Everything you always wanted to know about LaTeXing an article and I grew tired of telling

version 3.16227

smagt@brml.org

February 20, 2014

You don't know who Donald Knuth is? Stop reading here. This document is not for you. But if you do, you may have more satisfaction from reading the TeXbook than these few lines. Do so, please. Here may be a surprise for you: it is probably on your computer.

1 Language

Beware your spelling bugs. You can squeeze them before giving your document to someone else, it saves a few iterations and many irritations. Note that you can decide whether you use US English or British English (BE) spelling. Some notable differences include the following words:

British English	US English
colour, flavour, humour, etc.	color, flavor, humor, etc.
centre, fibre, etc.	center, fiber, etc.
aluminium	aluminum

but this list is far from complete. See, e.g., a corresponding wiki page.

In BE, you can also choose for the "-ise" rather than "-ize" ending, depending on which dictionary you adhere to. To my best of knowledge, Oxford uses "minimize," "minimization," etc. while Chambers dictates "minimise," "minimising," etc. Statistically, "-ise" is preferred in BE.

An interesting anomaly is that, in US English, punctuation is *always* within quotes. I used that style in this document, but BE does not prefer that anymore.

1.1 ll

"All, till, full, and well,
always have double l
except in compounds."

(So why does TeX write `overfull`? And what about underfull? Overfulfil, or is it overfulfill? Woeful, all of this.)

1.2 a or an?

Everyone knows that "an" is used before a vowel, "a" before a consonant. But that refers to the *sound* that follows, not to the spelling! So please write, "modern physics discovered that the constant indicated by **an** h is no longer relevant."

1.3 Dash

The dash rule in English is really very, very simple (I am talking hyphens here). Generally, adverbs and adjectives are connected by a dash. To help you refresh your grammar: an adjective describes a noun. An adverb describes an adjective.

So, for instance, when we write "the red-beaten man," we refer to a man who was beaten until red. The adverb "red" describes a property of "beaten," not of man. If, however, we write "the red beaten man" (or, more clearly, "the red, beaten man"), the poor chap was red before being beaten.

So, write "the two-dimensional manifold" but "the two dimensional manifolds" (though I would not know what a "dimensional manifold" is). Write the "light-weight robot" (although the whole name is clearly given by individuals who have no understanding of the English language: a weight is not light, but it can be low).

2 LaTeX typesetting

2.1 Commands

A quick one about commands. Many LaTeX and TeX commands take one or more arguments. How does TeX know when a command ends? The rules, which can be redefined, are simple in normal ("plain") mode: a command starts with a \ followed by a single non-alpha character, or one or more alpha characters. So, `\!` is defined as being one character (after the backslash) long only, but `\ashdhf` only ends when the first non-alpha character is seen.

If an alpha command is followed by one or more spaces, these spaces are interpreted as an end-of-command character, and does not render a space in the text. So, when I write `\TeX is great` I don't get what I want.

You can use this rule to your advantage. For instance, `\frac2n` yields exactly the same as `\frac{2}{n}`, but `\fracn2` goes bitterly wrong. In math mode, why not write `$n^2$` rather than `$n^{2}$` (yes, `^` is also a special control character, as are `$#%&{}_~`. Which did I miss?)

2.2 Space control

It is worth understanding the spacing algorithm in TeX. The intricacies go beyond this short section, but the principle is as follows. TeX defines glue between words, sentences, etc. This glue can be compared with a spring, having a base length, which can be compressed by a certain amount, and which can be extended by a certain amount. Normally, two words are separated by some horizontal glue, as are two lines, by vertical glue. TeX then goes on defining the length of a sentence, for instance, and of a page, which also have some flexibility (in the order of picas).

Since the amount of glue is not infinite, there are cases where TeX cannot fill a page or line with the amount of glue available. You may find an example in Sec. 5.1. So, TeX gives up and outputs an error message (`overfull \hbox`).

You really need to scan your log file for such errors. And repair them! There are a few things you can do. The best thing is to reformulate your sentence, or help TeX break words with the `\-` command. If you are lazy, however, you can use `\tolerance10000` in the preamble of the document. For TeX, 10,000 is the same as infinity, and this means that the corresponding glues can be extended up "infinity." You will never end up with overfull hboxes—but you will probably get `underfull` ones!

After a full stop *not* following a capital letter, TeX assumes the end of a sentence and adds more glue (i.e., more stretch). Sometimes you don't want this, and you can prevent this extra

glue with a \ (followed by a space) or a ~ (*not followed by a space*). The latter does not break along lines. To make things more intuitive, this also holds beyond a " or a). So, type `I think, viz.\ I am` to yield "I think, viz. I am" rather than "I think, viz. I am." Do you notice the difference? I do. Depending on how much the space has to be stretched to fill a line, TEX will do that after a full stop.

`\raggedright` and `\raggedbottom` switch off TEX's line and page filling algorithms. The latter is, by default, off in article mode.

