

# FIRST STEPS WITH GENERAL TYPESETTING

RICHARD KOCH

## 1. AN APOLOGY

TeXShop's Help system is organized into three pieces. The first gives information about TeXShop itself, placing emphasis on simple matters every user needs to know. The second gives information about  $\text{\TeX}$  for those who need its superior typesetting features but do not particularly need its mathematical capabilities. The third gives information about  $\text{\TeX}$  for users whose documents contain a lot of mathematics.

Each of the last pieces is to contain a short introduction, and then a fairly long instruction manual which teaches enough  $\text{\TeX}$  to begin using the program seriously.

In the case of mathematical users, such an instruction manual exists. The fourth edition of George Grätzer's well known book *More Math Into  $\text{\LaTeX}$*  will be published in September. This book begins with an excellent Short Course with exactly the right amount of information, and Grätzer graciously gave permission to include the Short Course in TeXShop.

However, at the moment the right manual has not been found for General Typesetting, so no manual is included. Temporarily, users interested in general typesetting should use Grätzer's course and ignore the chapters on mathematics.

There is a reason for the lack of an appropriate manual on general typesetting. Over the last year, a new variant of  $\text{\LaTeX}$  named  $\text{\XeLaTeX}$  has appeared; this variant solves two major  $\text{\TeX}$  problems general users run into, so I believe it is the appropriate starting point for new users. This conviction is so strong that I am not willing to provide a manual which leads in a different direction.

Several free manuals on  $\text{\LaTeX}$  exist. What is needed is time to edit one of these manuals so it refers to  $\text{\XeLaTeX}$  from the start. Such a manual should appear soon and become part of TeXShop.

If you want  $\text{\TeX}$  for general typesetting, but find that there is not enough information in George Grätzer's book, and cannot wait for the appropriate  $\text{\XeLaTeX}$ -based manual, go to [www.uoregon.edu/~koch/texshop/documentation.html](http://www.uoregon.edu/~koch/texshop/documentation.html) and download ``The Not So Short Introduction to  $\text{\LaTeX} 2_{\epsilon}$ '' by Tobias Oetiker.

In the remaining portion of this document I'll give some preliminary information about  $\text{\XeLaTeX}$ , and explain how to use it to typeset standard  $\text{\LaTeX}$  source and all examples in Grätzer's book.

## 2. SOME $\text{\TeX}$ HISTORY

$\text{\TeX}$  was invented before several of the important computing developments of the

twentieth century. In 1978 there was no Macintosh and no generally available computer with a graphical interface. Personal computers printed output using a single bitmapped font in which all characters had the same width; some computers only printed upper case letters.

Consequently, Knuth had to invent outline fonts at the same time that he invented  $\text{\TeX}$ . In his documentation, Knuth always talks about *two* programs, the typesetting program  $\text{\TeX}$  and the outline font program  $\text{\METAFONT}$ .

Later in the decade, other players introduced their own outline fonts using different formats: Adobe's Type 1 Postscript fonts, Apple's TrueType fonts, and Microsoft's OpenType fonts. The history of these fonts is rather complicated; suffice it to say that each is now supported on the Mac and available on other operating systems as well. A vast number of beautiful fonts are available in these new formats.

Adapting these fonts for use in  $\text{\TeX}$  is a complicated task, which has only been successfully completed in a small number of cases. One of the most important complaints about  $\text{\TeX}$  is the limited number of available fonts.

### 3. $\text{\XeTeX}$ AND FONTS

$\text{\XeTeX}$  completely solves the font problem. It can use Knuth's fonts, of course, but it can mix these with system fonts in Adobe Type 1, TrueType, and OpenType formats. These fonts need not be adapted for use in  $\text{\TeX}$ ; instead they are immediately available.

### 4. MORE HISTORY: UNICODE

$\text{\TeX}$  source manuscripts contain standard characters available on any typewriter. Indeed, originally  $\text{\TeX}$  could only accept the 128 standard ASCII input characters. Later this was expanded to 256 characters. This is enough for standard English and for Western European Languages, but it is certainly not enough for Japanese, Chinese, and a multitude of other languages. Over the years, many projects have attempted to adopt  $\text{\TeX}$  so it can typeset these other languages.

Later the same problem arose in the entire computer industry, since a large number of sales occur in countries which do not use Western scripts. The solution invented by the industry is Unicode, an open standard which can theoretically encode the characters of all languages on earth. Most computer manufacturers have adopted this standard. For example, the edit class in Cocoa (and consequently  $\text{\TeX}$ Shop's editor) accepts arbitrary Unicode characters.

### 5. $\text{\XeTeX}$ AND UNICODE

$\text{\XeTeX}$  modifies  $\text{\TeX}$  to accept any Unicode character; source documents are saved in UTF-8 Unicode format. In conjunction with full support for OpenType fonts, this makes it possible to write  $\text{\TeX}$  documents in virtually any language on earth.

