Boris Doubrov, Dual Lab

electronic document CONFERENCE

Accessible Math in PDF

Students, Scientists, Publishers, Standards, Tools

Students

What do they need?

Personal story

- Alexander Kozlovsky, Minsk, BY
- Bachelor of Science, Mathematics
- *“I was willing to continue with Master in Mathematics or Computer Science, but unfortunately, all advanced books were not accessible to me”*
- Working now at Dual Lab on the development of a new LaTeX package for accessibility

Copyright © 2018, PDF Association

Scientists

Writing articles

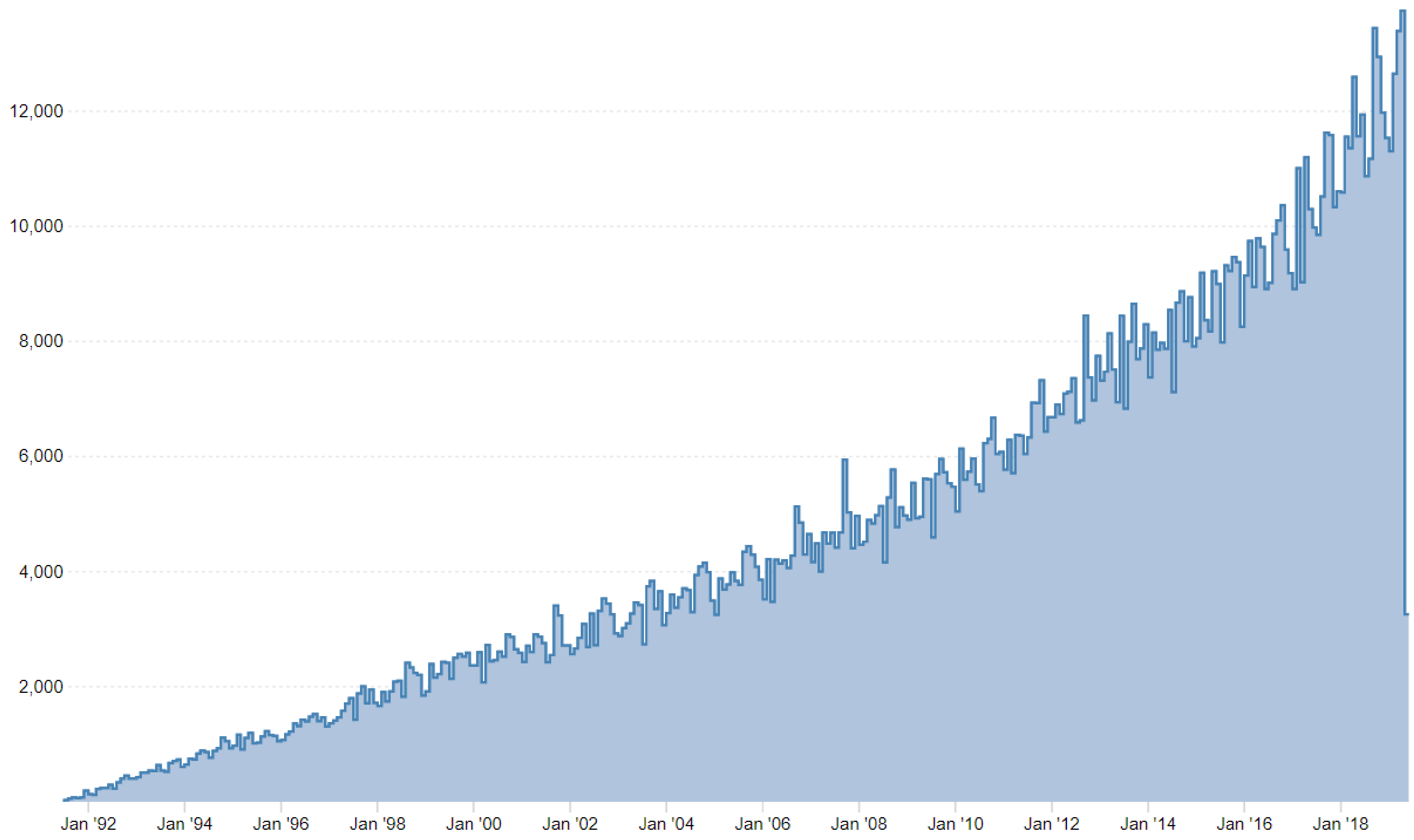
Generating scientific content

- arXiv.org: open access to 1,549,559 e-prints in the fields of physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering and systems science, and economics

