

1st Leibniz MMS Days

**How could the formation of the
Global Digital Mathematics Library
support mathematical modelling and
simulation?**

Olaf Teschke

FIZ Karlsruhe

Berlin, January 28th, 2016



What is a digital mathematics library?

Traditionally, a library has been a dynamic collection of publications, arranged in an accessible form.

“Mathematics” obviously refers to a restriction to math content.

Digital? (Scans vs. sophisticated XML, access options,...)

Content? Organization?

Naively (~2000): “ Make whole math literature digitally available“

> Shaped into World Digital Mathematics Library (WDML): Digitization and accessibility projects, distributed and centralized frameworks



Components of digitization

- Basic layers: HTML, (La)TeX, PDF, TCP/IP, XML,...
- (Retro-)Digitization efforts: Gallica, GDZ, JSTOR, Numdam, Publishers...
- E-born documents: arXiv, Electronic Publishing, ELibM, Euclid, EuDML...
- Superstructures: CrossRef, MathSciNet, MSC, zbMATH,...

By now, about 2.4 million math documents in zbMATH have links to electronic versions (out of 3.6 millions).

FIZ services in the classical DML framework

About Contact General Help Reviewer Service Subscription Preferences

EuDML | The EUROPEAN DIGITAL
MATHEMATICS LIBRARY

English (en) Login Register (Why Register?)

Title, Author, Keyword, Citation, Date... Search

zbMATH
the first resource for mathematics

Documents Authors Journals Classification Software Formulae

Structured Search

Search for documents Fields Operators Help

Welcome to the zbMATH interface.
Explore our new features:

- Feedback to improve document assignments can be given directly at the author profile pages or via [this link](#) (if you already know the author ID).
- The formula facet in the [structured search](#) now allows for a free combination of formula search with other query types. Formula queries can now also be refined by filters.

Home Advanced Search Browse by Subject Browse by Journals Refs Lookup

Search

Enter your search terms to get started

Title, Author, Keyword, Citation, Date...

Advanced Search

[Embed EuDML Search on Your Website](#)

Search Tips

- search is case and diacritics insensitive (**Bézout** = **bezout**)
- search is performed on exact words as typed (**theorem # theorems**)
- phrases are supported with quote notation ("Uniformization theorem" # Uniformization theorem = uniformization AND theorem)
- wildcards * and ? can be used (except in phrases)

EuDML is currently indexing **234125** items across **14** collections [more statistics](#)

- ~3.6 million publications (almost complete in mathematics since 1868) indexed; ~2.4 million full text links (doi, arXiv, EuDML...); ~900,000 authors, ~16 million references
- classified by Mathematical Subject Classification
- most of core articles are reviewed independently
- TeX encoded, MathML display

- >240,000 math full-texts
- almost 300 journals available via (sometimes moving wall) OA
- Metadata TeX encoded, documents mostly pdf

 **FIZ Karlsruhe**
Leibniz Institute for Information Infrastructure

 The European
Mathematical Society

 HEIDELBERG ACADEMY OF
SCIENCES AND HUMANITIES

© 2014 FIZ Karlsruhe GmbH Privacy Policy Legal Notices Terms & Conditions

 W3C

What is a digital mathematics library? (II)

More recent digital libraries concept as collection of digital objects with a management -> more granular math entities may be contained, e.g.:

Let p be a prime factor with multiplicity n of the order of a finite group G , so that the order of G can be written as $p^n m$, where $n > 0$ and p does not divide m . Let n_p be the number of Sylow p -subgroups of G . Then the following hold:

- (1) p^n divides m .
- (2) $n_p \equiv 1 \pmod{p}$.
- (3) $n_p = |G : N_G(P)|$, where P is any Sylow p -subgroup of G and N_G denotes the normalizer.

```
[/\ \ \forall P, [max P | p.-subgroup(G) P] = p.-Sylow(G) P,  
  [transitive G, on 'Syl_p(G) | 'JG],  
  \forall P, p.-Sylow(G) P \rightarrow \#| 'Syl_p(G) | = \#| G : 'N_G(P) |  
  & prime p \rightarrow \#| 'Syl_p(G) | %% p = 1%N].
```

Coq formalization,
by Freek Wiedijk

The shape of things to come?

