# 摘 要

对于论文排版而言，学界通常会采用LaTeX方案，LaTeX虽然有着种种优点，然而却并不是那么容易上手，很多时候我们也不得不提交Word格式的论文。Word虽然上手简单，可要排版出一篇格式工整的论文也并非那么容易。

那么有没有一种方案可以兼得二者的好处呢？或者说，在这中间同时保持对二者的支持呢？在有了pandoc这个工具之后，这个问题似乎有了方向。pandoc作为通用的文档转换器，可以将一种名为Markdown的纯文本格式转换为LaTeX，又可以把Markdown转换为Word文件；恰好LaTeX和Markdown因为都是纯文本，转换过程几乎不会受到阻碍。而Markdown本身，又有着两者无可替代的优越性——随着电子设备的不断普及，我们需要在纸张上排版的情况越来越少了，而Markdown正好符合了现代化电子记录的需求：实时渲染，设备无关，轻量化……所以乍一看似乎Markdown就是我们的救星。

然而Markdown因为其本身只是为记录而设计，并没有“格式”的概念，所以对于其而言，想要输出成一篇符合格式要求的Word文档，是很困难的。但是pandoc可以支持插件，在转换过程中进行处理！由此，我们的点子便借着这个插件功能出来了——

只需要**制作一个符合格式规范的Word模版**，而后**编写插件**处理中途Word在功能、格式上的问题，即可输出满足要求的成品Word文档

现在，你们看到的，便是由我们半年时间研究出的方案，使用 pandoc 生成出的接近成品的 word 文档。

该文档中的格式，均是来自《重庆大学博士、硕士学位论文撰写格式标准及要求（2019年修订）》、《重庆大学普通本科毕业设计（论文）撰写规范要求》，具体见后文

关键词：排版，论文，格式转换

# **Abstract**

按照排版规定英文标题需要加粗

Let us not wallow in the valley of despair, I say to you today, my friends.

And so even though we face the difficulties of today and tomorrow, I still have a dream. It is a dream deeply rooted in the American dream.

I have a dream that one day this nation will rise up and live out the true meaning of its creed: “We hold these truths to be self-evident, that all men are created equal.”

I have a dream that one day on the red hills of Georgia, the sons of former slaves and the sons of former slave owners will be able to sit down together at the table of brotherhood.

I have a dream that one day even the state of Mississippi, a state sweltering with the heat of injustice, sweltering with the heat of oppression, will be transformed into an oasis of freedom and justice.

I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character.

I have a dream today!

I have a dream that one day, down in Alabama, with its vicious racists, with its governor having his lips dripping with the words of “interposition” and “nullification” – one day right there in Alabama little black boys and black girls will be able to join hands with little white boys and white girls as sisters and brothers.

I have a dream today!

Keywords: pandoc, Markdown, Word

目 录

[摘 要 I](#_Toc61219675)

[Abstract II](#_Toc61219676)

[1 功能介绍 1](#_Toc61219677)

[1.1 标题 1](#_Toc61219678)

[1.2 English in title should be Bold 1](#_Toc61219679)

[1.2.1 三级标题**Level Three 3** 1](#_Toc61219680)

[1.3 脚注 1](#_Toc61219681)

[这是一个没有编号的二级标题 1](#_Toc61219682)

[1.4 标题的交叉引用 1](#_Toc61219683)

[2 其他自定义功能 2](#_Toc61219684)

[2.1 制表符 2](#_Toc61219685)

[2.1.1 普通制表符 2](#_Toc61219686)

[2.1.2 特殊制表符 2](#_Toc61219687)

[2.2 换行符分页符空段落生成 2](#_Toc61219688)

[2.2.1 换行符 2](#_Toc61219689)

[2.2.2 分页符 2](#_Toc61219690)

[2.2.3 空段落生成 3](#_Toc61219691)

[尚未实现的功能 3](#_Toc61219692)

[3 公式 4](#_Toc61219693)

