

Вологдин Э.И.

**КОМПЬЮТЕРНЫЙ ПРАКТИКУМ ПО
ЦИФРОВОЙ АУДИОТЕХНИКЕ**

**Санкт Петербург
2010**

СОДЕРЖАНИЕ

ЛАБОРАТОРНАЯ РАБОТА N 1 АТ	3
ИЗМЕРЕНИЕ АБСОЛЮТНОГО ПОРОГА СЛЫШИМОСТИ	3
ЛАБОРАТОРНАЯ РАБОТА N 2 АТ	15
ИССЛЕДОВАНИЕ МАСКИРОВКИ.....	15
ТОНАЛЬНЫХ ЗВУКОВ БЕЛЫМ ШУМОМ	15
ЛАБОРАТОРНАЯ РАБОТА N 3 АТ	26
ИССЛЕДОВАНИЕ ДИНАМИЧЕСКИХ ХАРАКТЕРИСТИК.....	26
ЦИФРОВЫХ ФОНОГРАММ	26
ЛАБОРАТОРНАЯ РАБОТА N 4 АТ	36
ИССЛЕДОВАНИЕ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК	36
ЦИФРОВЫХ ФОНОГРАММ	36
ЛАБОРАТОРНАЯ РАБОТА N 5 АТ	47
ИССЛЕДОВАНИЕ ДИНАМИЧЕСКОЙ ОБРАБОТКИ	47
ЦИФРОВЫХ ФОНОГРАММ	47
ЛАБОРАТОРНАЯ РАБОТА N 6 АТ	64
ИССЛЕДОВАНИЕ ЧАСТОТНЫХ ХАРАКТЕРИСТИК	64
ШЕЛЬФОВЫХ ФИЛЬТРОВ.....	64
ЛАБОРАТОРНАЯ РАБОТА N 7 АТ	74
ИССЛЕДОВАНИЕ ЧАСТОТНЫХ ХАРАКТЕРИСТИК	74
КОДЕКОВ ФОРМАТОВ MPEG-1 Layer 2 и Layer 3.....	74
ЛАБОРАТОРНАЯ РАБОТА N 8 АТ	85
ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИК КОДЕРА MP3 Lame.....	85
В РЕЖИМАХ РАБОТЫ CBR, ABR И VBR.....	85
Программные пакеты Sound Forge 9c и Audition 3.....	100

ЛАБОРАТОРНАЯ РАБОТА N 1 АТ

ИЗМЕРЕНИЕ АБСОЛЮТНОГО ПОРОГА СЛЫШИМОСТИ

Руководство к выполнению лабораторной работы по Аудиотехнике

1. Цель работы

Целью лабораторной работы является ознакомление студентов с методикой измерения аудиограмм абсолютного порога слышимости с использованием компьютерной программы SpectraLab.

2. Задание к лабораторной работе

- Ознакомиться с программой SpectraLab в объеме, необходимом для измерения порога слышимости по звуковому давлению в шкале Sound Pressure Level (SPL)
- Измерить частотную зависимость уровня спектральной плотности мощности и интегральное значение суммарного уровня акустических и электрических шумов измерительного тракта
- Измерить аудиограммы абсолютного порога слышимости в диапазоне 30...16000 Гц с использованием головных стерео телефонов
- Измерить порог слышимости белого шума
- По результатам измерений сформулировать выводы

2. Функциональная схема измерений

В приведенной на рис.1.1 схеме измеряется частотная зависимость порога слышимости тональных звуков по звуковому давлению в шкале SPL. Тестовые сигналы в частотном диапазоне 20... 20000 Гц формируются генератором, их уровни устанавливаются дискретным регулятором Level генератора и двумя плавными регуляторами Play Control и Wave/Direct (звук). В качестве источников звуковых колебаний используются стерео телефоны. Уровень порогов слышимости измеряются по шкале спектроанализатора и по индикатору Power Level.

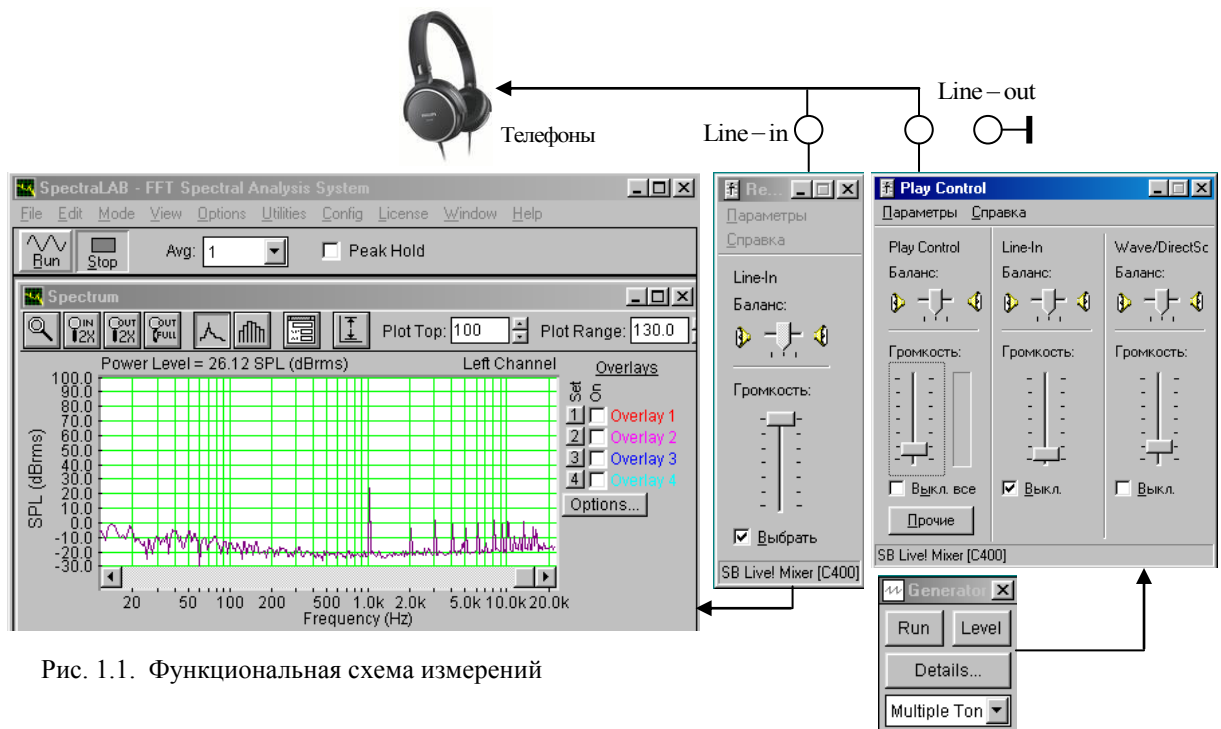


Рис. 1.1. Функциональная схема измерений

4. Методические указания

4.1. Характеристики звуковых колебаний. Основными характеристиками звука являются давление P и интенсивность I . Давление характеризует амплитуду звуковых колебаний, а интенсивность – их мощность. Звуковое давление измеряется в Паскалях (Па), а интенсивность – в Вт/м²,

$$1 \text{ Па} = \text{Ньютон} / \text{м}^2. \quad (1-1)$$

Звуковое давление и интенсивность связаны квадратичной зависимостью

$$I = \frac{P^2}{Z_s}, \quad (1-2)$$

где $Z_s = Q_s \cdot C_s$ – удельное акустическое сопротивление воздушной среды,

$Q_s = 1,23 \text{ кг/м}^3$ – плотность воздушной среды, $C_s = 340 \text{ м/с}$ – скорость распространения звука.

При измерениях напряжение на выходе микрофона пропорционально звуковому давлению, а интенсивность, если необходимо, рассчитывается по формуле (1). Слух воспринимает давления от 10^{-5} до 200 Па , что соответствует изменению интенсивности звука от 10^{-12} до 100 Вт/м^2 .

В аудиотехнике звуковое давление и интенсивность звуковых колебаний обычно измеряется в децибелах в шкале SPL (Sound Pressure Level), В этой шкале за 0 дБ принята интенсивность звука

$$I_0 = 10^{-12} \text{ Вт/м}^2. \quad (1-3)$$

Этой интенсивности соответствует звуковое давление $P_0 = 2 \cdot 10^{-5} \text{ Па}$.

Уровни звукового давления и интенсивности тональных звуков рассчитываются по формулам:

$$L_p = 20 \cdot \log \left(\frac{P}{P_0} \right), \quad (1-4)$$

$$L_{\text{int}} = 10 \cdot \log \left(\frac{I}{I_0} \right) = 20 \cdot \log \left(\frac{P}{P_0} \right). \quad (1-5)$$

Из этих формул следует, что уровень звукового давления всегда равен уровню интенсивности звука, и поэтому можно использовать просто термин «уровень». Интенсивность звука через уровень звукового давления можно определить по формуле

$$I = 10^{\frac{L_p}{10}} \cdot (1-6)$$

4.2. Порог слышимости – это минимальное звуковое давление, при котором еще существует слуховое ощущение звука. Наличие порога слышимости связано со строением слуховой системы человека. Порог соответствует звуковому давлению, при котором базилярная мембрана улитки лишь слегка касается нервных волосковых клеток. Он характеризует чувствительность уха к интенсивности звуковой энергии. Величина порога слышимости зависит от характера тестовых звуков и условий измерений.

Знание величины слуховых порогов на различных частотах крайне важно для современных систем компрессирования цифровых звуковых потоков, так как звуковые сигналы ниже порога слышимости в цифровой тракт могут не передаваться. Это позволяет сократить скорость передачи данных по каналам связи и увеличить время записи на дисковых носителях.

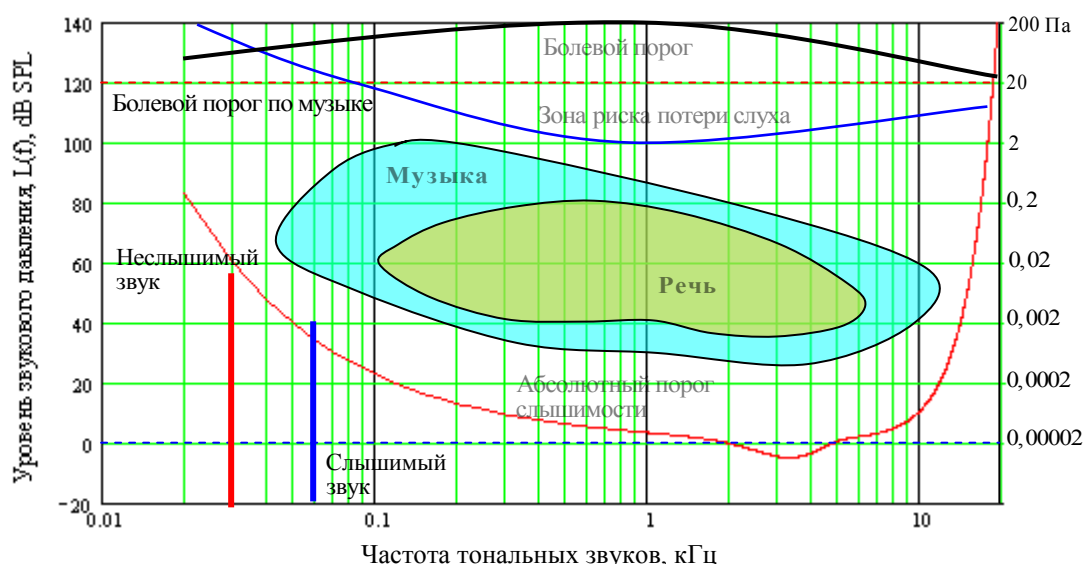


Рис.1.2. Частотная зависимость абсолютного порога слышимости

4.3. Абсолютный порог слышимости измеряется на тональном звуке в полной тишине. Определяется он для людей в возрасте 18...20 лет при размещении моноурального источника звука (один громкоговоритель) перед слушателем. Длительность воздействия звука должна быть не меньше 250 мс.

График частотной зависимости уровня абсолютного порога слышимости $L(F_t)$ в логарифмической шкале частот приведен на рис.1.2. На этом графике по оси абсцисс отложена частота тональных звуков F_t , а по осям ординат - звуковое давление P в Паскалях и уровень звукового давления $L(F_t)$ в децибелах. Частотная зависимость порога слышимости конкретного человека называется **аудиограммой**. Для некоторых применений кривая порога слышимости снимается с помощью треть октавных полос шума, которые формируется 30-ти полосным эквалайзером из широкополосного белого шума.

Чувствительность слуха максимальна на частотах 2...5,0 кГц. Это повышение чувствительности связано со строением слуховой системы, в которой в слуховом проходе и в среднем ухе возникают резонансы в этой области частот. На этих частотах

абсолютный порог слышимости меньше $2 \cdot 10^{-5}$ Па. Усредненные пороги слышимости по международному стандарту ISO/R-226 приведены в табл.1.1

Табл.1.1. Абсолютные пороги слышимости (основная треть октавная шкала)					
Частота, Гц	Порог, дБ	Частота, Гц	Порог, дБ	Частота, Гц	Порог, дБ
20	77	1000	4	6300	5
31	55	2000	1	8000	12
63	30	2500	-3	10000	16
125	17	3150	-3,6	12500	4
250	15	4000	-4	16000	30
500	9	5000	-1,1	20000	90

Аналитически график зависимости абсолютного порога слышимости от частоты описывается равенством

$$L(f)_{\text{пор}} = 3,64 \cdot F_t^{-0,8} - 6,5 \cdot e^{-0,6(F_t-3,3)^2} + 10^{-3} \cdot F_t^4 \quad (1-7)$$

где F_t – частота звукового колебания (кГц).

Следует заметить, что звуковое давление, возникающее вследствие броуновского движения молекул при температуре 25 градусов, составляет $5 \cdot 10^{-6}$ Па. Если бы ухо было вдвое чувствительней, оно слышало бы непрерывный шум флуктуаций молекул воздуха и тока крови. Таким образом, чувствительность слуха человека находится на пределе биологической целесообразности.

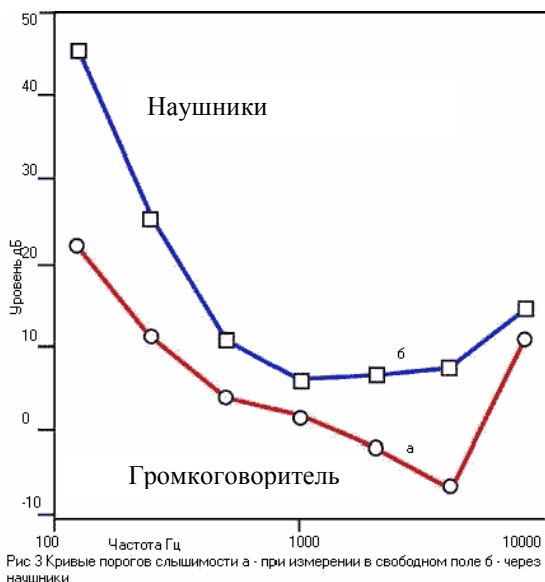


Рис.1.3. Аудиограммы

Пороги слышимости разные для левого и правого уха, поэтому результаты измерений зависят от того, что используется при измерениях: громкоговоритель или телефоны. При использовании телефонов абсолютный порог слышимости выше на 5...10 дБ (рис.1.3). Пороги слышимости очень сильно меняются в зависимости от азимутального положения источника звука и от длительности звуков. Необходимое время звучания 250 мс обусловлено постоянной времени слуха примерно 50...60 мс. При длительности звуков менее 250 мс пороги слышимости очень быстро возрастают на всех частотах. Например, сокращение длительности звука с 200 до 20 мс приводит к увеличению порога слышимости на

10 дБ. На результаты измерения порога слышимости сильное влияние оказывают акустический шум помещения прослушивания и шумы измерительных приборов, которые маскируют тональный сигнал.

С возрастом слух людей притупляется и быстрее всего на высоких частотах. Обычно на частоте 10 кГц чувствительность уха у 60-летнего человека на 20 дБ ниже, чем у 20-летнего. Однако, основная энергия музыкальных звуков находится в частотной области от 40 до 5000 Гц, поэтому возрастное изменение частотных порогов приводит к некоторому уменьшению яркости звучания обертонов, но не мешает слушать музыку и речь.

При звуковом давлении от 60 до 80 Па возникает ощущение давления на уши – эта величина называется *порогом осязания*. При уровне давления около 100 дБ возникают неприятные ощущения. Давление от 120 до 140 Па причиняет боль и называется *болевым порогом*. При звуковом давлении выше 165 Па разрывается барабанная перепонка

Табл.1. 2. Звуковые давления		
Звуки	Макс. уровень, дБ	Макс. давление, Па
Шум ракеты носителя	200	100000
Громовые раскаты	120	20
Шум самолета	110...140	
Рок музыка	110	
Шум поезда	100	2
Шумная улица	80	0,2
Разговор в комнате	40...60	0,002...0,02
Шепот	20...30	0,0002

Наибольшее влияние на сдвиг порога чувствительности оказывают высокие частоты. Понижение слуха по этим причинам лечению не поддается. Исследования показали, что в последние годы резко снизился порог слышимости у молодых людей, что возможно связано с образом их музыкальной жизни. Слуховая система человека приспособлена к звукам малой и средней интенсивности с уровнем давления не выше 90 дБ (табл.1.2.). Звуки с уровнем давлением уже больше 75 дБ приводят к изменению порогов слышимости и даже к полной глухоте. Степень повреждения пропорциональна времени воздействия. Иногда порог чувствительности восстанавливается через 16...20 часов.

Области слышимости звуков. Кривые абсолютного порога слышимости и порога болевого ощущения на рис.1.2 ограничивают области слышимости музыки и речи.

Диапазон давлений : речь от 40 до 84 дБ,
 музыка от 35 до 100 дБ.

Полоса частот: речь от 100 до 7000 Гц,
 музыка от 31 до 15000 Гц.

В этой области чувствительность уха на граничных частотах понижается на 60...80 дБ. Поэтому очевидно, что совершенно бесполезно при цифровом кодировании передавать составляющие спектра звука ниже абсолютного порога слышимости. Так как этот порог существенно зависит от частоты, то целесообразно кодирование производить в частотной области с помощью субполосной фильтрации.

Табл.1.3. Частотный и динамический диапазоны музыки		
Источник звука	Частотный диапазон, Гц	Динамический диапазон, дБ
Музыка – основной тон	100...3400	
Музыка - обертона	До 10000...20000	
Большой симфонический оркестр	31...15000	60...70
Малый симфонический оркестр	31... 15000	40...50
Духовой оркестр	50...10000	30...40
Эстрадный оркестр	25...16000	
Голос (основной тон)	70...450	35
Сопрано	262...4318	
Контральто	175...784	
Тенор	321...587	
Баритон	110...440	
Бас	82...398	

Диапазон слышимых звуков составляет 120...130 дБ, однако, уровень звукового давления на музыкальной программе не превышает 100 дБ, а для речи – 84 дБ. Причем этот диапазон существенно зависит от частоты (табл.1.3). Это необходимо учитывать при выборе числа разрядов для различных программ и в зависимости от частотной полосы кодирования. Целесообразно также при кодировании иметь переключение речь/музыка, что позволит при передаче речевых сигналов существенно сократить скорости цифровых потоков.

В цифровой аудиотехнике порог слышимости необходимо учитывать при выборе числа разрядов аналого-цифрового преобразователя (АЦП), которое определяет максимально возможный динамический диапазон сигналов DR и отношение сигнал/шум SNR . Если при осуществлении АЦП используются импульсно-кодовая модуляция и технология *dithering*, то в этом случае шум квантования является белым, а динамический диапазон сигналов определяется широко известным равенством

$$DR = SNR = 6,02 \cdot q - 3,01 \text{ dB}, \quad q > 1 \quad (1-8)$$

где q – число разрядов АЦП.

Измерение SNR производится в электрическом тракте в шкале FS (Full Scale), в которой 0 дБ соответствует максимально возможное значение сигнала при выбранном

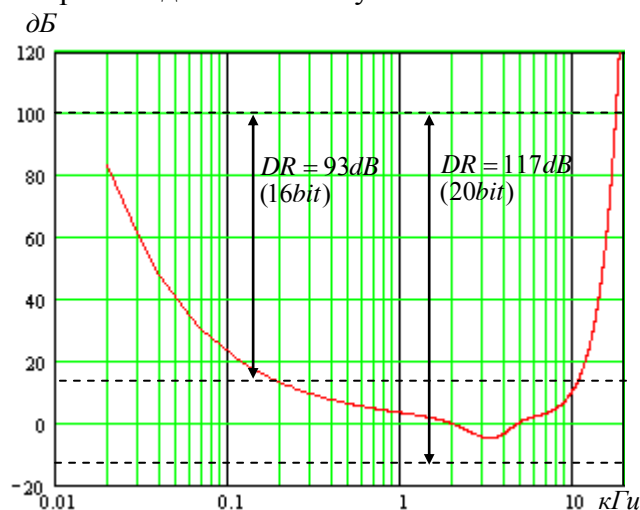


Рис.1.4. Влияние порога слышимости на динамический

числе двоичных разрядов. При воспроизведении цифровой звукозаписи максимальному значению электрического сигнала в шкале FS соответствует максимальный уровень звукового давления L_{\max} в шкале SPL. При достаточной мощности громкоговорителя, это уровень может легко меняться с помощью усилителя мощности теоретически без изменения отношения сигнал/шум.

В закрытых помещениях для комфортного восприятия музыки звуковое давление не должно превышать 100...104 dB SPL. При этом реализуемый динамический диапазон звуковых давлений из-за порога

слышимости на низких и высоких частотах существенно меньше (рис.1.4). Так при $q=16$ динамический диапазон 93 дБ из-за порога слышимости реализуется в частотном диапазоне от 100 Гц до 10 кГц. При увеличении числа разрядов, например до 20, расчетное значение динамического диапазона увеличивается до 117 дБ, но, как видно из рис.1.4., из-за порога слышимости по звуковому давлению он не может превышать 105 дБ.

Таким образом, чтобы уменьшить влияние абсолютного порога слышимости на динамический диапазон сигналов необходимо увеличивать максимальный уровень звукового давления до 120 дБ, как это имеет место на рок концертах, проводимых на открытом пространстве, и все равно при этом возникает угроза нарушения слуха у некоторых людей.

Важный практический интерес представляет интегральный порог слышимости шума L_{wn} с равномерной спектральной плотностью в звуковом диапазоне частот, равный

$$\bar{L}_{wn} = +12 \text{ dB SPL} \quad (1-9)$$

Таким шумом являются ошибки квантования в цифровых звуковых трактах. Этот порог слышимости шума позволяет рассчитать число разрядов и максимальный уровень звукового давления, при которых будут слышны шумы квантования. Основой этого расчета является формула (1-2) для отношения сигнал/шум цифрового тракта. Из этого равенства определяется уровень шума квантования по отношению к амплитуде звукового сигнала в электрическом тракте в шкале FS

$$\bar{L}_{nq}(FS) = -\frac{1}{SNR}, dB FS \quad (1-10)$$

Уровень шума квантования по звуковому давлению зависит от L_{\max} и определяется равенством

$$\bar{L}_{nq}(SPL) = L_{\max} + \bar{L}_{nq}(FS), dB SPL \quad (1-11)$$







Очевидно, что этот уровень будет ниже порога слышимости только при выполнении условия

$$\bar{L}_{nq}(SPL) < \bar{L}_{nw} \quad (1-12)$$

В музыкальных программах обычно $L_{\max} = 80...120 dB SPL$, поэтому при $q = 16$ бит условие (1-12) выполняется только при $L_{\max} > 105 dB SPL$. Если $q = 20$ бит, то условие (1-12) выполняется при $L_{\max} > 130 dB SPL$. Это значит, что увеличение числа разрядов свыше 16 не имеет практической пользы.

5. Порядок выполнения лабораторной работы

5.1. Функции кнопок управления:

- кнопкой с пиктограммой  осуществляется автоматическая установка пределов измерения по вертикальной оси шкалы анализатора спектра, фазометра или осциллографа
- кнопкой с пиктограммой  устанавливается частотный диапазон шкалы от 1 Гц до частоты Найквиста
- кнопкой с пиктограммой  вызывается маркер, которым устанавливается желаемый частотный диапазон шкалы
- при нажатии кнопки с пиктограммой  графики рисуются тонкими линиями
- при нажатии кнопки с пиктограммой  графики рисуются вертикальными линиями или столбиками
- кнопкой с пиктограммой  открывается меню анализатора спектра, фазометра или осциллографа, в котором вводятся необходимые опции.
- кнопкой **Run** включаются все измерительные приборы
- кнопкой **Stop** выключаются все измерительные приборы и запоминаются результаты измерений
- в меню **Avg** задается число усредняемых результатов измерения
- в окошке **Peak Hold** курсором включается функция удержания пиковых значений напряжения

5.2. Подготовка к измерениям


Конфигурация схемы измерений и ввод необходимых опций производится с помощью загрузочных файлов **LR- 1.1 AT.cfg- LR- 1.3 AT.cfg**. На экран монитора вызываются окна: программа SpectraLab, анализатор спектра, регулятор уровня выходного сигнала синтезатора тестовых сигналов и генератор сигналов. Измерения производятся в режиме моно, в реальном времени в частотном диапазоне 20...20000 Гц с использованием тестовых *тональных сигналов и белого шума* в логарифмической шкале частот.


В лабораторной работе измеряются на дискретных частотах измеряются пороги слышимости тональных звуков с использованием головных телефонов.

- Скопируйте с рабочего стола файл «Отчет LR-1 АТ» на свою флешкарту, дайте ему свое имя в виде: «Группа»_ «Фамилия (работавшего на клавиатуре)»_ «Номер ЛР»_ «Дата».
- Откройте этот файл и сверните его окно.


В файл этого отчета вы будете копировать все экспериментальные графики, и заносить в таблицы отчета результаты измерений сразу в процессе их получения. В день выполнения лабораторной работы отчет в электронном виде необходимо сдать преподавателю, только тогда работа считается выполненной.



- Кнопкой с пиктограммой  откройте окно программы **SpectraLab**. Не меняйте размеры этого окна, они согласованы с Вашим отчетом
- Сделайте обязательные установки в микшере **Play Control**

Для этого двойным *резким* нажатием кнопки с пиктограммой  откройте его окно. Затем после команд **Параметры** \Rightarrow **Свойства** в появившемся окне **Свойства** выберите режим «*воспроизведение*». В выпадающем меню этого окна курсором отметьте регуляторы **Play Control**, **Wave** и **Line in**, по команде **ОК** окна этих регуляторов появляются на экране монитора. Курсором включите первые два регулятора и поставьте их ползунки в верхнее положение шкалы. Регулятор **Line in** *обязательно выключите*.

- Сделайте обязательные установки в микшере **Record Control**

Для этого двойным *резким* нажатием кнопки с пиктограммой  откройте окно микшера **Play Control**. Затем после команд **Параметры** \Rightarrow **Свойства** в появившемся окне **Свойства** выберите режим «*запись*». В выпадающем меню этого окна курсором отметьте регуляторы **Microphone** и **Line in**. По команде **ОК** окна этих регуляторов появляются на экране монитора. Курсором поставьте их ползунки в верхнее положение шкалы.

- Курсором сначала включите регулятор **Microphone**, затем переключите в режим **Line in**

5.3. Измерение спектра и уровня шума измерительного тракта

В этом разделе лабораторной работы измеряется уровень шума измерительного тракта, из-за которого могут возникать значительные ошибки измерения порогов слышимости на различных частотах. Акустический шум при измерениях порождается наушниками, которые выполняют функцию микрофонов. Этот шум суммируется с шумами электронного тракта и маскирует тестовый тональный сигнал при измерении порога слышимости. Шум характеризуется частотной зависимостью его спектральной плотности мощности и интегральным значением мощности, уровень которого индицируется на шкале PowerLevel.

- Загрузите файл **LR-1.1.AT.cfg**, который конфигурирует схему измерений и устанавливает все необходимые опции для выполнения настоящего раздела лабораторной работы.

Для этого в меню **Config** выберите опцию **Load Configuration** и в выпадающем меню выберите файл **LR-1.1. AT. cfg**.

Следующие первые три пункта задания, набранные курсивом, выполняются автоматически при загрузке файла **LR-1.1. AT.cfg**, нужно только проверить, что эти установки выполнены.

- *Устанавливаются пределы измерений уровня: Plot Top 100 dB SPL и Plot Range 130 дБ.*
- *В меню Avg выбирается число усреднений –20*
- *В меню Options в окне Scaling Control выбирается логарифмическая шкала частот.*
- По команде **Run** начните измерение. Через 5...10 секунд зафиксируйте результаты измерений кнопкой **Stop**.
- Скопируйте график спектра шума генератора (рис.1.5) в файл вашего отчета.

Для этого левой клавишей мышки направьте курсор на график, затем нажмите правую клавишу мышки, в появившемся на экране меню выберите команду “Copy as a bitmap”. Откройте файл вашего отчета и вставьте в него график, подпишите название рисунка

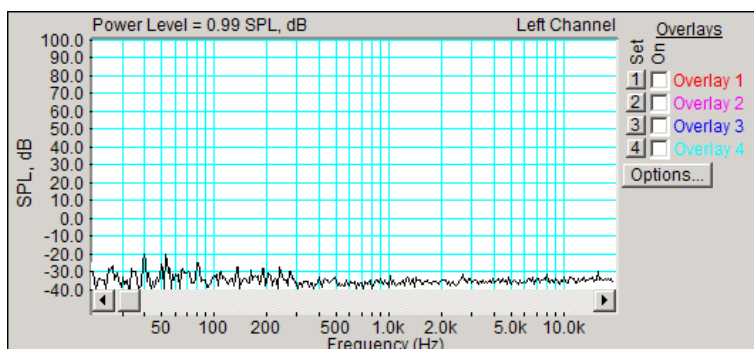


Рис.1.5. Спектр шума измерительного тракта

Обратите внимание на уровень шума квантования, сравните его с порогом слышимости белого шума, оцените возможную погрешность измерения. Уровень среднего значения спектральной плотности мощности определяет погрешность измерения абсолютного порога слышимости.

5.4. Измерение частотной зависимости порога слышимости

Измерение порогов слышимости производится на частотах, приведенных в табл.1.4. На результаты этих измерений существенное влияние оказывает неравномерность частотной характеристики телефонов, которая в настоящей работе никак не корректируется. В частотном диапазоне от 10 до 20 кГц частота испытательного сигнала на графике определяется по ее положению на шкале анализатора спектра. Сначала свою аудиограмму снимает первый исполнитель лабораторной работы, затем - второй.

- Загрузите файл **LR-1.2.AT.cfg**, который конфигурирует схему измерений и устанавливает все необходимые опции для выполнения настоящего раздела лабораторной работы.

Для этого в меню **Config** выберите опцию **Load Configuration** и в выпадающем меню выберите файл **LR-1.2. AT. cfg**.

Следующие первые четыре пункта задания, набранные курсивом, выполняются автоматически при загрузке файла **LR-1.2. AT.cfg**, нужно только проверить, что эти установки выполнены.

- *Устанавливаются пределы измерений уровня: Plot Top 100 dB SPL и Plot Range 130 дБ.*
- *В меню **Avg** выбирается число усреднений –5*
- *В меню **Options** в окне **Scaling Control** выбирается **логарифмическая шкала частот**.*
- *В меню генератора **Multiple Tones** выбирается сигнал с частотой 1 кГц и уровнем минус 20 дБ.*
- *Оденьте наушники и по команде **Run** начните измерение порога слышимости на частоте 1 кГц. Для этого с помощью регуляторов **Play Control** и **Wave** (Звук) уменьшайте звуковое давление до полного исчезновения звучания на слух.*

Для повышения точности измерения можно уменьшить уровень выходного сигнала генератора в меню **Level** с 0 до минус 20 или 40 дБ. Точность измерения порога

слышимости также повышается, если курсором периодически выключать (**Stop**) и включать (**Run**) тестовый сигнал.

Измерение порога слышимости производится по показаниям измерителя Power Level и контролируется (более точно) по шкале анализатора спектра (при малых уровнях эти показания могут не совпадать).

- Через 10...30 секунд зафиксируйте результаты измерения кнопкой **Stop**. Результат измерения порога слышимости занесите в табл.1.4 отчета.
- Повторите измерения на всех частотах табл.1.4. Если Вы слышите в диапазоне 10...20 кГц, можно сделать дополнительные измерения на выбранных вами частотах.
- В отчете по лабораторной работе постройте графики двух аудиограмм.

Табл.1.4. Аудиограммы исполнителей			
Частоты, Гц	Абс. порог, дБ SPL	Аудиограмма 1, дБ SPL	Аудиограмма 2, дБ SPL
20	82		
30	60		
50	40		
100	22		
200	12		
500	7		
1000	4		
2000	0		
3000	-3		
4000	-3.1		
5000	0		
6000	1.5		
7000	2		
8000	3		
9000	6		
10000	10		
16000	90		

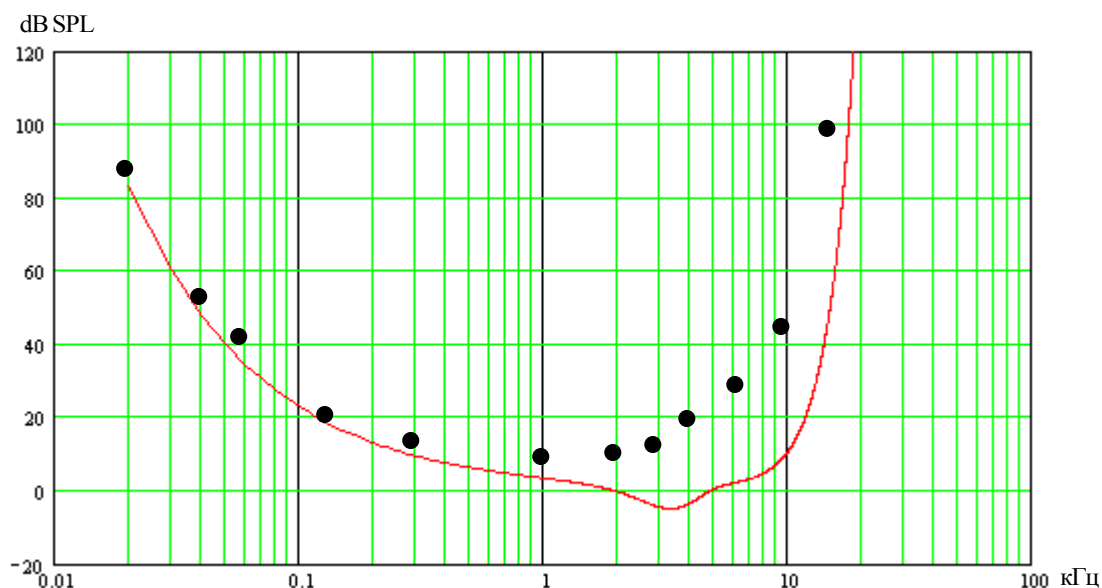


Рис.1.6. Аудиограммы исполнителей лабораторной работы

5.5. Измерение порога слышимости белого шума

- Загрузите файл **LR-1.3.AT.cfg**, который конфигурирует схему измерений и устанавливает все необходимые опции для выполнения настоящего раздела лабораторной работы.

Для этого в меню **Config** выберите опцию **Load Configuration** и в выпадающем меню выберите файл **LR-1.3. AT. cfg**.

Следующие первые четыре пункта задания, набранные курсивом, выполняются автоматически при загрузке файла **LR-1.3. AT.cfg**, нужно только проверить, что эти установки выполнены.

- Устанавливаются пределы измерений уровня: Plot Top 100 dB SPL и Plot Range 130 дБ.*
- В меню **Avg** выбирается число усреднений –20*
- В меню **Options** в окне **Scaling Control** выбирается *логарифмическая шкала частот*.*
- В меню генератора выбирается сигнал белого шума White Noise с уровнем 0 дБ.*
- Оденьте наушники и по команде **Run** начните измерение порога слышимости на белом шуме. Для этого с помощью регуляторов **Play Control** и **Wave** (Звук) уменьшайте звуковое давление до полного исчезновения звучания на слух.

