FIZ Karlsruhe

Leibniz Institute for Information Infrastructure

The publication-based approach for mathematical models: Identification of mathematical models

Wolfram Sperber

CICM 2018

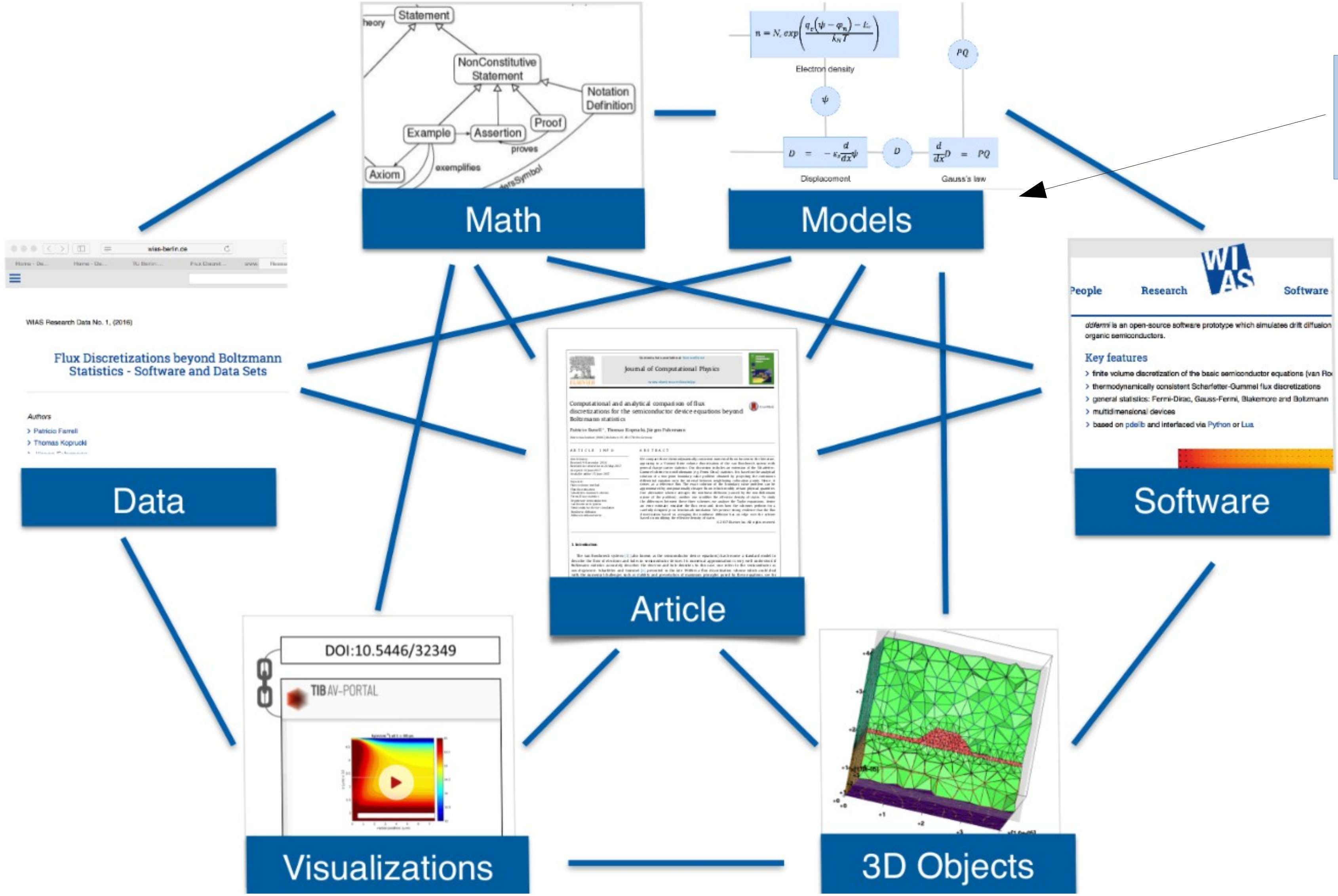
RISC, Hagenberg

2018-08-13

INTRODUCTION
AN EXAMPLE
THE PUBLICATION-BASED APPROACH
MODEL INFORMATION IN PUBLICATIONS
THE MODEL TENSOR
OUTLOOK

The central role of mathematical models

A slide
of Thomas



Applications: Modeling

rlsruhe
te for Information Infrastructure

The goal and the idea

Building up a useful service for mathematical models:

- Thomas/Michael: Models play a central role for mathematical research
- How can we build a useful service for mathematical models?

The problem

- Typically, up to now models are part of publications and not marked as single objects

Difficulties with mathematical models

``Mathematical model” is not a exactly defined term. Mathematical models can be formulated in a real-world context (an increasing number of mathematical models has a non-mathematical background) or context-free.

- Mathematical models are given by definitions, only a small number of models has also a name.

- More pragmatic:
  The definitions of mathematical models in many publications are incomplete.

Challenges for the publication-based approach

- Identification of mathematical models in publications and the analysis of the information about models is closely related:
  Definitions contain specific content information.

  Technically:
  - In general, the definitions of the mathematical model which is investigated is given in the fulltexts of publications, not in the review or abstract.
  - Mathematical definitions contain mathematical expressions (symbols, formulae) plus text. Therefore we need more than the fulltext in PDF format.
    At least TeX-encoded sources of the publications are necessary.
    arXiv seems to be a first resource.

A first approach

1) Where can we find information about mathematical models?

   → analysis of the document structure

2) How can we extract and analyze information about mathematical models?

   Which information about mathematical models is possible?

   → weak description model for mathematical models

Hybrid quantum-classical modeling of quantum dot devices

Markus Kantner,* Markus Mittnenzweig, and Thomas Koprucki
Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin, Germany

The structure (TOC):

Hybrid quantum-classical modeling of quantum dot devices

Markus Kantner,* Markus Mittnenzweig, and Thomas Koprucki
Weierstrass Institute for Applied Analysis and Stochastics,
Mohrenstr. 39, 10117 Berlin, Germany

For humans:
In Title: Modeling
in Abstract: modeling: `By coupling the van Roosbroeck equation with a quantum master equation in Lindblad form, ...'
in Introduction: → Model Equations in Sec. II

$$-\nabla \cdot \varepsilon \nabla \psi = q\left(p - n + C + Q\left(\rho\right)\right), \tag{1}$$

$$\partial_t n - \frac{1}{q}\nabla \cdot \mathbf{j}_n = -R - S_n\left(\rho; n, p, \psi\right), \tag{2}$$

$$\partial_t p + \frac{1}{q}\nabla \cdot \mathbf{j}_p = -R - S_p\left(\rho; n, p, \psi\right), \tag{3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho = \mathcal{L}\left(\rho; n, p, \psi\right) = -\frac{i}{\hbar}\left[H, \rho\right] + \mathcal{D}\left(\rho; n, p, \psi\right) \tag{4}$$

on the domain $\Omega \subset \mathbb{R}^3$. The system (1)–(4) is subject to initial conditions and boundary conditions modeling electrical contacts and other interfaces [32]. See Appendix A

→ Model Consistency III
→ Application

What is the model in the paper?

in textual form:
`hybrid quantum-classical model that self-consistently couples semi-classical transport theory to a kinetic equation for the quantum mechanical density matrix.'

as mathematical expressions (equations):

$$-\nabla \cdot \varepsilon \nabla \psi = q\left(p - n + C + Q\left(\rho\right)\right), \tag{1}$$

$$\partial_t n - \frac{1}{q}\nabla \cdot \mathbf{j}_n = -R - S_n\left(\rho; n, p, \psi\right), \tag{2}$$

$$\partial_t p + \frac{1}{q}\nabla \cdot \mathbf{j}_p = -R - S_p\left(\rho; n, p, \psi\right), \tag{3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho = \mathcal{L}\left(\rho; n, p, \psi\right) = -\frac{i}{\hbar}\left[H, \rho\right] + \mathcal{D}\left(\rho; n, p, \psi\right) \tag{4}$$

on the domain $\Omega \subset \mathbb{R}^3$. The system (1)–(4) is subject to initial conditions and boundary conditions modeling electrical contacts and other interfaces [32]. See Appendix A

But

- the model is incomplete: (1)-(4) are the partial differential equations. They must be completed by initial- and boundary equations
  →appendix (This is trivial for the mathematicians.)

- the model is specified (some hypotheses for the application case):
- (1)-(4) are the basic equations
- (1)-(3) van Roosbroeck (semi-classical transport model for electrons and holes)
- (4) quantum master equation
- there are additional equations and principles which are presented in the subsection A, B, C and section III (evaluation of modeling: correctness accordance with the principles of thermodynamics).

- *Remark: This papers describes the modeling process (the model is the result of this paper), not the mathematical treatment of the model.*

A first summary

- The model is given by text and formulae.
  For the mathematical treatment the formulae are the most important part. Often the textual environment gives some information about the meaning of the expression and the background of the model (origin / use cases) but can contain also structural information as in the example (van Roosbroeck, quantum master equatiion in Lindblad form, links to scientific theories, algorithms, software, ...)
- Both the textual information and the mathematical expressions which describe a model are relevant.
- Therefore we should try to extract both
  - the mathematical expressions describing the model
  - the corresponding terms and phrases for the mathematical expressions

- *Remark: Terms, phrases and mathematical expressions are ambigious, but mathematical expressions, especially complex objects as equations, inequalities, structural diagrams contain precise information about their structure which allow to compare these objects and enhanced retrieval features.*

What can we do automatically?

- **The publication-based approach for mathematical models**

  Up to now, models are part of publications. →Therefore we have to develop methods to extract the
- mathematical expressions describing the mathematical model and the corresponding terms and
- phrases.

Remarks:
- *Fulltexts:*
  The most abstracts or reviews, the zbMATH data, don't contain precise model descriptions. Typically, the mathematical expressions are not part of the abstract or reviews. Therefore, we need the fulltexts of the models.
- *Formats:*
  Up to now, no powerful converter exists for transforming PDF formatted documents to TeX or MathML. At least, TeX encoded documents are needed. Therefore the arXiv e-prints are a first resource to develop methods for model analysis.

The publication-based approach

The publication-based approach covers three steps:

1) Identification of the model-relevant parts of the publication

- Model relevant sections or subsections
  which can be identified by characteristic terms in the titles, especially `model*' or `problem*'

alternatively

- the beginning of a document, e.g., the first 3,000 characters

A statistic

Journal	(Sub-)sections containg the terms \`model‘ or \`problem‘
International Journal of Robust and Nonlinear Control, v. 10, n. 28	8 (9)
International Journal of Robust and Nonlinear Control v. 10, n. 29	6 (8)
Journal of System Science & Complexity v. 31, n. 3	7 (13)
Cybernetics and System Analysis v. 54, n. 2	8 (18)
International Journal of Control and Signal Processing v. 32, n. 6	7 (10)
Automation and Remote Control, v. 79, n. 5	6 (12)

Identification of model-relevant parts: results and open questions

by eliminating the non-relevant parts we get reduced (TeX-encoded) artifacts which contain only the model-relevant parts

- *Remark: It is not clear that the model-relevant parts contain only one model information.*
  The reduced document can also contain previous models, similar models, specific models, generalized models, transformed models, etc.). In other words, it may contain a lot of additional information.
  How can we eliminate redundant / irrelevant information?
  (If yes, we could markup the model in the original file by a `model tag')

  Our example: The section II contains ~ 100 mathematical expressions
  9 / A:22 /B:29 / C:18 /fig.2:(15) text:10) (moreover, the mathematical expression from appendix A must be added).
  The frequencies of some expressions is > 1.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Analysis of the model-reduced documents

2) Analysis of our artifacts:

1) extraction of all mathematical expressions (TeX or MathML encoded expressions) and the corresponding terms and phrases

   → creating a `model-related' matrix A_{i3} (1 <= i <= n, n – number of mathematical expression in the model-reduced document) for a document with the components :

   a_{i1}: mathematical expression

   a_{i2}: left-hand textual neighborhood (k terms)

   a_{i3}: right-hand textual neighborhood (k terms)

Example:	C	doping profile	0
	Q(p)	expectation value of the charge density	0
	n	free electron density	0
	p	free hole density	0
	(1)-(3)	0	van Roosbroeck system

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Analysis of the model-reduced documents: results and open questions

storing in a database (A_{i3}^{k} with k as the number of k-th document)

Remarks: As result we get a tensor which contains mathematical expressions plus their neighborhood. The mathematical expressions are listed in the sequence of their occurrence. This can make difficulties, e.g., parts of the model are presented at different positions in the document.
Can we identify the input variables of the model?
How can we organize it in a better way (sorting by length is not a real good idea)?
How can we identify the system structure of a model?
How can we find relevant textual information for a mathematical expression which is not in the left- or in the right-hand neighborhood of a mathematical expression?
What do we do with this information (e.g, storing this information in a special field of the tensor)?