NRC report “Developing a 21st Century Global Library for Mathematics Research”

(The National Academies Press, http://www.nap.edu/catalog.php?record_id=18619, <http://arxiv.org/abs/1404.1905>)

Some theses:

- Digitization in (classical) sense is already at a mature state (arXiv, EuDML, ...), diminishing returns from retrodigitization
- More value is expected from creating a quality digital mathematics information resource “of greater value than the sum of its contributing parts”
- Automated recognition of mathematical objects (e.g., theorems, proofs, sequences, groups) is not yet possible but should be seen as long-term goal
- Existing scalable methods and algorithms can be improved to assist intellectual analysis of mathematical concepts in the research literature

IMU GDML working group

2014 IMU GDML Panel (ICM Seoul; Bouche, Daubechies, Greuel, Ion, Zhang) related to the installment of a working group (Ion (chair), Bouche, Buchberger, Kohlhase, Pitman, Teschke, Watt, Weisstein) to identify and pursue realizable components of a GDML architecture

Mission statement: The GDML should

- enhance openness and accessibility of all mathematical knowledge world-wide
- serve research mathematics, education and the scientific and technological use of mathematics.
- be a resource for developing tools to promote use and development of mathematics.
- facilitate creation, dissemination and archiving of semantically annotated mathematical material.
- encourage the collaborative development of services based on semantic annotation.
- be a truly global resource, shaped as a distributed service which matches the highest possible standards of independence from national interests, of reliability, and of data protection

Disadvantages of classical publications for mathematical modelling and simulation

- Inadequate presentation of software
- Inadequate presentation of data
- Inadequate presentation of models
- Incomplete presentation of simulation results
- ...

Necessary improvements in a GDML framework

Linked open data concepts should facilitate

- Software and data repositories in appropriate formats
- Sustainable links
- Semantic layers to enable appropriate organization and retrieval

Measures to gain and improve semantic information:

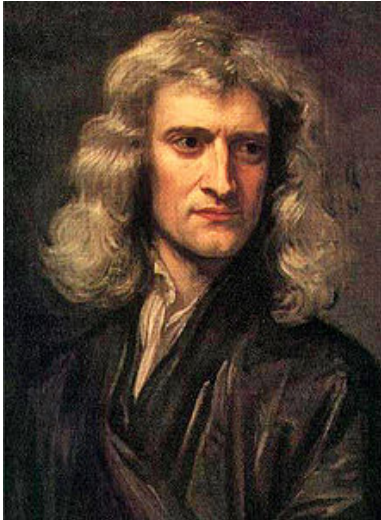
- Appropriate model annotation
- Link-generated enrichment
- Formats for human-computer interaction

Most important:

Community should define appropriate policies for research data

User-driven (bottom-up) approaches complement heuristic (top-down) information

Semantics as modelling layer



$\dot{y}$

$\frac{dy}{dt}$



Transport of Heat

$$\partial_t(\rho^{\text{bulk}} c_p^{\text{bulk}} T) + \text{div}(\vec{u} \rho^{\text{fluid}} c_p^{\text{fluid}} T) - \text{div}(\lambda \text{grad} T) - S^T = 0$$

Hydraulics

$$\rho S \partial_t p - \text{div}\left(\rho \frac{K}{\mu} \text{grad}(p)\right) + \text{div}\left(\rho^2 \frac{K}{\mu} \vec{g}\right) - S^m = 0$$

Mechanics

$$\sigma_{kl} = D_{ijkl} \epsilon_{kl}$$

Chemistry

$$\partial_t(C^j \rho^{\text{fluid}}) + \text{div}(\vec{u} \rho^{\text{fluid}} C^j) - \text{div}(D^j \text{grad} C^j) - S^{C^j} = 0$$

The shape of things to come? (II) - Algorithms, data, and semantics

- Increasingly powerful generic algorithms
- Increasing value of the appropriate data fed into these algorithms
- Even more increasing value of the appropriate model



Thank you !