Для повышения точности измерения можно уменьшить уровень выходного сигнала генератора в меню **Level** с 0 до минус 20 или 40 дБ. Точность измерения порога слышимости также повышается, если курсором периодически выключать (**Stop**) и включать (**Run**) тестовый сигнал.

Измерение порога слышимости производится по показаниям измерителя Power Level.

- Через 10...30 секунд зафиксируйте результаты измерения кнопкой **Stop**. Скопируйте график спектра (рис.1.7) белого шума на пороге слышимости в файл вашего отчета.

Для этого левой клавишей мышки направьте курсор на график, затем нажмите правую клавишу мышки, в появившемся на экране меню выберите команду **“Copy as a bitmap”**. Откройте файл вашего отчета **LR-1 AT** и вставьте в него график, подпишите название рисунка.

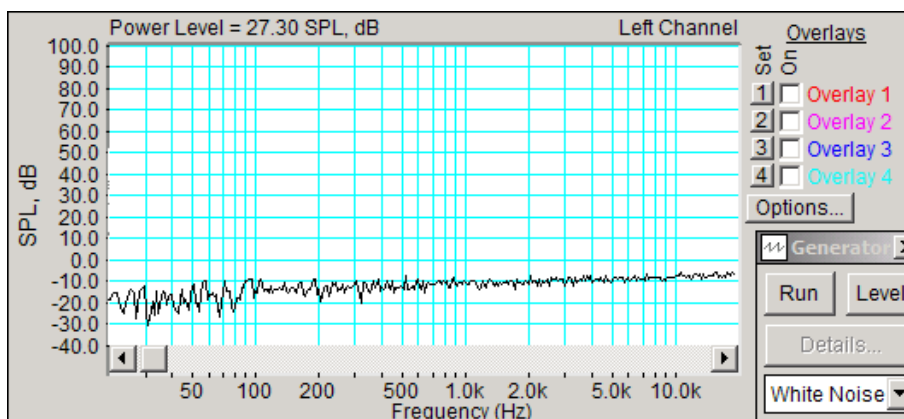


Рис.1.7. Спектр и уровень белого шума на пороге слышимости

6. Контрольные вопросы

- Какими физическими особенностями слуха определяется порог слышимости
- Почему чувствительность слуха максимальна в диапазоне 2...5 кГц и минимальна на границах звукового диапазона
- Что понимается под термином «Абсолютный порог слышимости»

4. Как и почему зависит порог слышимости от выбора источника звуковых колебаний и его местоположения при снятии аудиограммы
5. Как зависит порог слышимости от возраста, как меняется восприятие музыки с возрастом
6. Что понимается под терминами : «порог осязания» и «болевого порог». При каком уровне звукового давления возникает потеря слуха и происходит разрыв барабанной перепонки
7. Чему равны границы по частоте и уровню звукового давления музыки и речи
8. Как в технике компрессирования используется эффект порога слышимости
9. Как и почему зависит динамический диапазон восприятия звуковых колебаний на различных частотах от порога слышимости
10. Поясните, почему из-за порога слышимости не имеет смысла увеличивать число разрядов цифровых сигналов свыше 16..18.
11. Как связана слышимость шума квантования с порогом слышимости белого шума

7. Содержание отчета

- 7.1. Титульный лист
- 7.2. Цель работы
- 7.3. Задание к лабораторной работе
- 7.4. Функциональная схема измерений
- 7.5. Результаты измерений в виде графиков
- 7.6. Таблицы с результатами измерений и расчетов
- 7.7. Выводы по результатам измерений

8. Литература

Вологдин Э.И. Слух и восприятие звука. Учебное пособие по курсу «Аудиотехника» СПб. ГУТ, 2004.

ЛАБОРАТОРНАЯ РАБОТА N 2 АТ

ИССЛЕДОВАНИЕ МАСКИРОВКИ ТОНАЛЬНЫХ ЗВУКОВ БЕЛЫМ ШУМОМ

Руководство к выполнению лабораторной работы по Аудиотехнике

1. Цель работы

Целью лабораторной работы является ознакомление студентов с методами исследования маскировки тональных звуков широкополосным белым шумом с использованием компьютерной программы SpectraLab.

2. Задание к лабораторной работе

- Ознакомиться с программой SpectraLab в объеме, необходимом для измерения порога слышимости тональных звуков при маскировке их белым шумом
- Измерить частотную зависимость уровня спектральной плотности мощности и интегральное значение суммарного уровня акустических и электрических шумов измерительного тракта
- Измерить частотную зависимость порога слышимости тональных звуков в диапазоне 30...16000 Гц с использованием головных стерео телефонов
- Измерить порог слышимости белого шума
- По результатам измерений сформулировать выводы

3. Функциональная схема измерений

В приведенной на **рис.2.1** схеме тестовый измерительный сигнал формируется с помощью двух программных генераторов и сумматора. Генератор программы SpectraLab вырабатывает сигнал белого шума, вспомогательный генератор FR может генерировать тональный сигнал в диапазоне от 20 до 20000 и имеет собственный ступенчатый регулятор уровня.

Суммарный сигнал с двух генераторов подается на регулятор уровня Play Control, который одновременно изменяет уровни шумового и тонального сигнала. С линейного выхода звуковой карты тестовый сигнал через регулятор уровня Record Control подается на вход анализатора спектра и телефоны. В анализаторе спектра используется амплитудная шкала SPL (Sound Pressure Level), в которой можно измерять уровни звуковых давлений тонального сигнала в диапазоне от минус 20 до плюс 100 дБ SPL.

В приведенной схеме регуляторами двумя регуляторами Play Control устанавливается

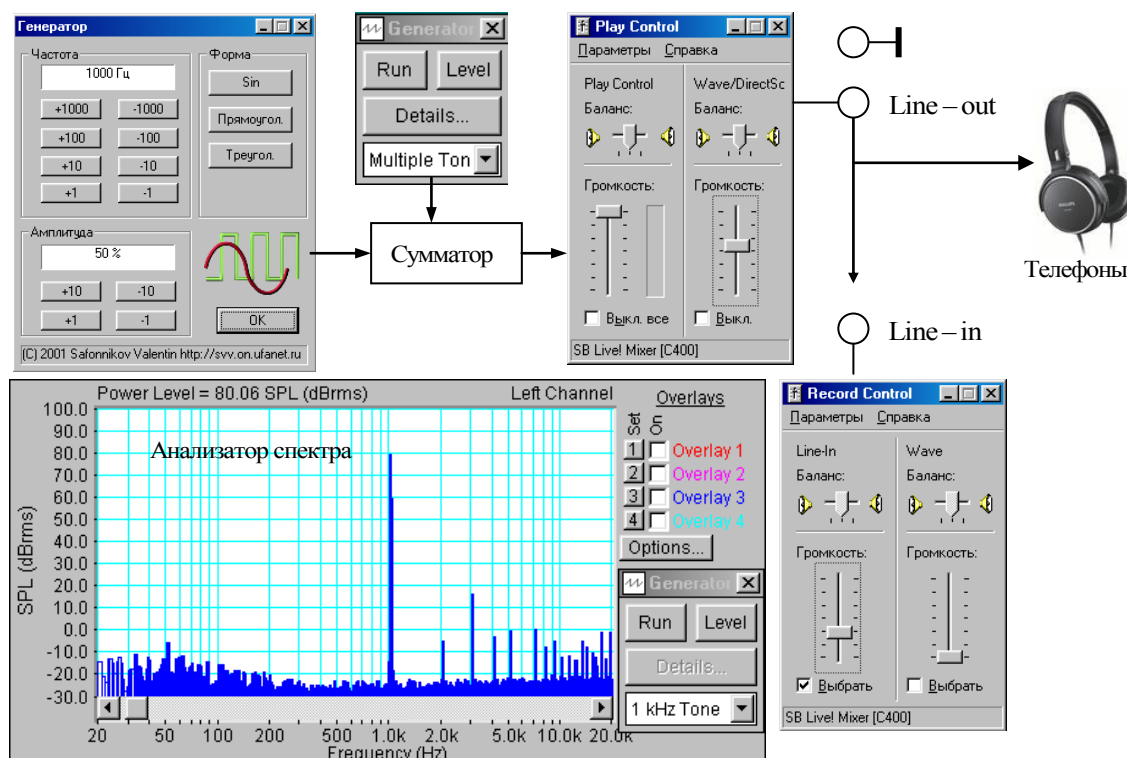


Рис. 2.1. Функциональная схема измерений

уровень спектральной плотности мощности маскирующего белого шума по шкале анализатора спектра. Порог слышимости подбирается линейным регулятором уровня генератора F_g на слух с помощью стерео телефонов. Измеряется он с помощью индикатора уровня Power Level и шкалы анализатора спектра.

4. Методические указания

4.1. Характеристики звуковых колебаний. Звуковое давление измеряется в Паскалях (Па)

$$1 \text{ Па} = \text{Ньютон} / \text{м}^2.$$

Слух воспринимает давления от 10^{-5} до 200 Па. Интенсивность звуковых колебаний измеряется в $\text{Вт}/\text{м}^2$. Звуковое давление P и интенсивность I связаны квадратичной зависимостью

$$I = \frac{P^2}{Z_s}$$

где $Z_s = Q_s \cdot C_s$ – удельное акустическое сопротивление воздушной среды,

$Q_s = 1,23 \text{ кг}/\text{м}^3$ – плотность воздушной среды, $C_s = 340 \text{ м}/\text{с}$ – скорость распространения звука.

В аудиотехнике звуковое давление и интенсивность звуковых колебаний обычно измеряется в децибелах в шкале SPL (Sound Pressure Level), В этой шкале за 0 дБ принята интенсивность звука

$$I_0 = 10^{-12} \text{ Вт}/\text{м}^2$$

Этой интенсивности соответствует звуковое давление $P_0 = 2 \cdot 10^{-5} \text{ Па}$.

Уровни звукового давления и интенсивности тональных звуков рассчитываются по формулам:

$$L_p = 20 \cdot \log \left(\frac{P}{P_0} \right),$$

$$L_{\text{int}} = 10 \cdot \log \left(\frac{I}{I_0} \right) = 20 \cdot \log \left(\frac{P}{P_0} \right)$$

Из этих формул следует, что уровень звукового давления всегда равен уровню интенсивности звука, и поэтому можно использовать просто термин «уровень». Интенсивность звука через уровень звукового давления можно определить по формуле

$$I = 10^{\frac{L_p}{10}}$$

Основной характеристикой акустического шума является спектральная плотность мощности D_s , равная интенсивности шума в полосе шириной 1 Гц. Для белого шума S_D не зависит от частоты и определяется равенством

$$D_s = \frac{I}{\Delta F},$$

где ΔF – полная ширина спектра шума. Уровень спектральной плотности мощности рассчитывается по формуле

$$L(D_s) = 10 \cdot \log \left(\frac{D_s}{D_{s0}} \right).$$

В шкале SPL $D_{s0} = 10^{-12}$ Вт/м² Гц. Уровень $L(S_D)$ обычно измеряется, а интегральный уровень шума в заданной полосе частот определяется расчетным путем через произведение спектральной плотности мощности на ширину полосы частот. В логарифмической шкале этот уровень аналитически записывается в виде

$$\bar{L}_n = L(D_s) + 10 \log \Delta F$$

где ΔF – ширина выбранной полосы частот.

Экспериментально график спектральной плотности шума снимается с помощью быстрого преобразования Фурье (БПФ), при котором всегда возникают ошибки измерения, которые надо корректировать. Эти ошибки связаны с тем, что при БПФ полоса сканирования не равна 1 Гц, она зависит числа выборок N_{fft} , используемых для анализа, частоты дискретизации f_s и выбора оконной функции. Погрешность измерения уровня спектральной плотности в децибелах рассчитывается по формуле

$$\Delta L_{fft} = 10 \cdot \log \left(\frac{N_{fft}}{K_w \cdot f_s} \right)$$

где K_w – коэффициент оконной функции. Для измерения спектра обычно рекомендуется оконная функция Hanning, для которой этот коэффициент равен 1,5. При выборе этой функции и частоты дискретизации 48 кГц погрешность близка к нулю когда $N_{fft} \approx 72000$.

4.2. Эффект частотной маскировки тона шумом. При одновременном воздействии на слух тона и шума один из них может быть не слышен на фоне другого. *Этот эффект называется частотной маскировкой.* В настоящей лабораторной работе изучается эффект маскировки тона белым шумом. Механизм маскировки в этом случае основывается на том, что органы слуха реагируют не на общую мощность маскирующего звука, а на его мощность в критической полосе слуха, в которой находится маскируемый тон. Слух расчленяет широкополосный маскирующий шум на частотные группы и

распознает испытательный тон с частотой F_T когда его уровень ниже уровня интенсивности маскирующего шума в критической полосе тестового тона.

При таком механизме эффект маскировки оценивается с помощью *коэффициента маскировки* K_M , показывающего, на сколько децибел порог слышимости маскируемого тона L_T с частотой F_T ниже интегрального уровня интенсивности маскирующего шума \bar{L}_N в критической полосе на этой частоте

$$K_M = L_T - \bar{L}_N$$

Этот коэффициент маскировки всегда отрицательный, это означает, что слышен тональный звук, уровень интенсивности которого ниже уровня мощности маскирующего шума. Чем больше абсолютное значение K_M , тем выше эффект маскировки.

На **рис.2.2.** поясняется идея определения коэффициента маскировки.. Маскируемый тон с частотой F_T имеет критическую полосу шириной B_C . Приблизительно она

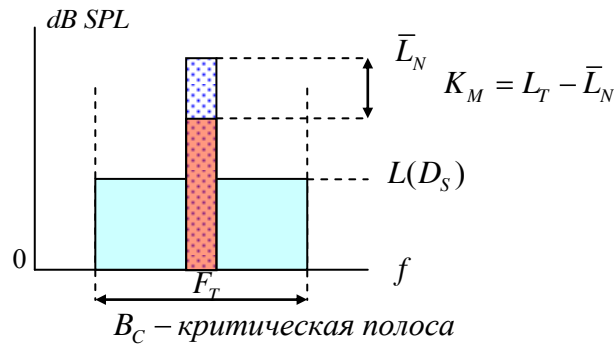


Рис.2.2. Маскировка тона шумом

определяется равенствами

$$\begin{aligned} B(f)_C &= 100 \text{ Hz}, & F_T < 500 \text{ Hz} \\ B(f)_C &= 0,2 \cdot F_T \text{ Hz}, & F_T > 500 \text{ Hz} \end{aligned}$$

более точно рассчитывается по формуле

$$B(f)_C = 25 + 75[1 + 1,4(F_T / 1000)^2]^{0,69}, \text{ Hz}$$

В этой полосе уровень спектральной плотности маскирующего белого шума $L(D_s)$, а

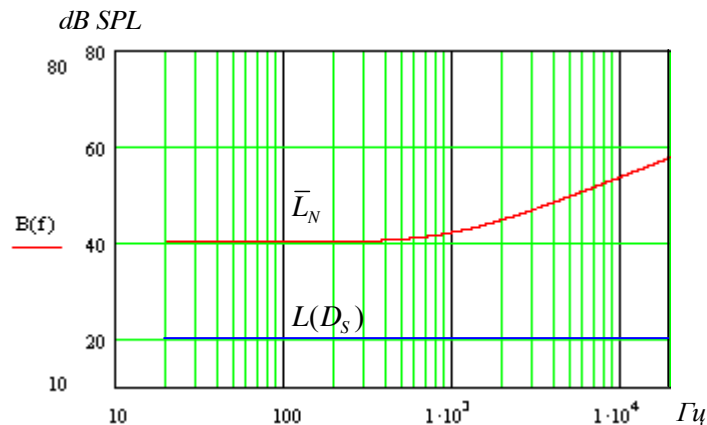


Рис.2. 3. Зависимость уровня маскирующего шума от частоты

его интенсивность I_N определяется произведением спектральной плотности D_s на

ширину критической полосы. При этом уровень интенсивности шума в критической полосе \bar{L}_{NC} определяется равенством

$$\bar{L}_N(f) = L(D_s) + 10 \cdot \log B(f)_c$$

Как видно из рис. 2.3, на частотах ниже 500 Гц интегральный уровень шума в критической полосе \bar{L}_{NC} больше уровня спектральной плотности мощности $L(D_s)$ на 20 дБ, на частотах выше 500 Гц с расширением критической полосы уровень интенсивности маскирующего шума увеличивается почти линейно с крутизной примерно 3,6 дБ/октава.

На рис.2.4 приведены экспериментальные кривые частотной маскировки тона белым шумом для разных уровней спектральной плотности мощности маскирующего шума

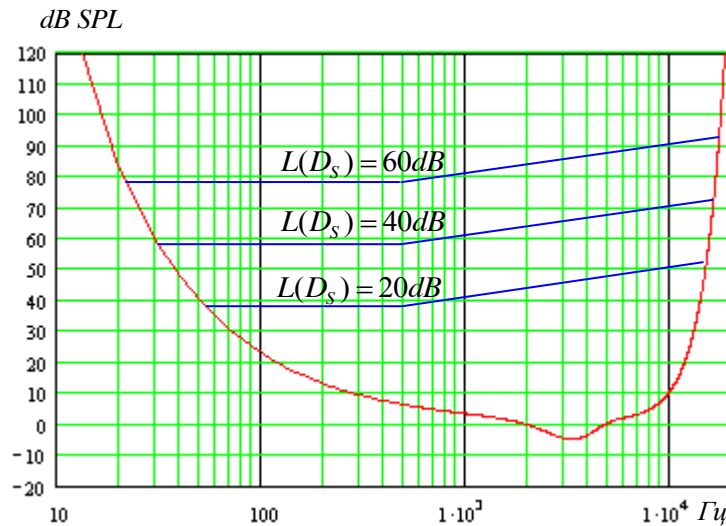


Рис.2.4 Частотная зависимость порога слышимости тона маскируемого белым шумом

$L(D_c)$. На этом рисунке по оси ординат откладываются уровни порога слышимости маскируемого тона L_T и уровня интенсивности шума в критической полосе \bar{L}_N , а по оси абсцисс – частота маскируемого тона F_T . Как видно, в диапазоне до 500 Гц порог слышимости на 18 дБ выше уровня спектральной плотности мощности маскирующего шума и на 2 дБ ниже уровня мощности шума в критической полосе.

4.3. Особенности маскировки тона белым шумом. Эти особенности могут быть сформулированы следующим образом:

- основной характеристикой маскирующего шума является уровень его спектральной плотности мощности $L(D_c)$ в шкале SPL;
- до частоты 500 Гц порог слышимости маскируемого тона L_{THT} от его частоты не зависит и при любых значениях $L(D_c)$ его уровень на 20 дБ ниже уровня интенсивности шума ;
- на частотах выше 500 Гц порог слышимости маскируемого тона повышается примерно с крутизной 3 дБ/октава, при увеличении его частоты в 10 раз порог возрастает на 10 дБ.
- эффект маскировки измеряется только в области частот, где порог слышимости тона выше абсолютного порога слышимости.

Ниже 500 Гц все критические полосы слуха имеют одинаковую ширину 100 Гц, поэтому интенсивность маскирующего шума в критических полосах одинаковая и порог

слышимости маскируемого тона от частоты не зависит. В области частот, где ширина критических полос слуха пропорциональна средней частоте критической полосы, при увеличении частоты в 10 раз интенсивность маскирующего шума в критических полосах также возрастает в 10 раз, порог слышимости увеличивается на 10 дБ.

Графики частотной зависимости порога слышимости и интенсивности маскирующего шума в критической полосе имеют разную крутизну нарастания с увеличением частоты. По этой причине возникает частотная зависимость порога слышимости. Экспериментальный порог слышимости хорошо аппроксимируется функцией

$$L(f)_T = -2 - \log \left[1 + \left(\frac{F_T}{502} \right)^{2,5} \right].$$

Ее график приведен на [рис.2.5](#). Из этого графика видно, что коэффициент маскировки до 500 Гц от частоты не зависит и равен минус 2 дБ. На частотах выше 500 Гц коэффициент маскировки увеличивается почти линейно и на частоте 20 кГц достигает

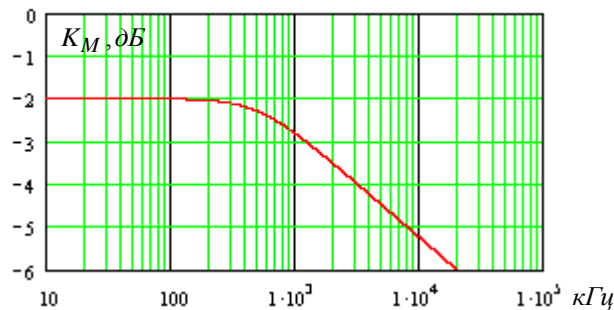


Рис.2.5. Частотная зависимость коэффициента маскировки тона шумом

значения минус 6 дБ.

4.4. Функции маскировки в шкале барк. Кривые частотной маскировки достаточно сложны для математического моделирования, поэтому при разработке алгоритмов компрессии рассчитывается зависимость коэффициента маскировки от высоты тона

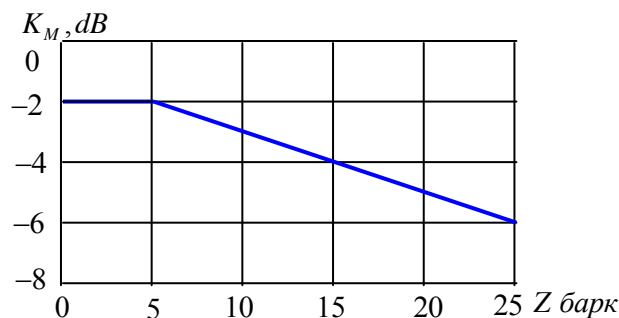


Рис.2.6. График коэффициента маскировки тона шумом







звука в барках (Z). Барк, как известно, равен критической полосе слуха, а связь между частотой сигнала F и высотой тона в барках Z определяется равенством

$$Z = 13 \cdot \arctan(0,00076 \cdot F) + 3,5 \cdot \arctan \left[\left(\frac{F}{7500} \right)^2 \right]. \quad (2-4)$$

При такой замене частотному диапазону 20...20000 Гц соответствует диапазон в барках от 0 до 25. При этом функцию маскировки часто возможно аппроксимировать двумя наклонными прямыми линиями разной крутизны ([рис.2.6](#)), относительно абсолютного порога слышимости. Иногда используется более сложная аппроксимация с помощью полиномов.

5. Порядок выполнения лабораторной работы

5.1. Функции кнопок управления:

- кнопкой с пиктограммой  осуществляется автоматическая установка пределов измерения по вертикальной оси шкалы анализатора спектра, фазометра или осциллографа
- кнопкой с пиктограммой  устанавливается частотный диапазон шкалы от 1 Гц до частоты Найквиста
- кнопкой с пиктограммой  вызывается маркер, которым устанавливается желаемый частотный диапазон шкалы
- при нажатии кнопки с пиктограммой  графики рисуются тонкими линиями
- при нажатии кнопки с пиктограммой  графики рисуются вертикальными линиями или столбиками
- кнопкой с пиктограммой  открывается меню анализатора спектра, фазометра или осциллографа, в котором вводятся необходимые опции.
- кнопкой **Run** включаются все измерительные приборы
- кнопкой **Stop** выключаются все измерительные приборы и запоминаются результаты измерений
- в меню **Avg** задается число усредняемых результатов измерения
- в окошке **Peak Hold** курсором включается функция удержания пиковых значений напряжения


5.2. Подготовка к измерениям


Конфигурация схемы измерений и ввод необходимых опций производится с помощью загрузочных файлов **LR- 2.1 AT.cfg** и **LR- 2.2 AT.cfg**. На экран монитора вызываются окна: программа SpectraLab, анализатор спектра, регулятор уровня выходного сигнала синтезатора тестовых сигналов, генератор белого шума и вспомогательный генератора синусоидальных сигналов. Измерения производятся в режиме моно, в реальном времени в частотном диапазоне 20...20000 Гц с использованием тестовых *тональных сигналов и белого шума* в логарифмической шкале частот.

В лабораторной работе на слух с помощью телефонов на дискретных частотах измеряются пороги слышимости тональных звуков, маскируемых широкополосным белым шумом.


- Скопируйте с рабочего стола файл «Отчет LR-2 AT» на свою флешкарту, дайте ему свое имя в виде: «Группа»_ «Фамилия (работавшего на клавиатуре)»_ «Номер ЛР»_ «Дата».
- Откройте этот файл и сверните его окно.

В файл этого отчета вы будете копировать все экспериментальные графики, и заносить в таблицы отчета результаты измерений сразу в процессе их получения. В день выполнения лабораторной работы отчет в электронном виде необходимо сдать преподавателю, только тогда работа считается выполненной.


- Кнопкой с пиктограммой  откройте окно программы **SpectraLab**. Не меняйте размеры этого окна, они согласованы с Вашим отчетом
- Сделайте обязательные установки в микшере **Play Control**

Для этого двойным *резким* нажатием кнопки с пиктограммой  откройте его окно. Затем после команд **Параметры** \Rightarrow **Свойства** в появившемся окне **Свойства** выберите режим **Воспроизведение**. В выпадающем меню этого окна курсором отметьте регуляторы **Play Control**, **Wave** и **Line in**, по команде **ОК** окна этих регуляторов появляются на экране монитора. Курсором включите первые два регулятора и поставьте их ползунки в верхнее положение шкалы. Регулятор **Line in** *обязательно выключите*.

- Сделайте обязательные установки в микшере **Record Control**

Для этого двойным *резким* нажатием кнопки с пиктограммой  откройте окно микшера **Play Control**. Затем после команд **Параметры** \Rightarrow **Свойства** в появившемся окне **Свойства** выберите режим **Запись**. В выпадающем меню этого окна курсором отметьте регуляторы **Microphone** и **Line in**. По команде **ОК** окна этих регуляторов появляются на экране монитора. Курсором поставьте их ползунки в верхнее положение шкалы.

- Курсором сначала включите регулятор **Microphone**, затем переключите в режим **Line in**

- Кнопкой с пиктограммой  откройте окно генератора **Fr** и курсором переключите генератор в режим синусоидальных колебаний.

5.3. Измерение спектра и уровня шума измерительного тракта

В этом разделе лабораторной работы измеряется уровень шума измерительного тракта, из-за которого могут возникать значительные ошибки измерения порогов слышимости на различных частотах. Акустический шум при измерениях порождается наушниками, которые выполняют функцию микрофонов. Этот шум суммируется с шумами электронного тракта и маскирует тестовый тональный сигнал при измерении порога слышимости. Шум характеризуется частотной зависимостью его спектральной плотности мощности и интегральным значением мощности, уровень которого индицируется на шкале **PowerLevel**.

- Загрузите файл **LR-2.1.AT.cfg**, который конфигурирует схему измерений и устанавливает все необходимые опции для выполнения настоящего раздела лабораторной работы.

Для этого в меню **Config** выберите опцию **Load Configuration** и в выпадающем меню выберите файл **LR-2.1. AT. cfg**.

Следующие первые три пункта задания, набранные курсивом, выполняются автоматически при загрузке файла **LR-2.1. AT.cfg**, нужно только проверить, что эти установки выполнены.

- *Устанавливаются пределы измерений уровня: **Plot Top 100 dB SPL** и **Plot Range 130 дБ**.*
- *В меню **Avg** выбирается число усреднений –20*
- *В меню **Options** в окне **Scaling Control** выбирается *логарифмическая шкала частот*.*
- По команде **Run** начните измерение. Через 5...10 секунд зафиксируйте результаты измерений кнопкой **Stop**.
- Скопируйте график спектра шума генератора в файл вашего отчета (**рис.2.7.**).

Для этого левой клавишей мышки направьте курсор на график, затем нажмите правую клавишу мышки, в появившемся на экране меню выберите команду **“Copy as a bitmap”**. Откройте файл вашего отчета и вставьте в него график, подпишите название рисунка

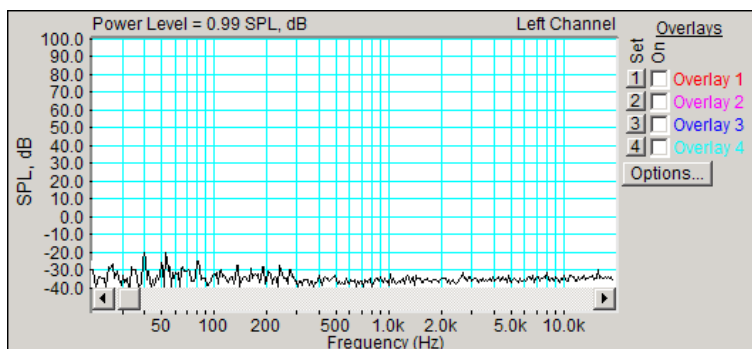


Рис.2.7. Спектр шума измерительного тракта

Обратите внимание на уровень шума квантования, сравните его с порогом слышимости белого шума, оцените возможную погрешность измерения. Уровень среднего значения спектральной плотности мощности определяет погрешность измерения абсолютного порога слышимости.

5.4. Измерение частотной зависимости порога слышимости маскируемых тональных звуков

Измерение порогов слышимости производится на частотах, приведенных в [табл.2.1](#). На результаты этих измерений существенное влияние оказывает неравномерность частотной характеристики телефонов, которая в настоящей работе никак не корректируется. Сначала частную зависимость снимает первый исполнитель лабораторной работы, затем - второй.

- Загрузите файл **LR-2.2.AT.cfg**, который конфигурирует схему измерений и устанавливает все необходимые опции для выполнения настоящего раздела лабораторной работы.

Для этого в меню **Config** выберите опцию **Load Configuration** и в выпадающем меню выберите файл **LR-2.2. AT. cfg**.

Следующие первые четыре пункта задания, набранные курсивом, выполняются автоматически при загрузке файла **LR-1.2. AT.cfg**, нужно только проверить, что эти установки выполнены.

- *Устанавливаются пределы измерений уровня: Plot Top 100 dB SPL и Plot Range 130 дБ.*
- *В меню **Avg** выбирается число усреднений –5*
- *В меню **Options** в окне **Scaling Control** выбирается **логарифмическая шкала частот**.*
- *В меню генератора выбирается сигнал белого шума White Noise с уровнем 0 дБ.*
- Командой **Run** включите схему измерений и регулятором Record Control установите по шкале анализатора спектра заданный преподавателем (20...80 дБ SPL) уровень спектральной плотности шума
- Курсором установите на шкале вспомогательного генератора частоту 1000 Гц
- Оденьте наушники и определите порог слышимости тонального звука

Для этого курсором ступенями через 10 и 1 % подберите уровень сигнала с частотой 1 кГц, когда он равен порогу слышимости. Точность измерения порога слышимости повышается, если курсором периодически изменять уровень сигнала $\pm 1\%$. Измерение порога слышимости производится по шкале анализатора спектра с помощью курсора.

- Зафиксируйте результаты измерения кнопкой **Stop**. Результат измерения порога слышимости занесите в **табл.2.1** отчета.
- Скопируйте сигналограмму анализатора спектра в файл вашего отчета (**рис.2.8.**)

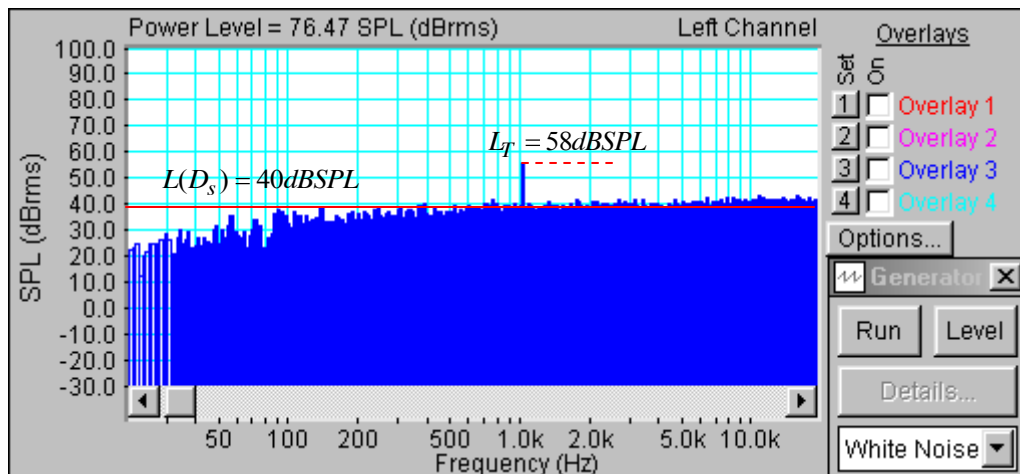


Рис.2. 8. К определению порога слышимости тона 1 кГц

Для этого левой клавишей мышки направьте курсор на график, затем нажмите правую клавишу мышки, в появившемся на экране меню выберите команду **“Copy as a bitmap”**. Откройте файл вашего отчета и вставьте в него график, подпишите название рисунка

- Повторите измерения на всех частотах табл.1, сигналограммы копировать не надо. Если Вы слышите в диапазоне 10...20 кГц, можно сделать дополнительные измерения на выбранных вами частотах.

Табл.2.1. Таблицы порога слышимости тонов, маскируемых шумом		
Частоты, Гц	Аудиограмма 1, дБ SPL	Аудиограмма 2, дБ SPL
20		
30		
50		
100		
200		
500		
1000		
2000		
3000		
4000		
5000		
6000		
7000		
8000		
9000		
10000		
16000		

- На графике **рис.2.9** в отчете по лабораторной работе постройте экспериментальный график порога слышимости тона, маскируемого белым шумом.
- На графике **рис.2.9** в отчете по лабораторной работе постройте теоретический график порога слышимости тона, маскируемого белым шумом для заданного значения уровня спектральной плотности мощности шума..

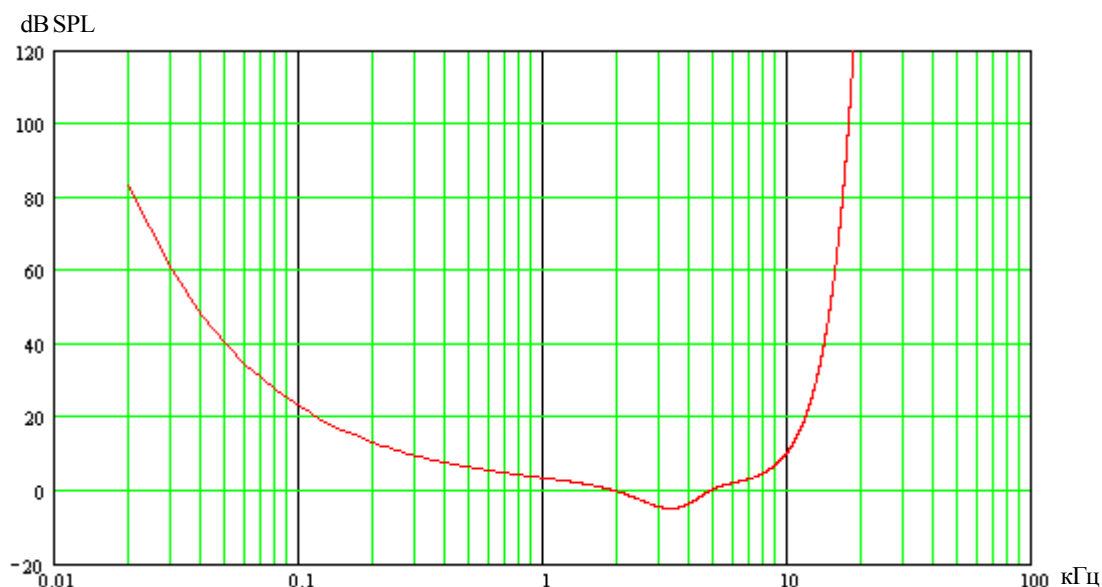


Рис. 2.9. Графики порога слышимости тонов, маскируемых белым шумом

6. Контрольные вопросы

1. Какой вид имеет график спектральной плотности белого шума в логарифмической шкале частот
2. Дайте определение понятию «критическая полоса» слуха, как она зависит от частоты
3. Опишите порядок расчета мощности белого шума в критической полосе
4. Каким образом связаны мощность шума в критической полосе и уровень спектральной плотности мощности шума
5. Как рассчитывается коэффициент маскировки в критической полосе
6. Сформулируйте порядок определения порога слышимости тона, маскируемого шумом
7. Как связаны порог слышимости тона и уровень спектральной плотности мощности шума
8. Опишите характер зависимости порога слышимости от частоты
9. Какой вид имеет зависимость коэффициента маскировки тона шумом от частоты

7. Содержание отчета

- 7.1. Титульный лист
- 7.2. Цель работы
- 7.3. Задание к лабораторной работе
- 7.4. Функциональная схема измерений
- 7.5. Результаты измерений в виде графиков
- 7.6. Таблицы с результатами измерений и расчетов
- 7.7. Выводы по результатам измерений

8. Литература

Вологдин Э.И. Слух и восприятие звука. Учебное пособие по курсу «Аудиотехника» СПб. ГУТ, 2004.

