



DML-CZ a jiné digitální matematické archívy

Katedra matematiky FAV ZČU v Plzni
Kolokvium 31. 3. 2011

- Kdyby se někomu podařilo zničit knihovny a všechny knihy, o třicet let později by žádný zámečnick neuměl ani utáhnout šroub.

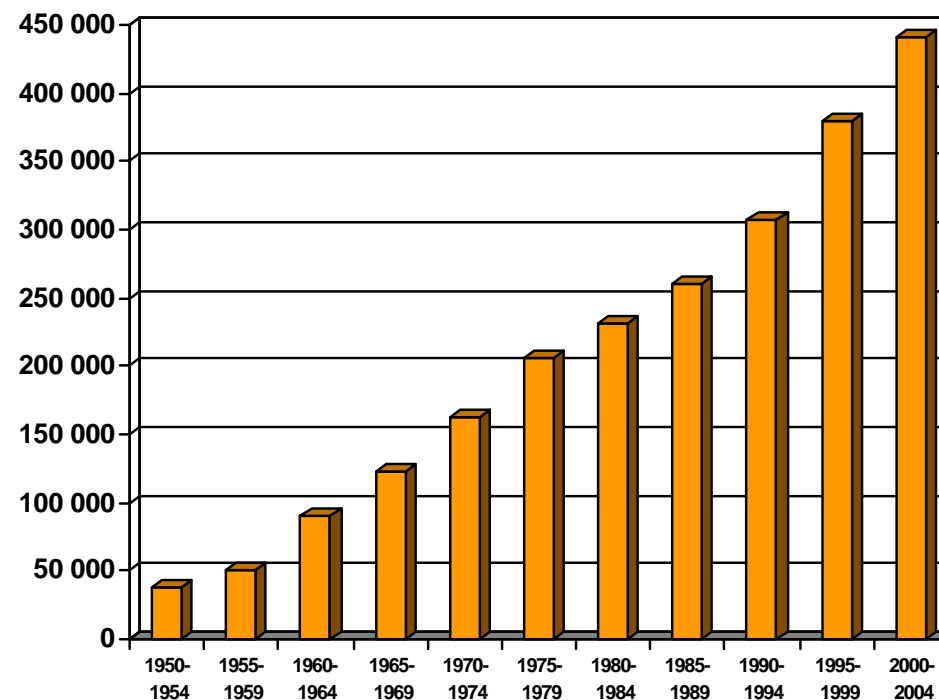
Thomas Alva Edison

- Hromada knih není knihovna a disk plný digitálních dokumentů není digitální knihovna.

Matematická literatura: nárůst

- současný rozsah zhruba
3 000 000 publikací
100 000 000 stránek
- nárůst se zrychluje
1/5 před r. 1900
1/2 po r. 1950
- [Zentralblatt MATH](#)
3 041 495 záznamů celkem
29 949 odkazů od začátku r. 2011
- 50 % odkazů směřuje 10 let zpět
25 % odkazů směřuje 20 let zpět

Pětileté přírůstky podle Zentralblatt MATH



Matematická literatura: milníky

- 1665 zrození mat. časopisů (*Journal des sçavans, Philosophical transactions*)
- 1800 asi 200 časopisů obsahuje matematické články
- 1810 první ryze matematický časopis (*Annales de mathématiques pures et appliquées, Annales de Gergonne*)
- 1850 ročně vychází asi 1000 matematických vědeckých článků
- 1872 *Časopis pro pěstování matematiky a fysiky*
- 1950 ročně vychází asi 6000 matematických vědeckých článků
- 1978–1986 T_EX – počítačový typografický systém
- 1992 arXiv, matematické preprinty (fyzikální od 1991)
- 1994 první nesespecializovaný čistě matematický elektronický časopis, volně přístupný (*New York Journal of Mathematics*)
- 2008 4,5 mil. stran digitalizováno, 65 % hlavních časopisů je dostupných digitálně

