

How To Write a Paper and Format it Using L^AT_EX

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: July 30, 2020)

The goal of this document is to demonstrate how to write a paper. We walk through the process of outlining, writing, formatting in L^AT_EX, making figures, referencing, and checking style and content. Source files are available at: <http://hoffman.physics.harvard.edu/example-paper/>.

I. GETTING STARTED

You should start writing your paper *while* you are working on your experiment. Prof. George Whitesides says: “A paper is not just an archival device for storing a completed research program; it is also a structure for planning your research in progress. If you clearly understand the purpose and form of a paper, it can be immensely useful to you in organizing and conducting your research. A good outline for the paper is also a good plan for the research program. You should write and rewrite these plans/outlines throughout the course of the research. At the beginning, you will have mostly plan; at the end, mostly outline. The continuous effort to understand, analyze, summarize, and reformulate hypotheses on paper will be immensely more efficient for you than a process in which you collect data and start to organize them only when their collection is ‘complete’.” Here are some concrete steps to get started.

1. Read George Whitesides’ “How to Write a Paper” [1].
2. Read through *at least* one full paper in your target journal, to get a sense of the content and writing style.
3. To clarify in your own head the purpose of your paper, start by writing your abstract.
4. 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.
5. 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 example, for a standard 4-page scientific letter, aim for 13 paragraphs (generally, you can estimate about 200 words per paragraph). Figure out how to tell your entire story (not numbers, just story!) in 13 stand-alone sentences.
6. 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.
7. If you think of references as you go, you can include the minimal identifying information in parentheses to trigger your memory later, e.g. “(WhitesidesAdvMat)”, so all of the information is compact.

8. Rewrite your abstract, taking into account what you have learned from the process of writing the paper. As you fine-tune your abstract, you may find it helpful to refer to Nature’s instructions for writing an abstract [2] and for clear communication more generally [3].

As you contemplate the paper you have just written, put yourself 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, the references are hyperlinked for easy access, 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 figures & 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 ease of understanding, regardless of the scientific merits of the work. (See Ref. 4 for more information on how formatting, even as simple as font choice, will influence the reader’s “cognitive ease” and ultimately their judgment of the report.) Down the road, consider a reader who might cite the paper and launch you to fame and glory: the potential citer’s decision will be influenced by their 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. L^AT_EX

L^AT_EX is a formatting language that allows professional, publication-quality presentation of scientific text, equations, figures, and hyperlinked references. REVTeX is a specific set of L^AT_EX macro packages assembled by the American Physical Society (APS) to standardize manuscript formatting for the various Physical Review journals. The APS author website [5] lists acceptable submission formats as: REVTeX (**preferred**), L^AT_EX, or MSWord. It is sensible to draft your paper in L^AT_EX/REVTeX from the beginning. You will need a

* jhoffman@physics.harvard.edu

reference manager, a text editor, and a \LaTeX compiler. Typically, you will be working closely with other authors, so you should pick a platform that enables simple sharing and collaborative workflow.

Here we recommend the online \LaTeX compiler Overleaf (<http://www.overleaf.com/>) for text editing and compiling. Overleaf (formerly known as Share \LaTeX) is an online platform that is great for collaborative writing: it tracks changes and even has a real-time chat function. Note that LyX is a WYSIWYG (what you see is what you get) \LaTeX editor that may seem tempting if you initially find tex formatting daunting, but LyX may come back to bite you in the long run, because it hides the guts of the tex code from the user, making it harder to control formatting details.

1. Choose & install a reference management software such as Mendeley: <http://www.mendeley.com/> (See appendix A for usage tips.)
2. Make an account on Overleaf: <http://www.overleaf.com/>
3. Download the source files for this example paper: <http://hoffman.physics.harvard.edu/example-paper/>
4. Copy the tex and bib files into a new overleaf project, and share with your collaborators.

