Writing Scientific Documents Using LaTeX

Andrew J. Bennieston

Fifth Edition
February 19, 2009

Contents

1 Introduction

LaTeX is the de-facto standard for typesetting scientific documents. It is an accepted format for submission to a number of scientific journals, as well as many professional publishers. Its popularity stems from a history of flexibility, power and just 'doing the right thing'. LaTeX automatically numbers your equations and figures, lays content out according to strict style rules so as to obtain a readable, polished finish, and rarely gets in the way of the task of writing!

In this document, I introduce many of the LaTeX commands and concepts required to produce a reasonable scientific document of any length, emphasising some of the more complicated aspects while skipping over the simpler things. In some cases, these simple things are silently included as part of a more involved example, so look closely at the example code provided!

2 Starting a LaTeX Document

A LaTeX document must have a preamble (the stuff that goes before
`\begin{document}`). If your LaTeX is split across multiple files, only the top-level file (the one from which you include the others, and run the `latex` program against) needs the preamble. The preamble is where you specify the document class, include packages to provide useful functionality, and define any custom commands you need. At the very least, you must specify a document class.

```
\documentclass[10pt,a4paper]{article}
\usepackage{amsmath,amsthm,amssymb}
```

The LaTeX above starts a document with the *article* document class. Several document classes exist, but *article* serves for most purposes. Two others see some use; these are *book* and *report*. I generally prefer to modify the *article* class to suit my purposes, since it is usually the most sensible. The main argument to the command is enclosed in `{` and `}`, and arguments which are passed on to the *article* class are in `[` and `]`. In this case, they specify to use the A4 paper size and a 10 point font. The `\usepackage` command imports LaTeX style files; in this case the American Mathematical Society styles, which extend and improve the mathematical typesetting provided by LaTeX. Common packages you may want to include are listed below, with short descriptions.

amsmath provides additional mathematical typesetting

amssymb provides mathematical symbols

amsthm provides theorem formatting for AMS publications

fancyhdr provides customisable headers and footers

graphics provides graphics support

graphicx provides extended graphics support (use instead of **graphics**)

hyperref provides navigation support for PDF documents (clickable references, etc.)

url provides clickable URL support

Many of the packages listed require arguments and customisations to function correctly. Each one has online documentation which may be consulted.

The title, authors and date are also usually included in the preamble. This is achieved as follows:

```
\title{My Report}
\author{C. Turk \and J. Dorian}
\date{\today}
```

The contents of those fields can be anything you like — you could use the date field to insert a lecture number as well as a date, for example
`\date{\today -- Lecture 2}`. If the date field is omitted, the current date will be included automatically, but explicitly stating it makes it clearer!

Finally, the document itself is started with `\begin{document}` and ended with `\end{document}`. The title, author and date fields are typeset with the `\maketitle` command, and a table of contents can be included with the cunningly named `\tableofcontents` command. A complete LaTeX document template might look something like this:

```
\documentclass[12pt,a4paper]{article}
\usepackage{amsmath,amssymb,fancyhdr,url}
\usepackage[pdftex]{graphicx}

\title{My Report, Mk. II}
\author{J. Dorian \and B. Kelso}
\date{Before time began\ldots}

\begin{document}
\maketitle
\newpage
\tableofcontents
\newpage

\section{Life at Sacred Heart}
This is the story of \ldots

\end{document}
```

There are a few new things introduced here; sections will be covered later, but let us briefly mention the `\newpage` command, which does exactly what it says on the tin, and the `\ldots` command, which produces: …

3   Sections and Subsections

Sections are started with the `\section{section-title}` command. The section title goes between the `{}`. Similarly, subsections are started with the

`\subsection{sub-title}` command, and sub-subsections with the `\subsubsection` command. By default, all sections and nested subsections are automatically numbered. This behaviour can be turned off on a per-section basis by using the `\section*{section-title}` form, where the `*` tells LaTeX not to number these sections according to the document numbering scheme. The `*` can be used on (almost) anything which would normally be numbered; (sub)sections, equations, etc. Figures and tables use the `*` option in a different way because their naming and numbering is controlled through the `\caption` command.

4 Typesetting Considerations

LaTeX is a little more pedantic than most word-processors, and requires certain punctuation marks to be typeset in slightly different ways. This section discusses the most commonly encountered issues.

4.1 Single Quotes

LaTeX requires left and right quotes, rather than straight quotes. Using the apostrophe, ', always produces a right quote, so you need to use the *backtick*, ', to obtain a left-quote. Quoted text then looks something like *the 'flavour' of a particle* and not like *the 'flavour' of a particle.* The former produces the correct result when typeset, the latter begins and ends with end-quotes!