ЛАБОРАТОРНАЯ РАБОТА N 3 АТ

ИССЛЕДОВАНИЕ ДИНАМИЧЕСКИХ ХАРАКТЕРИСТИК ЦИФРОВЫХ ФОНОГРАММ

Руководство к выполнению лабораторной работы по Аудиотехнике

3. Цель работы

Целью лабораторной работы является ознакомление студентов с методиками измерения динамического диапазона и пик-фактора цифровых фонограмм по их статистическим характеристикам с использованием компьютерной программы Audition 3 (Cool Edit Pro).

2. Задание к лабораторной работе

- Ознакомиться с программой Audition 3 в объеме, необходимом для измерения статистических характеристик цифровых фонограмм
- Измерить мгновенные динамические характеристики 3 фонограмм при времени интегрирования 1 мс
- Измерить усредненные динамические характеристики 3 фонограмм при времени интегрирования 60 мс
- Исследовать зависимость результатов измерений динамического диапазона от времени интегрирования в диапазоне от 1 до 60 мс
- Исследовать зависимость динамического диапазона от жанров музыкальных произведений
- По результатам измерений сформулировать выводы

3. Функциональная схема измерений

Приведенной на **рис.3.1** схема поясняет порядок измерения динамического диапазона и пик-фактора цифровой фонограммы в программе Cool Edit Pro (Audition 3). Измерение начинается с того, что раскрывается файл фонограммы в виде сигналограммы. Затем по команде **Анализ** рассчитываются статистические характеристики фонограммы, из которых основными являются: Peak Amplitude (L_{pic}), (Maximum RMS Power (L_{max}), Minimum RMS Power (L_{min}) и Average RMS Power (L_{avr}) (уровни максимального, минимального и среднего среднеквадратического (эффективного) значения мощности сигнала).

Динамический диапазон фонограммы определяется как

$$DR_m = L_{pic} - L_{min}, \quad (3-1)$$

пик-фактор рассчитывается по формуле

$$PF_m = L_{pic} - L_{avr} \quad (3-2)$$

Динамический диапазон может быть также определен по гистограмме распределения уровней фонограммы.

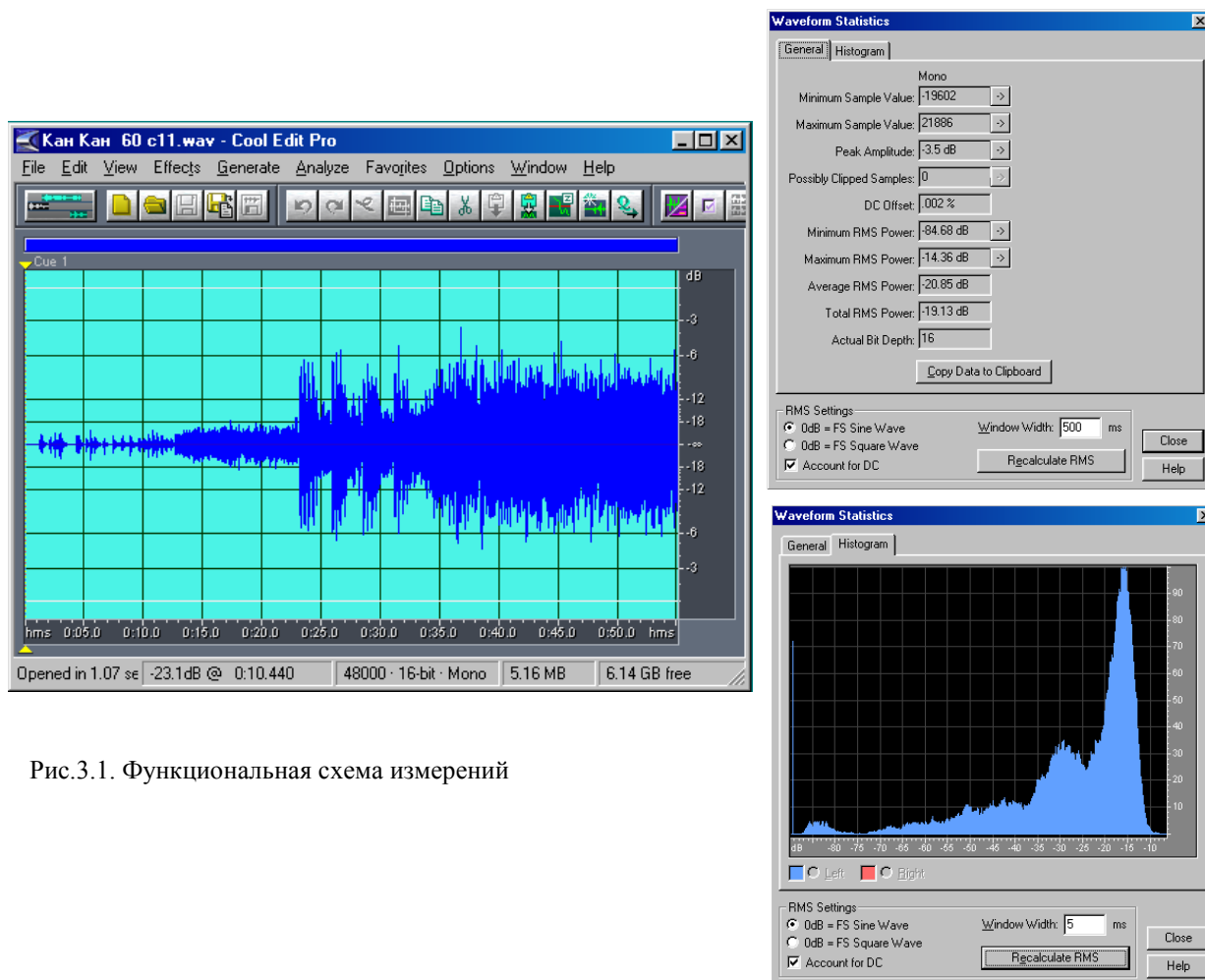


Рис.3.1. Функциональная схема измерений

4. Методические указания

4.1. К *динамическим характеристикам* цифровой фонограммы относятся: динамический диапазон, пик-фактор и функция распределения уровней фонограммы (гистограмма). К ним часто относят и динамический диапазон музыки фонограммы, который иногда существенно меньше динамического диапазона фонограммы.

4.2. *Динамический диапазон цифровой фонограммы* обычно определяется как отношение в децибелах максимального и минимального *мгновенных уровней* электрического сигнала фонограммы. В программе Audition 3 динамический диапазон фонограммы определяется через статистические характеристики фонограммы по формуле (3-1) через пиковое и минимальное среднеквадратическое значение уровня сигнала фонограммы. Эта характеристика связывает пиковое значение сигнала, шум фонограммы и число разрядов. Для определения этой характеристики *время интегрирования* при измерении уровней фонограммы не имеет существенного значения. Однако, оно сильно влияет на вид гистограммы, поэтому его принимают равным 1...5 мс.

4.3. *Динамический диапазон музыки* определяется отношением в децибелах самого громкого звука (фортиссимо) и самого тихого звука (пианиссимо). *Динамический диапазон симфонической музыки* составляет 65...75 дБ, а на концертах рок-музыки он возрастает до 105 дБ, при этом пики звуковых давлений могут достигать 122...130 дБ SPL.

Табл.3.1. Динамический диапазоны музыки и голоса	
Источник звука	Динамический диапазон, дБ
Большой симфонический оркестр	60...80
Малый симфонический оркестр	40...50
Духовой оркестр	30...40
Эстрадный оркестр	35...40
Хор	30...45
Голос (основной тон)	25...35

Динамический диапазон вокальных исполнителей - не превышает 35...45 дБ (табл.3.1).

При оценке динамического диапазона музыки на слух принципиальное значение имеют такие свойства слуха человека как постоянная времени, равная 50...60 мс, и эффект адаптации слуха. Благодаря этому эффекту в тишине чувствительность слуха повышается, а при громких звуках она понижается. Эти факторы должны учитываться при осуществлении объективных измерений динамического диапазона музыки техническими средствами. Поэтому динамический диапазон музыки измеряется через максимальное и минимальное среднеквадратическое значения уровней фонограммы по формуле

$$DR_p = L_{\max} - L_{\min} \quad (3-3)$$

при времени интегрирования 50...60 мс.

4.4. *Пик-фактор фонограммы* определяется через ее статистические характеристики: пиковое и среднее среднеквадратическое значения уровней фонограммы по формуле (2). Чем меньше пик-фактор, то громче звучит фонограмма. *Пик-фактор музыки* определяется через максимальное и минимальное среднеквадратические значения уровней фонограммы по формуле

$$PF_p = L_{\max} - L_{avg} \quad (3-4)$$

4.5. *Статистические характеристики* фонограмм являются основой для измерения динамических характеристик. В звуковом редакторе Adobe Audition 3 к ним относятся :

- максимальное положительное значение выборки (в десятичном коде)
- максимальное отрицательное значение выборки (в десятичном коде),
- пиковое значение уровня сигнала в шкале FS ,
- текущее минимальное эффективное значение уровня сигнала,
- текущее максимальное эффективное значение уровня сигнала,
- среднее значение эффективного уровня сигнала за время измерения,
- эффективное значение уровня сигнала за время измерения,
- гистограмма распределения уровней фонограммы.

Измерение уровней напряжений производится в шкале FS от 0 дБ до минус бесконечности. За 0 dBFS принимается напряжение с уровнем + 4 Dbu (эффективное напряжение 1,23В).

Гистограмма — это способ графического представления табличных данных случайных событий в виде прямоугольников, высота которых равна числу (частоте) событий, попадающих в отдельные интервалы общего диапазона изменения этих событий. Например, распределение солнечных дней может быть представлено по годам, временам года или по месяцам.

Гистограмма распределения уровней фонограммы. В такой гистограмме горизонтальная ось градуируется в уровнях фонограммы от L_{\max} до L_{\min} и разбивается на интервалы ΔL , шириной от 5 до 0,1 дБ. По вертикальной шкале откладывается число отсчетов, попадающих в различные интервалы по шкале уровней. Функция распределения уровней может быть представлена в виде

$$F(i) = \frac{n_i}{\Delta L} \quad (3-5)$$

где n_i - число отсчетов, попадающих в i -й интервал.

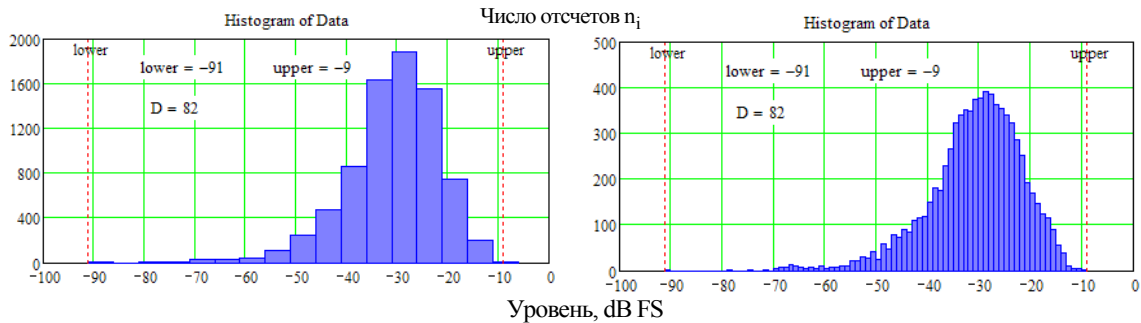


Рис.3.2. Гистограммы распределения уровней фонограммы с интервалом усреднения 5 и 1 дБ.

В качестве примера на рис.3.2 приведены гистограммы распределения уровней одной и той же фонограммы с интервалом ΔL , равным 5 и 1 дБ ($n = 7755$).

Нормализованная гистограмма. При изменении длительности фонограммы, времени интеграции и интервала гистограммы пропорционально меняется масштаб по вертикальной оси. Чтобы исключить эти зависимости используется нормализованная

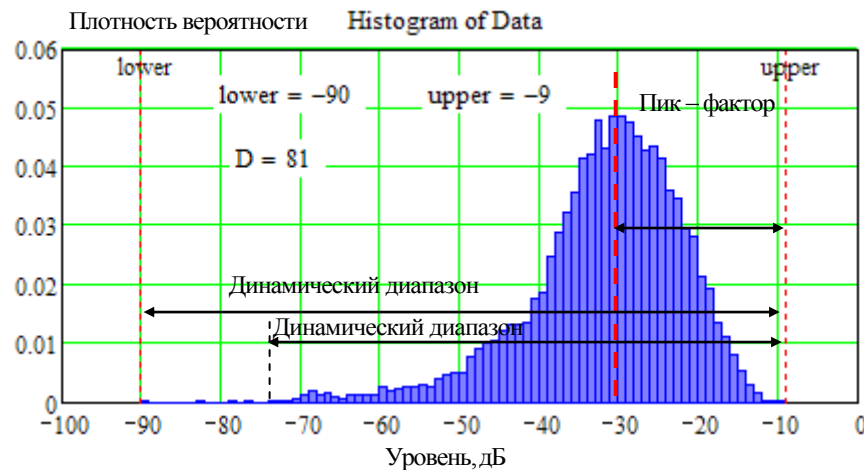


Рис. 3.3 Распределение плотности вероятности уровней фонограммы с интервалом 1 дБ функция распределения

$$F(i) = \frac{n_i}{n \cdot \Delta L} \quad (3-6)$$

В этой формуле общее число отсчетов уровней за время измерения n определяется по формуле

$$n = \frac{T_{ph}}{T_{in}}$$

где T_{ph} - длительность фонограммы, T_{in} - время интеграции при измерении уровней.

На **рис.3.3** приведена нормализованная гистограмма фонограммы, в ней по вертикальной оси откладывается плотность вероятности распределения уровней фонограммы, время интеграции равно 21 мс. С помощью нормализованной гистограммы можно определить такие статистические характеристики как динамический диапазон, наиболее вероятный уровень сигналов и пик-фактор фонограммы. Важно, что эти характеристики измеряются с определенной вероятностью. Так, например, нижнюю границу диапазона уровней фонограммы можно считать равной от -90 до -75 дБ, так как вероятностей таких уровней ничтожно мала, соответственно при этом уменьшается и динамический диапазон.

В звуковом редакторе Adobe Audition 3 используется нормализация гистограммы, при которой максимальной вероятности событий всегда соответствует значение 100. Такая гистограмма описывает распределение вероятностей уровней сигналов фонограммы относительно максимального значения. При ее построении автоматически подбирается масштаб по оси X, поэтому затруднительно сравнивать гистограммы.

При измерении гистограмм распределения уровней аналоговых фонограмм используются квазипиковые детекторы, тогда как в случае цифровых фонограмм определяется *эффе́ктивное на́пряжение* с временем интеграции от 1 мс до нескольких секунд. Это время никак не регламентируется и выбирается самим исполнителем в зависимости от задачи исследования. Если, например, важным является диапазон мгновенных значений уровней фонограммы, то время интеграции должно быть менее 5 с. Если измеряется динамический диапазон музыки с учетом слухового восприятия, то время интеграции выбирается равным 60 мс, это постоянная времени слуха.

Практическое применение. Кому и зачем нужна статистическая информация и гистограмма фонограммы. Эти данные прежде всего оказывают неоценимую помощь при динамической обработке фонограммы, так как они позволяют обоснованно выбрать характеристики компрессора и экспандера. Статистические результаты обработки фонограмм с музыкой различных жанров позволяют определить необходимый динамический диапазон электроакустического тракта, сформировать требования по пиковой и средней мощности головок акустических систем. Они играют существенную роль при разработке алгоритмов компрессии звуковых сигналов.


5. Порядок выполнения лабораторной работы

5.1. Подготовка к измерениям


- Скопируйте с рабочего стола файл «Отчет LR-3 АТ» на свою флешкарту, дайте ему свое имя в виде: «Группа»_ «Фамилия (работавшего на клавиатуре)»_ «Номер ЛР»_ «Дата».
- Откройте этот файл и сверните его окно.

В файл этого отчета вы будете копировать все экспериментальные графики, и заносить в таблицы отчета результаты измерений сразу в процессе их получения. В день выполнения лабораторной работы отчет в электронном виде необходимо сдать преподавателю, только тогда работа считается выполненной.



- Кнопкой с **пиктограммой**  **откройте** окно программы **Audition 3**. Не меняйте размеры этого окна, они согласованы с Вашим отчетом



- Кнопкой с пиктограммой  **откройте** и затем сверните окно программы **HyperSnap-DX**, предназначенной для вырезки и копирования графических изображений

5.2. Измерение динамических характеристик фонограмм

5.2.1. Кнопкой **File** откройте меню тестовых музыкальных сигналов и выберите в нем тестовую музыку: Тест 1. Вагнер. « Полет Валькирий ». По команде **Open** в окне осциллографа появится сигналограмма данного музыкального фрагмента.

- Нажмите кнопку **Analyze**. В выпадающем меню выберите **Statistics**. По этой команде на экране появится окно **Waveform Statistics – General** со статистическими характеристиками фонограммы
- В окне **Window Width** введите время интегрирования 1 мс и нажмите кнопку **Recalculate**
- Результаты измерений L_{pic} , L_{min} и L_{avr} занесите в табл.3.2.

Табл.3.2. Динамические характеристики фонограмм при времени интегрирования 1 мс						
Номер теста	L_{pic}, dB	L_{min}, dB	DR_m, dB	DR_g, dB	L_{avr}, dB	PF_m, dB
1						
2						
3						

- По формулам

$$DR_m = L_{pic} - L_{min}$$

$$PF_m = L_{pic} - L_{avr}$$

рассчитайте динамический диапазон и пик-фактор фонограммы

- Нажмите кнопку **Histogram** - в окне появится гистограмма распределения мгновенных уровней фонограммы. Скопируйте этот график (рис.3.4) в Ваш отчет.

Для этого в окне программы **HyperSnap-DX** нажмите кнопку **Capture** и в открывшемся окне выберите **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу окна гистограммы и нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол окна гистограммы и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** нажмите кнопку **File** и в открывшемся меню выберите **Save As**. Установите путь копирования на Вашу флешку, дайте название файлу и сохраните его в своем отчете.

- Определите динамический диапазон DR_g по гистограмме, результат занесите в табл.3.2.

5.2.2. Кнопкой **File** откройте меню тестовых музыкальных сигналов и выберите в нем тестовую музыку: Тест 2. Бетховен. « К Элизе ». По команде **Open** в окне осциллографа появится сигналограмма данного музыкального фрагмента.

- Повторите все измерения по п.5.2.1.

5.2.3. Кнопкой **File** откройте меню тестовых музыкальных сигналов и выберите в нем тестовую музыку: Тест 3. Марина Ладынина. Каким ты был... (Песня из к/ф «Кубанские казаки». По команде **Open** в окне осциллографа появится сигналограмма данного музыкального фрагмента.

- Повторите все измерения по п.5.2.1. Скопируйте гистограммы (рис.3.5) в Ваш отчет

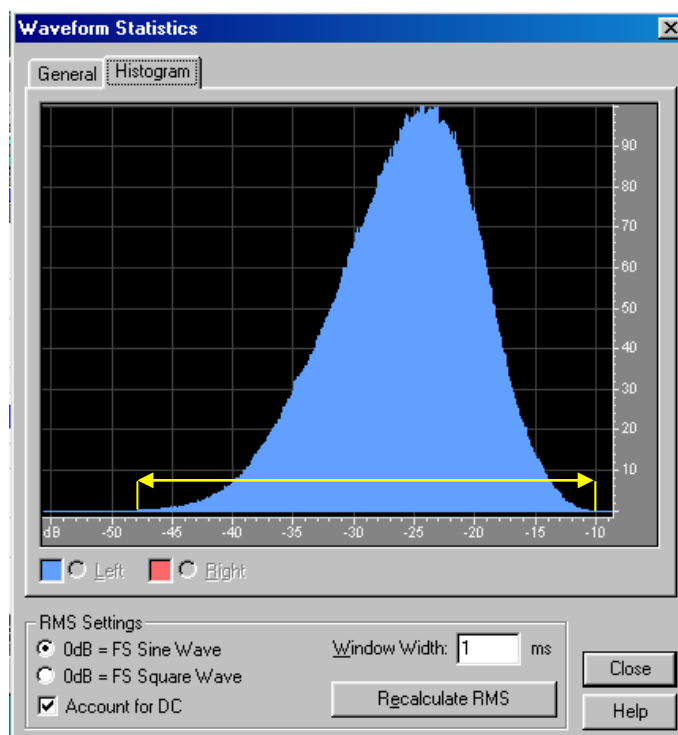


Рис.3.4. Гистограмма. Тест 1. (Вагнер «Полет Валькирий»)

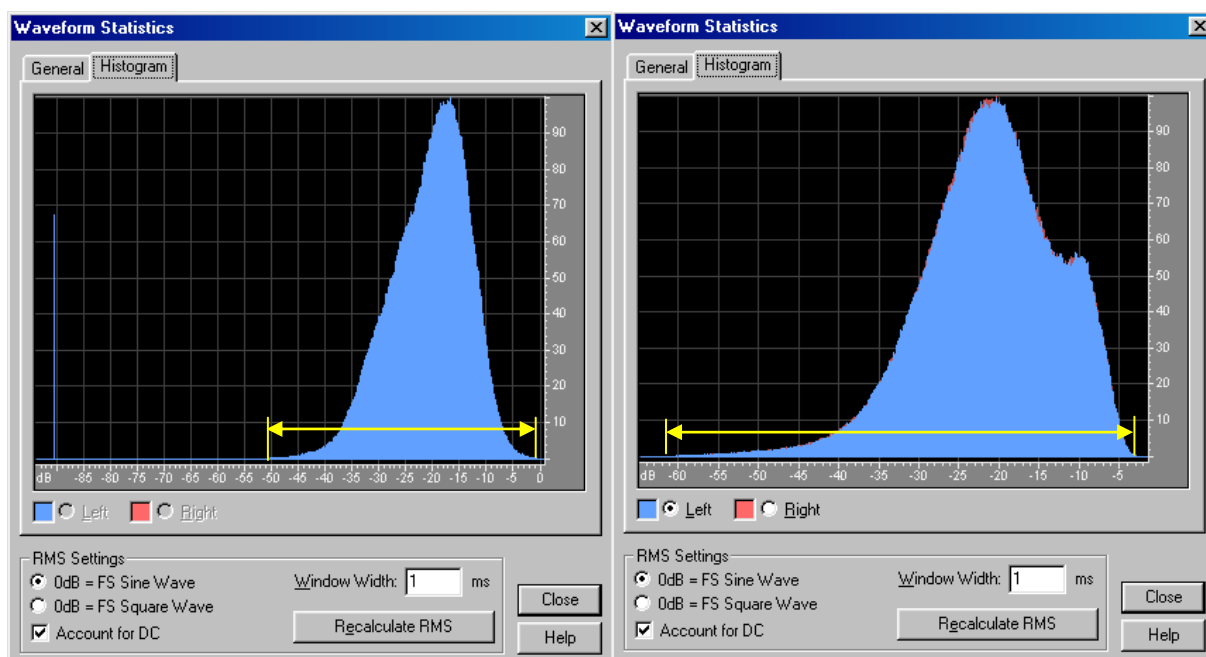


Рис.3.5. К определению динамического диапазона фонограмм по гистограммам (Тест 1 слева, Тест 2 справа)

5.3. Измерение динамических характеристик музыкальных фрагментов

5.3.1. Кнопкой **File** откройте меню тестовых музыкальных сигналов и выберите в нем тестовую музыку: Тест 1. Вагнер. «Полет Валькирий». По команде **Open** в окне осциллографа появится сигналограмма данного музыкального фрагмента.

- Нажмите кнопку **Analyze**. В выпадающем меню выберите **Statistics**. По этой команде на экране появится окно **Waveform Statistics – General** со статистическими характеристиками фонограммы

- В окне **Window Width** введите время интегрирования 60 мс и нажмите кнопку **Recalculate**
- Результаты измерений L_{\max} , L_{\min} и L_{avr} занесите в табл.3.3.
- По формулам

$$DR_p = L_{\max} - L_{\min}$$

$$PF_p = L_{\max} - L_{avr}$$

рассчитайте динамический диапазон и пик-фактор фонограммы

Табл.3.3. Динамические характеристики фонограмм при времени интегрирования 60 мс						
Номер теста	L_{\max}, dB	L_{\min}, dB	DR_p, dB	DR_g, dB	L_{avr}, dB	PF_p, dB
1						
2						
3						

- Нажмите кнопку **Histogram** - в окне появится гистограмма распределения уровней фонограммы. Скопируйте этот график (рис.3.6) в Ваш отчет.

Для этого в окне программы **HyperSnap-DX** нажмите кнопку **Capture** и в открывшемся окне выберите **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу окна гистограммы и нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол окна гистограммы и снова нажмите левую клавишу мышки (

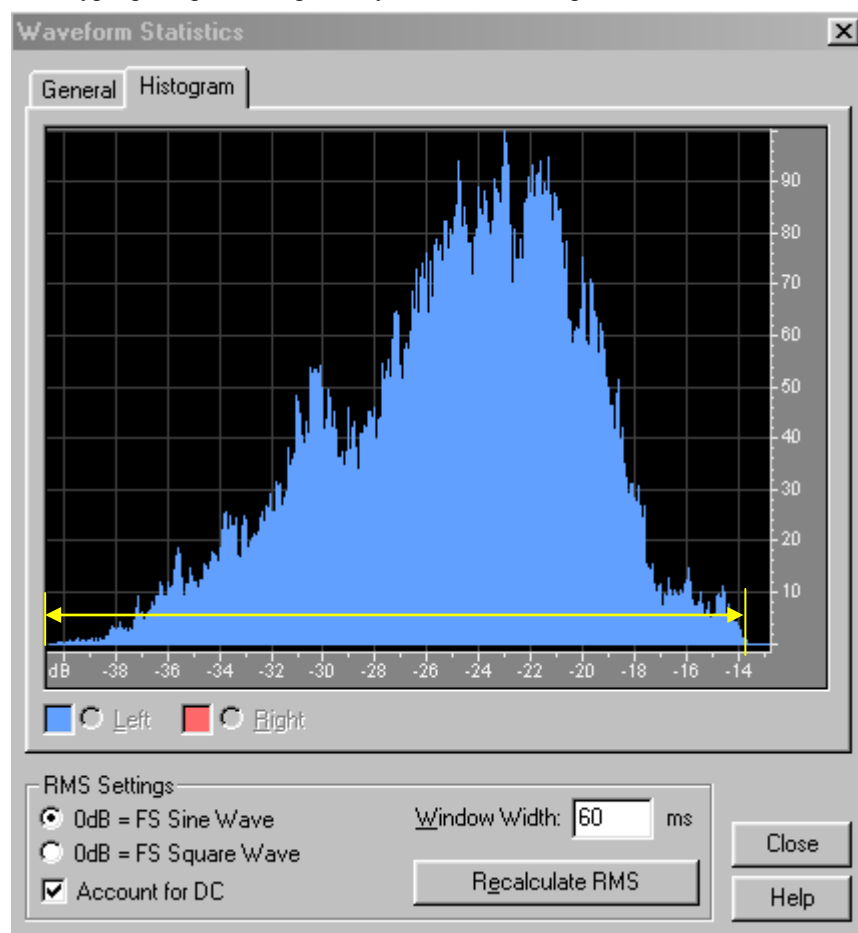


Рис.3.6. К определению динамического диапазона музыки (Тест 1)

этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** нажмите кнопку **File** и в открывшемся меню выберите **Save As**. Установите путь копирования на Вашу флешкарту, дайте название файлу и сохраните его в своем отчете.

- Определите динамический диапазон DR_g по гистограмме, результат занесите в табл.3.3.

- 5.3.2. Кнопкой **File** откройте меню тестовых музыкальных сигналов и выберите в нем тестовую музыку: Тест 2. Бетховен. «К Элизе». По команде **Open** в окне осциллографа появится сигналограмма данного музыкального фрагмента.
- Повторите все измерения по п.5.3.1.
- 5.3.3. Кнопкой **File** откройте меню тестовых музыкальных сигналов и выберите в нем тестовую музыку: Тест 3. Марина Ладыгина. Каким ты был... (Песня из к/ф «Кубанские казаки»). По команде **Open** в окне осциллографа появится сигналограмма данного музыкального фрагмента.
- Повторите все измерения по п.5.3.1. Скопируйте гистограммы (рис3.7) в Ваш отчет

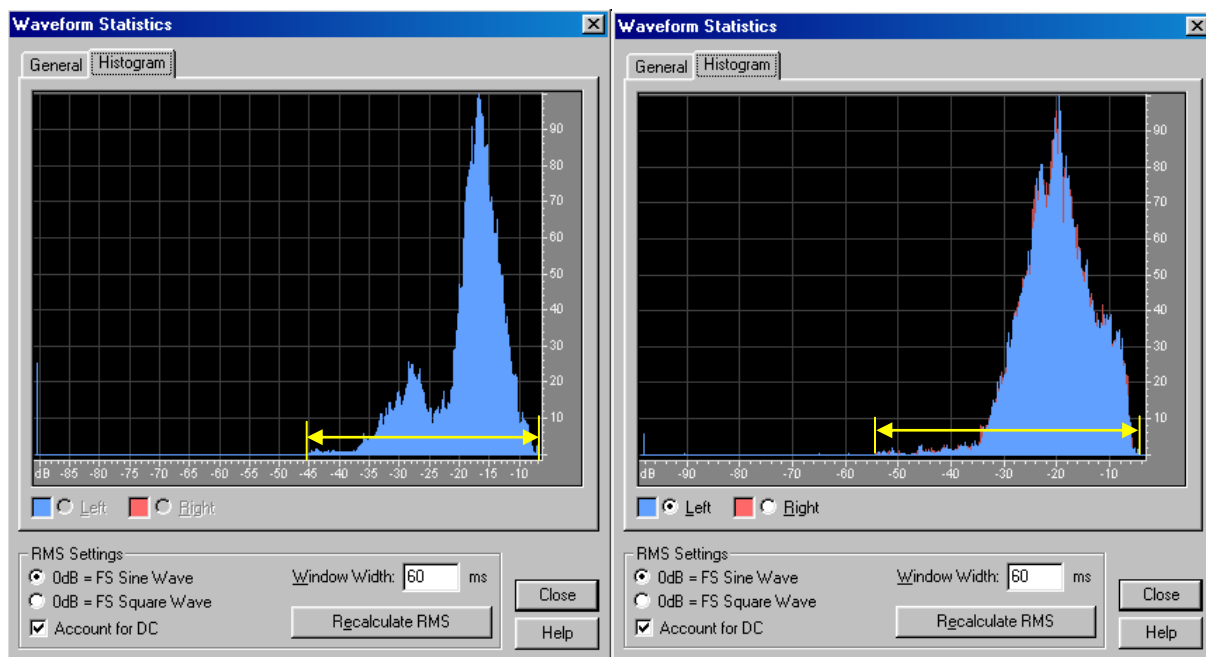


Рис.3.7. К определению динамического диапазона музыки (Тест 2 слева, Тест 3 справа)

6. Контрольные вопросы

12. Как отличается восприятие музыки с широким и узким динамическим диапазоном
13. Какую роль играет пик-фактор музыки на ее восприятие
14. Что понимается под термином «динамический диапазон фонограммы»
15. Как определяется динамический диапазон фонограммы через ее статистические характеристики
16. Что понимается под термином «пик-фактор фонограммы»
17. Как измеряется пик-фактор фонограммы
18. В чем проявляются отличия динамического диапазона фонограммы и динамического диапазона музыки, записанной на этой же фонограмме
19. Что характеризует гистограмма фонограммы
20. Как рассчитывается гистограмма

7. Содержание отчета

- 7.1. Титульный лист
- 7.2. Цель работы
- 7.3. Задание к лабораторной работе
- 7.4. Функциональная схема измерений
- 7.5. Результаты измерений в виде графиков

7.6. Таблицы с результатами измерений и расчетов

7.7. Выводы по результатам измерений

8. Литература

Вологдин Э.И. Методы и алгоритмы обработки звуковых сигналов. Учебное пособие. Часть 1 и 2. СПб ГУТ. 2009.

ЛАБОРАТОРНАЯ РАБОТА N 4 АТ

ИССЛЕДОВАНИЕ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК ЦИФРОВЫХ ФОНОГРАММ

Руководство к выполнению лабораторной работы по Аудиотехнике

4. Цель работы

Целью лабораторной работы является ознакомление студентов с методиками измерения спектральных характеристик цифровых фонограмм с использованием компьютерной программы SpectraLab.

5. Задание к лабораторной работе

- Ознакомиться с программой SpectraLab в объеме, необходимом для измерения текущего спектра, 3-х мерного мгновенного спектра и спектрограмм цифровых фонограмм.
- Измерить текущие спектры двух фонограмм различного музыкального жанра
- Измерить 3-х мерные спектры двух фонограмм различного музыкального жанра
- Измерить спектрограммы двух фонограмм различного музыкального жанра
- Провести сравнительный анализ и по результатам измерений сформулировать выводы

3. Функциональная схема измерений

В приведенной на **рис.4.1.** схеме цифровая фонограмма, спектральные характеристики которой необходимо исследовать, хранится на жестком диске компьютера. При ее воспроизведении цифровая информация передается в оперативную память осциллографа и отражается на его дисплее. По отдельным командам с помощью спектроанализатора вычисляются и отражаются на отдельных дисплеях текущий спектр, мгновенный 3D спектр (3-х мерный) и мгновенный двумерный спектр в виде спектрограммы.