Matematická literatura: databáze

- 1868 *Jahrbuch über die Fortschritte der Mathematik*
- 1894 *Répertoire bibliographique des sciences mathématiques*
(„hodnotné” odkazy na literaturu z 19. století)
- 1931 *Zentralblatt für Mathematik und ihre Grenzgebiete*
- 1940 *Mathematical Reviews*
- 1990 elektronické verze databází (MathSci Disc, CompactMath)
a on-line přístup (telnet. . .)
- 1995 internetový přístup (MathSciNet, Zentralblatt MATH)

Matematická literatura

- nestárne
- ohromná stavba: každý stavebný kámen zůstává její součástí navždy – starší výsledky nejsou nahrazovány novými
- má smysl pouze jako celek propojený širokou sítí referencí
- je nadčasově užitečná ostatním vědním oborům
- musí být pečlivě archivována, indexována a uchovávána
- musí být dlouhodobě přístupná

Potřebujeme knihovnu

- co nejúplnější
- aktuální
- dobře uspořádanou
- široce přístupnou
- snadno použitelnou i pro nematematiky

NUMDAM

- Numérisation de documents anciens mathématiques
- <http://www.numdam.org/?lang=en>
- 32 časopisů, 29 seminárních sérií
- 763 000 stran, 43 000 článků
- většinou francouzské
- Ann. Scuola Norm. Sup. Pisa Cl. Sci.
- Compositio Mathematica
- Rend. Semin. Mat. Univ. Padova



- <http://gdz.sub.uni-goettingen.de/en/gdz/>
- matematická literatura: 6094 svazků

DigiZeitschriften

- <http://www.digizeitschriften.de/>
- matematika: 39 časopisů, 2518 svazků

RusDML

- Russian Digital Mathematics Library
- <http://www.rusdml.de/>
- 11 ruských časopisů od r. 1866
- 603 svazků

Polska Wirtualna Biblioteka Nauki

- **Kolekcja Matematyczna**
<http://matwbn.icm.edu.pl/>
- 14 sérií časopisů a monografií



The logo for the Spanish Digital Mathematics Library (DML) features the letters 'DML' in a stylized, handwritten font. A small red square with a white letter 'e' is positioned to the right of the 'L', indicating it is an electronic library.

- Spanish Digital Mathematics Library
- http://dmle.cindoc.csic.es/en/portada_en.php
- 22 španělských časopisů, 2 knihy

Srbská eLibrary

- <http://elibrary.matf.bg.ac.rs/>
- srbské matematické časopisy, knihy, disertace, ...

DML: Digital Mathematics Library

- http://www.math.uni-bielefeld.de/~rehmann/DML/dml_links.html
- Ulf Rehmann
- pozoruhodná databáze digitálních knihoven a archivů
- linky na 4514 digitalizovaných knih (> 643597 stran)
a 318 digitalizovaných časopisů a seminářů (> 4303824 stran)

EMIS

- European Mathematical Information Service
- <http://www.emis.de/>
- databáze: Zentralblatt, Jahrbuch, MathEduc, STMA-Z, io-port, MPRESS
- ELibM: > časopisů, > sborníků, 20 knih, software, obrázky, slovníky, ...

JSTOR



- <http://www.jstor.org/>
- 66 matematických časopisů

Projekt Euclid

- <http://projecteuclid.org/>
- Cornell University Library
- The Andrew W. Mellon Foundation
- 64 časopisů
- 1 446 751 stran (1 025 992 open access)
108 131 časopiseckých článků (78 827 open access)
161 knih (2 353 kapitol)
23 sborníků

Projekt DML-CZ

- program *Informační společnost (Akademie věd ČR)*
- 2005–2009
- 8,1 mil. Kč
- řešitelé
 - Matematický ústav AV ČR v Praze
 - Ústav výpočetní techniky MU v Brně
 - Fakulta informatiky MU v Brně
 - Matematicko-fyzikální fakulta UK v Praze
 - Knihovna AV ČR