- Primary e-print format:
  - PDF
- Formats for text submission:
  - (La)TeX, AMS(La)TeX, PDFLaTeX
  - PDF
  - HTML with JPEG/PNG/GIF images

Copyright © 2018, PDF Association

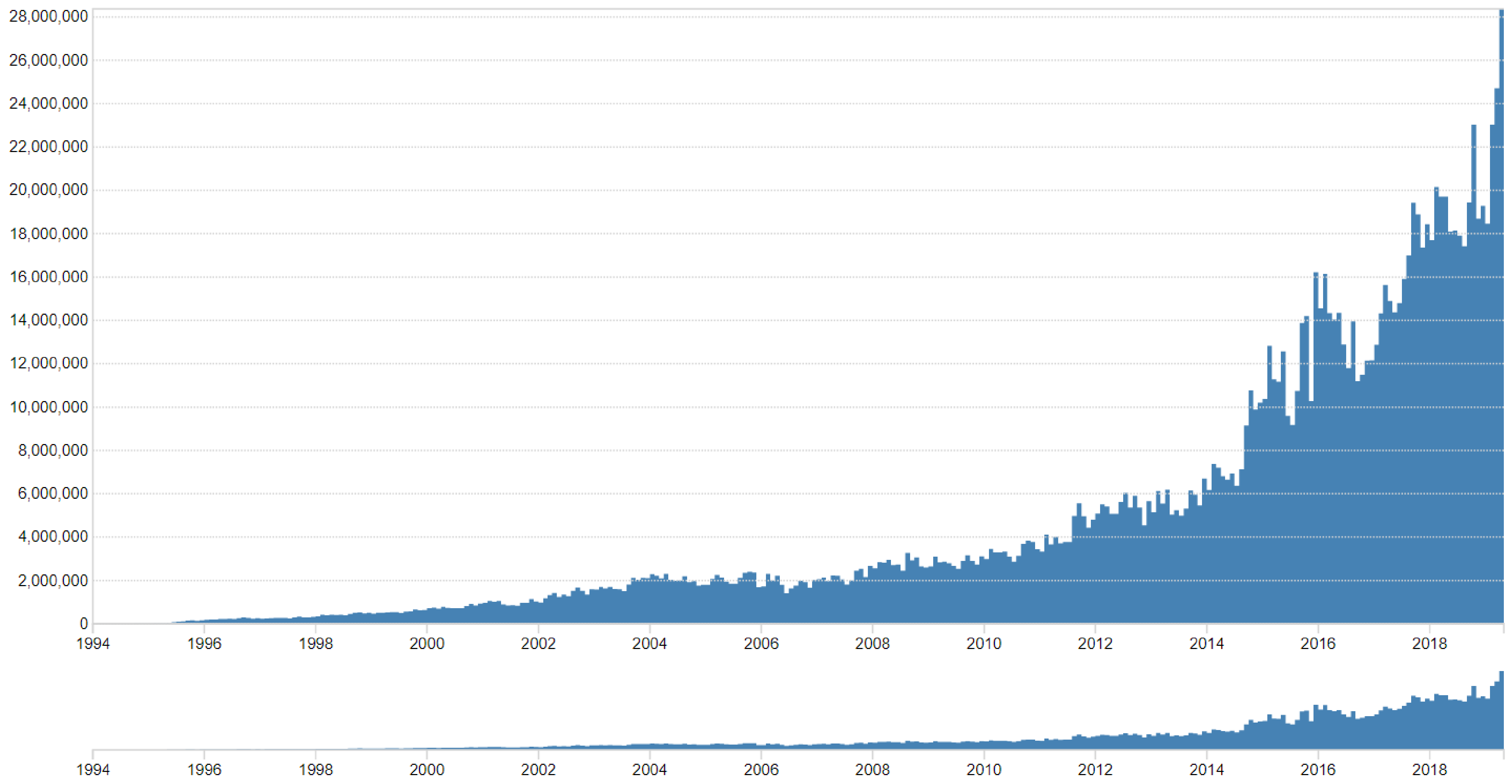
arXiv Monthly Submission Rates [CSV]



Blue: Number of new submissions received during each month since August 1991.

