How To Write a Paper and Format it Using LaTeX

Jennifer E. Hoffman[1, 2, *]

[1] *Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*
[2] *School of Engineering & Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*
(Dated: August 14, 2018)

The goal of this document is to demonstrate how to write a paper. We walk through the process of outlining, writing, formatting in LaTeX, making figures, referencing, and checking style and content.

I. GETTING STARTED

1. Read George Whitesides' "How to Write a Paper" (Ref. 1). If you've read it before, read it again.
2. Read through at least one full paper in your target journal, to get a sense of the content and writing style.
3. It's probably easiest to begin the outlining process in a wysiwyg editor such as Word, where you can include rough figures and see everything at a glance.
4. To clarify in your own head the purpose of your paper, start by writing your abstract.
5. Before you tackle the body of the paper, sketch block outlines of the figures. Decide what images and plots you will put in the paper, and how the panels will be arranged.
6. Outline at the paragraph level before you write. Look at how many paragraphs there will be in the style of paper you are trying to write. For PRL, aim for 13 paragraphs. Figure out how to tell your entire story (not numbers, just story) in 13 stand-alone sentences.
7. Make each of those sentences into the first sentence of a paragraph, and fill into each paragraph only details that are relevant to that first sentence. If you find yourself writing details about the figures, cut and paste them into the captions.
8. If you think of references as you go, just include the minimal identifying information in parentheses to trigger your memory later, e.g. "(WhitesidesAdvMat)", so all of the information is compact.
9. Rewrite your abstract, taking into account what you have learned from the process of writing the paper.

As you contemplate the paper you have just written, put yourselves in the shoes of the reviewers (including your collaborators). You already work many, many hours/week, and you don't really want to spend more time reading this paper. So you're going to be very happy if the figures are pretty, the text flows logically, and you can understand the paper quickly. But you're going to be very grumpy if you can't get the main points of the paper from scanning through the figs & captions. You're going to be even grumpier if you invest time in reading the paper but you still can't get it. Your evaluation of this paper is likely to be swayed by your happy-grumpy level, regardless of the scientific merits of the paper.

Down the road, consider a reader who might cite the paper and launch you to fame and glory: the decision will be influenced by the potential citer's ability to easily understand your paper.

Your paper should be fractal. Somebody with 30 seconds to look at it should be able to get the main idea just from reading the abstract. Somebody with 5 minutes should be able to look at the figures and captions and get more out of it. Somebody with 10 minutes should be able to get the story from the first sentence of each paragraph.

II. CONVERTING TO LaTeX

After you've received high-level feedback on your outline, it's time to convert to LaTeX for easier readability, i.e. compact presentation of text and figures, with hyperlinks to all references.

1. Download & install a LaTeX compiler, e.g. http://miktex.org/
2. Download & install REVTex 4-1 from here: http://authors.aps.org/revtex4/
3. Choose & install a reference management software such as:

   - Mendeley: http://www.mendeley.com/
   - Papers: http://guides.library.harvard.edu/cite/papers
   - Endnote: http://guides.library.harvard.edu/cite/endnote

4. Pick an editor such as:

   - WinEdt: http://downloads.fas.harvard.edu/download (You will need to request a license; typically requires 24 hour turnaround.)
   - VI: http://www.ccsf.edu/Pub/Fac/vi.html

Alternative to the steps above: use the online latex compiler such as ShareLaTeX (http://www.sharelatex.com). Note that LyX (http://www.lyx.org/) doesn't seem like a great alternative, because it hides much of the guts of the tex code from the user, and makes it harder to control formatting details.

Download the paper template from http://hoffman.physics.harvard.edu/policies/example-paper/.

III.   FORMATTING

Whether you are using a compiler on your computer or online, please use the latest version of REVTeX, and check your formatting carefully.

☐ Check math & symbolic formatting, as in Tab. I.

TABLE I. Formatting mathematical symbols.

