|  |  |
| --- | --- |
| Федеральное агентство железнодорожного транспорта  Омский государственный университет путей сообщения  Кафедра «Автоматика и системы управления»  ФОРМАТ DOCX  Лабораторная работа № 7  по дисциплине «Информационные технологии» | |
|  | Студент гр. 22м                            Н.И. Яковлев  «    »                  2022 г.  Руководитель –  старший преподаватель кафедры «АиСУ»                            Т. В. Васеева  «    »                  2022 г. |
| Омск 2022 | |

Цель работы

Разобраться в устройстве формата и способах взаимодействия с ним.

# Выполнение работы

## Предустановки к работе

Для начала необходимо конвертировать md-файл из второй лабораторной работы в формат docx. Для этого воспользуемся программой «pandoc». После конвертации, получим следующей результат (рисунок 1).



Рисунок  – Файл второй лабораторной работы в формате docx

## Сравнение документов

Сравнение приведено в таблице 1.

Таблица 1 – Сравнение markdown и docx документов

|  |  |  |
| --- | --- | --- |
| Описание | markdown | docx |
| Заголовок 1 | # ... | Heading 1 |
| Заголовок 2 | ## ... | Heading 2 |
| Заголовок 3 | ### ... | Heading 3 |
| Заголовок 4 | #### ... | Heading 4 |
| Заголовок 5 | ##### ... | Heading 5 |
| Заголовок 6 | ###### ... | Heading 6 |
| Нумерованный список | 1. …  2. …  …  n. … | Compact + number list |
| Курсив | \*…\* | Italic |
| Полужирный | \*\*…\*\*, \_\_...\_\_ | Bold |
| Зачеркнутый | ~~…~~ | Strikethrought |
| Полужирный курсив | \*\*\*…\*\*\* | Bold + Italic |
| Цитата | > | Block Text |

Продолжение таблицы 1

|  |  |  |
| --- | --- | --- |
| Описание | markdown | docx |
| Вложенная цитата | >> | Block Text |
| Маркированный список | \* …, - …, + … | Bullet list |
| Таск-лист | - [ ] …  - [x] …  … | Compact + bullet list |

## Преобразование на основе шаблона

Для начала необходимо создать эталонный файл (шаблон) и настроить его стили. Затем, к ранее использованной команде для преобразования .md в .docx, добавим параметр командной строки «--reference-doc=reference.docx», где после символа «=» указывается полное название файла-шаблона. Полученный результат представлен на рисунке 2.

