



Mathdoc



La bibliothèque numérique des mathématiques

Thierry Bouche

Cellule Mathdoc, Université Joseph-Fourier (Grenoble 1)

IRCAM

Paris, jeudi 13 juin 2013

Synth se

- 1 La biblioth que num rique de math matiques
- 2 P rennit ?
- 3 Math matiques et num rique
- 4 \LaTeX et l' dition des maths
- 5 Et ensuite?

1 La bibliothèque numérique de mathématiques

2 Pérennité ?

3 Mathématiques et numérique

4 \LaTeX et l'édition des maths

5 Et ensuite ?

La documentation en mathématiques

Enjeux spécifiques

« *Les mathématiciens se contentent de mettre leur production à la disposition de tous, comme sur des étagères où l'on peut venir se servir.* »

(Jean-Pierre Serre)

- La documentation mathématique *validée* ne se périmé pas
(Les résultats anciens ne sont pas remplacés par les nouveaux : ils sont leur fondation)
- Elle fonctionne comme un *tout*, qui forme un vaste réseau interdépendant
- Elle est utile pour d'autres sciences, de façon *asynchrone*

⇒ Elle doit donc être soigneusement validée, rangée, indexée et conservée

⇒ Elle doit rester accessible sur le très long terme

La documentation en mathématiques

Enjeux spécifiques

« Les mathématiciens se contentent de mettre leur production à la disposition de tous, comme sur des étagères où l'on peut venir se servir. »

(Jean-Pierre Serre)

- La documentation mathématique *validée* ne se périmé pas
(Les résultats anciens ne sont pas remplacés par les nouveaux : ils sont leur fondation)
- Elle fonctionne comme un *tout*, qui forme un vaste réseau interdépendant
- Elle est utile pour d'autres sciences, de façon *asynchrone*

⇒ Elle doit donc être soigneusement validée, rangée, indexée et conservée

⇒ Elle doit rester accessible sur le très long terme

La documentation en mathématiques

Enjeux spécifiques

« Les mathématiciens se contentent de mettre leur production à la disposition de tous, comme sur des étagères où l'on peut venir se servir. »

(Jean-Pierre Serre)

- La documentation mathématique *validée* ne se périmé pas
(Les résultats anciens ne sont pas remplacés par les nouveaux : ils sont leur fondation)
- Elle fonctionne comme un *tout*, qui forme un vaste réseau interdépendant
- Elle est utile pour d'autres sciences, de façon *asynchrone*

⇒ Elle doit donc être soigneusement validée, rangée, indexée et conservée

⇒ Elle doit rester accessible sur le très long terme

La biblioth que de r f rence

Nous avons donc besoin d'une biblioth que *de r f rence*

- exhaustive
-   jour
- bien rang e
- p renne
- grande ouverte
- facile d'acc s pour les non-math maticiens

Et num rique !

⇒ Le projet WDML
The World Digital Mathematical Library

En m me temps que :

- Une infrastructure pour les pr publications (arXiv, HAL, DRIVER, etc.)
- Un syst me d' dition qui s lectionne, valide et produit le corpus de r f rence

La biblioth que de r f rence

Nous avons donc besoin d'une biblioth que *de r f rence*

- exhaustive
-   jour
- bien rang e
- p renne
- grande ouverte
- facile d'acc s pour les non-math maticiens

Et num rique !

⇒ Le projet WDML
The World Digital Mathematical Library

En m me temps que :

- Une infrastructure pour les pr publications (arXiv, HAL, DRIVER, etc.)
- Un syst me d' dition qui s lectionne, valide et produit le corpus de r f rence

La bibliothèque de référence

Nous avons donc besoin d'une bibliothèque *de référence*

- exhaustive
- à jour
- bien rangée
- pérenne
- grande ouverte
- facile d'accès pour les non-mathématiciens

Et numérique !

⇒ Le projet WDML

La bibliothèque numérique de mathématiques

En même temps que :

- Une infrastructure pour les prépublications (arXiv, HAL, DRIVER, etc.)
- Un système d'édition qui sélectionne, valide et produit le corpus de référence

La biblioth  que de r  f  rence

Nous avons donc besoin d'une biblioth  que *de r  f  rence*

- exhaustive
-    jour
- bien rang  e
- p  renne
- grande ouverte
- facile d'acc  s pour les non-math  maticiens

Et num  rique !