Alternatively, you may prefer to run \LaTeX directly on your own laptop. In that case, follow the steps below:

1. Choose & install a reference management software:
 - Mendeley: <http://www.mendeley.com/>
 - Papers: <http://www.papersapp.com/>
 - Endnote: <http://www.endnote.com/>
2. Download & install a \LaTeX compiler such as MikTeX: <http://miktex.org/>
3. Download & install REVTeX 4-2 (see appendix B): <http://authors.aps.org/revtex4/>
4. Pick an editor such as WinEdt: <http://www.winedt.com/> (You may need to request a university license; may require ~ 24 hour turnaround.)
5. Download the source files for this example paper: <http://hoffman.physics.harvard.edu/example-paper/>, open the tex and bib files in WinEdt, and get started writing!

It's important not to lose sight of your outline, as you fill in the details of your paper. This \LaTeX template file allows you to title each paragraph using the `\ptitle{}` command. You should keep these titles in place throughout the entire paper-writing process; they will serve as a constant reminder to keep each paragraph focused on a single point. You should be able to skim through these bold paragraph titles, without reading any of the intervening sentences, and still understand the basic logical flow of the paper. At the final step before submission, comment out the line `\ptitletrue` in the header, to hide the paragraph titles. But do not delete the paragraph titles, because they will remain useful to you in the inevitable paper revision process down the road!

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 Table I.

TABLE I. Formatting mathematical symbols.

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

- 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, as shown in Table II. If you want to prevent a linebreak, you can use `~` instead of `\`.

TABLE II. Spacing.

	\LaTeX	Output
Incorrect	<code>e.g. incorrect</code>	e.g. incorrect
Incorrect	<code>Fig. 2</code>	Fig. 2
Correct	<code>e.g. \ correct</code>	e.g. correct
Correct	<code>Fig.\ 2</code>	Fig. 2
Correct	<code>Fig.~2</code>	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,

$$a^2 + b^2 = c^2. \quad (1)$$

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 $c = 12$.

$$\begin{aligned} d &= \sqrt{a^2 + b^2 + c^2} \\ &= \sqrt{3^2 + 4^2 + 12^2} = 13 \end{aligned} \quad (2)$$

IV. STYLE & GRAMMAR

- 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 (e.g. if you wonder whether acronyms are appropriate in the abstract, check a few recent published examples in your target journal).
- “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 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 Prof. Margo Seltzer: <http://www.seltzer.com/margo/pet-peeves/>

V. FIGURES

It is worth taking 2-3 hours to read the definitive guide to “The Visual Display of Quantitative Information” by Edward Tufte [6]. Tufte defines several metrics for figure optimization:

$$\begin{aligned} \text{Data-ink ratio} &= \frac{\text{data-ink}}{\text{total ink used to print the graphic}} \\ &= 1.0 - \text{fraction of a graphic that can be} \\ &\quad \text{erased without loss of data-information} \end{aligned}$$

$$\text{Data density} = \frac{\text{number of entries in a data matrix}}{\text{area of data graphic}}$$

A clear and visually pleasing figure should:

- Maximize data-ink ratio.
- Maximize data density.
- Avoid “chart-junk”, i.e. hatching patterns that interact with the natural motion of the eye to promote the distracting perception of vibration in a static graphic.
- Avoid excessive colors.
- Avoid red-green combinations (5-10% of people are red-green colorblind!)
- Use concise but clear words (not inscrutable abbreviations) directly on the graphic, so the reader doesn’t have to dig through a lengthy caption or text to understand the components of the figure.
- Orient words horizontally whenever possible.

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.

