
Experiences converting from PDF-only to paper

Jim Hefferon and Lon Mitchell

Abstract

Offering a textbook for free download has become common. With the growth of on-demand printing, adding a paper option is an easy way to distribute the work in a format that some users prefer. We will give an overview of today's landscape of print on demand and discuss some differences in delivering works in the two media.

1 Background

The text *Linear Algebra*¹ by the first author has been available since 1996, with L^AT_EX source, under a license that is Free, in this case either the GNU Free Documentation License (GFDL) or Creative Commons CC-BY-SA. It covers a standard US first course, has extensive question sets with worked answers for every question, and is supplemented by classroom beamer slides and a lab manual. In electronic form it is widely accessed; during the academic year it has 30K–40K downloads per week. In the past it has only been available online but it is now also offered as a paper version. The experience of that addition is detailed here.

Until now, to get paper copies to students an instructor would download a PDF, make a master copy, produce bound copies of that, and then students paid to defray the cost, usually at their college bookstore.

This model has implications. First, it suggested that the page size should be US letter, 8.5×11 inches, since that is the most convenient form for printing, copying, and binding here in the US. Second, typically instructors would use a black and white printer so the book design couldn't be ambitious with color (hyperlinks are in blue, which prints to a dark gray). Finally, print resolution was an issue. In 1996 printers were typically 300 dots per inch. That is low enough that some shading and graphics vanished so illustrations had to limit those. Another impact of the low resolution is that, since in this model students got copies of copies, Computer Modern characters often had dropouts, particularly in subscripts or superscripts. Consequently the book style was switched to use Concrete, with Euler for math.

Thus, even with a distribution model that was online-only, printer capabilities influenced book design decisions.

¹ <http://joshua.smcvt.edu/linearalgebra>

2 Desire for paper

Through the almost twenty years the book has been available, requests for paper copies have come in regularly.² Some came from self-studiers who were simply used to a certain format. More worrying was that sometimes instructors who were potential adopters said that they would not use the text because of the hassle of having it printed.

Despite those requests, the prospect of dealing with bounced checks, return copies, and lost-in-the-mails kept the book officially available online only.

3 Publishing

Those who lay out their own text using T_EX and post the result for free download on a website have provided (in a basic way) two of the main services historically provided by publishers: typesetting and marketing. Although many books produced this way could benefit from additional effort and advice in both areas, T_EX and the Internet can make self-publishing feasible and attractive.

A print edition requires a printer, distribution, and retail (all of which are disjoint from what typically defines a publisher: owning an ISBN). Multiple companies now offer the convenience of a combination of two or all three, but beware those who would also bundle in other services such as marketing or book layout; such services, if needed, are almost always better when acquired separately. Perhaps the most important caveat is that companies that offer all three services may also act as the publisher (because they own and provide the ISBN). This is not necessarily bad, as it can save time and effort, but it will mean that you may lose control over some decisions normally made by a publisher, such as those discussed below.

Publishers choose a list price and a 'discount' for each title. Retailers pay the list price less the discount, and publishers must pay the print cost and distribution fees from this revenue. The industry standard discount is 55%, but some distributors will allow less. Smaller discounts can discourage retailers from carrying your title. For many years, online retailers and university bookstores seemed content with discounts as low as 25%, but that may be changing. This is particularly unfortunate as the higher costs of print on demand (when compared to offset printing) are amplified by higher discounts.

² Studies suggest up to 75% of students prefer a paper copy: <http://www.studentpirgs.org/reports/cover-cover-solution>

In our opinion, the ideal print edition of a Free book should have a (very) low price, be easily available to college bookstores and sold by multiple online retailers, and require little or no effort of the author beyond initial setup. Unfortunately, to the best of our knowledge, there is not currently an easy way for an author working alone to achieve all of these simultaneously. There is a great need for a new model in the book business that is attuned to open-source projects, but that business is dominated by a few large companies and change seems far off.

With *Linear Algebra*, using a “low-profit” company (L3C) created by the second author to act as the publisher, we have achieved (for the moment) each of the three goals above.

4 Why do it yourself?

The GFDL has a fascinating premise, providing specific language that allows anyone to publish your book. This turns the usual relationship between author and publisher around—anyone can grab a GFDL-licensed work and offer a version for sale, including a print version. That has happened a number of times with *Linear Algebra*. This may be attractive to authors who would like to have a print edition but would rather not have to do any of the work. The GFDL may not be to everyone’s liking, however, as it does not place many limits on what the publisher can and cannot do; for example, previous print versions of *Linear Algebra* offered for sale have not been updated when the online text has been updated. One could easily adapt a more restrictive license to a similar function by offering specific exceptions to reputable potential publishers.

If you are interested in having your work published in this way, you can be of great help to potential publishers in your use of T_EX. For example, it is likely a publisher will need to reformat your work to an industry-standard page size; using already provided descriptive commands (such as `\emph` instead of `\textit`) and/or providing new commands that can easily be redefined in the preamble (such as `\beforeexercisesskip` instead of a `\medskip`) is highly recommended.

Open source books need not be new material; if you own the copyright to an out-of-print book (or can acquire it from the publisher), you can make a quality contribution to this growing movement.

5 Money

With reputable printers, costs are significantly less for industry-standard book sizes. The original size of *Linear Algebra*, US letter 8.5 × 11 inches, is not standard in printing, and is expensive. Lower-level

textbooks often have a larger page size than upper level ones, so we went with one of the larger standards, 7.5 × 9.25 inches. One advantage of this size is that we can retain L^AT_EX’s default text width and so downloaders have the option to print onto US letter.