Obsah

- časopisy, sborníky konferencí, monografie vydávané na území Česka
 - Časopis pro pěstování matematiky a fysiky od r. 1872
 - dílo Bernarda Bolzana
- v současnosti zpracováno
 - 275 000 stran
 - 30 000 článků/kapitol
 - 35 knih (připravuje se dalších 30)
- připravuje se
 - další knihy
 - sekce osobnosti české matematiky: O. Borůvka
 - časopisy pro výuku matematiky

Zdroje

- tisk na papíře
 - 1810–1990
 - předlohy pouze v tištěné podobě
 - klasická digitalizace (scan)
- retro-born-digital
 - 1991–2007
 - předlohy zčásti v digitální formě
 - různé formáty (ps, pdf, TEX, HTML) – konverze
 - v případě problémů OCR, popř. skenování
- digital-online
 - 2008–
 - automatizované přebírání nové literatury do DML-CZ

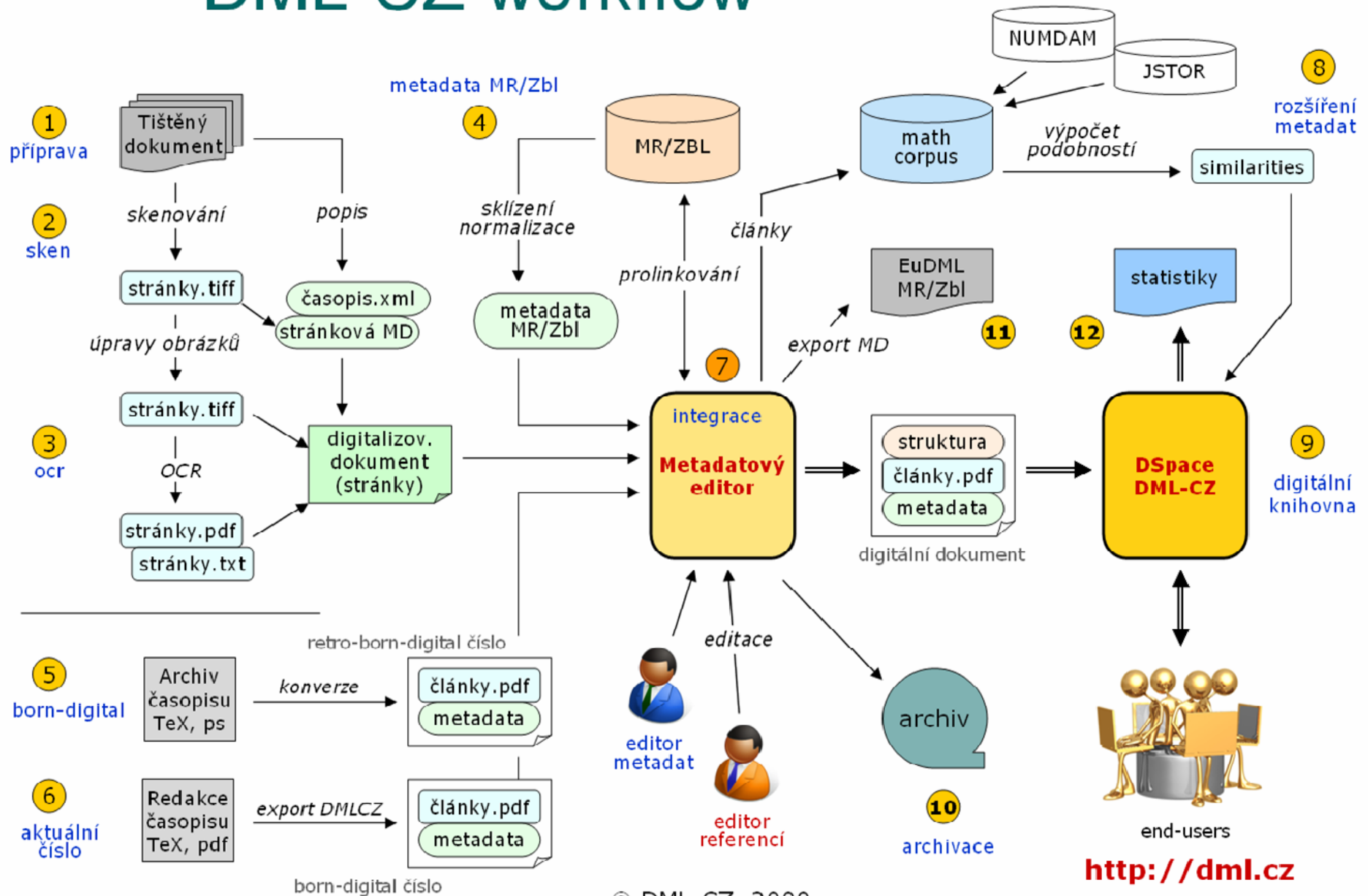
DML-CZ nabízí

- plné texty článků a kapitol cca 10 000 autorů ve formátu PDF