Incorrect	Correct
$cos\theta$	$\cos\theta$
T_{sample}	T_{sample}
V$_{rms}$, V (rms)	V_{rms}
E_{x}, x direction	E_x, x direction
$\vec{B}_{\mathrm{app}}$	$\vec{B}_{\mathrm{app}}$
Sb_2Te_3, Sb2Te3	$\mathrm{Sb_2Te_3}$
$\mathrm{Sb_{2-x}V_xTe_3}$	$\mathrm{Sb_{2-x}V_xTe_3}$
dI/dV	dI/dV
$B = 5T$, B=5T	$B = 5$ T
x direction, X direction	x direction
0th, 0th, 1st, 1st	0th, 1st

☐ Use \label{tab:name} and \ref{tab:name} to refer to tables.

☐ Check hyphenation. Sometimes LaTeX likes to divide a single letter off the beginning or end of a word, for line wrapping. The default settings for LaTeX's hyphenation of English-language words are \righthyphenmin=3 and \lefthyphenmin=2, but apparently they can be mysteriously reset to allow single dangling letters.

☐ Check spacing. When a period falls in the middle of a sentence, use a \ (backslash) to prevent LaTeX from thinking it's the end of the sentence and thus adding extra space. If you want to prevent a linebreak, you can use ~ instead of \.

TABLE II. Spacing.

	LaTeX	Output
Incorrect	`e.g. incorrect`	e.g. incorrect
Incorrect	`Fig. 2`	Fig. 2
Correct	`e.g.\ correct`	e.g. correct
Correct	`Fig.\ 2`	Fig. 2
Correct	`Fig.~2`	Fig. 2

☐ Number all equations. But do not separately number each line of a single multi-line equation.

☐ Use \label{eqn:name} and \ref{eqn:name} to refer to equations.

☐ If the equation is mid-paragraph, use \noindent at the beginning of the first line following the equation.

Here is an improperly-labeled equation in the middle of a paragraph.

$$1 + 1 = 2$$

Use \noindent to prevent indentation mid-paragraph. Here is a more interesting example of a properly labeled equation: the Pythagorean theorem relates the 3 sides of a right triangle according to Eqn. 1

$$a^2 + b^2 = c^2 \,. \tag{1}$$

Eqn. 2 shows one more example of a multi-line equation extending the Pythagorean theorem to find the diagonal d of a rectangular prism of sides $a = 3$, $b = 4$, and $c = 12$.

$$d = \sqrt{a^2 + b^2 + c^2}$$
$$= \sqrt{3^2 + 4^2 + 12^2} = 13 \,. \tag{2}$$

IV.   STYLE

☐ Use the active voice. If you use the passive voice, it is very hard to tell the difference between what you personally worked hard to do in your experiment, vs. what you are citing as a prior result.

☐ Avoid pronouns if at all possible (e.g. not "that" but "that voltage signal").

☐ Use adjectives if there is any doubt (e.g. not "the modulation" but "the z modulation").

☐ Define all acronyms and symbols at first use; then use the acronym consistently from that point on.

☐ Do not use the word "significantly" - unless you mean it in the true statistical sense and are prepared to back it up quantitatively.

☐ Remove redundancy, including redundancy between main text and figure captions. When in doubt, the information probably belongs in the caption but not the main text.

☐ Check that all equations are dimensionally correct.

☐ For each equation, define all symbols in a previous equation or in the surrounding text.

☐ Report each quantity consistently throughout the text (e.g. don't exaggerate a quantity in one place, give an exact version of the same quantity in another place, and round to the nearest 100 in a third place).

☐ Check that all numbers have units.

☐ Use reasonable significant figures, report errors where appropriate, and clearly explain the method of error determination.

☐ When in doubt, check examples in the journal (e.g. if you plan to submit to APL and you wonder whether acronyms are appropriate in the abstract, check a few recent published examples).

☐ "Only" can be an adjective or an adverb, so its meaning can be ambiguous. It should be placed immediately before the noun, adjective, or verb that it is modifying. For example "I only bought groceries at the store" means I didn't run, jump, dance, or sing

at the store, I only bought. But "I bought only groceries at the store" means I didn't buy STMs, valves, or screws at the store.

☐ "Its" is possessive;
  "it's" is a contraction of "it" and "is".