### 6. SOLD. HOW DO I USE $\text{\XeTeX}$ ?

Simple. Take any standard  $\text{\LaTeX}$  document, for example, any sample document from Grätzer's book. Somewhere close to the top, say within the first ten lines, include the lines

```

%!TEX TS-program = xelatex
%!TEX encoding = UTF-8 Unicode

```

These lines begin with the  $\TeX$  comment character, so  $\TeX$  will ignore them. However, TeXShop recognizes the lines and does something special when you open, save, or typeset the file. The first line tells TeXShop to call `xelatex` rather than `pdflatex` when you ask it to typeset the document. The second line tells TeXShop to load and save the file with `UTF-8 Unicode` encoding, rather than the standard Mac OS Roman encoding.

Since  $\XeLaTeX$  supports standard  $\TeX$  fonts and standard  $\TeX$  commands, typesetting will proceed without change.

#### 7. HOW DO I SWITCH TO ALTERNATE FONTS IN $\XeLaTeX$ ?

A style package written by Will Robertson makes this task simple. Add the following lines to the  $\TeX$  preamble:

```

\usepackage{fontspec,xltxtra,xunicode}
\defaultfontfeatures{Mapping=tex-text}
\setromanfont{Hoefler Text}
\setsansfont[Scale=MatchLowercase]
               {Gill Sans}
\setmonofont[Scale=MatchLowercase]
               {Andale Mono}

```

Change the three font names to any system fonts you'd like to use. The first font is the default serif font, the second is the default sans serif font, and the third is the default monospaced font for typewriter-like output. If your entire document is in a serif font, it is enough to change the first font.

To find appropriate font names, open Apple's Font Book application, look at the font

samples, and choose one of the font names listed by the application.

#### 8. HOW DO I INSERT UNICODE IN THE SOURCE MANUSCRIPT?

Go to Apple's System Preferences and choose the International panel. Select the Input Menu tab. A large number of keyboards layouts are provided, as well as more complicated input methods for Chinese and Japanese.

Choose, for example, Arabic, Nepali, and Russian. A flag icon will appear at the right side of the menu bar. Open a blank TeXShop document. Start with a US or Western European flag and type some standard text.

Then select the Russian flag and type some more. Notice that you are inputting Cyrillic characters. Select Nepali and type another line, noticing the distinct characters.

Finally select Arabic and type again. Arabic is written from right to left; notice that the characters you type in TeXShop appear from right to left. In Arabic, character shapes change when the character ends a word. Notice that previous characters tend to change shape as you add additional characters.

None of this is very useful unless you know one of these languages!

#### 9. HOW DO I SEE THESE CHARACTERS IN THE $\TeX$ OUTPUT?

Ah, that is the tricky step. Some Macintosh fonts appear to contain only standard Roman characters, but have been extended to contain additional Unicode characters.

You must typeset in a font which contains characters for the language you have selected.

For example, the Geeza Pro font contains Arabic characters, the Lucida Grande font contains Hebrew characters, and the Osaka font contains Japanese characters. Some experimentation may be needed to find a font which contains characters for the language you want to use.

Suppose, for example, that you want to output Arabic. The following line defines a font family which contains Arabic characters:

```
\newfontfamily{\A}{Geeza Pro}
```

To insert Arabic, type the following and insert Arabic characters in the spot currently containing dots:

```
This is Arabic text: {\A ....}
```

as in {\A السلام عليكم}.

Incidentally, this document was typeset with Xe<sub>La</sub>TeX. In ordinary L<sup>A</sup>T<sub>E</sub>X, the last clause of the previous paragraph would have been difficult to typeset.

#### 10. IF Xe<sub>La</sub>TeX TYPESETS STANDARD L<sup>A</sup>T<sub>E</sub>X, WHY NOT USE IT FOR MATHEMATICS?

In standard T<sub>E</sub>X, the mathematical fonts match the computer modern fonts designed by Knuth, so the document has a standardized pleasing appearance. Currently there are no, or almost no, OpenType fonts containing the full set of mathematical symbols required by T<sub>E</sub>X. Indeed, the Unicode

standard has only recently been extended to full mathematical support, and experimental fonts are being constructed at the moment.

Consequently, a document with a large amount of mathematics will have to use Knuth's mathematical fonts for the mathematics, even if it uses an open type Macintosh font for the text. The result is a slight mismatch between the text font and the mathematical font. This makes using unusual fonts less attractive in mathematical texts.

There are a small number of commercial T<sub>E</sub>X font sets containing redesigned text characters and matching mathematical characters. Consult the Internet for details. The usual text fonts in these packages are Lucida and Times Roman.