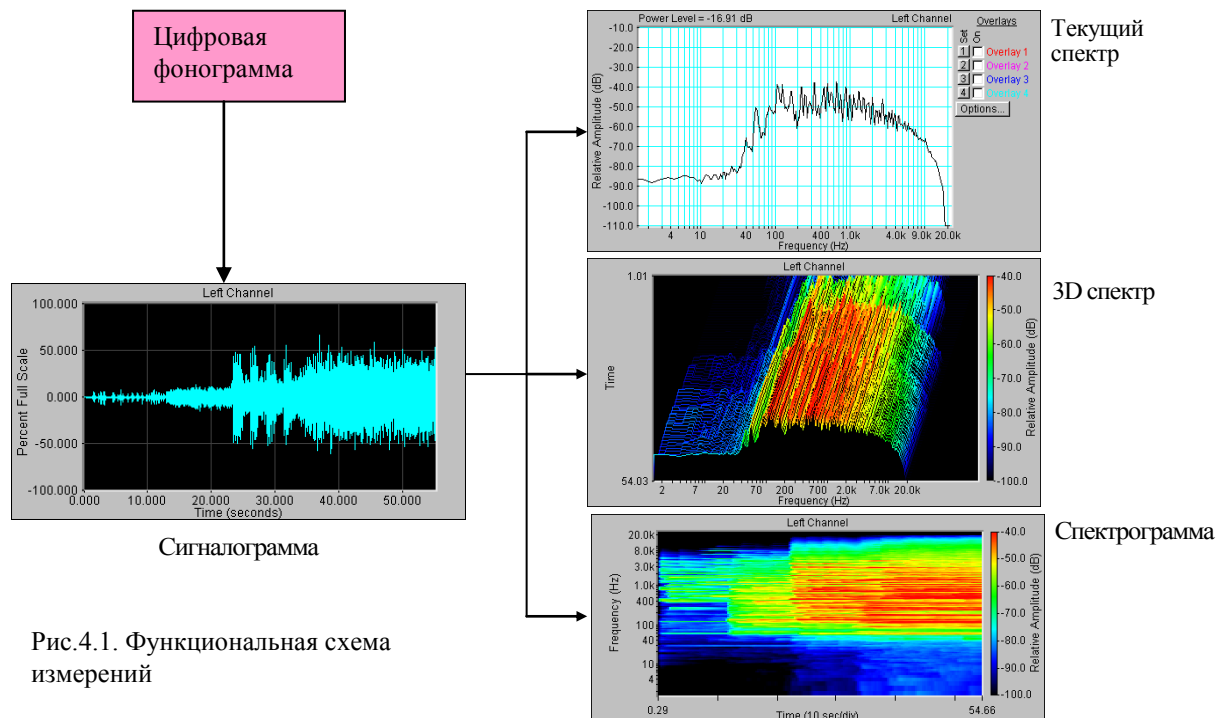


Рис.4.1. Функциональная схема измерений

4. Методические указания

4.1. Спектральная плотность и спектр случайных сигналов. Музыкальные сигналы относятся к группе нестационарных случайных процессов, однако, при длительности больше 2...3 мин с некоторым приближением они могут рассматриваться как стационарные. Такие сигналы описываются парой преобразования Фурье в виде функции времени t

$$f(t) = \int_{-\infty}^{\infty} S(\omega) \cdot e^{j\omega t} d\omega \quad (4-1)$$

или в виде функции круговой частоты ω , называемой комплексной спектральной плотностью

$$S(\omega) = \int_{-\infty}^{\infty} f(t) \cdot e^{-j\omega t} dt \quad (4-2)$$

Частотная зависимость модуля спектральной плотности $|S(\omega)|$ называется **амплитудным спектром**. Вычисления по формулам (4-1) и (4-2) называют соответственно прямым и обратным преобразованиями Фурье. Если известна функция $S(\omega)$, то по ней можно рассчитать функцию времени, и наоборот.

Смысл формулы (4-1) состоит в том, что функция времени $f(t)$ представляется суммой синусоидальных составляющих, у которых комплексная амплитуда dA и частотный интервал между ними $d\omega$ чрезвычайно малы. Спектральная плотность и амплитуды этих составляющих связаны равенством

$$S(\omega) = \pi \cdot \frac{dA}{d\omega} \quad (4-3)$$

Текущий спектр. Любой звуковой сигнал ограничен по времени от 0 до \bar{T} , и поэтому определение спектра принимает вид

$$S_{\bar{T}}(\omega) = \int_0^{\bar{T}} f(t) \cdot e^{j\omega t} dt \quad (4-4)$$

В этом случае частотная зависимость $|S_{\bar{T}}(\omega)|$, являющаяся функцией не только частоты, но и времени, носит название *текущего спектра*.

Мгновенный спектр. Этот спектр определяется как спектр отрезка процесса длительностью T , непосредственно предшествующего данному моменту. При измерении мгновенного спектра интервал интегрирования имеет постоянную длину, но перемещается по оси времени. Аналитически этот спектр записывается в виде

$$S_T(\omega) = \int_{\bar{T}-T}^T f(t) \cdot e^{j\omega t} dt. \quad (4-5)$$

Энергетический спектр. Случайные стационарные процессы, обладающие свойством эргодичности, более правильно характеризовать *энергетическим спектром* $G(\omega)$. Этот спектр также называют статистическим, так как при его вычислении используется функции вероятности мгновенных значений сигналов. Математически он связан со спектром амплитуд $|S_{\bar{T}}(\omega)|$ равенством

$$G(\omega) = \frac{1}{\pi} M \left(\frac{d}{dt} |S_{\bar{T}}(\omega)|^2 \right), \quad (4-6)$$

где M – символ математического ожидания. В соответствии с этой формулой энергетический спектр случайного процесса определяется как среднеквадратическое значение по множеству от амплитудного спектра.

Спектральная плотность мощности. В соответствии с формулой (6) в практике измерений за единицу спектральной плотности принята мощность широкополосного сигнала в полосе шириной 1 Гц, она имеет размерность Вт/Гц. Эта единица называется спектральной плотностью мощности $\bar{D}(\omega)$. Часто используется и другая характеристика – просто *спектральная плотность* $D(\omega)$, она определяется эффективным значением напряжения широкополосного сигнала в полосе шириной 1 Гц, и имеет размерность В/Гц^{0,5}. Обычно $D(\omega)$ измеряется, а $\bar{D}(\omega)$ – рассчитывается. Очевидно, что числовые значения этих характеристик существенно разные, однако, это не имеет практического значения, так как в аудиотехнике все измерения осуществляются в децибелах, и уровень $\bar{D}(\omega)$ всегда равен уровню $D(\omega)$, также как это имеет место при измерении уровней звукового давления и интенсивности.

Спектральный анализ. Только в своем воображении можно представить, что текущий амплитудный спектр измеряется с помощью набора из 20000 фильтров с шириной полосы 1 Гц и средней частотой от 20 до 20000 Гц, и на выходе каждого такого фильтра включен вольтметр *средних значений* с временем интеграции равным \bar{T} . На практике число фильтров всегда ограничено, и их полоса может быть существенно больше 1 Гц, и тогда измеряется *средний текущий спектр*. Если использовать вольтметры эффективных значений, то будет измеряться частотная зависимость спектральной плотности мощности (*энергетический спектр*).

4.2. Быстрое преобразование Фурье (БПФ). В настоящее время основным инструментом измерения спектра цифровых фонограмм является быстрое преобразованием Фурье, которое позволяет одновременно измерять мгновенный и текущий спектр. Результаты измерений мгновенного спектра хранятся в оперативной

памяти, а текущего - в долговременной. Наиболее эффективно БПФ применяется для вычисления спектра фонограмм, осуществляемого в *режиме пост обработки*, так как при этом не предъявляются высокие требования к быстродействию, как это имеет место при измерении спектра в *реальном времени*.

Модель БПФ упрощенно можно представить в виде набора фильтров с одинаковой полосой пропускания ΔF , которые перекрывают частотный диапазон от 0 до частоты Найквиста. *Ширина полосы пропускания* определяется частотой дискретизации f_s и числом выборок N_{fft} , используемых при анализе

$$\Delta F = \frac{f_s}{N_{fft}} \quad (4-7)$$

Так, при $f_s = 48$ кГц и $N_{fft} = 2048$ $\Delta F = 23,4$ Гц, и она сокращается вдвое при каждом удвоении числа выборок. Число элементов оперативной и долговременной памяти равно гипотетическому числу полосовых фильтров, и рассчитывается оно путем деления частоты Найквиста на ширину полосы пропускания. Чем больше разрешение, тем больше требуется объем памяти. Для визуализации спектра каждый элемент памяти привязывается к средней частоте полосы пропускания фильтра.

Нижняя граница анализируемой полосы спектра F_l связана с шириной полосы пропускания фильтров примерным равенством

$$F_l = 2 \cdot \frac{f_s}{N_{fft}} \quad (4-8)$$

и при $N_{fft} = 2048$ $F_l = 50$ Hz

Время спектрального анализа T блока данных длиной N_{fft} определяется равенством

$$T = \frac{N_{fft}}{f_s} = \frac{1}{\Delta F} \quad (4-9)$$

Для приведенного выше примера оно составляет 42 мс. За время T вычисляется мгновенный спектр фонограммы, относящийся к текущему времени $i \cdot T$.

В процессе измерения мгновенного спектра рассчитывается *эффективное значение напряжения* звукового сигнала в каждой из полос, и результаты фиксируются в оперативной памяти. При этом время интеграции равно времени анализа T .

Длительность фонограммы может быть от нескольких секунд до 5...6 часов. В процессе спектрального анализа с интервалом T будут повторяться измерения мгновенного спектра. Результаты измерения мгновенного спектра из оперативной памяти передаются в долговременную, в которой со временем интеграции, равным длительности фонограммы \bar{T} , вычисляется текущий спектр. Если при измерении мгновенного спектра используется усреднение по времени, то при измерении текущего спектра производится усреднение по множеству.

Погрешность БПФ. Заметим, что результат вычислений БПФ является спектральной плотностью мощности только при условии, что $\Delta F = 1$ Гц. В противном случае необходимо вводить коррекцию. Погрешность измерения уровня спектральной плотности рассчитывается по формуле

$$\Delta L_w = 10 \cdot \log \left\{ \frac{N_{fft}}{K_w \cdot f_s} \right\}, \quad (4-10)$$

где K_w - коэффициент, зависимый от выбора оконной функции. Если, например, используется оконная функция Blackman, то $K_w = 2,004$. В этом случае погрешность измерения близка к нулю при $N_{fft} = 65536$, когда $\Delta F = 0,7$ Гц.

Для иллюстрации разрешающей способности БПФ по частоте на [рис.4.2.](#) приведены измеренные с использованием оконной функции Blackman спектры тестового сигнала, состоящего из 6 одинакового уровня дискретных составляющих от 20 Гц до 8 кГц. Как

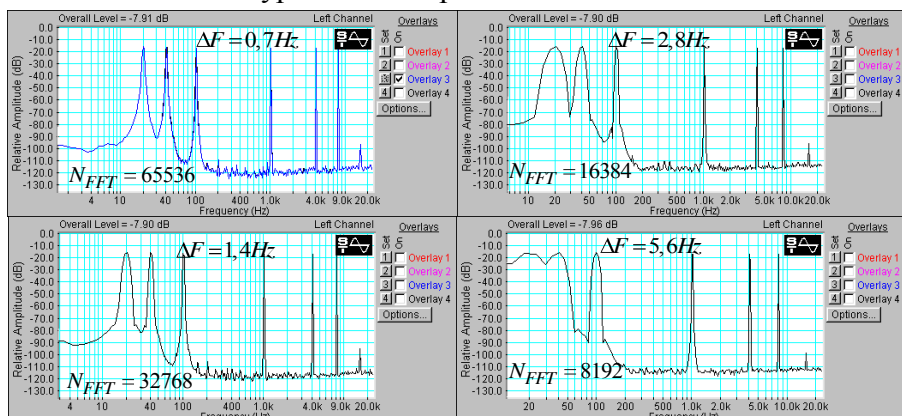


Рис. 4.2. Зависимость разрешающей способности БПФ от N_{fft}

видно, с уменьшением числа N_{fft} быстро уменьшается разрешение в области низких частот, и при $N_{fft} = 8192$ частоты 20 и 40 Гц уже практически не различимы. На этом основании можно считать, что при измерении спектров фонограмм должно быть не меньше 8192 при использовании оконных функций Hanning или Blackman.

4.3. Измерение текущего и мгновенного спектров. Теоретики и практики в области

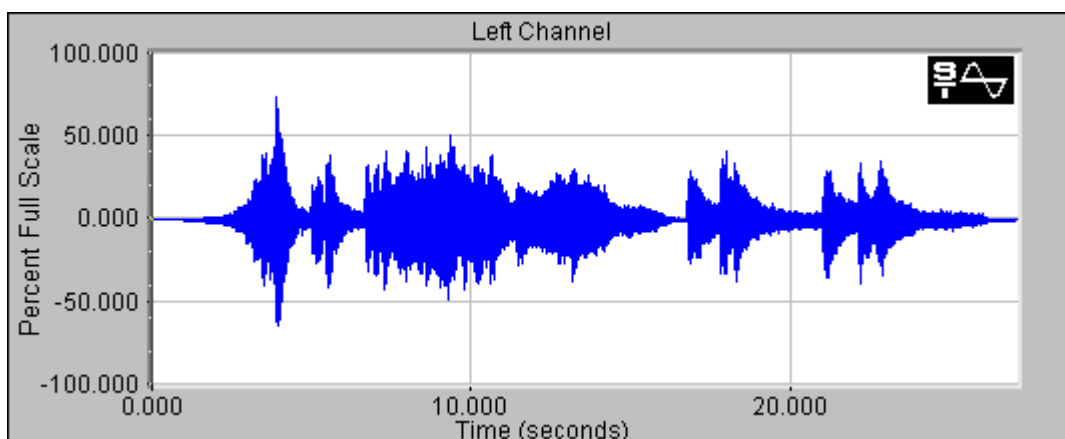


Рис.4.3. Сигналограмма музыкального фрагмента длительность 27 секунд

спектрального анализа не дают рекомендаций по выбору времени интеграции при измерении мгновенного спектра с использованием БПФ, поэтому приходится его обосновывать. Когда речь идет об измерении спектра фонограмм, очевидно, что время интеграции не может быть меньше постоянной времени слуха, равной примерно 60 мс, так как иначе теряется практический смысл понятия спектра. Восприятие звука и его спектра связано также с памятью человека, поэтому можно допустить время анализа от 90 мс до 1...2 с.

На **рис.4. 4.** приводится трехмерная зависимость мгновенного спектра фонограммы от разрешающей способности БПФ (частота, время и уровень в децибелах). Сигналограмма этого фрагмента приведена на **рис.4. 3.** Из **рис.4.4** видно, что с уменьшением

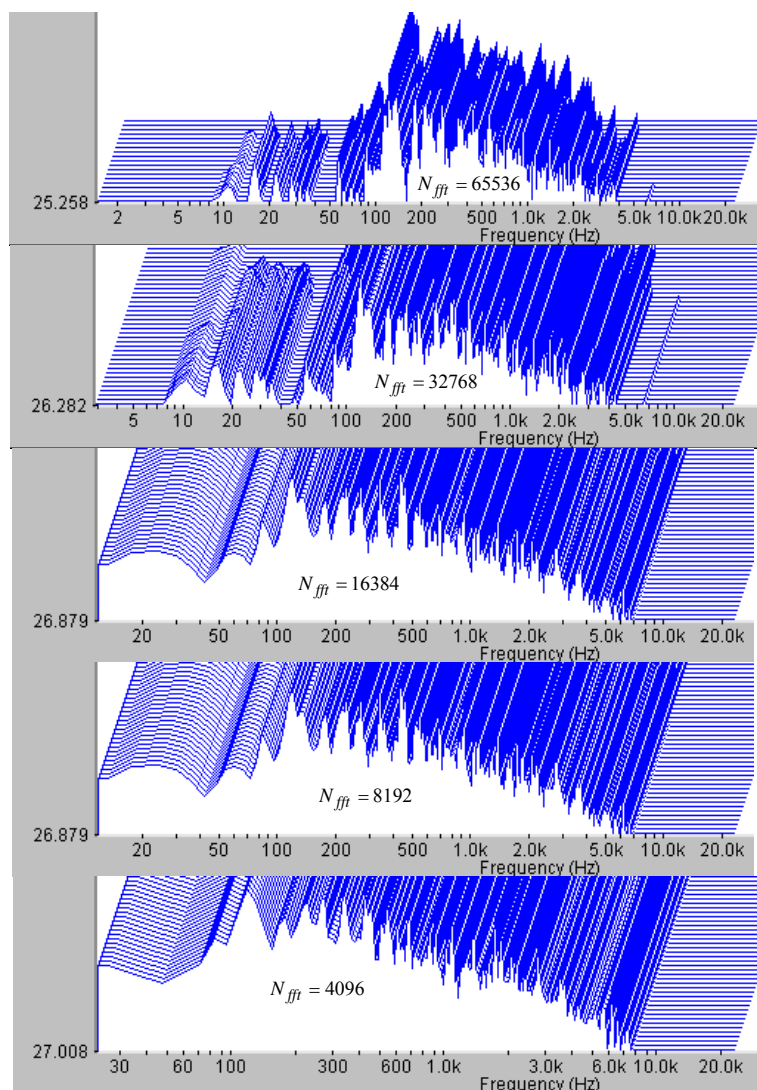


Рис. 4.4 Трехмерные мгновенные спектры музыкального фрагмента длительностью 27 с для разных значений N_{fft}

разрешающей способности БПФ увеличиваются уровни всех составляющих спектра и расширяется измеренный спектр как в область низких, так и высоких частот. И то и другое объясняется тем, что с уменьшением N_{fft} увеличивается полоса частот ΔF , в которой вычисляется спектр. При этом результаты измерения спектральной плотности необходимо корректировать в сторону уменьшения, так как полоса, в которой она измеряется, существенно больше 1 Гц.

Чем шире полоса измерения, тем в большем частотном диапазоне усредняются составляющие спектра. Таким образом, чем уже полоса ΔF , тем ближе к реальному измеренный спектр, но, естественно, время анализа должно быть существенно меньше длительности музыкального фрагмента.

Эти соображения подтверждаются **рис.4.5**, на котором приведены текущие спектры того же музыкального фрагмента, измеренные для двух значений N_{fft} . Рассчитываются они как усреднение по множеству мгновенных спектров, что справедливо для эргодических случайных процессов.

4.4. Практическое применение. Кому и зачем нужна информация о мгновенных и текущих спектрах музыки различных жанров. Прежде всего, это необходимо для проектирования электроакустических трактов, в частности, для обоснования

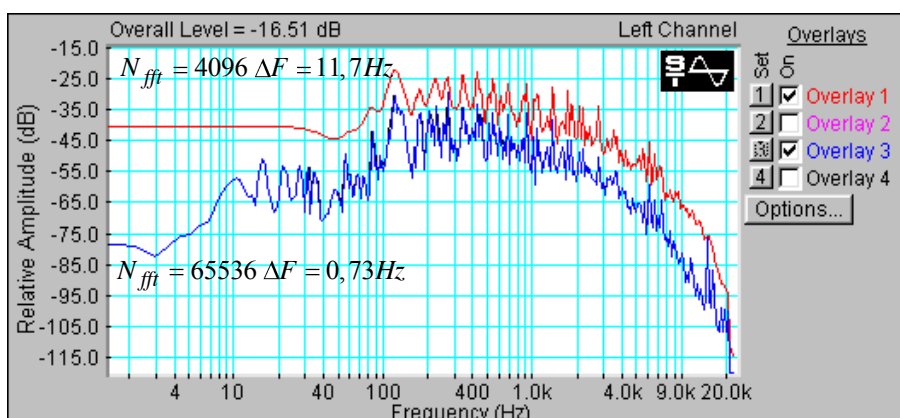


Рис.4.5 Текущий спектр музыкального фрагмента

необходимой полосы пропускания и требуемой мощности в различных участках звукового диапазона, например, в кроссоверах. Мгновенным спектром очень интересуются звукорежиссеры. Спектры музыки детально изучаются при создании систем компрессирования звука. Некоторые результаты таких исследований приведены в

Табл.4.1. Частотные диапазоны музыки

Источник звука	Частотный диапазон, Гц
Музыка – основной тон	100...3400
Музыка - обертона	До 10000...20000
Большой симфонический оркестр	31...15000
Малый симфонический оркестр	31... 15000
Духовой оркестр	50...10000
Эстрадный оркестр	25...16000
Голос (основной тон)	70...450
Сопрано	262...4318
Контральто	175...784
Тенор	321...587
Баритон	110...440
Бас	82...398

табл.4.1.







Заметим, что измерение спектра цифровых фонограмм производится в шкале FS, поэтому графики спектров одних и тех музыкальных произведений могут существенно отличаться от приводимых в учебной литературе, так как они измерялись в шкалах dBu.

4.5. Сигналограмма, текущий спектр, 3-х мерный спектр и спектрограмма. Для пояснения разницы между этими спектрами все они приведены на одном рис.4.6. Сигналограмма иллюстрирует как меняется мгновенный уровень сигнала в функции времени. В каждый момент времени этот уровень определяется всеми музыкальными инструментами. Текущий спектр дает усредненное распределение уровней по частоте. Как видно, если уровни сигналов фонограммы достигают -6 дБ, то уровни текущего спектра не превышают - 40 дБ. Текущий спектр позволяет определить частотный диапазон фонограммы. 3D спектр представляет собой 3-мерный график (время, частота, уровень). Он определяет зависимость мгновенного спектра от времени. На этом графике цвет несет дополнительную информацию об уровне спектральных составляющих (графики могут быть одного цвета). Спектрограмма – это тоже график 3-х мерного мгновенного спектра в одной плоскости, в котором уровень спектральных составляющих индицируется только

цветом. Графики мгновенного спектра иллюстрируют зависимость ширины спектра и уровня спектральных составляющих от времени.

5. Порядок выполнения лабораторной работы

5.1. Функции кнопок управления:

- кнопкой с пиктограммой  осуществляется автоматическая установка пределов измерения по вертикальной оси шкалы анализатора спектра, фазометра или осциллографа
- кнопкой с пиктограммой  устанавливается частотный диапазон шкалы от 1 Гц до частоты Найквиста
- кнопкой с пиктограммой  вызывается маркер, которым устанавливается желаемый частотный диапазон шкалы
- при нажатии кнопки с пиктограммой  графики рисуются тонкими линиями
- при нажатии кнопки с пиктограммой  графики рисуются вертикальными линиями или столбиками
- кнопкой с пиктограммой  открывается меню анализатора спектра, фазометра или осциллографа, в котором вводятся необходимые опции.
- кнопкой **Run** включаются все измерительные приборы
- кнопкой **Stop** выключаются все измерительные приборы и запоминаются результаты измерений
- в меню **Avg** задается число усредняемых результатов измерения
- в окошке **Peak Hold** курсором включается функция удержания пиковых значений напряжения



5.2. Подготовка к измерениям

Конфигурация схемы измерений и ввод необходимых опций производится с помощью загрузочного файла **LR-4 AT.cfg**. На экран монитора вызываются окна: программа **SpectraLab (SL)**, анализатор спектра (Spectrum), осциллограф (Time Series), 3-х мерный анализатор спектра (3D Surface) и измеритель спектрограммы (Spectrogram).

В лабораторной работе измеряются текущий спектр, 3-х мерный мгновенный спектр и мгновенная спектрограмма в частотном диапазоне от 2 до 20000 кГц.

- Включите компьютер
- Скопируйте с рабочего стола файл «Отчет LR-4 AT» на свою флешкарту, дайте ему свое имя в виде: «Группа»_ «Фамилия (работавшего на клавиатуре)»_ «Номер ЛР»_ «Дата».
- Откройте этот файл и сверните его окно.

В файл этого отчета вы будете копировать все экспериментальные графики, и заносить в таблицы отчета результаты измерений сразу в процессе их получения. В день выполнения лабораторной работы отчет в электронном виде необходимо сдать преподавателю, только тогда работа считается выполненной.

- Кнопкой с пиктограммой  откройте окно программы **SpectraLab**. Не меняйте размеры этого окна, они согласованы с Вашим отчетом
- Кнопкой с пиктограммой  откройте и затем сверните окно программы **Paint Brush**
- Загрузите файл **LR-4 AT .cfg**, который конфигурирует схему измерений и устанавливает все необходимые опции для выполнения настоящего раздела лабораторной работы.

Для этого в меню **Config** выберите опцию **Load Configuration** и в выпадающем меню выберите файл **LR-4 AT. cfg**.

5.3. Измерение спектральных характеристик фонограммы «Бетховен. К Элизе»

- В меню **File** выберите и откройте музыкальный тест 1 «Бетховен. К Элизе»
- Нажмите кнопку **Run**. По этой команде будут вычислены графики текущего и

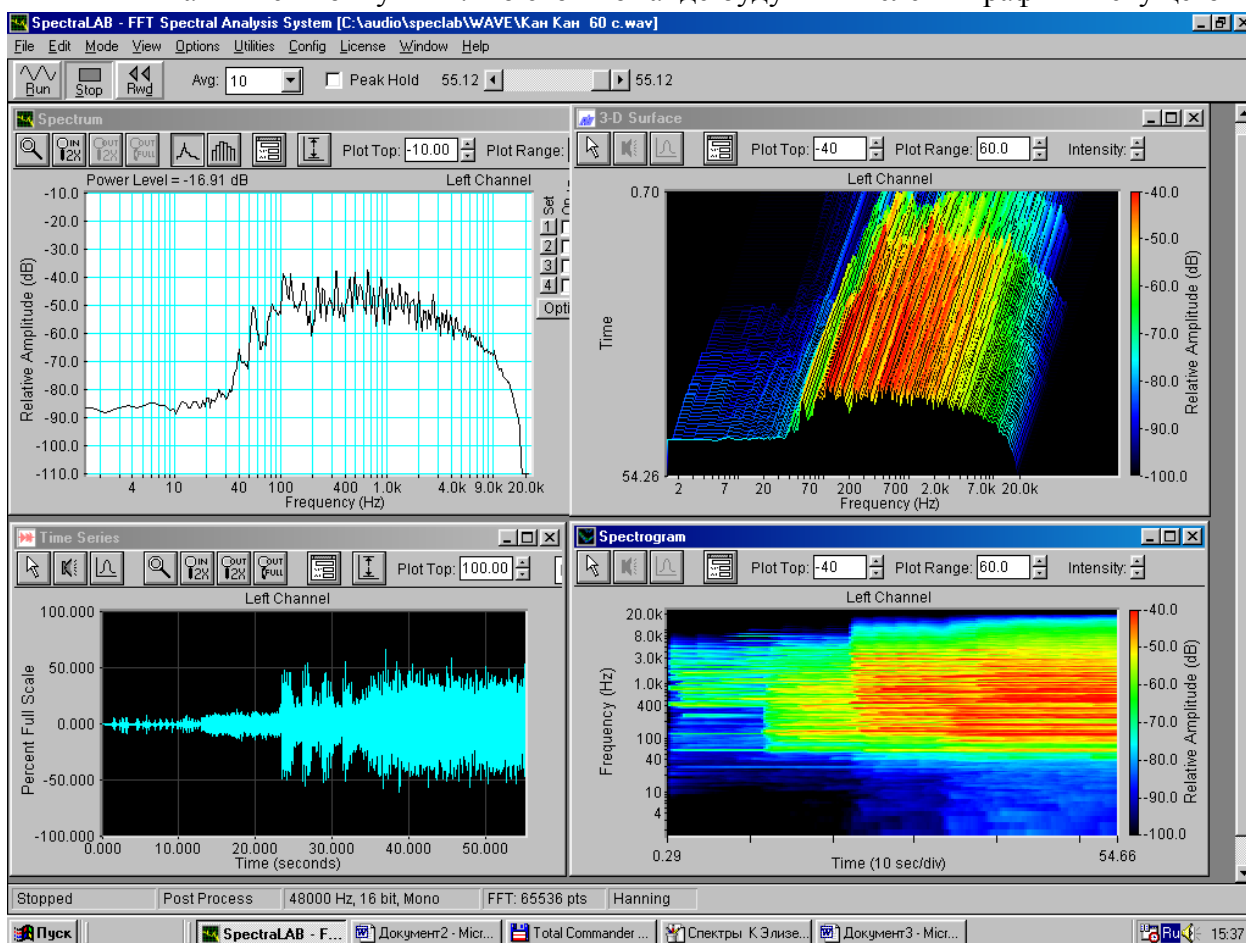


Рис.4.6. Сигналограмма, текущий спектр, 3D спектр и спектрограмма цифровой фонограммы

3D спектров.

- Дайте команду для вычисления спектрограммы

Для этого в окне **Time Series** курсором нажмите кнопку с пиктограммой в виде наклонной стрелки – появится маркер в виде прямоугольника со знаком +. Нажав левую клавишу мышки, этим маркером выделите всю сигналограмму она потемнеет. Направьте курсор на окно **Time Series** и нажмите правую клавишу мышки. В появившемся меню выберите **Compute and Display Spectrogram** и нажмите левую кнопку мышки

- Скопируйте все графики с дисплея монитора в Ваш отчет

Для этого на клавиатуре нажмите кнопку **PrtSc** и откройте курсором окно программы **Pbrush**. Командами **Правка-Вставка-Копировать файл** сохраните графики в Вашем отчете (рис.4.6)

5.4. Измерение спектральных характеристик фонограммы «Каким ты был.. (песня из к/ф Кубанские казаки)»

- В меню **File** выберите и откройте музыкальный тест 2 «Каким ты был.. (песня из к/ф Кубанские казаки)»
- Повторите все операции по п. 5.3. Скопируйте все графики с дисплея монитора в Ваш отчет
- Проведите сравнительный анализ спектральных характеристик двух фонограмм, сформулируйте выводы

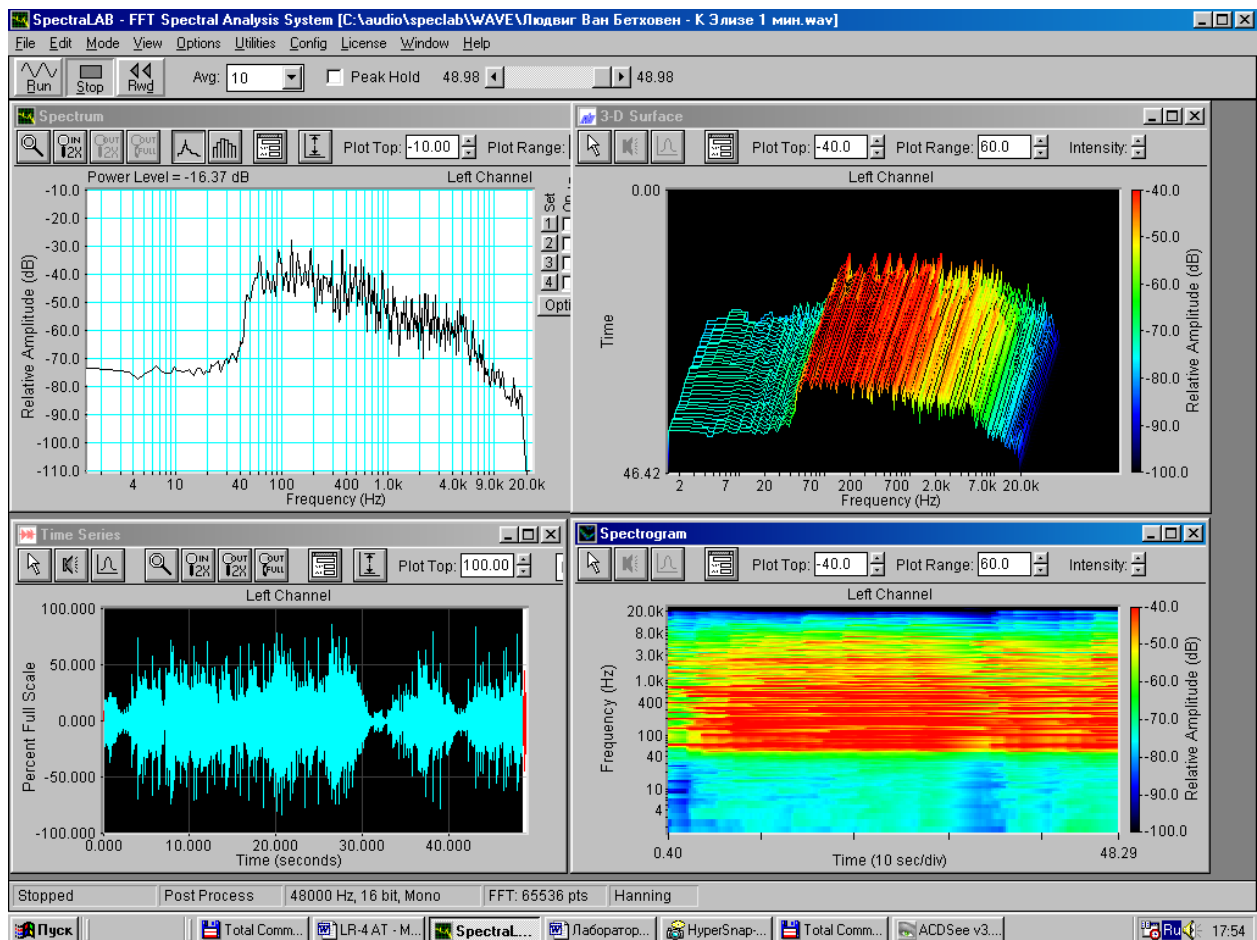


Рис. 4.7. Спектральные характеристики фонограммы «Бетховен. К Элизе»

6. Контрольные вопросы

1. Дайте определение понятию «текущий спектр фонограммы»
2. Каким параметром задается время измерения мгновенного спектра
3. Что понимается под термином «мгновенный спектр»
4. Поясните понятия «прямое и обратное преобразования Фурье»
5. Что характеризует энергетический спектр
6. Как связаны разрешающая способность и длительность спектрального анализа БПФ
7. С какой целью измеряется текущий спектр фонограммы
8. С какой целью измеряется мгновенный спектр фонограммы
9. Как на графиках 3D спектра отражаются время, частота и уровень спектральных составляющих
10. Как на графиках спектрограмм отражаются время, частота и уровень спектральных составляющих
11. Чем отличаются графики 3D спектра и спектрограммы

7. Содержание отчета

- 7.1. Титульный лист
- 7.2. Цель работы
- 7.3. Задание к лабораторной работе
- 7.4. Функциональная схема измерений
- 7.5. Результаты измерений в виде графиков
- 7.6. Таблицы с результатами измерений и расчетов
- 7.7. Выводы по результатам измерений

8. Литература

Вологдин Э.И. Методы и алгоритмы обработки звуковых сигналов. Учебное пособие. Часть 1 и 2. СПб ГУТ. 2009.

ЛАБОРАТОРНАЯ РАБОТА N 5 АТ

ИССЛЕДОВАНИЕ ДИНАМИЧЕСКОЙ ОБРАБОТКИ ЦИФРОВЫХ ФОНОГРАММ

Руководство к выполнению лабораторной работы по Аудиотехнике

Цель работы

Целью лабораторной работы является ознакомление студентов с методиками компрессирования музыкальных цифровых фонограмм с помощью динамического графического компрессора и нормализатора с использованием компьютерной программы Audition

2. Задание к лабораторной работе

- Ознакомиться с программой Audition в объеме, необходимом для компрессирования музыкальных фонограмм с помощью динамического компрессора
- Ознакомиться с программой Audition в объеме, необходимом для компрессирования музыкальных фонограмм с использованием нормализатора
- Скомпрессировать фонограмму с музыкой Бетховена «Элизе» с несколькими степенями компрессирования:
- Исследовать зависимость динамического диапазона и пик-фактора фонограммы от степени компрессирования
- С помощью телефонов на слух определить какая фонограмма звучит лучше
- По результатам измерений сформулировать выводы

6. Функциональная схема измерений

В приведенной на **рис.5.1** схеме тестовый музыкальный сигнал хранится в памяти компьютера. К схеме измерения динамического диапазона фонограммы и ее пик-фактора он может подключаться непосредственно или через динамический процессор программы Audition. Исходный файл или компрессированный подаются на цифровой осциллограф с авметью. Затем по команде рассчитываются статистические характеристики фонограммы, из которых основными являются: Peak Amplitude (L_{pic}), (Maximum RMS Power (L_{max}), Minimum RMS Power (L_{min}) и Average RMS Power (L_{avr}) (уровни

максимального, минимального и среднего среднеквадратического (эффективного) значения мощности сигнала).

