Formatting Mathematical Articles with Cascading Style Sheets

George Chavchanidze

Department of Theoretical Physics, A. Razmadze Institute of Mathematics, 1 Aleksidze Street, Tbilisi 0193, Georgia

Abstract. This page shows how to format mathematical articles with Cascading Style Sheets (CSS). Simple XML 1.0 markup is used to capture basic structure of math expressions while rendering is specified via CSS 2.0/2.1. Abilities and limitations of present approach are discussed.

1. Introduction

CSS is the simplest way to control formatting of XML and SGML documents. Today it is mainly used in web design to render and style (X)HTML documents, but its applications spread far beyond ordinary web design and one can use XML and CSS to embed mathematical expressions in web pages. Generally speaking it is easy to control general layout and style of scientific documents with CSS, the only problem is rendering of complex mathematical formulæ. So in this article we will mainly focus on this problem and suggest relatively simple approach for rendering of mathematical expressions that may appear in scientific papers. We use simple XML 1.0 markup to capture basic structure of math expressions and CSS 2.0/2.1 to specify their rendering (see [2]-[4] for specs).

Each mathematical expression may be formatted as inline equation like $z = 1/c$ or as block level (displayed) equation

$$z = 1/c \tag{1}$$

Rendering of inline equations is more specific as one has to care about height of line box.

2. Indices in display mode

In mathematical expressions indices are probably the most widespread objects so it is essential to keep corresponding markup as simple as possible, otherwise compactness and readability of XML source will be seriously undermined. The simpliest way is to use CSS relative positioning to shift indices up/downwards. Most of the browsers support relative positioning and thus can easily process simple indices. Here are some examples of top

$$dz^k = dx^k + idy^k \tag{2}$$

$$ch^2(x) - sh^2(x) = 1 \tag{3}$$

$$ch(3x) = 4ch^3(x) - 3ch(x) \tag{4}$$

$$Na - e^- = Na^+ \tag{5}$$

and lower indices taken from mathematics and chemistry

$$T_{mn} = - T_{nm} \tag{6}$$

$$Fe_3O_4 + 4H_2 \rightleftarrows 3Fe + 4H_2 \tag{7}$$

Simple indices may be combined. Subscripts may precede superscripts and vise versa.

$$T_{mn} = \varepsilon_{nm}{}^k S_k \tag{8}$$

$$T^{mn} = \varepsilon^{nm}{}_k S^k \tag{9}$$

$$_{13}Al^{27} + {}_2He^4 = {}_{15}P^{30} + {}_0n^1 \tag{10}$$

Indices may be also positioned one over another in this way

$$\nabla_m R^n_{ikl} + \nabla_l R^n_{imk} + \nabla_k R^n_{ilm} = 0 \tag{11}$$

However to archive correct alignment of such an indices one has to put them in [inline table](). Here browser compatibility issues show up as some browsers does not support inline tables. Nesting of simple indices is allowed.

$$dz^{k_1} = dx^{k_1} + i dy^{k_1} \tag{12}$$

$$dz^{k^{(1)}} = dx^{k^{(1)}} + i dy^{k^{(1)}} \tag{13}$$

$$T_{k_1 k_2} = - T_{k_2 k_1} \tag{14}$$

$$T_{k^{(1)} k^{(2)}} = - T_{k^{(2)} k^{(1)}} \tag{15}$$

$$e^{A^{(1)}} e^{A^{(2)}} = e^{A^{(1)} + A^{(2)}} \tag{16}$$

$$X_h(F) = L_{X_h} F \tag{17}$$

$$\nabla_{k^1} R^{k^5}_{k^2 k^3 k^4} + \nabla_{k^4} R^{k^5}_{k^2 k^1 k^3} + \nabla_{k^3} R^{k^5}_{k^2 k^4 k^1} = 0 \tag{18}$$

$$\nabla_{\hat{s}_1} R^{\hat{s}_5}_{\hat{s}_2 \hat{s}_3 \hat{s}_4} + \nabla_{\hat{s}_4} R^{\hat{s}_5}_{\hat{s}_2 \hat{s}_1 \hat{s}_3} + \nabla_{\hat{s}_3} R^{\hat{s}_5}_{\hat{s}_2 \hat{s}_4 \hat{s}_1} = 0 \tag{19}$$

So rendering indices with CSS2 is not a big problem. Their rendering can be controlled in details by both author and user through adjusting relative positioning offsets, font-size and style in author and user style sheets.