Copyright © 2018, PDF Association

arXiv Monthly Download Rates [CSV]



Blue: Number of downloads per month.

Copyright © 2018, PDF Association

arXiv accessibility check

arXiv:1906.03246v1 [math.RT] 7 Jun 2019

Notes on Jordan-Hölder property for exact categories

Souheila Hassoun and Sunny Roy

Abstract

We prove a generalised version of Jordan-Hölder theorem for pre-abelian exact categories, which allow us to study the relative length function.

1 INTRODUCTION

In 1869 Camille Jordan announced, and demonstrated in 1870, that for two composition sequences of the same finite group, the order sequence (the number of elements) of quotients is the same, up to permutation. In 1889 Otto Hölder reinforced this result by proving the theorem known as the Jordan-Hölder-Schreier theorem, which states that any two composition series of a given group are equivalent, that is, they have the same composition length and the same composition factors, up to permutation and isomorphism. It is proved using the Schreier refinement theorem (which uses the Zassenhaus lemma), see [Rot95] for the standard proof. In [Ba06], Baumslag gives a short proof of the Jordan–Hölder theorem, for *groups*, by intersecting the terms in one subnormal series with those in the other series.

In this article we follow Baumslag's method and we prove a generalised version of Jordan-Hölder theorem for *pre-abelian exact* categories, which are a generalisation of the abelian categories, namely pre-abelian additive categories with a choice of a Quillen exact structure [Qu73] (a class of short exact sequences, called admissible pairs of morphisms, satisfying Quillen's axioms.

To do so we generalise the definition of simple modules to relative simple objects corresponding to a fixed exact structure on an additive category, and using this we generalise the *Schur lemma* to pre-abelian exact categories. Then we study the notion of intersection and sum for pre-abelian categories and prove some of their properties on exact categories. We also study *the isomorphism theorems* in our more general context.

Summary Report

This PDF file is not PDF/UA compliant.

Checkpoint	Passed	Warned	Failed
PDF Syntax	28	0	0
Fonts	44	0	2
Content	2	0	30734
Embedded Files	0	0	0
Natural Language	0	0	15138
Structure Elements	180	0	180
Structure Tree	0	0	0
Role Mapping	0	0	0
Alternate Descriptions	0	0	90
Metadata	0	0	6
Document Settings	2	0	30

PAC3 accessibility report

Copyright © 2018, PDF Association

Publishers

Do they care?

Academic publishing world

- Fifty-seven major publishing groups bring in a combined revenue of **€59.328 billion**
- **Elsevier**, **Springer**, **Wiley-Blackwell**, **Taylor & Francis** and **SAGE** published more than half of 2013's peer-reviewed academic papers
- "These five corporations control academic publishing", report by Joshua A. Krisch, 2015
- Springer: more than 2,900 journals and 300,000 books
- Elsevier: over 2,960 journals, 48,300 books

Copyright © 2018, PDF Association

Classical book: looks better than it is

TABLE OF NOTATIONS

$A, B, \ldots, G, H, \ldots$	groups (or their subsets)
$a, b, c, g, h, u, v, w, x, y, z$	elements of groups
$i, j, k, l, m, n, p, q, r, s, t$	rational integers (p a prime)
$\mathfrak{J}$	ring of integers
$\mathfrak{P}$	ring of p-adic integers
ω	first infinite ordinal
$\aleph_\alpha$ ($\alpha = 0, 1, \ldots$)	αth infinite cardinal ($\aleph =$ continuum)
$[a, \ldots]$	set of elements $a, \ldots$
$\{a, \ldots\}$	subgroup generated by $a, \ldots$
$\{a, \ldots\}_*$	pure subgroup generated by $a, \ldots$
nG	set of all ng with $g \in G$
$G[n]$	set of all $g \in G$ with $ng = 0$
$\mathfrak{R}^+$	additive group of the ring $\mathfrak{R}$
$\mathfrak{F}^\times$	multiplicative group of the field $\mathfrak{F}$
$O(a)$	order of a
$E(a)$	exponent of a
$H(a)$ ($H_p(a)$)	height of a (at the prime p)
$T(a)$	type of a