4.2 Double Quotes

Double left-quotes are input using two single left quotes (the backtick), while double right-quotes are input using two single right quotes (the apostrople). Double quoted text is typeset something like `\ldots said ''the project is on schedule and an investigation has been launched''`, which looks something like this when typeset: ...said "the project is on schedule and an investigation has been launched"

4.3 An Aside on Quotation Marks

In British English, the usual practice is to enclose quoted matter between single quotation marks, reserving double quotation marks for a quotation within a quotation. In US English, the reverse is true. For quotations nested within the second quotation, one should revert to the original mark, alternating accordingly for subsequent levels.

4.4 Hyphenation and Dashes

Hyphens and dashes in printed text have three sizes; the soft hyphen, used to split a word across a line, is the shortest. This is automatically inserted by LaTeX when required. The hard hyphen is the same length as its soft sibling. It is used to hyphenate words within a line, and is typeset using a single dash (- $\Rightarrow$ -). The en-rule (or en-dash) marks a range (often of numbers), such as 3–5. This is slightly longer (one *en* in length), and is typeset using two hyphens (-- $\Rightarrow$ –). Finally, the em-rule (or em-dash) is used to insert a pause in text —

which should probably be replaced by a comma or semicolon anyway — and is the longest (one *em* in length). It is typeset using three hyphens (--- ⇒ —).

4.5   Special Characters

There are some characters which are used by LaTeX as control sequences. These must be escaped to be used in normal text. The characters most often encoutered are \ _ & % { } # (begin command, underscore, table separator, comment, left curly brace, right curly brace, hash or "pound" sign) and they can be produced in the text verbatim by 'escaping' with a leading backslash: _ \& \% \{ \} \#, with the exception of the backslash, which can be inserted with the \textbackslash command: \.

4.6   Emphasising Text

To place emphasis on a certain phrase, use the \emph{something} command. This usually produces *italic* text, but the effect depends on the document style used. The advantage of using \emph is that emphasised text within an already emphasised section is typeset in Roman (non-italic) text. For example, The \emph{President of the \emph{United States} of America} produces: The *President of the* United States *of America.*

In addition, it is possible to directly specify the typeface required. The command \textit{italic text} produces *italic text*, while \textbf{bold text} produces **bold text** and \texttt{monospaced text} produces monospaced text.

5   Mathematical Typesetting

LaTeX provides several environments for mathematics and I will not cover them all here. Instead, I will mention those which are most important in short documents and reports.

5.1   Inline Mathematics

To insert mathematics inline with the paragraph text you can use the $...$ environment. This is achieved as follows:

```
The solar neutrino flux was found to be
$2.1 \pm 1$ solar neutrino units (SNU).
```

When typeset, this looks like: The solar neutrino flux was found to be 2.1 ± 1 solar neutrino units (SNU).

The use of $...$ to insert math-mode LaTeX inline with the paragraph text allows for many $effects$; I used it to produce the right-arrows ($\Rightarrow$) in the hyphenation section, above.

5.2   Display Mathematics

The display mathematics environment sets mathematical content on a line of its own, centred on the page. The shortcut form is the \[...\] environment,

which enters display mode but does not add equation numbering.

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

The Schrödinger equation, above, was typeset using:

```
\[ \frac{\hbar^2}{2m}\nabla^2\Psi + V(\mathbf{r})\Psi
= -i\hbar \frac{\partial\Psi}{\partial t} \]
```

To get equation numbering, one has to be slightly more verbose:

```
\begin{equation}\label{rw-metric}
ds^2 = c^2 dt^2 \left( \frac{d\sigma^2}{1-k\sigma^2}
+ \sigma^2\left[ d\theta^2 + \sin^2\theta d\phi^2 \right] \right)
\end{equation}
```

. . . which looks like:

$$ds^2 = c^2 dt^2 \left(\frac{d\sigma^2}{1 - k\sigma^2} + \sigma^2 \left[d\theta^2 + \sin^2\theta d\phi^2 \right] \right) \tag{1}$$

Note that the `\label` command was used here, which sets a label that can used in references to that equation. For example, typing `\eqref{rw-metric}` at any point in the document will yield: (1). As with sections, using the "starred form" `\begin{equation*}` suppresses equation numbering (which is precisely what `\[...\]` does).