[4 图片 5](#_Toc61219694)

[4.1 图片的插入 5](#_Toc61219695)

[4.2 图片的标记&尺寸定义 5](#_Toc61219696)

[4.3 双语题注 5](#_Toc61219697)

[4.4 图片的引用 6](#_Toc61219698)

[5 代码和文本块引用 7](#_Toc61219699)

[5.1 代码 7](#_Toc61219700)

[5.1.1 内联代码 7](#_Toc61219701)

[5.1.2 代码块 7](#_Toc61219702)

[5.2 引用文本块 8](#_Toc61219703)

[6 参考文献的插入 10](#_Toc61219704)

[参考文献 11](#_Toc61219705)

[附录 12](#_Toc61219706)

[A 12](#_Toc61219707)

[B 12](#_Toc61219708)

👆上面通过\toc{目 录}生成标题为目 录的目录

下面这里通过\newSection命令来生成一个Word中的分节符。在Word中，只有通过分节之后，才能给不同的节设定不同的页眉页脚以及纸张大小等参数。分节符的作用域为其前方的内容，这里的参数UpperRoman表示页码采用“大写罗马数字”表示，也就是从文档最开始的摘要到目录的部分，页码采用大写罗马数字表示。👇

# 功能介绍

这里开始介绍基本功能的使用。

## 标题

这里是各级标题，根据论文排版要求，本科论文理科规定了6级标题样式的编号和格式，文科则规定了5级，且编号格式与理科不同；而根据研究生的要求，文理科都只能按照本科理科的样子编号。其中英文标题要求应用加粗效果。

## **English in title should be Bold**

### 三级标题**Level Three 3**

#### 四级标题**Level Four 4**

##### 五级标题**Level Five 5**

###### 六级标题**Level Five 6**

更多级别的标题在模版中未调整样式，请勿使用

根据需求，“摘要”“参考文献”“附录”等样式虽为一级标题，但是并没有编号，可以通过在标题尾部加上{-}来取消编号。考虑到实际情况，模版里只实现了对一级标题和二级标题的支持。

## 脚注

按照规定，脚注是被允许的，[[1]](#footnote-1)可以直接按照Markdown的脚注格式进行书写[[2]](#footnote-2)，这里就有了几个脚注了[[3]](#footnote-3)。脚注的编号可以乱写[[4]](#footnote-4)，只要一一对应就好[[5]](#footnote-5)，最后会自动按顺序编好。

## 这是一个没有编号的二级标题

## 标题的交叉引用

测试标题的交叉引用：1.4

页码：1

标题文本：标题的交叉引用

# 其他自定义功能

## 制表符

制表符可以用来对齐文字，在Word中，有普通的制表符（按Tab键插入）和特殊的制表符（需要通过菜单插入）。

### 普通制表符

长江，长城， 黄山，黄河， 在我心中重千斤

无论何时， 无论何地， 心中一样亲

普通制表符默认在文档中按照“两个字符”的位置进行对齐（可在“标尺”中查看），所以这里恰好就把“在我心中重千斤”给对齐到了下一个位置，需要手动在“心中一样亲”之前再加一个制表符才能对齐。

### 特殊制表符

这是一串居中的文字。现在我跑到下一行的最左边了。然后我又到了这一行的最右边。

床前明月光，疑是地上霜。

举头望明月，低头思故乡。

*——李白《静夜思》*

## 换行符分页符空段落生成

### 换行符

有的时候我们需要在一段话的中间换行，而不进行分段，这时候就需要用到换行符，也就是LaTeX中的\\，可以用\newLine{}实现。👉  
👈比如这里就到了一个新行当中而没有因为分段产生首行缩进。后面的双语题注也是用的换行符实现的换行不分段。

### 分页符

这也是Word的中的特殊功能，等同于LaTeX中的\newpage，但是这里为了和LaTeX以示区分，采用了更贴近于程序员的“小驼峰命名法”，即应当写成\newPage