Further processing: Transformation of the TeX-encoded mathematical expressions

3) It is difficult to compare TeX-encoded mathematical expressions
(the TeX codes for a PDF document differ; zbMATH experience: mathematical expressions with more than 6 characters occur in the database zbMATH only once)
better (more standardized): XML / MathML encoded mathematical expression
(tree structured documents)
MathML transformation of mathematical expressions via **LaTeXML**
→ tree-structured documents which allow to compare the structure of mathematical formulae

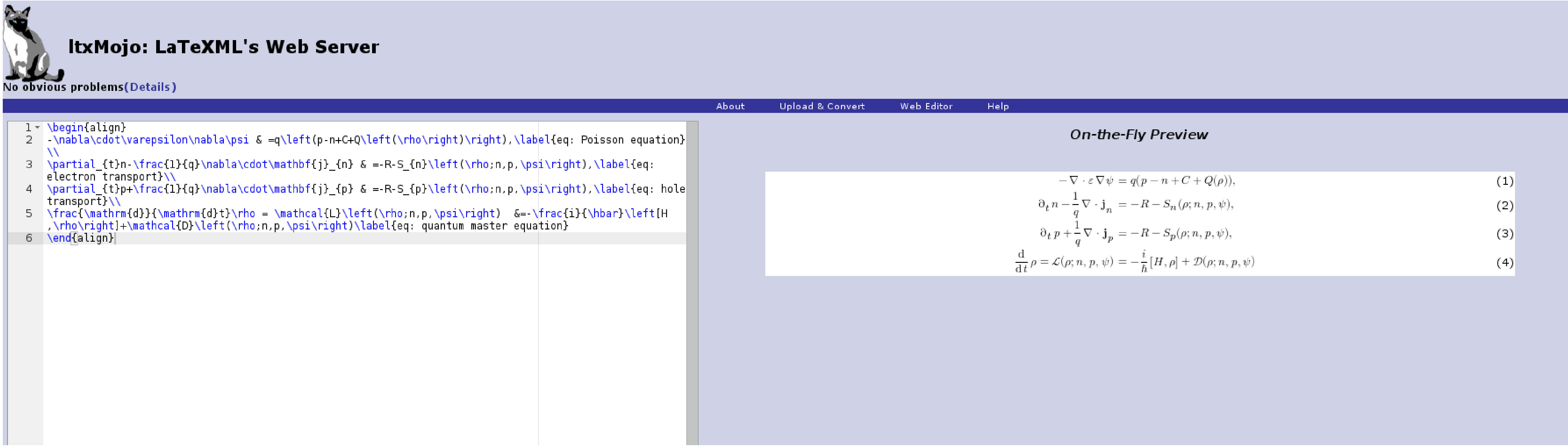
Remarks: LaTeXML transform the TeX code to XML. By postprocessing, the XML code can be transformed to MathML
(Presentation Markup, Content Markup, or OpenMATH)
Which format should be used for storing, all or only MathML Content Markup?
(If the authors would write their documents in STeX, MathML or OpenMATH should be used).

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Example (the PDE system 1 - 4)

```
\begin{align}

-\nabla\cdot\varepsilon\nabla\psi & =q\left(p-n+C+Q\left(\rho\right)\right),
\label{eq: Poisson equation}\\

\partial_{t}n-\frac{1}{q}\nabla\cdot\mathbf{j}_{n} & =-R-S_{n}\left(\rho;n,p,\psi\right),\label{eq:
electron transport}\\

\partial_{t}p+\frac{1}{q}\nabla\cdot\mathbf{j}_{p} & =-R-S_{p}\left(\rho;n,p,\psi\right),\label{eq: hole
transport}\\

\frac{\mathrm{d}}{\mathrm{d}t}\rho = \mathcal{L}\left(\rho;n,p,\psi\right)  &=-\frac{i}
{\hbar}\left[H,\rho\right]+\mathcal{D}\left(\rho;n,p,\psi\right)\label{eq: quantum master equation}

\end{align}
```