☐ "which" vs. "that": http://www.writersdigest.com/online-editor/which-vs-that

☐ "fewer" vs. "less": http://www.quickanddirtytips.com/education/grammar/less-versus-fewer

☐ "affect" vs. "effect": http://grammarist.com/usage/affect-effect/

☐ See additional tips from Margo Seltzer: http://www.eecs.harvard.edu/margo/writing.html

V. FIGURES

Figures should typically be made in Python, Adobe Illustrator, or other program that allows vector format export, so that all fonts, arrows, etc. will scale cleanly when zoomed. Most journals prefer to stay away from Microsoft Powerpoint (although it can be exported to eps or pdf) because the fonts are often not transcribed correctly in publication format. A bigger problem with Microsoft is that it does not faithfully reproduce the pixelation of data images. Microscope images are acquired with a specific pixel resolution, and that pixelation should be honestly communicated to the reader without interpolation. Fig. 1 illustrates this point.
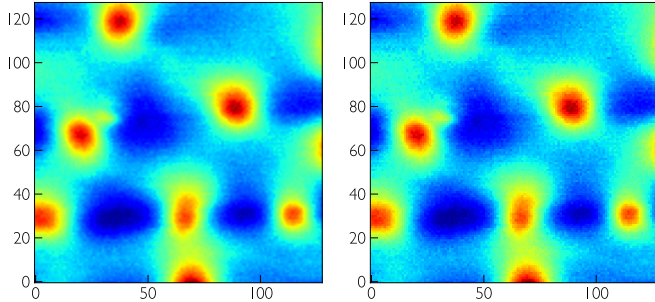

FIG. 1. Comparison between blurry pixels (dishonest interpolation occurs when the image is processed in Microsoft Powerpoint) vs. clean pixels (honest representation is preserved when the image is processed in Python and Adobe Illustrator). MFM images of vortices in $NdFeAsO_{1-x}F_x$ (Ref. 2).

Note that faithful representation of images in vector format usually also results in a smaller figures size. This can be important, because the arXiv places an upper bound of 5 MB on each submitted manuscript.

All figure fonts should be at least size 8 in the final publication figure. Note also that san-serif fonts are preferred by most journals (e.g. Arial, Helvetica). To achieve the appropriate font size, please start by measuring the desired final figure size (e.g. one or two column width) in the desired journal. Then make a box in AI (or other program) of exactly the final size, and build your figure within it, using no fonts smaller than size 8. Although

some journals do prefer that you initially submit your figure at full-page size, you can easily scale up your figure for this purpose. But if you start with a page-size figure and arbitrary font sizes, it becomes harder to later scale it down while maintaining adequate font size.

Figure checklist:

☐ Use consistent sans-serif font, at least size 8.

☐ Label all axes, with units.

☐ Each plot should have a legend that describes *all* symbols and lines.

☐ Each image (or set of same-scale images) should have an accurate length scalebar, with numerical label. (Note that some journals discourage or "forbid" superimposing the numerical length on the image. But our goal is clarity: we want the reader to understand the image at a glance, without digging through a lengthy caption to find the necessary number. Journals will generally accept this argument for keeping the number on the image.)

☐ Each image (or set of same-palette images) should have a colorbar. The colorbar should be labeled with numerical values and units if possible.

☐ If using a waterfall plot to display a set of spectra: clearly state the offset of the waterfall plot, or use small horizontal lines to denote the true zero reference for each individual spectrum.

☐ The caption should describes all figure sub-parts, in order. Each and every mark on the figure should be described; there should be no mysterious unexplained arrows or other features.

☐ If any analysis has been performed (i.e. if it's not raw data), then all analysis steps should be clearly divulged, usually in the caption (rather than main text).

☐ Clearly explain the origin of all error bars, usually in the caption (rather than main text).

☐ For STM images: give setup conditions in the caption ($V_{sample} = 100$ mV; $R_J = 1$ GΩ).

☐ For STM spectra: give sample bias modulation in the caption ($V_{rms} = 2$ mV).

☐ For all data: clarify temperature and field conditions in the caption.

☐ Appropriately cite all copied figures or data, in the caption of the figure.

☐ Use \label{fig:name} within the caption, and use \ref{fig:name} in the paper to refer to it.

VI. REFERENCES

Referencing should be done using BibTeX.

