

СПб ГУТ)))

Формирование и обработка звуковых сигналов

ПРАКТИЧЕСКОЕ ЗАДАНИЕ № 2

Сведение речевого сигнала и сигнала музыкального сопровождения

Выполнил:

Балан К. А.

Студент группы:

РЦТ-22

Преподаватель:

Ишугина О. Ю.

Санкт-Петербург

2024

1. Запись испытательного звукового сигнала

1. Подключим к ПК головные телефоны и микрофон. Установим в настройках драйвера звуковой карты для обоих устройств частоту дискретизации 48000 Гц, формат и разрядность квантования 16 бит/отсчет.

2. Запустим ПО Audacity. Создадим новый проект и в нем создадим новую монофоническую дорожку. Установим дорожке частоту дискретизации 48000 Гц, формат и разрядность квантования 16 бит/отсчет. Назовем дорожку “original speech”.

3. Запишем отрывок речевого сигнала длительностью от 3 до 5 минут.

4. Сохраним проект.

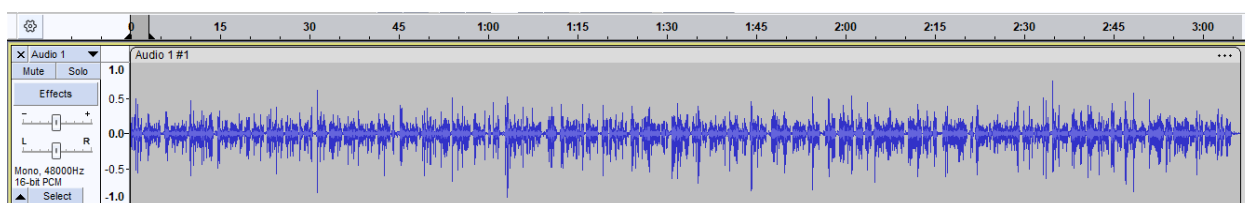


Рисунок 1 – Полученный испытательный звуковой сигнал.

2. Удаление нежелательных фрагментов

1. Продублируем звуковую дорожку “original speech” и назовем ее “edited speech”.
2. Прослушаем скопированный сигнал и удалим нежелательные фрагменты.
3. Зафиксируем осциллограмму полученного речевого сигнала в отчет.
4. Сохраним проект.

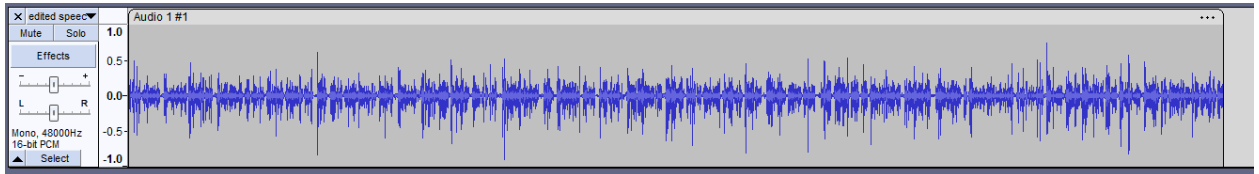


Рисунок 2 – Отредактированный испытательный звуковой сигнал.