The Online Editor of LaTeXML



```
          </bibref></cite>. The model equations read</p>
      <equationgroup class="ltx_eqn_align" xml:id="A6.EGx1">
        <equation frefnum="(1)" labels="LABEL:eq:_Poisson_equation" refnum="1" xml:id="S2.E1">
          <MathFork>
            <Math tex="\displaystyle-\nabla\cdot\varepsilon\nabla\psi=q\left(p-n+C+Q\left(\rho\right)%&#10;\right)," text="- (nabla cdot varepsilon) * nabla@(psi) = q * ((p - n) + C + Q * rho)" xml:id="S2.E1.m3">
              <XMath>
                <XMDual>
                  <XMRef idref="S2.E1.m3.2"/>
                  <XMWrap>
                    <XMApp xml:id="S2.E1.m3.2">
                      <XMTok meaning="equals" role="RELOP">=</XMTok>
                      <XMApp>
                        <XMTok meaning="minus" role="ADDOP">-</XMTok>
                        <XMApp>
                          <XMTok meaning="times" role="MULOP"></XMTok>
                          <XMApp>
                            <XMTok name="cdot" role="MULOP">·</XMTok>
                            <XMTok name="nabla" role="OPERATOR">∇</XMTok>
                            <XMTok font="italic" name="varepsilon" role="UNKNOWN">ε</XMTok>
                          </XMApp>
                          <XMApp>
                            <XMTok name="nabla" role="OPERATOR">∇</XMTok>
                            <XMTok font="italic" name="psi" role="UNKNOWN">ψ</XMTok>
                          </XMApp>
                        </XMApp>
                      </XMApp>
                      <XMApp>
                        <XMTok meaning="times" role="MULOP"></XMTok>
                        <XMTok font="italic" role="UNKNOWN">q</XMTok>
                        <XMDual>
                          <XMRef idref="S2.E1.m3.2.1"/>
                          <XMWrap>
                            <XMTok role="OPEN" stretchy="true">(</XMTok>
                            <XMApp xml:id="S2.E1.m3.2.1">
                              <XMTok meaning="plus" role="ADDOP">+</XMTok>
                              <XMApp>
                                <XMTok meaning="minus" role="ADDOP">-</XMTok>
                                <XMTok font="italic" role="UNKNOWN">p</XMTok>
                                <XMTok font="italic" role="UNKNOWN">n</XMTok>
                              </XMApp>
                              <XMTok font="italic" role="UNKNOWN">C</XMTok>
                              <XMApp>
                                <XMTok meaning="times" role="MULOP"></XMTok>
                                <XMTok font="italic" role="UNKNOWN">Q</XMTok>
                                <XMDual>
                                  <XMRef idref="S2.E1.m3.1"/>
                                  <XMWrap>
                                    <XMTok role="OPEN" stretchy="true">(</XMTok>
                                    <XMTok font="italic" name="rho" role="UNKNOWN" xml:id="S2.E1.m3.1">ρ</XMTok>
                                    <XMTok role="CLOSE" stretchy="true">)</XMTok>
                                  </XMWrap>
                                </XMDual>
                              </XMApp>
                            </XMApp>
                            <XMTok role="CLOSE" stretchy="true">)</XMTok>
                          </XMWrap>
                        </XMDual>
```

```
ref>)</cite>. The model equations read</p>
      <equationgroup class="ltx_eqn_align" xml:id="A6.EGx1" fragid="A6.EGx1">
        <equation frefnum="(1)" labels="LABEL:eq:_Poisson_equation" refnum="1" xml:id="S2.E1" fragid="S2.E1">
          <MathFork>
            <Math tex="\displaystyle-\nabla\cdot\varepsilon\nabla\psi=q\left(p-n+C+Q\left(\rho\right)%&#10;\right)," text="- (nabla cdot varepsilon) * nabla@(psi) = q * ((p - n) + C + Q * rho)" xml:id="S2.E1.m3" fragid="S2.E1.m3"><m:math alttext="\displaystyle-\nabla\cdot\varepsilon\nabla\psi=q\left(p-n+C+Q\left(\rho\right)%&#10;\right)," display="inline"><m:mrow><m:mrow><m:mrow><m:mo>-</m:mo><m:mrow><m:mrow><m:mo>∇</m:mo><m:mo>⋅</m:mo><m:mi>ε</m:mi></m:mrow><m:mo></m:mo><m:mrow><m:mo>∇</m:mo><m:mo></m:mo><m:mi>ψ</m:mi></m:mrow></m:mrow></m:mrow><m:mo>=</m:mo><m:mrow><m:mi>q</m:mi><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mrow><m:mrow><m:mi>p</m:mi><m:mo>-</m:mo><m:mi>n</m:mi></m:mrow><m:mo>+</m:mo><m:mi>C</m:mi><m:mo>+</m:mo><m:mrow><m:mi>Q</m:mi><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math>
            <MathBranch>
              <td align="right"><Math mode="inline" tex="\displaystyle-\nabla\cdot\varepsilon\nabla\psi" text="- (nabla cdot varepsilon) * nabla@(psi)" xml:id="S2.E1.m1" fragid="S2.E1.m1"><m:math alttext="\displaystyle-\nabla\cdot\varepsilon\nabla\psi" display="inline"><m:mrow><m:mo>-</m:mo><m:mrow><m:mrow><m:mo>∇</m:mo><m:mo>⋅</m:mo><m:mi>ε</m:mi></m:mrow><m:mo></m:mo><m:mrow><m:mo>∇</m:mo><m:mo></m:mo><m:mi>ψ</m:mi></m:mrow></m:mrow></m:mrow></m:math></Math></td>
              <td align="left"><Math mode="inline" tex="\displaystyle=q\left(p-n+C+Q\left(\rho\right)\right)," text="absent = q * ((p - n) + C + Q * rho)" xml:id="S2.E1.m2" fragid="S2.E1.m2"><m:math alttext="\displaystyle=q\left(p-n+C+Q\left(\rho\right)\right)," display="inline"><m:mrow><m:mrow><m:mi/><m:mo>=</m:mo><m:mrow><m:mi>q</m:mi><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mrow><m:mrow><m:mi>p</m:mi><m:mo>-</m:mo><m:mi>n</m:mi></m:mrow><m:mo>+</m:mo><m:mi>C</m:mi><m:mo>+</m:mo><m:mrow><m:mi>Q</m:mi><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math></td>
            </MathBranch>
          </MathFork>
        </equation>
        <equation frefnum="(2)" labels="LABEL:eq:_electron_transport" refnum="2" xml:id="S2.E2" fragid="S2.E2">
          <MathFork>
            <Math tex="\displaystyle\partial_{t}n-\frac{1}{q}\nabla\cdot\mathbf{j}_{n}=-R-S_{n}\left(%&#10;\rho;n,p,\psi\right)," text="(partial-differential _ t)@(n) - ((1 / q) * nabla) cdot j _ n = - R - S _ n * list@(rho, n, p, psi)" xml:id="S2.E2.m3" fragid="S2.E2.m3"><m:math alttext="\displaystyle\partial_{t}n-\frac{1}{q}\nabla\cdot\mathbf{j}_{n}=-R-S_{n}\left(%&#10;\rho;n,p,\psi\right)," display="inline"><m:mrow><m:mrow><m:mrow><m:mrow><m:msub><m:mo>∂</m:mo><m:mi>t</m:mi></m:msub><m:mo></m:mo><m:mi>n</m:mi></m:mrow><m:mo>-</m:mo><m:mrow><m:mrow><m:mstyle displaystyle="true"><m:mfrac><m:mn>1</m:mn><m:mi>q</m:mi></m:mfrac></m:mstyle><m:mo></m:mo><m:mo>∇</m:mo></m:mrow><m:mo>⋅</m:mo><m:msub><m:mi>𝐣</m:mi><m:mi>n</m:mi></m:msub></m:mrow></m:mrow><m:mo>=</m:mo><m:mrow><m:mrow><m:mo>-</m:mo><m:mi>R</m:mi></m:mrow><m:mo>-</m:mo><m:mrow><m:msub><m:mi>S</m:mi><m:mi>n</m:mi></m:msub><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>;</m:mo><m:mi>n</m:mi><m:mo>,</m:mo><m:mi>p</m:mi><m:mo>,</m:mo><m:mi>ψ</m:mi><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow></m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math>
            <MathBranch>
              <td align="right"><Math mode="inline" tex="\displaystyle\partial_{t}n-\frac{1}{q}\nabla\cdot\mathbf{j}_{n}" text="(partial-differential _ t)@(n) - ((1 / q) * nabla) cdot j _ n" xml:id="S2.E2.m1" fragid="S2.E2.m1"><m:math alttext="\displaystyle\partial_{t}n-\frac{1}{q}\nabla\cdot\mathbf{j}_{n}" display="inline"><m:mrow><m:mrow><m:msub><m:mo>∂</m:mo><m:mi>t</m:mi></m:msub><m:mo></m:mo><m:mi>n</m:mi></m:mrow><m:mo>-</m:mo><m:mrow><m:mrow><m:mstyle displaystyle="true"><m:mfrac><m:mn>1</m:mn><m:mi>q</m:mi></m:mfrac></m:mstyle><m:mo></m:mo><m:mo>∇</m:mo></m:mrow><m:mo>⋅</m:mo><m:msub><m:mi>𝐣</m:mi><m:mi>n</m:mi></m:msub></m:mrow></m:mrow></m:math></Math></td>
              <td align="left"><Math mode="inline" tex="\displaystyle=-R-S_{n}\left(\rho;n,p,\psi\right)," text="absent = - R - S _ n * list@(rho, n, p, psi)" xml:id="S2.E2.m2" fragid="S2.E2.m2"><m:math alttext="\displaystyle=-R-S_{n}\left(\rho;n,p,\psi\right)," display="inline"><m:mrow><m:mrow><m:mi/><m:mo>=</m:mo><m:mrow><m:mrow><m:mo>-</m:mo><m:mi>R</m:mi></m:mrow><m:mo>-</m:mo><m:mrow><m:msub><m:mi>S</m:mi><m:mi>n</m:mi></m:msub><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>;</m:mo><m:mi>n</m:mi><m:mo>,</m:mo><m:mi>p</m:mi><m:mo>,</m:mo><m:mi>ψ</m:mi><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow></m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math></td>
            </MathBranch>
          </MathFork>
        </equation>
        <equation frefnum="(3)" labels="LABEL:eq:_hole_transport" refnum="3" xml:id="S2.E3" fragid="S2.E3">
          <MathFork>
            <Math tex="\displaystyle\partial_{t}p+\frac{1}{q}\nabla\cdot\mathbf{j}_{p}=-R-S_{p}\left(%&#10;\rho;n,p,\psi\right)," text="(partial-differential _ t)@(p) + ((1 / q) * nabla) cdot j _ p = - R - S _ p * list@(rho, n, p, psi)" xml:id="S2.E3.m3" fragid="S2.E3.m3"><m:math alttext="\displaystyle\partial_{t}p+\frac{1}{q}\nabla\cdot\mathbf{j}_{p}=-R-S_{p}\left(%&#10;\rho;n,p,\psi\right)," display="inline"><m:mrow><m:mrow><m:mrow><m:mrow><m:msub><m:mo>∂</m:mo><m:mi>t</m:mi></m:msub><m:mo></m:mo><m:mi>p</m:mi></m:mrow><m:mo>+</m:mo><m:mrow><m:mrow><m:mstyle displaystyle="true"><m:mfrac><m:mn>1</m:mn><m:mi>q</m:mi></m:mfrac></m:mstyle><m:mo></m:mo><m:mo>∇</m:mo></m:mrow><m:mo>⋅</m:mo><m:msub><m:mi>𝐣</m:mi><m:mi>p</m:mi></m:msub></m:mrow></m:mrow><m:mo>=</m:mo><m:mrow><m:mrow><m:mo>-</m:mo><m:mi>R</m:mi></m:mrow><m:mo>-</m:mo><m:mrow><m:msub><m:mi>S</m:mi><m:mi>p</m:mi></m:msub><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>;</m:mo><m:mi>n</m:mi><m:mo>,</m:mo><m:mi>p</m:mi><m:mo>,</m:mo><m:mi>ψ</m:mi><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow></m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math>
            <MathBranch>
              <td align="right"><Math mode="inline" tex="\displaystyle\partial_{t}p+\frac{1}{q}\nabla\cdot\mathbf{j}_{p}" text="(partial-differential _ t)@(p) + ((1 / q) * nabla) cdot j _ p" xml:id="S2.E3.m1" fragid="S2.E3.m1"><m:math alttext="\displaystyle\partial_{t}p+\frac{1}{q}\nabla\cdot\mathbf{j}_{p}" display="inline"><m:mrow><m:mrow><m:msub><m:mo>∂</m:mo><m:mi>t</m:mi></m:msub><m:mo></m:mo><m:mi>p</m:mi></m:mrow><m:mo>+</m:mo><m:mrow><m:mrow><m:mstyle displaystyle="true"><m:mfrac><m:mn>1</m:mn><m:mi>q</m:mi></m:mfrac></m:mstyle><m:mo></m:mo><m:mo>∇</m:mo></m:mrow><m:mo>⋅</m:mo><m:msub><m:mi>𝐣</m:mi><m:mi>p</m:mi></m:msub></m:mrow></m:mrow></m:math></Math></td>
              <td align="left"><Math mode="inline" tex="\displaystyle=-R-S_{p}\left(\rho;n,p,\psi\right)," text="absent = - R - S _ p * list@(rho, n, p, psi)" xml:id="S2.E3.m2" fragid="S2.E3.m2"><m:math alttext="\displaystyle=-R-S_{p}\left(\rho;n,p,\psi\right)," display="inline"><m:mrow><m:mrow><m:mi/><m:mo>=</m:mo><m:mrow><m:mrow><m:mo>-</m:mo><m:mi>R</m:mi></m:mrow><m:mo>-</m:mo><m:mrow><m:msub><m:mi>S</m:mi><m:mi>p</m:mi></m:msub><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>;</m:mo><m:mi>n</m:mi><m:mo>,</m:mo><m:mi>p</m:mi><m:mo>,</m:mo><m:mi>ψ</m:mi><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow></m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math></td>
            </MathBranch>
```


XML Tab Width: 8 Ln 142, Col 253 INS

Some remarks and questions

All mathematical expressions of our model description tensor could be transformed on this way to MathML PM or CM or OpenMATH. This allows to use graph-based models, e.g. for comparing models or retrieval.

The advantage of the approach:
All steps of the publication-based approach can be done widely automatically.
But the method has also some disadvantages:
In principle our model tensors of publications should contain the model information, but it is possible that the information is too much. How can we remove model-irrelevant information?
Big tensors describing a model are not very impressive for the user. How can we compactify the model representations

(A first idea:
A grouping of complex mathematical expressions with the involved notations should be possible. Often long mathematical expression are complex expressions (especially PDEs) containing special symbols which we could connect).

Further questions

- What is with retrieval (access for mathematicans and for non-mathematians)?
- How can we detect similar models (ranking of the mathematical expressions of the models)?
- How can we make the service interactive (for improving the model, linking to further model information)?
- How can we make the modeling process understandable?

Our ambitioned goal: Building up a knowledge base for mathematical models

Full semantic
formalization for
automatic processing

Visual semantic presentation
of models
(Model Pathway Diagrams)

Publication-based descriptions
of mathematical models

Example 2:
The model
description
in the e-print
arXiv:1806.07795

1. Introduction

In this paper, we consider a system of N spherical particles $(B_i)_{1\leq i\leq N}$ with identical radii R immersed in a viscous fluid satisfying the following Stokes equation:

$$\left\{\begin{array}{rcl} -\Delta u^N + \nabla p^N &=& 0, \\ \operatorname{div} u^N &=& 0, \end{array}\right. \text{ on } \mathbb{R}^3 \setminus \bigcup_{i=1}^N \overline{B_i}, \tag{1}$$

completed with the no-slip boundary conditions:

$$\left\{\begin{array}{rcl} u^N &=& V_i + \Omega_i \times (x - x_i), \text{ on } \partial B_i, \\ \lim\limits_{|x|\to\infty} |u^N(x)| &=& 0, \end{array}\right. \tag{2}$$

where $(V_i, \Omega_i) \in \mathbb{R}^3 \times \mathbb{R}^3$, $1 \leq i \leq N$ represent the linear and angular velocities,

$$B_i := B(x_i, R).$$

We describe the intertialess motion of the rigid spheres $(B_i)_{1\leq i\leq N}$ by adding to the instantaneous Stokes equation the classical Newton dynamics for the particles $(x_i)_{1\leq i\leq N}$:

$$\left\{\begin{array}{rcl} \dot{x}_i &=& V_i, \\ F_i + mg &=& 0, \\ T_i &=& 0, \end{array}\right. \tag{3}$$

where m denotes the mass of the identical particles adjusted for buoyancy, g the gravitational acceleration, F_i (resp. T_i) the drag force (resp. the torque) applied by the the fluid

on the i^{th} particle B_i defined as

$$\begin{array}{rcl} F_i &:=& \displaystyle\int_{\partial B_i} \sigma(u^N, p^N)n, \\ T_i &=& \displaystyle\int_{\partial B_i} (x - x_i) \times [\sigma(u^N, p^N)n], \end{array}$$

with n the unit outer normal to ∂B_i and $\sigma(u^N, p^N) = 2D(u^N) - p^N\mathbb{I}$, the stress tensor where $2D(u^N) = \nabla u^N + \nabla u^{N\top}$.