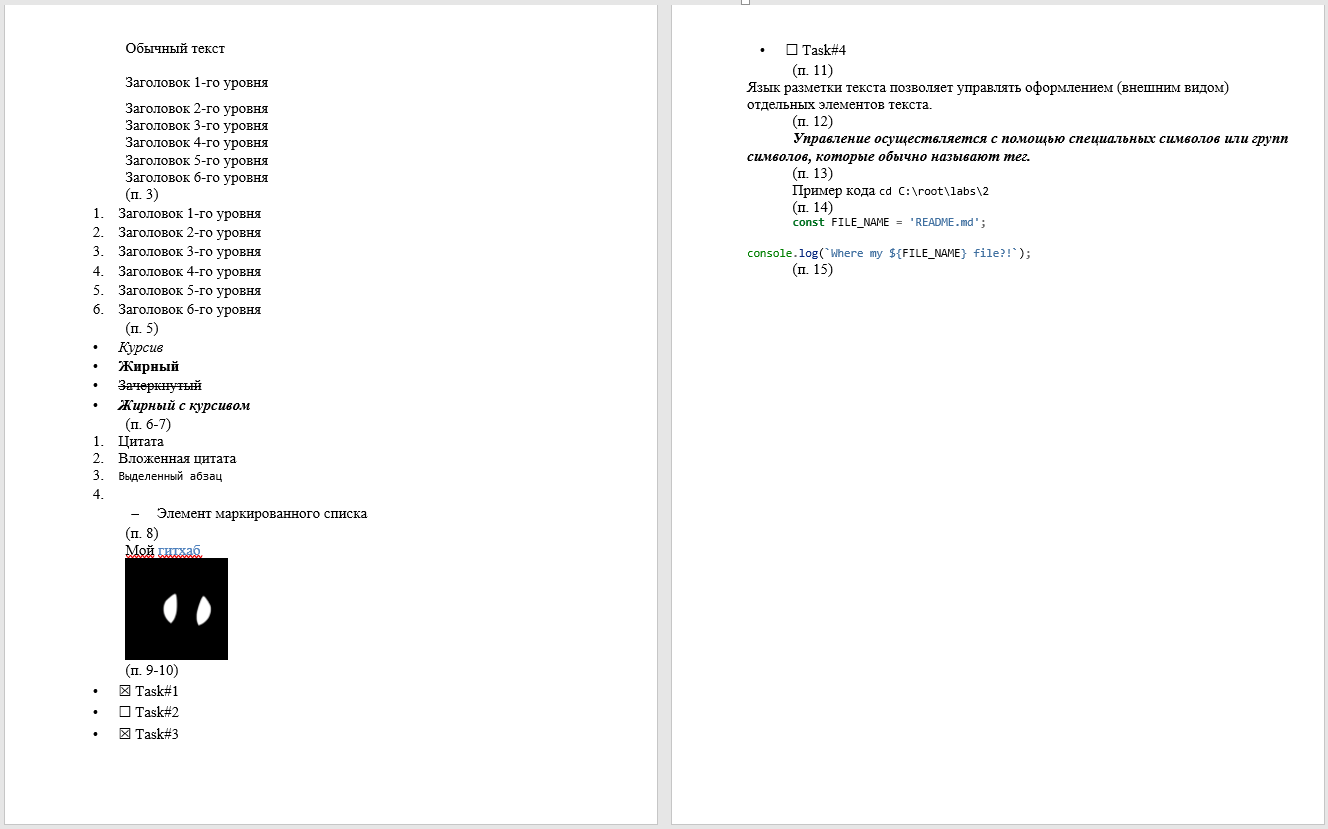


Рисунок  – Результат преобразования

Анализируя полученный результат, можно сказать о том, что стили из ранее созданного шаблона были применены в процессе текущего преобразования.

## Распакованный .docx

При распаковке файла .docx через «7-zip», получаем папку, приведенную на рисунке 3.

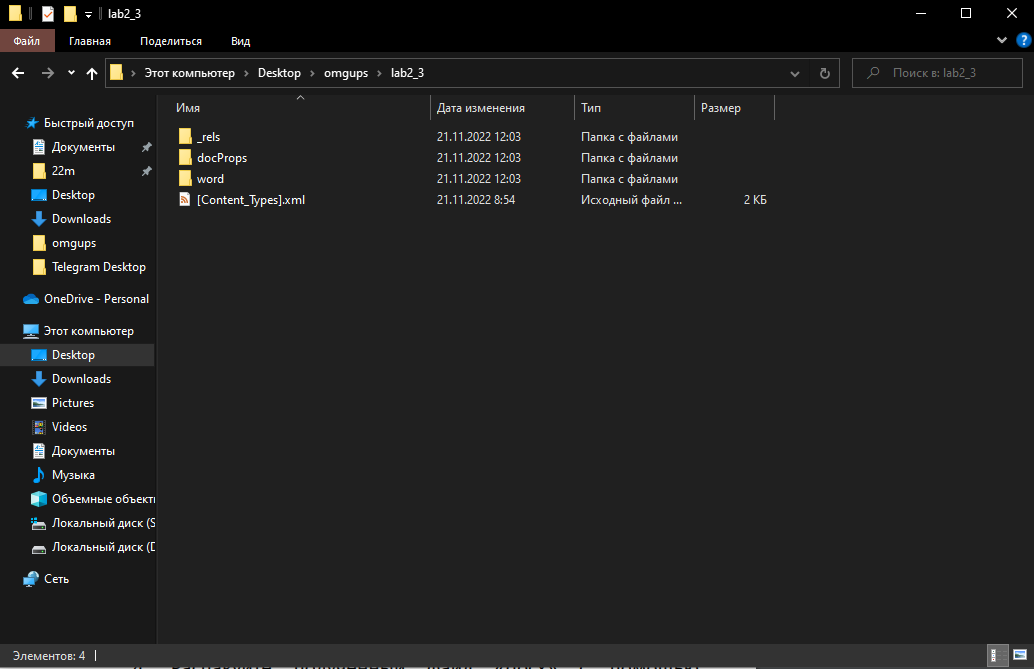


Рисунок  – Распакованный.docx

Перейдем в папку word и откроем в текстовом редакторе файлы «document.xml» и «styles.xml». Рассмотрим рисунок 4.