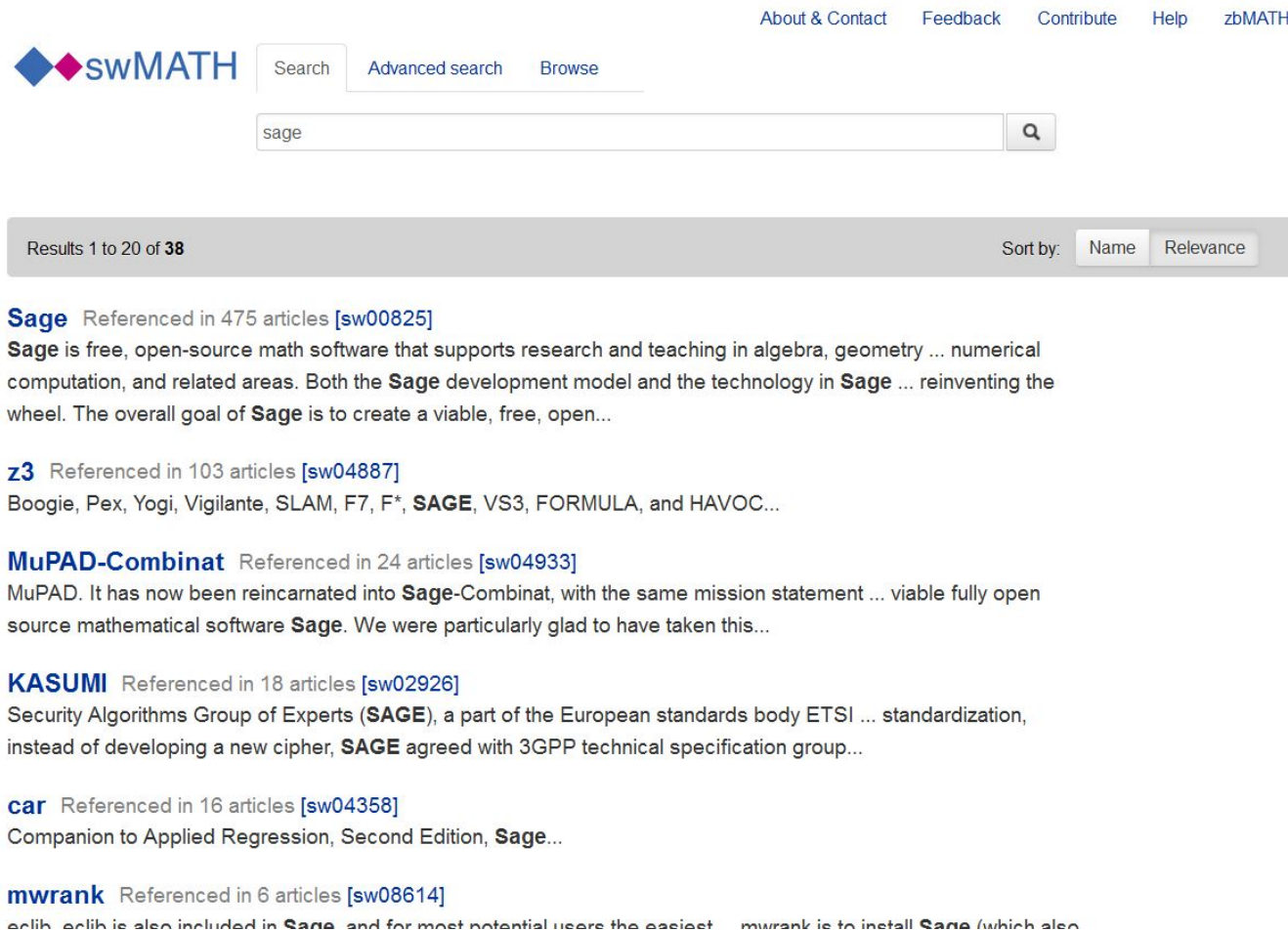
A top-down view on semantics: Content Analysis in Mathematics – Why?

- Embedding a publication in its general mathematical context
 - Standardized content description
-
- Easy searching for content also if no further content description like abstracts available
 - Possibility for visualization of embedded content
 - Possibility for data analysis
 - Possibility for search interaction (proposal of search terms)

Mathematical software as additional layer (I)

Since 2011, software information is extracted from the corpus mainly by heuristics + intellectual refinement

As of Jun 2015, 10.064 systems and packages have been linked to 89.124 publications



The screenshot displays the swMATH search interface. At the top right, there are links for 'About & Contact', 'Feedback', 'Contribute', 'Help', and 'zbMATH'. The swMATH logo is on the left. Below it, there are search options: 'Search', 'Advanced search', and 'Browse'. A search bar contains the text 'sage' with a magnifying glass icon. Below the search bar, it says 'Results 1 to 20 of 38'. On the right, there are sorting options: 'Sort by: Name' and 'Relevance'. The search results list several items:

- Sage** Referenced in 475 articles [sw00825]
Sage is free, open-source math software that supports research and teaching in algebra, geometry ... numerical computation, and related areas. Both the Sage development model and the technology in Sage ... reinventing the wheel. The overall goal of Sage is to create a viable, free, open...
- z3** Referenced in 103 articles [sw04887]
Boogie, Pex, Yogi, Vigilante, SLAM, F7, F*, SAGE, VS3, FORMULA, and HAVOC...
- MuPAD-Combinat** Referenced in 24 articles [sw04933]
MuPAD. It has now been reincarnated into Sage-Combinat, with the same mission statement ... viable fully open source mathematical software Sage. We were particularly glad to have taken this...
- KASUMI** Referenced in 18 articles [sw02926]
Security Algorithms Group of Experts (SAGE), a part of the European standards body ETSI ... standardization, instead of developing a new cipher, SAGE agreed with 3GPP technical specification group...
- car** Referenced in 16 articles [sw04358]
Companion to Applied Regression, Second Edition, Sage...
- mwrnk** Referenced in 6 articles [sw08614]
... mwrnk is to install Sage (which also

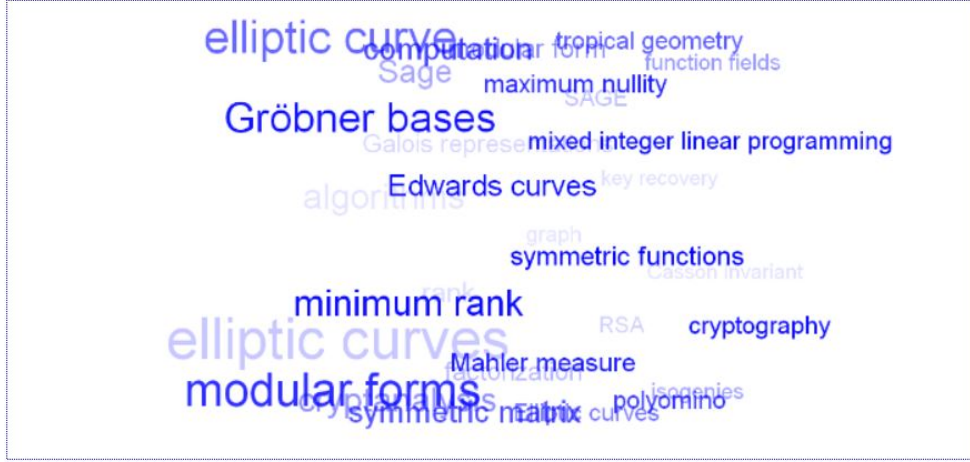
Mathematical software as additional layer (II)

 Search Advanced search Browse

Sage
Sage is free, open-source math software that supports research and teaching in algebra, geometry, number theory, cryptography, numerical computation, and related areas. Both the Sage development model and the technology in Sage itself are distinguished by an extremely strong emphasis on openness, community, cooperation, and collaboration: we are building the car, not reinventing the wheel. The overall goal of Sage is to create a viable, free, open-source alternative to Maple, Mathematica, Magma, and MATLAB.

 This software is also referenced in ORCID.

 Keywords for this software



elliptic curves, computation, tropical geometry, function fields, Gröbner bases, maximum nullity, mixed integer linear programming, Edwards curves, key recovery, algorithms, symmetric functions, minimum rank, RSA, cryptography, elliptic curves, Mahler measure, modular forms, symmetric matrix, isogenies, polyomino

References in zbMATH (referenced in 475 articles):

Showing results 1 to 20 of 475.

1 2 3 ... 22 23 24 next

1. Albrecht, Martin R.; Cid, Carlos; Faugère, Jean-Charles; Fitzpatrick, Robert; Perret, Ludovic: On the complexity of the BKW algorithm on UWE (2015)
2. Altman, Joshua; Lian, Carl; Tran, Brandon: Circular planar electrical networks, posets and positivity (2015)
3. Anastassiou, George A.; Medel, Rocco A.: Numerical analysis using Sage (2015)
4. Ash, Armin; Gundersen, Paul E.; McConnell, Mark: Mod 2 homology for GL(4) and Galois representations (2015)
5. Ayer, Arvind; Schilling, Anne; Steinberg, Benjamin; Thiéry, Nicolas M.: Directed nonabelian sandpile models on trees (2015)

While the area of application is well reflected by corresponding documents, there is up to now no software classification/ontology representational typical facets like dependencies or functions (e.g., computer algebra systems, solvers, interfaces...)

