Для разбиения отдельного pdf документа на страницы воспользуемся следующей программой:

**from PyPDF2 import PdfFileReader, PdfFileWriter**

**pdf\_document = "D:\User\Костерин\Поддомены\source\Формула включений и исключений.pdf"**

**pdf = PdfFileReader(pdf\_document)**

**for page in range(pdf.getNumPages()):**

**pdf\_writer = PdfFileWriter()**

**current\_page = pdf.getPage(page)**

**pdf\_writer.addPage(current\_page)**

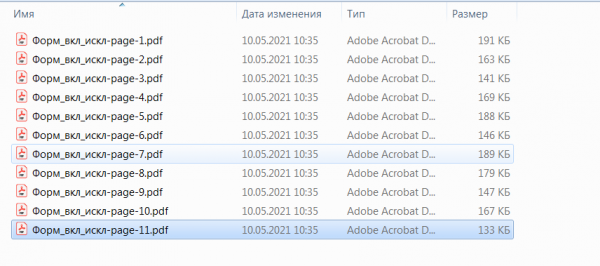
**outputFilename = "dist/Форм\_вкл\_искл-page-{}.pdf".format(page + 1)**

**with open(outputFilename, "wb") as out:**

**pdf\_writer.write(out)**

**print("created", outputFilename)**

Результат представлен на следующем рисунке:



Найти все страницы, где есть заданный текст:

**import fitz**

**filename = "D:\User\Костерин\Поддомены\source\Формула включений и исключений.pdf"**

**search\_term = "множество"**

**pdf\_document = fitz.open(filename)**

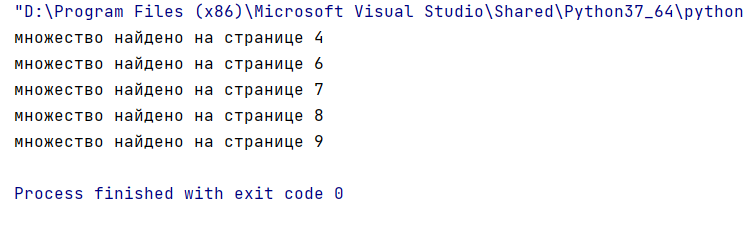
**for current\_page in range(len(pdf\_document)):**

**page = pdf\_document.loadPage(current\_page)**

**if page.searchFor(search\_term):**

**print("%s найдено на странице %i" % (search\_term, current\_page+1))**

Осуществим поиск в документе Формула включения и исключения.pdf слова «множество».  
Результат работы программы:



Добавление водяного знака с помощью PyPDF2  
Добавим в pdf файл водяной знак «Черновик». Текст программы:

**# Добавление водяного знака в одностраничный PDF**

**import PyPDF2**

**input\_file = "D:\User\Костерин\Поддомены\source\Задание 12.3.pdf"**

**output\_file = "dist/Водяной\_знак-page-drafted.pdf"**

**watermark\_file = "source/Черновик.pdf"**

**with open(input\_file, "rb") as filehandle\_input:**

**# читать содержимое исходного файла**

**pdf = PyPDF2.PdfFileReader(filehandle\_input)**

**with open(watermark\_file, "rb") as filehandle\_watermark:**

**# читать содержание водяного знака**

**watermark = PyPDF2.PdfFileReader(filehandle\_watermark)**

**# получить первую страницу оригинального PDF**

**first\_page = pdf.getPage(0)**

**# получить первую страницу водяного знака PDF**

**first\_page\_watermark = watermark.getPage(0)**

**# объединить две страницы**

**first\_page.mergePage(first\_page\_watermark)**

**# создать объект записи PDF для выходного файла**

**pdf\_writer = PyPDF2.PdfFileWriter()**

**# добавить страницу**

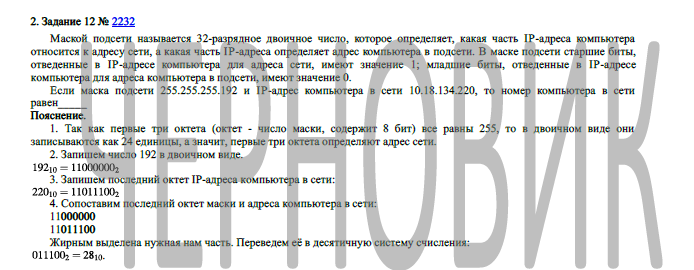
**pdf\_writer.addPage(first\_page)**

**with open(output\_file, "wb") as filehandle\_output:**

**# записать файл с водяными знаками в новый файл**

**pdf\_writer.write(filehandle\_output)**

Результат работы программы:



Удаление страниц с помощью pdfrw  
Для удаления страниц необходимо установить библиотеку pdfrw. Текст программы:

**# Удалите первые две страницы (титульный лист) из PDF**

**from pdfrw import PdfReader, PdfWriter**

**input\_file = "D:\User\Костерин\Поддомены\source\Формула включений и исключений.pdf"**

**output\_file = "dist/Удаление\_страниц-page-drafted.pdf"**

**# Определить объекты чтения и записи**

**reader\_input = PdfReader(input\_file)**

**writer\_output = PdfWriter()**

**# Перейти на страницу один за другим**

**for current\_page in range(len(reader\_input.pages)):**

**if current\_page > 1:**

**writer\_output.addpage(reader\_input.pages[current\_page])**

**print("adding page %i" % (current\_page + 1))**

**# Записать измененный контент на диск**

**writer\_output.write(output\_file)**

Результат работы программы – сформирован новый файл, у которого вырезано 2 первые страницы.