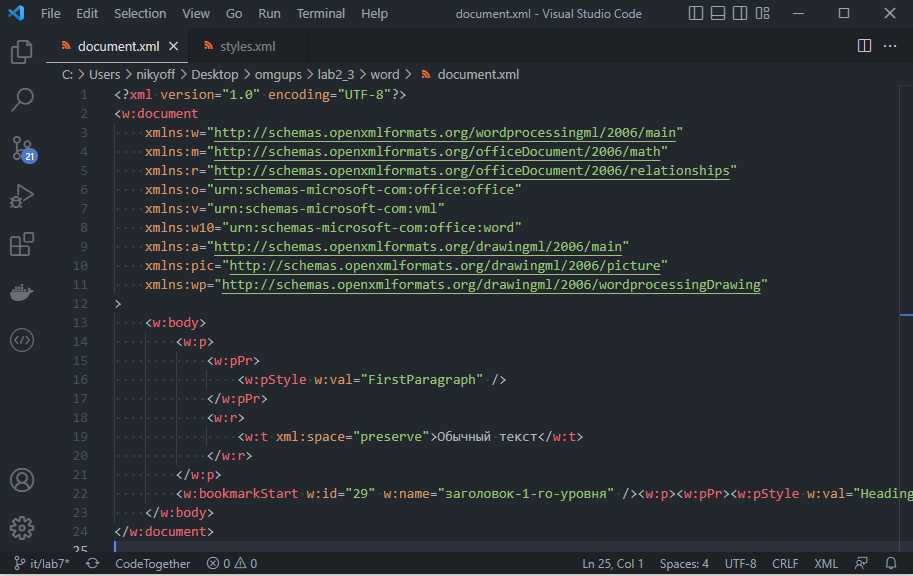


Рисунок  – Файл «document.xml» в текстовом редакторе

Выделить какие-либо преимущества возможно при сравнении одного языка разметки с другим, поэтому имеет смысл рассматривать язык разметки Word опираясь на подобный, например HTML.

Основным недостатком языка разметки Word является избыток тегов для какой-либо метаинформации, подобное можно переместить в атрибуты, что в свою очередь дает преимущество при чтении самой структуры документа. В свою очередь преимуществом является простота самого языка xml, который дает структурированный подход к разметке текста.

## Теги стилей, активных полей в Word

Распакуем файл .docx из предыдущей лабораторной работы и найдем в файле word.xml необходимые нам теги. Для указания стиля (например заголовка) указывается тег «w:pStyle» обратимся к рисунку 5, а для указания активного поля (формул) используется группа тегов, обратимся к рисунку 6.