妈妈：不听话就滚到下一页去，眼不见心不烦

我：呜呜~~妈妈不要我了😥

### 空段落生成

pandoc无法将多余的空行识别为空的段落产生一段垂直间距（比如留出数学题的解题间隔），LaTeX中可以采用生成垂直间距的命令来实现，但Word没有这样的功能，只能用空段落来生成间距。这里采用命令\newPara{n}来生成n个空段落。👇

👆这里就有了3个空段落。排版规范要求“摘要”部分的“关键词”与正文间隔一行，内部就是用的这个逻辑来实现。

## 尚未实现的功能

1. 对指定段落套用自定义样式
   * 实现起来不复杂，单纯还没做而已
   * LaTeX中的“引理”“定理”“证明”等环境可通过该功能实现
2. 图片的并排子图
   * 这个功能得用表格进行嵌套，理解起来很简单，做是能做的
   * 就是逻辑比较复杂，代码写起来应该会很酸爽
3. 附录的公式表格编号
   1. 排版要求规定一级标题“附录”不编号，附录里面的二级标题以ABCD编号
   2. 而附录里的图片和表格等用图A1的形式编号
   3. 实现起来虽然不复杂，但是要考虑到和普通的题注编号的共存，所以判断逻辑较为复杂
4. 图目录，表目录，符号表
5. 任意域代码的插入
   * 同自定义样式

# 公式

勾股定理：设顶点，，的对边分别为 ，，，若成立

|  |  |  |
| --- | --- | --- |
|  |  | (3.1) |

则 是直角三角形，且

|  |  |  |
| --- | --- | --- |
|  |  | (3.2) |

通过预定义的开关开启公式编号功能之后，“显示公式”默认会被编号，实现原理是用一个三列的表格嵌套公式，中间放置公式，右边放置编号，由于该表格的对称性，公式被对齐在正中间。这里可以简单的用一用LaTeX的\newcommand命令，由pandoc提供原生支持。

|  |  |  |
| --- | --- | --- |
|  |  | (3.3) |
|  |  |  |
|  |  | (3.4) |
|  |  | (3.5) |

上述4个公式中，第二个公式后的{-}同标题一样，表示不编号；  
 第三个公式后的{#eq1}表示对公式打上名为eq1的标记，就可以进行交叉引用了，使用[@eq1]见公式(3.4)完成对该编号的引用

如果要在公式后面输入{#a}这样的文字，可以先用空的花括号{}隔开

|  |  |  |
| --- | --- | --- |
|  |  | (3.6) |

{#a}

|  |  |  |
| --- | --- | --- |
|  |  | (3.7) |

{#a}

# 图片

## 图片的插入



图 4.1 这里需要一个图片

通过上面这样![题注](图片路径)的语法可以插入一张图片。和公式一样，在有题注的情况下，图片默认编号。

## 图片的标记&尺寸定义

同样也可以采用和公式一样的语法来引用图片的编号，图片甚至可以设置宽度和高度，完整语法为：![题注](图片路径){#fig1 weight=3cm height=2cm}，例如：



图 4.2 这是一张设置了尺寸的图片

👆上面这张图有编号有标签fig0，  
下面这张图无编号无标签👇



这是一张设置了尺寸的图片

## 双语题注

有时候图片的题注需要同时有中文和英文，使用语法![中文题注\Caption2{fig}英文题注](){}的方式生成英文题注。其中中文题注的编号自动完成，\Caption2命令通过在中文题注后面插入换行符和英文题注的前缀及序号来实现英文题注的插入



图 4.3 这里是中文题注  
Fig. 4.3 English Caption is also needed

## 图片的引用

与公式不同的是，由于图片的复杂要求，以对上图为例，对其编号的引用语法为[@fig-DoubleCaption-no]，对其中文题注内容的引用语法为[@fig-DoubleCaption-zh]，对其英文题注的引用语法为[@fig-DoubleCaption-en]。