3. Indices in inline mode

Here similar math expression appear inline. Note that relatively positioned indices should not affect line height $T_{mn} = - T_{nm}$. Alternatively one can shift indices using CSS vertical-align property used to control vertical alignment. However, unlike relative positioning, vertical alignment may affect height of line box. So it is better to use relative positioning $\varepsilon_{nm}{}^{k} S_k$. Line breaks inside equations are prohibited as browsers may generate line break in inappropriate place (after all, there is no

reliable algorithm for handling line breaks in mathematical equations). Thus the simplest solution is to prohibit all line breaks inside math expressions $\varepsilon^{nm}{}_{k}S^{k}$ (this can be done by setting CSS ['white-space'](#) property to 'nowrap') and mark those few points where line breaks are allowed with an element that has CSS 'white-space' property set to 'auto'. More complex indices $\nabla_{m}R^{n}_{ikl}$ can be used in inline mode as well, but again some browsers does not like them. This happens due to weak CSS support in those browsers. One can nest indices as follows $\nabla_{k^1}R^{k^5}_{k^2k^3k^4}$. Nested indices are also shifted using relative positioning and thus they should not change height of line box. Here are another samples of nested superscripts $e^{A^{(1)}}e^{A^{(2)}}$ and nested subscripts $X_{h}(F) = L_{X_{h}}F$.

4. Fractions in display mode

Fractions are quite frequent in mathematical articles, so it is preferable to have simple markup for fractions. Making it too simple results poor rendering quality so we think that the minimal reasonable approach is to take two elements, to mark fraction and its rows (numerator/denominator). Fractions can be formatted as vertically centered [inline tables](#). To reduce markup [anonymous table objects](#) may be exploited. Below you see example from statistics

$$\overline{A} = \frac{Tr(AW)}{Tr(W)} \tag{20}$$

another example from gravity

$$R_{nm} - \frac{1}{2}Rg_{nm} = 0 \tag{21}$$

and some samples from analysis and trigonometry

$$\ln(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots \tag{22}$$

$$B_6 = \frac{3617}{510} \tag{23}$$

$$\mathrm{ch}(x) = \frac{1}{2}(e^x + e^{-x}) \tag{24}$$

$$\mathrm{th}(x \pm y) = \frac{\mathrm{th}(x) \pm \mathrm{th}(y)}{1 \pm \mathrm{th}(x)\mathrm{th}(y)} \tag{25}$$

Fractions may contain nested indices. Here are some samples from mechanics

$$I = \frac{m_1 m_2}{m_1 + m_2} x^2 \tag{26}$$

$$D = \frac{4B^3}{\omega^2} \tag{27}$$

examples from field theory

$$F_{mn} = \frac{\partial A_m}{\partial x^n} - \frac{\partial A_n}{\partial x^m} \tag{28}$$

$$F^{mn} = \frac{\partial A^m}{\partial x_n} - \frac{\partial A^n}{\partial x_m} \tag{29}$$

and other mathematical expressions

$$\mathrm{th}(2x) = \frac{2\mathrm{th}(x)}{1 + \mathrm{th}^2(x)} \tag{30}$$

$$G = 1 - \frac{1}{3^2} + \frac{1}{5^2} - \frac{1}{7^2} + \ldots \tag{31}$$

Nesting of fractions is possible but limited to second order

$$g = 1 - \frac{1}{1 - \frac{a}{r}} \tag{32}$$

$$m = \frac{1}{\frac{1}{m_1} + \frac{1}{m_2}} \tag{33}$$

More deep nesting is also possible, but may require extra markup so at the moment it is reasonable to limit nesting to second order as in real world articles deeper nesting is rarely used and can be always avoided.