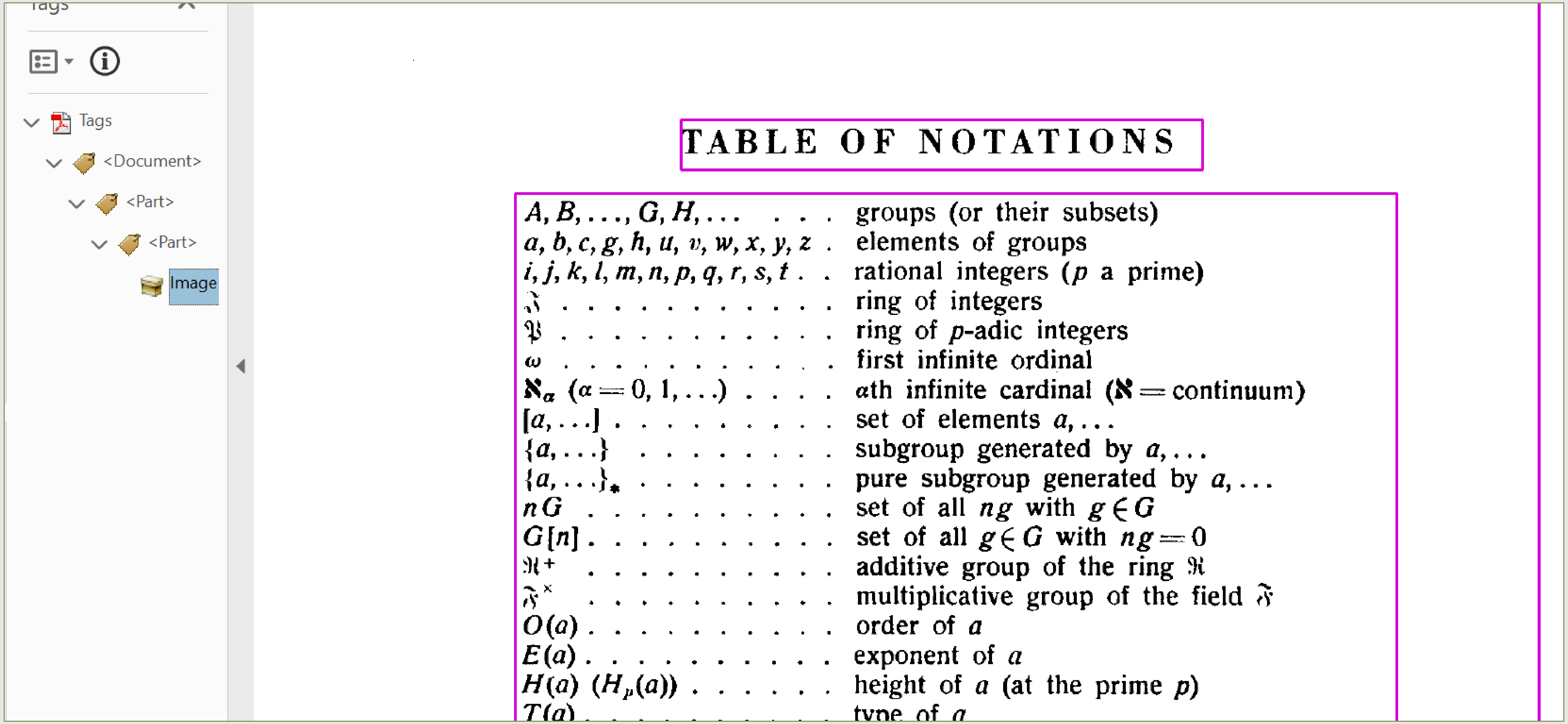
This PDF file is not PDF/UA compliant.

