What Is New in LaTeX? I. Breaking Free

G. Grätzer

This is the first of a series of columns updating the mathematical community about some current developments in TeX and TeXing.

—Andy Magid

Jerry Seinfeld?

On this, we can all agree: LaTeX is the most important tool of a mathematician. So everybody wants to know: What is new in LaTeX?

In two words: *Not much.* Donald Knuth corrects obscure bugs in TeX every few years, and 2008 was such a year (next: 2013). TeX was updated from version 3.141592 to version 3.1415926. Chances are you have never come across any of the bugs exterminated this year…

LaTeX is supposed to be updated in December of every year. It was not updated last year, and in the years before, the changes were rather small. So LaTeX is extremely stable.

The AMS packages, which are so important for mathematical typesetting, are grouped under the name `amsmath`—not changed in eight years, and `amscls`—not changed in four years.

Is my series of articles like a Jerry Seinfeld episode: articles about nothing?

So why is it that any LaTeX expert you talk to is so excited about the changes that are taking place?

Looking Back

Donald E. Knuth's multivolume work, *The Art of Computer Programming* [4], caused its author a great deal of frustration because the proofs of the second edition of Volume 2 turned out to be so awful. To solve this problem, Knuth decided in 1978 to create his own typesetting language. The result is described in *The TeXbook* [5]. We can say that TeX was designed by an American mathematician to be used in his own work, and then later "generalized" to be used by other American mathematicians also. TeX had only one font family, Computer Modern, designed in Knuth's Metafont [7], based—indirectly—on the work of the great French font designers Didot (father and son), who published romantic novels in the early nineteenth century.

To work with TeX, you need a platform. LaTeX, developed by Leslie Lamport [8] in the early 1980s, provided

- the use of logical units to separate the logical and the visual design of an article;
- automatic numbering, cross-referencing;
- bibliographic databases.

If you

- write in English (accents are difficult to type and the default hyphenation works only in English),
- use an American keyboard,
- do not need sophisticated mathematical typesetting of the type developed by the AMS,
- do not dislike the CM fonts,

then you had a very capable system as early as 1982.

What Is a Font?

When you work at your computer, a font is selected for you. When you hit a key, your keyboard transmits to the computer an 8-bit number (that is, an integer between 0 and 255). The computer looks up in a table what character corresponds to that number in the font used and displays the character on your monitor. The same way, when you print your document, that character is transformed into an "outline", which will produce the required output on your printer or when viewing a `pdf` file. The output is produced using a set of tables, mapping the character into a "glyph" (the

G. Grätzer is Distinguished Professor of Mathematics at the University of Manitoba. His email address is `gratzer@me.com`.

drawing), specifying sizes, ligatures, hinting (adjusting the glyph to the screen pixels), the spacing between the characters (kerning), and so on.

Until 1989 TEX only accepted 128 numbers (0 to 127); Version 3 accepted all 256. This made it possible to extend TEX to foreign languages. Johannes Braams' `babel` package (released soon after Version 3 became available) provided the framework.

What happens if you type a character not accepted by LATEX? Type

```
Grätzer György
```

in a LATEX document. It will typeset as

Grtzer Gyrgy

The accented characters will be omitted (and the `log` file will record it).

PostScript Fonts

By this time, the basic 35 Adobe Type 1 PostScript fonts became available on most laser printers. All computers came equipped with the appropriate files to utilize them. The New Font Selection Scheme, NFSS, of Frank Mittelbach and Rainer Schöpf, written in 1989, allows the easy integration of the PostScript fonts into LATEX.

A LATEX document class specifies three standard font families (see Section 5.6.2 of [3]):

- a roman (or serif or main) font family,
- a sans serif font family,
- a typewriter style font family.

For instance, the `times` package in the PSNFSS distribution makes Times-Roman the roman font family, Helvetica the sans serif font family, and Courier the typewriter style font family. In the `times` package these are named `ptm`, `phv`, and `pcr`, respectively.

In the preamble of your document, type

```
\usepackage{times}
```

after the `\documentclass` line. Then Times becomes the roman, Helvetica the sans serif, and Courier the typewriter style document font family.

MathTıme and Lucida

Looking at a mathematical article typeset with the Times text font, you may find that the Computer Modern math symbols look too thin. To more closely match Times and other PostScript fonts, Michael Spivak created the *MathTıme* fonts (the most recent version is called *MathTıme Pro 2*). You can purchase these fonts from Personal TEX:

`http://pctex.com/fonts.html`

If you do not like CM or Times fonts, you may want to consider the Lucida fonts designed by Bigelow & Holmes, with support for both mathematics and text. You can obtain the font set from the TEX User Group (also from Personal TEX):

`http://tug.org/lucida`

Problems, Problems

When I get into trouble with LATEX, I turn to the TeX on Mac OS X Mailing List, where experts help me out. A substantial part of the ongoing discussion is about fonts:

- How do I install …
- Why does the installation not work …

Different types of fonts in the three basic operating systems require differing—and nontrivial—installation processes, and the user is confused.

Font technology develops very fast; LATEX has always had a hard time catching up.

Apple Computer developed TrueType fonts in the late 1980s as a competitor to Adobe's Type 1 fonts. Adobe and Microsoft in 1996 introduced OpenType fonts, which in 2007 became ISO Standard ISO/IEC 14496-22. By 2005 there were more than 10,000 OpenType fonts.

And how many of these can we use natively in LATEX? None.