Рисунок  – Структура тегов, в которой определен стиль

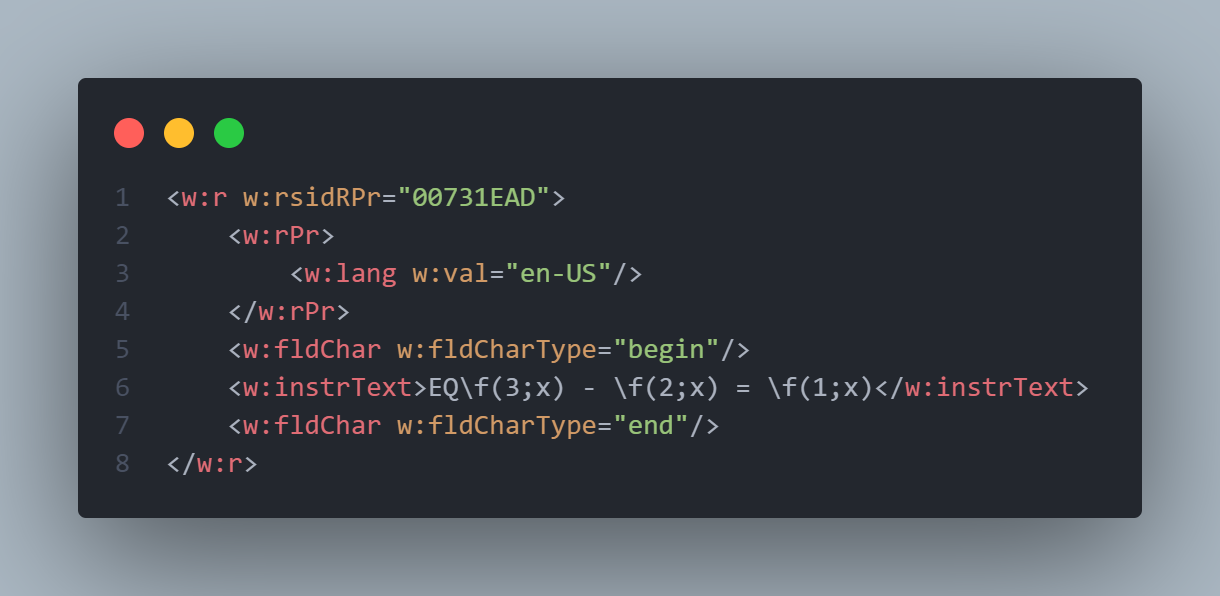


Рисунок  – Структура тегов для определения активного элемента (формулы)

# Контрольные вопросы

## Convert Markdown(pandoc) to LaTeX

Для конвертации markdown в LaTeX, следует указать соответствующие расширение выходного файла для команды конвертации pandoc: «pandoc -o latex\_file.tex lab2.md». Ключ «o» расшифровывается, как «output», он нужен для указания выходного файла. После конвертации получим следующий результат (рисунок 7).