The publicly available TeX-code

```
\section{Introduction}
In this paper, we consider a system of $N$ spherical particles $(B_i)_{1\leq i \leq N}$ with
identical radii $R$ immersed in a viscous fluid satisfying the following Stokes equation:
\begin{equation}\label{eq_stokes}
\left \{
\begin{array}{rcl}
-\Delta u^N + \nabla p^N & = &0, \\
\div u^N & = & 0,
\end{array}
\text { on $\mathbb{R}^3 \setminus \underset{i=1}{\overset{N}{\bigcup}}\overline{B_i}$, }
\right.
\end{equation}
completed with the no-slip boundary conditions:
\begin{equation}\label{cab_stokes}
\left \{
\begin{array}{rcl}
u^N & = &  V_i +\Omega_i \times (x-x_i), \text{ on } \partial {B_i}, \\
\underset{|x| \to \infty}{\lim} |u^N(x)| & = & 0,
\end{array}\right.
\end{equation}
where $(V_i,\Omega_i) \in \mathbb{R}^3\times \mathbb{R}^3 \,,\, 1\leq i \leq N$ represent the
linear and angular velocities,
\begin{eqnarray*}
B_i:= B(x_i,R ).
\end{eqnarray*}
We describe the intertialess motion of the rigid spheres $(B_i)_{1\leq i \leq N}$ by adding to
the instantaneous Stokes equation the classical Newton dynamics for the particles $
(x_i)_{1\leq i\leq N}$:
\begin{equation}\label{ode}
\left \{
\begin{array}{rcl}
\dot{x}_i &= & V_i,\\
F_i+ mg &=& 0,\\
T_i&=& 0,
\end{array}
\right.
\end{equation}
where $m$ denotes the mass of the identical particles adjusted for buoyancy, $g$ the
gravitational acceleration, $F_i$ (resp. $T_i$) the drag force (resp. the torque) applied by
the the fluid on the $i^{th}$ particle $B_i$ defined as
\begin{eqnarray*}
F_i & := & \int_{\partial B_i} \sigma(u^N,p^N) n, \\
T_i & = &  \int_{\partial B_i} (x-x_i) \times [\sigma(u^N,p^N) n],
\end{eqnarray*}
with $n$ the unit outer normal to $\partial B_i$ and
$\sigma(u^N,p^N)= 2 D(u^N) - p^N \mathbb{I}, $ the stress tensor where $2D(u^N)=\nabla u^N +{
\nabla u^N }^{\top} $.\\
```

The XML snippet of the Stokes equation (via application of LaTeXML)

```
        </Math> immersed in a viscous fluid satisfying the following Stokes equation:
<!--  %**** mecherbet˙sedimentation.tex Line 50 **** --></p>
      <equation frefnum="(1)" labels="LABEL:eq_stokes" refnum="1" xml:id="S1.E1">
        <Math mode="display" tex="\left\{\begin{array}[]{rcl}-\Delta u^{N}+\nabla p^{N}&amp;=&amp;0,\\&#10;\operatorname{div}u^{N}&amp;=&amp;0,\end{array}\text{ on $\mathbb{R}^{3}\s
\overset{N}{\bigcup}}\overline{B_{i}}$, }\right." text="cases@(Array[[(- Delta * u ^ N) + nabla@(p ^ N), =, 0], [div@(u ^ N), =, 0]]) * [ on \R3=i1N⋃¯Bi, ]" xml:id="S1.E1.m1">
          <XMath>
            <XMApp>
              <XMTok meaning="times" role="MULOP"></XMTok>
              <XMDual>
                <XMApp>
                  <XMTok meaning="cases"/>
                  <XMRef idref="S1.E1.m1.3"/>
                </XMApp>
                <XMWrap>
                  <XMTok role="OPEN" stretchy="true">{</XMTok>
                  <XMArray role="ARRAY" vattach="middle" xml:id="S1.E1.m1.3">
                    <XMRow>
                      <XMCell align="right">
                        <XMApp>
                          <XMTok meaning="plus" role="ADDOP">+</XMTok>
                          <XMApp>
                            <XMTok meaning="minus" role="ADDOP">-</XMTok>
                            <XMApp>
                              <XMTok meaning="times" role="MULOP"></XMTok>
                              <XMTok name="Delta" role="UNKNOWN">Δ</XMTok>
                              <XMApp>
                                <XMTok role="SUPERSCRIPTOP" scriptpos="post8"/>
                                <XMTok font="italic" role="UNKNOWN">u</XMTok>
                                <XMTok font="italic" fontsize="70%" role="UNKNOWN">N</XMTok>
                              </XMApp>
                            </XMApp>
                          </XMApp>
```

A MathML (PM) snippet of the model (via LaTeXML)


```
<title><tag close=". ">1</tag>Introduction</title>
<toctitle><tag close=" ">1</tag>Introduction</toctitle>
<para xml:id="S1.p1" fragid="S1.p1">
  <p>In this paper, we consider a system of <Math mode="inline" tex="N" text="N" xml:id="S1.p1.m1" fragid="S1.p1.m1"><m:math alttext="N" display="inline"><m:mi>N</m:mi></m:math></Math> spher
<Math mode="inline" tex="(B_{i})_{1\leq i\leq N}" text="B _ i _ (1 less= i less= N)" xml:id="S1.p1.m2" fragid="S1.p1.m2"><m:math alttext="(B_{i})_{1\leq i\leq N}" display="inline"><m:msub><m:mrow
y="false">(</m:mo><m:msub><m:mi>B</m:mi><m:mi>i</m:mi></m:msub><m:mo stretchy="false">)</m:mo></m:mrow><m:mrow><m:mn>1</m:mn><m:mo>≤</m:mo><m:mi>i</m:mi><m:mo>≤</m:mo><m:mi>N</m:mi></m:mrow></m:
/Math> with identical radii <Math mode="inline" tex="R" text="R" xml:id="S1.p1.m3" fragid="S1.p1.m3"><m:math alttext="R" display="inline"><m:mi>R</m:mi></m:math></Math> immersed in a viscous flu
he following Stokes equation:
<!-- %**** mecherbet sedimentation.tex Line 50 **** --></p>
  <equation frefnum="(1)" labels="LABEL:eq_stokes" refnum="1" xml:id="S1.E1" fragid="S1.E1">
   <Math mode="display" tex="\left\{\begin{array}[]{rcl}-\Delta u^{N}+\nabla p^{N}&amp;=&amp;0,\\&#10;\operatorname{div}u^{N}&amp;=&amp;0,\end{array}\text{ on $\mathbb{R}^{3}\setminus%&#10;
\overset{N}{\bigcup}\overline{B_{i}}$, }\right." text="cases@(Array[[(- Delta * u ^ N) + nabla@(p ^ N), =, 0], [div@(u ^ N), =, 0]]) * [ on \R3=i1N⋃¯Bi, ]" xml:id="S1.E1.m1" fragid="S1.E1.m1"><m
"\left\{\begin{array}[]{rcl}-\Delta u^{N}+\nabla p^{N}&amp;=&amp;0,\\&#10;\operatorname{div}u^{N}&amp;=&amp;0,\end{array}\text{ on $\mathbb{R}^{3}\setminus%&#10;\underset{i=1}{\overset{N}{\bigcup
{i}}$, }\right." display="block"><m:mrow><m:mrow><m:mo>{</m:mo><m:mtable columnspacing="5pt" displaystyle="true" rowspacing="0pt"><m:mtr><m:mtd columnalign="right"><m:mrow><m:mrow><m:mo>-</m:mo>
athvariant="normal">Δ</m:mi><m:mo></m:mo><m:msup><m:mi>u</m:mi><m:mi>N</m:mi></m:msup></m:mrow></m:mrow><m:mo>+</m:mo><m:mrow><m:mo>∇</m:mo><m:mo></m:mo><m:msup><m:mi>p</m:mi><m:mi>N</m:mi></m:
m:mrow></m:mtd><m:mtd columnalign="center"><m:mo>=</m:mo></m:mtd><m:mtd columnalign="left"><m:mrow><m:mn>0</m:mn><m:mo>,</m:mo></m:mrow></m:mtd></m:mtr><m:mtr><m:mtd columnalign="right"><m:mrow>
><m:mo></m:mo><m:msup><m:mi>u</m:mi><m:mi>N</m:mi></m:msup></m:mrow></m:mtd><m:mtd columnalign="center"><m:mo>=</m:mo></m:mtd><m:mtd columnalign="left"><m:mrow><m:mn>0</m:mn><m:mo>,</m:mo></m:mr
mtr></m:mtable><m:mi/></m:mrow><m:mo></m:mo><m:mrow><m:mtext> on </m:mtext><m:mrow><m:msup><m:mi>ℝ</m:mi><m:mn>3</m:mn></m:msup><m:mo>\</m:mo><m:mrow><m:munderover accent="true" accentunder="true
splaystyle="false"><m:mo largeop="true" mathsize="160%" stretchy="false" symmetric="true">⋃</m:mo></m:mstyle><m:mrow><m:mi mathsize="142%">i</m:mi><m:mo mathsize="142%" stretchy="false">=</m:mo>
"142%">1</m:mn></m:mrow><m:mo>N</m:mo></m:munderover><m:mo></m:mo><m:mover accent="true"><m:msub><m:mi>B</m:mi><m:mi>i</m:mi></m:msub><m:mo>¯</m:mo></m:mover></m:mrow></m:mrow><m:mtext>, </m:mte
m:mrow></m:math></Math>
  </equation>
  <p>completed with the no-slip boundary conditions:</p>
  <equation frefnum="(2)" labels="LABEL:cab_stokes" refnum="2" xml:id="S1.E2" fragid="S1.E2">
   <Math mode="display" tex="\left\{\begin{array}[]{rcl}u^{N}&amp;=&amp;V_{i}+\Omega_{i}\times(x-x_{i}),\text{ on }%&#10;\partial{B_{i}},\\&#10;\underset{|x|\to\infty}{\lim}|u^{N}(x)|&amp;=&
ay}\right." text="cases@(Array[[u ^ N, =, list@(V _ i + Omega _ i * (x - x _ i), [ on ] * partial-differential@(B _ i))], [(absolute-value@(x) to infinity)@(limit) * absolute-value@(u ^ N * x), =
="S1.E2.m1" fragid="S1.E2.m1"><m:math alttext="\left\{\begin{array}[]{rcl}u^{N}&amp;=&amp;V_{i}+\Omega_{i}\times(x-x_{i}),\text{ on }%&#10;\partial{B_{i}},\\&#10;\underset{|x|\to\infty}{\lim}|u^
mp;0,\end{array}\right." display="block"><m:mrow><m:mo>{</m:mo><m:mtable columnspacing="5pt" displaystyle="true" rowspacing="0pt"><m:mtr><m:mtd columnalign="right"><m:msup><m:mi>u</m:mi><m:mi>N<
</m:mtd><m:mtd columnalign="center"><m:mo>=</m:mo></m:mtd><m:mtd columnalign="left"><m:mrow><m:mrow><m:mrow><m:msub><m:mi>V</m:mi><m:mi>i</m:mi></m:msub><m:mo>+</m:mo><m:mrow><m:msub><m:mi mathva
>Ω</m:mi><m:mi>i</m:mi></m:msub><m:mo>×</m:mo><m:mrow><m:mo stretchy="false">(</m:mo><m:mrow><m:mi>x</m:mi><m:mo>-</m:mo><m:msub><m:mi>x</m:mi><m:mi>i</m:mi></m:msub></m:mrow><m:mo stretchy="fal
:mrow></m:mrow></m:mrow><m:mo>,</m:mo><m:mrow><m:mtext> on </m:mtext><m:mo></m:mo><m:mrow><m:mo>∂</m:mo><m:mo></m:mo><m:msub><m:mi>B</m:mi><m:mi>i</m:mi></m:msub></m:mrow></m:mrow></m:mrow><m:mo
ow></m:mtd></m:mtr><m:mtr><m:mtd columnalign="right"><m:mrow><m:munder accentunder="true"><m:mo movablelimits="false">lim</m:mo><m:mrow><m:mrow><m:mo stretchy="false">|</m:mo><m:mi>x</m:mi><m:mo
e">|</m:mo></m:mrow><m:mo>→</m:mo><m:mi mathvariant="normal">∞</m:mi></m:mrow></m:munder><m:mo></m:mo><m:mrow><m:mo stretchy="false">|</m:mo><m:mrow><m:msup><m:mi>u</m:mi><m:mi>N</m:mi></m:msup>
:mrow><m:mo stretchy="false">(</m:mo><m:mi>x</m:mi><m:mo stretchy="false">)</m:mo></m:mrow></m:mrow><m:mo stretchy="false">|</m:mo></m:mrow></m:mrow></m:mtd><m:mtd columnalign="center"><m:mo>=</
:mtd columnalign="left"><m:mrow><m:mn>0</m:mn><m:mo>,</m:mo></m:mrow></m:mtd></m:mtr></m:mtable><m:mi/></m:mrow></m:math></Math>
  </equation>
```

So far some ideas about the use of the publications-based approach for mathematical information.

Thanks for your attention!