Unicode

The problem is that a font does not contain enough room for the large number of characters we would need to typeset all languages (and also math!). It is easy to see that the 8-bit font tables are at the root of the problem. Unicode is supposed to fix this problem.

In 1988 Joe Becker published a draft proposal with a 64-bit font table: "Unicode is intended to address the need for a workable, reliable world text encoding … to encompass the characters of all the world's living languages."

The Unicode standard in 1991 defined 16 "planes" (see [10]), each containing 65,534 characters. Plane 0 is the Basic Multilingual Plane and all the often-used characters from the vast majority of living languages can be found there.

But what good is Unicode for LATEX?

XƎLATEX

If instead of LATEX, you use XƎLATEX, all your font problems go away. It is easy to switch from LATEX to XƎLATEX. In WinEdt 6, the TEX icon is a pull-down menu; select XƎLATEX. In TeXShop, `LaTeX` is a pull-down menu; pull it down and choose `XeLaTeX`. The next time you typeset, you do it with XƎLATEX.

XƎLATEX is the brainchild of Jonathan Kew. It was introduced in 2004, and since 2007, it is included in the TEX distributions TEX Live and MiKTEX. So you probably have it, even if you have never heard about it.

Now take your Hungarian (German) keyboard, and in a LATEX document type:

```
Grätzer György
```

(the Hungarian keyboard has all these keys) and XƎLATEX will typeset this as

Grätzer György
(provided you use a font family, such as Lucida, that contains the characters ä and ö).

The real magic of XeLaTeX is its handling of fonts. We illustrate some simple uses with a short sample article:

```
\documentclass{amsart}
\usepackage{amssymb,latexsym}
\usepackage{fontspec,xltxtra,xunicode}
\setmainfont[Mapping=tex-text]
   {Garamond Premier Pro}
\setmonofont[Scale=MatchLowercase]
   {Courier}
\defaultfontfeatures{Mapping=tex-text,
   Scale=MatchLowercase}
\setsansfont{Lucida Grande}
\newfontfamily{\Bic}{Bickham Script Pro}
\begin{document}
\title{Illustrating document fonts\\
and font switching}
\maketitle
This is set in Garamond Premier Pro,
the roman document font.---

\texttt{This is set in Courier, the
typewriter style document font.}

We set up a command, \verb+\Bic+,
for switching to Bickham Script Pro:\\
{\Bic This is Bickham Script Pro.}
\end{document}
```

In the preamble, we invoke the necessary packages and specify the three document fonts (you can pick any three system fonts installed on your computer). No more special installation for LaTeX! You no longer care whether the fonts are PostScript Type 1, or TrueType, or OpenType fonts; it just works.

The last command of the preamble sets up the font switching command, `\Bic`.

ILLUSTRATING DOCUMENT FONTS AND FONT SWITCHING

This is set in Garamond Premier Pro, the roman document font.—

`This is set in Courier, the typewriter style document font.`

We set up a command, `\Bic`, for switching to Bickham Script Pro:

This is Bickham Script Pro.

The typeset sample article.

XeLaTeX accommodates all Latin alphabets, with all the accents, Cyrillic, Arabic, Chinese… It can print left-to-right or right-to-left, horizontal or vertical, as required by the language.

We have broken free from the LaTeX font constraints.

For a technical description of XeLaTeX, read Michel Goossens' *The XeTeX Companion* [2]. To view a recent lecture by Jonathan Kew on "What is new in the XeTeX world?" go to `http://river-valley.tv/conferences/bachotex2008/`

If you are curious about the commands, `Mapping=tex-text` (note how `---` became an m-dash) and `Scale=MatchLowercase` (adjusting the font size), look them up in the `fontspec` package documentation by Will Robertson [9].

A Warning

LaTeX is easy to work with, because it uses standard 8-bit encoding for the characters. So if I write a paper on my Mac, send it to my coauthor, who works on a PC, and we submit it to a journal that uses Unix, there is no problem.

Mathematicians, as a rule, have no great need for fancy fonts, so there is little incentive to switch to XeLaTeX. However, soon there will be Unicode math fonts to choose from, and then we should reevaluate the situation.

On the other hand, if you are a linguist, or want to work in non-European languages, switching to XeLaTeX is a no-brainer.

References

[1] Johannes Braams, Babel, a multilingual package for use with LaTeX's standard document classes, *TUGboat* **12** (1991), 291–301.

[2] Michel Goossens, *The XeTeX Companion*, `http://cern.ch/XML/lgc2/xetexmain.pdf`

[3] George Grätzer, *More Math into LaTeX*, Springer-Verlag, New York, 2007. xxxiv+619 pp.

[4] Donald E. Knuth, *The Art of Computer Programming*, Volumes 1–3, Addison-Wesley, Reading, MA, 1968–1998.

[5] ———, *The TeXbook. Computers and Typesetting*, Vol. A., Addison-Wesley, Reading, MA, 1984, 1990.

[6] ———, *The METAFONTbook. Computers and Typesetting*, Vol. C., Addison-Wesley, Reading, MA, 1986.

[7] ———, *Computer Modern Typefaces. Computers and Typesetting*, Vol. E., Addison-Wesley, Reading, MA, 1987.

[8] Leslie Lamport, *LaTeX: A Document Preparation System*, 2nd edition, Addison-Wesley, Reading, MA, 1994.

[9] Will Robertson, The fontspec package, `http://ctan.org/tex-archive/macros/xetex/latex/fontspec/`

[10] The Unicode Consortium, *Unicode Standard*, Version 5.0, Addison-Wesley Professional, Reading, MA, 2006.