

Министерство науки и высшего образования Российской Федерации
Санкт-Петербургский политехнический университет Петра Великого
Физико-механический институт

Работа допущена к защите
Руководитель ОП
_____ К.Н. Козлов
« _____ » _____ 2023 г.

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
РАБОТА БАКАЛАВРА**

**ОПТИМИЗАЦИЯ КОЛИЧЕСТВА ВЫЗОВОВ ОТРИСОВКИ В
СОВРЕМЕННЫХ ГРАФИЧЕСКИХ СИСТЕМАХ**

по направлению подготовки 01.03.02 "Прикладная математика и информатика"
Направленность (профиль) 01.03.02_02 "Системное программирование"

Выполнил
студент гр. 5030102/90201

В.А. Парусов

Руководитель
доцент ВШПМиВФ,
степень, звание

С.Ю. Беляев

Консультант
старший преподаватель, ВШПМиВФ

В.С. Чуканов

Санкт-Петербург
2023

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ПЕТРА ВЕЛИКОГО
Физико-механический институт

УТВЕРЖДАЮ

Руководитель образовательной программы «Прикладная математика и информатика»

К.Н. Козлов

«__» _____ 202__ г.

ЗАДАНИЕ

на выполнение выпускной квалификационной работы
студенту Парусову Владимиру Алексеевичу, гр. 5030102/90201

1. Тема работы: «Оптимизация количества вызовов отрисовки в современных графических системах»
2. Срок сдачи студентом законченной работы: июнь 2023 г.
3. Исходные данные по работе:

Инструментальные средства:

- язык программирования C++
- среда разработки Visual Studio 2022
- программная библиотека для работы с видеокартой DirectX 12
- система контроля версий git


Ключевые источники литературы:

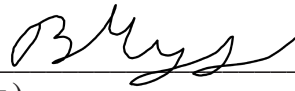
1. Riccio C., Lilley S. Introducing the programmable vertex pulling rendering pipeline //GPU Pro. – 2013. – Т. 4. – С. 21-37
2. Park H., Han J. H. Fast rendering of large crowds using GPU //International Conference on Entertainment Computing. – Springer, Berlin, Heidelberg, 2008. – С. 197-202.

4. Содержание работы (перечень подлежащих разработке вопросов):

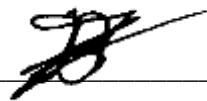
- 1) Введение. Обоснование актуальности
- 2) Постановка задачи
- 3) Обзор существующих решений
- 4) Предлагаемое решение
- 5) Результаты и их сравнительный анализ
- 6) Заключение

5. Дата выдачи задания 07.12.2022

Руководитель ВКР  _____ С.Ю. Беляев
(подпись)

Консультант ВКР  _____ В.С. Чуканов
(подпись)

Задание принял к исполнению

Студент  _____ В.А. Парусов
(подпись)

РЕФЕРАТ

На 18 с., 9 рисунков, 1 таблицу, 2 приложения

КЛЮЧЕВЫЕ СЛОВА: TBD: СТИЛЕВОЕ ОФОРМЛЕНИЕ САЙТА, УПРАВЛЕНИЕ КОНТЕНТОМ, PHP, MYSQL, АРХИТЕКТУРА СИСТЕМЫ.

Тема выпускной квалификационной работы: «Оптимизация количества вызовов отрисовки в современных графических системах»

Данная работа посвящена разработке и реализации архитектуры конвейера отрисовки трехмерной компьютерной графики, снижающей нагрузку на центральный процессор, тем самым увеличивая объем ресурсов компьютера, доступных для разработчиков ПО.

В качестве основной демонстрационной сцены используется совокупность различных трехмерных геометрических объектов, освещённых при помощи подхода, основанного на физической модели микрограней. Весь исходный код проекта написан на языке C++ с применением графической библиотеки DirectX 12. Средством для программирования шейдеров является HLSL.