5.3 The Align Environment

To typeset equations over multiple lines, use the *align* environment, of the *amsmath* package. Here is an example with the *align* environment:

```
\begin{align}
X(t) = 1 + \frac{t^3}{2t} \\
= 1 + \frac{t^2}{2}
\end{align}
```

$$X(t) = 1 + \frac{t^3}{2t} \tag{2}$$

$$= 1 + \frac{t^2}{2} \tag{3}$$

The ampersand character, &, is used to set the alignment points. Note that using the *align* environment does not squash fractions into a single line height! Line numbers can be suppressed with `\nonumber`, and the environment and *align** provides an unnumbered version of the same functionality.

5.4 Symbols

There is a lot of documentation on the LaTeX mathematical command set. For the most part it is obvious; Greek letters are produced using `\lettername`, for example γ is `$\gamma$`. Uppercase Greek letters are obtained by capitalising the first letter: Ω is `$\Omega$`...

There are also a variety of arrows and other symbols: `$\rightarrow$` produces →. Many more exist, and the *Comprehensive LaTeX Symbol List*[1] covers

---

[1] http://www.ctan.org/tex-archive/info/symbols/comprehensive/symbols-a4.pdf

a huge number of mathematical, scientific and other symbols.

Summation symbols can be inserted either with `\sum` or `\Sigma`, though `\sum` is preferred, since it scales as required. Integrals can be represented using `\int` and variants for double integration, closed loop integration, etc.

```
\[ \sum_{i=1}^N \rightarrow \int_1^N\,\mathrm{d}i \]
```

$$\sum_{i=1}^N \rightarrow \int_1^N \mathrm{d}i$$

5.5 Fractions

Fractions can be typeset in the form 1/2: `$1/2$`, or with horizontal lines $\frac{1}{2}$: `$\frac{1}{2}$`. Fractions and many other symbols are sized differently depending on whether the equation is inline (part of the paragraph text) or displayed. To use the display sizes for inline symbols, add the `\displaystyle` command to the inline equation: `$\displaystyle\frac{1}{2}$` renders: $\displaystyle\frac{1}{2}$

5.6 Brackets

Brackets can, for the most part, be inserted normally. To scale the bracket (in displayed equations) so that it fits around the content, use `\left` and `\right` before the bracket character itself; for an example see equation (1), the Robertson-Walker metric. Angle brackets such as those used in Dirac notation can be produced with `\langle` and `\rangle`: `$\langle\psi|\hat{H}|\psi\rangle$` produces $\langle\psi|\hat{H}|\psi\rangle$. Again, `\left` and `\right` can be used where necessary. In order to suppress a bracket on either the left or the right, use `\left.` or `\right.` to prevent LaTeX complaining about unbalanced brackets.

5.7 Differentials

Differentials can be typeset in a number of ways. The simplest method is to use `\frac{dy}{dx}`: $\frac{dy}{dx}$. Some people like the d to be typeset in Roman (rather than italic) or bold font: `\frac{\mathrm{d}y}{\mathrm{d}x}` or `\frac{\mathbf{d}y}{\mathbf{d}x}`, producing $\frac{\mathrm{d}y}{\mathrm{d}x}$ and $\frac{\mathbf{d}y}{\mathbf{d}x}$ respectively.

Partial derivatives can be typeset using the `\partial` symbol in place of the d: `\displaystyle\frac{\partial y}{\partial x}` which gives $\displaystyle\frac{\partial y}{\partial x}$. The various vector differentials can be typeset using `\nabla` and decorations: ∇^2, $\nabla \times \mathbf{A}$.

5.8 Matrices

The *array* environment, used within one of the mathematical environments, can be used to typeset row and column vectors, matrices, lists of equations, etc. One must specify a column alignment for each column in the array, then use it just as the *eqnarray* environment:

```
\begin{equation}
\mathbf{V}_{CKM} = \left(
\begin{array}{ccc}
```

```
V_{ud} & V_{us} & V_{ub} \\
V_{cd} & V_{cs} & V_{cb} \\
V_{td} & V_{ts} & V_{tb}
\end{array}
\right)
\end{equation}
```

$$\mathbf{V}_{CKM} = \left(\begin{array}{ccc} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{array} \right) \tag{4}$$

Determinants can be typeset in a similar way by using `\left\vert` and `\right\vert`:

$$\det \mathbf{I} = \left| \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right| = 1 \tag{5}$$