Checkpoint	Passed	Warned	Failed
PDF Syntax	8	0	0
Fonts	32	0	0
Content	2324	0	0
Embedded Files	0	0	0
Natural Language	0	0	1160
Structure Elements	4	0	0
Structure Tree	4	2	0
Role Mapping	6	0	0
Alternate Descriptions	12	0	0
Metadata	4	0	2
Document Settings	2	0	4

PAC3 accessibility report

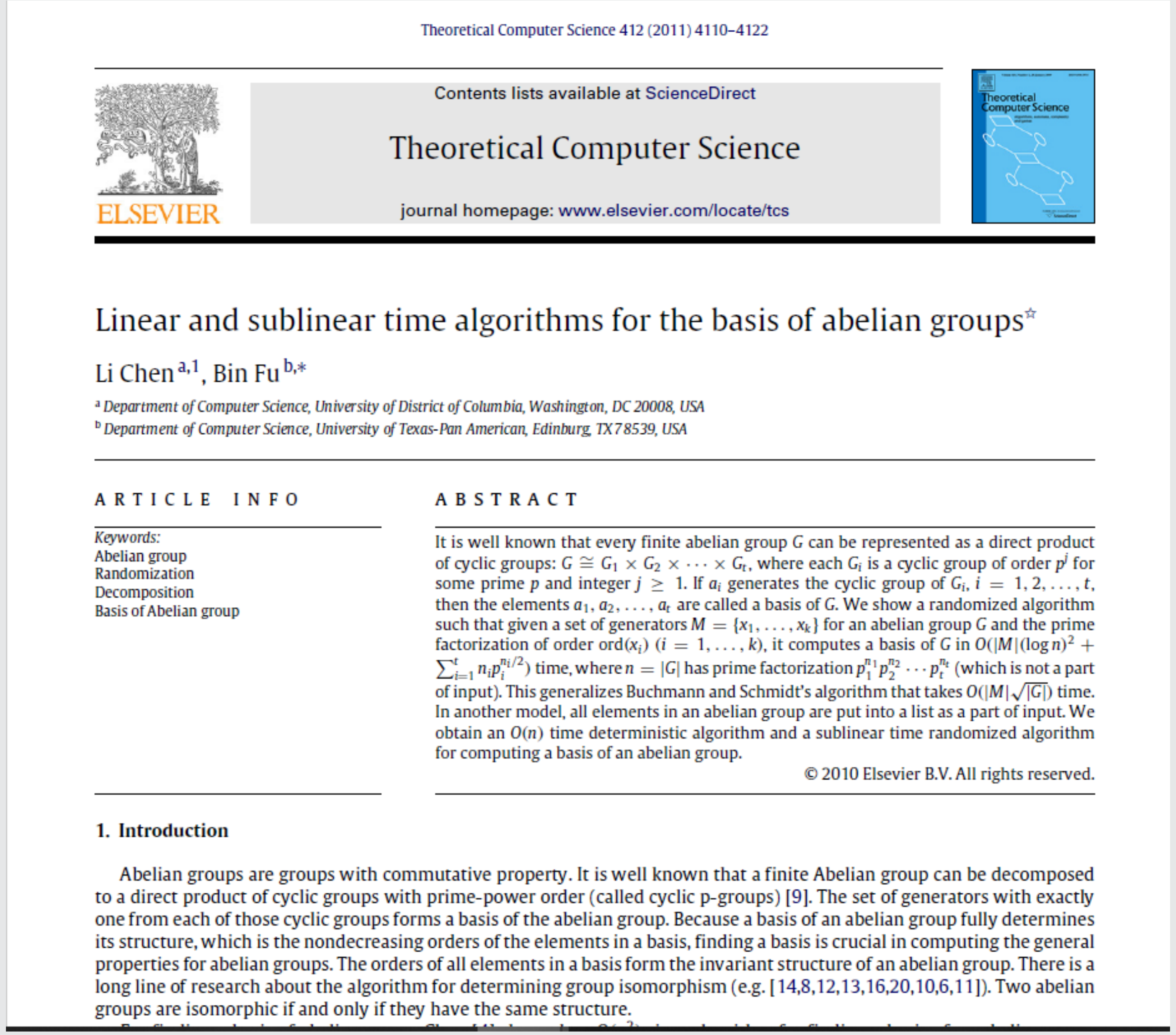
Copyright © 2018, PDF Association

But is this really an accessible tagging?