Contact

wolfram@zbmath.org

www.fiz-karlsruhe.de

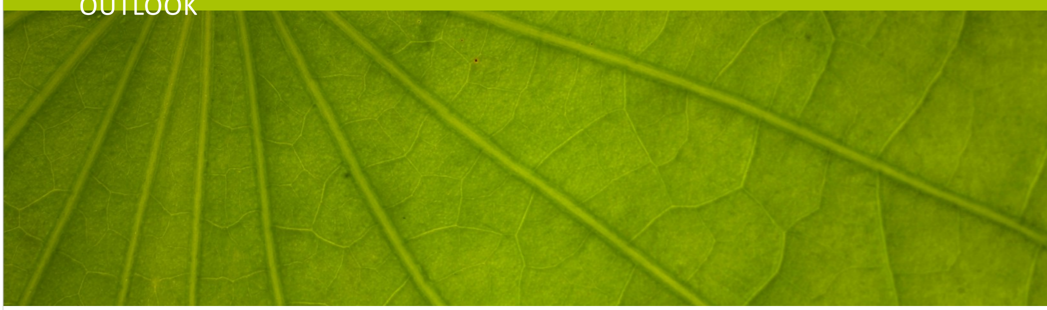
FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Leibniz
Association

The publication-based approach for mathematical models: Identification of mathematical models

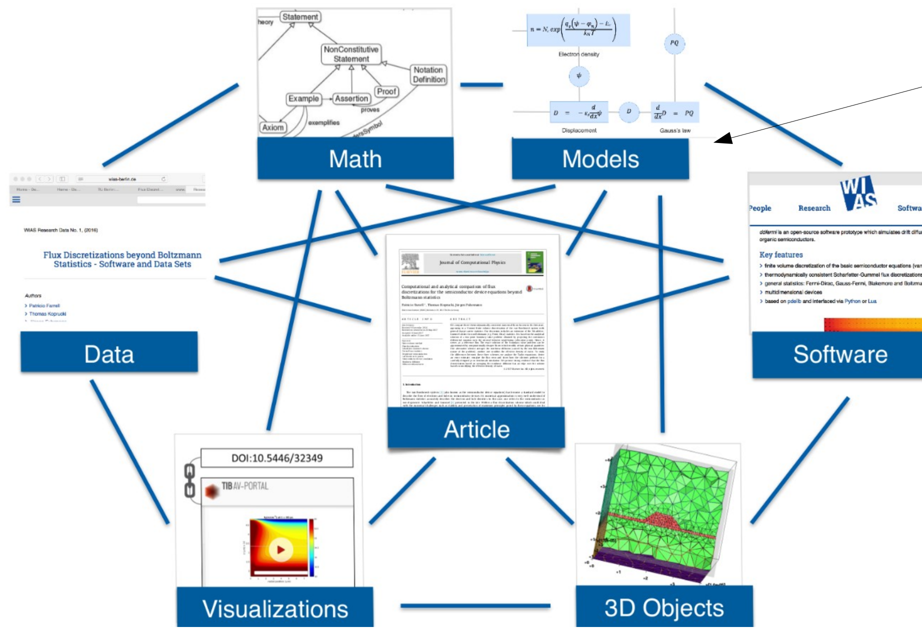
RISC, Hagenberg
2018-08-13

INTRODUCTION
AN EXAMPLE
THE PUBLICATION-BASED APPROACH
MODEL INFORMATION IN PUBLICATIONS
THE MODEL TENSOR
OUTLOOK



The central role of mathematical models

A slide
of Thomas



Application
Modeling

rlsruhe
te for Information Infras

The goal and the idea

Building up a useful service for mathematical models:

- Thomas/Michael: Models play a central role for mathematical research
- How can we build a useful service for mathematical models?

The problem

- Typically, up to now models are part of publications and not marked as single objects

Difficulties with mathematical models

``Mathematical model" is not a exactly defined term. Mathematical models can be formulated in a real-world context (an increasing number of mathematical models has a non-mathematical background) or context-free.

- Mathematical models are given by definitions, only a small number of models has also a name.
- More pragmatic:
  The definitions of mathematical models in many publications are incomplete.

Challenges for the publication-based approach

- Identification of mathematical models in publications and the analysis of the information about models i closely related:
  Definitions contain specific content information.

  Technically:
  - In general, the definitions of the mathematical model which is investigated is given in the fulltexts of publications, not in the review or abstract.
  - Mathematical definitions contain mathematical expressions (symbols, formulae) plus text. Therefore we need more than the fulltext in PDF format. At least TeX-encoded sources of the publications are necessary. arXiv seems to be a first resource.

FIZ Karlsruhe
Leibniz Institute for Information Infras

A first approach

1) Where can we find information about mathematical models?
   → analysis of the document structure

2) How can we extract and analyze information about mathematical models?
   Which information about mathematical models is possible?
   → weak description model for mathematical models

Hybrid quantum-classical modeling of quantum dot devices

Markus Kantner,[*] Markus Mittnenzweig, and Thomas Koprucki
Weierstrass Institute for Applied Analysis and Stochastics,
Mohrenstr. 39, 10117 Berlin, Germany

The structure (TOC):

Hybrid quantum-classical modeling of quantum dot devices

Markus Kantner,* Markus Mittnenzweig, and Thomas Koprucki
Weierstrass Institute for Applied Analysis and Stochastics,
Mohrenstr. 39, 10117 Berlin, Germany

For humans:
In Title: Modeling
in Abstract: modeling: `By coupling the van Roosbroeck equation with a quantum master equation in Lindblad form, ...'
in Introduction: → Model Equations in Sec. II

$$-\nabla \cdot \varepsilon \nabla \psi = q\left(p - n + C + Q(\rho)\right), \qquad (1)$$

$$\partial_t n - \frac{1}{q} \nabla \cdot \mathbf{j}_n = -R - S_n(\rho; n, p, \psi), \qquad (2)$$

$$\partial_t p + \frac{1}{q} \nabla \cdot \mathbf{j}_p = -R - S_p(\rho; n, p, \psi), \qquad (3)$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho = \mathcal{L}(\rho; n, p, \psi) = -\frac{i}{\hbar}[H, \rho] + \mathcal{D}(\rho; n, p, \psi) \qquad (4)$$

on the domain $\Omega \subset \mathbb{R}^3$. The system (1)–(4) is subject to initial conditions and boundary conditions modeling electrical contacts and other interfaces [32]. See Appendix A

→ Model Consistency III
→ Application

What is the model in the paper?

in textual form:
`hybrid quantum-classical model that self-consistently couples semi-classical transport theory to a kinetic equation for the quantum mechanical density matrix.'

as mathematical expressions (equations):

$$-\nabla \cdot \varepsilon \nabla \psi = q\left(p - n + C + Q(\rho)\right),$$

$$\partial_t n - \frac{1}{q}\nabla \cdot \mathbf{j}_n = -R - S_n(\rho; n, p, \psi),$$

$$\partial_t p + \frac{1}{q}\nabla \cdot \mathbf{j}_p = -R - S_p(\rho; n, p, \psi),$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho = \mathcal{L}(\rho; n, p, \psi) = -\frac{i}{\hbar}[H, \rho] + \mathcal{D}(\rho; n, p, \psi)$$

on the domain $\Omega \subset \mathbb{R}^3$. The system (1)–(4) is subject t
initial conditions and boundary conditions modeling ele
trical contacts and other interfaces [32]. See Appendix

FIZ Karlsruhe
Leibniz Institute for Information Infras

But

- the model is incomplete: (1)-(4) are the partial differential equations. They must be completed by initial- and boundary equations
  →appendix (This is trivial for the mathematicians.)

- the model is specified (some hypotheses for the application case):
- (1)-(4) are the basic equations
- (1)-(3) van Roosbroeck (semi-classical transport model for electrons and holes)
- (4) quantum master equation
- there are additional equations and principles which are presented in the subsection A, B, C and section III (evaluation of modeling: correctness accordance with the principles of thermodynamics).

- *Remark: This papers describes the modeling process (the model is the result of this paper), not the mathematical treatment of the model.*

A first summary

- The model is given by text and formulae.
  For the mathematical treatment the formulae are the most important part. Often the textual environment gives some information about the meaning of the expression and the background of the model (origin / use cases) but can contain also structural information as in the example (van Roosbroeck, quantum master equatiion in Lindblad form, links to scientific theories, algorithms, software, ...)
- Both the textual information and the mathematical expressions which describe a model are relevant.
- Therefore we should try to extract both
  - the mathematical expressions describing the model
  - the corresponding terms and phrases for the mathematical expressions

- *Remark: Terms, phrases and mathematical expressions are ambigious, but mathematical expressions, especially complex objects as equations, inequalities, structural diagrams contain precise information about their structure which allow to compare these objects and enhanced retrieval features.*

FIZ Karlsruhe
Leibniz Institute for Information Infras

What can we do automatically?

- **The publication-based approach for mathematical models**

  Up to now, models are part of publications. →Therefore we have to develop methods to extract the
- mathematical expressions describing the mathematical model and the corresponding terms and
- phrases.

Remarks:

- *Fulltexts:*
  The most abstracts or reviews, the zbMATH data, don't contain precise model descriptions. Typically, the mathematical expressions are not part of the abstract or reviews. Therefore, we need the fulltexts of the models.
- *Formats:*
  Up to now, no powerful converter exists for transforming PDF formatted documents to TeX or MathML. At least, TeX encoded documents are needed. Therefore the arXiv e-prints are a first resource to develop methods for model analysis.

The publication-based approach

The publication-based approach covers three steps:

1) Identification of the model-relevant parts of the publication
   - Model relevant sections or subsections
     which can be identified by characteristic terms in the titles, especially `model*' or `problem*'
   alternatively
   - the beginning of a document, e.g., the first 3,000 characters

FIZ Karlsruhe
Leibniz Institute for Information Infras

A statistic

Journal	(Sub-)sections containg the terms `model‘ or `problem‘
International Journal of Robust and Nonlinear Control, v. 10, n. 28	8 (9)
International Journal of Robust and Nonlinear Control v. 10, n. 29	6 (8)
Journal of System Science & Complexity v. 31, n. 3	7 (13)
Cybernetics and System Analysis v. 54, n. 2	8 (18)
International Journal of Control and Signal Processing v. 32, n. 6	7 (10)
Automation and Remote Control, v. 79, n. 5	6 (12)

FIZ Karlsruhe
Leibniz Institute for Information Infras

Identification of model-relevant parts: results and open questions

by eliminating the non-relevant parts we get reduced (TeX-encoded) artifacts which contain only the model-relevant parts

- *Remark: It is not clear that the model-relevant parts contain only one model information.*
  The reduced document can also contain previous models, similar models, specific models, generalized models, transformed models, etc.). In other words, it may contain a lot of additional information.
  How can we eliminate redundant / irrelevant information?
  (If yes, we could markup the model in the original file by a `model tag')

  Our example: The section II contains ~ 100 mathematical expressions
  9 / A:22 /B:29 / C:18 /fig.2:(15) text:10) (moreover, the mathematical expression from appendix A must be added).
  The frequencies of some expressions is > 1.

Analysis of the model-reduced documents

2) Analysis of our artifacts:

1) extraction of all mathematical expressions (TeX or MathML encoded expressions) and the corresponding terms and phrases

→ creating a `model-related' matrix A_{i3} (1 <= i <= n, n – number of mathematical expression in the model-reduced document) for a document with the components :

a_{i1}: mathematical expression

a_{i2}: left-hand textual neighborhood (k terms)

a_{i3}: right-hand textual neighborhood (k terms)

Example:	C	doping profile	0
	Q(p)	expectation value of the charge density	0
	n	free electron density	0
	p	free hole density	0
	(1)-(3)	0	van Roosbroeck system

Analysis of the model-reduced documents: results and open questions

storing in a database (A_{i3}^{k} with k as the number of k-th document)

Remarks: As result we get a tensor which contains mathematical expressions plus their neighborhood. The mathematical expressions are listed in the sequence of their occurrence. This can make difficulties, e.g., parts of the model are presented at different positions in the document.
Can we identify the input variables of the model?
How can we organize it in a better way (sorting by length is not a real good idea)?
How can we identify the system structure of a model?
How can we find relevant textual information for a mathematical expression which is not in the left- or in the right-hand neighborhood of a mathematical expression?
What do we do with this information (e.g, storing this information in a special field of the tensor)?