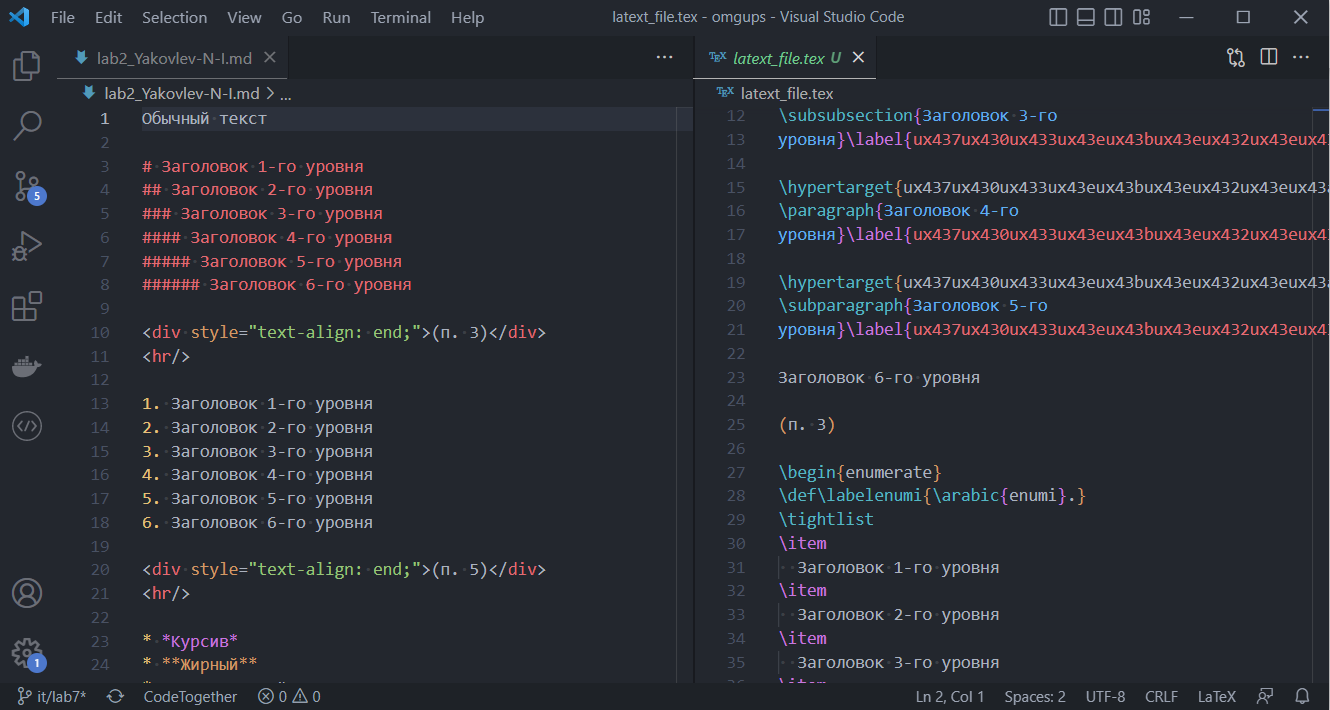


Рисунок  – Результат конвертации Markdown в LaTeX

## Convert Markdown(pandoc) to EPUB версии 2 или 3

Для конвертации markdown в EPUB 3 версии, следует указать соответствующие расширение выходного файла для команды конвертации pandoc: «pandoc -o epub\_file.epub lab2.md», по умолчанию будет выбрана третья версия epub. Ключ «o» расшифровывается, как «output», он нужен для указания выходного файла. После конвертации получим следующий результат (рисунок 8).

