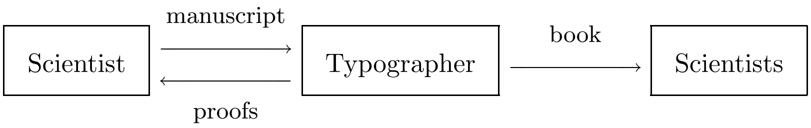
Natural TEX notation in mathematics

Michal Marvan

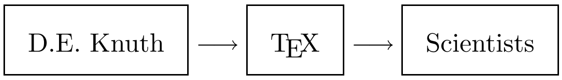
The current TEX/LATEX notation for math expressions encodes presentation, while mathematicians generally wish to communicate the content. This lecture introduces Nath, a new LATEX 2.09/2_ε style. Nath provides a very coarse context-independent notation (a "NAtural maTH notation"), from which presentation is derived in a context-dependent manner. The natural notation once again exploits the key principle of LATEX typography – separation of presentation and content.

The traditional scheme of scientific publication



Controversy: Typographers shape publications without understanding their content.

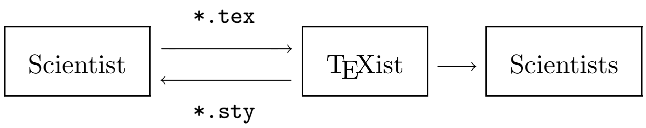
Knuth's scheme of scientific publication



Go forth now and create masterpieces of the publishing art!
– D.E. Knuth, The TEXbook (1984)

Lamport's scheme of scientific publication

or



While technology may change, what constitutes good typography does not
– J. Craig, Basic Typography. A Design Manual (1990)

The ugly (in-line style)

(i) Built-up in-line fractions, e.g., $\frac{\partial f}{\partial x}$. Should be either $\partial f/\partial x$ or displayed. Numeric fractions are O.K.: $\frac{1}{2}$.

The `\over` and `\frac` encode presentation, which may be wrong in some contexts. Proper translation from display to in-line requires understanding of mathematics (adding parentheses whenever needed for preservation of the mathematical meaning).

(ii) Wrong spacing after math operators: $\lambda\,\mathrm{id}\, {-g}$ should be $\lambda\,\mathrm{id} - g$.

The TEXbook, p. 170: "...such cases never arise, because Bin atoms must be preceded and followed by atoms compatible with the nature of binary operations."

The ugly (display style)

(iii) Composed fractions:

$$\frac{1+\frac{u}{v}}{1-\frac{u}{v}} \quad \text{should be} \quad \frac{1+\dfrac{u}{v}}{1-\dfrac{u}{v}} \quad \text{or} \quad \frac{1+u/v}{1-u/v}.$$

(iv) Wrong size of delimiters: Both

$$(\sum_{i\in I} a_i) \quad \text{and} \quad \left(\sum_{i\in I} a_i\right)$$

are ugly; should be

$$\biggl(\sum_{i\in I} a_i\biggr).$$

(v) Big numeric fractions in displaystyle:

$$A = \frac{1}{2}B \quad \text{should be} \quad A = \tfrac{1}{2}B.$$

The “impossible”

(vi) Asymmetric positioning of delimiters:

$$\left(\frac{1}{1 - \frac{u}{v}} \right).$$

(vii) Alignments and line breaks disable `\left` and `\right`.

(viii) Alignments and line breaks interfere with the spacing mechanism:

$$\begin{array}{rcl} A & = & B + C \\ & & +D + E. \end{array}$$

Both (vii) and (viii) have been already addressed by the `breqn` package from M. Downes, Breaking equations, *TUGboat* 18 (1997) 182–194.

A puzzle

(ix) Which of the delimiters

$$C_4 \left| \left| f \left| \tilde{S}_{a,-}^{-1,0} W_2 \Omega \right| \right| \left| |u| \rightarrow W_2^{\tilde{A}} \Omega \right| \right|$$

is left and which is right?

Natural notation

The aim is to give authors an opportunity to encode the mathematical meaning instead of the presentation.

A natural notation is defined to be the coarsest (context independent) notation such that an unthinking machine can derive the (context dependent) presentation.

E.g., all fractions are written via `\frac` and the style decides on their type (built-up, case or solidus). The style adds parentheses whenever needed.

An editor who does not understand the mathematics and
does not know what to do, should ask someone who does.
– The CBE Manual for Authors, Editors, and Publishers (1984)

Fractions

Fractions indicate division in a very broad sense (cf. $\partial f/\partial x$). They occur in three shapes:

$$\text{built-up}\ \ \frac{A}{B}, \qquad \text{piece}\ \ \tfrac{1}{2}, \qquad \text{solidus}\ \ A/B.$$

Besides of the obvious slash `/`, Nath provides a unique command `\frac`.

Displayed fractions

All non-numeric fractions come out as built up.

