The Printing of Mathematics:  
Some Hints for Non-mathematicians

Workshop, SFEP Conference, Eastbourne, September 1994

The usual convention is that italic is used for mathematical variables. How do you know if something is a variable? One answer is that it ought to be defined when it is first introduced. In a mathematical context the definition may be yet more maths, so this may not help. Variables can be manipulated mathematically, for example

\[(x + y)(x - y) = x^2 - y^2.\]

The \(x\) and \(y\) are variables. These letters are quite arbitrary and do not stand for anything. In expressions such as:

For a rectangle of length \(l\) and height \(h\), the area is \(a = lh\)

the letters obviously do stand for something, but they too are variables. Computer scientists often use whole words as variables:

\[\text{Area} = \text{Length} \times \text{Height}\]

In biology and medicine multi-letter abbreviations are often used as variables. For example:

Let the muscle strength ratio (MSR) be defined as

\[\text{MSR} = \frac{\text{MM}}{\text{DR}}\]

where MM is muscle mass and DR is the dynamometer reading (in arbitrary units).

Logically these should be italic too, but roman is often used. This avoids any possible confusion about \(MIR\) meaning \(M \times I \times R\); also, if the same terms have been used in the text as abbreviations, it seems odd to italicize them when they are used as variables.

Italic is also used for physical constants such as \(g\) (the gravitational constant), or generic mathematical constants which as often designated \(c\).

Roman is used for operators. These include abbreviations that are now considered a part of conventional mathematical notation:

\[\log, \cos, \sin, \tan, \exp, \ldots\]

There are quite a lot of these: lists are to be found in reference books such as Mathematics into Type. Some of them have variant forms. You may be familiar with cosec and cotan, but would you recognize csc and cotn?

It is now standard to use roman for the differential operator \(d\) and the exponential operator \(e\) (although many mathematicians and US usage haven’t yet caught up). You can generally recognize a differential \(d\): it comes in contexts like

\[\int f(x) \, dx.\]
Mathematics Workshop

Where there are integral signs, you can expect to find differentials too. The exponential e can appear as a subscript (\(\log_e\) meaning logarithm to the base e, same as natural logarithm ln). When it denotes an exponential, it may have a very cumbersome superscript:

\[
\frac{1}{e^{2\pi i^2}}
\]

or even worse. This is the reason for the use of the exp notation,

\[
\exp \left\{ \frac{1}{2\pi i^2} \right\}
\]

Mathematical constants such as i (square root of \(-1\)) are also roman. The use of an upright rather than slanting Greek font for \(\pi\) is also recommended when it is used as a constant. This is because \(i\) and \(\pi\) can also be defined as variables.

Another standard use of roman is for labels – often subscripts or superscripts. For example:

- Let the mass of the cat be \(m_C\) and the mass of the mouse be \(m_M\), and let their velocities be \(V_C\) and \(V_M\) respectively.

The purpose of the labels is often self-evident. Symbols such as \(T_{\text{max}}\) and \(T_{\text{min}}\) may not be explicitly defined because their meaning is obvious enough.

Bold is generally used for vectors and matrices. Sometimes these are specifically introduced:

- the vector \(\mathbf{x}\), with components \(z_1, z_2, z_3, \ldots\)

The reference to ‘components’ is helpfully diagnostic when the text is not so specific, as is the coexistence of subscripted (scalar, light face) and unsubscripted (vector, bold) versions of the same letter. Sometimes the author will indicate vector notation by a typed underline (\(\underline{z}\)), sometimes by an arrow over the letter (\(\vec{z}\)). This might have to be standardized in a multi-author work, or as a matter of house style. Modern typesetting systems can easily cope with these embellishments, but the use of bold is preferable in printed mathematics.

Italic and roman bold, serif and sans serif, are available. Publishers’ house style may call for roman or italic. See BS 5261 for further details if it is necessary to use different bold fonts for different purposes.