Note that faithful representation of images in vector format usually also results in a smaller figures size. This

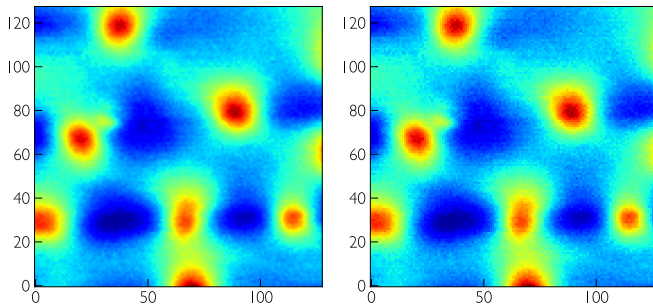


FIG. 1. Comparison between blurry pixels (dishonest interpolation occurs when the image is processed in Microsoft PowerPoint) vs. clean pixels (honest representation is preserved when the image is processed in Python and Adobe Illustrator). MFM images of vortices in $\text{NdFeAsO}_{1-x}\text{F}_x$ [7].

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 6 in the final publication figure [8]. 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 Adobe Illustrator (or other program) of exactly the final size, and build your figure within it, using no fonts smaller than size 6. 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 font, at least size 6.
- 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 ($V_{\text{sample}} = 100 \text{ mV}$; $R_J = 1 \text{ G}\Omega$).
- For STM spectra: give sample bias modulation in the caption ($V_{\text{rms}} = 2 \text{ mV}$).
- 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) [9] and Huang *et al*, Phys. Rev. B 93, 125129 (2016) [10]. 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 L^AT_EX formatting of any special characters in the author’s names, e.g. `S\{a}nchez`
- Check the L^AT_EX 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 documentclass at the top 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 [11]. But if you're using a superscript citation style, and 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 $\text{NdFeAsO}_{1-x}\text{F}_x$ (Ref. 7).

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

VII. PROFESSIONAL INTEGRITY CHECKLIST

- Authorship – are all major contributors and collaborators included? See the American Physical Society guidelines for authorship at http://www.aps.org/policy/statements/02_2.cfm
- Plagiarism – have you been careful to distinguish between your own work and ideas, as opposed to those of others?
- Citations – have you properly cited prior work, and references that you used?
- Data Integrity – have you clearly described the data analysis methods, and justified any data points that were excluded?
- Image Processing – have you clearly described any processing that was applied to images?
- Acknowledgments – have you given appropriate credit and thanks to collaborators and other individuals or organizations who deserve recognition?
- Clarity of collaborative structure – if this is a joint effort, have you identified people who you worked with on this project? Acknowledgments should clearly state who did which parts of the experiment & analysis, and who wrote the paper.
- Conflicts of Interest – do you have any conflicts of interest where you or someone close to you stands to gain, financially or otherwise, from this work?

VIII. 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

We acknowledge advice from Jessie Zhang and Harry Pirie to produce Fig. 1.

Appendix A: Mendeley

Mendeley provides a convenient (although not 100% bug-free) database structure for storing, sorting, and annotating the papers you read. Mendeley also provides an export function to automatically create your bib file. Here are some tips to use Mendeley most effectively.

1. Download Mendeley: <http://www.mendeley.com/>
Launch it on your desktop.
2. Set Mendeley options: Make sure your Mendeley database is set correctly to include the DOI (digital object identifier) field. Go to Tools → Options → Document Details. Scroll down and make sure the DOI box is checked.
3. Import paper: In the upper left corner of Mendeley Desktop, click the drop-down menu for “Add” and select the bottom option “Add Entry Manually”. In the dialog box that pops up, scroll down until you find the DOI field. Paste the DOI into the field, and click the little magnifying glass icon to the right of the field. This will auto-populate all of the relevant paper information such as author names, title, etc., without risk of typos due to manual copying.
Note 1: Mendeley also allows you to import directly from a PDF file, and it tries to pull all of the metadata from the PDF, but the process is imperfect. So it's safest to use the DOI for an error-free import.
Note 2: Even if you use the DOI, some journal titles will not import correctly with special characters, so you may need to manually correct.
4. Add tags: It's useful to add tags to help sort your imported papers. For example, if you are going to be writing a manuscript in 2019 on superconductivity, you might add the tag “sc19” to all the relevant papers that you will be citing in your manuscript.
5. Export bib file: Select all of the references that you want to include, and go to File → Export. Name your file, and it will add a citation key to each paper (e.g. Whitesides2004) and automatically export to a bib file.
6. Resolve redundant citation keys: At this point, you may have several references with the same citation key, e.g. Huang2016a and Huang2016b [9, 10]. For your future convenience, you should manually change the redundant citation keys to be more informative, e.g. HuangNanoLett2016 and HuangPRB2016. Now re-export the bib file.
7. Open the bib file in your tex file editor. By default, Mendeley exports all fields, including long ones like the abstract. To reduce clutter in your bib file, and make it easier to debug any errors, it's a good idea to remove the abstracts and other unnecessary fields. For example, in WinEdt go to Search → Replace, check the regular expressions box, search for “<abstract**\},>,” and replace it with nothing.