⇒ Le projet WDML

La biblioth  que num  rique de math  matiques

En m  me temps que :

- Une infrastructure pour les pr  publications (arXiv, HAL, DRIVER, etc.)
- Un syst  me d'  dition qui s  lectionne, valide et produit le corpus de r  f  rence

La vision EuDML



The Digital Mathematics Library should assemble **as much as possible** of the digital mathematical corpus in order to

- make it **available online**
 - with **eventual open access**,
 - in the form of an **authoritative** and **enduring** digital collection,
 - **growing** continuously with publisher supplied new content,
 - **augmented** with sophisticated search interfaces and interoperability services,
 - developed and curated by a network of **institutions**
 - interconnected with relevant infrastructures
- + and **preserve** it over the long term

La vision EuDML



La bibliothèque numérique de mathématiques devrait s'efforcer de réunir un corpus mathématique **aussi vaste que possible** pour

- le rendre **disponible en ligne**
 - en accès **libre à terme**,
 - sous la forme d'une collection **de référence**,
 - **alimentée** en continu par les nouveautés des éditeurs,
 - **valorisée** par des outils de recherche et référencement sophistiqués,
 - développée et entretenue par un réseau d'**institutions**
 - interconnecté avec les infrastructures pertinentes
- + et le **préserver** à très long terme

EuDML : La conjecture

D'o  le nom

$$U = \frac{W}{2}$$



$$\text{Europe} = W(\text{orld})/2$$



EuDML is half the effort to build the WDML

- Conjecture optimiste !
- Projet financ  par CIP ICT PSP « Libre acc s   l'information scientifique »
- Dur e 3 ans ; Budget 1,6 M  ; Travail 487 PM
- Ach v  le 31 janvier 2013,  valu  (positivement) le 15 mars 2013.

EuDML : La conjecture

D'o  le nom

$$U = \frac{W}{2}$$

↓

$$\text{Europe} = W(\text{orld})/2$$

↓

EuDML is half the effort to build the WDML

- Conjecture optimiste !
- Projet financ  par CIP ICT PSP « Libre acc s   l'information scientifique »
- Dur e 3 ans ; Budget 1,6 M  ; Travail 487 PM
- Achev  le 31 janvier 2013,  valu  (positivement) le 15 mars 2013.

Les résultats

- Une masse critique de contenus
225 000 textes de tous les types prévus
- Un réseau de 12 partenaires
en croissance lente
- Une bibliothèque numérique dédiée pour les maths
métadonnées MathML, recherche math, MSC, liens bases biblio
- Un site Web avec un *look* et une ergonomie innovante
- Un certain nombre d'outils permettant au service de fonctionner
certaines robustes, d'autres plus expérimentaux

⇒ Accès facilité et visibilité accrue

⇒ Pérennité « subsidiarisée », donc incertaine

Quelle conservation pour ce corpus numérique ?

Les résultats

- Une masse critique de contenus
225 000 textes de tous les types prévus
 - Un réseau de 12 partenaires
en croissance lente
 - Une bibliothèque numérique dédiée pour les maths
métadonnées MathML, recherche math, MSC, liens bases biblio
 - Un site Web avec un *look* et une ergonomie innovante
 - Un certain nombre d'outils permettant au service de fonctionner
certaines robustes, d'autres plus expérimentaux
- ⇒ Accès facilité et visibilité accrue
- ⇒ Pérennité « subsidiarisée », donc incertaine
Quelle conservation pour ce corpus numérique ?

1 La bibliothèque numérique de mathématiques

2 Pérennité ?

3 Mathématiques et numérique

4 \LaTeX et l'édition des maths

5 Et ensuite ?

Pérennité du corpus « français »

- Gallica/BNF \Rightarrow SPAR
- TEL/CCSD \Rightarrow CINES/PAC
- NUMDAM : en attente
 - Label BSN \Rightarrow CINES ?
 - Problème des coûts récurrents
 - Réseau CLOCKSS EuDML ?

Une archive numérique est maintenue tant qu'on s'en sert...

OAIS est-il un modèle sûr ?

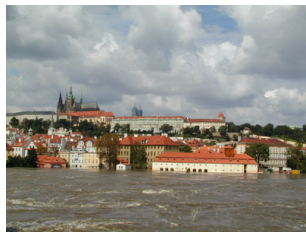
Pérennité du corpus « français »

- Gallica/BNF \Rightarrow SPAR
- TEL/CCSD \Rightarrow CINES/PAC
- NUMDAM : en attente
 - Label BSN \Rightarrow CINES ?
 - Problème des coûts récurrents
 - Réseau CLOCKSS EuDML ?

Une archive numérique est maintenue tant qu'on s'en sert...

OAIS est-il un modèle sûr ?