To omit the brackets on one side, use `\left.` or `\right.`:

```
\begin{equation}\label{sim-eq}
\left.
\begin{array}{c}
x + y = 3 \\
x^2 + y = 5
\end{array}\right\} \Rightarrow x=2~,~y=1
\end{equation}
```

$$\left. \begin{array}{c} x + y = 3 \\ x^2 + y = 5 \end{array} \right\} \Rightarrow x = 2 \ , \ y = 1 \tag{6}$$

5.9 Math-mode Spacing

It is possible to add spaces of varying lengths to your mathematical typesetting. In the simultaneous equations example, equation (6), I used `~,~` to add spaces either side of the comma. I also typeset integrals using `\int f(x)\,dx`, where the `\,` adds a small space before the dx symbol:

```
\[\langle\psi|\psi\rangle =
\int_{-\infty}^{\infty} \psi^\star(x)\psi(x) \, dx\]
```

$$\langle\psi|\psi\rangle = \int_{-\infty}^{\infty} \psi^\star(x)\psi(x) \, dx$$

6 Graphics

Graphics can be inserted into LaTeX documents from a number of formats. Depending on whether you use `latex` or `pdflatex`, the recommended formats are Encapsulated PostScript (EPS) and Portable Document Format (PDF), respectively. These formats are good because they are scalable, though non-scalable bitmap formats are supported, including JPEG. The *graphicx* environment allows inclusion of such graphics, as well as modification in-place (rotating, scaling, etc.) Full documentation for *graphicx* is available online.[2]

---

[2]`http://www.ctan.org/tex-archive/macros/latex/required/graphics/grfguide.pdf`
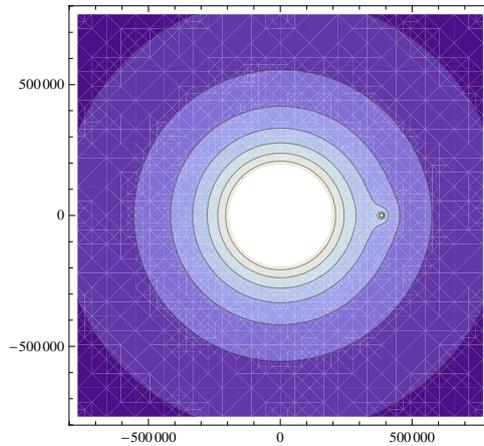
Figure 1: The combined gravitational potentials of the Earth and Moon.

```
\begin{figure}
\centering
\includegraphics[scale=0.75]{earth-moon.pdf}
\caption{\label{earth-moon}The combined gravitational
potentials of the Earth and Moon.}
\end{figure}
```

The `\centering` command causes the contents of the figure to be centred (horizontally) on the page. The `\label` command is used to create a reference to the figure so that we can later refer to it in the text using the command `Figure \ref{earth-moon}`, i.e. Figure 1. Note also that the command `\includegraphics` takes arguments which allow us to manipulate the image; here we've scaled it to 75% of its original size. The `\caption` command should be fairly self-explanatory!

Note that this document was produced using the *graphicx* package; the *graphics* package requires scaling and rotating to be done in a different way, but since almost all sites now have the *graphicx* package installed, I'm going to ignore the *graphics* package here.

LaTeX will try to position graphics on a page according to a set of rules about the minimum number of lines of text on a page, and other typographic concerns. If it cannot satisfy all of these requirements the figure will be placed on a page of its own. It is possible to provide LaTeX with hints for positioning a figure, given as options to the *figure* environment. The hint is provided in square brackets after the environment is opened, for example `\begin{figure}[t]` to request that a figure is placed at the top of a page. The following placement specifiers are understood by LaTeX:

h places the floating figure approximately at the point it occurs in the source text.

t places the floating figure at the top of the page.

b places the floating figure at the bottom of the page.

Natural	$\mathbb{N}$
Integer	$\mathbb{Z}$
Real	$\mathbb{R}$
Rational	$\mathbb{Q}$
Complex	$\mathbb{C}$

Table 1: A table!

p places the figure on a special page for floated items only.

! overrides LaTeX rules for determining figure positions.

H places the floating figure at precisely the point it occurs in the source text, and requires the *float* package.

7 Tables

To insert a table into a LaTeX document, you must enter the *table* environment, then define your table as required, using the *tabular* environment!

```
\begin{table}
\centering
\begin{tabular}{|c|c|}
\hline
Natural & $\mathbb{N}$ \\
\hline
Integer & $\mathbb{Z}$ \\
\hline
Real & $\mathbb{R}$ \\
\hline
Rational & $\mathbb{Q}$ \\
\hline
Complex & $\mathbb{C}$ \\
\hline
\end{tabular}
\caption{\label{my-table}A table!}
\end{table}
```

The *tabular* environment acts much like the *array* math environment, and the | character instructs LaTeX to draw a vertical line between columns. You can use the \hline command to draw horizontal lines. Such a table can, of course, be referred to using the \ref{my-table} command: table 1.