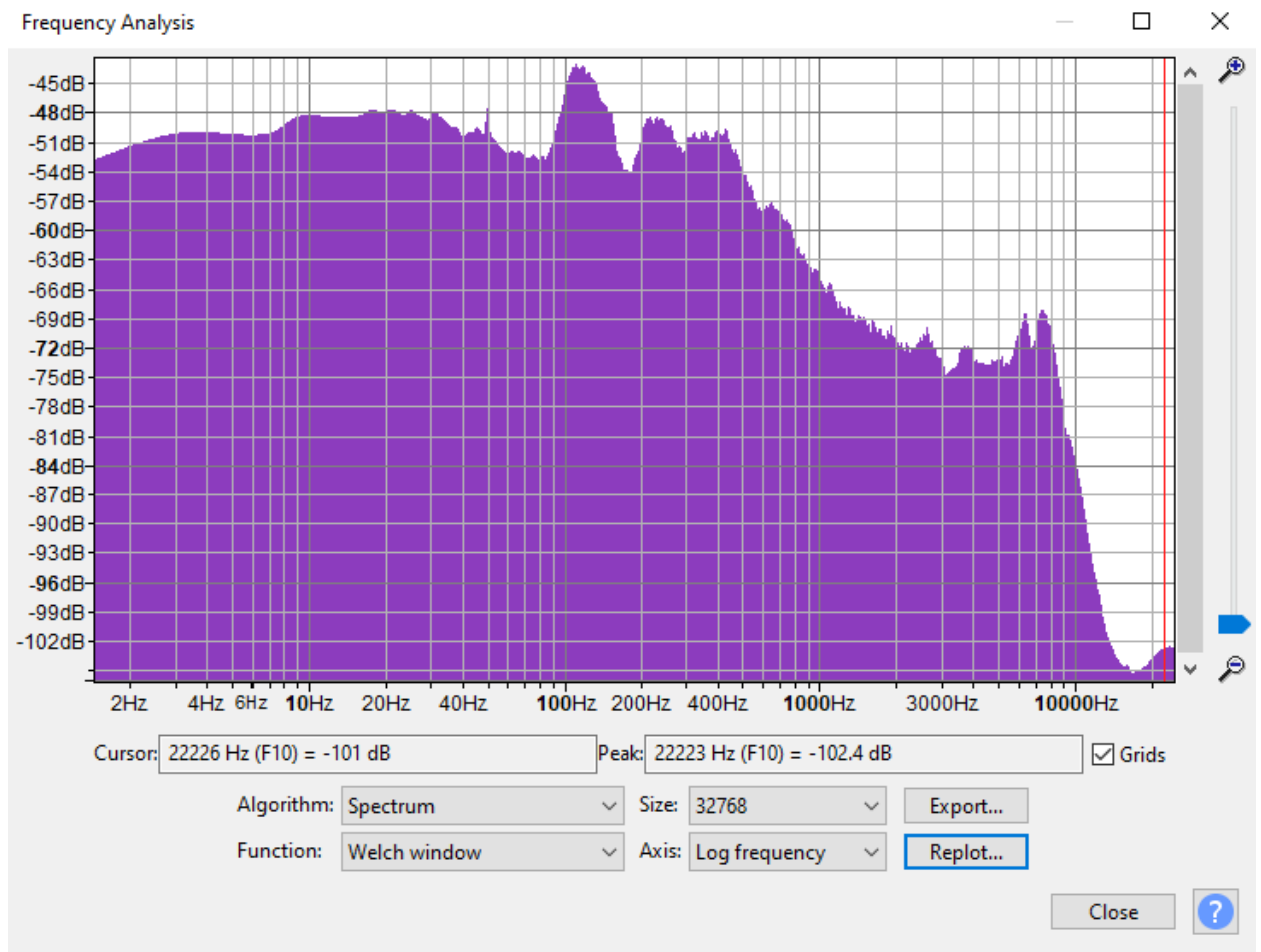


Рисунок 3 – Осциллограмма отредактированного испытательного звукового сигнала.

3. Выравнивание громкости речевого сигнала

1. Продублируем звуковую дорожку “edited speech” и назовем ее “compressed speech”.
2. Для сигнала “compressed speech” включим опцию Make-up gain for 0 dB after compressing и отключим опцию Compressed based on peaks.
3. Зафиксируем сигналограмму и спектрограмму полученного сигнала.
4. Заполним таблицу 1.
5. Сохраним проект.

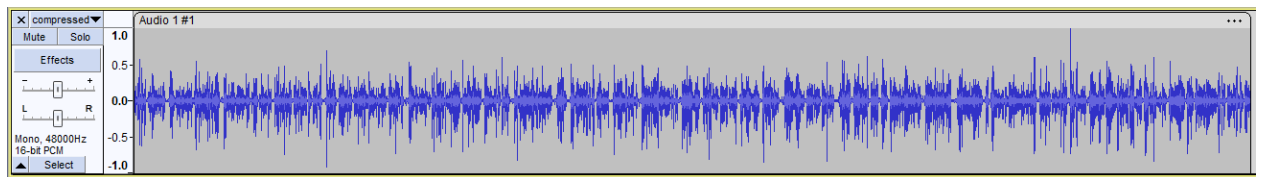


Рисунок 4 – Сигналограмма звуковой дорожки “compressed speech”.