Le paradoxe



Certaines collections ont  t  num ris es pour « sauver » des corpus papier (papier acide, inondations   Prague, etc.)

Avertissements

CAHIERS
GUTenberg

[Table des matières de ce fascicule](#) | [Article précédent](#)

[Présentation](#)

[Comité de rédaction](#)

[Instructions aux auteurs](#)

[Recherche avancée](#)

[Tous les articles en ligne](#)

[Derniers articles](#)

Christian Rossi
De la diffusion à la conservation des documents numériques
(From dissemination to preservation of digital documents)
Cahiers Gutenberg no. 49 (2007), p. 47-61
[Article PDF](#)

Résumé

Cet article a pour thème le problème de la conservation des documents numériques. Il se propose principalement de décrire le problème de la conservation en termes d'intermédiaires matériels et logiciels entre les données et l'utilisateur plutôt qu'en termes de format de fichier. Il est surtout consacré aux documents textuels et aborde différents aspects du problème : matériel et logiciel, migration, formats de fichiers, typographie, XML, conversion, durée.

Abstract

This article addresses the issues of conservation of (mostly textual) electronic documents. Its main objective is to describe the problems in terms of hardware and software mediators between the data and the user rather than focussing on file formats. The point of view is that of a skeptic.

Rechercher un article

[Afficher les formules en TeX](#)

[À propos de MathML](#)

Édité par le Groupe francophone des utilisateurs de TeX
produit par des bénévoles et hébergé par
la Cellule MathDoc avec les outils libres
du CEDRAM

Cahiers GUTenberg (en
ligne) ISSN 2118-8254

Saurons-nous « sauver » les corpus électroniques ?

Remarques de Christian Rossi

- Facilité d'usage \times espérance de vie = constante
- Il ne faut pas trop penser en termes de formats pérennes
- Les formats ne sont pas pérennes, les logiciels non plus et le matériel encore moins
- Plus il y a d'intelligence dans le logiciel exploitant un format, plus le risque de perte est grand
- Les formats « sémantiques » requièrent un environnement pour extraire le sens des données :

`<compositeur>Pierre Henry</compositeur>`

- 18 propositions de balisage différentes pour la musique, combien pour les maths ?

Remarques de Christian Rossi

- Facilité d'usage \times espérance de vie = constante
- Il ne faut pas trop penser en termes de formats pérennes
- Les formats ne sont pas pérennes, les logiciels non plus et le matériel encore moins
- Plus il y a d'intelligence dans le logiciel exploitant un format, plus le risque de perte est grand
- Les formats « sémantiques » requièrent un environnement pour extraire le sens des données :

$$\langle \mathcal{N} \rangle \mathcal{P} \rangle | \nabla \nabla | _ \mathcal{H} | \nabla \dagger \langle / \mathcal{N} \rangle$$

- 18 propositions de balisage différentes pour la musique, combien pour les maths ?

Remarques de Christian Rossi

- Facilité d'usage \times espérance de vie = constante
- Il ne faut pas trop penser en termes de formats pérennes
- Les formats ne sont pas pérennes, les logiciels non plus et le matériel encore moins

- Plus il y a d'intelligence dans le logiciel exploitant un format, plus le risque de perte est grand
- Les formats « sémantiques » requièrent un environnement pour extraire le sens des données :

```
<compositeur>Pierre Henry</compositeur>
```
- 18 propositions de balisage différentes pour la musique, combien pour les maths ?

1 La bibliothèque numérique de mathématiques

2 Pérennité ?

3 **Mathématiques et numérique**

4 \LaTeX et l'édition des maths

5 Et ensuite ?

Le problème des formules

- Il n'existe pas de représentation universelle d'une formule de maths
- Mais un babel de codes et formats adaptés à des usages précis
- Et des passerelles plus ou moins heuristiques qui marchent dans certains cas

⇒ Un exemple

« L'intégrale de zéro à π de la fonction sinus vaut 2 »

Le problème des formules

- Il n'existe pas de représentation universelle d'une formule de maths
- Mais un babel de codes et formats adaptés à des usages précis
- Et des passerelles plus ou moins heuristiques qui marchent dans certains cas

⇒ Un exemple

« L'intégrale de zéro à π de la fonction sinus vaut 2 »

Les formules...