☐ Consistently use reference tags that will be easily recognizable and editable from the tex file, e.g. AuthorJournalYear. Suppose you will be citing Huang *et al*, Nano. Lett. 16, 4224 (2016) (Ref. 3) and Huang *et al*, Phys. Rev. B 93, 125129 (2016) (Ref. 4). Instead of Huang2016a and Huang2016b, use

`HuangNanoLett2016` and `HuangPRB2016`. (Note: You can make these bib tags within Mendeley.)

☐ Spend 5 minutes to use find-replace to delete the abstracts, keywords, and other useless info from your bib file, to make it easier to read.

☐ Alphabetize the bib entries by author last name, so that it will be easy to notice if there are duplicates. (Note: Mendeley can automatically export them in alphabetical order.)

☐ The hyperlink should be generated from the doi, so make sure every bib entry has a valid doi. You should delete the URL field (unless you are specifically citing a website), because it may contain a non-general URL such as `https://journals-aps-org.ezp-prod1.hul.harvard.edu/prb/abstract/10.1103/PhysRevB.93.125129`.

☐ Check the LaTeX formatting of any special characters in the author's names, e.g. `S\'{a}nchez`

☐ Check the LaTeX formatting in the titles. You need to use extra {} around any letters that should be capitalized, e.g. `title = {Quantum Anomalous {H}all Effect}`

☐ Chemical formulas should also be enclosed in {} so element abbreviations will be capitalized, e.g. `title = {Experiments on {Bi$_2$Se$_3$}}`. Don't use excessively complicated formatting for chemicals. Just put the subscripts in $$.

☐ Use `longbibliography` in the document options of the tex file, so that the title of each paper appears in the references.

☐ For most citations, you can just use `\cite{AuthorJournalYear}`, e.g. Jeehoon used conducting force microscopy to measure a metal-insulator transition[5]. But if the citation comes directly after a number or chemical formula, use `Ref.\ \onlinecite{AuthorJournalYear}` instead to avoid confusion, e.g. Jessie measured the pinning force on vortices in $NdFeAsO_{1-x}F_x$ (Ref. 2).

☐ In the compiled PDF, check all the references carefully to make sure they have correctly formatted authors, titles, and hyperlinks.

VII.   FINAL CHECKLIST

☐ Think critically about all of your own claims *and* about all of the claims made by your coauthors. If you do not understand something that your coauthor has written in the draft, push back until you do understand, then suggest an alternative phrasing to clarify the manuscript or figure.

☐ Acknowledgments should clearly state who did which parts of the experiment & analysis, and who wrote the paper.

* jhoffman@physics.harvard.edu

[1] G. M. Whitesides, "Whitesides' group: Writing a paper," Advanced Materials **16**, 1375–1377 (2004).

[2] Jessie T Zhang, Jeehoon Kim, Magdalena Huefner, Cun Ye, Stella Kim, Paul C. Canfield, Ruslan Prozorov, Ophir M Auslaender, and Jennifer E. Hoffman, "Single-vortex pinning and penetration depth in superconducting $NdFeAsO_{1-x}F_x$," Physical Review B **92**, 134509 (2015).

[3] Dennis Huang, Tatiana A. Webb, Can-Li Song, Cui-Zu Chang, Jagadeesh S. Moodera, Efthimios Kaxiras, and Jennifer E. Hoffman, "Dumbbell defects in FeSe films: A scanning tunneling microscopy and first-principles investi-

gation," Nano Letters **16**, 4224–4229 (2016).

[4] Dennis Huang, Tatiana A. Webb, Shiang Fang, Can-Li Song, Cui-Zu Chang, Jagadeesh S. Moodera, Efthimios Kaxiras, and Jennifer E. Hoffman, "Bounds on nanoscale nematicity in single-layer $FeSe/SrTiO_3$," Physical Review B **93**, 125129 (2016).

[5] Jeehoon Kim, Changhyun Ko, Alex Frenzel, Shriram Ramanathan, and Jennifer E. Hoffman, "Nanoscale imaging and control of resistance switching in VO_2 at room temperature," Applied Physics Letters **96**, 213106 (2010).