The technical community uses T<sub>E</sub>X extensively and exchanges source documents over the Internet. The fact that Xe<sub>La</sub>TeX source files are written in UTF-8 rather than standard ASCII may confuse colleagues.

Thus users in the technical community have little incentive to use unusual fonts, and some incentive to avoid confusion while exchanging source documents.

#### 11. SAMPLE Xe<sub>La</sub>TeX DOCUMENTS

The remaining pages show sample Xe<sub>T</sub>E<sub>X</sub> source manuscripts, and the resulting Xe<sub>La</sub>TeX output. This gives some feel for the advantages of Xe<sub>La</sub>TeX in the general non-scientific community.

% These lines tell TeXShop to typeset with xelatex, and to open  
% and save the source with Unicode encoding.

%!TEX TS-program = xelatex  
%!TEX encoding = UTF-8 Unicode

\documentclass[12pt]{article}  
\usepackage{geometry}  
\geometry{letterpaper}  
\usepackage{graphicx}  
\usepackage{amssymb}

% Will Robertson's fontspec.sty can be used to simplify font choices.  
% To experiment, open /Applications/Font Book to examine the fonts  
% provided on Mac OS X, and change "Zapfino" to any of these choices.

\usepackage{fontspec,xltxtra,xunicode}  
\defaultfontfeatures{Mapping=tex-text}  
\setromanfont[Mapping=tex-text]{Zapfino}  
\setsansfont[Scale=MatchLowercase,Mapping=tex-text]{Gill Sans}  
\setmonofont[Scale=MatchLowercase]{Andale Mono}

\title{Brief Article}  
\author{The Author}

\begin{document}  
\maketitle

This is a standard  $\text{\TeX}$  document using Zapfino type. It uses  $\text{\LaTeX}$  and  
the standard  $\text{\TeX}$  graphic packages.

\begin{figure}[htbp]  
  \centering  
  \includegraphics[width=2in]{example.pdf}  
  \caption{from Mathematica}  
\end{figure}

\end{document}

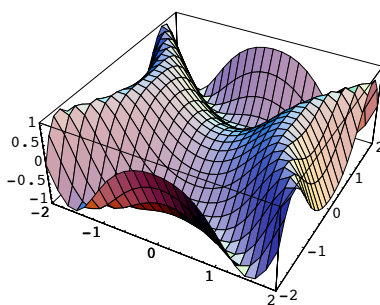
# Brief Article

The Author

May 5, 2007

*This is a standard  $\text{\TeX}$  document using Zapfino type.*

*It uses  $\text{\LaTeX}$  and the standard  $\text{\TeX}$  graphic packages.*



*Figure 1: from Mathematica*

% These lines tell TeXShop to typeset with xelatex, and to open and  
% save the source with Unicode encoding.

%!TEX TS-program = xelatex  
%!TEX encoding = UTF-8 Unicode

\documentclass[12pt]{article}  
\usepackage{geometry}  
\geometry{letterpaper}  
\usepackage{graphicx}  
\usepackage{amssymb}

% Will Robertson's fontspec.sty can be used to simplify font choices.  
% To experiment, open /Applications/Font Book to examine the fonts  
% provided on Mac OS X, and change "Hoefler Text" to any of these choices.

\usepackage{fontspec,xltxtra,xunicode}  
\defaultfontfeatures{Mapping=tex-text}  
\setromanfont[Mapping=tex-text]{Hoefler Text}  
\setsansfont[Scale=MatchLowercase,Mapping=tex-text]{Gill Sans}  
\setmonofont[Scale=MatchLowercase]{Andale Mono}

\title{Brief Article}  
\author{The Author}

\begin{document}  
\maketitle

This is a standard TeX document using Hoefler Text. It uses \LaTeX\ and  
the standard \TeX\ graphic packages.

% For many users, the previous commands will be enough.  
% If you want to directly input Unicode, add an Input Menu or Keyboard to the menu bar  
% using the International Panel in System Preferences.  
% Unicode must be typeset using a font containing the appropriate characters.

\newfontfamily{\A}{Geeza Pro}  
\newfontfamily{\H}[Scale=0.9]{Lucida Grande}  
\newfontfamily{\J}[Scale=0.85]{Osaka}

Here are some multilingual Unicode fonts: this is Arabic text: {\A السلام عليكم},  
this is Hebrew: {\H שלום}, and here's some Japanese: {\J 今日は}.

\end{document}

# Brief Article

The Author

April 24, 2007

This is a standard TeX document using Hoefler Text. It uses  $\LaTeX$  and the standard TeX graphic packages.

Here are some multilingual Unicode fonts: this is Arabic text: السلام عليكم, this is Hebrew: שלום, and here's some Japanese: 今日は.