L^AT_EX $\int_0^\pi \sin x \, dx = 2$

PDF $\int_0^\pi \sin x \, dx = 2$

DOC/Web $\int_0^\pi \sin x \, dx = 2$

MathML (pr sentation)

```
<math xmlns:mml="http://www.w3.org/1998/Math/MathML">
  <mml:mrow>
    <mml:msubsup>
      <mml:mo>&int;</mml:mo>
      <mml:mn>0</mml:mn>
      <mml:mn>&pi;</mml:mn>
    </mml:msubsup>
    <mml:mo form="prefix">sin</mml:mo>
    <mml:mi>x</mml:mi>
    <mml:mspace width="0.166667em"/>
    <mml:mi mathvariant="normal">d</mml:mi>
    <mml:mi>x</mml:mi>
    <mml:mo>=</mml:mo>
    <mml:mn>2</mml:mn>
  </mml:mrow>
</math>
```

Les formules...

Maple $\text{int}(\sin(x), x=0..Pi) = 2$

MathML (contenu)

```
<math xmlns="http://www.w3.org/1998/Math/MathML">
  <apply><eq/>
    <apply><int/>
      <bvar>
        <ci>x</ci>
      </bvar>
      <lowlimit>
        <cn>0</cn>
      </lowlimit>
      <uplimit>
        <pi/>
      </uplimit>
      <apply>
        <sin/>
        <ci>x</ci>
      </apply>
    </apply>
    <cn>2</cn>
  </apply>
</math>
```

Les formules...

OpenMath

```

<OMOBJ>
  <OMA>
    <OMS cd = "relation1" name="eq"/>
    <OMA>
      <OMS cd="calculus1" name="defint"/>
      <OMA>
        <OMS cd="interval" name="interval"/>
        <OMI>0</OMI>
        <OMS cd = "nums" name = "pi"/>
      </OMA>
    <OMBIND>
      <OMS cd="fns1" name="lambda"/>
      <OMBVAR><OMV name="x"/></OMBVAR>
      <OMA>
        <OMS cd="transc1" name="sin"/>
        <OMV name="x"/>
      </OMA>
    </OMBIND>
  </OMA>
  <OMI>2</OMI>
</OMA>
</OMOBJ>

```

1 La bibliothèque numérique de mathématiques

2 Pérennité ?

3 Mathématiques et numérique

4 \LaTeX et l'édition des maths

5 Et ensuite ?

Documentation mathématique & électronique

Souhaits pour un environnement pérenne

Prépublications Simple & efficace (pour les auteurs et les lecteurs)

- Gestion des versions
- Ouvert
- Centralisé
- Gratuit

Édition Durable : qui réponde aux besoins actuels sans compromettre la capacité à répondre aux besoins des scientifiques dans l'avenir

Archives \simeq Bibliothèques : Service public indépendant des éditeurs pouvant exercer des choix scientifiques et archiver les originaux

Synthèses Wiki ?

Encyclopédies Wolfram Enc., Wikipedia, PlanetMath, Mathoverflow ?
Wolfram alpha

Documentation mathématique & électronique

Souhaits pour un environnement pérenne

Prépublications Simple & efficace (pour les auteurs et les lecteurs)

Édition Durable : qui réponde aux besoins actuels sans compromettre la capacité à répondre aux besoins des scientifiques dans l'avenir

- Professionnelle & efficace
- Pas de frais pour publier ; coût modeste & proportionné
- Possibilité de sélectionner les titres sans surcoût dissuasif
- Liberté d'annuler un abonnement à tout moment
- Libre accès à terme aux textes intégraux
- Métadonnées et navigation libres
- Formats ouverts
- Archivage intégral de la production chez un tiers sans but lucratif (électronique *et* papier pérenne)

Archives \approx Bibliothèques : Service public indépendant des éditeurs pouvant exercer des choix scientifiques et archiver les originaux

Synthèses Wiki ?

Encyclopédies Wolfram Enc., Wikipedia, PlanetMath, Mathoverflow ?

Documentation mathématique & électronique

Souhaits pour un environnement pérenne

Prépublications Simple & efficace (pour les auteurs et les lecteurs)

Édition Durable : qui réponde aux besoins actuels sans compromettre la capacité à répondre aux besoins des scientifiques dans l'avenir

Archives \simeq Bibliothèques : Service public indépendant des éditeurs pouvant exercer des choix scientifiques et archiver les originaux

- Une archive globale (distribuée) acquérant ce qui paraît au fil de l'eau (par numérisation ou transfert nativement numérique)
- Un registre à jour de toutes les ressources disponibles
- Des outils pour résoudre des références et fournir des identifiants uniques
- Une navigation sans frontières dans la totalité du corpus
- Un accès instantané
- Une disponibilité éternelle

Synthèses Wiki ?

Encyclopédies Wolfram Enc., Wikipedia, PlanetMath, Mathoverflow ?

