Textbooks for the Web from MathBook XML

Joint Mathematics Meetings WEBSIGMAA Guest Lecture

Rob Beezer

January 6, 2017
Atlanta, Georgia

Department of Mathematics and Computer Science
University of Puget Sound
beezer@pugetsound.edu

Open Source Publishing

Mathematics on the Web

- The Internet is an amazing thing
- Online Encyclopedia of Integer Sequences
- Also: kitten videos and fake news
- Your own:
  - printing press
  - shipping and distribution network
  - marketing (i.e. Google Page Rank)

- If you want to be read,
  you want your writing on the Internet
- MathJax is the enabling technology
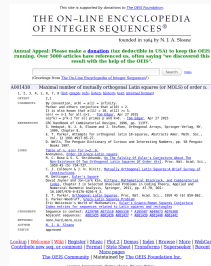
- Copyright is a government-granted monopoly
- An open license provides additional freedoms
    - Unlimited copying
    - In perpetuity
    - Modifications for personal use
    - Possibly: the right to distribute modifications
    - "Copyleft"

> "freely available"
> $\neq$
> distribute with an explicit open license

Web Versions of Open Textbooks

- Portable: 64 GB is
  - 64 Encyclopedia Britannica (text)
  - 1 English Wikipedia (text)
  - 10,000 400-page math textbooks
- Ubiquitous: laptop, tablet, or phone
- Up-to-Date: correct, and refresh, at will
- Accurate: crowd-sourced proof-reading
- Open: never out-of-print
- Intellectually Honest:
  no pressure to satisfy market segments
- FREE!!!!!

THIS WAS GOOGLE
WHEN I WAS A KID

Proprietary Electronic Textbooks

- Expire after course
- Limited printing
- Inconvenient interfaces
- Difficult to resell
- Page-oriented (ala PDF)

MathBook XML Design

Progenitor: Open Source Linear Algebra

- 2004: A First Course in Linear Algebra
- Highly structured LaTeX, PDF output
- Very semantic LaTeX math macros
- 2006: HTML + MathML via `tex4ht`
- 2008: HTML + jsMath via `tex4ht`
- 2012: Conversion to experimental XML
- 2013: Initiated MathBook XML
- Today: 500,000 annual unique visits
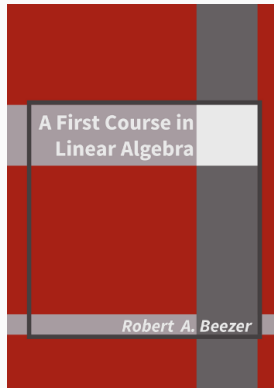
A First Course in
Linear Algebra

Robert A. Beezer

MathBook XML

- Philosophy: **rigorously** separate
  - structure and content
  - presentation

MathBook XML

- Philosophy: **rigorously** separate
    - structure and content
    - presentation
- Realization:
    - structure: an XML vocabulary
    - not Markdown, ASCIIDoc, JSON
      YAML, MediaWiki, Pandoc, ...
    - math content: still LaTeX (AMS Math)

MathBook XML

- Philosophy: **rigorously** separate
  - structure and content
  - presentation
- Realization:
  - structure: an XML vocabulary
  - not Markdown, ASCIIDoc, JSON
    YAML, MediaWiki, Pandoc, ...
  - math content: still LaTeX (AMS Math)
- Payoff:
  - multiple outputs from a single source
  - powerful and flexible processing with
    eXtensible Stylesheet Language (XSL)
  - author with your favorite text editor
  - cross-platform open-source toolchain

New Author-Friendly XML Vocabulary

- book, chapter, section, subsection
- theorem: title, statement, proof
- p, ul, ol, dl, li, q, em
- m, me, md/mrow

- Extensive cross-reference support
- Excellent index creation (more coming)
- Multiple bibliographies (needs work)
- Consistent element use
  - title
  - introduction
  - xref, xml:id
- Only three dangerous characters: &, <, >

MathBook XML Example

```xml
<theorem xml:id="power-rule">
    <title>Power Rule</title>
    <index>power rule</index>

    <statement>
        <p>The derivative of <m>f(x)=x^n</m>
        is <m>f'(x)=nx^{n-1}</m>.</p>
    </statement>

    <proof>
        <p>Apply induction to the product
        <me>f(x)=x^n=x\cdot x^{n-1}</me>
        using <xref ref="product-rule"/>.</p>
    </proof>
</theorem>
```