- úplná metadata včetně bibliografických referencí
- propojení článků i referencí do MathSciNet a Zentralblatt MATH
- procházení podle rejstříků
 - kolekce, názvy, autoři, MSC
- jednoduché i pokročilé vyhledávání
 - v metadatach, v plných textech
- vyhledávání obsahově podobných článků

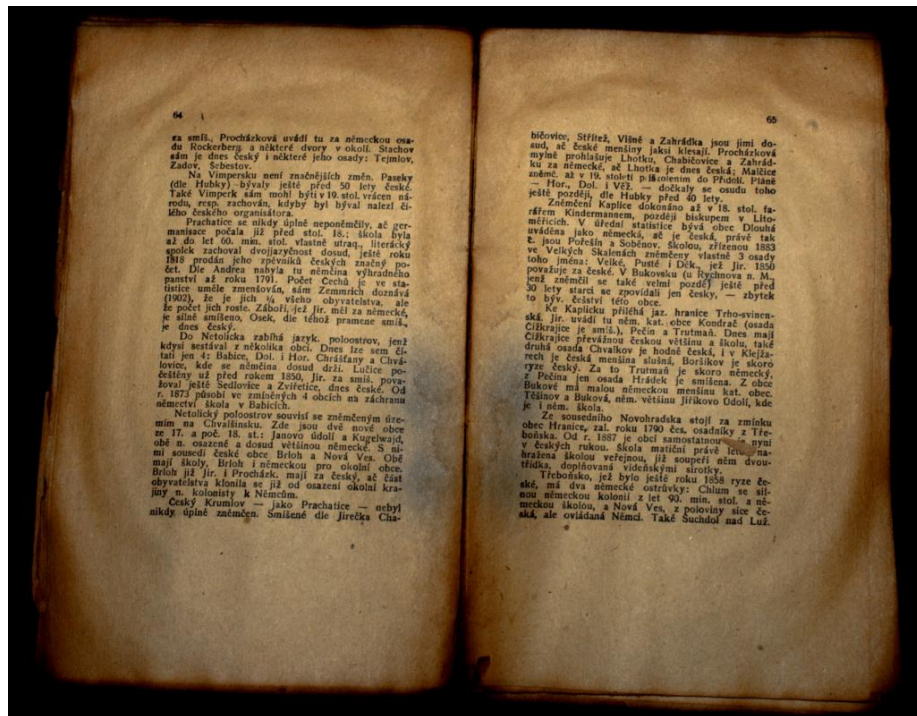
Jak se dělá digitální knihovna

1. Příprava materiálů
2. Digitalizace
 - a) Optical Character Recognition
 - b) Referenční metadata (MathSciNet, Zentralblatt MATH)
3. Retro-born-digital
4. Nově vznikající produkce (new-born-digital)
5. Integrace (Metadatový editor)
6. Podobné články
7. Digitální knihovna
8. Archivace
9. Exporty metadat do MR/ZBL, miniDML, EuDML

DML-CZ workflow



Rovnění, čištění



za smíš., Procházková uvádí tu za německou osadu Rockerberg a některé dvory v okolí. Stachov sám je dnes český i některé jeho osady: Tejmlov, Zadov, Šebestov.