FIZ Karlsruhe
Leibniz Institute for Information Infras

Further processing: Transformation of the TeX-encoded mathematical express

3) It is difficult to compare TeX-encoded mathematical expressions
(the TeX codes for a PDF document differ; zbMATH experience: mathematical expressions with more than 6 characters occur in the database zbMATH only once)
better (more standardized): XML / MathML encoded mathematical expression
(tree structured documents)
MathML transformation of mathematical expressions via **LaTeXML**
→ tree-structured documents which allow to compare the structure of mathematical formulae

Remarks: LaTeXML transform the TeX code to XML. By postprocessing, the XML code can be transformed to M
(Presentation Markup, Content Markup, or OpenMATH)
Which format should be used for storing, all or only MathML Content Markup?
(If the authors would write their documents in STeX, MathML or OpenMATH should be used).

Example (the PDE system 1 - 4)

```
\begin{align}
-\nabla\cdot\varepsilon\nabla\psi & =q\left(p-n+C+Q\left(\rho\right)\right),
\label{eq: Poisson equation}\\
\partial_{t}n-\frac{1}{q}\nabla\cdot\mathbf{j}_{n} & =-R-S_{n}\left(\rho;n,p,\psi\right),\label{eq:
electron transport}\\
\partial_{t}p+\frac{1}{q}\nabla\cdot\mathbf{j}_{p} & =-R-S_{p}\left(\rho;n,p,\psi\right),\label{eq: hole
transport}\\
\frac{\mathrm{d}}{\mathrm{d}t}\rho = \mathcal{L}\left(\rho;n,p,\psi\right)  &=-\frac{i}
{\hbar}\left[H,\rho\right]+\mathcal{D}\left(\rho;n,p,\psi\right)\label{eq: quantum master equation}
\end{align}
```

The Online Editor of LaTeXML

```xml
    </bibref></cite>. The model equations read</p>
  <equationgroup class="ltx_eqn_align" xml:id="A6.EGx1">
    <equation frefnum="(1)" labels="LABEL:eq:_Poisson_equation" refnum="1" xml:id="S2.E1">
      <MathFork>
        <Math tex="\displaystyle-\nabla\cdot\varepsilon\nabla\psi=q\left(p-n+C+Q\left(\rho\right)%&#10;\right)," text="- (nabla cdot varepsilon) * nabla@(psi) = q * ((p - n) + C + Q * rho)" xml:id="S2.E1.m3">
          <XMath>
            <XMDual>
              <XMRef idref="S2.E1.m3.2"/>
              <XMWrap>
                <XMApp xml:id="S2.E1.m3.2">
                  <XMTok meaning="equals" role="RELOP">=</XMTok>
                  <XMApp>
                    <XMTok meaning="minus" role="ADDOP">-</XMTok>
                    <XMApp>
                      <XMTok meaning="times" role="MULOP"></XMTok>
                      <XMApp>
                        <XMTok name="cdot" role="MULOP">·</XMTok>
                        <XMTok name="nabla" role="OPERATOR">∇</XMTok>
                        <XMTok font="italic" name="varepsilon" role="UNKNOWN">ε</XMTok>
                      </XMApp>
                      <XMApp>
                        <XMTok name="nabla" role="OPERATOR">∇</XMTok>
                        <XMTok font="italic" name="psi" role="UNKNOWN">ψ</XMTok>
                      </XMApp>
                    </XMApp>
                  </XMApp>
                  <XMApp>
                    <XMTok meaning="times" role="MULOP"></XMTok>
                    <XMTok font="italic" role="UNKNOWN">q</XMTok>
                    <XMDual>
                      <XMRef idref="S2.E1.m3.2.1"/>
                      <XMWrap>
                        <XMTok role="OPEN" stretchy="true">(</XMTok>
                        <XMApp xml:id="S2.E1.m3.2.1">
                          <XMTok meaning="plus" role="ADDOP">+</XMTok>
                          <XMApp>
                            <XMTok meaning="minus" role="ADDOP">-</XMTok>
                            <XMTok font="italic" role="UNKNOWN">p</XMTok>
                            <XMTok font="italic" role="UNKNOWN">n</XMTok>
                          </XMApp>
                          <XMTok font="italic" role="UNKNOWN">C</XMTok>
                          <XMApp>
                            <XMTok meaning="times" role="MULOP"></XMTok>
                            <XMTok font="italic" role="UNKNOWN">Q</XMTok>
                            <XMDual>
                              <XMRef idref="S2.E1.m3.1"/>
                              <XMWrap>
                                <XMTok role="OPEN" stretchy="true">(</XMTok>
                                <XMTok font="italic" name="rho" role="UNKNOWN" xml:id="S2.E1.m3.1">ρ</XMTok>
                                <XMTok role="CLOSE" stretchy="true">)</XMTok>
                              </XMWrap>
                            </XMDual>
                          </XMApp>
                        </XMApp>
                        <XMTok role="CLOSE" stretchy="true">)</XMTok>
                      </XMWrap>
                    </XMDual>
```

```
ref>)</cite>. The model equations read</p>
      <equationgroup class="ltx_eqn_align" xml:id="A6.EGx1" fragid="A6.EGx1">
        <equation frefnum="(1)" labels="LABEL:eq:_Poisson_equation" refnum="1" xml:id="S2.E1" fragid="S2.E1">
          <MathFork>
            <Math tex="\displaystyle-\nabla\cdot\varepsilon\nabla\psi=q\left(p-n+C+Q\left(\rho\right)%&#10;\right)," text="- (nabla cdot varepsilon) * nabla@(psi) = q * ((p - n) + C + Q * rho)" xml:id="S2.E1.m3"
fragid="S2.E1.m3"><m:math alttext="\displaystyle-\nabla\cdot\varepsilon\nabla\psi=q\left(p-n+C+Q\left(\rho\right)%&#10;\right)," display="inline"><m:mrow><m:mrow><m:mrow><m:mo>-</m:mo><m:mrow><m:mrow><m:mo>∇</
m:mo><m:mo>·</m:mo><m:mi>ε</m:mi></m:mrow><m:mo></m:mo><m:mrow><m:mo>∇</m:mo><m:mo></m:mo><m:mi>ψ</m:mi></m:mrow></m:mrow></m:mrow><m:mo>=</m:mo><m:mrow><m:mi>q</m:mi><m:mo></m:mo><m:mrow><m:mo>(</
m:mo><m:mrow><m:mrow><m:mi>p</m:mi><m:mo>-</m:mo><m:mi>n</m:mi></m:mrow><m:mo>+</m:mo><m:mi>C</m:mi><m:mo>+</m:mo><m:mrow><m:mi>Q</m:mi><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>)</m:mo></m:mrow></
m:mrow></m:mrow><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math>
            <MathBranch>
              <td align="right"><Math mode="inline" tex="\displaystyle-\nabla\cdot\varepsilon\nabla\psi" text="- (nabla cdot varepsilon) * nabla@(psi)" xml:id="S2.E1.m1" fragid="S2.E1.m1"><m:math
alttext="\displaystyle-\nabla\cdot\varepsilon\nabla\psi" display="inline"><m:mrow><m:mo>-</m:mo><m:mrow><m:mrow><m:mo>∇</m:mo><m:mo>·</m:mo><m:mi>ε</m:mi></m:mrow><m:mo></m:mo><m:mrow><m:mo>∇</m:mo><m:mo></
m:mo><m:mi>ψ</m:mi></m:mrow></m:mrow></m:mrow></m:math></Math></td>
              <td align="left"><Math mode="inline" tex="\displaystyle=q\left(p-n+C+Q\left(\rho\right)\right)," text="absent = q * ((p - n) + C + Q * rho)" xml:id="S2.E1.m2" fragid="S2.E1.m2"><m:math
alttext="\displaystyle=q\left(p-n+C+Q\left(\rho\right)\right)," display="inline"><m:mrow><m:mrow><m:mi/><m:mo>=</m:mo><m:mrow><m:mi>q</m:mi><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mrow><m:mrow><m:mi>p</m:mi><m:mo>-
</m:mo><m:mi>n</m:mi></m:mrow><m:mo>+</m:mo><m:mi>C</m:mi><m:mo>+</m:mo><m:mrow><m:mi>Q</m:mi><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>)</m:mo></m:mrow></m:mrow><m:mo>)</m:mo></m:mrow></
m:mrow></m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math></td>
            </MathBranch>
          </MathFork>
        </equation>
        <equation frefnum="(2)" labels="LABEL:eq:_electron_transport" refnum="2" xml:id="S2.E2" fragid="S2.E2">
          <MathFork>
            <Math tex="\displaystyle\partial_{t}n-\frac{1}{q}\nabla\cdot\mathbf{j}_{n}=-R-S_{n}\left(%&#10;\rho;n,p,\psi\right)," text="(partial-differential _ t)@(n) - ((1 / q) * nabla) cdot j _ n = - R - S _ n
* list@(rho, n, p, psi)" xml:id="S2.E2.m3" fragid="S2.E2.m3"><m:math alttext="\displaystyle\partial_{t}n-\frac{1}{q}\nabla\cdot\mathbf{j}_{n}=-R-S_{n}\left(%&#10;\rho;n,p,\psi\right),"
display="inline"><m:mrow><m:mrow><m:mrow><m:mrow><m:msub><m:mo>∂</m:mo><m:mi>t</m:mi></m:msub><m:mo></m:mo><m:mi>n</m:mi></m:mrow><m:mo>-</m:mo><m:mrow><m:mrow><m:mstyle displaystyle="true"><m:mfrac><m:mn>1</
m:mn><m:mi>q</m:mi></m:mfrac></m:mstyle><m:mo></m:mo><m:mo>∇</m:mo></m:mrow><m:mo>·</m:mo><m:msub><m:mi>j</m:mi><m:mi>n</m:mi></m:msub></m:mrow></m:mrow><m:mo>=</m:mo><m:mrow><m:mrow><m:mo>-</m:mo><m:mi>R</m:mi></
m:mrow><m:mo>-</m:mo><m:mrow><m:msub><m:mi>S</m:mi><m:mi>n</m:mi></m:msub><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>;</m:mo><m:mi>n</m:mi><m:mo>,</m:mo><m:mi>p</m:mi><m:mo>,</m:mo><m:mi>ψ</
m:mi><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow></m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math>
            <MathBranch>
              <td align="right"><Math mode="inline" tex="\displaystyle\partial_{t}n-\frac{1}{q}\nabla\cdot\mathbf{j}_{n}" text="(partial-differential _ t)@(n) - ((1 / q) * nabla) cdot j _ n" xml:id="S2.E2.m1"
fragid="S2.E2.m1"><m:math alttext="\displaystyle\partial_{t}n-\frac{1}{q}\nabla\cdot\mathbf{j}_{n}" display="inline"><m:mrow><m:mrow><m:msub><m:mo>∂</m:mo><m:mi>t</m:mi></m:msub><m:mo></m:mo><m:mi>n</m:mi></
m:mrow><m:mo>-</m:mo><m:mrow><m:mrow><m:mstyle displaystyle="true"><m:mfrac><m:mn>1</m:mn><m:mi>q</m:mi></m:mfrac></m:mstyle><m:mo></m:mo><m:mo>∇</m:mo></m:mrow><m:mo>·</m:mo><m:msub><m:mi>j</m:mi><m:mi>n</m:mi></
m:msub></m:mrow></m:mrow></m:math></Math></td>
              <td align="left"><Math mode="inline" tex="\displaystyle=-R-S_{n}\left(\rho;n,p,\psi\right)," text="absent = - R - S _ n * list@(rho, n, p, psi)" xml:id="S2.E2.m2" fragid="S2.E2.m2"><m:math
alttext="\displaystyle=-R-S_{n}\left(\rho;n,p,\psi\right)," display="inline"><m:mrow><m:mrow><m:mi/><m:mo>=</m:mo><m:mrow><m:mrow><m:mo>-</m:mo><m:mi>R</m:mi></m:mrow><m:mo>-</m:mo><m:mrow><m:msub><m:mi>S</
m:mi><m:mi>n</m:mi></m:msub><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>;</m:mo><m:mi>n</m:mi><m:mo>,</m:mo><m:mi>p</m:mi><m:mo>,</m:mo><m:mi>ψ</m:mi><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow></
m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math></td>
            </MathBranch>
          </MathFork>
        </equation>
        <equation frefnum="(3)" labels="LABEL:eq:_hole_transport" refnum="3" xml:id="S2.E3" fragid="S2.E3">
          <MathFork>
            <Math tex="\displaystyle\partial_{t}p+\frac{1}{q}\nabla\cdot\mathbf{j}_{p}=-R-S_{p}\left(%&#10;\rho;n,p,\psi\right)," text="(partial-differential _ t)@(p) + ((1 / q) * nabla) cdot j _ p = - R - S _ p
* list@(rho, n, p, psi)" xml:id="S2.E3.m3" fragid="S2.E3.m3"><m:math alttext="\displaystyle\partial_{t}p+\frac{1}{q}\nabla\cdot\mathbf{j}_{p}=-R-S_{p}\left(%&#10;\rho;n,p,\psi\right),"
display="inline"><m:mrow><m:mrow><m:mrow><m:mrow><m:msub><m:mo>∂</m:mo><m:mi>t</m:mi></m:msub><m:mo></m:mo><m:mi>p</m:mi></m:mrow><m:mo>+</m:mo><m:mrow><m:mrow><m:mstyle displaystyle="true"><m:mfrac><m:mn>1</
m:mn><m:mi>q</m:mi></m:mfrac></m:mstyle><m:mo></m:mo><m:mo>∇</m:mo></m:mrow><m:mo>·</m:mo><m:msub><m:mi>j</m:mi><m:mi>p</m:mi></m:msub></m:mrow></m:mrow><m:mo>=</m:mo><m:mrow><m:mrow><m:mo>-</m:mo><m:mi>R</m:mi></
m:mrow><m:mo>-</m:mo><m:mrow><m:msub><m:mi>S</m:mi><m:mi>p</m:mi></m:msub><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>;</m:mo><m:mi>n</m:mi><m:mo>,</m:mo><m:mi>p</m:mi><m:mo>,</m:mo><m:mi>ψ</
m:mi><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow></m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math>
            <MathBranch>
              <td align="right"><Math mode="inline" tex="\displaystyle\partial_{t}p+\frac{1}{q}\nabla\cdot\mathbf{j}_{p}" text="(partial-differential _ t)@(p) + ((1 / q) * nabla) cdot j _ p" xml:id="S2.E3.m1"
fragid="S2.E3.m1"><m:math alttext="\displaystyle\partial_{t}p+\frac{1}{q}\nabla\cdot\mathbf{j}_{p}" display="inline"><m:mrow><m:mrow><m:msub><m:mo>∂</m:mo><m:mi>t</m:mi></m:msub><m:mo></m:mo><m:mi>p</m:mi></
m:mrow><m:mo>+</m:mo><m:mrow><m:mrow><m:mstyle displaystyle="true"><m:mfrac><m:mn>1</m:mn><m:mi>q</m:mi></m:mfrac></m:mstyle><m:mo></m:mo><m:mo>∇</m:mo></m:mrow><m:mo>·</m:mo><m:msub><m:mi>j</m:mi><m:mi>p</m:mi></
m:msub></m:mrow></m:mrow></m:math></Math></td>
              <td align="left"><Math mode="inline" tex="\displaystyle=-R-S_{p}\left(\rho;n,p,\psi\right)," text="absent = - R - S _ p * list@(rho, n, p, psi)" xml:id="S2.E3.m2" fragid="S2.E3.m2"><m:math
alttext="\displaystyle=-R-S_{p}\left(\rho;n,p,\psi\right)," display="inline"><m:mrow><m:mrow><m:mi/><m:mo>=</m:mo><m:mrow><m:mrow><m:mo>-</m:mo><m:mi>R</m:mi></m:mrow><m:mo>-</m:mo><m:mrow><m:msub><m:mi>S</
m:mi><m:mi>p</m:mi></m:msub><m:mo></m:mo><m:mrow><m:mo>(</m:mo><m:mi>ρ</m:mi><m:mo>;</m:mo><m:mi>n</m:mi><m:mo>,</m:mo><m:mi>p</m:mi><m:mo>,</m:mo><m:mi>ψ</m:mi><m:mo>)</m:mo></m:mrow></m:mrow></m:mrow></
m:mrow><m:mo>,</m:mo></m:mrow></m:math></Math></td>
            </MathBranch>
```