Наиболее значимым результатом является система для вывода трёхмерных сцен, количество вызовов отрисовки в которой не зависит от числа выводимых объектов. В ходе работы был также разработан гибридный алгоритм отрисовки прозрачных объектов, который не требует их упорядочивания.

Предложенная архитектура пригодна для применения в современных графических приложениях и может быть развита в дальнейшем.

ABSTRACT

18 pages, 9 figures, 1 tables, 2 appendices

KEYWORDS: TBD: STYLE REGISTRATION, CONTENT MANAGEMENT, PHP, MYSQL, SYSTEM ARCHITECTURE.

The subject of the graduate qualification work is «Optimizing the number of draw calls on modern graphics systems».

The work aims to develop and implement graphics pipeline architecture, which will reduce usage of CPU and thereby increase amount of resources available to developers.

Main demonstration scene is a set of different geometric shapes, which uses physically-based lighting with microfacet model. All source code of project is written in C++ using the DirectX 12 graphics library. HLSL is used for shaders code.

The most significant result is a system for rendering three-dimensional scenes, the number of draw calls in which does not depend on the number of rendered objects. In the course of the work, a hybrid algorithm for rendering transparent objects was also developed, which does not require object sorting.

The proposed architecture is suitable for use in modern graphics applications and can be further developed to obtain more efficient results.

СОДЕРЖАНИЕ

РЕФЕРАТ	4
ABSTRACT	5
СОДЕРЖАНИЕ	6
ВВЕДЕНИЕ.....	7
ГЛАВА 1. ПОСТАНОВКА ЗАДАЧИ	9
1.1. Техническое задание	9
1.2. Ожидаемый результат	9
ГЛАВА 2. ОБЗОР СУЩЕСТВУЮЩИХ РЕШЕНИЙ	9
2.1. Использование Geometry Instancing	9
2.2. Использование Programmable Vertex Pulling	12
ГЛАВА 3. Предлагаемое решение.....	13
3.1. Название параграфа	13
3.2. Название параграфа	13
3.3. Выводы	13
ГЛАВА 4. Результаты и их сравнительный анализ	14
4.1. Название параграфа	14
4.2. Название параграфа	14
4.3. Выводы	14
Заключение	15
Список сокращений и условных обозначений	16
Словарь терминов.....	17
Список использованных источников.....	18
Приложение 1. Краткие инструкции по настройке издательской системы L ^A T _E X.....	19
Приложение 2. Некоторые дополнительные примеры	23

ВВЕДЕНИЕ. ОБОСНОВАНИЕ АКТУАЛЬНОСТИ

В большинстве современных компьютеров установлено и одновременно используются два вида вычислительных процессоров: центральный процессор (CPU) и графический (GPU). Из-за того что они работают одновременно, а также имеют обособленную друг от друга физическую память, возникает необходимость в синхронизации и передаче данных. Эту необходимость решает программа-драйвер, которую реализуют производители графических вычислительных процессоров, предоставляя разработчикам графических приложений специальные API, такие как OpenGL, Vulkan, DirectX 11 и тд. Однако, из-за того что заранее предугадать архитектурные особенности итоговых графических приложений невозможно, программа-драйвер может добавлять дополнительные синхронизации, что приводит к ухудшению производительности, так как центральный и графический процессоры начинают работать последовательно, как представлено на рис.0.1.

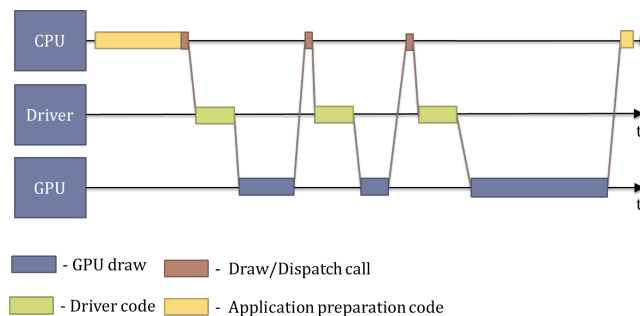


Рис.0.1. Схема работы приложения при неоптимальном драйвере DirectX 11

Во избежание подобных ситуаций, в более поздних и новых API (таких как DirectX 12 и Vulkan) задача синхронизации была снята с программы драйвера и была передана разработчику. Для этого было введено понятие “списка команд”, который разработчик мог заполнять и затем отправлять на исполнение, как показано на рис.0.2.