The *booktabs*[3] package documentation includes a number of points for correctly typesetting tables, particularly for professional book layouts.

8 Including LaTeX Files

If your document is split across several files, they can be included with the command \include{filename}. The .tex extension is not required — LaTeX

---

[3]ftp://ftp.tex.ac.uk/tex-archive/macros/latex/contrib/booktabs/booktabs.pdf

adds it automatically. `\include` will add a page-break before the contents of the included file.

The `\include` command will not work inside a file which has already been `\include`d by another file. The `\input` command is also useful.

9 Bibliographies & Citations

9.1 Bibliographies & References

A simple bibliography can be created as follows:

```
\begin{thebibliography}{99}
\bibitem{Fey49} R. P. Feynman,
Space-Time Approach to Quantum Electrodynamics,
Phys. Rev. \textbf{76} 6 (1949)
\bibitem{Cas48} H. B. Casimir, On the attraction
between two perfectly conducting plates,
Proc. K. Ned. Akad. Wetensch. \textbf{51} (1948)
\end{thebibliography}
```

The `{99}` tells LaTeX how many digits to reserve for bibliography reference numbers — twi, in this case. If you have over 100 references, use a three-digit number.

The section title will depend on the document class used; in the *article* class, it is 'References', in the *book* or *report* class, it is 'Bibliography'. You can use `\begin{thebibliography*}` to suppress a title and use your own.

9.2 Citations

Citing bibliography items is achieved through the use of the `\cite{label}` command. This adds the corresponding reference number in square brackets, at that point in the text.

9.3 BibTeX

For more complicated bibliographies, or larger documents, the BibTeX system allows a lot more flexibility and control over the presentation of bibliographic items. Documentation is available online.[4]

10 Processing LaTeX Documents

Once a LaTeX document has been written, it must be processed by LaTeX to produce a file in the required format (e.g. DVI, PDF, PS). I prefer the `pdflatex` command, which produces PDF output directly, and supports hyper-linked PDFs, where the relevant packages (*url* and *hyperref*) have been used.

The `pdflatex` command should be run at least twice, since LaTeX passes through once, enumerating the sections, equations, figures, tables and references, then the second run substitutes in the correct numbers at each point

---

[4]`http://www.bibtex.org/`

they are referenced. In this way, LaTeX is able to produce the correct number for references which have not been defined prior to their point of use (e.g. bibliography items and forward-references to figures or equations). The command I run is usually something like the one below, for a file called `report.tex`

```
pdflatex report && pdflatex report && pdflatex report
```

It is possible to create a DVI file by processing the document with the `latex` command. Again, you need to run LaTeX twice or more to get all of the correct references, and a DVI can be converted to PostScript or PDF using `dvips` or `dvipdf`. Directly producing PDF through the `pdflatex` command is, however, my preferred method.

A final point about processing the documents is that BibTeX requires the `bibtex` command to be run between calls to `pdflatex` (or just `latex`). I put it after the first call to `pdflatex`:

```
pdflatex report && bibtex report &&
pdflatex report && pdflatex report
```

11 About This Document

This document is copyright © 2007–2009 Andrew J. Bennieston. It may be freely distributed according to the terms of the LaTeX Project Public License (details below). The document was produced in the Vim text editor on the OpenSUSE Linux operating system. Earlier versions were produced on an Apple Macbook using the TextMate text editor, and the Slackware Linux operating system using the Vim text editor.

11.1 License

This work may be distributed and/or modified under the conditions of the LaTeX Project Public License, either version 1.3 of this license or (at your option) any later version. The latest version of this license is in `http://www.latex-project.org/lppl.txt` and version 1.3 or later is part of all distributions of LaTeX version 2005/12/01 or later.

This work has the LPPL maintenance status 'maintained'. The Current Maintainer of this work is Andrew J. Bennieston. This work consists of the files scidoc.tex, Makefile, README, earth-moon.pdf and the derived file scidoc.pdf.

11.2 History

- First Edition: February 4, 2007

- Second Edition: March 8, 2007

- Third Edition: June 13, 2008

- Fourth Edition: September 13, 2008

    - Licence changed to LPPL (September 16, 2008).

- Fifth Edition: February 18, 2009

    - Minor corrections (February 19, 2009)