Динамический диапазон фонограммы определяется как

$$DR_m = L_{pic} - L_{min},$$

пик-фактор рассчитывается по формуле

$$PF_m = L_{pic} - L_{avr}$$

По результатам статистического анализа строится график гистограммы, про которому можно также определить динамический диапазон фонограммы. Результаты измерений

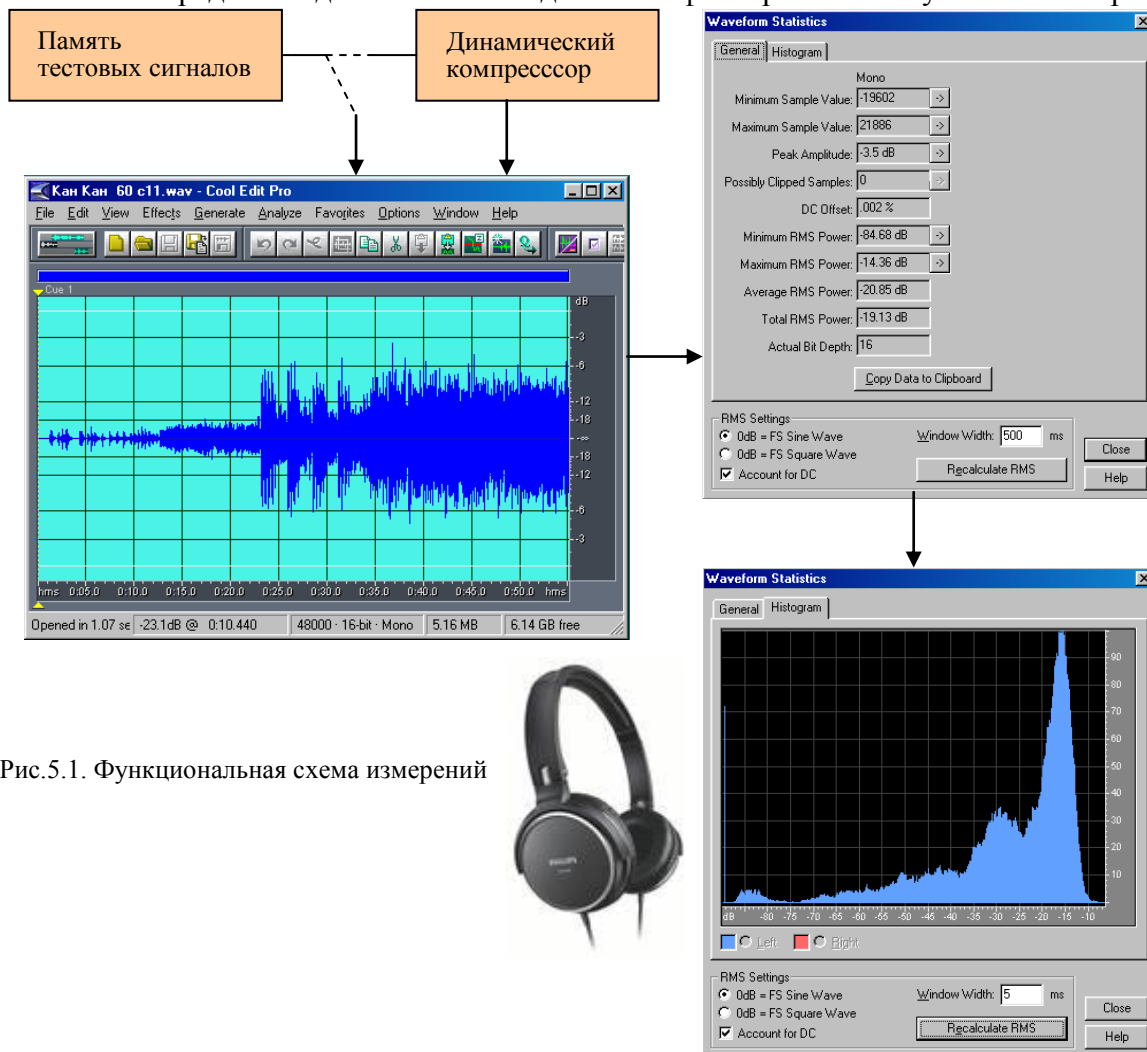


Рис.5.1. Функциональная схема измерений

могут существенно отличаться из-за неточностей определения на взгляд начала и конца динамического диапазона.

4. Методические указания

4.1. Динамическая обработка. Общие сведения

Динамическая обработка предназначена для сокращения динамического диапазона звуковых сигналов. Звуковые сигналы в радиовещании, телевидении и звукозаписи всегда подвергаются такой обработке, независимо аналоговые они или цифровые. Это связано с тем, что часто динамический диапазон природных звуков, звуков музыки и речи значительно шире динамического диапазона электроакустических трактов современной

аппаратуры. Если не использовать динамическую обработку и передавать полностью натуральный динамический диапазон оркестра (70...80 дБ), то чтобы не мешали шумы помещения, уровень сигнала пианиссимо должен быть не ниже 40 дБ SPL. Фортиссимо при этом будет достигаться при уровне звукового давления 120 дБ SPL, что близко к болевому порогу.

Динамическая обработка звуковых сигналов позволяет слушать радио и телевизионные передачи со звуковым давлением не выше 100 дБ SPL в обычных жилых помещениях, в которых обычно уровень шума не менее 40...50 SPL и, даже в автомобилях, где уровень шума значительно больше. Без такой обработки при пиковом звуковом давлении не выше 80 дБ SPL (как в театре) звукозаписи можно было бы слушать только в полной тишине, закрыв окна и двери, иначе любой посторонний шум заглушит тихие места фонограммы.

Динамическая обработка звуковых сигналов производится с помощью *лимитеров, максимайзеров, компрессоров, экспандеров и гейтов*. Это все пороговые устройства, в которых при достижении сигнала установленного уровня их коэффициент передачи меняется скачком. Таким способом можно как сжать, так и расширить динамический диапазон сигнала. В системах передачи звуковых сигналов по линиям связи сокращение и обратное расширение динамического диапазона производится с помощью компандерной системы компрессирования. На входе линии устанавливается компрессор, а на выходе – экспандер, поэтому в такой системе сигнал компрессирован только в линии связи. Названия приборов такие же, как и при динамической обработке, но принцип их работы совершенно иной, на это нужно обратить внимание.

4.2. Динамический диапазон сигналов и трактов

Человек слышит звук в чрезвычайно широком диапазоне звуковых давлений. Этот диапазон простирается от абсолютного порога слышимости до болевого порога 140 дБ

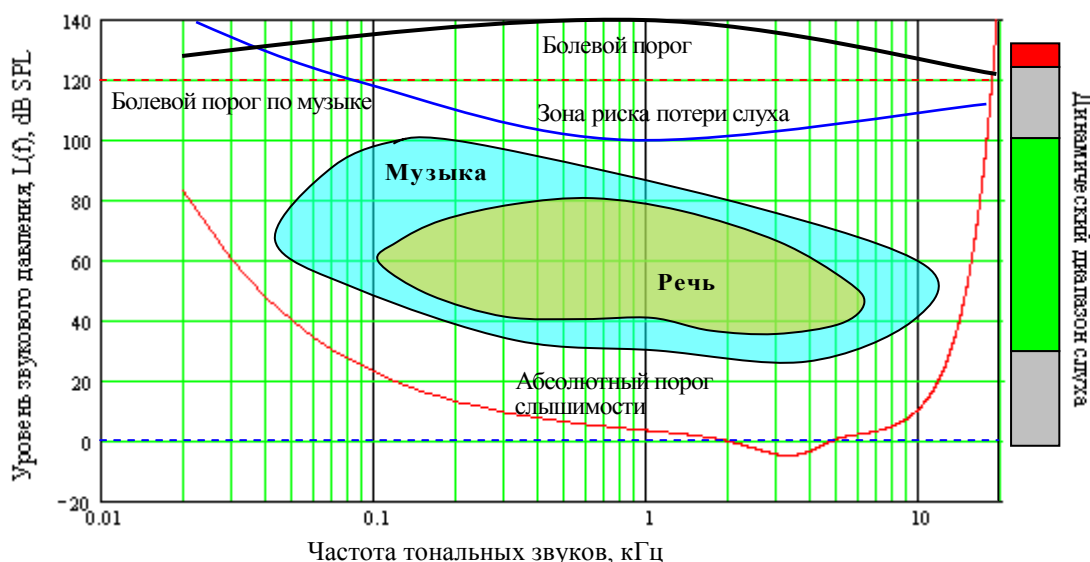


Рис. 5.2. Области слышимости слуха

SPL относительно нулевого уровня, за который принято давление 0,00002 Па (рис.5.2.). Зона риска на этом рисунке обозначает область звуковых давлений, которые при длительном воздействии могут привести к полной потери слуха. Болевой порог для тональных звуков зависит от частоты, для звуков с произвольным спектром за болевой порог принят уровень давления 120 дБ SPL.

В тишине чувствительность слуха человека повышается, а в атмосфере громких звуков – понижается, слух адаптируется к окружающей звуковой среде, поэтому

динамический диапазон слуха не такой большой – около 70..80 дБ. Сверху он ограничен давлением 100 дБ SPL, а снизу шумом с уровнем -30...35 дБ SPL. Этот динамический диапазон может сдвигаться вверх и вниз до 20 дБ. Для комфортного восприятия музыки рекомендуется, чтобы звуковое давление не превышало 104 дБ SPL в домашних условиях и 112 дБ SPL, в специально оборудованных помещениях.

Динамический диапазон музыки определяется отношением в децибелах самого громкого звука (фортиссимо) и самого тихого звука (пианиссимо). *Динамический диапазон симфонической музыки* составляет 65...75 дБ, а на концертах рок-музыки он возрастает до 105 дБ, при этом пики звуковых давлений могут достигать 122...130 дБ SPL. Динамический диапазон вокальных исполнителей - не превышает 35...45 дБ (рис.5.2).

Шум помещения маскирует звук и этим уменьшает динамический диапазон. Влиянием шума можно полностью пренебречь, если его уровень на 10...20 дБ SPL ниже минимального уровня музыкальных звуков. В студиях звукозаписи уровень шума не превышает 20...30 дБ SPL, ночью в квартирах «тихих» домов этот уровень равен 40 дБ

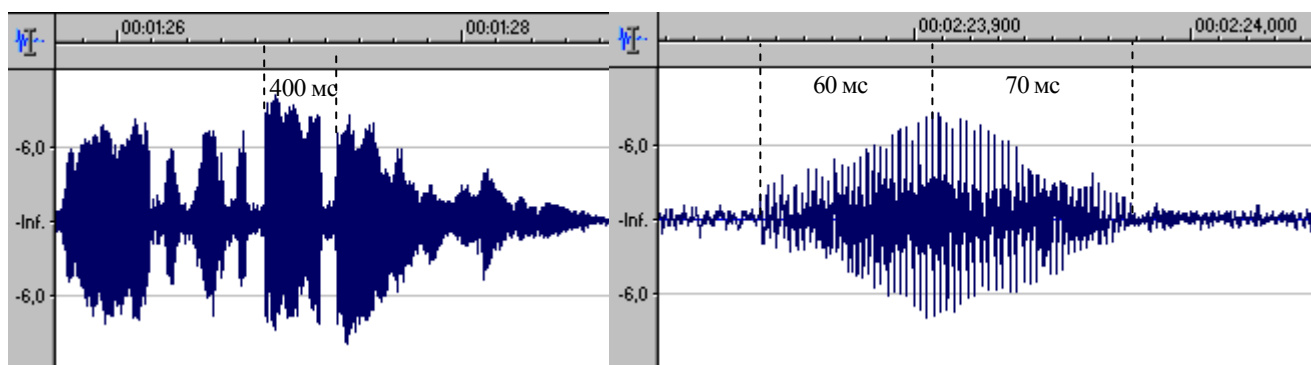


Рис.5.3. Фрагменты фонограмм музыкальных отрывков различной длительности

SPL, любой разговор увеличивает уровень шума до 60дБ SPL. Именно поэтому тихая музыка часто тонет в шумах помещения прослушивания и невольно возникает желание увеличить громкость.

Музыкальный и речевой сигналы представляют собой последовательность быстро нарастающих и более медленно затухающих звуковых импульсов (рис.5.3.). Такой сигнал

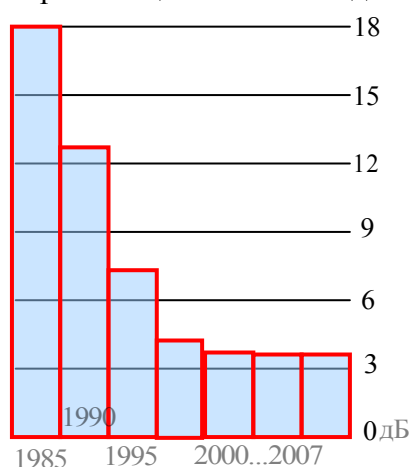


Рис. 5.4. Пик-фактор CD дисков

характеризуется *среднеквадратическим и пиковым значениями уровней*, разность этих уровней называется *пик-фактором*. Прямоугольная волна (меандр) имеет единичный пик-фактор 0 дБ, пик-фактор синусоиды равен 3 дБ. Фонограммы музыкальных и речевых сигналов имеют пик-фактор до 20 дБ и более. *Время определения пик-фактора* связано временем интегрирования при вычислении среднеквадратического значения сигнала, и обычно, оно равно 50 мс.

Эмоциональную музыку с широким динамическим диапазоном и большим пик-фактором можно слушать только на высококачественной дорогой аппаратуре с хорошими акустическими колонками. На дешевых плеерах с наушниками и в автомобилях из-за шумов динамический диапазон сокращается и она звучит просто отвратительно. Поэтому широким спросом такие записи не пользуются и, неизбежно, с каждым годом динамический диапазон и пик-фактор фонограмм преднамеренно изготовителями уменьшается (рис.5.4.). На современных CD дисках в большинстве случаев динамический диапазон не превышает 20 дБ, а пик-фактор- чуть больше 3 дБ, что вполне достаточно для танцевальной музыки. На рис.5.4. приведена картинка современной фонограммы с компакт диска.

В программе Audition динамический диапазон и пик-фактор музыкальной фонограммы определяют путем статистической обработки мгновенных значений

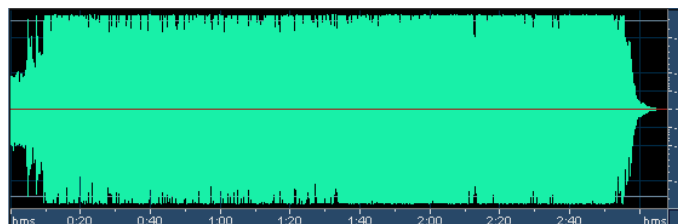


Рис.5.5. Фонограмма песни "I'll Be There For You"

сигналов. Затем по команде **Анализ** рассчитываются статистические характеристики фонограммы (рис 5.6.), из которых основными являются: Peak Amplitude (L_{pic}), (Maximum RMS Power (L_{max}), Minimum RMS Power (L_{min}) и Average RMS Power (L_{avr}) (уровни максимального, минимального и среднего среднеквадратического (эффективного) значения мощности сигнала).

Динамический диапазон фонограммы определяется как

$$DR_m = L_{pic} - L_{min}$$

пик-фактор рассчитывается по формуле

$$PF_m = L_{pic} - L_{avr}$$

DR

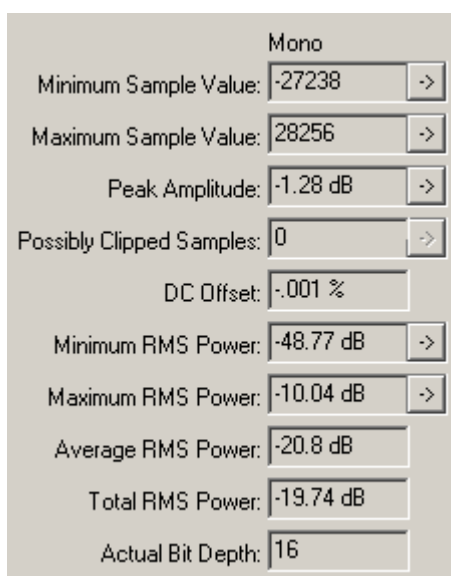


Рис.5.6. Статистические характеристики фонограммы музыки Бетховена «Элизе»

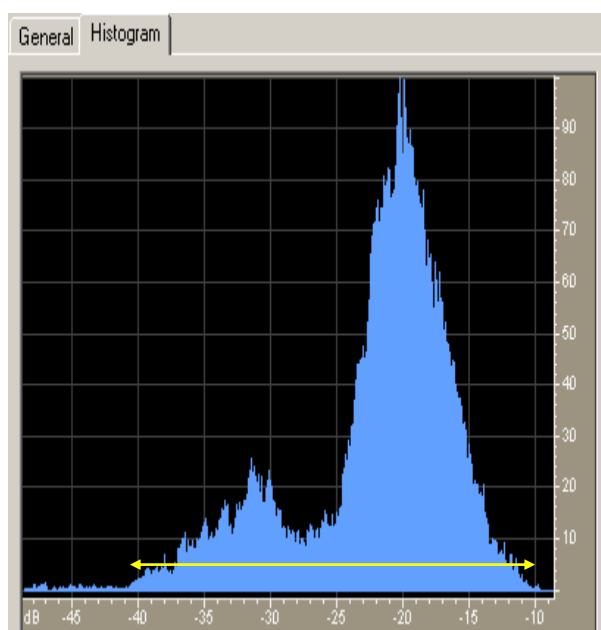


Рис.5.7. Гистограмма распределения музыки Бетховена «Элизе»

Динамический диапазон может быть также определен по гистограмме распределения уровней фонограммы, приведенной на рис. 5.7..

Динамический диапазон микрофонов определяется так же, как это обычно делается в электрических трактах. Верхняя граница ограничивается допустимой величиной нелинейных искажений, а нижняя - уровнем собственных шумов. Современные студийные микрофоны допускают максимальное звуковое давление 125...145 дБ SPL, при этом нелинейные искажения не превышают 0,5% ... 3%. Уровень собственных шумов микрофонов составляет 15...20 дБА, динамический диапазон – от 90 до 112 дБА, а

отношение сигнал/шум- от 70 до 80 дБА. Эти микрофоны с запасом перекрывают весь диапазон слуха человека от 120 дБ SPL до уровня шума студии 20 дБ SPL. В современных студиях запись производится с использованием 22 или 24 разрядных АЦП, иногда используется квантование с плавающей запятой, поэтому проблем с динамическим диапазоном не возникает. Стоит такая аппаратура крайне дорого.

4.3. Громкость и динамический диапазон звука

С динамическим диапазоном звукового сигнала и его пик-фактором тесно связана такая важная характеристика фонограммы как громкость звучания. *Громкость звука* является субъективной характеристикой. Она наиболее сильно зависит от двух факторов - звукового давления и спектрального состава звука. В качестве одной из оценок громкости используется понятие уровня громкости, измеряемого в фонах. В

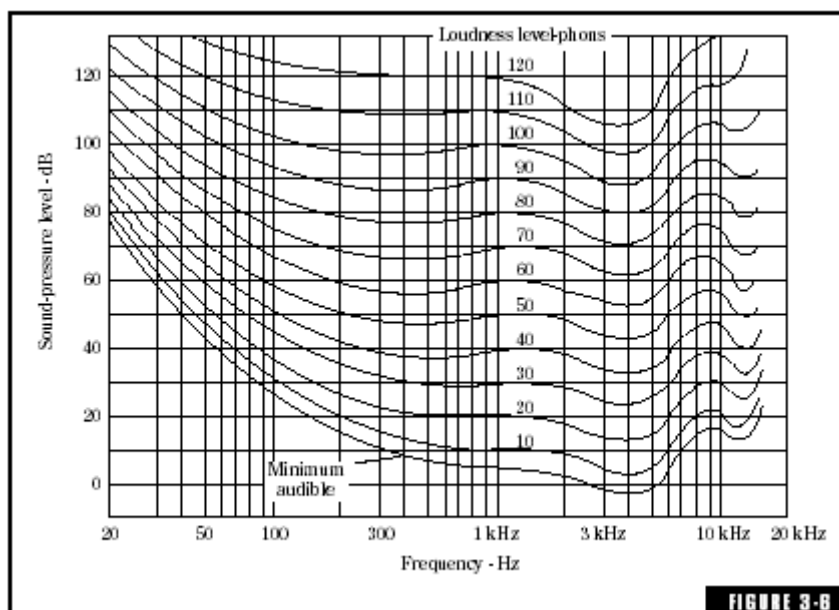


Рис.5.8. Кривые равной громкости

соответствии со стандартом ISO на частоте 1 кГц *уровень громкости в фонах* полностью определяется уровнем звукового давления в децибелах (например, 40 дБ SPL – 40 фон). Связь между уровнем громкости и звуковым давлением на других частотах определяется

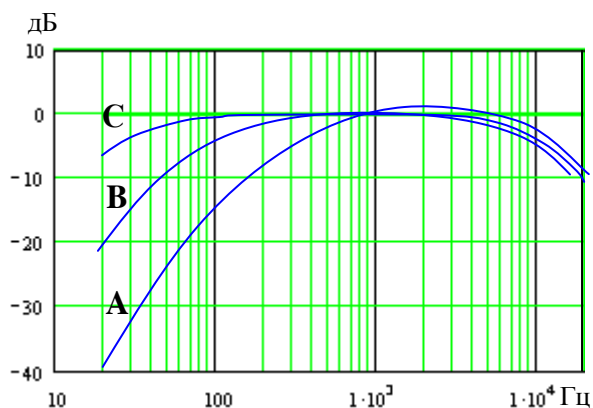


Рис. 5.9. Взвешивающие кривые

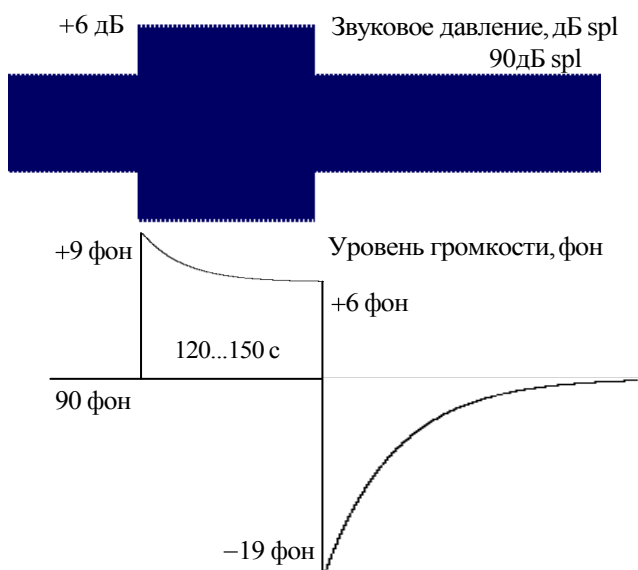
экспериментальными графиками, называемыми *кривыми равной громкости* или *изофонами* (рис.5.8).

Из приведенных графиков кривых равной громкости следует, что для натурального звучания запись должна воспроизводиться при том же уровне громкости, при котором она была осуществлена. Когда запись, сделанная на высоком уровне громкости, воспроизводится на более низких уровнях, низкие и высокие частоты как бы пропадают, частотный баланс нарушается.

Если запись сделана на низком уровне громкости, а воспроизводится на высоком - пение и речь кажутся бубнящими, так как слушатель воспринимает в них низкие частоты значительно более громкими, чем при прослушивании естественного источника на более «тихих» уровнях.

Изофоны в перевернутом виде (рис.5.9.), называемые весовыми характеристиками, представляют собой частотную зависимость чувствительности слуха человека от уровня громкости звука. Кривые с индексами С, В и А соответствуют изофонам с уровнями громкости 100, 70 и 30 фон. Из этих графиков следует, что частотные искажения слуховой системы тем меньше, чем больше уровень громкости звука. Именно поэтому *громкая музыка чаще всего кажется "красивее"*, чем тихая, и больше привлекает внимание. На этом основании большинство продюсеров всеми силами стремятся повысить уровень записи фонограммы при мастеринге, так как от этого может зависеть ее коммерческий успех.

При динамической обработке звуковых сигналов следует учитывать *эффект адаптации слуха*, который возникает при звуковых давлениях 90 и более децибел. Этот



эффект заключается в том, что изменение уровня громкости при повышении и понижении звукового давления на одну и ту же величину, допустим на 6 дБ, воспринимается совершенно по разному. При резком повышении давления на 6 дБ уровень громкости скачком увеличивается на 9 фон и затем медленно убывает по экспоненциальному закону (рис.5.10) до значения +6 фон. При резком уменьшении звукового давления, также на 6 дБ, уровень громкости скачком уменьшается на 19 фон и затем очень медленно нарастает до стационарного значения.

Время адаптации слуха к уровню громкости при перепаде звукового давления на 6 дБ примерно равно 120

...150 с, максимально оно может достигать 260 с. При спаде звукового давления время адаптации существенно больше, и скачок уровня громкости составляет 19 фон. Все это учитывается при выборе временных параметров динамических процессоров.

График огибающей уровня громкости на рис.5.10. свидетельствует, что при больших уровнях звукового давления слуховой аппарат человека осуществляет своеобразную компрессию звуковых сигналов, подобную динамической обработке. Такое происходит благодаря нелинейности растяжения барабанной перепонки внешнего уха. Этим слуховой аппарат самостоятельно старается защитить собственные уши.

Таким образом, представление о возможности естественного восприятия всего динамического диапазона музыки является совершенно иллюзорным. Динамическая обработка звуковых сигналов в электрических трактах избавляет наши уши от ненужной работы. Важно только чтобы осуществлялось это так же, как это делает слуховой аппарат.

Как при аналоговой, так и цифровой звукозаписи увеличение громкости звучания фонограммы путем увеличения уровня записи в значительной мере ограничивается пик-фактором звукового сигнала. В первом случае важно, чтобы пики не приводили к сильному насыщению магнитного носителя, а во втором - пики могут вызвать клипирование сигнала.

При сведении на цифровой носитель фонограмма без динамической обработки имеет пик-фактор в районе 20 дБ. Это означает, что она будет звучать на 6 дБ (в два раза!) тише, чем та же фонограмма, сведенная на аналоговую магнитную ленту. Поэтому цифровая звукозапись может звучать *менее привлекательно*, чем ее оригинал на студийной

магнитной ленте. Для компенсации двукратной потери громкости приходится использовать довольно жесткую динамическую обработку звуковых сигналов.

В настоящее время самая современная аудиотехника не в состоянии передать полной динамики, которую воспринимает человеческое ухо. Поэтому для практических целей при записи всегда приходится осуществлять сокращение динамического диапазона с использованием динамической обработки. При выполнении этой операции решаются как технические, так и художественные задачи. Прежде всего, необходимо полностью использовать динамический диапазон носителя. Также важно максимально использовать динамический диапазон воспроизводящего устройства, чтобы запись не тонула в шумах. Третьей, наиболее сложной задачей, является *повышение громкости звучания за счет сокращения динамического диапазона* без заметного ухудшения качества звучания. Это достигается тем, что при меньшем пик-факторе звукового сигнала более высокая громкость звучания фонограммы достигается путем увеличения уровня записи.

4.4. Принцип динамической обработки

Динамическая обработка звуковых сигналов производится для сужения или расширения динамического диапазона. Прибор или программный модуль, выполняющий эту функцию, называется *динамическим процессором (DP)*. Динамическая обработка применяется только для нестационарных сигналов с подъемами и провалами огибающей звукового сигнала. Уровень огибающей определяется с помощью пикового детектора, при этом одновременно вычисляется среднеквадратическое значение сигнала с временем интегрирования примерно равным постоянной времени слуха – 50...60 мс.

Принцип цифровой *динамической обработки* основан на почти *мгновенном изменении коэффициента передачи* сигналов, когда уровень огибающей звукового сигнала становится выше (ниже) установленного порога. Порог может задаваться как пиковым значением уровня, так и среднеквадратическим.

Основными параметрами DP являются: *передаточная характеристика* (в виде графика или функции), порог срабатывания, время установления (attack) T_y и время восстановления T_b коэффициента передачи. *Перелом* в передаточной характеристике называется точкой *перегиба*. Входной уровень, соответствующий перегибу, называется *порогом*, задается он в децибелах. *Угол наклона графика передаточной характеристики выше (ниже) точки перегиба* определяет коэффициент передачи подъемов (провалов)

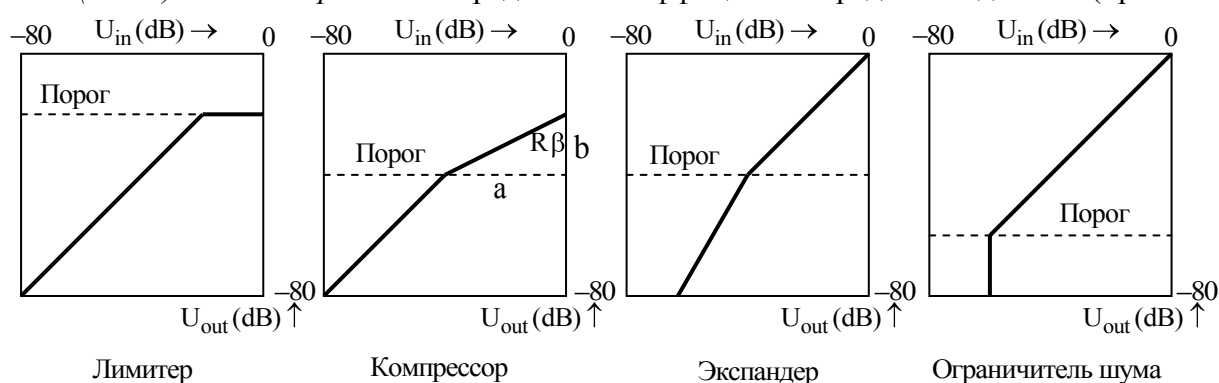


Рис.5.11. Передаточные функции динамических процессоров

огибающей звукового сигнала.

В зависимости от вида передаточной характеристики DP он может выполнять функции лимитера, компрессора, экспандера или гейта. Все эти функции могут быть реализованы в одном устройстве или с помощью одного программного модуля. На *рис.5.11* приведены типовые передаточные характеристики этих устройств.

Передаточная характеристика DP однозначно определяет *регулирующую характеристику*, в соответствии с которой меняется коэффициент передачи

регулирующего элемента ДР. Изгибы и изломы передаточной характеристики предопределяют дискретный характер изменения коэффициента передачи регулирующего элемента ДР, и они в значительной мере определяют форму *амплитудной огибающей* динамически обработанного звукового сигнала.

Лимитер ограничивает динамический диапазон сверху – пики и подъемы огибающей, значение которых превышают пороговый уровень, он ослабляет до уровня, равного установленному порогу. *Компрессор* также ограничивает динамический диапазон сверху – подъемы огибающей, величина которых превышает пороговый уровень, он ослабляет в соответствии с установленным коэффициентом передачи выше точки перегиба.

Разница в работе лимитера и компрессора особенно заметна при малой скорости нарастания и спада огибающей входного звукового сигнала (рис.5.12). Все изломы

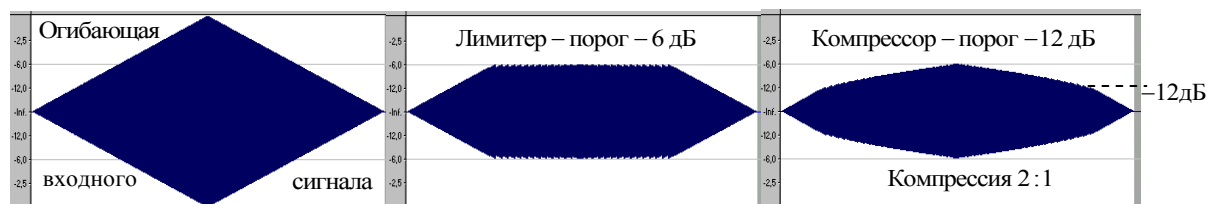


Рис. 5.12. Амплитудные огибающие выходного сигнала ДР при $T_y = T_v = 0$

амплитудной огибающей выходного сигнала полностью соответствуют моментам скачкообразного изменения коэффициента передачи. Как видно, при компрессировании склоны амплитудной огибающей более плавные. С увеличением крутизны нарастания и спада огибающей входного сигнала изломы амплитудной огибающей становятся все резче, вплоть до полного разрыва, а разница в работе лимитера и компрессора нивелируется.

Экспандер расширяет динамический диапазон снизу – провалы огибающей, уровень которых ниже порогового он увеличивает путем уменьшения коэффициента передачи ниже точки перегиба. *Гейт* ограничивает динамический снизу – все сигналы в провалах огибающей, уровень которых ниже порогового он не пропускает.

Степень сжатия (компрессии) определяется отношением коэффициентов передачи ниже и выше точки перегиба передаточной функции. Поэтому она может задаваться как отношение приращений входного и выходного сигналов в децибелах, например, 18 : 6. Коэффициент передачи определяется также углом наклона передаточной функции β , поэтому степень сжатия может задаваться в виде отношения сторон прямоугольного треугольника a и b (рис.5.12.). Если, например, $a : b = 1 : 1$, угол наклона равен 45 градусам- это значит, что сжатия нет. Если $a : b = 2 : 1$ - это значит, что при увеличении входного уровня на 2 дБ выше порога выходной уровень вырастет лишь на 1 дБ и, следовательно, степень сжатия равна 2.

В работе ДР очень важную роль играют его временные параметры: время установления T_y , часто называемое временем срабатывания, и время восстановления T_v , в значительной мере определяющие длительность переходных процессов. Эти параметры измеряются в единицах времени, тогда как скорость установления и восстановления измеряется в дБ/с. Последнее означает, что время срабатывания и время восстановления не являются величинами постоянными, а зависят от уровня звукового сигнала. Эта зависимость проявляется в том, что чем больше уровень сигнала, тем больше время установления и меньше время восстановления.

Выбор значений этих параметров зависит от режима работы динамического процессора (компрессор, экспандер, лимитер или гейт). Существенное значение имеет характер музыкального материала. Иногда эти параметры выбираются исходя из необходимости решения каких либо творческих задач. Современные ДР позволяют плавно менять время установления от 0 до 500 мс, а время восстановления - от 0 до 5 с.

Чем меньше время установления и восстановления, тем громче получается результирующий звук, но тем более изломанной получается амплитудная огибающая выходного сигнала DP, и тем больше возникает интермодуляционных искажений. Ясно, что чем больше время установления и восстановления, тем медленнее будет меняться во времени амплитудная огибающая, тем более гладкой она будет. Однако, при большом времени установления и восстановления начинает проявляться эффект "проваливания громкости" (pumping). Провалы громкости возникают вокруг кратковременных пиков в сигнале. Таким образом, выбор времени установления/восстановления - это компромисс между интермодуляционными искажениями и эффектом проваливания громкости

На многих моделях компрессоров есть возможность включения автоматического режима установки одной или обеих постоянных времени. Более того, существуют модели компрессоров, у которых одна из постоянных времени вообще не имеет ручной установки (устанавливается автоматически), а еще бывают и полностью автоматические компрессоры - вообще без регуляторов времени срабатывания и восстановления.

Как правило, в автоматическом режиме время срабатывания уменьшается при увеличении уровня входного сигнала, время же восстановления при этом увеличивается. В результате удастся добиться большей ровности и мягкости амплитудной огибающей в моменты резкого изменения уровня входного сигнала.

Время установления T_y определяет время между моментом нарастания или спада огибающей входного сигнала до порогового значения (снизу или сверху) и временем, через которое устанавливается заданный уровень выходного сигнала динамического процессора.

Чем меньше время установления - тем на более коротких пиках уровня срабатывает DP. При малом времени установления можно практически исключить возникновение пиков сигнала при скачкообразном увеличении его уровня. Однако при этом звучание может стать недостаточно акцентированным. Так же важно помнить, что малое время установления может вызвать искажения, особенно на низких частотах. Низкочастотные сигналы имеют большой период, и компрессия, действующая посреди периода, звучит очень неестественно.

При большом времени установления компрессор будет пропускать короткие атаки звукового сигнала, превышающие порог, т.к. не будет успевать на них реагировать. Это может быть нежелательным, например, для лимитеров. В сигнале на выходе компрессора будут присутствовать пики.

Для лимитеров чаще всего выбирают время установления в пределах 0,5 - 1,0 мс, для речевых компрессоров время установления должно составлять 1... 2 мс, в обычных компрессорах время установления - порядка 10...100 мс.