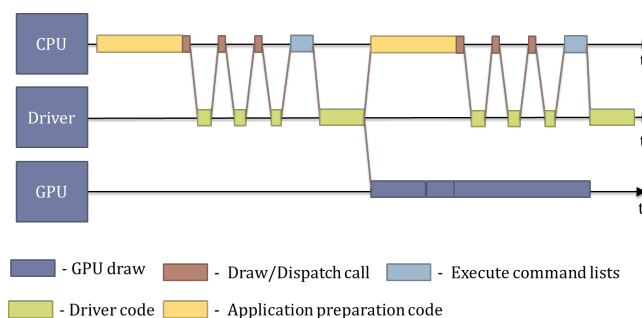


Рис.0.2. Схема работы приложения при DirectX 12

Однако несмотря на то, что данный подход даёт возможность явно управлять синхронизациями между процессорами, часть времени центрального процессора уходит на построение вышеупомянутого “списка”. Причём время построения и передачи будет зависеть от количества выводимых объектов (см. рис.0.3).

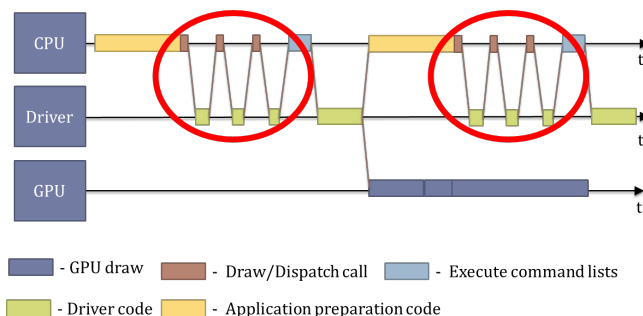


Рис.0.3. Схема работы приложения при DirectX 12, с выделенным временем на построение списка выводимых объектов

Целью работы является построение архитектуры графического приложения таким образом, чтобы время построения и передачи, а также размер “списка команд” не зависели от количества выводимых объектов (см. рис.0.4). Таким образом разработчику ПО будет предоставлено больше времени центрального процессора, а значит более трудоёмкие алгоритмы можно будет использовать для достижения желаемых результатов.

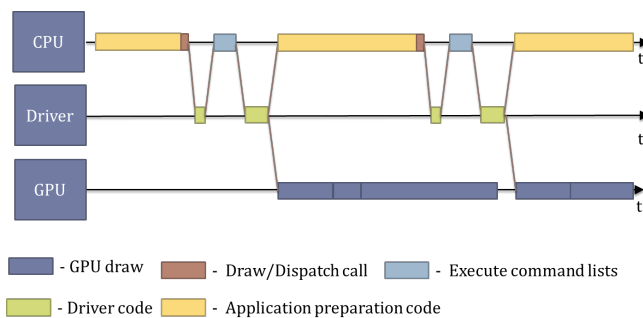


Рис.0.4. Схема работы приложения при DirectX 12 с использованием предлагаемого конвейера

ГЛАВА 1. ПОСТАНОВКА ЗАДАЧИ

1.1. Техническое задание

Требуется:

- разработать архитектуру графического конвейера использующего константное количество вызовов отрисовки относительно количества объектов сцены
- реализовать приложение, позволяющее использовать как разработанную архитектуру, так и традиционную.

1.2. Ожидаемый результат

Ожидаемым результатом работы является демонстрационный проект, позволяющий сравнить производительность традиционной и разработанной архитектуры на тестовой сцене. Тестовая сцена должна содержать большое число объектов разной геометрии, положения которых изменяются на каждом кадре. Помимо этого, ожидаемым результатом работы является отчёт об производительности каждой архитектуры, и их сравнительный анализ.