XML ▾ Tab Width: 8 ▾ Ln 142, Col 253 ▾ INS

Some remarks and questions

All mathematical expressions of our model description tensor could be transformed on this way to MathML PM or CM or OpenMATH. This allows to use graph-based models, e.g. for comparing models or retrieval.

➢

The advantage of the approach:
All steps of the publication-based approach can be done widely automatically.
But the method has also some disadvantages:
In principle our model tensors of publications should contain the model information, but it is possible that the information is too much. How can we remove model-irrelevant information?
Big tensors describing a model are not very impressive for the user. How can we compactify the model representations

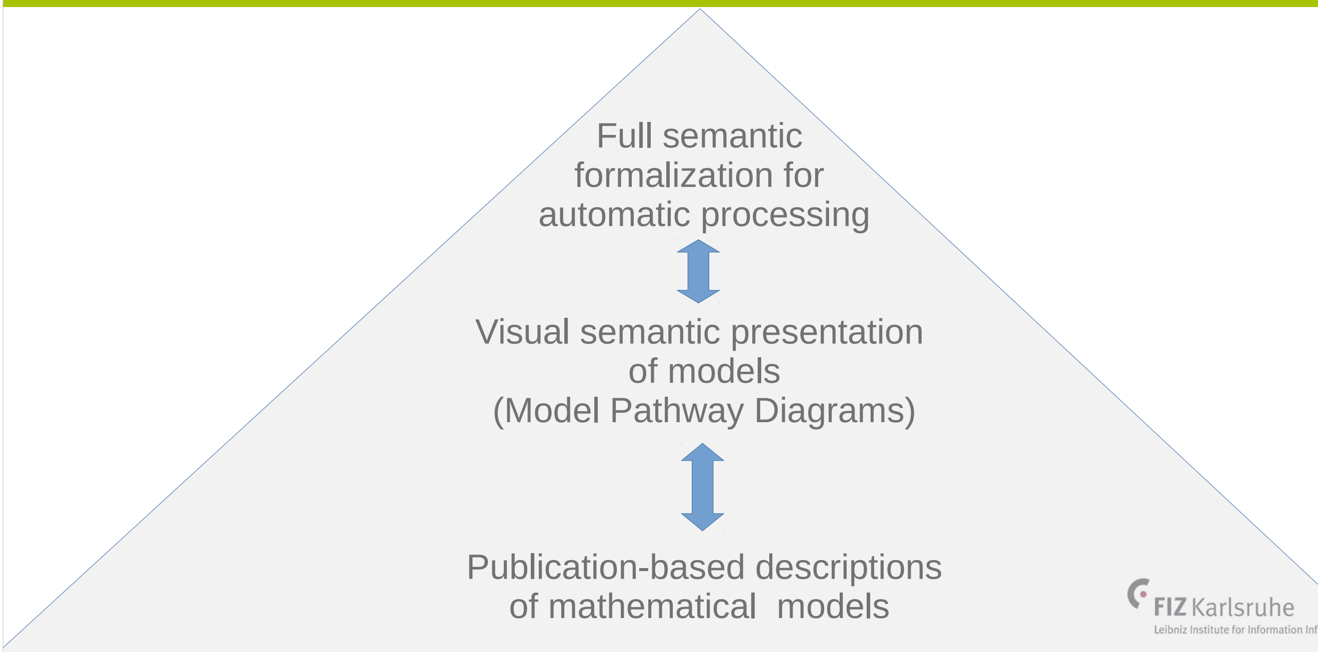
(A first idea:
A grouping of complex mathematical expressions with the involved notations should be possible. Often long mathematical expression are complex expressions (especially PDEs) containing special symbols which we could connect).

FIZ Karlsruhe
Leibniz Institute for Information Infras

Further questions

- What is with retrieval (access for mathematicans and for non-mathematians)?
- How can we detect similar models (ranking of the mathematical expressions of the models)?
- How can we make the service interactive (for improving the model, linking to further model information)?
- How can we make the modeling process understandable?

Our ambitioned goal: Building up a knowledge base for mathematical models



FIZ Karlsruhe
Leibniz Institute for Information Infras

Example 2:
The model
description
in the e-print
arXiv:1806.07795

1. Introduction

In this paper, we consider a system of N spherical particles $(B_i)_{1\leq i\leq N}$ with identical radii R immersed in a viscous fluid satisfying the following Stokes equation:

$$\left\{\begin{array}{rcl} -\Delta u^N + \nabla p^N & = & 0, \\ \operatorname{div} u^N & = & 0, \end{array}\right. \text{ on } \mathbb{R}^3 \setminus \bigcup_{i=1}^{N} \overline{B_i}, \tag{1}$$

completed with the no-slip boundary conditions:

$$\left\{\begin{array}{rcl} u^N & = & V_i + \Omega_i \times (x - x_i), \text{ on } \partial B_i, \\ \lim\limits_{|x|\to\infty} |u^N(x)| & = & 0, \end{array}\right. \tag{2}$$

where $(V_i, \Omega_i) \in \mathbb{R}^3 \times \mathbb{R}^3$, $1 \leq i \leq N$ represent the linear and angular velocities,

$$B_i := B(x_i, R).$$

We describe the intertialess motion of the rigid spheres $(B_i)_{1\leq i\leq N}$ by adding to the instantaneous Stokes equation the classical Newton dynamics for the particles $(x_i)_{1\leq i\leq N}$:

$$\left\{\begin{array}{rcl} \dot{x}_i & = & V_i, \\ F_i + mg & = & 0, \\ T_i & = & 0, \end{array}\right. \tag{3}$$

where m denotes the mass of the identical particles adjusted for buoyancy, g the gravitational acceleration, F_i (resp. T_i) the drag force (resp. the torque) applied by the the fluid

on the i^{th} particle B_i defined as

$$\begin{array}{rcl} F_i & := & \displaystyle\int_{\partial B_i} \sigma(u^N, p^N) n, \\ T_i & = & \displaystyle\int_{\partial B_i} (x - x_i) \times [\sigma(u^N, p^N) n], \end{array}$$

with n the unit outer normal to ∂B_i and $\sigma(u^N, p^N) = 2D(u^N) - p^N \mathbb{I}$, the stress tensor where $2D(u^N) = \nabla u^N + \nabla {u^N}^{\top}$.