Wolfram alpha

Documentation mathématique & électronique

Souhaits pour un environnement pérenne

Prépublications Simple & efficace (pour les auteurs et les lecteurs)

Édition Durable : qui réponde aux besoins actuels sans compromettre la capacité à répondre aux besoins des scientifiques dans l'avenir

Archives \simeq Bibliothèques : Service public indépendant des éditeurs pouvant exercer des choix scientifiques et archiver les originaux

Vision partiellement implémentée en France par Mathdoc (NUMDAM, CEDRAM) et portée à l'échelon européen

Synthèses Wiki ?

Encyclopédies Wolfram Enc., Wikipedia, PlanetMath, Mathoverflow ?
Wolfram alpha

Documentation mathématique & électronique

Souhaits pour un environnement pérenne

Prépublications Simple & efficace (pour les auteurs et les lecteurs)

Édition Durable : qui répondre aux besoins actuels sans compromettre la capacité à répondre aux besoins des scientifiques dans l'avenir

Archives \simeq Bibliothèques : Service public indépendant des éditeurs pouvant exercer des choix scientifiques et archiver les originaux

Synthèses Wiki ?

Encyclopédies Wolfram Enc., Wikipedia, PlanetMath, Mathoverflow ?
Wolfram alpha

Documentation math  matique &   lectronique

Souhaits pour un environnement p  renne

Pr  publications Simple & efficace (pour les auteurs et les lecteurs)

  dition Durable : qui r  pondre aux besoins actuels sans compromettre la capacit      r  pondre aux besoins des scientifiques dans l'avenir

Archives \simeq Biblioth  ques : Service public ind  pendant des   diteurs pouvant exercer des choix scientifiques et archiver les originaux

Synth  ses Wiki ?

Encyclop  dies Wolfram Enc., Wikipedia, PlanetMath, Mathoverflow ?
Wolfram alpha

CEDRAM

Un modèle de publication assez réac

- Les articles sont des travaux originaux de recherche publiés sous la responsabilité d'un comité scientifique identifié
- Une fois publié, le contenu mathématique n'est jamais modifié (recours aux *corrigenda* si nécessaire)
- La navigation (y compris bibliographies) et les métadonnées sont libres
- La plateforme est neutre par rapport aux modèles économiques *durables* des revues (abonnement, bouquets, durée du créneau mobile).
- Les revues recouvrent elles-mêmes leurs coûts
- Chaque revue gère les soumissions et la préparation des articles acceptés au format de production (\LaTeX un peu spécial)
- Les articles sont (physiquement) transférés dans l'archive NUMDAM (en vue d'un accès à long terme)
- Les sources \LaTeX sont archivées (en vue d'une préservation à long terme)

LaTeX

- \LaTeX est un logiciel libre de mise en page scientifique
- Les sources sont du texte agrémenté de commandes qui définissent notamment la structure
- \LaTeX peut être comparé à XML sur le plan de la séparation de la mise en forme et du contenu (on balise une structure)
- Le langage de macro-commandes est très puissant ce qui donne des documents vivants, voire dynamiques
- Des milliers de contributions étendent les possibilités
- On peut même faire des partitions (MusiX \TeX) !
- La syntaxe d'un document donné est donc imprévisible : il faudrait sauver tout l'environnement de production pour interpréter un fichier source dans un futur lointain

Exemple LaTeX

```

\documentclass[10pt, a4paper]{article}
\usepackage[utf8]{inputenc}
\usepackage[T1]{fontenc}
\usepackage[french]{babel}
\begin{document}
\section{Une section}
\subsection{Une sous-section}
Il  tait une fois dans une lointaine
galaxie...
\subsection{Une deuxi me sous-section}
La guerre des clones faisait alors
rage...
\end{document}

```

1

1 Une section

1.1 Une sous-section

Il  tait une fois dans une lointaine galaxie...

1.2 Une deuxi me sous-section

La guerre des clones faisait alors rage...

Exemple math LaTeX

Simple align is used to align two or more formulas. To obtain the formulas

$$(2) \quad r^2 = s^2 + t^2$$

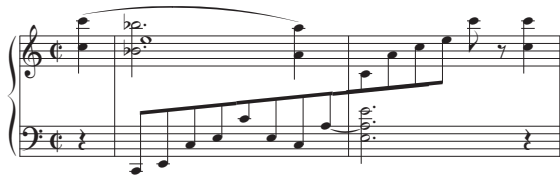
$$(3) \quad 2u + 1 = v + w^\alpha$$

$$(4) \quad x = \frac{y + z}{\sqrt{s + 2u}}$$

type (using `\\` as a line separator)

```
\begin{align}
  r^{2}   &=& s^{2} + t^{2} & \quad \quad \quad \backslashlabel{E:eqn1}\\
  2u + 1 &=& v + w^{\alpha} & \quad \quad \quad \backslashlabel{E:eqn2}\\
  x       &=& \frac{y + z}{\sqrt{s + 2u}} & \quad \quad \quad \backslashlabel{E:eqn3}
\end{align}
```