ГЛАВА 2. ОБЗОР СУЩЕСТВУЮЩИХ РЕШЕНИЙ

2.1. Использование Geometry Instancing

Когда необходимо придумать что-то новое, сначала необходимо понять каким образом работает текущее. Если посмотреть на "список команд который порождается в традиционной архитектуре для отрисовки одного объекта, то можно будет увидеть алгоритм представленный на рис.2.1

Algorithm

1. Установить формат хранения вершин
2. Установить формат соединения вершин
3. Установить программу-шейдер, в которой написан алгоритм обработки входных данных
4. Установить из какого буфера будут поступать вершины
5. Установить из какого буфера будут поступать номера вершин, для соединения
6. Установить из какого буфера брать текстуру
7. Установить преобразование объекта
8. Выполнить команду отрисовки

Рис.2.1. Примерный псевдокод алгоритма вывода одного объекта

Нетрудно заметить, что если повторить описанный выше алгоритм для большого числа различных объектов, то они отрисуются корректно, однако некоторые параметры выглядят излишними. Например, если используется только один формат хранения и соединения вершин, то нет необходимости выполнять эти(1 и 2 на рис.2.1) команды каждый раз. При рассмотрении крайнего случая, когда выводится один и тот же объект много раз, алгоритм будет выглядеть как показано на рис.2.2

Algorithm

1. Установить формат хранения вершин
2. Установить формат соединения вершин
3. Установить программу-шейдер, в которой написан алгоритм обработки входных данных
4. Установить из какого буфера будут поступать вершины
5. Установить из какого буфера будут поступать номера вершин, для соединения
6. Установить из какого буфера брать текстуру
7. **for** $\forall 1 \leq i \leq N$ **do**
8. Установить преобразование объекта i
9. Выполнить команду отрисовки

Рис.2.2. Примерный псевдокод алгоритма вывода множества одинаковых объектов

Как можно заметить, количество команд в данном алгоритме, линейно зависит от количества объектов которые рисуются(7 на рис.2.2). Возникает желание уменьшить этот список, заменив многократные вызовы команд, одним вызовом, чтобы алгоритм выглядел как показано на рис.2.3

Algorithm

1. Установить формат хранения вершин
2. Установить формат соединения вершин
3. Установить программу-шейдер, в которой написан алгоритм обработки входных данных
4. Установить из какого буфера будут поступать вершины
5. Установить из какого буфера будут поступать номера вершин, для соединения
6. Установить из какого буфера брать текстуру
7. Установить из какого буфера брать преобразования объектов
8. Выполнить команду отрисовки N раз

Рис.2.3. Примерный псевдокод алгоритма вывода множества одинаковых объектов при помощи Geometry Instancing

Именно поддержка такого алгоритма(рис.2.3) и называется Geometry Instancing, и его использование отрывает большие возможности. Например в статье Park-a[4] описывается архитектура графического конвейера, позволяющего выводить большие толпы объектов, при помощи технологии Geometry Instancing (рис.2.4). Из особенностей можно отметить:

1. Все объекты имеют несколько типов геометрии, разнящиеся по уровню детализации(рис.2.5)
2. Для каждого объекта, при помощи вычислительных шейдеров, определяется уровень детализации и находится ли вообще объект в области видимости.
3. В "списке команд"находятся команды отрисовки для каждого типа геометрии, но количество раз отрисовки (для каждого типа) указывается уже исходя из работы предыдущего вычислительного шейдера.