To paragraphs. Creating a new paragraph in TEX is simply done by leaving a line blank. The previous line will be terminated, a new paragraph will be started after a length of `\parskip` and indented by `\parindent`. Use this. If you really cannot use the indent on the next line, don't use `\\` but instead write `\noindent` before the next paragraph. Use `\\` only if you *really* want a newline.

Of course, don't create a new paragraph when, for instance, your text goes on with a displayed equation! It will add spurious space, your fault.

2.3 Dashes

Watch the use of dashes. For numerical ranges, use `--`. Use three dashes to isolate—as in this example—a subsentence. In equations, the dash is transformed to a minus. So never write `-1` yielding -1, instead use `$-1$` giving -1.

No spaces around dashes, except preceding a minus sign, of course!

2.4 Dots

Use `\dots` for $1, 2, \dots, n$, I'm sure[1]... But for $1 + 2 + \cdots + n$ you should use `[1+2+\cdots+n]`.

In math mode there are also $\vdots$ (`$\vdots$`), $\ddots$ (`$\ddots$`) and $\ldots$ (`$\ldots$`).

2.5 Quotes

British English prefers single quotes, but you can do double. Whatever you do—be consistent throughout your text!

To type quotes in LATEX, use ‘‘ and ’’ (or, of course, their single variants). The more general `\lq` (for left quote) and `\rq` will render the right quotes in the Babel language that you are typesetting. Alternatively, you can use `"‘` and `"’`—but only if the Babel package is loaded.

3 Figures and tables

Most authors cannot drive LATEX well in placing these. In fact, LATEX offers a very nice set of macros for placing and naming these. You need just write `\begin{figure} ... \end{figure}` and remember that the label must always come *after* the caption, and you are fine. Same for tables, but here the caption should *precede* the table contents.

Now, since the table caption comes before the table and the figure caption after the figure, typesetting rules dictate that, where possible, figures are at the top of a page and tables at the bottom. TEX tries to do that for you, and in most cases you need never to worry about this anymore.

[1] a bit tricky: the sentence ends with `\dots\` to make sure there is a bit of space after the dots.

Please refrain from using these `[hbtp!]` position modifiers if you don't need to. By default, TEX does a very good job in placing them. Also, please refrain from putting a figure in the middle of a column, unless you have a very good reason!

Refer to figures with `Fig.~\ref{...}` and to tables with `Tab.~\ref{...}`.

3.1 Table formatting

Even I am daunted by the range of possibilities to format a table in LATEX. The standard set of macros is not very good, but the number of extensions that are available—even if one only counts the good and useful ones—is challenging. See the below as my personal taste.

The issue of table formatting is, to a large extent, related to taste and design rather than typography—yet the line is thin. A succinct text on the matter can be found behind this link.

Apart from the looks, there are a few packages that you may need when working with more-than-trivial tables in LATEX. These include:

```
\usepackage{longtable} % must be loaded to use {longtabu}
\usepackage{tabu} % more general tabular environment, use longtabu for long tables
\usepackage{numprint} % formats numbers in a nice way
\usepackage{spreadtab} % allows calculations in tables, uses numprint for printing output
\usepackage[table]{xcolor} % alternative to tabu to create coloured cells
\usepackage{booktabs} % to make prettier tables, use \toprule...\midrule...\bottomrule
```

You only need spreadtab if you want to compute values in cells. Tabu gives a good alternative to the myriad of tabularx, tabulary, longtabular, and all of these older versions, plus does many things one does not need. Numprint prints scientific notation, is a good package to pretty-print tables with numbers, and to format the output of spreadtab. xcolor may be needed to colour cells, but tabu does it, too; so if you use tabu, don't use xcolor.

Check the manuals of these packages to enjoy their strength. A separate file `table-example` in this directory gives an example complex table.

4 Citing

There are two things to correct citing in LATEX. First, there is putting the cite in the paper. Quite an easy thing to do, by typing `\cite{ref1,ref2,ref3}`. Note that multiple cites are all included in one cite command.

The BibTEX entry itself must be carefully constructed. Many people like to use bib management programs, such as the public-domain tool JabRef. Fine, but you still must feed it with the right info:

- for a journal publication, you must add title, author, year, volume number, pages, and possibly issue number. Pages are written as `5--10`, but if you write `5-10` in a BibTEX entry, it will automatically translate that to `5--10` for you. Under no circumstance use three dashes!

- for a conference publication, don't add info on where the conference was. The `address` field is to specify the address of the publisher house, and quite irrelevant for these kind of publications.

  If you have them, specify page numbers. Then, of course, title, authors, name of the conference.

5 Equations

5.1 Referring to equations

When having a displayed equation such as

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{r})\psi \tag{1}$$

we can say that thanks to Schrödinger we know Eq. (1) which is typed as `Eq.~(\ref{eq:schroedinger})`.

5.2 Punctuation in equations

These things are easy to answer, as easy as

$$\pi r^2.$$

If an equation, for instance

$$E = mc^2,$$

is part of a sentence, of course normal punctuation must be applied. It makes no sense adding a comma after an equation such as

$$f(\vec{x}) = \sum_i x_i$$

when a comma has no place there.