5. Fractions in inline mode

In inline mode fractions are rendered in the same manner as in block level equations $\frac{Tr(AW)}{Tr(W)}$. The only difference is that inline fractions are slightly compactified (height of numerator/denominator is reduced) $\frac{th(x) \pm th(y)}{1 \pm th(x)th(y)}$ to avoid possible line-height enlargement and nesting of fractions in inline mode is not allowed but nesting of indices is possible so inline fractions may contain nested indices $\frac{m_1 m_2}{m_1 + m_2} x^2$. Here is another example with nested superscripts $\frac{2th(x)}{1 + th^2(x)}$

6. Operators in display mode

Rendering of indexed operators is much more complicated issue as in general number of under and over scripts may be arbitrary and in addition one needs to rearrange all this stuff when operators are nested in fractions or appear inline. Also shape and baseline of glyphs may vary. At the moment we will assume that only single under and over scripts are used (multiple under scripts are really rarely used while multiple over scripts are almost never used). One can format operators as inline tables. Below are some sample integrals

$$B(m, n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx \tag{34}$$

Here are some sums

$$p = \sum_{m=1}^{n} p_m \tag{35}$$

$$G = \sum_{p,\, s \,\in\, Z} A^p B^s \tag{36}$$

Product

$$\Omega = \prod_{m \,\in\, N} \Omega^{(m)} \tag{37}$$

Unification operator

$$A = \bigcup_{m=1}^{\infty} A^{(m)} \tag{38}$$

and intersection

$$B(s) = \bigcap_{m=1}^{2n} B^{(m)} \tag{39}$$

Direct sum and tensor products are rendered in the same manner

$$Q\left[\bigoplus_{k=1}^{n} x_k \right] = \bigoplus_{k=1}^{n} x_{Q[k]} \tag{40}$$

$$\hat{A} = \bigotimes_{k=1}^{n} \bigotimes_{s=1}^{m} N_{ks} M_s \tag{41}$$

Under and over scripts may contain nested indices

$$G = \sum_{m_1 m_2} A^{m_1} B^{m_2} \tag{42}$$

$$F = \sum_{m_1 \neq m_2} P_{m_1} P_{m_2} \tag{43}$$

Operators can be combined with indices and fractions to form complex expressions

$$\int_0^{\infty} \frac{e^{-ax} - e^{-bx}}{x}\, dx = \ln \frac{b}{a} \tag{44}$$

$$\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}} \tag{45}$$

$$\int_0^1 \frac{dx}{1 + 2x + x^2} = \frac{1}{2} \tag{46}$$

$$\int_0^1 \frac{\sin^2 mx}{c^2 + x^2} dx = \frac{\pi}{4c}(1 - e^{-2mc}) \tag{47}$$

$$\int_0^1 x^{2p-1} \ln(1 + x) dx = \frac{1}{2p} \sum_{n=1}^{2p} \frac{(-1)^{n-1}}{n} \tag{48}$$

Nested fractions should not affect vertical alignment of parent fraction

$$\int_0^1 x^m (1 - x^n)^p dx = \frac{\Gamma(p + 1)\Gamma(\frac{m+1}{n})}{n\Gamma(p + 1 + \frac{m+1}{n})} \tag{49}$$

$$\int_0^\infty \frac{\sin(mx)}{e^{ax} + 1} dx = \frac{1}{2m} - \frac{\pi}{2a\,\text{sh}(\frac{\pi m}{a})} \tag{50}$$

$$\int_0^\infty \frac{\text{th}(\frac{\pi x}{2})}{1 + x^2} \sin(mx) dx = \frac{m}{e^m} - \text{sh}(m)\ln(1 - e^{-2m}) \tag{51}$$

$$\int_0^\infty \frac{\cos(mx)\cos(nx)}{\text{ch}(ax)} dx = \frac{\pi \text{ch}(\frac{\pi m}{2a})\,\text{ch}(\frac{\pi n}{2a})}{a(\text{ch}(\frac{\pi m}{a}) + \text{ch}(\frac{\pi n}{a}))} \tag{52}$$

Operators can be nested inside fractions. But in this case it is better to replace under/over scripts with sub/superscripts placed after operator

$$\overline{A} = \frac{\sum_{k=1}^n W^{(k)} A^{(k)}}{\sum_{k=1}^n A^{(k)}} \tag{53}$$

$$S = \frac{\prod_{k=1}^n P_k}{\prod_{k=1}^n Q_k} \tag{54}$$

$$\bar{F} = \frac{\sum_{k=1}^{m} F^{(m)} W^{(m)}}{\sum_{k=1}^{m} W^{(m)}} = \frac{\sum_{k=0}^{m} F^{(m)} W^{(m)}}{\hat{W}} \tag{55}$$

So it is possible to render common operators with CSS.