Рис.2.4. Примеры работы приложения описанного в статье Park-a[4]

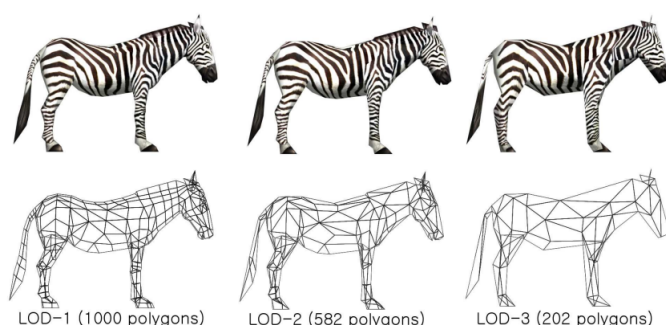


Рис.2.5. Уровни детализации, описанные в статье Park-a[4]

Основным недостатком является то, что предложенная Park-ом[4] архитектура требует добавлять в "список команд" команды отрисовки для каждого выводимого типа геометрии. В случае, если на экран выводится много разных объектов, предложенная архитектура ничем не будет отличаться от традиционной.

2.2. Использование Programmable Vertex Pulling

В своей статье Ricco[5] предлагает способ избавиться от основной проблемы Geometry Instancing-a - отрисовки объектов с одинаковой геометрией. Для этого он предлагает сделать следующее

1. Ввести понятие "пачки"треугольников - набор из N вершин(число N задаётся на центральном процессоре).
2. Все объекты, которые требуется вывести на экран, разделяются на "пачки".
3. Все вершины всех объектов складываются в один большой буфер.
4. В вычислительном шейдере, для всех объектов определяется сколько "пачек"соответствуют каждому объекту и для каждой "пачки"записываются индексы вершин из буфера со всеми вершинами.

5. В "списке команд" находится лишь одна команда отрисовки, которая рисует "пачку" столько раз, сколько было решено на вычислительном шейдере.

Такой подход уменьшает количество вызовов команд отрисовок до одной, что является впечатляющим результатом. Однако этот подход не лишён недостатков:

1. Несмотря на уменьшение количества отрисовок в "списке команд количество вызовов отрисовки на графическом процессоре увеличивается, что может негативно сказаться на производительности.
2. Число N , используемое для задания размера "пачки должно подбираться эмпирически. Если сделать его слишком маленьким, то число вызовов отрисовки на графическом процессоре возрастет (см п.1). Если сделать его слишком большим, то будут отрисовываться "лишние"треугольники из "пачки".
3. Буфер (упомянутый в пункте 3. прошлого списка) создаётся не как буфер для вершин, а как буфер общего назначения. Это приводит к тому, что к нему не применяются оптимизации программы-драйвера(Например способ кеширования или разложение в видеопамяти).
4. Из-за фрагментации, которая может происходить в видеопамяти, буфер упомянутый выше, может не поместиться в видеопамять, хотя объекты по отдельности могли бы.

ГЛАВА 3. ПРЕДЛАГАЕМОЕ РЕШЕНИЕ

Хорошим стилем является наличие введения к главе. Во введении может быть описана цель написания главы, а также приведена краткая структура главы.

3.1. Название параграфа

3.2. Название параграфа

3.3. Выводы

Текст выводов по главе 3.

ГЛАВА 4. РЕЗУЛЬТАТЫ И ИХ СРАВНИТЕЛЬНЫЙ АНАЛИЗ

Хорошим стилем является наличие введения к главе. Во введении может быть описана цель написания главы, а также приведена краткая структура главы.

4.1. Название параграфа

4.2. Название параграфа

Пример ссылки на литературу [1—3; 6].

4.3. Выводы

Текст выводов по главе 4.

ЗАКЛЮЧЕНИЕ

Заключение (2 – 5 страниц) обязательно содержит выводы по теме работы, *конкретные предложения и рекомендации* по исследуемым вопросам. Количество общих выводов должно вытекать из количества задач, сформулированных во введении выпускной квалификационной работы.

Предложения и рекомендации должны быть органически увязаны с выводами и направлены на улучшение функционирования исследуемого объекта. При разработке предложений и рекомендаций обращается внимание на их обоснованность, реальность и практическую приемлемость.