$A, B, \ldots, G, H, \ldots$	groups (or their subsets)
$a, b, c, g, h, u, v, w, x, y, z$	elements of groups
$i, j, k, l, m, n, p, q, r, s, t$	rational integers (p a prime)
$\mathfrak{Z}$	ring of integers
$\mathfrak{P}$	ring of p-adic integers
ω	first infinite ordinal
$\aleph_\alpha$ ($\alpha = 0, 1, \ldots$)	αth infinite cardinal ($\aleph$ = continuum)
$[a, \ldots]$	set of elements $a, \ldots$
$\{a, \ldots\}$	subgroup generated by $a, \ldots$
$\{a, \ldots\}_*$	pure subgroup generated by $a, \ldots$
nG	set of all ng with $g \in G$
$G[n]$	set of all $g \in G$ with $ng = 0$
$\mathfrak{R}^+$	additive group of the ring $\mathfrak{R}$
$\mathfrak{F}^\times$	multiplicative group of the field $\mathfrak{F}$
$O(a)$	order of a
$E(a)$	exponent of a
$H(a)$ ($H_p(a)$)	height of a (at the prime p)
$T(a)$	type of a

Modern articles: no improvements



Theoretical Computer Science 412 (2011) 4110–4122

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

ELSEVIER

Linear and sublinear time algorithms for the basis of abelian groups[☆]

Li Chen[a,1], Bin Fu[b,*]

[a] Department of Computer Science, University of District of Columbia, Washington, DC 20008, USA
[b] Department of Computer Science, University of Texas-Pan American, Edinburg, TX 78539, USA