即图👉4.3👈的中文题注内容为👉这里是中文题注👈，英文题注内容为👉English Caption is also needed👈。无论图片是否具有英文题注，其中文题注的引用都需要有后缀-zh

图1：单语有标签 fig1（编号标签fig1-no，中文标题标签fig1-zh）



图 4.4 我是单语中文题注

图2：双语有标签 fig2（同图 1，外加英文标题标签fig2-en）



图 4.5 我是双语的中文题注  
Fig. 4.5 I am Thanox

图3：无编号双语标签



中文  
English

上面3张图片中，图4.4，其题注为我是单语中文题注；图4.5，其中文题注为：我是双语的中文题注，英文题注为：I am Thanox。

# 代码和文本块引用

这两个语法属于Markdown的语法，学校排版要求并未做出明确规定

## 代码

### 内联代码

首先是一行内部的内联代码👉这里是内联代码some code👈

### 代码块

接着是单独成段出现的代码块

if a == 1:  
 print("str")

纯文本测试  
这里和上面的两行python代码会在一个框里  
plain text

下面是一段xml代码

<w:lvl w:ilvl="3">  
 <w:start w:val="1"/>  
 <w:numFmt w:val="decimalEnclosedCircleChinese"/>  
 <w:pStyle w:val="4"/>  
 <w:lvlText w:val="%4"/>  
 <w:lvlJc w:val="left"/>  
 <w:pPr>  
 <w:ind w:left="0" w:firstLine="0"/>  
 </w:pPr>  
</w:lvl>

下面是一段python代码

# 绝对素数.py  
# 将自然数区间10-1000中所有绝对素数挑选出来，每行打印输出5个。绝对素数是：一个数和其反序数都是素数。  
List = [] # 创建空数组  
flag = 1  
count = 0 # 计数器初始状态为0  
for i in range(10, 1001): # 自然数i在10-1000之间  
 for j in range(2, i):  
 if i % j == 0: # 自然数i是否是素数  
 flag = 0  
 break # 只要遇到一个因数，就可以不用进行后面的循环了  
 if flag == 1: # 如果经过上述判断flag还是1，说明正向是素数，现在判断反向素数  
 s = str(i)  
 m = int(s[::-1])  
 for n in range(2, m):  
 if m % n == 0: # 自然数i的反序数m是否是素数  
 flag = 0  
 break # 只要遇到一个因数，就可以不用进行后面的循环了  
 if flag == 1: # 将绝对素数全部添加至列表  
 List.append(i)  
 flag = 1  
for k in List: # 将list打印输出，五个一换行  
 print(k, end=' ')  
 count = count+1  
 if count % 5 == 0:  
 print('\n')

## 引用文本块

下面是一段对《歌剧魅影》剧本唱段台词的引用块。（这里是借助文本编辑器通过正则表达式批量生成的规范化文本，仅作炫技展示之用）

In dreams he came  
♪ 梦境中他近我身旁

That voice to me  
♪ 他的声音召唤着我

And speaks my name  
♪ 喃喃轻诉我的名字

And do I dream again  
♪ 我是否又重回梦境

For now I find  
♪ 此时此刻我才发现

The Phantom of the Opera is there  
♪ 歌剧魅影就在这里

Inside my mind  
♪ 潜伏在我心灵深处

Sing once again with me  
♪ 再次与我一起

Our strange duet  
♪ 同唱这奇妙的旋律

# 参考文献的插入

参考文献的插入和LaTeX类似，通过一个.bib文件来实现，该文件可通过任意文献管理软件导出，为纯文本格式，通过[@xxx]这样的语法来进行引用。

而参考文献的格式由一个.csl格式的文件来控制，该文件可从citationstyle.org组织的github仓库下载<https://github.com/citation-style-language/styles>。也就是我们可以简单的替换一下这个文件，实现参考文献样式的“一键”切换

这里右上角会出现参考文献的引用[1]

一条参考文献不够，再来两个一起试试[2–4]

# 参考文献

pandoc默认参考文献会生成在文档的最后方，但根据论文规范，参考文献之后还有“附录”，所以这里通过命令\Reference来实现在任意位置插入参考文献👇，根据排版要求，参考文献的内容采用五号字体。

[1] JIFA G, LINGLING Z. Data, DIKW, Big Data and Data Science[J]. Procedia Computer Science, 2014, 31(0): 814–821. DOI:[10.1016/j.procs.2014.05.332](https://doi.org/10.1016/j.procs.2014.05.332).

