

   If you have an account with a participating JSTOR publisher, please use the Individual User logon page.

If you are not affiliated with a participating institution, you may choose to request a copy of this article by contacting the publisher. You may also wish to check your library's print holdings or visit a JSTOR participating institution (US or International) near you.

Access to the JSTOR archive is available only through affiliation with a participating institution (US or International) or through an individual account with a participating publisher. Access from on-campus is determined either by IP addresses provided by participating sites, or via ATHENS for affiliates of participating UK and Ireland sites.

Acknowledgement

The author kindly acknowledges that his study group has been financially supported by the Deutsche Forschungsgemeinschaft since April 1, 2000 for two years to work on the project "Prototyp einer automatisch verlinkten, suchbaren digitalen Fachzeitschriftenbibliothek".

He also thanks Professor M. Suzuki (Kyushu University, Fukuoka, Japan) and Mr. A. Busch (Lizardtech, Amsterdam, Holland) for providing great technical support for using the OCR engine "ExpressReader" of Toshiba Ltd. and the high quality compression technique "DjVu", respectively, in this project. He is also very grateful to the Birkhäuser Verlag (Basel, Switzerland) for its constant legal support.

Finally, the author owes great thanks to his collaborators Mrs. B. Kratzer, Mrs. D. Mirrouch, Dr. M. Kratzer, Dr. R. Nörenberg and Dr. R. Staszewski who have written all the computer programs mentioned in this article.

References

[1] Crossref, the central source for refernce linking, http://crossref.org.
[2] ExpressReader OCR Engine, version 3,5, Toshiba Corporation, Tokyo (confidential).
[3] FineReader OCR Engine, Developer's Guide, ABBYY Software House (BIT Software), Moscow, 1993-1997.
[4] H. Gollan, F. Lützenkirchen, D. Nastoll, "MILESS - a learning and teaching server for multimedia documents", Lecture Notes in Control and Information Sciences **249** (1999), 143-150.
[5] P. Haffner, L. Bottou, P. G. Howard and Y. LeCun, DjVu: Analysing and compressing scanned documents for internet distributions, AT & T Labs-Research, 1999.
[6] IBM DB2 Digital Library, http://www.software.ibm.com./is/dig-lib/
[7] JSTOR, Journal storage, the scholarly journal archive, http://www.jstor.org/journals/
[8] H. Kasuga, M. Okamoto, H. Yamamoto, "Extraction of characters from color documents", in (eds. D.P. Lopresti, J. Zhou) Document Recognition and Retrieval VII, Proceedings of SPIE - International Society for Optical Engineering, **3967**, 278-285 (2000).
[9] J. Kister, "Reference Lists and Citations in the Mathematical Reviews Database", Notices of the AMS **48** (2001), p. 965.
[10] G. Michler, "A prototype of a combined digital and retrodigitized searchable mathematical journal", Lecture Notes in Control and Information Sciences **249** (1999), 219-235.
[11] G. Michler, "Report on the retrodigitization project "Archiv der Mathematik" ", Archiv der Mathematik **77** (2001), 116-128.
[12] MILESS - Multimedialer Lehr- und Lernserver Essen, http://miless.uni-essen.de
[13] American Mathematical Society MR Lookup - A reference tool for linking, http://www.ams.org/batchmrlookup
[14] M. Okamoto, H. Higasi, "Mathematical expression recognition by the layout of symbols (in Japanese)", Trans. IEICE Japan, J78-D-11, No. 3, pp. 474-482 (1995-3).
[15] M. Okamoto, H. Msafiri, "Mathematical expression recognition by projection profile characteristics", Trans. IEICE Japan, J78-D-11, No. 2, pp. 366-370 (1995-2).
[16] T. A. Phelps, "Multivalent Documents: Anytime, Anywhere, Any Type, Every Way User-Improvable Digital Documents and Systems", dissertation, UC Berkeley, 1998, http://www.cs.berkeley.edu/phelps/
[17] T. A. Phelps, R. Wilensky, "Multivalent Documents: Inducing Structure and Behaviors in Online Digital Documents", in *Proceedings of the 29th Hawaii International Conference on System Sciences* Maui, Hawaii, January 3-6, 1996.

Institut für Experimentelle Mathematik, Universität GH Essen, Ellernstrasse 29, D-45326 Essen, Germany