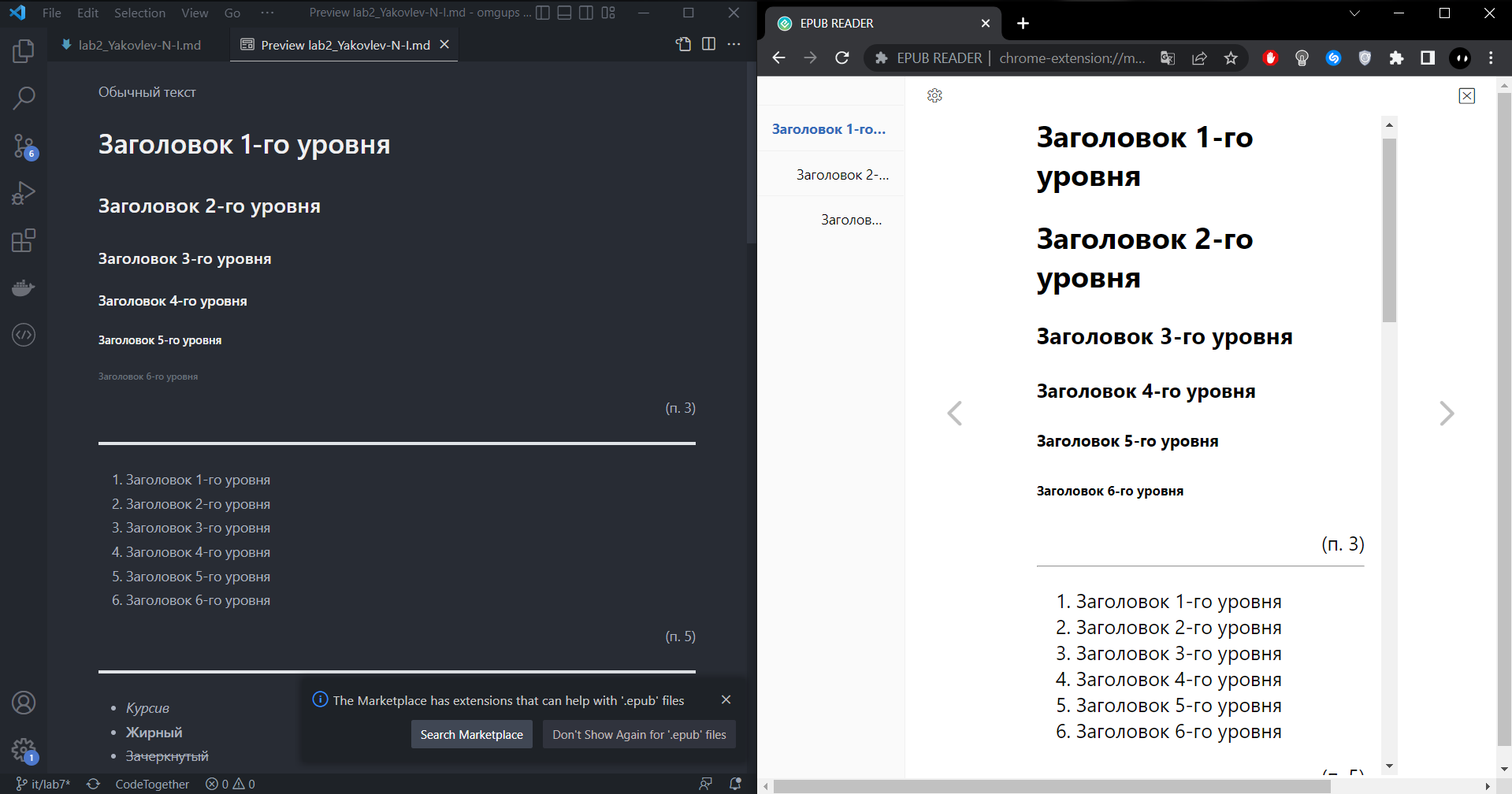


Рисунок  – Результат конвертации Markdown в EPUB (превью)

## Convert Markdown(pandoc) to HTML5

Для конвертации markdown в HTML5, следует указать соответствующие расширение выходного файла, а также дополнительный ключ –to=html5 для указания версии конвертации: «pandoc –to=html5 -o html5**\_**file.html lab2.md». Ключ «o» расшифровывается, как «output», он нужен для указания выходного файла. Ключ «to» нужен для указания версии конвертации. После конвертации получим следующий результат (рисунки 8-9).

