

typezh: XeTeX 排版中文辅助宏

XeTeX 是 TeX 排版系统的扩展，增加了对 Unicode 和 OpenType® 的支持。因此可被用于中文排版。

typezh 利用 \XeTeXinterchartoks 机制实现中文排版所需的字体自动切换和标点间距调整功能。Plain TeX 用户可以通过

```
\input typezh
```

加载这些功能。LaTeX 用户建议使用功能更加完善的 xeCJK 宏包或者 CTeX 套装。

typezh 以能进行中文排版为目标，提供尽可能少的功能，而不是成为一个万能的“宏包”。和 Plain TeX 一样，仅提供了一个字号（10 磅）的字体。用户如果需要更多的样式，应当创建自己的格式文件，并将 typezh 的代码（可以做一些本地修改）包含于其中。


1. 中西文字体切换

你可以自由地书写{\rm 罗马体 Roman}、{\it 意大利体 *Italic*}、{\sl 倾斜体 *Slanted*}、{\bf 粗体 **Boldface**} 和 {\tt 电报体 Teletype}。

typezh 默认使用宋体排版所有中文，这看起来比较单调，字体渲染器给出的粗体和倾斜体的效果也不是很好。另外，在不同的系统上，所需的字体名称可能并不是“宋体”。因此，使用者应当修改源文件，以符合自己的需求。例如，有些人可能希望使用楷体配合 *Italic*，黑体配合 **Boldface**。^{*}

另外注意，电报体属于等宽字体，因此 typezh 进行了特殊处理（将字体加宽 5%），使得一个汉字占据两个西文字符的宽度：

```
The quick brown fox jumps over the lazy dog. The quick...
那只敏捷的棕毛狐狸跃过了那条懒狗。那只敏捷的……
```

 当需要手动指定字体时，可以使用 \zhsinglefont 宏禁用字体自动切换功能，这样你就可以像正常使用 Plain TeX 那样，指定任意的字体。例如，上面的楷体文本可以使用

```
\zhsinglefont\tenkai 楷体
```

得到。（其中 \tenkai 是作者载入的 10 磅楷体字体。）但是这样也将会使用该字体排版西文字符，并很可能得到 **some ugly letters**. 使用

```
\zhdualfont<字体>
```

可以仅设置中文（实际上是 Unicode 字符集中的“中日韩表意文字”）及标点的字体，而西文字体仍然使用当前的字体。一个常见的用法是

```
\zhdualfont<中文字体><西文字体>
```

同时设置中西文的字体；这对于格式设计者而言是很有用的。作者的建议是将这两个宏封装在格式中，而不要出现在正文里。


^{*} 本文作者认为这是一种不恰当的设计，因为黑体属于非衬线体 (sans-serif)，不宜与衬线体 (serif) 搭配。

2. 标点压缩

由于中文字体中标点占据了一个字符的宽度,为了提高排版质量,需要对标点中的空白进行压缩。typezh 默认将中文标点压缩至半个汉字的宽度(正如你所见,本段的标点就是这样的风格)。这和 TeX 的西文标点采用了同样的约定:标点本身不含空白。作者建议用户在标点符号的旁边手动添加空格(尽管很多人并不习惯这样做),就和西文排版习惯一样。

另一些人使用另一种标点符号的风格。即不压缩标点,让它保持一个汉字的宽度(正如你所见,本段的标点就是这样的风格)。常在屏幕上阅读中文文本的用户可能会习惯于这种风格,但是从排版角度而言并不美观:“引号”和它左侧的冒号之间的空白看起来过大了。使用 \zhfullwidth 宏选择这种风格。要恢复默认的风格,使用 \zhhalfwidth。

一些用户可能习惯其他的风格,例如:标点符号默认为一个汉字的宽度,但是位于行尾的标点将被压缩到半个汉字的宽度。typezh 默认不支持这种风格,用户需要修改源代码来指定她所需要的具体的规则。

 当用户需要使用其他字号的字体时,应当在载入字体之后使用 \zhnormalskips 宏,重新设置标点压缩的大小。格式设计者在定义诸如 \tenpoint、\ninepoint 一类的宏时,尤其要注意这一点。

3. 标点禁则

由于中文用户大多不在标点符号旁边添加空格,除了带来空白压缩的问题以外,还带来了禁则的问题——TeX 不知道应该在哪里断行。typezh 处理了这些情况,使得 TeX 可以在左标点的左侧或者右标点的右侧断行;反过来的情况,则禁止断行。

4. 数学公式中使用中文

在技术文档中,偶尔会在公式中使用中文,例如作为变量的下标。typezh 允许直接在 TeX 的数学模式中使用中文,并以恰当的字体显示。


$$F_{\text{浮}} = \rho_{\text{液}} g V_{\text{排}}$$

在目前的 typezh 实现中,仅允许在公式中使用基本的 CJK 字符(Unicode 范围“4E00–9FFF”)。罕见的字符很少会出现在公式中。另外,不提倡在公式中使用中文标点,因为逗号的水平位置和西文字母不太匹配,而句号容易和小写字母 o 作下标弄混。

5. 字距调整

许多中文排版工具会在中西文之间自动添加空白,以免中文和 Western text 之间过于拥挤,影响美观。typezh 默认不做这样的事,简化了实现,而且把控制权完全交给用户。但是,用户也可以选择设置 \zhspace 参数,指定中西文之间自动插入的空白。

在每个汉字之间,typezh 会自动插入空白。如果不这样做,那么 TeX 将无法通过调整字间距,而使文本边界对齐。(对于西文而言,单词之间的空白是可以调整的。)设置 \zhglue 参数调节汉字之间的间距。


 但是,上述的机制在较复杂的情况中经常失效,尤其是遇到有分组或者 box 的情况。例如,设置 \zhspace=.2em plus .2em minus .1em, 则

使用 TeX 排版中文, 使用{\TeX}排版中文

给出的结果是

使用 TeX 排版中文, 使用TeX排版中文

这也是 typezh 默认不调整中西文间距的理由之一。在一些罕见的情况下,例如使用分组切换中文字体时,应当在分组前后添加 \hskip\zhglue, 以使得该文本和周围文本的间距被正确处理。

 例如,在排版这篇文档时,作者定义 \def\@{\hskip\zhglue}, 当作者在正文中插入**粗体字**, 而 TeX 又不得不调整字间距时,比较:

\@{\bf 粗体字\@} 在正文中插入**粗体字**时 ……
{\bf 粗体字} 在正文中插入**粗体字**时 ……

分组还会破坏汉字之间的断行,所以在汉字的上下文中使用分组时,请务必小心。