Время восстановления T_v определяет время между моментом спада пика огибающей входного сигнала до порогового уровня (сверху или снизу) и временем, через которое на выходе процессора восстановится исходный уровень звукового сигнала. При малом времени восстановления обеспечивается более существенное сглаживание. Если время восстановления слишком велико, то компрессор дольше находится в активном состоянии и воздействует на динамический диапазон даже тогда, когда это нежелательно. Это дает заметный на слух эффект пульсации звука, так как компрессия не приводит к сглаживанию сигнала.

Обычно рекомендуется для инструментальной музыки в качестве грубого приближения и отправной точки для более тонкой настройки выбирать время восстановления около 500 мс. Это соответствует промежутку между двумя тактами при темпе 120 четвертей в минуту. Для речевых компрессоров время восстановления - 300 мс, в обычных компрессорах время восстановления - порядка 100...1000 мс.

Для исследования и демонстрации работы различных DP больше всего подходит синусоидальный звуковой сигнал с амплитудно-импульсной модуляцией (АИМ), у которого огибающая имеет прямоугольную форму (рис.5.13.). Минимально возможная

крутизна фронтов таких импульсов (время атаки) определяется частотой звукового сигнала, и поэтому может принимать значения от 25 мкс до 10 мс. Это следует учитывать при выборе времени установления динамического процессора.

На **рис.5.13** приведена сигналограмма с выхода графического динамического процессора звукового редактора Sound Forge, работающего в режиме компрессора.

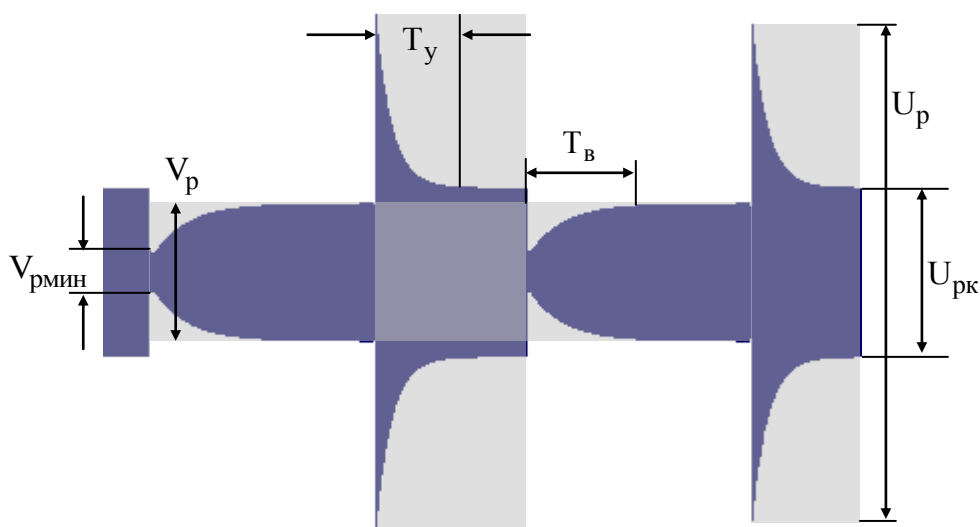


Рис. 5.13. Динамическая обработка АИМ сигнала с помощью компрессора

Устанавливаемый порог определяет максимальный размах выходного сигнала U_{pk} в установившемся режиме, а степень компрессии определяет изменение коэффициента передачи.

В момент нарастания входного пика огибающей сигнала его размах U_p на выходе компрессора почти такой же, как и на входе. Это обусловлено тем, что в рассматриваемом примере время установления существенно больше времени фронта импульсов. Спад огибающей выходного сигнала компрессора происходит по затухающей экспоненте. По окончании переходного процесса (статика) коэффициент усиления компрессора уменьшается в три раза, во столько же раз уменьшается размах выходного сигнала U_{pk} .

В момент спада пика огибающей установившийся в режиме коэффициент усиления не может мгновенно измениться, поэтому происходит резкое уменьшение размаха выходного сигнала V_p , тоже в три раза. Далее происходит постепенное восстановление коэффициента усиления до исходного значения по нарастающей экспоненте. Длительность этого процесса определяется значением T_b .

Из приведенного примера следует, что применение компрессии отнюдь не означает автоматического уменьшения динамического диапазона по пиковым значениям и не исключает возможности клиппирования звукового сигнала. Хорошая работа компрессора очень сильно зависит от выбора постоянных времени процессора. В некоторых случаях можно допустить моментальное срабатывание компрессора, но время восстановления не может быть нулевым. Это связано с тем, что любому акустическому затуханию, например, колебания струны или реверберирующий звук, сопутствуют флуктуации, которые в компрессоре могут приводить к спонтанному изменению громкости.

5. Порядок выполнения лабораторной работы

5.1. Подготовка к измерениям


- Скопируйте с рабочего стола файл «Отчет LR-5 АТ» на свою флешкарту, дайте ему свое имя в виде: «Группа»_ «Фамилия (работавшего на клавиатуре)»_ «Номер ЛР»_ «Дата».
- Откройте этот файл и сверните его окно.

В файл этого отчета вы будете копировать все экспериментальные графики, и заносить в таблицы отчета результаты измерений сразу в процессе их получения. В день выполнения лабораторной работы отчет в электронном виде необходимо сдать преподавателю, только тогда работа считается выполненной.



- Кнопкой с пиктограммой  откройте окно программы **Audition**. Не меняйте размеры этого окна, они согласованы с Вашим отчетом



- Кнопкой с пиктограммой  откройте окно программы **HyperSnap-DX**, предназначенной для вырезки и копирования графических изображений
- Задайте время интегрирования при измерении статистических характеристик.

Для этого кнопкой **File** программы **Audition** откройте меню тестовых музыкальных сигналов и выберите в нем Тест 1. Бетховен. «Элиза». По команде **Open** в окне осциллографа появится сигналограмма данного музыкального фрагмента. Нажмите кнопку **Analyze**. В выпадающем меню выберите пункт **Statistics**. По этой команде на экране появится окно **Waveform Statistics – General** со статистическими характеристиками фонограммы. В окне **Window Width** введите время интегрирования 50 мс и нажмите кнопку **Recalculate**

5.2. Измерение динамических характеристик фонограмм без компрессирования

- Кнопкой **File** программы **Audition** откройте меню тестовых музыкальных сигналов и выберите в нем Тест 1. Бетховен. «Элиза». По команде **Open** в окне осциллографа появится сигналограмма данного музыкального фрагмента.
- Скопируйте сигналограмму (рис.5.14) в Ваш отчет.

Для этого в окне программы **HyperSnap-DX** нажмите кнопку **Capture** и в открывшемся окне выберите **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу окна гистограммы и нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол окна гистограммы и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** нажмите кнопку **File** и в открывшемся меню выберите **Save As**. Установите путь копирования на Вашу флешкарту, дайте название файлу и сохраните его в своем отчете.

- Нажмите кнопку **Analyze**. В выпадающем меню выберите пункт **Statistics**. По этой команде на экране появится окно **Waveform Statistics – General** со статистическими характеристиками фонограммы. Результаты измерений L_{pic} , L_{min} и L_{avr} занесите в табл.5.1.
- По формулам

$$DR_m = L_{pic} - L_{min}$$

$$PF_m = L_{pic} - L_{avr}$$

рассчитайте динамический диапазон и пик-фактор фонограммы, результаты занесите в табл.5.1

Табл.5.1. Динамические характеристики фонограмм при времени интегрирования 50 мс						
Степень компрессии	L_{pic}, dB	L_{min}, dB	DR_m, dB	DR_g, dB	L_{avr}, dB	PF_m, dB
1:1						
2,4:1 (50\30)						
5:1 (50\40)						

- Нажмите кнопку **Histogram** - в окне появится гистограмма распределения мгновенных уровней фонограммы. Скопируйте гистограмму (рис.5.17) в Ваш отчет.

Для этого в окне программы **HyperSnap-DX** нажмите кнопку **Capture** и в открывшемся окне выберите **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу окна гистограммы и нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол окна гистограммы и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** нажмите кнопку **File** и в открывшемся меню выберите **Save As**. Установите путь копирования на Вашу флешкарту, дайте название файлу и сохраните его в своем отчете.

- Определите динамический диапазон DR_g по гистограмме, результат занесите в табл.5.1.

5.3. Компрессирование цифровых фонограмм

- Для компрессирования исходного файла «Элизе» в меню **Effects** программы **Audition** выберите пункт **Amplitude** (рис.5.20). В выпадающем меню выберите пункт **Dynamics Processing**. По этой команде на экране появится окно Графического динамического компрессора (рис.5.21)

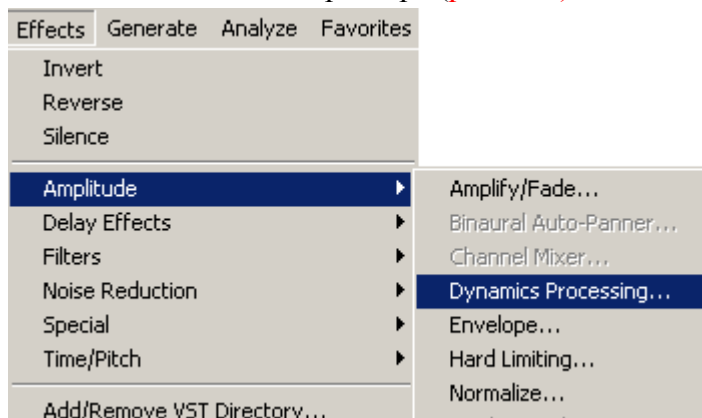


Рис.5.20. Меню компрессора

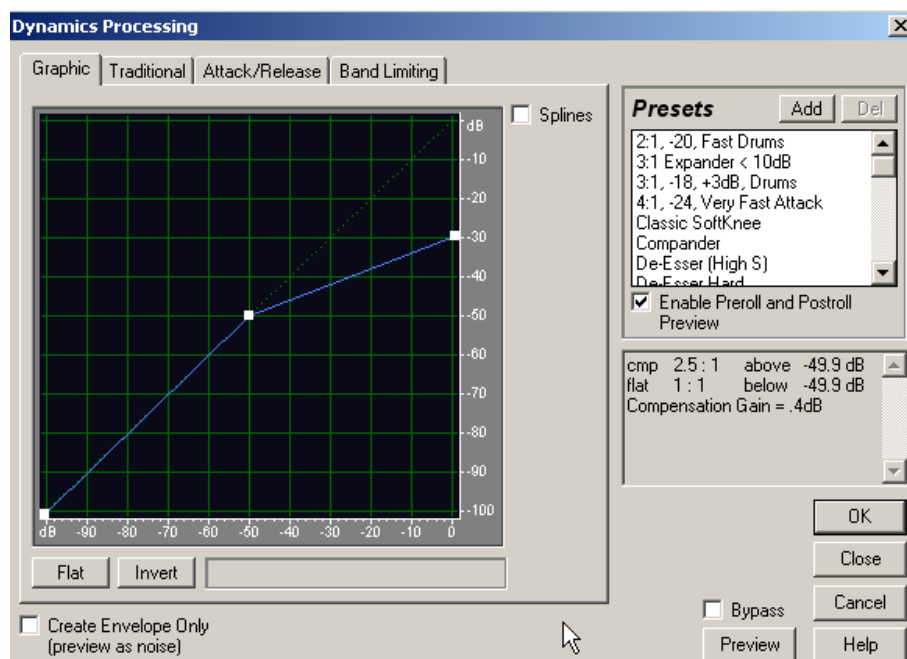


Рис.5.21. Окно динамического графического компрессора

- Компрессирование начинается с выбора уровня сигнала, ниже которого сигнал не компрессируется (допустим минус 50 дБ). С помощью курсора первая точка изгиба передаточной функции устанавливается по координатам +50 , –50 дБ
- Курсором вторая точка перемещается по вертикали 0 дБ пока в первой строчки измерителя степени компрессирования не установится заданное соотношение сmp 2,4:1. Компрессирование осуществляется по нажатию кнопки ОК- сигнал уменьшается по амплитуде (рис.5.15).
- Для повышения громкости звучания после компрессирования всегда производится нормализация фонограммы, при которой сигнал усиливается до заданного уровня. Для выполнения этой операции в меню **Effects** программы **Audition** выберите пункт **Amplitude** (рис.5.20). В выпадающем меню выберите **Normalize**. По этой команде скомпенсированный сигнал усиливается (рис.5.15) до значения близкого к максимально возможному.
- Скопируйте сигналограмму (рис.5.15) в Ваш отчет.
Для этого в окне программы **HyperSnap-DX** нажмите кнопку **Capture** и в открывшемся окне выберите **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу окна гистограммы и нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол окна гистограммы и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** нажмите кнопку **File** и в открывшемся меню выберите **Save As**. Установите путь копирования на Вашу флешкарту, дайте название файлу и сохраните его в своем отчете.
- Нажмите кнопку **Analyze**. В выпадающем меню выберите пункт **Statistics**. По этой команде на экране появится окно **Waveform Statistics – General** со статистическими характеристиками фонограммы . Результаты измерений L_{pic} , L_{min} и L_{avr} занесите в табл.5.1.

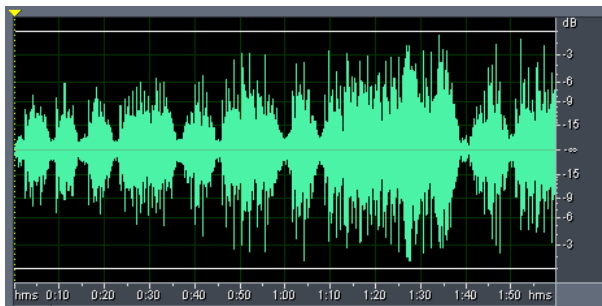


Рис. 5.14. Нет компрессии (исходная фонограмма)

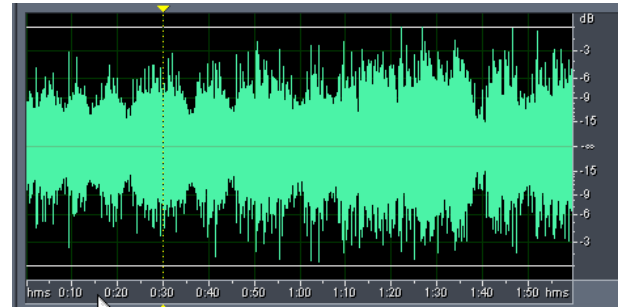
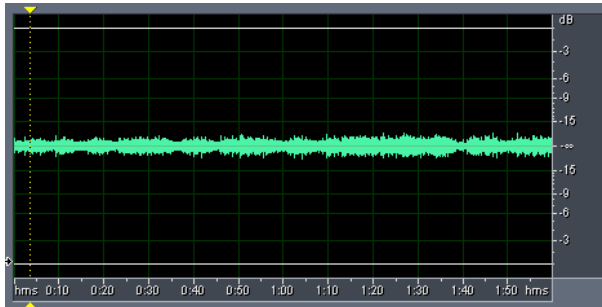


Рис. 5.15. Степень компрессии 2,4:1 (50-30) После компрессора

После нормализатора

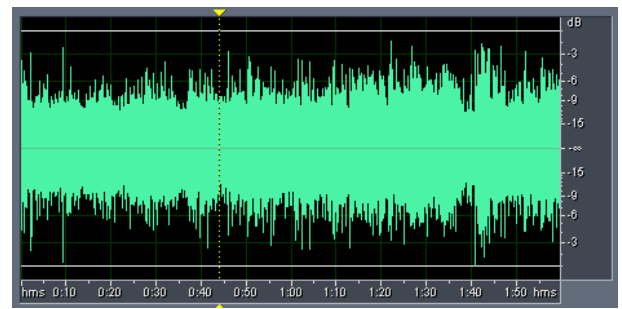
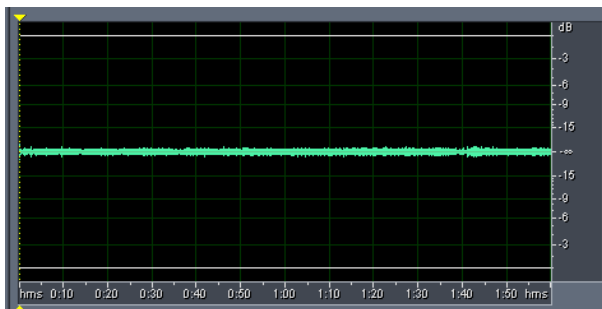


Рис.5.16. Степень компрессии 5:1 (50-40) После компрессора

После нормализатора

Фонограммы музыки Бетховена «Элизе». Компрессирование - нормализация

- По формулам

$$DR_m = L_{pic} - L_{min}$$

$$PF_m = L_{pic} - L_{avr}$$

рассчитайте динамический диапазон и пик-фактор фонограммы, результаты занесите в табл.5.1

Скопируйте гистограмму (рис.5.17 в Ваш отчет.

Для этого в окне программы **HyperSnap-DX** нажмите кнопку **Capture** и в открывшемся окне выберите **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу окна гистограммы и нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол окна гистограммы и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** нажмите кнопку **File** и в открывшемся меню выберите **Save As**. Установите путь копирования на Вашу флешку, дайте название файлу и сохраните его в своем отчете.

- Повторите операции по п.5.3 при степени компрессирования 5:1 (50-40) дБ. Скопируйте графики сигналограмм (рис.5.16) и гистограммы (рис.5.17-5.19)

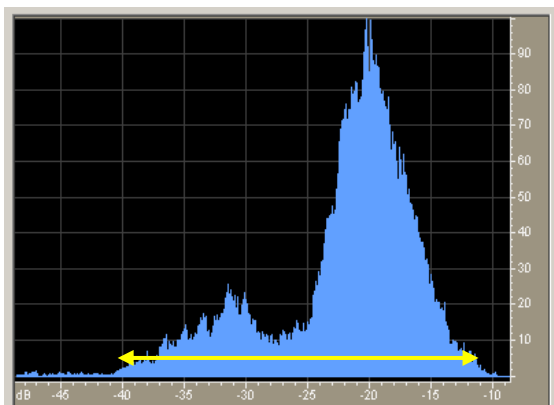


Рис.5.17. Гистограмма исходной
Фонограммы DR = 30 дБ

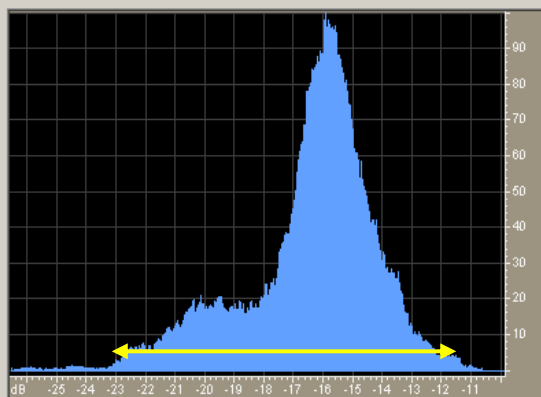


Рис.5.18. Гистограмма при степени
компрессирования 5 :1. DR= 11,5 дБ

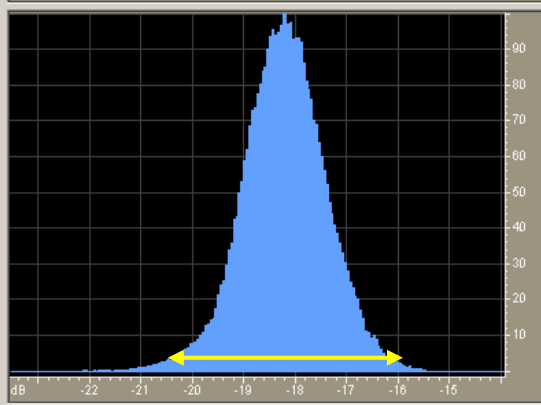


Рис.5.19. Гистограмма при
степени компрессии 5 :1
DR = 4,3 дБ

6. Контрольные вопросы

21. Как отличается восприятие музыки с широким и узким динамическим диапазоном
22. Какую роль играет пик-фактор музыки на ее восприятие
23. Что понимается под термином «динамический диапазон фонограммы»
24. Как определяется динамический диапазон фонограммы через ее статистические характеристики
25. Что понимается под термином «пик-фактор фонограммы»
26. Как измеряется пик-фактор фонограммы
27. В чем проявляются отличия динамического диапазона фонограммы и динамического диапазона музыки, записанной на этой же фонограмме
28. Что характеризует гистограмма фонограммы
29. Как рассчитывается гистограмма

7. Содержание отчета

- 7.1. Титульный лист
- 7.2. Цель работы
- 7.3. Задание к лабораторной работе

- 7.4. Функциональная схема измерений
- 7.5. Результаты измерений в виде графиков
- 7.6. Таблицы с результатами измерений и расчетов
- 7.7. Выводы по результатам измерений

8. Литература

Вологдин Э.И. Методы и алгоритмы обработки звуковых сигналов. Учебное пособие. Часть 1 . СПб ГУТ. 2009.

ЛАБОРАТОРНАЯ РАБОТА N 6 АТ

ИССЛЕДОВАНИЕ ЧАСТОТНЫХ ХАРАКТЕРИСТИК ШЕЛЬФОВЫХ ФИЛЬТРОВ

Руководство к выполнению лабораторной работы по Аудиотехнике

1. Цель работы

Целью лабораторной работы является исследование частотных характеристик (ЧХ) шельфовых фильтров низких и высоких частот с помощью тестового сигнала белого шума с использованием компьютерной программы **Sound Forge 9c**.

2. Задание к лабораторной работе

- Ознакомиться с программой Sound Forge 9c в объеме, необходимом для синтеза тестовых сигналов белого шума
- Ознакомиться с программой Sound Forge 9c в объеме, необходимом для знакомства с характеристиками шельфовых фильтров
- Ознакомиться с программой Sound Forge 9c в объеме, необходимом для измерения частотных характеристик шельфовых фильтров с помощью сигналов белого шума
- Исследовать зависимость ЧХ шельфовых фильтров от коэффициента их усиления (ослабления)
- Исследовать зависимость ЧХ шельфовых фильтров от ширины переходной зоны в пределах от 0,1 до 5 октав
- По результатам измерений сформулировать выводы

3. Функциональная схема измерений

На **рис.6.1** приведена функциональная схема измерения частотных характеристик шельфовых и других фильтров с помощью тестового сигнала белого шума. Этот тестовый сигнал формируется в синтезаторе с заданным значением уровня спектральной плотности мощности минус 18 дБ длительностью 50 с. Чем больше эта длительность, тем меньше неравномерность уровня спектральной плотности шума и выше точность измерений.

Этот тестовый сигнал хранится в долговременной памяти, и при измерения ЧХ подается на вход исследуемого фильтра. Выходной сигнал фильтра, отражающий его ЧХ, хранится в другой памяти. Сигналы с входа и выхода фильтра через цифровой осциллограф могут подаваться на анализатор спектра, который измеряет частотную характеристику.

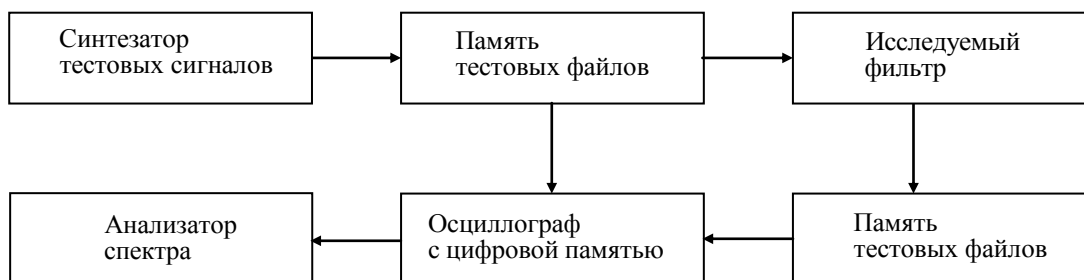


Рис. 6.1 Функциональная схема измерения частотных характеристик фильтров

4. Методические указания

Фильтры эквалайзеров типа Low-Shelf и High-Shelf предназначены для создания фиксированного подъема (спада) АЧХ в заданном интервале частот на краях частотного диапазона. В отечественной технической литературе эти фильтры часто называют шельфовыми (от английского слова shelf – полка). *Подъем АЧХ осуществляется в режиме усиления, а спад - в режиме ослабления.*

Шельфовые ФНЧ в режиме усиления на частотах ниже частоты среза подъем АЧХ нарастает плавно, и затем коэффициент передачи фильтра остается постоянным до нулевых частот (полка). В режиме ослабления ниже частоты среза происходит плавный спад АЧХ, и затем коэффициент передачи фильтра остается постоянным до нулевой частоты (полка). Частоту среза можно менять плавно во всем звуковом диапазоне без изменения крутизны нарастания (спада) АЧХ (рис.6.2.). Напомним, что под частотой среза шельфового фильтра понимается частота, уровень которой выше или ниже на 3 дБ относительно уровня плоской части полки.

Шельфовые ФВЧ в режиме усиления на частотах выше частоты среза подъем АЧХ нарастает плавно, и затем коэффициент передачи фильтра остается постоянным до частоты Найквиста (полка). В режиме ослабления выше частоты среза происходит плавный спад АЧХ, и затем коэффициент передачи фильтра остается постоянным до

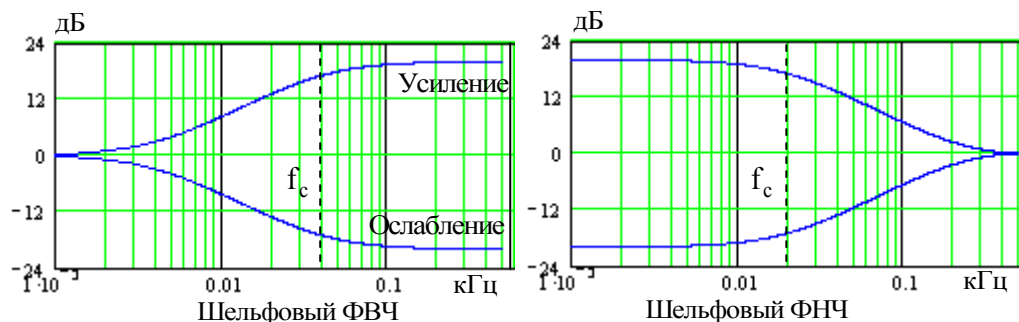


Рис.6.2. АЧХ шельфовых фильтров НЧ/ВЧ 1 порядка

частоты Найквиста (полка). Частоту среза можно менять плавно во всем звуковом диапазоне без изменения крутизны нарастания (спада) АЧХ.

Передаточные функции аналогового прототипа шельфового ФНЧ 1 порядка, полученные с помощью преобразования Лапласа,

$$H(s)_{\text{ФНЧ}} = \frac{s + V_0}{s + 1}, \quad V_0 > 1 \text{ (усиление)},$$

$$H(s)_{\text{ФНЧ}} = \frac{s + 1}{s + V_0}, \quad V_0 > 1 \text{ (ослабление)},$$

где V_0 – коэффициент передачи фильтра на нулевой частоте. В логарифмической шкале V_0 означает подъем и спад на одну и ту же величину, но разного знака (рис.6.2).

Передаточные функции шельфового ФВЧ 1 порядка

$$H(s)_{\text{ФВЧ}} = \frac{s \cdot V_0 + 1}{s + 1}, \quad V_0 > 1 \text{ (усиление)},$$

$$H(s)_{\text{ФВЧ}} = \frac{s + 1}{s \cdot V_0 + 1}, \quad V_0 > 1 \text{ (ослабление)},$$

где V_0 – коэффициент передачи фильтра на частоте, равной бесконечности. Как видно из этих формул переход из режима усиления в режим ослабления означает инвертирование передаточной функции. АЧХ этих фильтров (рис.6.2) отличаются очень плавным нарастанием (спадом) с крутизной не более 6 дБ/октава.

При цифровой реализации этих фильтров с помощью билинейного преобразования эти формулы трансформируются и принимают вид

$$H(z) = 1 + \frac{H_0}{2} (1 - A_1),$$

где $H_0 = 1 - V_0$, $A_1(z)$ – передаточная функция базового все пропускающего фильтра 1 порядка. В логарифмической шкале усиление фильтра задается в виде $V_0 = 10^{G/20}$, где G задается в децибелах. В этой формуле знаки – и + означают, что это, соответственно, шельфовые ФНЧ и ФВЧ. Этой формуле соответствует схема цифрового фильтра на рис.6.3. Она отличается введением дополнительной прямой связи с выхода на вход сумматора и множителя H_0 , определяющего усиление/ослабление фильтра.

Для увеличения крутизны подъема (спада) АЧХ на краях частотного диапазона используется шельфовые фильтры НЧ/ВЧ 2 порядка. Передаточные функции аналоговых прототипов этих фильтров определяются равенствами

Передаточные функции шельфового ФВЧ 1 порядка

$$H(s)_{\text{ФВЧ}} = \frac{s \cdot V_0 + 1}{s + 1}, \quad V_0 > 1 \text{ (усиление)},$$

$$H(s)_{\text{ФВЧ}} = \frac{s + 1}{s \cdot V_0 + 1}, \quad V_0 > 1 \text{ (ослабление)},$$

где V_0 – коэффициент передачи фильтра на частоте, равной бесконечности. Как видно из этих формул переход из режима усиления в режим ослабления означает инвертирование передаточной функции. АЧХ этих фильтров (рис.6.2) отличаются очень плавным нарастанием (спадом) с крутизной не более 6 дБ/октава.

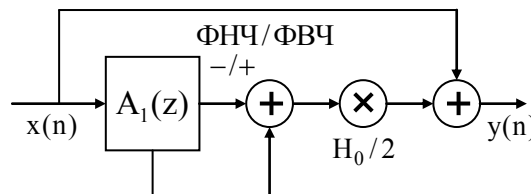


Рис.6.3. Шельфовый ФНЧ/ФВЧ 1 порядка

При цифровой реализации этих фильтров с помощью билинейного преобразования эти формулы трансформируются и принимают вид

$$H(z) = 1 + \frac{H_0}{2} (1 - A_1),$$

где $H_0 = 1 - V_0$, $A_1(z)$ – передаточная функция базового все пропускающего фильтра 1 порядка. В логарифмической шкале усиление фильтра задается в виде $V_0 = 10^{G/20}$, где G задается в децибелах. В этой формуле знаки – и + означают, что это, соответственно, шельфовые ФНЧ и ФВЧ. Этой формуле соответствует схема цифрового фильтра на рис.6.3.

На рис.6.4. приведены АЧХ цифровых шельфовых ФНЧ и ФВЧ 2 порядка с одной и той же частотой среза 200 Гц. На этих графиках иллюстрируется зависимость формы

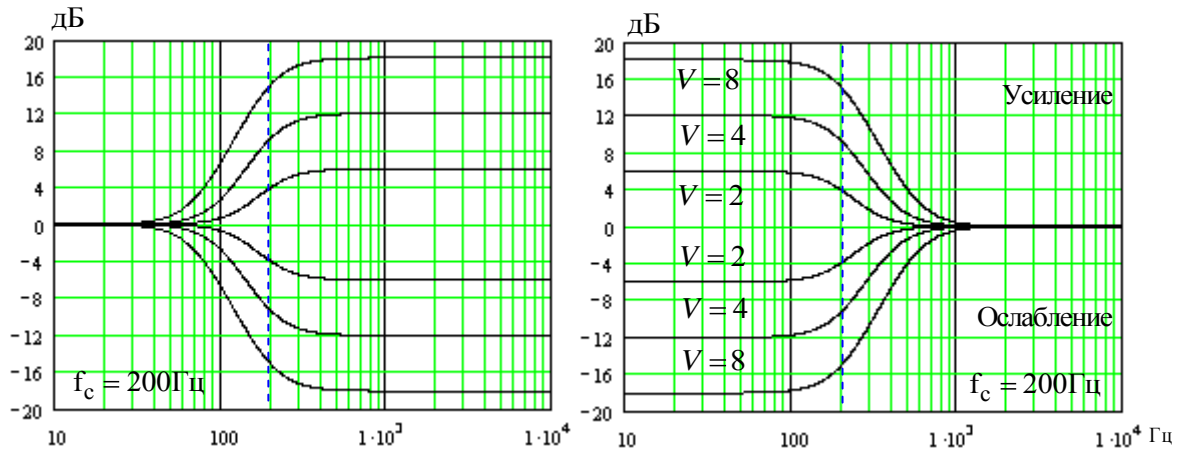


Рис.6.4. АЧХ шельфовых ФНЧ и ФВЧ 2 порядка.

АЧХ от коэффициента усиления (ослабления) фильтра.

Наиболее часто шельфовые фильтры используются для регулировки тембра на высоких и низких частотах. Частоты среза регуляторов тембра обычно жестко настраиваются производителем оборудования. Частоты максимального подъема и спада выбираются из соображений наибольшей заметности на слух вводимой коррекции. В современных плеерах обычно им соответствуют частоты 60 и 10000 Гц. В типовых корректорах величина регулируемого подъема и спада на этих частотах составляет ± 12 дБ на низких частотах и ± 15 дБ на высоких частотах.



Такие фильтры всегда присутствуют во всех пультах, они есть почти в каждом бытовом и автомобильном усилителе, приемнике, магнитофоне и плеере. С помощью регуляторов низких (Bass) и высоких (Treble) частот можно отрегулировать звучание системы таким образом, чтобы оно соответствовало вашим вкусам и - в некоторой степени - особенностям помещения, в котором вы находитесь.

5. Порядок выполнения лабораторной работы

5.1. Подготовка к измерениям

- Скопируйте с рабочего стола файл «Отчет LR-6 АТ» на свою флешку, дайте ему свое имя в виде: «Группа»_ «Фамилия (работавшего на клавиатуре)»_ «Номер ЛР»_ «Дата».
- Откройте этот файл и сверните его окно.

В файл этого отчета вы будете копировать все экспериментальные графики, и заносить в таблицы отчета результаты измерений сразу в процессе их получения. В день выполнения лабораторной работы отчет в электронном виде необходимо сдать преподавателю, только тогда работа считается выполненной.

- Кнопкой с пиктограммой  откройте окно программы **Sound Forge 9c (SF)**, содержащее окно сигналограммы и окно спектра левого звукового канала. Не меняйте размеры этого окна, они согласованы с Вашим отчетом
- Кнопкой с пиктограммой  откройте окно **Spectrum Setting** анализатора спектра, (рис.6.5) и сделайте предустановки, указанные на рисунке (если они не сделаны ранее)

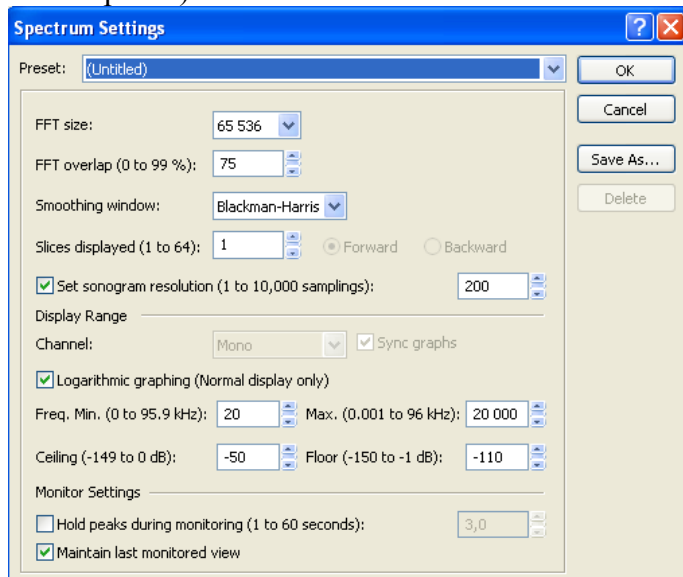



Рис.6.5. Окно предустановок анализатора

спектра

- Кнопкой с пиктограммой  откройте окно программы **HyperSnap-DX**, предназначенной для вырезки и копирования графических изображений.