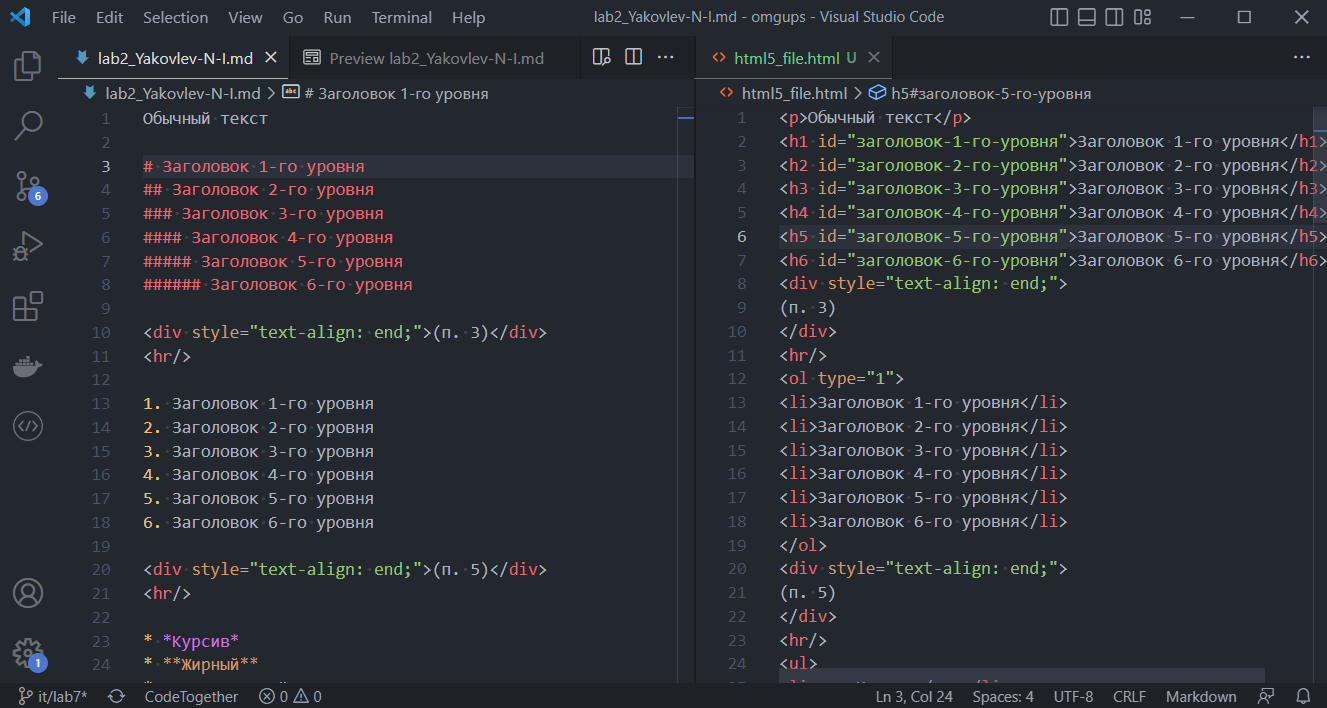


Рисунок  – Результат конвертации Markdown в HTML5

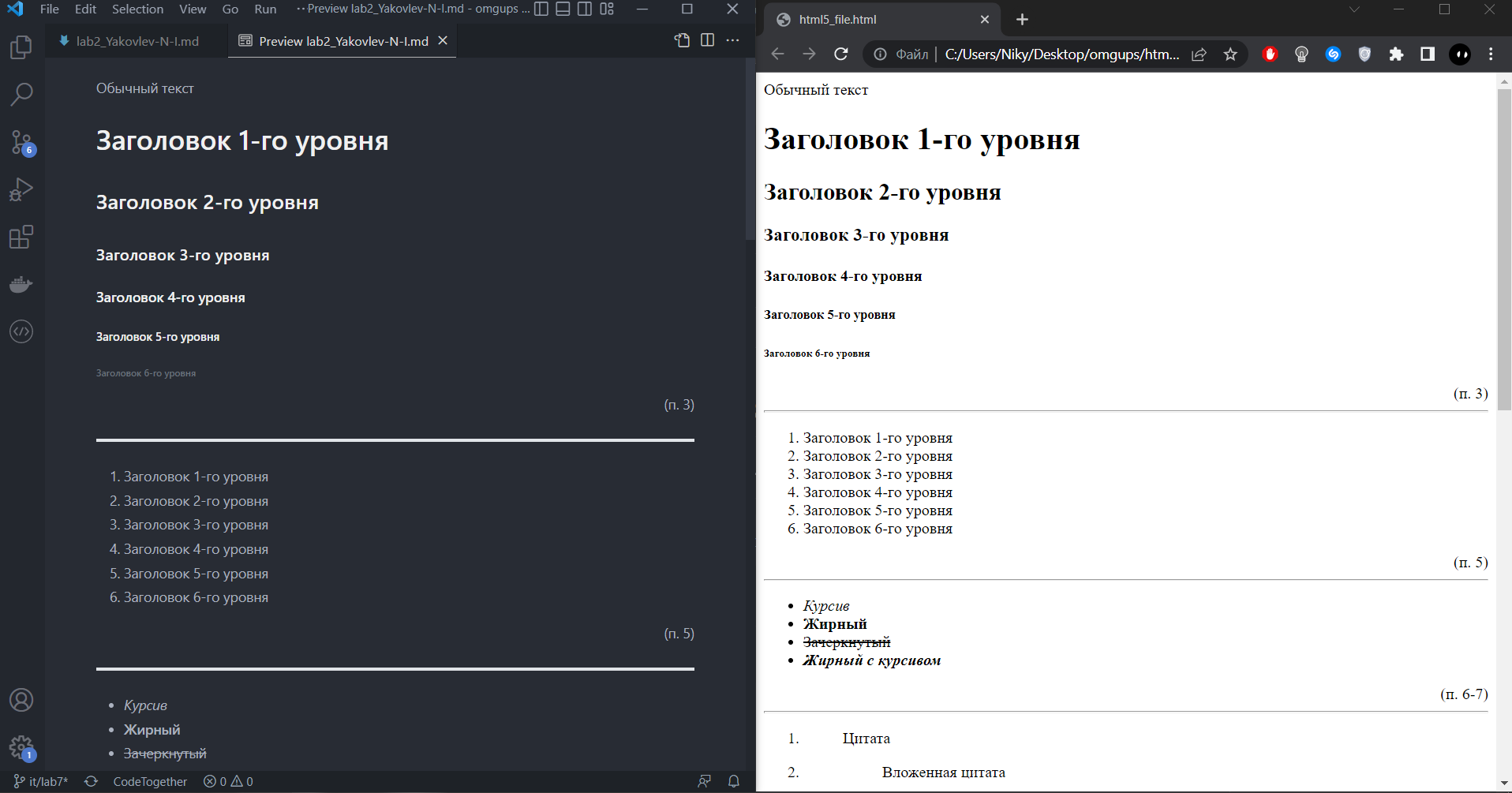


Рисунок  – Результат конвертации Markdown в HTML5 (превью)

Вывод

В данной лабораторной работе, мы научились работать с форматом файла DOCX, разобрались в способах конвертации документов в данный формат, а также затронули внутреннее устройство формата.