Other alphabets

- The only Hebrew letter you are likely to encounter is \(\aleph\) (aleph).
- Script letters are often used to denote sets of entities. Many word-processing systems can’t produce them, so deciphering the author’s intentions may present a problem.
- Fraktur, also known as German or Gothic, is quite commonly used by physicists and mathematicians (especially those from continental Europe), but not likely to be found outside heavily mathematical contexts. As with script, the main difficulty is to decipher the author’s handwriting.
- Blackboard capitals, also known as shadow, or special roman, these capitals are conventionally used to denote real numbers \(\mathbb{R}\), natural numbers \(\mathbb{N}\), integers \(\mathbb{Z}\) and a few others. It’s quite acceptable if they are printed as italic or bold instead (so long as it’s consistent), unless the same letters are also used as variables.
Mathematics Workshop

- Sans serif is conventional notation in some areas (mathematical physics, for example). It can also be used for matrices. Some publishers have a practice of labelling their artwork in a sans serif font: this is not suitable for mathematical work, where it is preferable for any notation on the artwork to correspond exactly to the text. In heavy maths, authors sometimes use sans serif or blackboard letters just because they run out of fonts: they may have no special significance, and could be changed if your typesetter has a difference range of fonts.

- Greek is widely used. If you are not familiar with the alphabet — capital and lower case — keep a reference book handy. Authors usually know what letter they want, but sometimes call it by the wrong name: \( \zeta \) (zeta), \( \xi \) (xi), and \( \chi \) (chi) often fall victim to this. It is also not uncommon to confuse \( l.c. \) and cap theta (\( \theta \) and \( \Theta \)). The latter often looks quite odd in handwriting: like a little capital \( H \) in a circle.

In copy-editing, it should be sufficient to identify each letter the first time it appears if it is clearly distinguishable from other characters. How can you tell whether something is meant to be (say) \( \alpha \) not \( \alpha \), \( \rho \) not \( \rho \), or \( \nu \) not \( \nu \)? It’s useful to remember that these things often come in pairs or sequences. If there are two quantities under discussion and one is \( \beta \), chances are that the other is \( \alpha \) rather than \( a \). If a third one is introduced, it is more likely to be \( \gamma \) than \( v \). Similarly \( \lambda \), \( \mu \), \( \nu \) or \( \sigma \), \( \tau \) are often found together. If a sequence of variables \( k, l, m, n \), has associated properties \( k, \lambda, \mu, \?, \), the problem letter is likely to be \( v \). (These are only hints, not rules!)

Spacing is something that can go badly wrong if an author (or a typesetter) is using a system not specifically designed for maths. An experienced mathematical typesetter will follow house rules, and can safely be left to do so if the copy is clear. Authorities and house styles vary on the spacing round signs such as \(+\), \(-\), \(\times\), etc. (binary operators): they recommend either a thin space or none at all. But in some systems (such as ordinary typing) where the choice is simply space or no space, \( a + b - c \) looks far too cramped and \( a + b - c \) is preferable. There is a thin space round roman ‘word’ notation: \( \cos A \sin B \), not \( \cos A \sin B \) but extra space is not necessary before parentheses or brackets (although some publishers’ house styles call for it). Space is also usual before (and in some styles between) terms in integral expressions:

\[
\int \int \int dxdydz
\]

but not between the differential \( d \) and the following letter. If you have more than one displayed equation on the same line, there should be adequate space between them: a double quad space (2 ems) is usual. One em is enough on either side of a word linking two equations, as in:

\[
\alpha = \beta + \gamma \quad \text{and} \quad a = b + c.
\]

If your author is sparing in his or her use of commas, misreadings are quite likely. For example, ‘for all \( a x^2 = b \to \infty \), meaning ‘for all \( a, x^2 = b \to \infty \)’ may come out as ‘for all \( ax^2 = b \to \infty \)’, which makes no sense. Sometimes authors are quite happy with the unpunctuated version in typescript, but not when they see it in proof because the word spacing is generally much tighter.