Exemple MusiXTeX



The coding is

```

\interstaff{13}
\instrumentnumber{1}
\setstaves1{2}
\setclef1\bass
\generalmeter\allabreve
\startextract
\Notes\qp\nextstaff\isluru0q\zq{q}\ql{j}\enotes
\bar
\nspace
\Notes\ibu0a1\qb0{CEJLcL}%
  \nextstaff\roff{\zw{1}}\pt{p}\zh{p}\pt{i}\hl{i}\enotes
\Notes\qb0J\itiedia\qb0a\nextstaff\tslur0o\zq{o}\ql{h}\enotes
\bar
\Notes\ttie1\zh{.L.a}\hl{.e}%
  \nextstaff\qb0{chj}\tbl0\qb01\c1{q}\ds\enotes
\Notes\qp\nextstaff\zq{q}\ql{j}\enotes
\endextract

```

Systeme de production *LaTeX* du CEDRAM

Le CEDRAM produit un fascicule de revue en une passe \LaTeX qui

- d termine les m tadonn es   partir des sources du volume et de chaque article, calcule les folios   la vol e, et produit les PDF pour l'imprimeur
- compile s par ment chaque article et produit le PDF  cran
- convertit toutes les m tadonn es \LaTeX en XML/MathML (tout en conservant l' criture \TeX des formules en parall le)
- Les sources \LaTeX assurent d'une certaine fa on la reproductibilit  du sens
- Ne sont exploitables qu'avec l'installation \LaTeX compl te correspondante
1240 fichiers sp cifiques, 77 846 fichiers « standard », 2,5 Go
- Les m tadonn es XML et les PDF des images fig es le jour de production
- Dans NUMDAM, on ajoute une version DjVu (bitmap),
plus « b te » donc plus durable ?

- 1 La bibliothèque numérique de mathématiques
- 2 Pérennité ?
- 3 Mathématiques et numérique
- 4 \LaTeX et l'édition des maths
- 5 Et ensuite ?

Un exemple de revue

Plus complexe   mettre en  uvre et archiver !



IPOL · Image Processing On Line

[HOME](#) · [ABOUT](#) · [ARTICLES](#) · [PREPRINTS](#) · [NEWS](#) · [SEARCH](#)

IPOL is a research journal of image processing and image analysis. Each article contains a text describing an algorithm and source code, with an online demonstration facility and an archive of online experiments. The text and source code are peer-reviewed and the demonstration is controlled. IPOL is an Open Science and Reproducible Research journal.

[✎ Editorial Policy](#) · [👤 Editorial Board](#) · [📄 Submit an Article](#)[📖 Index](#) · [Articles 2012](#) [2011](#) [2010](#) · [Preprints](#)

Latest Articles

➔ follow IPOL [✉](#) [🐦](#) [📷](#)

- **Automatic Color Enhancement (ACE) and its Fast Implementation**
2012-11-06 · Pascal Getreuer
- **Color and Contrast Enhancement by Controlled Piecewise Affine Histogram Equalization**
2012-10-17 · Jose-Luis Lisani, Ana-Belen Petro, Catalina Sbert
- **The Flutter Shutter Camera Simulator**
2012-10-17 · Yohann Tendo
- **Chan-Vese Segmentation**
2012-08-08 · Pascal Getreuer
- **An Analysis and Implementation of the BM3D Image Denoising Method**
2012-08-08 · Marc Lebrun
- **Total Variation Deconvolution using Split Bregman**
2012-07-30 · Pascal Getreuer

Polymath

Des projets de math massivement collaboratifs

The polymath blog



June 4, 2013

Polymath proposal: bounded gaps between primes

Filed under: [planning, polymath proposals](#) — Terence Tao @ 4:31 am

Two weeks ago, Yitang Zhang announced [his result](#) establishing that bounded gaps between primes occur infinitely often, with the explicit upper bound of 70,000,000 given for this gap. Since then there has been a *flurry of activity* in reducing this bound, with the current record being 4,802,222 (but likely to improve at least by a little bit in the near future).