The publicly available TeX-code

```latex
\section{Introduction}
In this paper, we consider a system of $N$ spherical particles $(B_i)_{1\leq i \leq N}$ with
identical radii $R$ immersed in a viscous fluid satisfying the following Stokes equation:
\begin{equation}\label{eq_stokes}
\left \{
\begin{array}{rcl}
-\Delta u^N + \nabla p^N & = &0, \\
\div u^N & = & 0,
\end{array}
\text { on $\mathbb{R}^3 \setminus \underset{i=1}{\overset{N}{\bigcup}}\overline{B_i}$, }
\right.
\end{equation}
completed with the no-slip boundary conditions:
\begin{equation}\label{cab_stokes}
\left \{
\begin{array}{rcl}
u^N & = &  V_i +\Omega_i \times (x-x_i), \text{ on } \partial {B_i}, \\
\underset{|x| \to \infty}{\lim} |u^N(x)| & = & 0,
\end{array}\right.
\end{equation}
where $(V_i,\Omega_i) \in \mathbb{R}^3\times \mathbb{R}^3 \,,\, 1\leq i \leq N$ represent the
linear and angular velocities,
\begin{eqnarray*}
B_i:= B(x_i,R ).
\end{eqnarray*}
We describe the intertialess motion of the rigid spheres $(B_i)_{1\leq i \leq N}$ by adding to
the instantaneous Stokes equation the classical Newton dynamics for the particles $
(x_i)_{1\leq i\leq N}$:
\begin{equation}\label{ode}
\left \{
\begin{array}{rcl}
\dot{x}_i &= & V_i,\\
F_i+ mg &=& 0,\\
T_i&=& 0,
\end{array}
\right.
\end{equation}
where $m$ denotes the mass of the identical particles adjusted for buoyancy, $g$ the
gravitational acceleration, $F_i$ (resp. $T_i$) the drag force (resp. the torque) applied by
the the fluid on the $i^{th}$ particle $B_i$ defined as
\begin{eqnarray*}
F_i & := & \int_{\partial B_i} \sigma(u^N,p^N) n, \\
T_i & = &  \int_{\partial B_i} (x-x_i) \times [\sigma(u^N,p^N) n],
\end{eqnarray*}
with $n$ the unit outer normal to $\partial B_i$ and
$\sigma(u^N,p^N)= 2 D(u^N) - p^N \mathbb{I}, $ the stress tensor where $2D(u^N)=\nabla u^N +{
\nabla u^N }^{\top} $.\\
```

The XML snippet of the Stokes equation (via application of LaTeXML)


```
    </Math> immersed in a viscous fluid satisfying the following Stokes equation:
<!--  %**** mecherbet sedimentation.tex Line 50 **** --></p>
      <equation frefnum="(1)" labels="LABEL:eq_stokes" refnum="1" xml:id="S1.E1">
        <Math mode="display" tex="\left\{\begin{array}[]{rcl}-\Delta u^{N}+\nabla p^{N}&amp;=&amp;0,\\&#10;\operatorname{div}u^{N}&amp;=&amp;0,\end{array}\text{ on $\mathb
\overset{N}{\bigcup}}\overline{B_{i}}$, }\right." text="cases@(Array[[(- Delta * u ^ N) + nabla@(p ^ N), =, 0], [div@(u ^ N), =, 0]]) * [ on \R3=i1N⋃¯Bi, ]" xml:id="S1.E1.
          <XMath>
            <XMApp>
              <XMTok meaning="times" role="MULOP"></XMTok>
              <XMDual>
                <XMApp>
                  <XMTok meaning="cases"/>
                  <XMRef idref="S1.E1.m1.3"/>
                </XMApp>
                <XMWrap>
                  <XMTok role="OPEN" stretchy="true">{</XMTok>
                  <XMArray role="ARRAY" vattach="middle" xml:id="S1.E1.m1.3">
                    <XMRow>
                      <XMCell align="right">
                        <XMApp>
                          <XMTok meaning="plus" role="ADDOP">+</XMTok>
                          <XMApp>
                            <XMTok meaning="minus" role="ADDOP">-</XMTok>
                            <XMApp>
                              <XMTok meaning="times" role="MULOP"></XMTok>
                              <XMTok name="Delta" role="UNKNOWN">Δ</XMTok>
                              <XMApp>
                                <XMTok role="SUPERSCRIPTOP" scriptpos="post8"/>
                                <XMTok font="italic" role="UNKNOWN">u</XMTok>
                                <XMTok font="italic" fontsize="70%" role="UNKNOWN">N</XMTok>
                              </XMApp>
                            </XMApp>
                          </XMApp>
```


FIZ Karlsruhe
Leibniz Institute for Information Infras

A MathML (PM) snippet of the model (via LaTeXML)

```xml
    <title><tag close=". ">1</tag>Introduction</title>
    <toctitle><tag close=" ">1</tag>Introduction</toctitle>
    <para xml:id="S1.p1" fragid="S1.p1">
      <p>In this paper, we consider a system of <Math mode="inline" tex="N" text="N" xml:id="S1.p1.m1" fragid="S1.p1.m1"><m:math alttext="N" display="inline"><m:mi>N</m:mi></m:math></
<Math mode="inline" tex="(B_{i})_{1\leq i\leq N}" text="B _ i _ (1 less= i less= N)" xml:id="S1.p1.m2" fragid="S1.p1.m2"><m:math alttext="(B_{i})_{1\leq i\leq N}" display="inline"><m:
y="false">(</m:mo><m:msub><m:mi>B</m:mi><m:mi>i</m:mi></m:msub><m:mo stretchy="false">)</m:mo></m:mrow><m:mrow><m:mn>1</m:mn><m:mo>≤</m:mo><m:mi>i</m:mi><m:mo>≤</m:mo><m:mi>N</m:mi></
/Math> with identical radii <Math mode="inline" tex="R" text="R" xml:id="S1.p1.m3" fragid="S1.p1.m3"><m:math alttext="R" display="inline"><m:mi>R</m:mi></m:math></Math> immersed in a
he following Stokes equation:
<!--  %**** mecherbet sedimentation.tex Line 50 **** --></p>
      <equation frefnum="(1)" labels="LABEL:eq_stokes" refnum="1" xml:id="S1.E1" fragid="S1.E1">
        <Math mode="display" tex="\left\{\begin{array}[]{rcl}-\Delta u^{N}+\nabla p^{N}&amp;=&amp;0,\\&#10;\operatorname{div}u^{N}&amp;=&amp;0,\end{array}\text{ on $\mathbb{R}^{3}\set
\overset{N}{\bigcup}}\overline{B_{i}}$, }\right." text="cases@(Array[[(- Delta * u ^ N) + nabla@(p ^ N), =, 0], [div@(u ^ N), =, 0]]) * [ on \R3=i1N⋃¯Bi, ]" xml:id="S1.E1.m1" fragid="S
"\left\{\begin{array}[]{rcl}-\Delta u^{N}+\nabla p^{N}&amp;=&amp;0,\\&#10;\operatorname{div}u^{N}&amp;=&amp;0,\end{array}\text{ on $\mathbb{R}^{3}\setminus%&#10;\underset{i=1}{\overse
{i}}$, }\right." display="block"><m:mrow><m:mrow><m:mo>{</m:mo><m:mtable columnspacing="5pt" displaystyle="true" rowspacing="0pt"><m:mtr><m:mtd columnalign="right"><m:mrow><m:mrow><m:
athvariant="normal">Δ</m:mi><m:mo></m:mo><m:msup><m:mi>u</m:mi><m:mi>N</m:mi></m:msup></m:mrow></m:mrow><m:mo>+</m:mo><m:mrow><m:mo>∇</m:mo><m:mo></m:mo><m:msup><m:mi>p</m:mi><m:mi>N<
m:mrow></m:mtd><m:mtd columnalign="center"><m:mo>=</m:mo></m:mtd><m:mtd columnalign="left"><m:mrow><m:mn>0</m:mn><m:mo>,</m:mo></m:mrow></m:mtd></m:mtr><m:mtr><m:mtd columnalign="righ
><m:mo></m:mo><m:msup><m:mi>u</m:mi><m:mi>N</m:mi></m:msup></m:mrow></m:mtd><m:mtd columnalign="center"><m:mo>=</m:mo></m:mtd><m:mtd columnalign="left"><m:mrow><m:mn>0</m:mn><m:mo>,</
mtr></m:mtable><m:mi/></m:mrow><m:mo></m:mo><m:mrow><m:mtext>_on_</m:mtext><m:mrow><m:msup><m:mi>ℝ</m:mi><m:mn>3</m:mn></m:msup><m:mo>\</m:mo><m:mrow><m:munderover accent="true" accen
splaystyle="false"><m:mo largeop="true" mathsize="160%" stretchy="false" symmetric="true">⋃</m:mo></m:mstyle><m:mrow><m:mi mathsize="142%">i</m:mi><m:mo mathsize="142%" stretchy="fals
"142%">1</m:mn></m:mrow><m:mo>𝑁</m:mo></m:munderover><m:mo></m:mo><m:mover accent="true"><m:msub><m:mi>B</m:mi><m:mi>i</m:mi></m:msub><m:mo>¯</m:mo></m:mover></m:mrow></m:mrow><m:mte
m:mrow></m:math></Math>
      </equation>
      <p>completed with the no-slip boundary conditions:</p>
      <equation frefnum="(2)" labels="LABEL:cab_stokes" refnum="2" xml:id="S1.E2" fragid="S1.E2">
        <Math mode="display" tex="\left\{\begin{array}[]{rcl}u^{N}&amp;=&amp;V_{i}+\Omega_{i}\times(x-x_{i}),\text{ on }%&#10;\partial{B_{i}},\\&#10;\underset{|x|\to\infty}{\lim}|u^{N
ay}\right." text="cases@(Array[[u ^ N, =, list@(V _ i + Omega _ i * (x - x _ i), [ on ] * partial-differential@(B _ i))], [(absolute-value@(x) to infinity)@(limit) * absolute-value@(u
="S1.E2.m1" fragid="S1.E2.m1"><m:math alttext="\left\{\begin{array}[]{rcl}u^{N}&amp;=&amp;V_{i}+\Omega_{i}\times(x-x_{i}),\text{ on }%&#10;\partial{B_{i}},\\&#10;\underset{|x|\to\inft
mp;0,\end{array}\right." display="block"><m:mrow><m:mo>{</m:mo><m:mtable columnspacing="5pt" displaystyle="true" rowspacing="0pt"><m:mtr><m:mtd columnalign="right"><m:msup><m:mi>u</m:
</m:mtd><m:mtd columnalign="center"><m:mo>=</m:mo></m:mtd><m:mtd columnalign="left"><m:mrow><m:mrow><m:mrow><m:msub><m:mi>V</m:mi><m:mi>i</m:mi></m:msub><m:mo>+</m:mo><m:mrow><m:msub>
>Ω</m:mi><m:mi>i</m:mi></m:msub><m:mo>×</m:mo><m:mrow><m:mo stretchy="false">(</m:mo><m:mrow><m:mi>x</m:mi><m:mo>-</m:mo><m:msub><m:mi>x</m:mi><m:mi>i</m:mi></m:msub></m:mrow><m:mo st
:mrow></m:mrow></m:mrow><m:mo>,</m:mo><m:mrow><m:mtext>_on_</m:mtext><m:mo></m:mo><m:mrow><m:mo>∂</m:mo><m:mo></m:mo><m:msub><m:mi>B</m:mi><m:mi>i</m:mi></m:msub></m:mrow></m:mrow></m
ow></m:mtd></m:mtr><m:mtr><m:mtd columnalign="right"><m:mrow><m:munder accentunder="true"><m:mo movablelimits="false">lim</m:mo><m:mrow><m:mrow><m:mo stretchy="false">|</m:mo><m:mi>x<
e">|</m:mo></m:mrow><m:mo>→</m:mo><m:mi mathvariant="normal">∞</m:mi></m:mrow></m:munder><m:mo></m:mo><m:mrow><m:mo stretchy="false">|</m:mo><m:mrow><m:msup><m:mi>u</m:mi><m:mi>N</m:m
:mrow><m:mo stretchy="false">(</m:mo><m:mi>x</m:mi><m:mo stretchy="false">)</m:mo></m:mrow></m:mrow><m:mo stretchy="false">|</m:mo></m:mrow></m:mrow></m:mtd><m:mtd columnalign="center
:mtd columnalign="left"><m:mrow><m:mn>0</m:mn><m:mo>,</m:mo></m:mrow></m:mtd></m:mtr></m:mtable><m:mi/></m:mrow></m:math></Math>
      </equation>
```

So far some ideas about the use of the publications-based approach for mathematical information.

Thanks for your attention!

Contact
wolfram@zbmath.org

Leibniz Association