Appendix B: REVTeX

How to install REVTeX 4-2 on Windows for MikTeX:

1. Download REVTeX 4-2:
<http://authors.aps.org/revtex4/>
2. Unzip the downloaded folder revtex4-2-tds
3. If you don't already have one, create a folder C:\TeX-local
4. Copy the four subfolders (bibtex, doc, source, tex) into C:\TeX-local
5. From the start menu, open MikTeX 2.9 → Maintenance (Admin) → Settings
6. On the root tab, add the path C:\TeX-local
7. On the general tab, click "Refresh FNDB" (You may need to close WinEdt, or whatever is using your MikTeX installation, in order to properly "Refresh FNDB".)
8. Download natbib:
<http://www.ctan.org/tex-archive/macros/latex/contrib/natbib/>
9. Unzip natbib.
10. Open natbib.ins, and run TeX on it (Shift+Ctrl+T in WinEdt)
11. Open bibentry.ins, and run TeX on it (Shift+Ctrl+T in WinEdt)
12. You may also need to get two more packages, i.e. download from ctan and run TeX on url.ins and textcase.ins

- [1] G. M. Whitesides, Whitesides' group: Writing a paper, *Advanced Materials* **16**, 1375 (2004).
- [2] [How to construct a Nature summary paragraph.](#)
- [3] How experts communicate, *Nature Neuroscience* **3**, 97 (2000).
- [4] D. Kahneman, *Thinking, Fast and Slow* (Farrar, Straus and Giroux, 2011).
- [5] [Physical Review journals web submission guidelines.](#)
- [6] E. R. Tufte, *The Visual Display of Quantitative Information*, 2nd ed. (Graphics Press LLC, Cheshire, CT, 2001).
- [7] J. T. Zhang, J. Kim, M. Huefner, C. Ye, S. Kim, P. C. Canfield, R. Prozorov, O. M. Auslaender, and J. E. Hoffman, Single-vortex pinning and penetration depth in superconducting NdFeAsO_{1-x}F_x, *Physical Review B* **92**, 134509 (2015).
- [8] B. de Bivort, [Figure style guide.](#)
- [9] D. Huang, T. A. Webb, C.-L. Song, C.-Z. Chang, J. S. Moodera, E. Kaxiras, and J. E. Hoffman, Dumbbell defects in FeSe films: A scanning tunneling microscopy and first-principles investigation, *Nano Letters* **16**, 4224 (2016).
- [10] D. Huang, T. A. Webb, S. Fang, C.-L. Song, C.-Z. Chang, J. S. Moodera, E. Kaxiras, and J. E. Hoffman, Bounds on nanoscale nematicity in single-layer FeSe/SrTiO₃, *Physical Review B* **93**, 125129 (2016).
- [11] J. Kim, C. Ko, A. Frenzel, S. Ramanathan, and J. E. Hoffman, Nanoscale imaging and control of resistance switching in VO₂ at room temperature, *Applied Physics Letters* **96**, 213106 (2010).