Theorem 4.4 (Power Rule). *The derivative of $f(x) = x^n$ is $f'(x) = nx^{n-1}$.*

Proof. Apply induction to the product

$$f(x) = x^n = x \cdot x^{n-1}$$

using Theorem 4.1. □

Theorem 4.4 (Power Rule)**.** *The derivative of $f(x) = x^n$ is $f'(x) = nx^{n-1}$.*
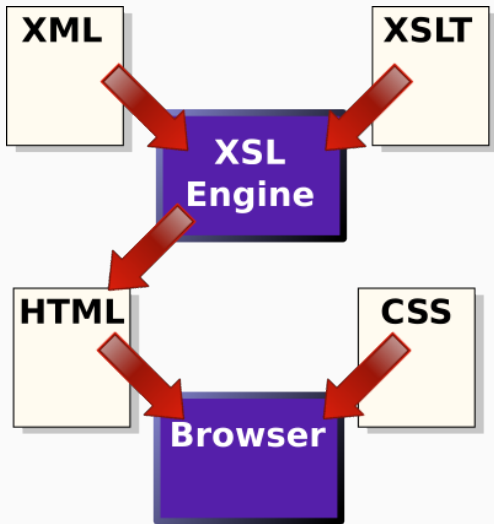
Proof. Apply induction to the product

$$f(x) = x^n = x \cdot x^{n-1}$$

using Theorem 4.1. $\square$

Corollary 4.5. *Suppose $f(x)$ is a continuous function. Then*

- Well-formed, clean, semantic LaTeX
- Extensive, thoroughly tested, preamble
- Supports `pdflatex`, `xelatex`
- Numbering by usual LaTeX mechanisms
- Extensive support for verbatim text
- Standard organization of front matter
- Emerging support for scripts
  (Latin, Greek, Cyrillic, Vietnamese)
- The "Data Liberation" guarantee
- (Sat AM OER talks: M. Boelkins, M. Keller)

Output: HTML for the Web

- Preserve the best parts of book design
- Use browser technology to improve other parts
- Preserve hierarchical organization
- Improve navigation, cross-referencing
- Consistent numbering with $\LaTeX$ version
- Information-hiding with knowls
- Information-duplication with knowls
- Improved index
- Semantic CSS for themes
- Special attention to accessibility

Available to every author, minimal technical expertise required

- Javascript
- CSS (American Inst. of Math)
- MathJax (AMS, others)
- Knowls (American Inst. of Math)
- Sage cells (SageMath Inc., NSF)
- WeBWorK exercises
- YouTube videos (Google)
- Online fonts (Google)
- Search (Google)
- Analytics (Google, StatCounter)
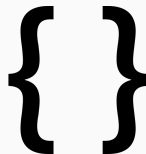
- ePUB
  - Portable, offline, reflowable PDF
  - Math via MathJax SVG
  - Beta quality now
- SageMathCloud Worksheets
  Jupyter Notebooks
  - Promising, limited functionality now
- Sage Doctesting
  - Test embedded Sage code
  - Works exceedingly well
- Office Open XML
  - Microsoft Office, LibreOffice
  - ISO 29500 standard, 1000's of pages
  - Feasible?

ePUB

Jupyter

Microsoft
Office

Demonstrations

Next Steps

- Course handouts: <article> element
- Lecture notes: <book> element
- Open textbook: for **your** course (no deadlines!)
- Get involved, join our community
    - 84 members in Google Group, `mathbook-xml-support`
    - 14 contributors on GitHub (35 forks), `rbeezer/mathbook`
    - ~30 authors, see "Gallery" at project site

`mathbook.pugetsound.edu`

Contributors

- Chris Godsil: Original motivation
- David Farmer: Design philosophy, HTML output
- Alex Jordan: WeBWorK, major code contributions
- Tom Judson: Lead Tester and Premier Exemplar
- William Stein: Sage, Sage Cell, SageMathCloud

- National Science Foundation
- University of Puget Sound
- American Institute of Mathematics
- OpenOregon
- Shuttleworth Foundation