ARTICLE INFO

Keywords:
Abelian group
Randomization
Decomposition
Basis of Abelian group

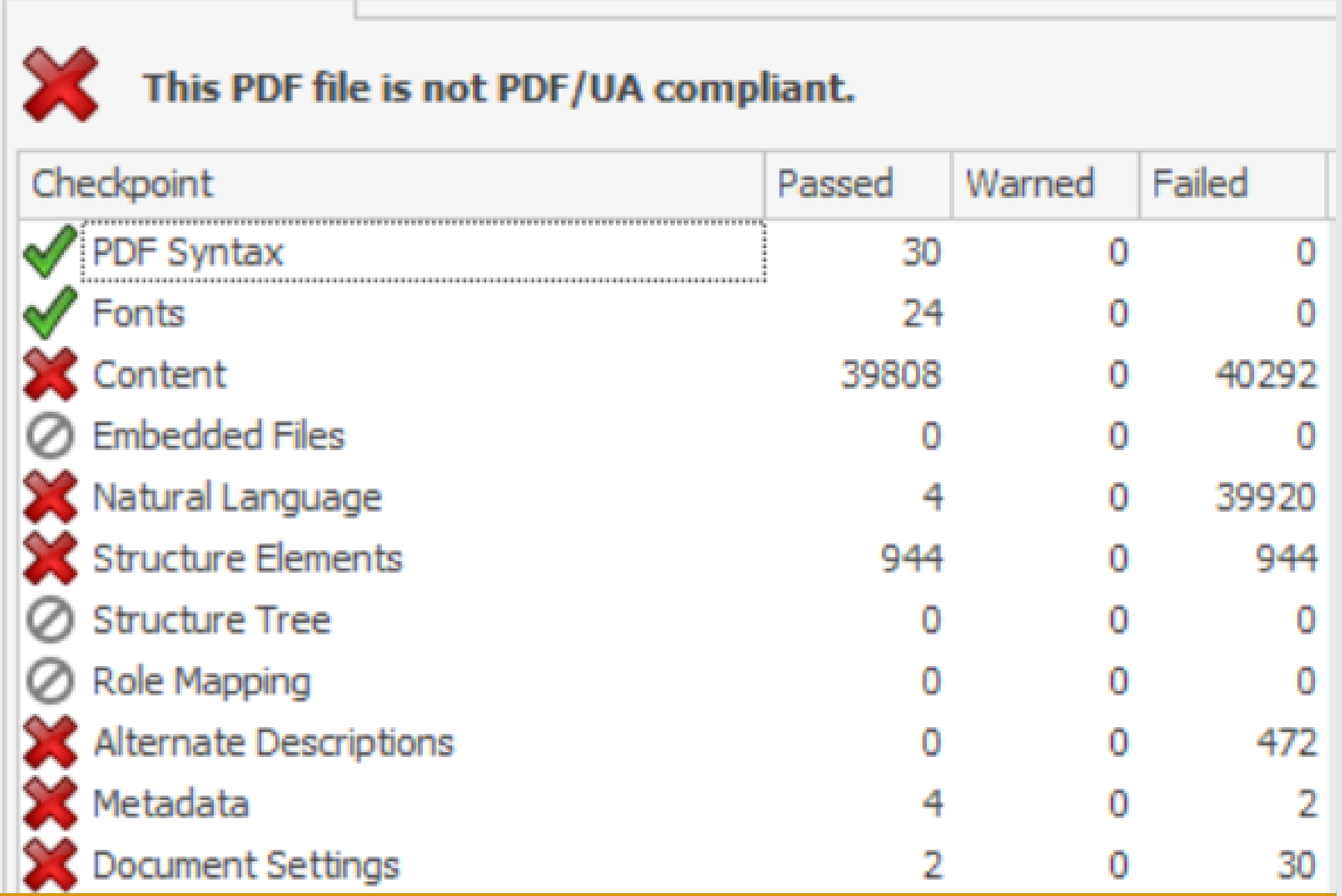
ABSTRACT

It is well known that every finite abelian group G can be represented as a direct product of cyclic groups: $G \cong G_1 \times G_2 \times \cdots \times G_t$, where each G_i is a cyclic group of order p^j for some prime p and integer $j \geq 1$. If a_i generates the cyclic group of G_i, $i = 1, 2, \ldots, t$, then the elements $a_1, a_2, \ldots, a_t$ are called a basis of G. We show a randomized algorithm such that given a set of generators $M = \{x_1, \ldots, x_k\}$ for an abelian group G and the prime factorization of order $\mathrm{ord}(x_i)$ $(i = 1, \ldots, k)$, it computes a basis of G in $O(|M|(\log n)^2 + \sum_{i=1}^{t} n_i p_i^{n_i/2})$ time, where $n = |G|$ has prime factorization $p_1^{n_1} p_2^{n_2} \cdots p_t^{n_t}$ (which is not a part of input). This generalizes Buchmann and Schmidt's algorithm that takes $O(|M|\sqrt{|G|})$ time. In another model, all elements in an abelian group are put into a list as a part of input. We obtain an $O(n)$ time deterministic algorithm and a sublinear time randomized algorithm for computing a basis of an abelian group.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Abelian groups are groups with commutative property. It is well known that a finite Abelian group can be decomposed to a direct product of cyclic groups with prime-power order (called cyclic p-groups) [9]. The set of generators with exactly one from each of those cyclic groups forms a basis of the abelian group. Because a basis of an abelian group fully determines its structure, which is the nondecreasing orders of all elements in a basis, finding a basis is crucial in computing the general properties for abelian groups. The orders of all elements in a basis form the invariant structure of an abelian group. There is a long line of research about the algorithm for determining group isomorphism (e.g. [14,8,12,13,16,20,10,6,11]). Two abelian groups are isomorphic if and only if they have the same structure.



This PDF file is not PDF/UA compliant.