We had some other costs: setup fees, an annual listing fee, and ISBNs. (Traditionally, if two copies of a book had the same ISBN then a buyer could depend on them being substantially the same, so any major or even many minor changes in a book should trigger a new ISBN.)

On the revenue side, having a little extra for each copy helps with things like being able to display and distribute free books at the Joint Math Meeting. In any event, rounding up to the nearest multiple of five is safer than rounding down.

Note that with a GFDL book, anyone could undercut the price. However, most readers interested in the book go to its web page, which points to the official paper edition. The price of that is so reasonable that we don’t see users looking for a non-official (and possibly outdated) version.

The takeaway is that we used Ingram’s print on demand service at a \$20 list price, which splits into (about) thirds for printing, the retailer, and us.

6 The outside

Coming out with a paper copy seemed to call for giving it a nice cover. With print on demand services, the cover is uploaded as one file that shows the front, back, and spine. Any revision that adds or removes pages changes the size of the spine (*Linear Algebra*’s 498 pages yield a spine of 1.001 inches). This means that an author should create the cover using a process that is easily redone. In this case, the elements in the cover file were arranged in a L^AT_EX `picture` environment to get exact locations.

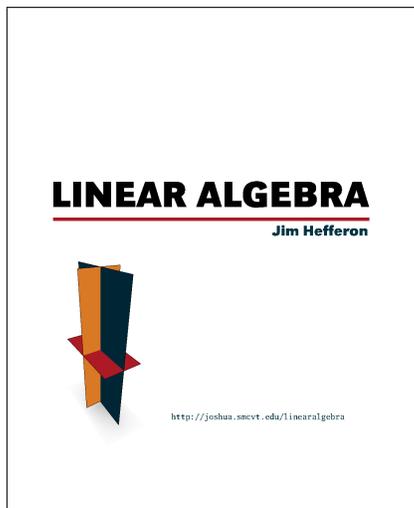
There is no printing on the inside of the cover. Consequently the driver file `book.tex` got a flag so that if the PDF is for paper then it makes an extra inside sheet, a title page, whose back is the page with the list of symbols. (This flag also prints hyperlinks in the book’s body in black instead of the electronic PDF’s blue.)

Traditionally the back cover has some marketing text, which is a challenge to write for an academic unused to selling. The ISBN and bar code go there also, in a box about an inch tall.

6.1 Art

The main difficulty with the cover was not technical. Getting a capable person to execute a graphic proved to be hard. For example, a colleague in Media Studies recommended a talented student who

seemed interested and agreed to produce it for a fair price. But it never appeared. After a half dozen such episodes, out of embarrassment at how long it was taking, ten minutes spent in Asymptote drawing the planes and a half hour spent on Kuler³ gave the result below.



7 The inside

Here are the basic dimensions.

```
\usepackage[papersize={7.5in,9.25in},
  textwidth=345pt,
  inner=.8in,
  top=.85in, bottom=.85in,
  headsep=12pt,
  bindingoffset=0.4in,
]{geometry}
```

The `top` and `bottom` setting squeezes more lines on a page. The more pages there are, the more the binding will use up page space, so the inside margins needed to be set a little larger.

Repaginating took a long time. The `microtype` package helped with linebreaks, and forbidding widows and orphans along with using `\raggedbottom` helped with page breaks. But each page required individual attention, including some rewriting.

7.1 Lurking culprits

At the first submission, besides the binding margin there was only one thing to fix: a non-embedded Type 3 font somewhere in the document. The program `pdffonts` gave a page number. The only suspicious things there were two graphs from *Sage*.

Some spelunking⁴ revealed that *Sage* required passing to the plot the parameter `typeset=latex`.

³ <http://kuler.adobe.com>

⁴ <http://trac.sagemath.org/ticket/14664>

With that, the fonts were embedded and the print on demand technical requirements were satisfied.

7.2 Shades of gray

The major issue remaining in the text is that many of the illustrations use shaded lines. One example pictures the linear map $h(x) = x + y$ from the plane to the line. The text has three drawings: first, a vector from the line $x + y = w_1$; second, another vector from the line $x + y = w_2$; and finally, the vector that is the sum of the two, drawn as ending at the line $x + y = w_1 + w_2$. Each of the three gives its vector in normal face with the associated line shaded. In the PDF a level of 0.85 gray seemed visually right (this is mostly white, with fifteen percent black). But on the printed page this sometimes came out too light, to the point where it was nearly invisible.

To bring out the lines on paper without compromising the online version, the illustrations were adjusted to make the lines thicker and to have the darkness 0.70. (A few illustrations in the book used this darkness level and it seemed to print fine.) However, even with this change, in the current paper version of the text some lines are too light and the quality of the illustration suffers. An author considering this issue may want to go with dashed lines, or some other larger stylistic change.

7.3 In print, mistakes look *really* bad

The book has been available for twenty years and there have been any number of corrections of typos, and some errors of fact as well. That was routine. But when the paper version came from the printer with a glaring error, it was startling. (A `LATEX` `picture` environment was picking up a `\setlength` that had not been put inside a group and so the environment was three times the size of the page.) The next PDF sent to the printer got very close scrutiny.

8 The result

The book is now available at the usual online retailers and can be ordered at wholesale pricing by college bookstores from Ingram's catalog. The download page gives a link. This answers the requests that have come in over the years, without requiring any book-toting by the author.

- ◇ Jim Hefferon
Saint Michael's College
`jhefferon` (at) `smcvt` dot `edu`
- ◇ Lon Mitchell
Orthogonal Publishing
`l3c` (at) `orthogonalpublishing` dot `com`