Na Vimpersku není značnějších změn. Paseky (dle Hubky) -bývaly ještě před 50 lety české. Také Vimperk sám mohl býti v 19. stol. vrácen národu, resp. zachován, kdyby byl býval nalezl čilého českého organisátora.

Prachatic se nikdy úplně neponěmčily, ač germanisace počala již před stol. 18.; škola byla až do let 60. mn. stol. vlastně utraq., literárský spolek zachoval dvojjazyčnost dosud, ještě roku 1818 prodán jeho zpěvníků českých značný počet. Dle Andrea nabyta tu němčina výhradného panství až roku 1791. Počet Cechů je ve statistice uměle zmenšován, sám Zemmrich doznává (1902), že je jich 1/4 všeho obyvatelstva, ale že počet jich roste. Zábofí, jež Jir. měl za německé, je silně smíšeno, Osek, die téhož pramene smíš., je dnes český.

Do Netolicka zabíhá jazyk. poloostrov, jenž kdysi sestával z několika obcí. Dnes lze sem čítati jen 4: Babice, Dol. i Hor. Chrástany a Chválovic, kde se němčina dosud drží. Lučice počestěny už před rokem 1850, Jir. za smíš. považoval ještě Sedlovice a Zvířetice, dnes české. Od r. 1873 působí ve zmíněných 4 obcích na záchranu německví škola v Babicích.

Netolický poloostrov souvisí se změněným územím na Chválsínsku. Zde jsou dvě nové obce ze 17. a poč. 18. st.: Janovo údolí a Kugelwajd, obě n. osazené a dosud většinou německé. S nimi sousedí české obce Brloh a Nová Ves. Obě mají školy, Brloh i německou pro okolní obce. Brloh již Jir. i Procházka mají za český, ač část obyvatelstva klonila se již od osazení okolní krajiny n. kolonisty k Němcům.

Český Krumlov — jako Prachatic — nebyl nikdy úplně zmáčen. Smíšeno dle Jirečka Cha-

za smíš., Procházková uvádí tu za německou osadu Rockerberg a některé dvory v okolí. Stachov sám je dnes český i některé jeho osady: Tejmlov, Zadov, Šebestov.

Na Vimpersku není značnějších změn. Paseky (dle Hubky) -bývaly ještě před 50 lety české. Také Vimperk sám mohl býti v 19. stol. vrácen národu, resp. zachován, kdyby byl býval nalezl čilého českého organisátora.

Prachatic se nikdy úplně neponěmčily, ač germanisace počala již před stol. 18.; škola byla až do let 60. mn. stol. vlastně utraq., literárský spolek zachoval dvojjazyčnost dosud, ještě roku 1818 prodán jeho zpěvníků českých značný počet. Dle Andrea nabyta tu němčina výhradného panství až roku 1791. Počet Cechů je ve statistice uměle zmenšován, sám Zemmrich doznává (1902), že je jich 1/4 všeho obyvatelstva, ale že počet jich roste. Zábofí, jež Jir. měl za německé, je silně smíšeno, Osek, die téhož pramene smíš., je dnes český.

Do Netolicka zabíhá jazyk. poloostrov, jenž kdysi sestával z několika obcí. Dnes lze sem čítati jen 4: Babice, Dol. i Hor. Chrástany a Chválovic, kde se němčina dosud drží. Lučice počestěny už před rokem 1850, Jir. za smíš. považoval ještě Sedlovice a Zvířetice, dnes české. Od r. 1873 působí ve zmíněných 4 obcích na záchranu německví škola v Babicích.