7. Operators in inline mode

In inline mode it is better to replace under/over scripts with indices placed after operator, $\sum_{m=1}^{n} P_m$ in this way possible line height distortions are reduced $\prod_{m=0}^{s} \Omega^{(m)}$. The additional problem is that in different fonts glyphs that correspond to mathematical operators like sums, products and intergals have different shapes, baselines and sizes $\sum_{m_1 \neq m_2} P_{m_1} P_{m_2}$ so if in one font operators look perfectly centered in others they may appear to be distorted. The problem will be partly resolved when we will have single font with good coverage of all mathematical ranges (maybe [STIX](#)).

8. Under and Over scripts

Rendering of generic under and over scripted expressions is easier then rendering of indexed operators. They are formatted as inline tables. Here is for example limit

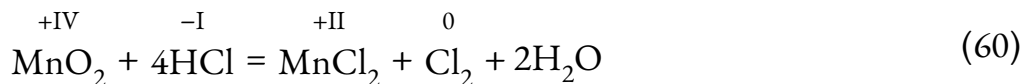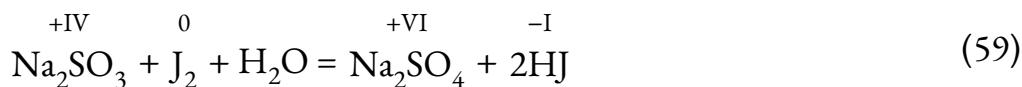$$c_{-3} = \lim_{x \to 0} x^3 s(x) \tag{56}$$

Residue

$$I(z_0) = \operatorname*{Res}_{z = z_0} F(z) \tag{57}$$

Under script with brace (rendering of such a simple square under/over braces is not a problem but making them round, or curly or allowing overlaps is much more difficult to achieve, so currently it is better to stick to simplest case).

$$V^{(n)} = \underbrace{W \wedge W \wedge \cdots \wedge W}_{n \text{ times}} \tag{58}$$

Below you see some samples from chemistry

$$\overset{+IV}{Na_2SO_3} + \overset{0}{J_2} + H_2O = \overset{+VI}{Na_2SO_4} + 2\overset{-I}{H}J \tag{59}$$

$$\overset{+IV}{MnO_2} + 4H\overset{-I}{Cl} = \overset{+II}{MnCl_2} + \overset{0}{Cl_2} + 2H_2O \tag{60}$$

$$Zn + H_2SO_4 \overset{+2e^-}{=} ZnSO_4 + H_2 \tag{61}$$

Here is example from nuclear physics

$$\underset{226}{Ra} \rightarrow \underset{222}{Rn} + \underset{4}{He} \tag{62}$$

Under over scripts are really rarely used. Below is rather artificial example from chemistry.

$$\overset{27}{\underset{13}{Al}} + \overset{4}{\underset{2}{He}} = \overset{30}{\underset{15}{P}} + \overset{1}{\underset{0}{n}} \tag{63}$$

9. Vectors, Matrices and Cases

Vectors, matrices and cases can be formatted as inline tables. Here are sample matrices and vectors. 'Hooks' are produced by inserting extra cells at edges using CSS generated content and removing inner borders via border collapse mechanism.