Oh, of course, punctuation should be within a display equation but outside an inline equation.

But there are some caveats. For instance, the comma in equations is a punctuation symbol. To change that, put it in braces to say, for instance, `$10{,}000|x|$ is not less than $|x|$`.

5.3 Give them space

Quite a tricky business, getting spacing in equations right. But doing it right will tremendously improve the readability of equations, which is key to have your reader follow your writing.

In general, put in a bit of space before a derivative symbol: `$a\,dx/dy$` is right. Also, deleting the space in `$x^2\!/2$` gives $x^2/2$, much nicer than $x^2/2$. As is adding some in `$\sqrt2\,x$` to yield $\sqrt{2}\,x$ (compare $\sqrt{2}x$).

5.4 Units

Units are, in general, written according to ISO style and that means, you use "s" for seconds, "m" for metres, etc. Make them roman (meaning, not slanted, not italics) and separate them from the number by half a space. So you write `I am 1.86\,m long, about 1,451,673,446,400\,ms old, and vibrate at 1\,Hz` to yield "I am 1.86 m long, about 1,451,673,446,400 ms old, and vibrate at 1 Hz" and don't use math mode!

But, we write 75%[2] And 25°C as `$25^\circ$C`[3].

[2] Not quite true—depending on the kind of style you prefer. There are enough style guides preferring 75 %, including the ISO 31-0 standard, while Knuth himself suggests a space, too. The Chicago Manual of Style disagrees. It's up in the air, really!

[3] Here, too, no consensus. 25°C (Chicago style) and 25 °C and 25° C are all acceptable—the latter not really in use anymore. No matter what: if you follow one style, be consistent.

5.5 italics or roman?

By default, equations are built from single-letter symbols, e.g., $f(x) = ax + b$. But sometimes it makes sense to write variables in a more wordy fashion. People sometimes like to write $vel = dpos/dt$. As you can see from the typesetting, this goes horribly wrong! If you really need to write multi-letter variables describing some physical quantity, use `\mathrm`, as in `$\mathrm{vel} = d\,\mathrm{pos} / dt$` to render $\mathrm{vel} = d\,\mathrm{pos}/dt$.

5.6 Multiplying

One sees many symbols for scalars multiplication. There is the `\times` displayed as $\times$, or the `\cdot` aka $\cdot$, sometimes even `\ast`, i.e., $\ast$.

Just don't. For engineers: $\cdot$ is, in general, the inner (or dot) product, $\times$ the outer product, and $\ast$ means convolution[4]. To write a times x, just type `$ax$` yielding ax.

5.7 Fractioning

Sure, you use `$\frac12$` to write $\frac{1}{2}$, while I myself prefer writing `$1\over2$` to obtain the same (note my use of parentheses—or rather, not using them where I need not). That is not worth discussing. But once things get more complex, such as $\frac{a}{\sqrt{1+q}}$, the whole fraction increases in length, and lines get cramped or, even worse, the baseline stretch is increased. At any rate, if your equation is so complex, by all means revert to writing $a/\sqrt{1+q}$. Suggestion: don't use fractions in inline equations, but use the simple / divider. What is easier to read, $\frac{1}{2}$ or 1/2?

5.8 Vectorising

Sometimes it is important to clear the difference between a vector and a scalar. The standard way to indicate vectors is as $\vec{x}$, and that works fine if it is just one old vector lying around. But when a whole bunch is involved, for instance

$$\vec{x}\vec{y} + \vec{b}\vec{a} = 2,$$

the result looks silly. Some people prefer

$$\underline{x}\underline{y} + \underline{b}\underline{a} = 2$$

but, well, judge yourself.

One preferred way out is using boldface symbols. If you do, do it right. Get the $\mathcal{AMS}$math package and write,

```
\usepackage{amsmath}
\newcommand{\vect}[1]{\boldsymbol{#1}} % I would write, \def\vect#1{\boldsymbol{#1}}
```

to render $\boldsymbol{abc\phi\chi}$.

5.9 Collected math typesetting hints

- write `p(x \mid y)` to yield $p(x \mid y)$ rather than `p(x|y)` which gives $p(x|y)$
- write `\|x\|` to yield $\|x\|$ rather than `||x||` which gives $||x||$

[4]But, of course, these are only general conventions and not laws. Indeed, there exist equations where throwing in a `\cdot` to indicate scalar multiplication may actually clarify things.

- sometimes you need to help: cf.
  - `\exp(-(x-y) / (2\pi) )` giving $\exp(-(x-y)/(2\pi))$
  - `\exp\left(-(x-y) / (2\pi) \right)` giving $\exp\left(-(x-y)/(2\pi)\right)$
  - `\exp\bigl(-(x-y) / (2\pi) \bigr)` giving $\exp\bigl(-(x-y)/(2\pi)\bigr)$

  of which the last one is arguably easier to read.