Numeric fractions are built up if and only if this will not extend any paired delimiters (the *principle of smallest fences*):

$$(\tfrac{1}{2} + x)\left(\frac{1}{2} + \frac{1}{x}\right),$$

$$\int x\,dx = \frac{1}{2}\,x^2.$$

In-line fractions

A `\frac` with numeric arguments results in a case fraction, such as $\frac{1}{10\,000}$.

Otherwise a solidus fraction results and parentheses are added whenever needed for preservation of the mathematical meaning. E.g.,

```
$\frac{\frac ab}{\frac cd}$
```

produces

$$(a/b)/(c/d).$$

References:

[1] J.W. Rayleigh, Address of the president, Lord Rayleigh, O.M., D.C.L., at the anniversary meeting on November 30, 1918, *Proc. Roy. Soc. London, Sect. A* 82 (1909) 1–17.

[2] K. Wick, *Rules for Typesetting Mathematics* (Mouton, The Hague, 1965); translated from the Czech original *Pravidla Matematické Sazby* (Academia, Praha, 1964).

Fencing rules (syntactically)

Binary operations other than slash have less binding power than the slash,

$$x + \frac{a+b}{c+d} \quad \rightarrow \quad x + (a+b)/(c+d),$$

$$\frac{\frac{a \cdot b}{c} \cdot d}{c \cdot d} \quad \rightarrow \quad ((a \cdot b)/c \cdot d)/(c \cdot d).$$

Juxtaposition has more binding power than the slash,

$$\frac{\partial}{\partial x} \frac{f}{g} \quad \rightarrow \quad (\partial/\partial x)(f/g)$$

$$\frac{a}{bc} \quad \rightarrow \quad a/bc.$$

Exceptions

$$\frac{\sin x}{2} + \sin\frac{x}{2} \qquad \rightarrow \qquad (\sin x)/2 + \sin(x/2),$$

$$-\frac{u}{v} + 2\frac{u}{v} + \frac{\pm u}{v} \qquad \rightarrow \qquad -u/v + 2u/v + (\pm u/v).$$

Fencing rules (TeXnically)

Fences around the whole fraction:

Type	Left neighbour	Example	Right neighbour	Example
Ord	Yes[1]	$x(a/b)$	Yes	$(a/b)x$
Op	Yes	$\sin(a/b)$	Yes	$(a/b)\sin x$
Bin*	No[2]	$1+a/b$	No	$a/b+1$
Rel	No	$=a/b$	No	$a/b=$
Open	No	$[a/b$	Yes	$(a/b)[$
Close	Yes	$](a/b)$	No	$a/b]$
Punct	No	$,a/b$	No	$a/b,$
Inner	Yes[1]	$\frac{1}{2}a/b$	Yes	$(a/b)\frac{1}{2}$

[1] No, if the left neighbour is a digit or a piece fraction (hence Inner) and at the same time A starts with neither Bin* nor digit nor decimal point. E.g., $\frac{1}{2}a/b$, but $\frac{1}{2}(-2a/b)$, $\frac{1}{2}(25a/b)$, $\frac{1}{2}(.5a/b)$.

[2] Yes, if A starts with Bin*, e.g., $1+(-a/b)$.

More examples

$$\frac{a}{(b+c)(b-c)} \qquad \to \qquad a/(b+c)(b-c),$$

$$\frac{\sqrt{a+b}}{\sqrt{c+d}} \qquad \to \qquad \sqrt{a+b}/\sqrt{c+d},$$

$$\frac{\partial^{k+l} f}{\partial x^k \, \partial y^l} \qquad \to \qquad \partial^{k+l} f/\partial x^k \, \partial y^l,$$

$$\frac{\sin(x+y)}{x+y} \qquad \to \qquad (\sin(x+y))/(x+y),$$

$$\frac{\bigl|\langle X_1, X_2\rangle\bigr|}{\|X_1\| \, \|X_2\|} \qquad \to \qquad |\langle X_1, X_2\rangle|/\|X_1\| \, \|X_2\|.$$

Delimiters

TEX's `\left` and `\right` produce rather poor results, especially when overused or underused.

Under natural notation every left or right fence is a left or right delimiter by default, and Nath does its best to match them properly to the material enclosed between, even asymmetrically. Example:

$$\frac{M}{\left(1 - \dfrac{x_1 + \cdots + x_n + pZ}{r}\right)\left(1 - p\dfrac{\dfrac{\partial Z}{\partial x_2} + \cdots + \dfrac{\partial Z}{\partial x_n}}{\rho}\right)}.$$

Sub- and superscripts are ignored:

$$\left(\sum_{i=0}^{\infty} a_i\right)^2.$$

Needless to say, a line break may occur between delimiters.

The modifiers `\left` and `\right` still must be used with symmetric delimiters (e.g., vertical lines | and ‖) or when intended to override the natural disposition (e.g., `\left]`).