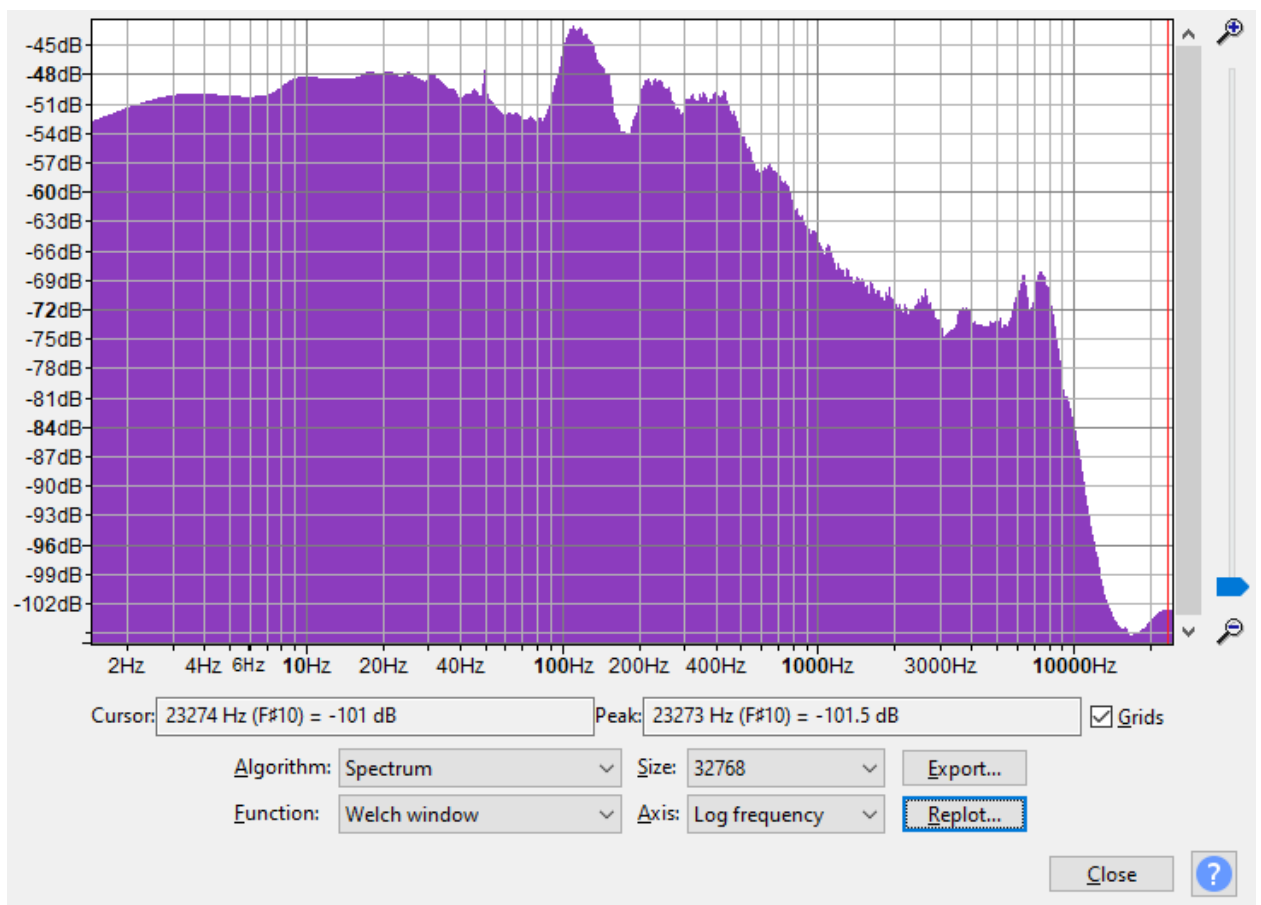


Рисунок 5 – Спектрограмма звуковой дорожки “compressed speech”.

Таблица 1 – Параметры инструмента Compressor.

Наименование параметра	Значение	Единицы измерения
Threshold	-12	dBFS
Noise Floor	-40	dBFS
Ratio	2:1	-
Attack time	0.2	сек
Release time	1	сек

4. Импорт музыкального сопровождения

1. Импортируем аудиофайл музыкального сопровождения (“Clean out of my closest (Instrumental)” Eminem) в проект.
2. Назовем новую звуковую дорожку “music”.
3. Зафиксируем осциллограмму и спектрограмму сигнала музыкального сопровождения в отчет.
4. Сохраним проект.

