PDF Days Europe 2025

Overcoming the sentiment that PDFs are evil

(at least for accessible math)

Tagging project page
https://latex3.github.io/tagging-project

General Challenges

- **Most (STEM[1]) documents are untagged**
  Without tagging, accessibility software can only guess at the content: usually this results in poor outcomes

  [1]STEM = science, technology, engineering, mathematics
- **Remediation of documents is slow**
  This adds a barrier to any change to a source once it's been tagged, and means that most documents never even get that far
- **Remediation of STEM documents is unreliable**
  Often it is not done by the original author and not by somebody with a mathematical or technical background, so misunderstandings are likely
- **Overall perception in the accessibility community**
  PDFs are evil at least for accessibility

Control over accessible math output

- The notation $|x|$ has several meanings, such as "absolute-value of x", or the "norm of x".
  It is hard for remediators unfamiliar with the content to correctly tag such an expression.
- LaTeX may automatically generate suitable MathML *intents*: translatable speech hints for the screen reader. This does not affect the braille, which is based on the notation not the spoken form.
- `\abs{x}` : $|x|$ is read as "absolute value of x"

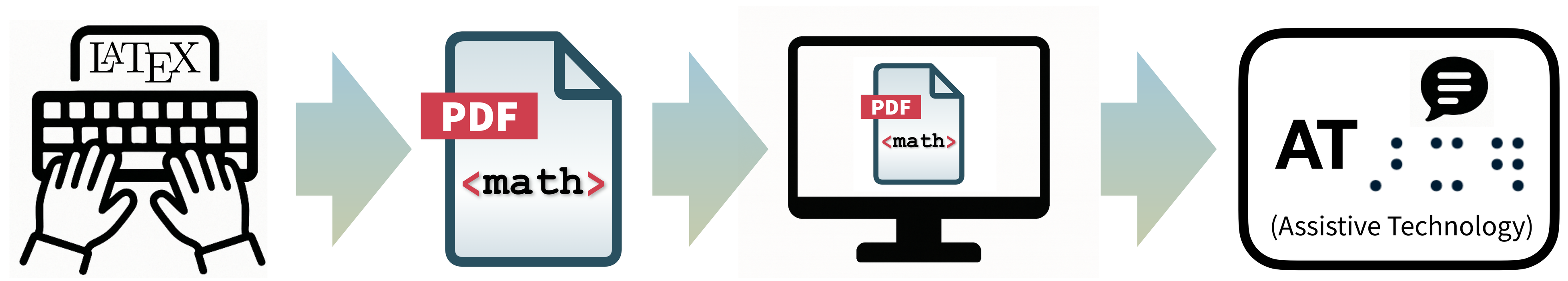
```
<mrow intent="absolute-value($x)">
  <mo>|</mo><mi arg="x">x</mi><mo>|</mo></mrow>
```

- `\norm{x}` : $|x|$ is read as "norm of x"

```
<mrow intent="norm($x)">
  <mo>|</mo><mi arg="x">x</mi><mo>|</mo></mrow>
```

An end-to-end workflow for truly accessible PDFs

The workflow:



1. 2. 3. 4.

Ingredients:

1. A suitable generator (e.g., LaTeX)
2. A modern PDF structure supporting the required features (PDF 2.0)
3. A PDF reader capable of processing MathML, (i.e., Foxit with AF and SE support, Acrobat with SE support, others ???)
4. Assistive technology that supports MathML in PDF context (NVDA + MathCAT)

Input:

```
% !TeX program = LuaLaTeX
\DocumentMetadata{tagging = on, lang = en}
\documentclass{article}
\usepackage{unicode-math}
\begin{document}
\[ \sum_{i=1}^{n} i = \frac{(n+1)n}{2} \]
\end{document}
```

Visual output:

$$\sum_{i=1}^{n} i = \frac{(n+1)n}{2}$$

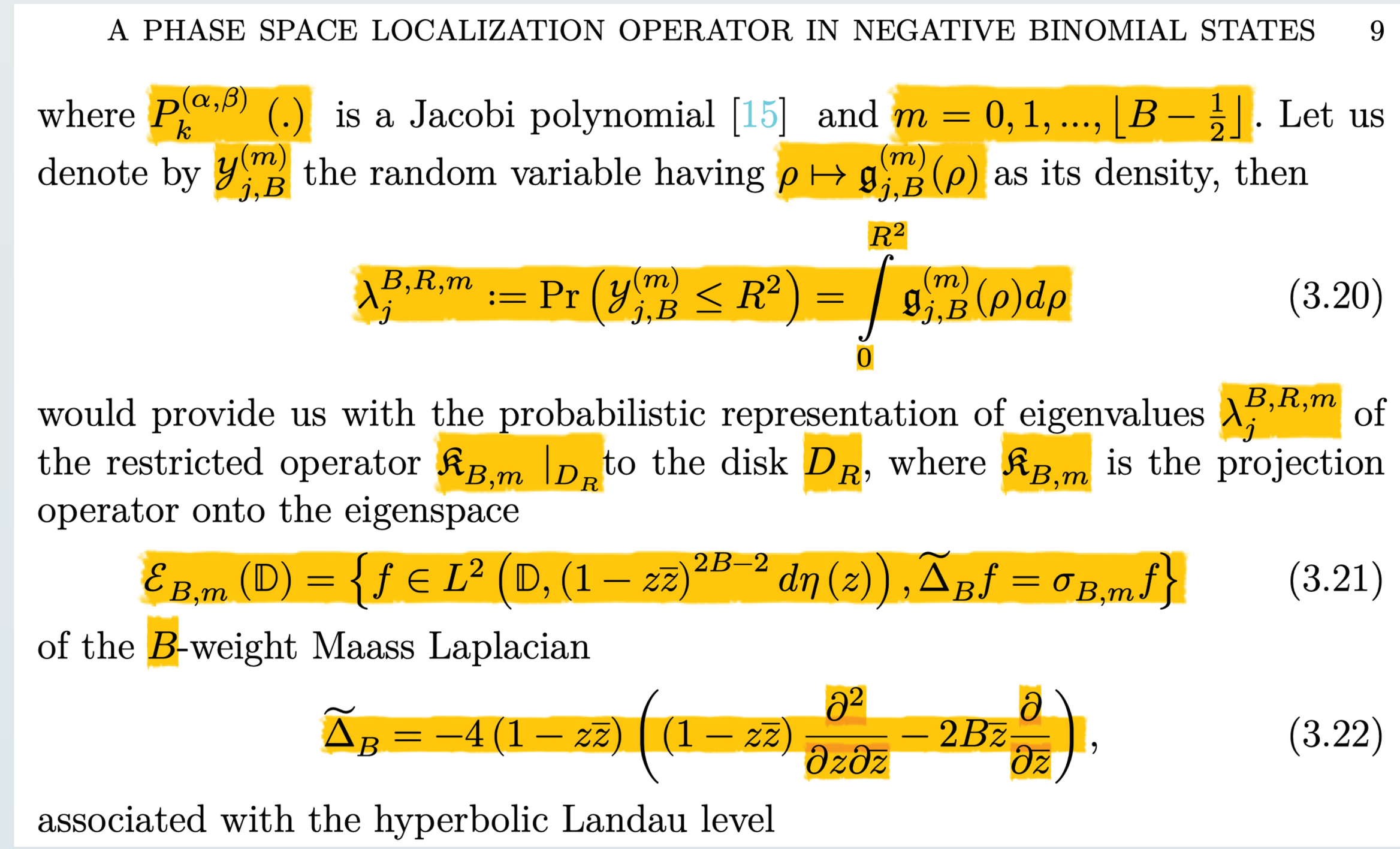
Braille: ⠐⠨⠠⠎⠩⠰⠊ ⠨⠅ ⠼⠂⠣⠰⠝⠻⠰⠊ ⠨⠅ ⠹⠷⠰⠝⠬⠂⠾⠰⠝⠌⠆⠼

Speech:

the sum from i is equal to 1 to n of i; is equal to; the fraction with numerator; open paren n plus 1, close paren; times n; and denominator 2; end fraction

PDF/UA-1 is unsuitable

- **PDF/UA-1 requires Alternative Text for math**
  This is unsuitable for math:
  – no Nemeth Braille output
  – reading loses punctuation and symbols
  – no navigation and highlighting of subterms
  – no reuse, e.g., export to HTML
  – speech styles and verbosity are not adjustable
- **STEM documents are math-heavy**
  Often several dozen math fragments per page



A PHASE SPACE LOCALIZATION OPERATOR IN NEGATIVE BINOMIAL STATES 9

where $P_k^{(\alpha,\beta)}(.)$ is a Jacobi polynomial [15] and $m = 0, 1, ..., \lfloor B - \frac{1}{2} \rfloor$. Let us denote by $y_{j,B}^{(m)}$ the random variable having $\rho \mapsto \mathfrak{g}_{j,B}^{(m)}(\rho)$ as its density, then

$$\lambda_j^{B,R,m} := \Pr\left(y_{j,B}^{(m)} \leq R^2\right) = \int_0^{R^2} \mathfrak{g}_{j,B}^{(m)}(\rho) d\rho \qquad (3.20)$$

would provide us with the probabilistic representation of eigenvalues $\lambda_j^{B,R,m}$ of the restricted operator $\mathfrak{K}_{B,m}|_{D_R}$ to the disk D_R, where $\mathfrak{K}_{B,m}$ is the projection operator onto the eigenspace

$$\mathcal{E}_{B,m}(\mathbb{D}) = \left\{ f \in L^2\left(\mathbb{D}, (1 - z\bar{z})^{2B-2} d\eta(z)\right), \widetilde{\Delta}_B f = \sigma_{B,m} f \right\} \qquad (3.21)$$

of the B-weight Maass Laplacian

$$\widetilde{\Delta}_B = -4(1 - z\bar{z})\left((1 - z\bar{z})\frac{\partial^2}{\partial z \partial \bar{z}} - 2B\bar{z}\frac{\partial}{\partial \bar{z}}\right), \qquad (3.22)$$

associated with the hyperbolic Landau level

- **Heavy manual work is required**
  Each of them needs to have Alternative Text provided by the author: as this is not simply the original input, mistakes and oversights are likely
- **Questionable seal of approval**
  PDF/UA-1 leads to STEM documents with a seal of approval — but in reality these documents are by no means accessible when passed to AT tools
- **Overall perception in the accessibility community**
  Alternative Texts on math are not useful