Not every mathematical expression needs to be displayed. There are two main criteria: significance and typographical complexity. Authors sometimes display every step in their calculations. The copy-editor may realize that this is unnecessary, and indicate some running on so that the more important equations stand out. Some expressions are simply too cumbersome to be run on in text, although (unfortunately) many word-processing systems will do their best to squeeze them in. Re-using a previous example, \( e^{\frac{1}{2} x^2} \) is ugly and hard to read. The second
version, \( \exp \left\{ \frac{1}{2 \pi^2} \right\} \), is not much better, although it can be made to fit by slashing and bracketing, \( \exp \left\{ \frac{1}{2 \pi^2} \right\} \), or by the use of a negative exponent, \( \exp \left\{ -1/(2 \pi^2) \right\} \). This is a relatively mild example, but it illustrates how solving one problem can introduce another. The slashed and bracketed expression is only one line deep, but it is longer than the stacked version and cannot be broken. Generally, fractions will be built up in display and slashed in run-on text.

The typesetter may be obliged to display an expression to avoid unacceptably bad line breaks. This is something to look out for in proof correction: any change between display and run-on may have consequences a few lines later on.

Conventions for layout differ between displayed and run-on expressions. For example, limits are normally set lateral in text and vertical in display: \( \int_0^x \), \( \sum_{k=1}^\infty \), \( \prod_{k=0}^{m-1} \), but

\[
\int_0^x, \sum_{k=1}^\infty, \prod_{k=0}^{m-1}, \text{ but}
\]

Some authors may indicate this difference as a ‘printer’s error’ in proofreading.

There are also differences in spacing: expressions that would be set as

\[ x^2 + y \quad (x = 1, 2, \ldots) \]

or

\[ x^2 + y \quad \text{as } x \to \infty \]

are run on \( (x^2 + y \ (x = 1, 2, \ldots) \) and \( x^2 + y \) as \( x \to \infty \)) with just the ordinary spacing of the line between the component parts.

Displayed equations can be either centred on the measure, or indented by a fixed amount (here 2.5 ems) from the left margin. The centred style is most common, but it may seem more logical to go for a left-aligned style if the other design elements (headings, for example) are all ranged left. This should be considered at the design stage.

A multi-line display is dealt with as a single entity. To quote from *Mathematics into Type*:

> Several lines of equations which occur in succession, with no intervening words, are considered one display. In such cases the first verb symbol occurring in each equation is lined up vertically with the one below: the set of equations is then centred.

Lines that begin with a + or − sign, for example, are further indented (usually 1 em).

The question of where it is permissible to break mathematical expressions either in display or in run-on text occupies several pages in the standard reference books, and is too complex to consider in detail here. Butcher (p.302) says that ‘Equations should, if possible, be broken before an operational sign and not within brackets’. Another possible solution is to set the equation in a smaller type size. The smaller type size is a confession of failure that can almost always be avoided. Breaking before an operational sign (+, −, ×) and not within brackets is a sound rule, but may not be enough if the terms in the brackets are lengthy.

There are two opposing views about the punctuation of displayed mathematics. One is that it shouldn’t be punctuated at all (often found in engineering mathematics): the other is that it should be punctuated just like any other kind of language. This is my own preference. Complex mathematical material can be virtually incomprehensible without punctuation to indicate where one clause or sentence ends and the next begins. If the mathematics is only an isolated displayed line, this probably isn’t a problem.

**Punctuation**
In scientific and mathematical work commas are not used to indicate thousands in large numbers: a thin space is used instead. For example:

9999999.888 888 888

not 9,999,999.888888

In some countries the convention is to use a comma where we would use a decimal point, and vice versa.

The use of symbolic notation should be avoided in the prose surrounding equations — this includes the use of =, <, >, ≤, ≥, as well as symbols such as ∃ (there exists) and ∀ (for all). It is considered better style to write

For y greater than 1, for all \( s \in \mathcal{S} \), \( z \) exists and is greater than zero

rather than the shorthand form

\[ y > 1, \forall s \in \mathcal{S} \exists z > 0 \]

even if you are writing for a mathematical audience.