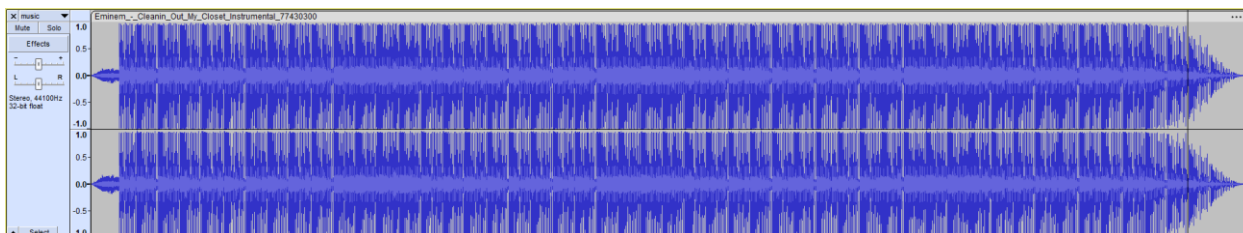


Рисунок 6 – Осциллограмма звуковой дорожки “music”.

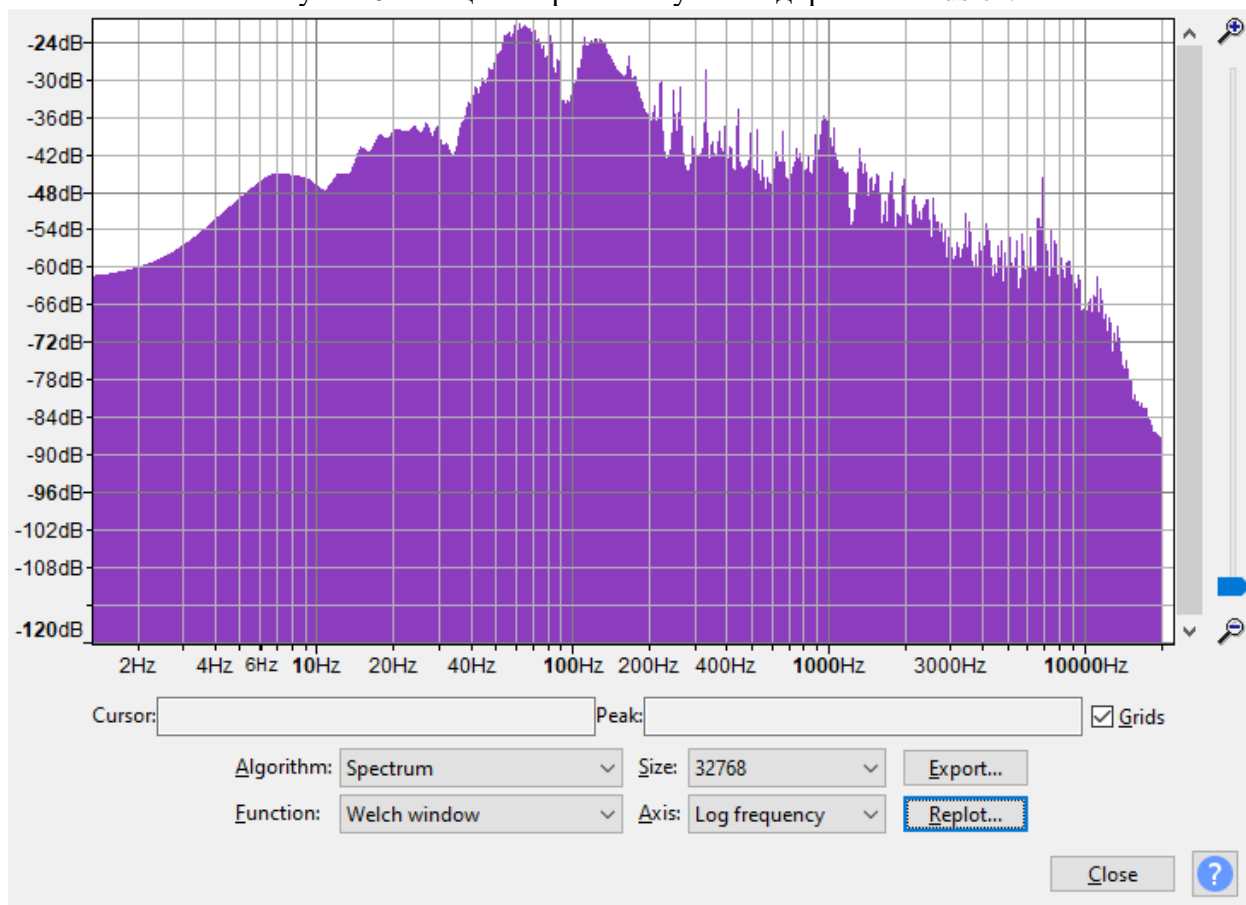


Рисунок 7 – Спектрограмма звуковой дорожки “music”.