Заключение не должно содержать новой информации, положений, выводов и т. д., которые до этого не рассматривались в выпускной квалификационной работе. Рекомендуются писать заключение в виде тезисов.

Последним абзацем в заключении можно выразить благодарность всем людям, которые помогали автору в написании ВКР.

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

DOI Digital Object Identifier.

WoS Web of Science.

ВКР Выпускная квалификационная работа.

ТГ-объект Текстово-графический объект.

СЛОВАРЬ ТЕРМИНОВ

TeX — язык вёрстки текста и издательская система, разработанные Дональдом Кнутом.

LaTeX — язык вёрстки текста и издательская система, разработанные Лэсли Лампортом как надстройка над TeX.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *Автономова Н. С.* Философский язык Жака Деррида. — М.: Российская политическая энциклопедия (РОССПЭН), 2011. — 510 с. — (Сер.: Российские Пропилеи).
2. *Котельников И. А., Чеботаев П. З.* LaTeX по-русски. — 3-е изд. — Новосибирск: Сибирский Хронограф, 2004. — 496 с. — URL: <http://www.tex.uniya.ac.ru/doc/kotelnikovchebotaev2004b.pdf> (дата обращения: 06.03.2019).
3. *Песков Н. В.* Поиск информативных фрагментов описаний объектов в задачах распознавания: дис. . . . канд. канд. физ.-мат. наук: 05.13.17 / Песков Николай Владимирович. — М., 2004. — 102 с.
4. *Park H., Han J.* Fast rendering of large crowds using GPU // Entertainment Computing-ICEC 2008: 7th International Conference, Pittsburgh, PA, USA, September 25-27, 2008. Proceedings 7. — Springer. 2009. — С. 197—202.
5. *Riccio C., Lilley S.* Introducing the programmable vertex pulling rendering pipeline // GPU Pro. — 2013. — Т. 4. — С. 21—37.
6. *Kotelnikov I. A., Chebotaev P. Z.* LaTeX in Russian. — 3rd ed. — Novosibirsk: Sibiskiy Hronograph, 2004. — 496 p. — URL: <http://www.tex.uniya.ac.ru/doc/kotelnikovchebotaev2004b.pdf> (visited on 06.03.2019); (in Russian).
7. MiKTeX web site. — URL: <https://miktex.org/> (visited on 06.03.2019).
8. *Peskov N. V.* Searching for informative fragments of object descriptions in the recognition tasks: diss. . . . cand. phys.-math. sci.: 05.13.17 / Peskov Nickolay Vladimirovich. — М., 2004. — 102 p. — (in Russian).
9. SPbPU photo gallery. — URL: <http://www.spbstu.ru/media/photo-gallery/> (visited on 06.03.2019).
10. TeXstudio web site. — URL: <https://www.texstudio.org/> (visited on 06.03.2019).

Приложение 1

Краткие инструкции по настройке издательской системы L^AT_EX

В SPbPU-BCI-template автоматически выставляются необходимые настройки и в исходном тексте шаблона приведены примеры оформления текстово-графических объектов, поэтому авторам достаточно заполнить имеющийся шаблон текстом главы (статьи), не вдаваясь в детали оформления, описанные далее. Возможный «быстрый старт» оформления главы (статьи) под Windows следующий^{П1.1}:

- A. Установка полной версии MikTeX [7]. В процессе установки лучше выставить параметр доустановки пакетов «на лету».
- B. Установка TexStudio [10].
- C. Запуск TexStudio и компиляция `my_chapter.tex` с помощью команды «Build&View» (например, с помощью двойной зелёной стрелки в верхней панели). Иногда, для достижения нужного результата необходимо несколько раз скомпилировать документ.
- D. В случае, если не отобразилась библиография, можно
 - воспользоваться командой Tools → Commands → Biber, затем запустив Build&View;
 - настроить автоматическое включение библиографии в настройках Options → Configure TexStudio → Build → Build&View (оставить по умолчанию, если сборка происходит слишком долго): `txs:///pdflatex | txs:///biber | txs:///pdflatex | txs:///pdflatex | txs:///view-pdf`.