Paired delimiters

Left delimiters		Right delimiters	
`(`	(	`)`	)
`[`,`\lbrack`	[	`]`,`\rbrack`	]
`\{`, `\lbrace`	{	`\}`, `\rbrace`	}
`<`, `\langle`	⟨	`>`, `\rangle`	⟩
`\lfloor`	⌊	`\rfloor`	⌋
`\lceil`	⌈	`\rceil`	⌉
`\lvert`, `\left|`	|	`\rvert`, `\right|`	|
`\lBrack`, `\double[`	⟦	`\rBrack`, `\double]`	⟧
`\lAngle`, `\double<`	⟪	`\rAngle`, `\double>`	⟫
`\lFloor`	⌊⌊	`\rFloor`	⌋⌋
`\lCeil`	⌈⌈	`\rCeil`	⌉⌉
`\lVert`, `\ldouble|`	‖	`\rvert`, `\rdouble|`	‖
`\triple[`	⟦[	`\triple]`	]⟧
`\triple<`	⟨⟨⟨	`\triple>`	⟩⟩⟩
`\ltriple|`	⦀	`\rtriple|`	⦀

Middle delimiters

Example:

$$\left\{ (x_i) \,\middle|\, \sum_{i<r} x_i \right\}$$

The full list of them:

`\mid`, `\middle|`	$\vert$
`\middle/`	$/$
`\Mid`, `\double|`	$\Vert$
`\double/`	$//$
`\triple|`	$\vert\vert\vert$
`\triple/`	$///$

Middle delimiters have the size of the nearest outer paired delimiters. Numeric fractions use the same principle.

Operators

1) Nath typsets $\lambda\,\mathrm{id} - g$.

2) Nath enables `\\` in subscripts of big operators:

$$\sum_{\substack{i<r \\ i\,\mathrm{odd}}} x_i.$$

3) `!` produces proper spacing around factorials, e.g., $(m!\,n!)$.

Abbreviations

Abbreviations are letter strings starting from the back quote '`'. They are typeset in roman. Multiletter abbreviations are `\mathop`'s.

For instance, `` `e^{2\pi`i} `` and `` `ad_x y `` typeset as

$$
\begin{aligned}
&\mathrm{e}^{2\pi \mathrm{i}} = -1,\\
&\mathop{\mathrm{ad}}\nolimits_x y.
\end{aligned}
$$

More examples:

$$
\begin{aligned}
&H' = H'_{\mathrm{symm}} + H'_{\mathrm{antisymm}},\\
&\bar{f} = f|_{\mathop{\mathrm{int}} U},\\
&a = \mathrm{const}_1\,,\\
&G = \mathrm{SO}(n),\\
&\mathop{\mathrm{span}}\{u, v\}.
\end{aligned}
$$

Displayed formulas

`$$` *stuff* `=` *stuff* `,` `\\` *stuff* `=` *stuff* `.` `$$` is a valid syntax and results in a left-aligned multiline formula:

$$\begin{array}{l} \rule{4cm}{1pt} = \rule{10cm}{1pt}, \\ \rule{8cm}{1pt} = \rule{4cm}{1pt}. \end{array}$$

A more sophisticated arrangement may be achieved by using `\quad`'s:

$$\begin{array}{l} \rule{8cm}{1pt} = \rule{2cm}{1pt} + (\rule{4cm}{1pt} \\ \qquad\qquad + \rule{10cm}{1pt}) \\ \qquad = \rule{14cm}{1pt} \\ \qquad = \rule{12cm}{1pt}\,. \end{array}$$

Walls

Display mode of delimiters clashes with alignments unless every cell has balanced delimiters (as is with matrices).

Walls are a convenient replacement for unbalanced alignments in displayed formulas. The syntax is `\wall` *stuff* `\\` *stuff* `\\` $\cdots$ `\\` *stuff* `\return`, and can be nested. The `\wall` makes every next line to start at the "wall" until `\return` turns it off. For instance,

$$\begin{aligned} \rule{2cm}{1pt} &= \rule{2cm}{1pt} + (-\rule{8cm}{1pt} \\ &\qquad\qquad\quad + \rule{10cm}{1pt}) \\ &= \rule{14cm}{1pt} \\ &= \rule{12cm}{1pt}\,. \end{aligned}$$

The typical placement of `\wall` is in front of a relation symbol or immediately after an opening delimiter.

Three dots

The Nath's rule: In math mode, three dots are raised to the level of axis

$$a_1 + \cdots + a_n,$$

unless they follow a comma or a semicolon:

$$a_1, \ldots, a_n.$$

The puzzle solved

```
\delimgrowth=1
```

$$C_4 \Big| |f| \tilde{S}_{a,-}^{-1,0} W_2 \Omega| \Big| \Big| |u| \to W_2^{\tilde{A}} \Omega| \Big|.$$