5.2. Синтез тестового сигнала белого шум

- В меню **File** программы **SF** выберите **New**. В открывшемся окне (рис.6.6) сделайте предустановки, указанные на рис. (если они не сделаны ранее)

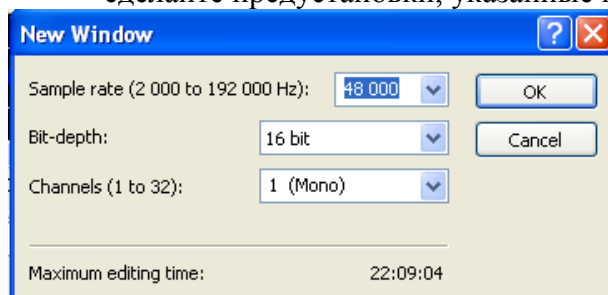


Рис. 6.6. Окно предустановок формата записи

- В меню **Tools** программы **SF** выберите опцию **Synthesis**, откроется еще одно окно, в нем выберите опцию **FM**. Тогда раскроется окно синтезатора (рис.6.7), сделайте в нем предустановки, указанные на рисунке (если они не сделаны ранее): белый шум, уровень минус 18 дБ, длительность 50 с.

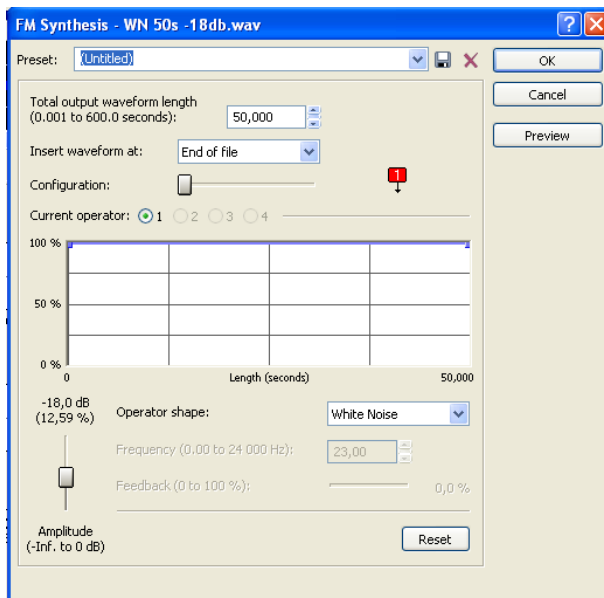


Рис.6.7. Окно синтезатора тестовых сигналов

- В меню **File** программы **SF** выберите опцию **Save As...** В открывшемся окне (рис.6.8.) напишите название и выберите указанный на рисунке тип файла. Сохраните его. Это имя появится на сигналограмме в окне **SF**

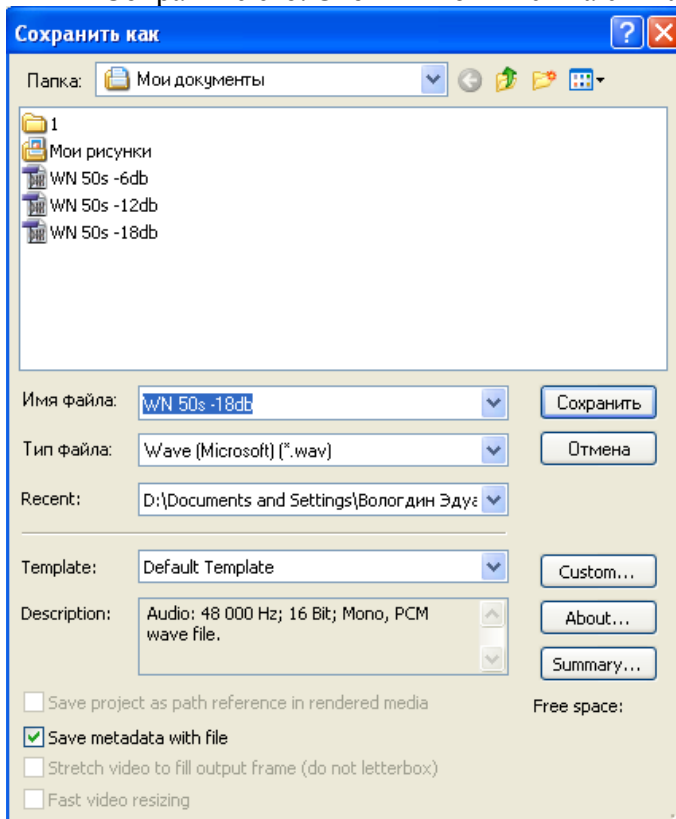


Рис.6.8.. Окно сохранения файла (имя, тип)

- Курсором активизируйте окно программы **Sound Forge 9c**
- В окне **SF** курсором выделите сигналограмму файла **WN50s-18db.wav**, ее фон должен стать черным (рис.6.9)

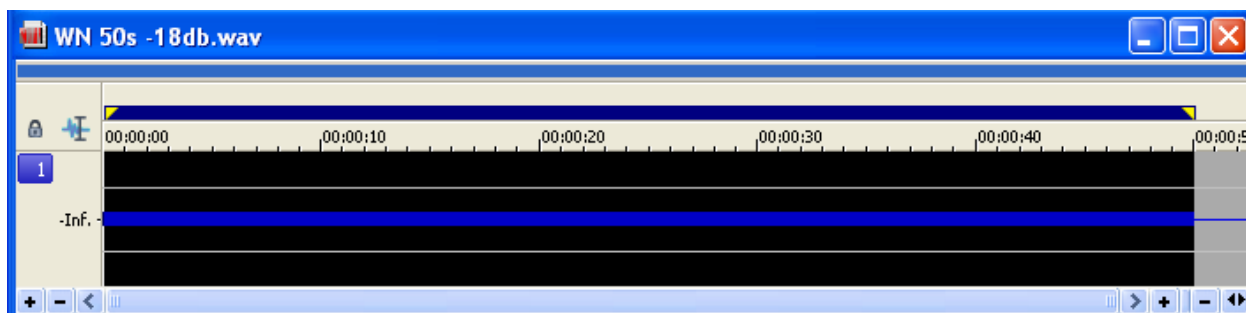


Рис.6.9. Сигналограмма выделенного тестового сигнала

5.3. Исследование частотных характеристик шельфовых НЧ фильтров

Для шельфового НЧ фильтра надо измерить 4 частотных характеристики: в режимах усиления на 20 dB (**Amount +20 dB**) и ослабления на 20 дБ (**Amount -20 dB**), частота среза 500-600 Гц. Эти измерения производятся при двух значениях ширины переходной зоны: **Transition Width (WT)**- шириной 0,1 и 5 октав.

5.3.1. Измерение АЧХ фильтра НЧ в режиме усиления с TW=0.1 октавы

- Курсором активизируйте окно программы **Sound Forge 9c**
- В меню программы **SF Process** откройте окно параметрического эквалайзера. Для этого курсором нажмите сначала **EQ** и затем **Paramet...** По этим командам раскроется окно параметрического эквалайзера (рис.6.10)

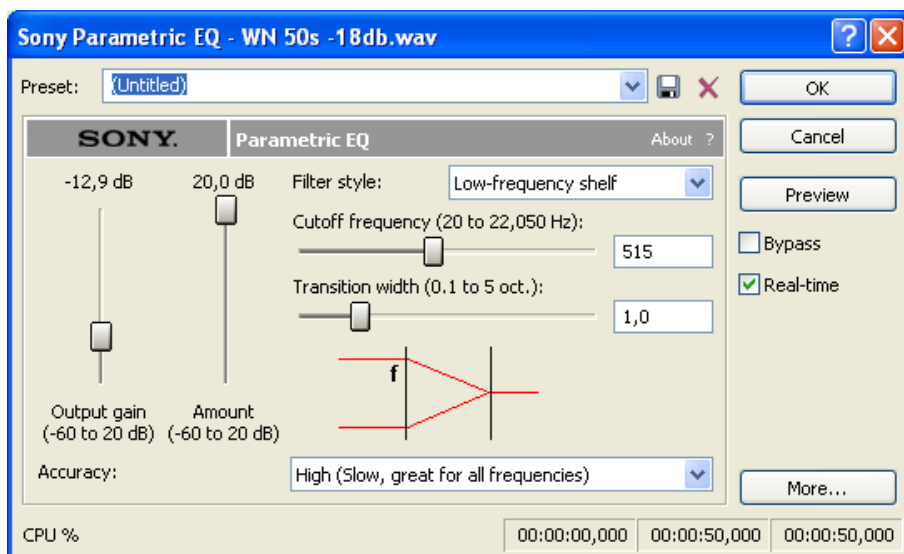





Рис.6.10. Окно параметрического эквалайзера


- В окне эквалайзера установить:
Filter Style – Low-frequency shelf (Шельфовый НЧ фильтр), **Accuracy – High (Slow.....)**, **Output gain** минус 12 dB., **Cutoff frequency** - 500...600 Гц.
- Переключите фильтр в режим усиления. Для этого установите курсором уровень **Amount +20 dB**
- Установите **TW = 0,1 октавы**
- Нажмите кнопку **OK**. После этого в верхнем окне программы **SF** появится новый файл под названием **WN50s-18db.wav*** -это выходной файл фильтра.
- Курсором выделите этот файл от начала до конца, фон фонограммы должен стать черным

- Для измерения частотной характеристики фонограммы этого файла нажмите кнопку с пиктограммой 
- Сохраните график АЧХ в 1 ячейке памяти. Для этого последовательно нажмите кнопки **Set** и 1.
- Замените файл **WN50s-18db.wav***- исходным файлом **WN50s-18db.wav**. Для этого нажмите кнопку с пиктограммой 

5.3.2. Измерение АЧХ фильтра в режиме усиления с $TW=5$ октав

- Повторите измерения по п. 5.3.1. , установив $TW = 5$ октав
- Сохраните график АЧХ в 2 ячейке памяти. Для этого последовательно нажмите кнопки **Set** и 2.
- Замените файл **WN50s-18db.wav***- исходным файлом **WN50s-18db.wav**. Для этого нажмите кнопку с пиктограммой 

5.3.4. Измерение АЧХ НЧ фильтра в режиме ослабления с $TW=0,1$ октавы

- Повторите измерения по п. 5.3.1. , переключив фильтр в режим ослабления. Для этого установите курсором уровень **Amount -20 dB** , $TW = 0,1$ октавы
- Сохраните график АЧХ в 3 ячейке памяти. Для этого последовательно нажмите кнопки **Set** и 3.
- Замените файл **WN50s-18db.wav***- исходным файлом **WN50s-18db.wav**. Для этого нажмите кнопку с пиктограммой 

5.3.5. Измерение АЧХ НЧ фильтра в режиме ослабления с $TW = 5$ октав

- Повторите измерения по п. 5.3.2 , переключив фильтр в режим ослабления. Для этого установите курсором уровень **Amount -20 dB**, $TW= 5$ октав
- Сохраните график АЧХ в 4 ячейке памяти. Для этого последовательно нажмите кнопки **Set** и 4.
- Скопируйте все 4 графика АЧХ шельфового фильтра НЧ в режимах усиления и ослабления (рис.6.11) в Ваш отчет.

Для этого активизируйте окно программы **HyperSnap-DX** и выберите в его меню **Capture** команду **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу графика спектра нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол графика спектра и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** в меню **File** выберите **Save As**. Установите путь копирования на Вашу флешкарту, дайте название файлу и сохраните его в своем отчете.

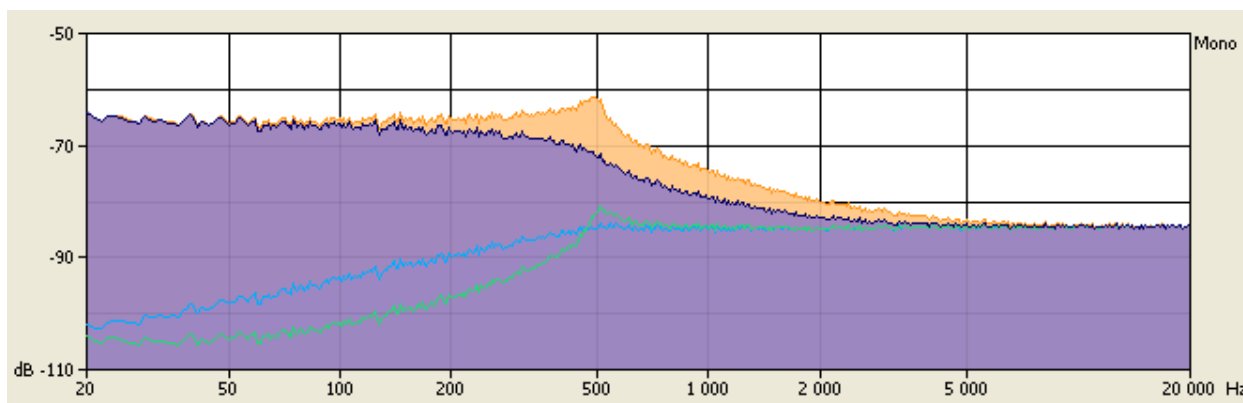


Рис. 6.11. Частотные характеристики шельфового фильтра низких частот

- С помощью курсора измерьте уровень АЧХ на частоте среза и на граничных частотах 20 и 20000 гц. Результаты измерений занесите в табл.5.1.


- Замените файл **WN50s-18db.wav***- исходным файлом **WN50s-18db.wav**. Для этого нажмите кнопку с пиктограммой 

Табл. 6.1 Характеристики шельфового НЧ фильтра					
Режим усиления +20 дБ					
Amount, dB	TW, октава	Усиление, дБ	L dB (Cutoff)	L dB (20 Hz)	L dB (20 kHz)
+20	0,1				
	5				
Режим ослабления -20 дБ					
-20	0,1				
	5				

5.4. Исследование частотных характеристик шельфовых НЧ фильтров

Для шельфового ВЧ фильтра надо измерить 4 частотных характеристики: в режимах усиления на 20 дБ (**Amount +20 dB**) и ослабления на 20 дБ (**Amount -20 dB**), частота среза 2000-2500 Гц. Эти измерения производятся при двух значениях ширины переходной зоны **Transition Width (WT)**- шириной 0,1 и 5 октав.

5.4.1. Частотные характеристики шельфовых ВЧ фильтров

- Повторите все измерения по п .5.3.1- 5.3.4 применительно к шлейфовым ВЧ фильтрам с частотой среза 2000 Гц. Результаты измерений характеристик фильтров занесите в табл.5.2

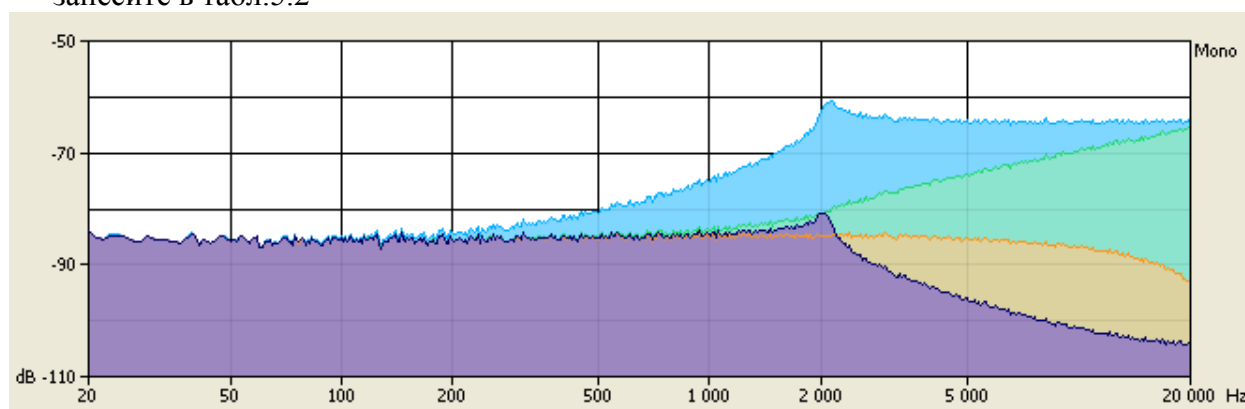


Рис. 6.12 Частотные характеристики шельфового фильтра высоких частот

- Скопируйте все 4 графики АЧХ шельфового фильтра ВЧ в режимах усиления и ослабления (**рис.6.12**) в Ваш отчет.

Для этого активизируйте окно программы **HyperSnap-DX** и выберите в его меню **Capture** команду **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу графика спектра нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол графика спектра и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** в меню **File** выберите **Save As**. Установите путь копирования на Вашу флешку, дайте название файлу и сохраните его в своем отчете.


- С помощью курсора измерьте уровень АЧХ на частоте среза и на граничных частотах 20 и 20000 гц. Результаты измерений занесите в табл.6.2.
- Замените файл **WN50s-18db.wav***- исходным файлом **WN50s-18db.wav**. Для этого нажмите кнопку с пиктограммой 

Табл.6. 2 Характеристики шельфового ВЧ фильтра					
Режим усиления +20 дБ					
Amount, dB	TW, октава	Усиление, дБ	L dB (Cutoff)	L dB (20 Hz)	L dB (20 kHz)
+20	0,1				
	5				
Режим ослабления -20 дБ					
-20	0,1				
	5				

6. Контрольные вопросы

1. Назначение шельфовых фильтров
2. Форма графиков шельфовых фильтров ВЧ и НЧ
3. Как определяется частота среза шельфовых фильтров
4. Как изменяется форма АЧХ шельфовых фильтров при изменении частоты среза
5. Как изменяется форма АЧХ шельфовых фильтров при изменении усиления
6. Как меняется форма АЧХ при изменении порядка фильтра
7. В каких пределах можно изменять усиление шельфовых фильтров
8. В отличие АЧХ реальных шельфовых ВЧ и НЧ фильтров от теоретических

7. Содержание отчета

- 7.1. Титульный лист
- 7.2. Цель работы
- 7.3. Задание к лабораторной работе
- 7.4. Функциональная схема измерений
- 7.5. Результаты измерений в виде графиков
- 7.6. Таблицы с результатами измерений и расчетов
- 7.7. Выводы по результатам измерений

8. Литература

Вологдин Э.И. Методы и алгоритмы обработки звуковых сигналов. Учебное пособие. Часть 1 . СПб ГУТ. 2009.

ЛАБОРАТОРНАЯ РАБОТА N 7 АТ

ИССЛЕДОВАНИЕ ЧАСТОТНЫХ ХАРАКТЕРИСТИК КОДЕКОВ ФОРМАТОВ MPEG-1 Layer 2 и Layer 3.

Руководство к выполнению лабораторной работы по Аудиотехнике

1. Цель работы

Целью лабораторной работы является исследование с помощью компьютерной программы Sound Forge 9с зависимости частотных характеристик кодеков форматов MPEG-1 Layer 2 и Layer 3 от уровня тестовых сигналов белого шума .

2. Задание к лабораторной работе

- Ознакомиться с программой Sound Forge 9с в объеме, необходимом для конвертирования тестовых файлов белого шума из формата Wave в форматы MPEG-1 Layer 2 и Layer 3
- Ознакомиться с программой Sound Forge 9с в объеме, необходимом для измерения частотной характеристики кодеков в с помощью тестового сигнала белого шума.
- Измерить спектр тестового сигнала белого шума в формате Wave с уровнем минус 6 дБ.
- Конвертировать тестовые сигналы белого шума с уровнями минус 6, 20, 40 , 60 и 80 дБ из формата Wave в формат MPEG-1 Layer 2
- Конвертировать тестовые сигналы белого шума с уровнями минус 6, 20, 40 , 60 и 80 дБ из формата Wave в формат MPEG-1 Layer 3
- Измерить частотные характеристики кодека формата MPEG-1 Layer 2 при уровнях тестового сигнала минус 6, 20, 40 , 60 и 80 дБ
- Измерить частотные характеристики кодека формата MPEG-1 Layer 3 при уровнях тестового сигнала минус 6, 20, 40 , 60 и 80 дБ
- Рассчитать скорость цифрового потока и коэффициент компрессии кодеков форматов MPEG-1 Layer 2 и Layer 3
- По результатам измерений сформулировать выводы

3. Функциональная схема измерений

На рис.7.1. приведена функциональная схема измерения частотных характеристик кодеков различного формата с помощью компьютерной программы Sound Forge 9с. В ЗУ хранятся записанные заранее файлы в формате WAVE тестовых сигналов белого шума с равномерной спектральной плотностью от 20 до 20000 кГц и уровнями минус 6, 20, 40, 60 и 80 дБ. Все эти файлы последовательно конвертируются в

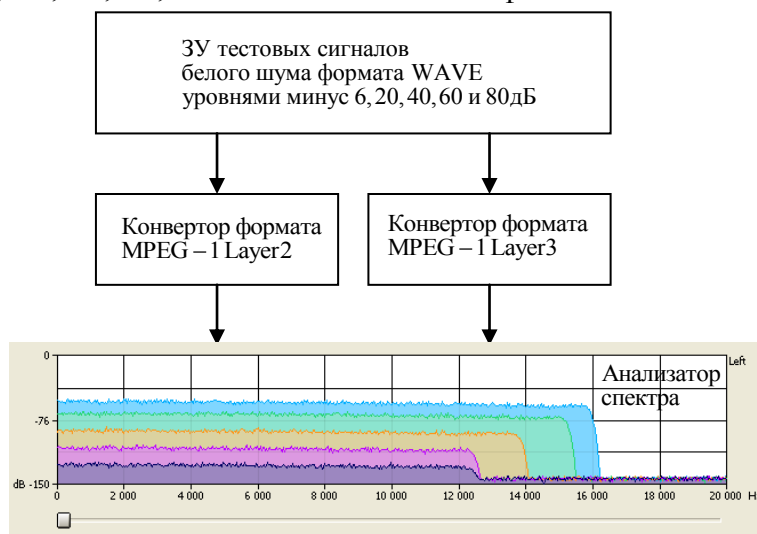


Рис.7.1. Функциональная схема измерения частотных характеристик кодеков

форматы MPEG-1 Layer 2 и Layer 3 и сохраняются в памяти компьютера. Сначала производится измерение частотных характеристик фонограмм файлов с расширением .mpg на одном графике. Затем производится измерение частотных характеристик фонограмм файлов с расширением .mp3 на другом графике.

4. Методические указания

4.1. Общие сведения

В настоящее время в цифровой аудиотехнике для записи и передачи звуковых сигналов по каналам связи используются два вида аудио систем: без компрессирования и с компрессированием звуковых сигналов. Студийная запись производится всегда без компрессирования с использованием линейных методов аналого-цифрового преобразования на основе импульсно-кодовой или сигма-дельта модуляции. Необходимость компрессирования возникает, прежде всего, при передаче звуковых сигналов по каналам связи и при необходимости увеличения длительности многоканальной звукозаписи на дисковых магнитных и оптических носителях.

В цифровых системах компрессирования основной задачей является сокращение скорости цифрового потока при записи или передачи звуковых сигналов. Такие системы с компрессированием по принципу работы разделяются на две группы:

- *перцептуальное* кодирование,
- *статистическое* кодирование.

Перцептуальное кодирование всегда осуществляется с потерями, так как при кодировании сокращение скорости цифрового потока достигается благодаря устранению информационной избыточности звуковых сигналов. При этом убирается информация, которую слух человека не воспринимает. Избыточность эта огромна, поэтому возможно сокращение скорости цифрового потока до 20...40 раз без заметных на слух искажений.

Статистическое кодирование в принципе осуществляется без потерь за счет сокращения статистической избыточности звуковых сигналов и применения оптимальной с точки зрения теории информации упаковки цифровых данных. Избыточность эта не так

велика, поэтому сокращение скорости цифрового потока возможно не более чем 2...4 раза.

В настоящей лабораторной работе исследуются две системы перцептуального компрессирования звуковых потоков по стандарту MPEG-1 Layer 2 и Layer 3 (MP3). Первая используется в цифровом радиовещании по стандарту DRM под названием Musicam, вторая, более известна под названием MP3, широко применяется для записи и передачи в сети Internet музыкальных файлов для любителей музыки.

4.2. Перцептуальные системы компрессирования звука

Основой перцептуальных систем компрессирования звуковых сигналов является применение субполосного кодирования, управляемого блоком, моделирующим психоакустические особенности слуха человека (слуховая модель). Упрощенная модель кодека формата MPEG-1 Layer 2 приведена на [рис.7.2](#). В этой схеме субполосная фильтрация при кодировании и реконструкции звуковых сигналов производится с помощью одинакового набора полосовых фильтров на передающей и приемной сторонах.

В приведенной на [рис.7.2](#) схеме субполосного кодека аналоговый сигнал с помощью линейной ИКМ преобразуется в цифровую последовательность с частотой дискретизации $f_s = 48$ кГц, которая организуется в блоки по 1152 выборки. С помощью набора из цифровых фильтров спектр звукового сигнала от 0 до частоты Найквиста, разделяется на $n = 32$ полосы шириной $\Delta F = 750$ Гц

В каждой полосе кодера с помощью блоков S (дискретизатор), Q (квантователь) и ВС (кодер) осуществляется дискретизация с понижением частоты в n раз, квантование и двоичное кодирование цифровых сигналов. Результаты кодирования входного блока данных во всех полосах мультиплексируются в один выходной блок данных, с которого

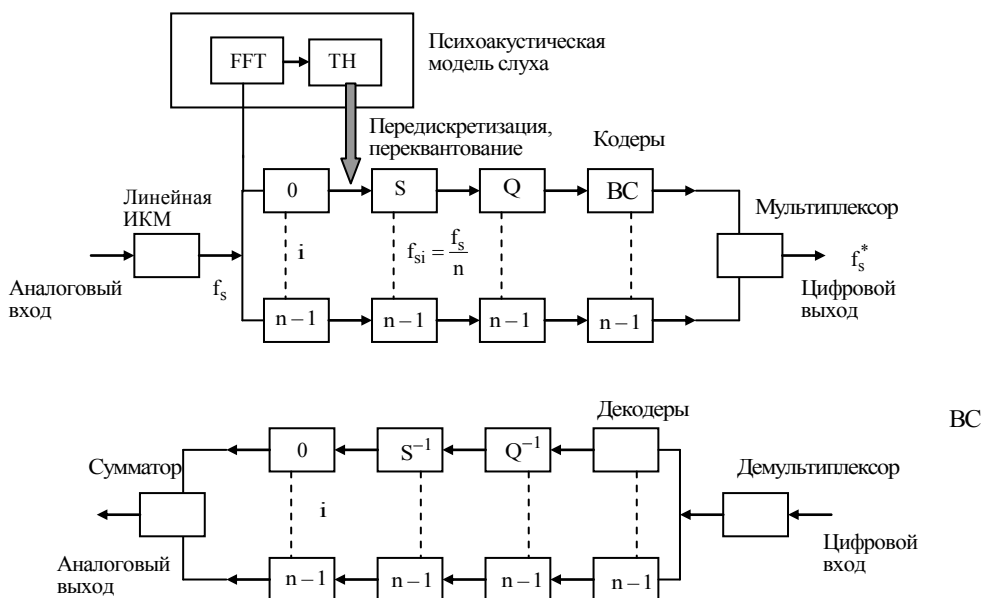


Рис.7.2. Схема кодека формата MPEG-1 Layer 2

передача происходит с частотой дискретизации f_s^* . Эта частота устанавливается кодером автоматически так, чтобы она не превышала половины скорости цифрового потока. При скорости 32 кбит/с $f_s^* = 16$ кбит/с. Она может принимать только дискретные значения из стандартного ряда.

В декодере, прежде всего, осуществляется демультимплексирование, при котором восстанавливаются цифровые данные в отдельных полосах. В схеме на [рис.7.2](#) блоки S^{-1} , Q^{-1} и декодеры в каждой полосе выполняют функцию цифроаналогового

преобразователя. Исходный аналоговый звуковой сигнал формируется на выходе сумматора.

При субполосном кодировании осуществляется перенос звукового сигнала из временной в спектральную область и при этом кодируется не сигнал, а его спектр. Чем больше спектральных полос n , тем точнее осуществляется это преобразование и тем выше может быть компрессия. В субполосном декодере происходит обратное преобразование сигналов из спектральной области во временную область.

Субполосное кодирование позволяет в каждой полосе использовать минимальное число разрядов, вплоть до нуля, обеспечивающих необходимое отношение сигнал/шум. Максимальная компрессия в таком кодере достигается в случае кодирования тональных сигналов, так как при этом сигнал присутствует только в одной полосе и поэтому скорость цифрового потока сокращается в n раз.

Компрессия минимальна при кодировании сигналов белого шума. В этом случае уровень сигнала во всех полосах одинаковый и определяется спектральной плотностью шума, которая на 43 дБ ниже уровня сигнала на входе кодера. Это значит, что в каждой полосе число используемых разрядов может быть сокращено на 7 бит.

Выбор частоты дискретизации при кодировании в частотной области основывается на том, что максимальная частота спектра огибающей звукового сигнала в функции времени F_{\max} и ширина полосы пропускания фильтра ΔF связаны неравенством

$$\frac{F_{\max}}{\Delta F} \leq 1$$

Из этого неравенства следует, что частота дискретизации в частотной области f_{si} должна, по крайней мере, вдвое превышать полосу фильтрации

$$f_{si} \geq 2\Delta F_i$$

Так, например, если звуковой диапазон от 0 до частоты Найквиста разбивается на 16 полос шириной по 1500 Гц, то минимальная частота дискретизации звуковых сигналов во всех полосах может быть равна 3 кГц.

Перцептуальное кодирование на [рис.7.2](#) осуществляется с помощью блока вычисления порога слышимости. Он включает в себя анализатор спектра FFT и вычислительный блок, рассчитывающий трехмерный (амплитуда, частота и время) так называемый глобальный порог слышимости ТН. Сигналы с этого блока управляют коэффициентом передачи фильтров субполосного кодера таким образом, что в полосах, в которых уровень сигналов меньше порога слышимости, далее по тракту не передаются. Благодаря этому и достигается значительное сокращение скорости цифрового потока.

Глобальный порог слышимости определяется абсолютным порогом и эффектами маскировки тона тоном, тона шумом или шума тоном. В случае использования тестовых тональных и шумовых сигналов глобальный порог слышимости тождественен абсолютному порогу, и он определяет частотную характеристику кодека, поэтому она и зависит от уровня тестовых сигналов.

На [рис. 7.3](#) приведен стандартный график абсолютного порога слышимости, из которого видно, что этот порог при малых уровнях звука приводит к достаточно сильному уменьшению полосы пропускания компрессируемого сигнала. Так при уровне звука 80 dB SPL слышимый диапазон простирается от 20 до 20000 Гц, а при его уровне 10 dB SPL этот диапазон равен 250...10000 Гц. Динамический диапазон сигналов (DR) в стандарте MPEG-1 рассчитывается в шкале FS, в которой за 0 дБ принят уровень звукового давления 96 дБ SPL.

В формате MPEG-1 Layer 2 используется линейное квантование и линейное двоичное кодирование с итерацией. В формате MPEG-1 Layer 3 используется нелинейное квантование и энтропийное кодирование с итерацией на основе кода Хаффмана,

благодаря чему достигается значительное увеличение коэффициента компрессии. Кроме того, на выходе каждого полосового фильтра кодера включен 18-полосный анализатор спектра FFT, благодаря чему число спектральных линий увеличивается до 576, что также

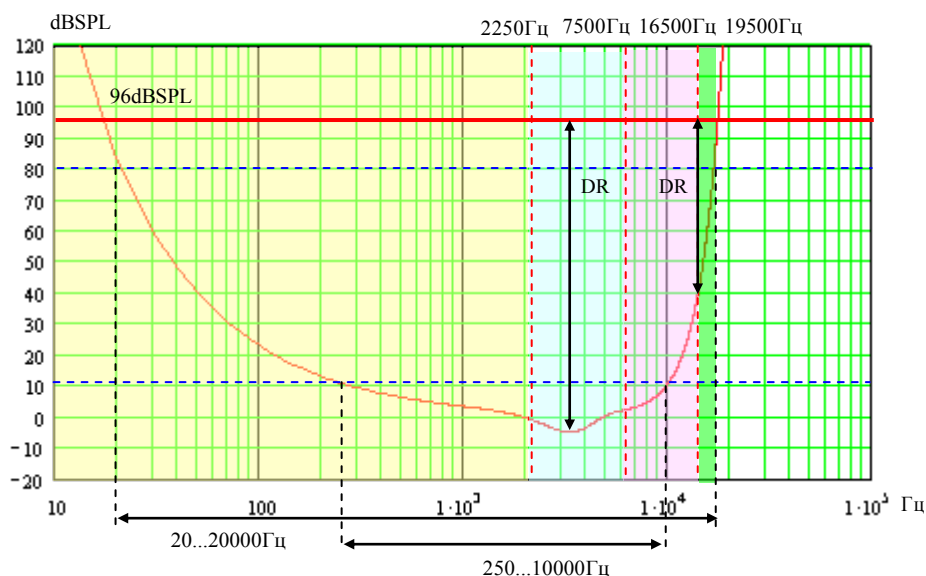


Рис.7.3.. К определению слышимой полосы частот

приводит к повышению эффективности компрессирования. Благодаря этому в формате MP3 перцептуальное кодирование можно не использовать и тогда полоса пропускания кодера может быть расширена до 20 кГц независимо от уровня сигнала.

4.3. Измерение частотной характеристики цифровых трактов

Измерение частотной характеристики кодеков форматов MPEG-1 Layer 2 и Layer 3 (MP3) возможно только с помощью программного пакета Sound Forge 9c. Для этого используется тестовый сигнал белого шума в формате Wave с равномерной спектральной плотностью от 20 до 20000 Гц в линейной шкале частот и уровнем от минус 6 до минус 80 дБ.

5. Порядок выполнения лабораторной работы

5.1. Подготовка к измерениям

- Скопируйте с рабочего стола файл «Отчет LR-7 АТ» на свою флешку, дайте ему свое имя в виде: «Группа»_ «Фамилия (работавшего на клавиатуре)»_ «Номер ЛР»_ «Дата».
- Откройте этот файл и сверните его окно.

В файл этого отчета вы будете копировать все экспериментальные графики, и заносить в таблицы отчета результаты измерений сразу в процессе их получения. В день выполнения лабораторной работы отчет в электронном виде необходимо сдать преподавателю, только тогда работа считается выполненной.



- Кнопкой с пиктограммой откройте окно программы **Sound Forge 9c (SF)**, содержащее окно стерео сигналограммы и окно спектров левого и правого звуковых каналов. Не меняйте размеры этого окна, они согласованы с Вашим отчетом




- Нажмите кнопку с пиктограммой и в появившемся меню:



Рис.7.4. Меню анализатора спектра

выберите пункты, указанные на рис. 7.4. При этом будет строиться только график спектра левого канала.

- Нажмите кнопку с пиктограммой . По этой команде откроется окно предустановок (рис.7.5.), их следует установить такими же как на рисунке. Обязательно используется линейная шкала частот

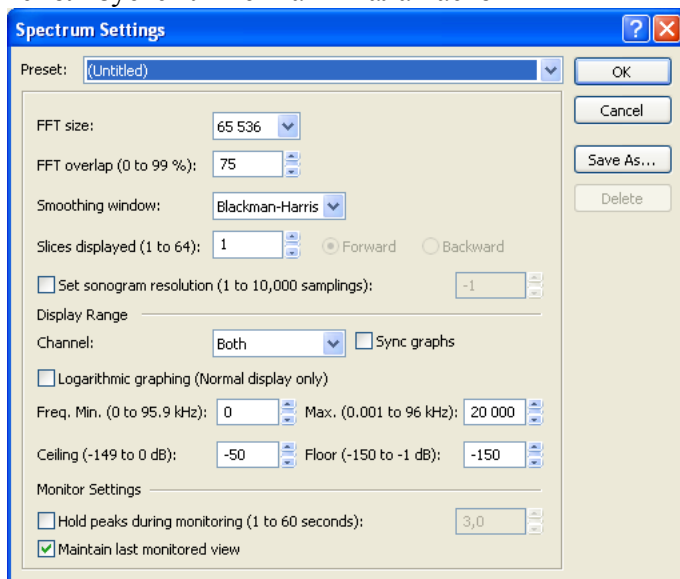




Рис.7.5. Окно предустановок анализатора спектра

- Графики спектров строятся в диапазоне 0...20 кГц в диапазоне уровней 0...-150 дБ, FFT size = 65536 (предустановки).