В случае возникновения ошибок, попробуйте скомпилировать документ до последних действий или внимательно ознакомьтесь с описанием проблемы в log-файле. Бывает полезным переход (по подсказке TexStudio) в нужную строку в pdf-файле или запрос с текстом ошибки в поисковиках. Наиболее вероятной проблемой при первой компиляции может быть отсутствие какого-либо установленного пакета L^AT_EX.

В случае корректной работы настройки «установка на лету» все дополнительные пакеты будут скачиваться и устанавливаться в автоматическом режиме. Если доустановка пакетов осуществляется медленно (несколько пакетов за один запуск

^{П1.1}Внимание! Пример оформления подстрочной ссылки (сноски).

компилятора), то можно попробовать установить их в ручном режиме следующим образом:

1. Запустите программу: меню → все программы → MikTeX → Maintenance (Admin) → MiKTeX Package Manager (Admin).
2. Пользуясь поиском, убедитесь, что нужный пакет присутствует, но не установлен (если пакет отсутствует воспользуйтесь сначала MiKTeX Update (Admin)).
3. Выделив строку с пакетом (возможно выбрать несколько или вообще все неустановленные пакеты), выполните установку Tools → Install или с помощью контекстного меню.
4. После завершения установки запустите программу MiKTeX Settings (Admin).
5. Обновите базу данных имен файлов Refresh FNDB.

Для проверки текста статьи на русском языке полезно также воспользоваться настройками Options → Configure TexStudio → Language Checking → Default Language. Если русский язык «ru_RU» не будет доступен в меню выбора, то необходимо вначале выполнить Import Dictionary, скачав из интернета любой русскоязычный словарь.

Далее приведены формулы (П1.2), (П1.1), рис.П1.2, рис.П1.1, табл.П1.2, табл.П1.1.

$$\pi \approx 3,141. \quad (\text{П1.1})$$



Рис.П1.1. Вид на гидробашню СПбПУ [9]

Представление данных для сквозного примера по ВКР [8]

G	m_1	m_2	m_3	m_4	K
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2

П1.1. Параграф приложения

П1.1.1. Название подпараграфа

Название подпараграфа оформляется с помощью команды `\subsection{...}`.
Использование подпараграфов в основной части крайне не рекомендуется.

П1.1.1.1. Название подподпараграфа

$$\pi \approx 3,141. \quad (\text{П1.2})$$



Рис.П1.2. Вид на гидробашню СПбПУ [9]

Таблица П1.2

Представление данных для сквозного примера по ВКР [8]

G	m_1	m_2	m_3	m_4	K
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2

Приложение 2

Некоторые дополнительные примеры

В приложении^{П2.1} приведены формулы (П2.2), (П2.1), рис.П2.2, рис.П2.1, табл.П2.2, табл.П2.1

$$\pi \approx 3,141.$$

(П2.1)



Рис.П2.1. Вид на гидробашню СПбПУ [9]

Таблица П2.1

Представление данных для сквозного примера по ВКР [8]

G	m_1	m_2	m_3	m_4	K
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2

^{П2.1}Внимание! Пример оформления подстрочной ссылки (сноски).

П2.1. Подраздел приложения

$$\pi \approx 3,141.$$

(П2.2)



Рис.П2.2. Вид на гидробашню СПбПУ [9]

Таблица П2.2

Представление данных для сквозного примера по ВКР [8]

<i>G</i>	<i>m</i> ₁	<i>m</i> ₂	<i>m</i> ₃	<i>m</i> ₄	<i>K</i>
<i>g</i> ₁	0	1	1	0	1
<i>g</i> ₂	1	2	0	1	1
<i>g</i> ₃	0	1	0	1	1
<i>g</i> ₄	1	2	1	0	2
<i>g</i> ₅	1	1	0	1	2
<i>g</i> ₆	1	1	1	2	2