5. Разделение файла речевого сигнала на отрывки

1. Сформируем дорожку меток. Дадим ей имя “speech Label” и разместим ее под дорожкой “compressed speech”.
2. Установим метки окончания для отрывков сигнала “compressed speech”.
3. Разделим речевой сигнал “compressed speech” на отдельные отрывки в соответствии со сформированной дорожкой меток.
4. Сохраним проект.

6. Синхронизация отрывков речевого сигнала и музыкального сопровождения

1. После разделения речевого сигнала на отрывки, определим моменты, в которые будут начинаться те или иные отрывки речевого сигнала.
2. Сформируем дорожку меток “timeline”, разместим ее под дорожкой сигнала музыкального сопровождения и установим метки в моменты начала и окончания каждого из отрывков речевого сигнала.
3. Используя инструмент “Time Shift Toolbar”, совместим отрывки сигнала на дорожке “compressed speech” с соответствующими метками начала их звучания на дорожке “timeline”.
4. Сохраним проект.

7. Ослабление музыкального сопровождения в присутствии речевого сигнала

1. Выделим дорожку “music” и создадим две копии. Назовем их “manual” и “automatic” соответственно.

2. Выполним ослабление уровня сигнала музыкального сигнала с помощью инструмента “Envelope Tool” для дорожки “manual” и с помощью инструмента “Auto Duck” для дорожки “automatic”.

3. Отключим все дорожки, кроме “compressed speech” и “manual”. Прослушаем полученный результат и убедимся, что звучание музыкального сигнала не приводит к снижению разборчивости речевого сигнала.

4. Отключим дорожку “manual” и включим дорожку “automatic”. Прослушаем полученный результат и убедимся, что звучание музыкального сигнала не приводит к снижению разборчивости речевого сигнала.

5. Заполним таблицу 2.

6. Сохраним проект.

Таблица 2 – Параметры инструмента Auto Duck.

Наименование параметра	Значение	Единицы измерения
Duck Amount	-17	дБ
Outer fade down length	0.5	сек
Inner fade down length	0	сек
Maximum pause	1	сек
Outer fade up length	0.	сек
Inner fade up length	0	сек
Threshold	-30	дБ

8. Сведение речевого сигнала и сигнала музыкального сопровождения

1. Сведем звуковые дорожки “compressed speech” и “manual”.
 2. Назовем полученную звуковую дорожку “manual mix”.
 3. Сведем звуковые дорожки “compressed speech” и “automatic”.
 4. Назовем полученную звуковую дорожку “automatic mix”.
 5. Прослушаем получившиеся в результате сведения композиции.
 6. Зафиксируем осциллограммы и спектрограммы полученных сигналов в отчет.
7. Сохраним проект.

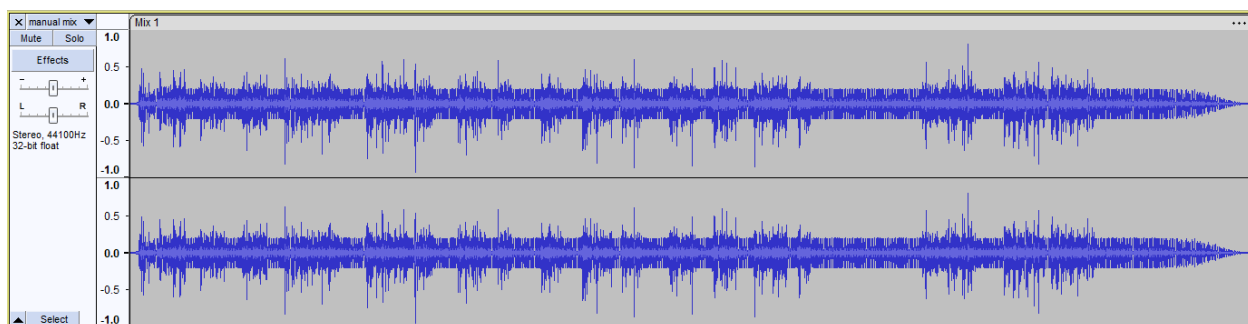


Рисунок 8 – Сигналограмма звуковой дорожки “manual mix”.