URL: www.sagemath.org
Authors: William Stein; David Joyner; David Kohel; Eric All, Burgh
Dependencies: Sage
Add information on this software.
Similar software:
Magma
GAP
SINGULAR
OEIS
RANKGP
Macaulay2
Maple
Mathematica
FGO
MUPAD
Show more...

Article statistics & filter:

Search for articles

MSC classification

 Top MSC classes
 65 Combinatorics

Formula search (I) - Overview

Motivation: Enable Access to formula content

Steps:

- Extract formula information from corpora
 - Statistical analysis
 - XML conversion (MathML 2.0 W3C standard), normalisation
- realized for zbMATH data to enable first prototype

Further challenges:

- Semantic analysis, disambiguation (example: ``what does dx/dy mean?'')
- Integration with textual information

Major tasks are the digitisation of math content, normalisation, and semantic enrichment

Formula search (II) - Example



Documents
Authors
Journals
Classification
Software
Formulæ

$$\int_a^b (f(x))^2 dx = r$$

Nowak, Werner Georg

Lattice points in a circle: an improved mean-square asymptotics. (English) Zbl 1092.11039

Acta Arith. **113**, No. 3, 259-272 (2004).

Summary: Let $P(x)$ denote the lattice point discrepancy (number of integer points minus area) of an origin-centered compact circular disc of radius $\sqrt{x}$. The question for the order of magnitude of $P(x)$, for large x , is known as the Gaussian circle problem. In fact, in this most classic topic, the sharpest known results have been established only quite recently [*M. N. Huxley*, *Proc. Lond. Math. Soc.* (3) **87**, 591–609 (2003; [Zbl 1065.11079](#))]: Huxley proved that

$$P(x) \ll x^{131/416} (\log x)^{18637/8320}, \quad \left(\frac{131}{416} = 0.3149\dots\right)$$

while *K. Soundararajan* [*Int. Math. Res. Not.* 2003, No. 36, 1987–1998; [Zbl 1130.11329](#)] showed that

$$P(x) \neq o\left((x \log x)^{1/4} (\log \log x)^{(3/4)(2^{1/3}-1)} (\log \log \log x)^{-5/8}\right).$$

It is known that $P(x) \ll x^{1/4}$ in mean-square. More precisely,

$$\int_0^x (P(x))^2 dx = CX^{3/2} + Q(X), \quad C \approx 1.69396.$$

For this new error term $Q(X)$, the present article provides the refined upper bound

$$Q(X) \ll X(\log X)^{3/2} \log \log X,$$

Formula search (III) – Usability



Documents Authors Journals Classification Software Formulae

|int 0 infy frac [sin x] [x] dx|

mark all display marked items

Page 1 of 1 first prev next last

Found 6 documents (Results 1–6)

Fischer, Hans

The history of the integral $\int_0^{\infty} \frac{\sin x}{x} dx$: a history of analysis in a nutshell. (Die Geschichte des Integrals $\int_0^{\infty} \frac{\sin x}{x} dx$: eine Geschichte der Analysis in der Nusschale.)

(German) [Zbl 1118.01014](#)

Math. Semesterber. 54, No. 1, 13-30 (2007).

MSC: 01A55 01A60 26-03 28-03 42-03

[BibTeX](#) [Full Text](#) [DOI](#) Reviewer: Ivor Grattan-Guinness (Middlesex)

Wasteels, C. E.

Méthode élémentaire de calcul d'une intégrale définie. (French) [JFM 42.0387.01](#)

Mathesis (4) 2, 8-10 (1912).

[BibTeX](#) Reviewer: Manson, Prof. (Gent) (Lampe, Prof. (Berlin))

Hardy, G. H.

The integral $\int_0^{\infty} \frac{\sin x}{x} dx$. (English) [JFM 40.0341.04](#)

Math. Gazette 5, 98-103 (1909).

[BibTeX](#) [Full Text](#) [DOI](#) Reviewer: Lampe, Prof. (Berlin)

Nanson, E. J.

Note on the integral $\int_0^{\infty} \frac{\sin x}{x} dx$. (English) [JFM 38.0340.04](#)

Messenger (2) 37, 113-114 (1907).

[BibTeX](#) Reviewer: Lampe, Prof. (Berlin)

Berry, A.

A note on the integral $\int_0^{\infty} \frac{\sin x}{x} dx$. (English) [JFM 38.0340.04](#)

Messenger (2) 37, 61-62 (1907).