Netolický poloostrov souvisí se změněným územím na Chválsínsku. Zde jsou dvě nové obce ze 17. a poč. 18. st.: Janovo údolí a Kugelwajd, obě n. osazené a dosud většinou německé. S nimi sousedí české obce Brloh a Nová Ves. Obě mají školy, Brloh i německou pro okolní obce. Brloh již Jir. i Procházka mají za český, ač část obyvatelstva klonila se již od osazení okolní krajiny n. kolonisty k Němcům.

Český Krumlov — jako Prachatic — nebyl nikdy úplně zmáčen. Smíšeno dle Jirečka Cha-

Metadata

- autor, název, časopis, ročník, číslo, stránky, abstrakt, klíčová slova, klasifikace oborů, jazyk textu, odkazy do databází, seznam citované literatury, rozměry, ISSN, ISBN, ...
- časově nejnáročnější, lze jen zčásti automatizovat
- specifické požadavky matematiky (symboly, formule, MSC, ...)
- vícejazyčnost textů
 - automatické nastavování OCR
 - překládání názvů
 - přepis jmen
- registr autorit
- různorodý obsah (matematika, fyzika, chemie, výuka, zprávy, oznámení, recenze, politika, obrázky, ...)

La torsion $\tau = A/H^2$ peut être discontinue en un point (t, λ) où $A = B = C = 0$, donc $\varrho = 0$. C'est ce qui arrive pour $\lambda = \frac{1}{4}$, $t = 0$. D'ailleurs, pour $\lambda = \frac{1}{4}$ on retrouve la courbe non-orientable d'ordre 2 que nous avons déjà présentée au No. 3. On doit remarquer que la torsion τ reste bornée sur cette courbe $\lambda = \frac{1}{4}$, mais $\tau(t, \lambda)$ n'est pas bornée au voisinage du point $t = 0$, $\lambda = \frac{1}{4}$. En effet, on trouve

$$\tau = \frac{6\lambda^2[(4\lambda - 1) + (\cos t - 1)(4\lambda - 2 \cos t - 2)]}{(4\lambda - 1)^2(5\lambda^2 - 4\lambda + 1) + (\cos t - 1)[12\lambda^2(\lambda^2 + 3)(\cos t + 1) - 4\lambda(3 + 26\lambda^2)]}$$

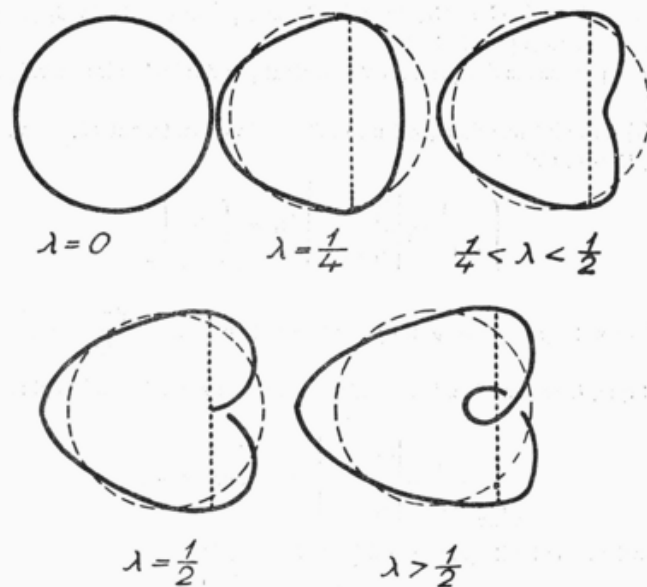


Fig. 9.

et l'on voit que pour $t \rightarrow 0$, $\lambda \rightarrow \frac{1}{4}$ la limite de ce rapport dépend de $\lim \frac{\cos t - 1}{4\lambda - 1}$;

cette limite est infinie si $\lim \frac{\cos t - 1}{4\lambda - 1} = 0$.