Whether things are confusable depends very much on the font (or handwriting) in which they are presented.

\[
\begin{align*}
\Sigma & \quad \Sigma \\
\Pi & \quad \Pi \\
\varepsilon & \quad \varepsilon \\
\Psi & \quad \Psi \\
\phi, \varphi & \quad \emptyset \\
\theta & \quad \varnothing \\
\kappa & \\
\chi & \quad \chi \\
y & \quad \gamma \\
\eta & \\
\cup & \quad \cup \\
\delta & \quad \delta \\
\alpha & \quad = \\
\nu & \quad \nu \\
\Xi & \quad \equiv \\
\Delta & \quad \Delta \\
\land & \quad \land \\
\langle \cdots \rangle & \quad \langle \cdots \rangle 
\end{align*}
\]

Sometimes the syntax of the mathematical expression makes it clear whether a particular character has to be a variable rather than an operator. Otherwise the hint above about families of characters may be helpful.
An equation number or a label such as (*) is included for ease of reference. If the equation is never referred to, it need not be numbered or labelled.

Numbering is preferable to labelling, which can be messy and hard to follow.

Numbering should be analogous to that of other numbered elements – by chapter, or by section.

House styles may demand equation numbers either ranged left or ranged right – the latter is more common.

The sequence of numbers should always be checked – it is something that often goes wrong in final revision of the text.

Some text processing systems will generate numbers automatically unless the option is turned off. They may be unnecessary.

The good news is that these fences come in pairs: the bad news is that they need not be matching pairs. The use of square/round brackets to denote open/closed intervals is common: \( x \in [0, 1) \), for example. The pairs do, however, have to match in size, and they have to be large enough to enclose everything within them. It is bad style to write

\[
\left( \sum_{x=1}^{\infty} (x + y)(x + z)(y + z) \right);
\]

you need

\[
\left( \sum_{x=1}^{\infty} (x + y)(x + z)(y + z) \right);
\]

The commonest delimiters are

\((\ldots), [\ldots], \{\ldots\}, (\ldots)(\ldots), ||\ldots||\)

in a wide range of sizes. There are also \([\ldots], ][\ldots]\) – these are not mutilated characters or optical illusions.

Careful traditional mathematicians observe a standard sequence of brackets: Chaundy, Barret and Batey specify \([\{\ldots\}]\) (working from the inside outwards).

Some copyeditors insist on observing this sequence, and will edit heavily to achieve it. There are reasons against doing so, however:

- Brackets are often used as part of notation, i.e. \{x\} or \[x\] may have a special meaning.
- The typesetter may use more brackets (or fewer) if he chooses to present a fraction in a different way: this could lead to considerable proof correction.
- It can sometimes be very difficult to ensure that analogous terms are similarly bracketed, and failure to do so can be confusing.

Imposing this sequence may be called for in house style, but should be queried if difficulties arise.
Mathematics Workshop

It is usual to have three type sizes available for mathematical work: text size, a smaller one for sub/superscripts, and an even smaller one for second-order sub/superscripts. If a third-order sub/superscript is wanted (a situation best avoided), the size remains the same as the second-order one but the position is suitably displaced.

For convenience in typing, you sometimes see

$$T_2^f = x_n^1 + y_n^1$$

This staggered appearance is undesirable, and the typesetter should be instructed to align subs and supers:

$$T_2^f = z_n^1 + y_n^1$$

How far the subs and supers need to be displaced vertically from the maths axis is a subtle typographical point – a good mathematical typesetting system will do it automatically. The size difference and the displacement are both important. Not enough displacement, and the eye misses the cue: too much, and the sub/super is optically detached from its parent letter. It may even be misattributed to the line above or below. Note the placing here:

$$a_p, b_q, c_r, d_s, p_A, q_B, r_C, \eta, \chi_1, \rho_2, \ldots$$

$$d^a, b^f, f^g, i^h, \lambda^p, \eta^q, \gamma^r, \eta^h, \chi^1, \rho^2, \ldots$$

Subscripts are not placed at the very base of a descender, nor superscripts at the tip of an ascender.