[BibTeX](#) Reviewer: Lampe, Prof. (Berlin)

Glaisher, J. W. L.

On $\sin \infty$ and $\cos \infty$. (English) [JFM 03.0002.01](#)

Messenger, V. 232-244 (1869).

MSC: 21810

[BibTeX](#) Reviewer: Glaisher, Prof. (Cambridge) (Ohrtmann, Dr. (Berlin))



About Contact General Help Reviewer Service Subscription Preferences Log-Out

Documents Authors Journals Classification Software Formulae

Structured Search

jan:05135828

Q Fields Operators

Neu anschauen Sing-N

Help

Fischer, Hans

The history of the integral $\int_0^{\infty} \frac{\sin x}{x} dx$: a history of analysis in a nutshell. (Die Geschichte des Integrals $\int_0^{\infty} \frac{\sin x}{x} dx$: eine Geschichte der Analysis in der Nusschale.) (German) [Zbl 1118.01014](#)

Math. Semesterber. 54, No. 1, 13-30 (2007).

Filter re:

Authors

Wasteels,

Nanson, E

Hardy, G

Glaisher, J

Fischer, H

Journals

Mathesis (

Math. Sem

Math. Gaz.

Classifica

42-XX (1)

26-XX (1)

28-XX (1)

24-XX (1)

01-XX (1)

Publicatio

2007 (1)

1912 (1)

1909 (1)

1907 (2)

1869 (1)

What with its oscillating integrand and the infinite range of its variable, this integral was rather hard for the power of the theory of integration of the early 19th century. As the author suggests in the sub-title of his paper, it proved a formidable test case in the development of the evaluation of definite integrals from Euler through Laplace, Fourier, Poisson and Cauchy to a more critical phase featuring Dirichlet and Bonnet and then to assessments at the turn of the centuries by figures such as de la Vallée Poussin and Hardy.

Reviewer: Ivor Grattan-Guinness (Middlesex)

MSC:

01A55 Mathematics in the 19th century

01A60 Mathematics in the 20th century

26-03 Historical (real functions)

28-03 Historical (measure and integration)

42-03 Historical (Fourier analysis)

[BibTeX](#)

Full Text: [DOI](#)

References:

- [1] Abel, N.H.: Recherches sur la série $1 + \frac{m}{1}x + \frac{m(m-1)}{1 \cdot 2}x^2 + \frac{m(m-1)(m-2)}{1 \cdot 2 \cdot 3}x^3 + \dots$. J. Reine Angew. Math. 1, 219–250 (1826). Alle Seitenangaben gemäß Wiederabdruck in OEuvres complètes, T. 1, 219–250, Grøndahl, Kristiania, 1881
- [2] Bonnet, O.: Remarques sur quelques intégrales définies. J. Math. Pure Appl. 14(1), 249–256 (1849)

Example:

Naive pure-text-searches may still provide the best results.

Formula search (IV) – usability (II)

Extra challenge: find a way to create a user-friendly gateway to formula retrieval!

SchemaSearch

derived from automated theorem proving: Substitution Tree Indexing. MWS performs mathematical full-text search, combining key phrase search with unification-based formula search.

SchemaSearch augments the power of MathWebSearch by providing faceted search capabilities. A math facet consists of a formula in which qvars replace nodes below a certain depth in its CMML representation.

$\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2}$

5



Search

Enter a keyword in the search box to receive a list of formula schemata which cover the math in the documents containing the keyword. Each formula schemata returned is accompanied by a group of formulae which are instantiations of it.

You can also enter a depth for the schemata (how deep the schemata should be) and check the R checkbox if you would like this depth to be relative. If you do not check the box, absolute depth is assumed.

If the depth is relative, its value should be entered in percentages, e.g. for a depth of 50%, 50 should be entered for the relative depth.

90

?a

$$\begin{aligned} M_1 \\ [-\partial_x^2 + p(x) + q(x) - \lambda]f &= 0. \\ x &> 0 \\ x &< 0 \\ &= 0 \\ \pi \\ &= 0 \\ R_*^2 \\ &\equiv 0 \\ \phi(x) &\in C^1(R, R) \end{aligned}$$

Unsuccessful
approach (so far):
Enable formula
browsing based
on statistics.

Thank you !





These documents are intended for presentation purposes only.
Copyright lies with FIZ Karlsruhe.
Any distribution or use of these documents or part thereof is
subject to FIZ Karlsruhe's express approval.

© FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur GmbH 2016