Notes on Scientific Writing

Carsten Burstedde

October 19, 2014

1 Organization

Diagrams require axis tics and units at all times.

The caption to a figure or a table should contain detailed information about how precisely it was obtained and what choices have been made (parameters, data, conventions). Text that explains how its information is to be decoded should go into the caption as well, and anything that refers to the colors, styles, and the legend. The caption can summarize the content of the figure, but should stay neutral and objective. The text, on the other hand, should not elaborate on the display of the figure/table but embed its meaning and relevance into the flow of the story, including the opinion of the author and a discussion on interpretation and judgement. The idea is that both caption and document text should be readable on their own, without requiring the reader to jump back and forth between the two. They should complement but never replicate each other.

Each and every variable must be introduced in the text whenever it occurs in a formula for the first time. The text must state how different formulas relate to each other, and which variables are known and unknown.

Carefully derived and well explained formulas help the reader understand better than an algorithm that is not backed by mathematics.

2 Language

Names are capitalized: Euclidian space, Jacobian matrix.

The meaning of that and which is different. Use that (without a preceding comma) for a definition that adds essential information to identify the subject. Use which (with preceding comma) to add additional information, which could also be omitted without changing the definition of the subject. Thus: I can see two cats. The cat that is white belongs to me. The cat (that is) sitting on the right hand side, which is white, belongs to me.

Avoid words of unclear definition like "tiny" and "huge," and plump words of judgement like "good" and "bad."

References to figures or citations can constutite a separate sentence, which is fine in parentheses (see here) but otherwise requires a semicolon instead of a comma; see here.

In English texts, use triple dashes—which signal an embedded sentence—like this. In German, use double dashes – surrounded by spaces – in this case.

A vector space has a basis. The base 2 logarithm of 1024 is 10.

Formulas displayed on one or more separate lines are part of the current sentence and need to be punctuated as if they were regular text. For instance, we know that

$$0 + a = a,$$

whereas it is not generally true that

$$0 + a = 1.$$

3 LaTeX

In LaTeX, a trailing dot introduces extra end-of-sentence rubber spacing. This can look extremely ugly, and demonstrates like nothing else that the writer is no LaTeX expert. In the middle of a sentence, use for example `e.g.\` to achieve regular spacing. Compare e.g. with e.g. with respect to the trailing space: The first one creates too much since it's missing the backslash in the source code.

Spaces after macros are eaten up by LaTeX. If required, add a pair of empty braces {} after the macro, as demonstrated here with LaTeX as the macro (see the source code).

Mathematical operators should be set in roman font. For the builtin LaTeX operators this happens automatically as in $\cos x$. If a function is not yet defined, this can be accomplished with the command `\DeclareMathOperator` from the package `amsmath`. This looks like $a = \operatorname{const}$, $\operatorname{span} x$, or $\operatorname{range}(A)$, for example. This package is recommended in general; see the user guide on the web. There is really no need to use `\eqnarray` anymore.

Words that are used in math mode, for example as sub- and superscripts, should be set inside the `\mathit` or `\mathrm` commands. Otherwise, the individual letters are interpreted as single-letter mathematical variables (like in a long multiplication). Compare for example $a_{notsonice}$ with $a_{verynice}$ and $a_{\mathrm{niceaswell}}$. Another example: *terrible* vs. *better*.

Source

```
\documentclass[a4paper,11pt]{article}

\usepackage[margin=1in]{geometry}

\usepackage{amsmath}
\usepackage{verbatim}

\title{Notes on Scientific Writing}
\author{Carsten Burstedde}

\DeclareMathOperator{\const}{const}
\DeclareMathOperator{\range}{range}
\DeclareMathOperator{\spann}{span}

\begin{document}
\maketitle

\section{Organization}

Diagrams require axis tics and units at all times.

The caption to a figure or a table should contain detailed information about
how precisely it was obtained and what choices have been made (parameters,
data, conventions).  Text that explains how its information is to be decoded
should go into the caption as well, and anything that refers to the colors,
styles, and the legend.  The caption can summarize the content of the figure,
but should stay neutral and objective.  The text, on the other hand, should not
elaborate on the display of the figure/table but embed its meaning and
relevance into the flow of the story, including the opinion of the author and a
discussion on interpretation and judgement.  The idea is that both caption and
```

document text should be readable on their own, without requiring the reader to jump back and forth between the two. They should complement but never replicate each other.

Each and every variable must be introduced in the text whenever it occurs in a formula for the first time. The text must state how different formulas relate to each other, and which variables are known and unknown.

Carefully derived and well explained formulas help the reader understand better than an algorithm that is not backed by mathematics.

\section{Language}

Names are capitalized: Euclidian space, Jacobian matrix.

The meaning of that and which is different. Use that (without a preceding comma) for a definition that adds essential information to identify the subject. Use which (with preceding comma) to add additional information, which could also be omitted without changing the definition of the subject. Thus: I can see two cats. The cat that is white belongs to me. The cat (that is) sitting on the right hand side, which is white, belongs to me.

Avoid words of unclear definition like ``tiny'' and ``huge,'' and plump words of judgement like ``good'' and ``bad.''

References to figures or citations can constutite a separate sentence, which is fine in parentheses (see here) but otherwise requires a semicolon instead of a comma; see here.

In English texts, use triple dashes---which signal an embedded sentence---like this. In German, use double dashes -- surrounded by spaces -- in this case.

A vector space has a basis. The base 2 logarithm of 1024 is 10.

Formulas displayed on one or more separate lines are part of the current sentence and need to be punctuated as if they were regular text. For instance, we know that
\begin{equation*}
  0 + a = a ,
\end{equation*}
whereas it is not generally true that
\begin{equation*}
  0 + a = 1 .
\end{equation*}

\section{\LaTeX}

In \LaTeX, a trailing dot introduces extra end-of-sentence rubber spacing. This can look extremely ugly, and demonstrates like nothing else that the writer is no \LaTeX{} expert. In the middle of a sentence, use for example \texttt{e.g.\textbackslash} to achieve regular spacing. Compare e.g. with

e.g.\ with respect to the trailing space: The first one creates too much since it's missing the backslash in the source code.

Spaces after macros are eaten up by \LaTeX. If required, add a pair of empty braces \{\} after the macro, as demonstrated here with \LaTeX{} as the macro (see the source code).

Mathematical operators should be set in roman font. For the builtin \LaTeX\ operators this happens automatically as in $\cos x$. If a function is not yet defined, this can be accomplished with the command \texttt{\textbackslash DeclareMathOperator} from the package \texttt{amsmath}. This looks like $a = \const$, $\spann x$, or $\range(A)$, for example. This package is recommended in general; see the user guide on the web. There is really no need to use \texttt{\textbackslash eqnarray} anymore.

Words that are used in math mode, for example as sub- and superscripts, should be set inside the \texttt{\textbackslash mathit} or \texttt{\textbackslash mathrm} commands. Otherwise, the individual letters are interpreted as single-letter mathematical variables (like in a long multiplication). Compare for example $a_{notsonice}$ with a_verynice and $a_\mathrm{niceaswell}$. Another example: $terrible$ vs.\ better.

\section*{Source}
\verbatiminput{scientificwriting.tex}

\end{document}