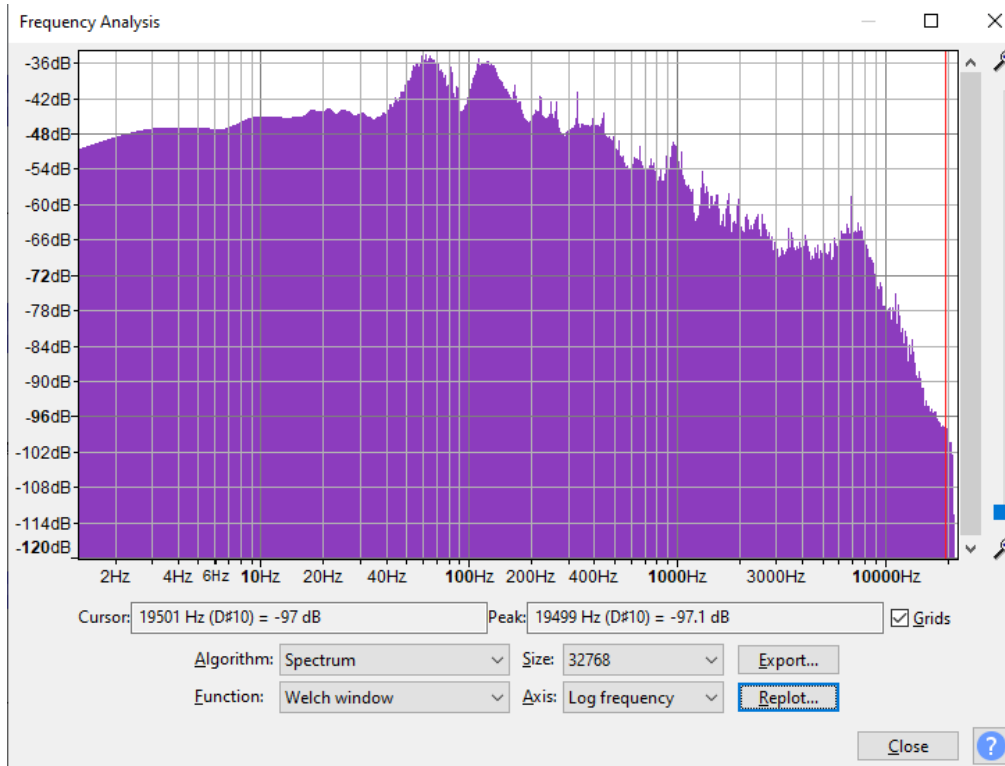


Рисунок 9 – Осциллограмма звуковой дорожки “manual mix”.



Рисунок 10 – Сигналограмма звуковой дорожки “automatic mix”.