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \tag{64}$$

Matrices may contain nested fractions, indices and operators

$$M_3 = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} = \begin{bmatrix} \frac{a^2}{b^2+c^2} & a^3 & -b^2 \\ c^4 & \frac{a^2+b^2}{12c} & 8a \\ 12c & b^4 & \frac{2a^2}{7b^3} \end{bmatrix} \tag{65}$$

rendering of determinants is much easier as no 'hooks' are necessary.

$$\det(M_3) = \begin{vmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{vmatrix} = \begin{vmatrix} \frac{a^2}{b^2+c^2} & a^3 & -b^2 \\ c^4 & \frac{a^2+b^2}{12c} & 8a \\ 12c & b^4 & \frac{2a^2}{7b^3} \end{vmatrix} \qquad (66)$$

Rendering of multivalues is slightly simplified.

$$\varepsilon(x) = \begin{vmatrix} +1 & \{\text{if } x > 0\} \\ 0 & \{\text{if } x = 0\} \\ -1 & \{\text{if } x < 0\} \end{vmatrix} \qquad (67)$$

10. Large Brackets

Simple square brackets can be imitated using generated content and borders. This is the simplest and most reliable way to handle them.

$$D = \left[G(z) + \frac{1}{S(z)} \right]^3 \qquad (68)$$

$$\hat{C}^{(m)} = \prod_{k=m}^{m^2} \left[kF^{k^3}(x) + \prod_{s=c^4}^{k^5} \frac{F^s(x) + \frac{1}{2}G^s(x)}{F^s(x) - \frac{1}{2}G^s(x)} \right]^{3k} \qquad (69)$$

$$\hat{S}^\pm = \frac{4\theta^2}{6z^3 + \frac{z^2}{3u}} \int_{2c^2}^{4b^2} \left[e^{\pm 3uc^2\hat{w}^2} + \sum_{m_1 \neq m_2} e^{\pm m_1 m_2 \hat{w}} \right] d\hat{w} \qquad (70)$$

As an alternative solution one can compose brackets by combining Unicode characters 239B-23AD (located in Misc. Technical range). Task can be simplified using CSS generated content.

11. Diacritical Marks

Rendering of diacritical marks is not governed by CSS. This issue is addressed by Unicode standard [1] that defines so called combining diacritical marks like ˙ ¨ ^ ˜ ¯ ^ that being combined with ordinary Unicode characters must produce characters with over dots, hats, tildes, bars etc.

$$\ddot{q} + \dot{q}^2 + q^4 = \tilde{c} \tag{71}$$

but at the moment browsers does not support combining diacritical marks properly and it is better to use precomposed characters

$$[\hat{A}, \hat{H}] = 0 \tag{72}$$

located in the following Unicode ranges: [Latin-1 Supplement](), [Latin Extended-A](), [Latin Extended-B](), [Latin Extended Additional]().

12. Radicals

The most reliable way to settle issue with radicals is to use more simple power notations

$$x = \frac{-b \pm (b^2 - 4ac)^{\frac{1}{2}}}{2a} \tag{73}$$

$$R^{\pm} = \left[m^2 \pm \frac{12c^5}{29w^7} \right]^{3/2} \tag{74}$$

In this way one avoids dependence on font metrics.

13. Nesting Limitations

Our present style sheet uses simple 'shrink to fit' nesting scheme to ensure that nested element does not affect height of its parent. Such a rendering naturally imposes limits on nesting of certain expressions. In overall these limitations are not severe and does not affect real world articles. For example simple nesting of fractions is allowed

$$Q = B_0 + \frac{A_0}{B_1 + \frac{A_1}{B_2 + A_2 / B_3}} \tag{75}$$

but those who try to nest fractions deeper will encounter nesting limitation. It is possible to remove this and other nesting limitations by using more complex markup, but at the moment we don't want to do this as limitation are not severe.

14. Conclusions

CSS2 is basically simple style language, but in spite of its simplicity it is powerful enough to afford rendering of most of mathematical expressions. So one can start using CSS for formatting of scientific and technical XML documents already today. Rendering of some things like for instance radicals, and deeply nested expressions are not fully addressed in the present article. However it happened just because current CSS2 based solutions for rendering of radicals are slightly artificial, while support for deep nesting requires slightly more detailed markup then we currently use, so taking into account that this issues are not urgent we think it is better to address them later when we will have more natural solutions (like CSS3 math module).

References

1. B. Beeton et al., Unicode Support for Mathematics, 2003
2. B. Bos et al., Cascading Style Sheets, level 2 revision 1, 2004
3. T. Bray et al., Extensible Markup Language (XML) 1.0, 2004
4. J. Clark, Associating Style Sheets with XML documents, 1999