It seems that this naturally suggests a Polymath project with two interrelated goals:

1. Further improving the numerical upper bound on gaps between primes; and
2. Understanding and clarifying Zhang's argument (and other related literature, e.g. the work of Bombieri, Fouvry, Friedlander, and Iwaniec on variants of the Elliott-Halberstam conjecture).

Part 1 of this project splits off into somewhat independent sub-projects:

1. Finding narrow **prime** admissible tuples of a given cardinality (or, dually, finding large **prime** admissible tuples in a given interval).
This part of the project would be relatively elementary in nature, relying on combinatorics, elementary number theory, computer search, and perhaps some clever algorithm design. (Scott Morrison has already been hosting a de facto project of this form at [this page](#), and is happy to continue doing so).
2. Solving a calculus of variations problem associated with the Goldston-Yildirim-Pintz argument (discussed at [this blog post](#), or in this older [survey of Soundararajan](#)) [in particular, this could lead to an improvement of a certain key parameter k_0 , currently at 341,640, even without any improvement in the parameter ϖ mentioned in part 3. below.]
3. Delving through the "hard" part of Zhang's paper in order to improve the value of a certain key parameter ϖ (which Zhang sets at 1/1168, but is likely to be enlargeable).

Recent Comments

Polymath8:
Bounded G... [on](#)
Polymath
proposal:
bounded gap...
List of



Misconception...
[on](#) Polymath
proposal:
bounded gap...
Bounded on prime
gaps... [on](#)
Polymath
proposal:
bounded gap...
John [on](#)



Polymath
proposal:
bounded gap...
Terence Tao [on](#)



Polymath
proposal:

La conjecture de Kepler

Une démonstration écrite en C++



The Kepler Conjecture

Background

[Announcement of a solution](#)
[Site summary](#)
[What is it?](#)
[History](#)
[Holyoke lecture](#)
[References](#)
[Ferguson's thesis](#)
[Dodecahedral Conj.](#)
[Serious Stuff](#)
[Notable Quotes](#)

Here it is at last!

Background

- This page is a summary of the site.
- The announcement of a [solution](#).
- A brief [statement](#) of the problem.
- The [history](#) of the problem.
- A 1996 [lecture](#) at Mt. Holyoke explains the key ideas that will be involved in the proof.
- [Academic papers](#), [popular articles](#), and general reference books on discrete geometry.
- Ferguson's Ph.D. [thesis](#) at the University of Michigan completes one of the five steps of the proof.
- Sean McLaughlin has proved Fejes Tóth's [dodecahedral conjecture](#)
- Follow this link to the [serious stuff](#). This link is intended for discrete geometers who wish to check the technical details of the solution.
- What have Kepler, Hilbert, Milnor, Coxeter, Fejes Tóth and [others said](#) about the Kepler Conjecture?

Software

[Interval Code](#)
[Graph Archive](#)
[Mathematica](#)

Software Tools

- [Interval Package](#). A C++ library especially designed to prove inequalities arising in the Kepler conjecture.
- [Graph Generator](#). An applet that generates all the combinatorial possibilities of the Kepler conjecture.
- [Mathematica](#) has been used extensively throughout this project for experimentation and exploration. Some of the most useful functions have been posted here.



FlyspeckFactSheet

Introduction to the Flyspeck Project (formal proof of the Kepler Conjecture). This is a good place to start.

Featured

Updated Feb 14, 2012 by TCHa...@gmail.com

The Flyspeck Project Fact Sheet

[Flyspeck Project Home](#)

Introduction

The purpose of the flyspeck project is to produce a formal proof of the Kepler Conjecture. This page gives some basic facts about this project.

What does the name mean?

The name 'flyspeck' comes from matching the pattern `/f.*p.*k/` against an English dictionary. FPK in turn is an acronym for "The Formal Proof of Kepler." The term 'flyspeck' can mean to *examine closely* or *in minute detail*; or to *scrutinize*. The term is thus quite appropriate for a project intended to scrutinize the minute details of a mathematical proof.

What is a formal proof?

A formal proof is understood in the sense of the [QED manifesto](#)

To learn more about formal proofs, see [Freek Wiedijk's web pages](#). Also see the complete [works of John Harrison](#).

How does a formal proof differ from a traditional mathematical proof?

Traditional mathematical proofs are written in a way to make them easily understood by mathematicians. Routine logical steps are omitted. An enormous amount of context is assumed on the part of the reader. Proofs, especially in topology and geometry, rely on intuitive arguments in situations where a trained mathematician would be capable of translating those intuitive arguments into a more rigorous argument.

In a formal proof, all the intermediate logical steps are supplied. No appeal is made to intuition, even if the translation from intuition to logic is routine. Thus, a formal proof is less intuitive, and yet less susceptible to logical errors.