If a complex expression is set as a sub/superscript, it is sensible to mark it explicitly for the typesetter.

Fractions in displayed formulae can easily be stacked (built up):

$$\frac{a + b}{c + d}$$

In run-on text, this produces either \(\frac{a + b}{c + d}\) or \(\frac{a + b}{c + d}\), both of which are undesirable (except for very simple fractions like \(\frac{3}{7}\)). The solidus (slash) is generally used here, so the fraction becomes \((a + b)/(c + d)\). Note the need for brackets to group the terms. The other option is the negative exponent: \((a + b)/(c + d)^{-1}\).

Accents and embellishments are widely used in some branches of mathematics and physics: the commonest ones are

\(\hat{a}, \tilde{a}, \ddot{a}, \check{a}, \breve{a}, \check{a}, \bar{a}\)

These require special care in proofreading. They are often difficult to distinguish, especially if the modified letters are subscripts or superscripts. Sometimes double accents are used: \(\hat{\check{a}}, \tilde{\check{a}}\).

This is not agreeable notation: it is hard to read, and may affect the line spacing. Wide accents are also called for occasionally: \(\check{a}, \bar{a}\). It is best to try to avoid these too, because the accents cannot stretch to cover a long expression. A form such as \((abxy)^n\) can be used instead.

House style decrees whether decimal points are centred on or on the line. Points are also used to indicate multiplication or grouping of terms. If the decimal point is centred the multiplication point is on the line, and vice versa. Sometimes when authors see multiplication points in proof they realize they are unnecessary and delete them. However, you cannot just edit them
all out – sometimes they are necessary. In some contexts it is preferable to replace them by a multiplication sign: $1 \times 10^6$ rather than $1 \cdot 10^6$.

A zero should be inserted before the decimal point for numbers less than 1: i.e. 0.345, not .345.

Dots are used in notation in various ways. Dots and double dots as accents have been mentioned already. They can also be found as subscripts and superscripts, or in brackets. This can be hard to read, especially if you are working from a spotty photocopy. It should be clarified for the typesetter.

Ellipses (...) come in threes. They can be centred on the line: one common convention is to centre them in $+\cdots+$ and to have them on the line in $a.(a+1)\ldots(a+n)$.

Amateurisms are the sort of makeshift that authors are reduced to when their text processing system (or their ability to use it) cannot cope: hand embellishments to make Greek letters, use of $x$ for $x$, $\infty$ for $\infty$, etc.

Blackboardisms are often used by mathematicians who do not realize that what is convenient with chalk may not lend itself to typesetting. The use of double accents, wide accents and overbar vector notation are examples. So is the excessive use of symbolic shorthand.

One common problem that has not so far been mentioned is the problem of 'invisible' bold – insufficient visual distinction between light face and bold. Another is the lack of distinction between hyphen, en-dash, minus sign, em-dash. Another is that many of the symbols required are missing from the ASCII and ANSI character sets.

Watch out for systems (often quite sophisticated word-processors) that don't adequately differentiate $a$ from $\alpha$, $v$ or $\gamma$ from $\nu$, or $\theta$ from $\Theta$ – sharp eyes are sometimes needed.

The problem may lie with the user rather than the system. There is no need to type ug instead of $\mu g$ when $\mu$ is readily available as part of the extended ASCII character set (ASCII 230, ANSI 0181). These hand-embellishments happen with other letters too: $a$ for $\alpha$, $B$ for $\beta$, $r$ for $\tau$. It's particularly unhelpful if you are editing on screen.

References


British Standards 5261 (Copy preparation and proof correction) and BS 5775 Part 11 (Mathematical signs and symbols for use in the physical sciences and technology)