On a $K = 0$ pour $0 \leq \lambda < \frac{1}{4}$ et $K = 1$ pour $\frac{1}{4} < \lambda \leq 1$, et ce saut est dû à la discontinuité de τ pour $\lambda = \frac{1}{4}$. Ce résultat est général:

Si C_1 et C_2 sont deux noeuds pour lesquels K prend des valeurs de parité différente ($K_2 - K_1 = 2h + 1$), pendant toute déformation continue (les dérivées des coordonnées jusqu'à l'ordre 3 étant continues) de C_1 en C_2 la torsion $\tau(s, \lambda)$ passe nécessairement par un point de discontinuité.



КЛЕМЕНТ ГОТВАЛЬД

1896—1953



ИОСИФ ВИССАРИОНОВИЧ СТАЛИН

1879—1953

PDF článků

- Spojení stránkových PDF do PDF-článku
 - 2-vrstvá pdf (obraz stránky + OCR text)
 - digitální podpis + ochrany
 - generovaná titulní stránka
 - název článku s vysázenou matematikou
 - bibliografická citace článku
 - PURL, MSC, Zbl a MR
 - terms of use
 - komprese
- Import do digitální knihovny
 - PDF + metadata + struktura



DML-CZ

DML-CZ

<http://dml.cz>

<http://project.dml.cz>

Proč tolik úsilí?

- Máme přece Google!!!
- ???

Jak dál?



*Eu*DML

The EUROPEAN DIGITAL
MATHEMATICS LIBRARY



*Eu*DML | *The* EUROPEAN DIGITAL MATHEMATICS LIBRARY

Komplexní řešení

- vytvořit společnou infrastrukturu pro bezbariérovou navigaci, vyhledávání a interakci v rámci hluboce prolinkované sítě ověřené digitální literatury dostupné v různých místech Evropy
- zajistit bezpečnou archivaci, aby se vydavatelé nemuseli starat o uchovávání a aby umožnili volné zpřístupnění
- splnit požadavek spolehlivé a dlouhodobé dostupnosti matematických vědeckých výsledků

Organizace, financování

- EC Competitiveness and Innovation Framework Programme
 - Information and Communications Technologies Policy Support Programme / Open Access to Scientific Information
- 3,2 M€ (1,6 M€ od Evropské komise)
- únor 2010 – leden 2013

Řešitelé

- Instituto Superior Técnico, Lisabon – General co-ordinator
- Cellule MathDoc, Université Joseph Fourier, Grenoble – Scientific co-ordinator
- University of Birmingham, School of Computer Science
- Fachinformationszentrum Karlsruhe / Zentralblatt MATH
- Masarykova univerzita Brno, Fakulta Informatiky
- Univerzytet Warszawski, Interdisciplinary Centre for Math. and Comput. Modelling
- Instituto de Estudios Documentales sobre Ciencia y Tecnología – IEDCYT, Madrid
- EDP Sciences, Paris
- University of Santiago de Compostela, Institute of Mathematics
- Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences
- Matematický ústav AV ČR
- Ionian University, Department of Informatics, Corfu
- Made Media Ltd, Birmingham
- Centre National de la Recherche Scientifique / Cellule MathDoc, Grenoble

Partneři

- Evropská matematická společnost
- Niedersächsische Staats- und Universitätsbibliothek Göttingen
- RusDML (Russian Digital Mathematics Library)

- eLibrary of Mathematical Institute of the Serbian Academy of Sciences and Arts
- SwissDML
- Biblioteca Digitale Italiana di Matematica
- ...



Vítejte v DML-CZ

<http://dml.cz>

Děkuji za pozornost