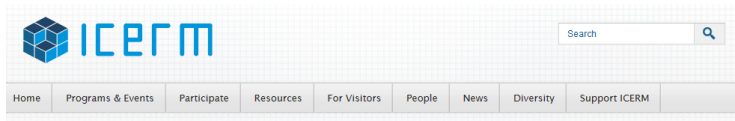
What are the specific tools that will be used in the flyspeck project?

All the formal proofs will be made by computer and programmed in the Objective CAML programming language.

Le problème de la reproductibilité

- Le calcul sur machine joue un rôle croissant dans des résultats théoriques ou expérimentaux
- Contrairement aux sciences expérimentales, il n'y a pas pour l'instant de culture du « cahier de laboratoire »
- Ni l'environnement permettant de reproduire, vérifier, réutiliser des résultats obtenues à l'aide de logiciels
- Imprédictibilité de certaines erreurs d'arrondi qui dépendent du matériel et de chaque composant logiciel
- Quid des calculs répartis « dans le nuage » . . .

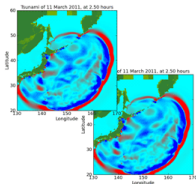
Conf rence sur la reproductibilit 



Reproducibility in Computational and Experimental Mathematics (*December 10-14, 2012*)

Description

In addition to advancing research and discovery in pure and applied mathematics, computation is pervasive across the sciences and now computational research results are more crucial than ever for public policy, risk management, and national security. Reproducibility of carefully documented experiments is a cornerstone of the scientific method, and yet is often lacking in computational mathematics, science, and engineering. Setting and achieving appropriate standards for reproducibility in computation poses a number of interesting technological and social challenges. The purpose of this workshop is to discuss aspects of reproducibility most relevant to the mathematical sciences among researchers from pure and applied mathematics from academics and other settings, together with interested parties from funding agencies, national laboratories, professional societies, and publishers. This will be a working workshop, with relatively few talks and dedicated time for breakout group discussions on the current state of the art and the tools, policies, and infrastructure that are needed to improve the situation. The groups will be charged with developing guides to current best practices and/or white papers on desirable advances.



[Click for code to create this image.](#)

Organizing Committee

- » [David H. Bailey](#)
(Lawrence Berkeley National Laboratory)
- » [Jon Borwein](#)
(Centre for Computer Assisted Research Mathematics and its Applications)
- » [Randall J. LeVeque](#)
(University of Washington)
- » [Bill Rider](#)
(Sandia National Laboratory)
- » [William Stein](#)
(University of Washington)
- » [Victoria Stodden](#)
(Columbia University)

Recommandations

Bonnes pratiques de publication

Pour faciliter la reproductibilit , les  l ments suivants devraient  tre fournis :

- Une description pr cise des  nonc s contenus dans l'article
- Une description de l'approche computationnelle, et de pourquoi elle constitue un test rigoureux de l' nonc  vis 
- Une description exhaustive de chaque algorithme employ 
- Une description d taill e des logiciels utilis s pour les calculs
- Une description d taill e de l'environnement de test (mat riel, syst me d'exploitation, nombre de processeurs)
- Une description d taill e des m thodes d'analyse statistique ou de r duction de donn es
- Une discussion de l'ad quation de param tres tels que la pr cision num rique
- Une description exhaustive des r sultats exp rimentaux
- Une validation des tests conduits par l'auteur
- La disponibilit  du code informatique, des donn es d'entr e et de sortie, le tout raisonnablement document 
- La description des strat gies et moyens pour la « curation » du code et des donn es
- Des instructions permettant de reproduire les exp riences de calcul d crites
- Les clauses de propri t  intellectuelle (id alement licence ouverte permettant la r utilisation)
- Une description des chemins explor s lors de la recherche, y compris les r sultats n gatifs
- La citation selon les normes acad miques de tout le code et toutes les donn es utilis es, y compris produits par l'auteur

Recommandations

Bons outils

- *Literate programming* (int grer code et documentation : CWEB. . .)
- Outils interactifs qui permettent d'int grer texte et calculs (Maple, TeXmacs. . .)
- Outils qui enregistrent les origines et la chronologie des  l ments participant au calcul (Sumatra, Madagascar. . .)
- Outils de contr le de versions et d veloppement collaboratif (CVS, SVN, GIT, Mercurial. . .)
- Outils permettant de sauvegarder un environnement logiciel (machines virtuelles, RunMyCode.org. . .)
- Assistants de preuve pour v rifier la logique ou certifier les calculs