[2] PERNER P. Mining Sparse and Big Data by Case-based Reasoning[J]. Procedia Computer Science, 2014, 35(0): 19–33. DOI:[10.1016/j.procs.2014.08.081](https://doi.org/10.1016/j.procs.2014.08.081).

[3] WEICHSELBRAUN A, GINDL S, SCHARL A. Enriching semantic knowledge bases for opinion mining in big data applications[J]. Knowledge-Based Systems, 2014, 69(0): 78–85. DOI:[10.1016/j.knosys.2014.04.039](https://doi.org/10.1016/j.knosys.2014.04.039).

[4] YANG S, GUO K, LI J, 等. Framework Formation of Financial Data Classification Standard in the Era of the Big Data[J]. Procedia Computer Science, 2014, 30(0): 88–96. DOI:[10.1016/j.procs.2014.05.385](https://doi.org/10.1016/j.procs.2014.05.385).

[5] XU M, RHEE S Y. Becoming data-savvy in a big-data world[J]. Trends in Plant Science, 2014, 19(10): 619–622. DOI:[10.1016/j.tplants.2014.08.003](https://doi.org/10.1016/j.tplants.2014.08.003).

[6] PHILIP CHEN C L, ZHANG C-Y. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data[J]. Information Sciences, 2014, 275(0): 314–347. DOI:[10.1016/j.ins.2014.01.015](https://doi.org/10.1016/j.ins.2014.01.015).

[7] 于秀清. F-数据簇与缺损数据修复-还原[J]. 计算机工程与应用, 2013, 49(6): 139–142. DOI:[10.3778/j.issn.1002-8331.1108-0086](https://doi.org/10.3778/j.issn.1002-8331.1108-0086).

# 附录

学校要求附录的字号为五号，因此这里需要通过在标题尾部写上{.appendix}来让插件识别并处理。

## A

这是附录A中的内容，对附录中图片和公式的编号处理目前尚未完成。

## B

附录B中的内容

由于附录的字号比正文小，所以这里测试一下在正文中的内联代码字号是否正常👉内联代码code code👈可见内联代码的字号也跟着正文变小了

文档最后这里也是有一个分节符的，这个分节符是文档自带的，在pandoc转换的过程中，文档尾部的分节符会从模版中继承下来，也就不需要手动设定了。这里的分节符在模版中已经设定好，采用阿拉伯数字从1开始编号。👇

1. 脚注：诶，上面的兄弟伙，你说你一根横线在我们脚注脑壳上面，应该跟我们一路对齐的撒，啷个子跟别个那些正文伙到一起去把脑壳歪起生起哟，大家一起对正标齐，看到起好威风哦 [↑](#footnote-ref-1)
2. 横线：嗐，老弟呀，不是我不想和你们对齐呀，我跟你讲嘛，是楞个的。虽然说我不是脚注，但是我和你们脚注一起走南闯北穿一条裤子的，按理说也该按脚注的待遇来撒，结果那些背时灾舅子上户口的，硬是把我给上成是正文的户口。结果嘞哈斗没得法的了，只有跟他们正文一路把脑壳歪起了 [↑](#footnote-ref-2)
3. 脚注：瓜兮兮，开飞机；神戳戳，骑摩托。嘞些个宝批龙怕是没读过书哦 [↑](#footnote-ref-3)
4. 横线：是啊是啊，唉不摆了不摆了 [↑](#footnote-ref-4)
5. …… [↑](#footnote-ref-5)