Checkpoint	Passed	Warned	Failed
PDF Syntax	30	0	0
Fonts	24	0	0
Content	39808	0	40292
Embedded Files	0	0	0
Natural Language	4	0	39920
Structure Elements	944	0	944
Structure Tree	0	0	0
Role Mapping	0	0	0
Alternate Descriptions	0	0	472
Metadata	4	0	2
Document Settings	2	0	30

PAC3 accessibility report

Standards

According to the rules…

ISO standards

- Tagged PDF in ISO 32000-1:2008 (PDF 1.7) and ISO 32000-2:2017 (PDF 2.0)
- PDF/UA-1 and PDF/UA-2 (under development)
- MathML is accepted as a the only other namespace that does not require additional role mapping
- PDF 2.0: Any additional data (e.g., MathML or LaTeX representations) can be linked with the <Formula> tag via Associated files.
- All well set?
- **No: as of today there are no authoring tools that can generate fully accessibly scientific publications!**

Copyright © 2018, PDF Association

Tools

Real world implementations

Challenges of the real world

Real world: end-user accessibility

- Screen readers - major players:
  - **NVDA** from NV Access: open source
  - **JAWS** from Freedom Scientific: closed source
  - **MathPlayer** from Design Science: universal MathML support
  - **MathJax** from AMS+SIAM: Math accessibility on Web

- Excessive set of **platform APIs** increases implementation complexity:
  - MSAA+UIA
  - MSAA+IAccessible2
  - ATK/AT-SPI
  - Mac OS X Accessibility
  - Android + iOS

Copyright © 2018, PDF Association

Authoring software

- **InDesign** is used mainly for textbooks up to High School or Bachelor levels
  - Only manual association of MathML with formulas is available (implemented as third-party addons)
- All other scientific publications are prepared by means of **LaTeX**:
  - Based on Knuth's TeX system, which dates back to 1978 (!)
  - PDF output was added by Hàn Thế Thành in 2000 along with the support of many PDF-specific features: hyperlinks, bookmarks, etc.
  - At the moment full support of PDF/X and PDF/A (Levels B and U) is available
  - Driven by the large open source community

Copyright © 2018, PDF Association

Modernizing LaTeX

- LuaTeX / LuaLaTeX
  - Based on C++ source code of PDFTeX
  - Adds Lua scripting and callbacks on all steps of formatting
  - Full Unicode support
  - Full modern fonts support
- ConTeXt
  - New markup, based on LuaTeX
  - However, not compatible with LaTeX
  - Tagged PDF support (!)

Copyright © 2018, PDF Association

LaTeX packages

- **tagpdf** by Ulrike Fischer (fischer@troubleshooting-tex.de)
  - Offers tools to experiment with tagging and accessibility
  - Low-level commands for adding structure elements and marked content sequences
- **accsupp** by Heiko Oberdiek (heiko.oberdiek@googlemail.com)
  - Minimal low-level interface to add alternative text, replacement text, language settings to <Span> marked content sequence
- **axessibility** by Università degli Studi di Torino (www.unito.it)
  - adds original LaTeX syntax as alternative text or replacement text to formulas

Copyright © 2018, PDF Association

Summary

- Despite all standardization efforts, **there is no clear guidance on how Mathematics should be placed inside PDF**, which is accepted by both TeX and Screen reader communities
- As a result:
  - no working implementations on the authoring software
  - no tools for individual scientists and academic publishers to generate accessible scientific content in PDF format
- Additional community challenges:
  - The TeX, PDF and Accessibility developers have very little contact
  - Funding challenges for the open source development

Copyright © 2018, PDF Association

Way forward

- Active ongoing project with the creators of **axessibility** package (**Uni Torino, Italy**)
- Prototype:
  - Automatic PDF tagging based on LaTeX semantics
  - Adding MathML / LaTeX syntax for formulas as Associated files (PDF 2.0)
  - Fully compatible with Derivation algorithm: is able to convert PDF to HTML with accessible formulas (via MathJax)
- To be used as a starting point for tests and discussion between different communities

Copyright © 2018, PDF Association

Questions?

Comments are welcomed.

Copyright © 2018, PDF Association

Thank you!

We appreciate your participation.

Copyright © 2018, PDF Association