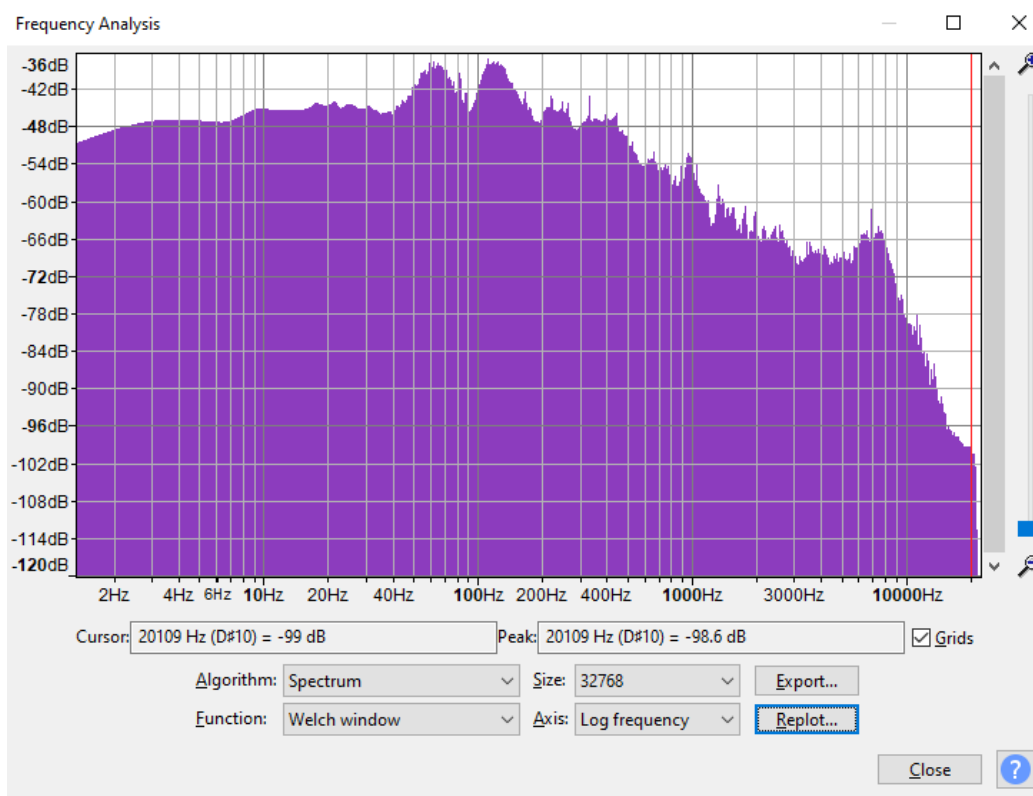


Рисунок 11 – Осциллограмма звуковой дорожки “ automatic mix”.

9. Анализ результатов выполнения задания

1. Какие нежелательные фрагменты присутствовали в записанном речевом сигнале? Что послужило причиной их возникновения? Как нужно изменить условия записи речевого сигнала, чтобы избежать их появления?

В сигнале присутствовали растяжения некоторых слогов, паузы в середине предложения и повторения некоторых слогов. Причиной их возникновения послужила плохая дикция и отсутствие постоянно практики прочтения вслух.

2. На основании каких критериев был произведен выбор оптимальных значений параметров инструмента Compressor?

Выбор оптимальных значений параметров инструмента Compressor был основан на субъективном восприятии звуковой дорожки.

3. Может ли сжатие динамического диапазона привести к повышению качества речевого сигнала? Как этого добиться?

Сжатие динамического диапазона помогает сделать тихие звуки громче, а громкие – тише, что делает звук более сбалансированным и легче воспринимаемым. Это особенно полезно в шумной обстановке.

Если нежелательные пики сигнала сжимаются, это помогает избежать искажений, которые могут возникать, когда уровень сигнала превышает максимально допустимый.

4. Может ли сжатие динамического диапазона привести к потере качества речевого сигнала? Как этого избежать?

Сильно сжатые сигналы могут потерять нюансы, что делает речь менее естественной. Также, при сильном сжатии могут возникнуть различные артефакты, такие как искажения и помехи. Сжатие может сделать речь плоской и менее выразительной.

Для избежания данных проблем необходимо использовать мягкие настройки сжатия, чтобы сохранить естественность звука.

5. Как сжатие динамического диапазона повлияло на громкость, разборчивость и тембральную окраску записанного речевого сигнала?

Сигнал стал несколько тише, остальные характеристики остались без ощутимых изменений.

6. На сколько уровень музыкального сигнала должен быть меньше уровня речевого сигнала для того, чтобы разборчивость речевого сигнала не ухудшилась в присутствии музыки?

Рекомендуется, чтобы уровень музыкального сигнала был ниже уровня речевого сигнала на 10–15 дБ.

7. Как влияют жанр и спектральный состав музыкального сопровождения на требуемую величину его ослабления в присутствии речевого сигнала?

1. Жанр музыки:

Инструментальная музыка как правило, требует меньшего ослабления, особенно если она содержит мало вокала и использует инструменты, которые не пересекаются по частотам с речью.

Вокальная музыка может потребовать большего ослабления, так как вокал в музыке будет конкурировать с речевым сигналом за внимание слушателя.

Экспериментальные жанры, такие как электронная музыка или авангард могут включать неожиданные частотные диапазоны, что требует индивидуального подхода к ослаблению.

2. Спектральный состав:

Музыка, содержащая много низких частот (басовые линии), может заглушать речь, что требует значительного ослабления. В то время как музыка с преобладанием высоких частот может быть менее мешающей.

Если в музыкальном сопровождении есть частоты, близкие к частотам речи (300–3000 Гц), это может значительно затруднить разборчивость. В таком случае может потребоваться дополнительное ослабление или использование эквалайзеров для фильтрации.

8. Какой музыкальный жанр, по вашему мнению, является наилучшим для использования в качестве фона для речевого сигнала?

Классическая музыка или джаз.