- Кнопкой с пиктограммой  откройте окно программы **HyperSnap-DX**, предназначенной для вырезки и копирования графических изображений.



- Кнопкой с пиктограммой  откройте окно программы **Total Commander**, на диске C с названием «Тест сигналы ЛР-7» и скопируйте все файлы папки в раздел «Мои документы». В этой папке записаны тестовые сигналы белого шума длительностью 60 секунд каждый с уровнями: - 6, -20, -40, -60 и -80 дБ без компрессирования в формате .wav, частота дискретизации 48 кГц, число разрядов 16. Файлы имеют названия: Т «уровень» (в дБ)

5.2. Измерение спектра тестовых сигналов в формате Wave

- В меню **File** программы **SF** Нажмите кнопку **Open**. По этой команде на экране ПК появится окно, содержащее файлы папки «Мои документы» Курсором выберите в ней тест **T6.wav** и нажмите кнопку **Открыть** (Open). По этой команде в окне **SF** появится спектрограмма тестового сигнала (рис.7.6)

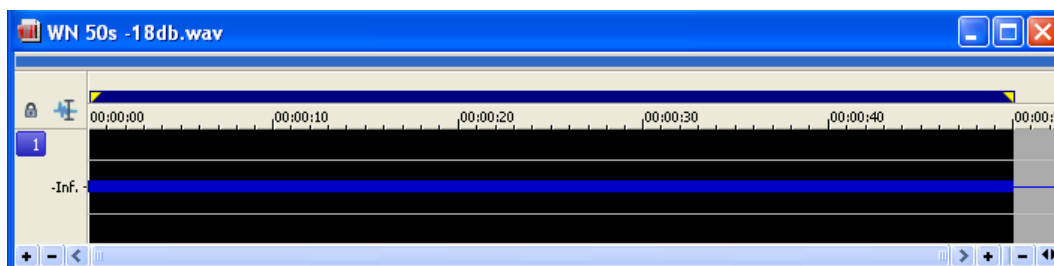



Рис.7.6. Выделенная фонограмма тестового сигнала

- Курсором выделите всю фонограмму от начала до конца, при этом ее фон становится черным
- Нажмите кнопку в окне спектрального анализа с пиктограммой . Через 10...60 с в окне появятся графики спектров левого и правого тестовых сигналов. Чтобы сохранить графики последовательно нажмите кнопки **Set** и **1** (сохранение в первой ячейке памяти, всего 4 ячейки)
- Скопируйте график спектра левого канала (рис.7.7) в Ваш отчет.

Для этого раскройте окно программы **HyperSnap-DX** и выберите в его меню **Capture** команду **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу графика спектра нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол графика спектра и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** в меню **File** выберите **Save As**. Установите путь копирования на Вашу флешкарту, дайте название файлу и сохраните его в своем отчете

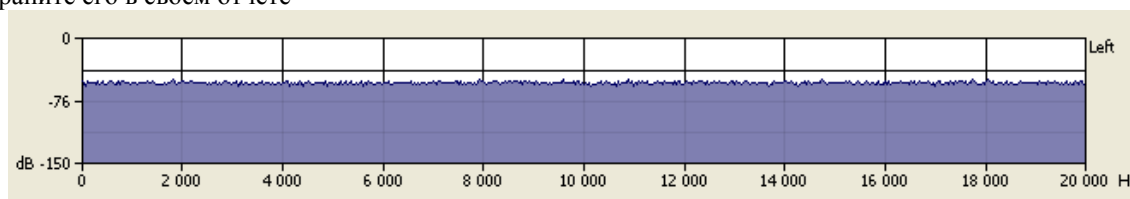


Рис.7.7. График спектра тестового широкополосного сигнала белого шума WN stereo 60 s -6dB в формате Wave (без компрессирования)

5.3. Преобразование тестовых сигналов из формата Wave в формат MPEG-1 layer

2

5.3.1.

- В меню **File** программы **SF** выберите команду **Open**. По этой команде на экране ПК появится окно, содержащее файлы папки «Мои документы» Курсором выберите в ней тест **T6.wav** и нажмите кнопку **Открыть** (Open).
- В меню **File** программы **SF** выберите команду **Save As**. По этой команде на экране ПК появится окно **Сохранить Как**. В этом окне установите тип сохраняемого файла **MPEG-1** (mpg,) и нажмите кнопку **Сохранить**

5.3.2.

Повторите измерения по п.5.3.1. для уровней тестовых сигналов -20, -40, -60 и -80 дБ

5.4. Преобразование тестовых сигналов из формата Wave в форматы MPEG-1 layer 3 (MP3)

5.4.1.

- В меню **File** программы **SF** выберите команду **Open**. По этой команде на экране ПК появится окно, содержащее файлы папки «Мои документы» Курсором выберите в ней тест **T6.wav** и нажмите кнопку **Открыть** (Open).

- В меню **File** программы **SF** выберите команду **Save As**. По этой команде на экране ПК появится окно **Сохранить Как**. В этом окне установите тип сохраняемого файла **MP3 Audio (mp3,)** и нажмите кнопку **Сохранить**

5.4.2.

Повторите измерения по п.5.4.1. для уровней тестовых сигналов –20, -40, -60 и –80 дБ

5.5. Измерение спектра тестовых сигналов белого шума различного уровня в формате MPEG-1 layer 2

5.5.1.


- В меню **File** программы **SF** нажмите кнопку **Open**. По этой команде на экране ПК появится окно, содержащее файлы папки «Мои документы» Курсором выберите в ней тест **T6.mpg** с расширением .mpg и нажмите кнопку **Открыть** (Open).
- Курсором выделите всю фонограмму от начала до конца, при этом ее фон становится черным
- Для измерения частотной характеристики кодека, в окне спектрального анализа нажмите кнопку с пиктограммой . Через 10...60 с в окне появится график спектра левого канала тестового сигнала. График этого спектра сохраните в ячейке памяти 1. Для этого последовательно нажмите кнопки **Set** и 1.
- По графику с помощью курсора определите верхнюю границу полосы пропускания, и результат занесите в **табл.7.1.**

Табл.7.1. Верхние границы полосы пропускания		
Формат файлов	Уровень сигнала, дБ	Верхняя граница, кГц
MPEG –1 Layer 2	-6	
	-20	
	-40	
	-60	
	-80	
MPEG –1 Layer 3	-6	
	-20	
	-40	
	-60	
	-80	


5.5.2.

Повторите измерения по п.5.5.1. для уровней тестовых сигналов –20, -40, -60 и –80 дБ с расширением .mpg, сохраняя их последовательно в ячейках памяти 2, 3 и 4. Сигнал с уровнем –80 дБ не сохраняется

- Скопируйте графики спектров левого канала(рис.7.8) в Ваш отчет.

Для этого раскройте окно программы **HyperSnap-DX** и выберите в его меню **Capture** команду **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу графика спектра, нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол графика спектра и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** в меню **File** выберите **Save As**. Установите путь копирования на Вашу флешкарту, дайте название файлу и сохраните его в своем отчете.

- Сотрите все графики спектров из памяти

Для этого нажмите кнопку с пиктограммой 

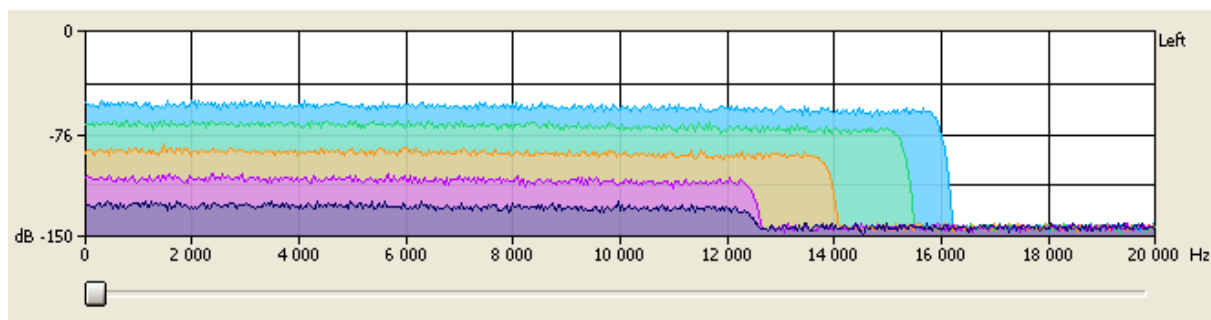



Рис.7.8. Графики частотных характеристик кодека формата MPEG-1 Layer 2 (уровни тестовых сигналов белого шума: -6, -20, -40, -60 и 80 дБ)

5.6. Измерение спектра тестовых сигналов белого шума различного уровня в формате MPEG-1 layer 3 (MP3)

5.6.1.

- В меню **File** программы **SF** нажмите кнопку **Open**. По этой команде на экране ПК появится окно, содержащее файлы папки «Мои документы». Курсором выберите в ней тест **T6.mp3** с расширением **.mp3** и нажмите кнопку **Открыть** (Open).
- Курсором выделите всю фонограмму от начала до конца, при этом ее фон становится черным
- Для измерения частотной характеристики кодека в окне спектрального анализа нажмите кнопку с пиктограммой . Через 10...60 с в окне появится график спектра левого канала тестового сигнала. График этого спектра сохраните в ячейке памяти 1. Для этого последовательно нажмите кнопки **Set** и 1.
- По графику определите верхнюю границу полосы пропускания, и результат занесите в [табл.7.1](#).

5.6.2.

Повторите измерения по п.5.5.1. для уровней тестовых сигналов -20, -40, -60 и -80 дБ с расширением **.mp3**, сохраняя их последовательно в ячейках памяти 2, 3 и 4. Сигнал с уровнем -80 дБ не сохраняется

- Скопируйте графики спектров левого канала([рис.7.9.](#)) в Ваш отчет.

Для этого раскройте окно программы **HyperSnap-DX** и выберите в его меню **Capture** команду **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу графика спектра нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол графика спектра и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** в меню **File** выберите **Save As**. Установите путь копирования на Вашу флешку, дайте название файлу и сохраните его в своем отчете

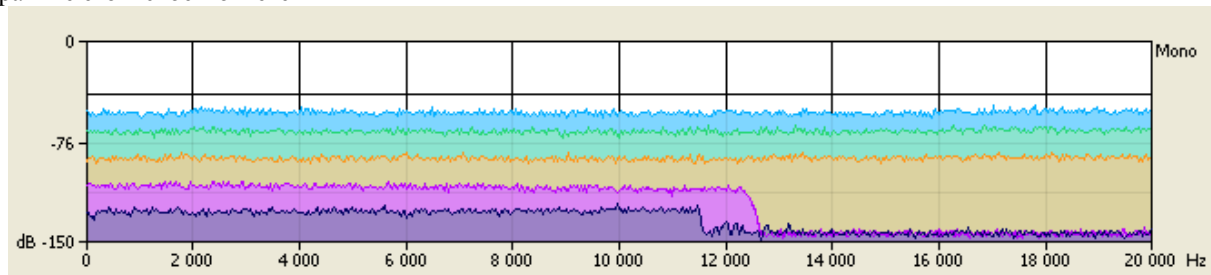


Рис.7.9. Графики частотных характеристик кодека формата MPEG-1 Layer 3 (MP3) (уровни тестовых сигналов белого шума: -6, -20, -40, -60 и 80 дБ)

5.7. Расчет скорости цифровых потоков и коэффициента компрессии

- Из папки «Мои документы» внесите в **табл.7.2** данные по размеру указанных файлов в байтах
- Исходя из длительности всех файлов 60 с рассчитайте скорости цифровых потоков в кбит/с, данные занесите в табл.2
- Рассчитайте коэффициенты компрессии для форматов MPEG-1 Layer 2 и MP3, сравните результаты, как результат деления входной скорости цифрового потока на выходную скорость кодера
- Сравните характеристики кодеков: скорость цифрового потока, коэффициент компрессии и ширина полосы пропускания

Табл.7.2. Расчет скорости цифровых потоков и коэффициента компрессии				
Формат	Уровень, дБ	Размер, байт	Скорость, кбит/с	Кэф. компрессии
Wave	-6			
MPEG-1 Layer 2	-6			
	-20			
	-40			
	-60			
	-80			
MP3	-6			
	-20			
	-40			
	-60			
	-80			

6.Контрольные вопросы

1. В чем различие кодеков MPEG-1 Layer 2 и MPEG-1 Layer 3
2. Как зависит частотная характеристика кодера MPEG Layer 2 от уровня тестового сигнала
3. Как зависит частотная характеристика кодера MP3 Audio от уровня тестового сигнала
4. Почему частотная характеристика кодеков MPEG-1 Layer 2 и MP3 может меняться от уровня тестового сигнала
5. Как исключить зависимость частотной характеристики кодеков MPEG-1 от уровня тестовых сигналов
6. Как рассчитать скорость цифрового потока
7. Как рассчитать коэффициент компрессии
8. Каким образом субполосное кодирование позволяет сокращать скорость цифрового потока
9. В чем сущность перцептуальных методов компрессирования звука

7. Содержание отчета

- 7.1. Титульный лист
- 7.2. Цель работы
- 7.3. Задание к лабораторной работе
- 7.4. Функциональная схема измерений
- 7.5. Результаты измерений в виде графиков
- 7.6. Таблицы с результатами измерений и расчетов
- 7.7. Выводы по результатам измерений

8. Литература

Ю.А. Ковалгин, Э.И. Вологдин Цифровое кодирование звуковых сигналов. СПб, Корона Принт, 2004.

ЛАБОРАТОРНАЯ РАБОТА N 8 АТ

ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИК КОДЕРА MP3 Lame В РЕЖИМАХ РАБОТЫ CBR, ABR И VBR

Руководство к выполнению лабораторной работы по Аудиотехнике

Цель работы

Целью лабораторной работы является исследование зависимости характеристик кодера MP3 Lame 3.98.4 от выбора скорости цифрового потока и режимов кодирования CBR, ABR и VBR с помощью тестовых сигналов белого шума с использованием компьютерной программы Sound Forge 9с. В режиме кодирования CBR скорость цифрового потока постоянная, в режиме ABR она равна среднему значению, а в режиме VBR скорость цифрового потока переменная, она зависит от музыкального материала и заданного значения качества от 0 до 9.

2. Задание к лабораторной работе

- Ознакомиться с программой Lame 3.98.4 в объеме, необходимом для конвертирования тестовых файлов белого шума из формата Wave в форматы MP3 CBR, ABR и VBR
- Ознакомиться с программой Sound Forge 9с в объеме, необходимом для измерения частотной характеристики кодеков MP3 с помощью тестового сигнала белого шума.
- Конвертировать тестовые сигналы белого шума с уровнем минус 6 дБ из формата Wave в формат MP3 CBR со скоростью выходного цифрового потока 256, 128, 64 и 32 кбит/с
- Конвертировать тестовые сигналы белого шума с уровнями минус 6, 40, 60 и 80 дБ из формата Wave в формат MP3 CBR со скоростью выходного цифрового потока 256 кбит/с
- Конвертировать тестовые сигналы белого шума с уровнем минус 20 из формата Wave в форматы MP3 CBR, ABR и VBR при V=0 и V=9
- Измерить частотные характеристики кодера формата Wave в формате MP3 CBR со скоростью выходного цифрового потока 256, 128, 64 и 32 кбит/с при уровне входного сигнала минус 6 дБ
- Измерить частотные характеристики кодера формата Wave в формате MP3 CBR при скорости выходного цифрового потока 256 кбит/с и уровне входного сигнала минус 6, 40, 60 и 80 дБ
- Рассчитать скорость цифрового потока и коэффициент компрессии кодеров форматов MP3 CBR, ABR и VBR при уровне тестового сигнала минус 20 дБ
- По результатам измерений сформулировать выводы

3. Функциональная схема измерений

На рис.8.1. приведена функциональная схема конвертирования файлов формата Wave в форматы MP3 CBR, ABR и VBR с помощью программы Lame 3,98.4 и

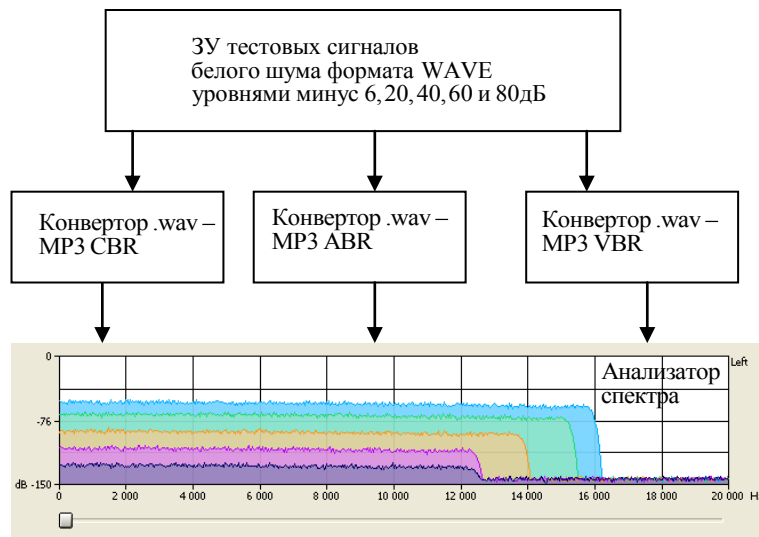


Рис.8.1. Функциональная схема измерения частотных характеристик кодеков

измерения частотных характеристик кодеков различного формата с помощью компьютерной программы Sound Forge 9с. В 3У хранятся записанные заранее файлы в формате WAVE тестовых сигналов белого шума с равномерной спектральной плотностью от 20 до 20000 кГц и уровнями минус 6, 20, 40, 60 и 80 дБ. Все эти файлы с помощью программы Lame 3,98.4 последовательно конвертируются в форматы MP3 CBR, ABR и VBR с различными скоростями цифрового потока и уровнями тестовых сигналов. Они сохраняются в памяти компьютера. Частотные характеристики фонограмм снимаются в программе Sound Forge 9с.

4. Методические указания

4.1. Компрессирование в формате MP3

Общие сведения

В настоящее время формат MP3 де-факто является стандартом, так как 90% музыкальных файлов передаются и скачиваются в интернете в этом формате, не менее важно, что файлы в таком формате могут воспроизводить все плееры массового производства, и программное обеспечение кодеров MP3 распространяется бесплатно. Компрессирование в этом формате основывается на рекомендациях MPEG 1, layer 3.

В формате MP3 при скорости цифрового потока 32 Кбит/с (моно) используются следующие алгоритмы компрессирования:

- субполосное кодирование,
- неравномерное квантование (компандерная система),
- энтропийное кодирование (код Хаффмана),
- перцептуальное кодирование.

С повышением скорости цифрового потока для улучшения качества звучания некоторые подпрограммы перцептуального кодирования исключаются. В результате уже при скорости 128 Кбит/с (стерео) качество звука такое же, как с CD, тогда как при этом в среднем на одну выборку приходится всего 1,17 бита. В настоящее время в

интернете распространяются файлы MP3 записанные со скоростью 192 Кбит/с (стерео), при этом формат MP3 уже можно отнести к компрессированию без потерь.

Основные технические характеристики кодера (при $f_s = 48$ кГц).

Входной сигнал

- моно, стерео ($m = 2$),
- линейная ИКМ,
- частота дискретизации $f_s = 48$ (44,1 и 32) кГц,
- число бит в выборке $q = 16$.
- скорость цифрового потока (стерео) $R = f_s q m = 1536$ Кбит/с
- размер фрейма аудиоданных 1152 выборки
- длительность фрейма

$$\tau = \frac{1152}{f_s} = 24 \text{ мс}$$

- размер блока данных 36 выборок

Выходной сигнал

- частота дискретизации $f_s^* = 48$ (44,1 и ниже) кГц,
- скорость цифрового потока (стерео) $R = 32, 48, 56, 64, 80, 96, 112, 128, 160, 192$ kbits/s,
- среднее число бит в выборке $q^* = 0,67 \dots 4$
- режимы кодирования:
 - CBR- constant bit rate (постоянная скорость)
 - VBR- variable bit rate (переменная скорость)
 - ABR- average bit rate (средняя скорость)

Субполосное кодирование

- число полос $n = 32$,
- ширина полосы

$$\Delta F_1 = \frac{f_s}{n} \text{ (750 Гц)}$$

- частота дискретизации

$$f_{si} = \frac{f_s}{n} \text{ (1,5 кГц)}$$

- ширина спектральных линий

$$\Delta F_2 = \frac{f_s}{n \cdot 18} = 41,667 \text{ Гц}$$

- число спектральных линий $n^* = 32 \cdot 18 = 576$

- квантование с неравномерным шагом с

адаптацией

- число разрядов квантователя $2 \dots 16$

- код Хаффмана

- число спектральных линий

Психоакустическая модель

$$n_m = \frac{FFTsize(1024)}{2} \text{ (512)}$$

- ширина спектральной линии

$$\Delta F_m = \frac{f_s}{FFTsize(1024)} \text{ (46,875 Гц)}$$

- время анализа блока данных 21,3 мс

- нижняя граница спектра

$$F_n = \frac{2f_s}{FFTsize(1024)} (93,75 \text{ Гц})$$

Режим CBR- является основным. В этом режиме весь файл передается с одной скоростью и это значит, что все части файла имеют одинаковое число бит. Поэтому качество их меняется, при этом сложная часть файла имеет худшее качество звучания, чем более простая. Команда конвертирования в режим CBR:

lame -h -b скорость имя.wav имя.mp3

В режиме ABR все части файла кодер передает с одинаковой средней для данного файла скоростью, от файла к файлу скорость меняется. При этом с увеличением скорости цифрового потока сопровождается увеличением числа бит. Поэтому качество звука в режиме ABR выше, чем в режиме CBR. Команда конвертирования в режим CBR:

lame --abr скорость имя.wav имя.mp3

В режиме VBR кодер пытается сохранить заданное качество всего файла, выбирая оптимальное число бит для каждого музыкального фрагмента файла. Качество задается числом V , которое может быть от 0 до 9. Команда конвертирования в режим CBR:

lame -h -V 0...9. имя.wav имя.mp3

Субполосное кодирование в формате MP3

Входной поток цифровых данных разделяется на блоки по 1152 выборки длительностью $\tau = 24$ мс. Это время выбрано примерно в 3 раза меньше постоянной времени слуха, поэтому возможно последовательное вычисление мгновенного спектра, который более точно отражает музыкальные образы.

Анализатор спектра субполосного тракта двухступенчатый, сначала звуковой спектр с помощью цифровых фильтров с крутизной спада 100 дБ/октава разбивается на 32 полосы шириной $\Delta F_1 = 750$ Гц. В каждую полосу для анализа попадают по 36 выборок, поэтому время анализа равно 8 мс.

Затем с помощью 32 анализаторов спектра каждая полоса разбивается еще на 18 более узких полос шириной $\Delta F_2 = 41,67$ Гц, называемых спектральными линиями, их

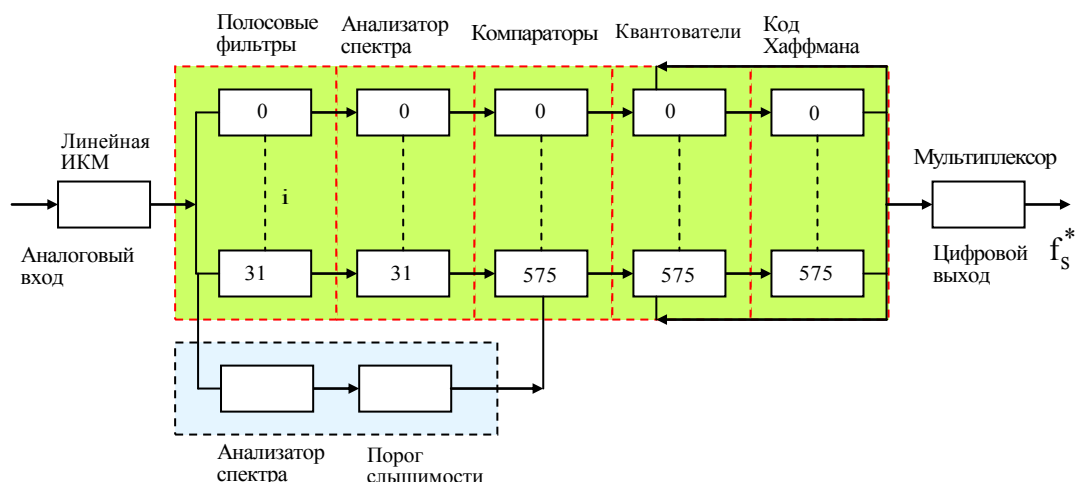


Рис.8.2. Схема субполосного кодера в формате MPEG-1, layer 3

число в полосе с 0 до частоты Найквиста 576 (рис.8.2.). Такая фильтрация осуществляется расчет путем чисто математически с помощью так называемого модифицированного косинусного преобразования, поэтому перекрытия между спектральными линиями нет. Анализ спектра производится в достаточной узких полосах

шириной 750 Гц, что при времени анализа 8 мс обеспечивает высокое разрешение, при котором спектр передается с 20 Гц.

Ширина спектральных линий ΔF_2 Гц выбрана равной 46,75 Гц, что примерно соответствует половине минимальной ширины критической полосы слуха ΔF_{bc} 100 Гц в диапазоне от 0 до 500 Гц. Такое решение позволяет с некоторой погрешностью моделировать критические полосы во всем звуковом диапазоне. Чем меньше ширина ΔF_2 , тем более тонко можно осуществлять регулирование в них коэффициента передачи в зависимости от порога слышимости в этих полосах.

Для регулировки коэффициента передачи по спектральным линиям предназначены *компараторы*, которые управляются сигналами из блока вычисления порога слышимости. Они пропускают сигнал только при условии, что он выше порога слышимости в данной полосе частот.

Психоакустическая модель

Блок вычисления порога слышимости, предназначен для управления коэффициентом передачи компараторов, включенных на выходах анализаторов спектра субполосного тракта. Он включает в себя многополосный анализатор спектра и вычислительный блок, рассчитывающий трехмерный (амплитуда, частота и время) так называемый глобальный порог слышимости по спектральным линиям анализатора спектра, которым управляются компараторы и осуществляется распределение бит.

С учетом, что в формат MP3 преобразуется 16-ти разрядный ИКМ сигнал, при вычислении порога слышимости за 0 дБ в шкале FS принят уровень звукового давления 96 дБ SPL (рис.8.3). Как видно из этого рисунка, учет порога слышимости приводит к достаточно сильному уменьшению воспроизводимого спектра, как на высоких, так и на

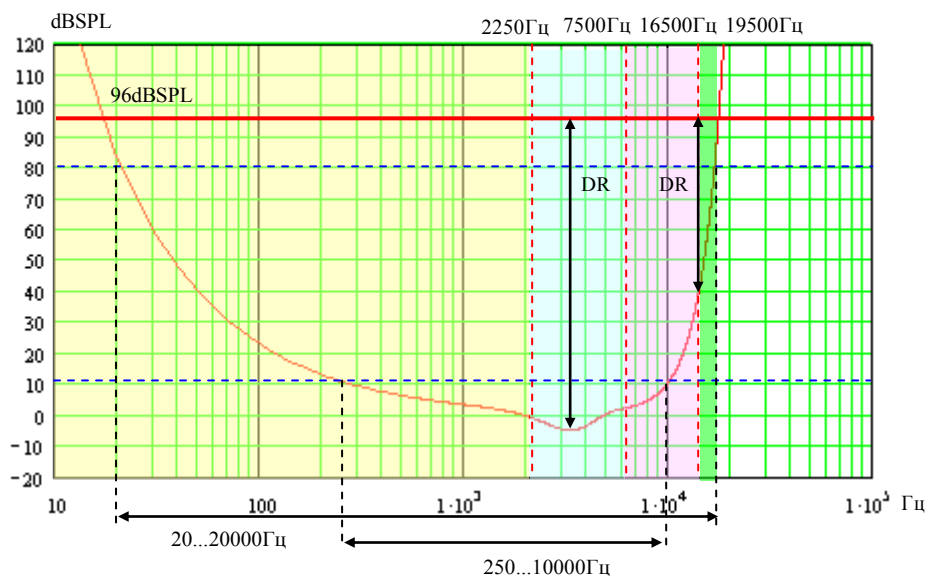


Рис.8.3. К определению слышимой полосы частот

низких частотах. Так при уровне сигнала около минус 80 дБ FS верхняя частота спектра понижается с 20 до 10 кГц, правда на слух это заметить невозможно. Сокращение частотного диапазона в области низких частот существенно меньше, не выше 250 Гц.

По идее *анализатор спектра блока вычисления порога слышимости* должен быть такой же, как в субполосном тракте, но это практически невозможно реализовать. Поэтому при анализе спектра используется БПФ с $FFTsize = 1024$, при котором в каждом блоке данных из 1152 выборок в расчетах спектра используются только 1024 с использованием оконной функции Hanning, уменьшающей вес выборок в начале и конце блока данных.

Ширина спектральных линий при БПФ определяется равенством

$$\Delta F_m = \frac{f_s}{FFTsize} = 46,875 \text{ Гц} (f_s = 48 \text{ кГц})$$

Нижняя граница анализируемого спектра определяется удвоенным значением ширины спектральной линии и равна 93,75 Гц. Это значит, что в формате MP3 ниже этой частоты использовать особенности слуха для компрессирования скорости цифрового потока невозможно. При выбранной ширине спектральных линий их число в диапазоне от 0 до частоты Найквиста равно 512. Однако, фактически при расчетах используются только 428 спектральных линий в частотном диапазоне до 20000 кГц. Причем вычисление порога слышимости производится всего на 126 спектральных линиях, которые охватывают 24 критические полосы слуха от 46,875 до 20250 Гц.

Для повышения скорости работы кодера MP3 в большинстве из них используется режим **-f fast mode**. В этом режиме производится только очень упрощенный психоакустический анализ: операция Noise shaping исключается и производятся расчеты, необходимые для распределения бит и обнаружения пре-эхо. При этих условиях обеспечиваются характеристики, приведенные в **таб.8.1**

Table 8.1: Bitrate versus sound quality		
Bitrate	Bandwidth	Quality comparable to or better than
16 kbps	4.5 kHz	shortwave radio
32 kbps	7.5 kHz	AM radio
96 kbps	11 kHz	FM radio
128 kbps	16 kHz	near CD
160-180 kbps (variable bitrate)	20 kHz	perceptual transparency
256 kbps	22 kHz	studio

Дискретизация

Дискретизации (точнее передискретизации) в субполосном тракте подвергаются сигналы на выходе анализатора спектра в полосах шириной 41,67 Гц. Теоретически частота дискретизации должна по крайней вдвое превышать ширину полосы фильтров, то есть быть выше 100 Гц. Она выбрана равной 1,5 кГц, то есть в 18 раз выше. Такое решение упрощает реализацию восстанавливающих фильтров декодера и одновременно решает проблему декорреляции выборок, что необходимо для оптимизации энтропийного кодирования.

Неравномерное квантование

Равномерное квантование является оптимальным только при равномерном распределении вероятностей уровней звуковых сигналов. При субполосном кодировании это распределение наиболее близко описывается законом Гаусса или Лапласа. При этом линейное квантование становится не эффективным и для минимизации средней ошибки квантования более подходит неравномерное квантование.

Неравномерно квантование в формате MP3 выполняется с использованием мгновенного компандирования, примерно так же, как это делается в цифровых телефонных линиях. На рис. 8.4 приведена упрощенная схема такого компандирования, на которой показаны графики передаточных функций компрессора, экспандера и

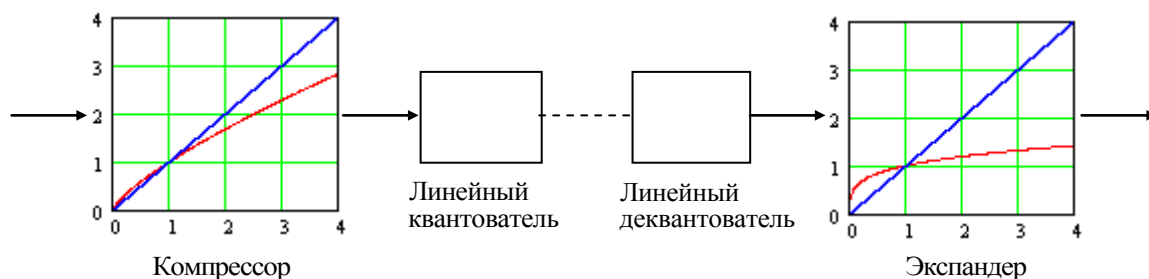


Рис. 8.4. Эквивалентная схема неравномерного квантования

сквозного тракта. В кодер этой схеме входят компрессор, который сжимает динамический диапазон сигналов, и линейный квантователь. В декодере сигнал после деквантования подается на экспандер, расширяющий динамический диапазон, сквозная характеристика-линейная. При таком квантовании большие амплитуды сигналов квантуются более грубо, а малые - более точно. Это очень важно, так как на слух ошибки квантования более заметны при малых и средних уровнях звуковых сигналов. Неравномерное квантование обеспечивает некоторое постоянство SNR по всему диапазону изменения уровней квантования звуковых сигналов.

В формате MP3 компрессирование и экспандирование сигналов (в аналоговой форме) осуществляется по законам, определяемым равенствами:

- компрессор $U(K)_{out} = (U(K)_{in})^{3/4}$
- экспандер $U(E)_{out} = (U(E)_{in})^{1/4}$

В реальном алгоритме работы кодера операции нелинейного преобразования и квантования выполняются одновременно чисто математическими средствами путем

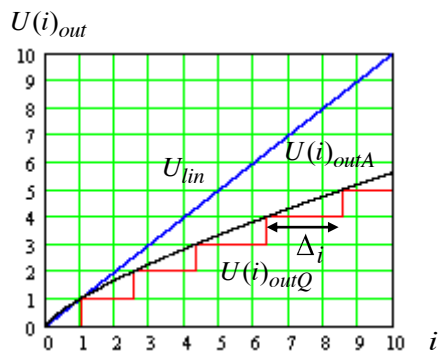


Рис.8.5. Квантование линейно нарастающего напряжения с переменным шагом квантования по закону $y = x^{3/4}$

возведения выборок входного сигнала в степень $3/4$ с использованием расчетного соотношения

$$U(i)_{outQ} = n \text{int} \left[\left(\frac{U(i)_{in}}{2^{\Delta/4}} \right)^{3/4} - 0,0946 \right], i = 0, 1 \dots 575$$

где $U(i)_{outQ}$ – квантованный выходной сигнал компрессора, $U(i)_{in}$ – входной сигнал компрессора, $n \text{int}()$ – ближайшее целое значение числового значения в скобках, Δ – шаг квантования, L – число уровней квантования

$$\Delta = \frac{U(i)_{in\max} - U(i)_{in\min}}{L}$$

На **рис.8.5.** иллюстрируется неравномерное квантование линейно нарастающего напряжения U_{lin} . Функция $U(i)_{out}$ – не квантованный сигнал на выходе компрессора, функция $U(K)_{outQ}$ – квантованный сигнал. Как видно, первый шаг квантования равен 1, чем больше уровень входного сигнала компрессора, тем больше шаг квантования Δ_i .

В схеме на **рис.8.5** используется линейный квантователь с передаточной характеристикой **Tread Mid**, осуществляет квантование сигналов с уровнем выше половины шага квантования. В этом квантователе число уровней квантования от пика до пика N_q всегда нечетное, и связано с числом разрядов q равенством

$$N_q = 2^q - 1$$

Зависимость отношения сигнал/шум (SNR) от числа разрядов рассчитывается по формуле

$$SNR = 6,02q + 1,76$$

Табл. 8.2. Квантование		
N_q	q	SNR, dB
3	2	7,00
5		11,00
7	3	16,00
9		20,84
15	4	25,28
31	5	31,59
63	6	37,85
127	7	43,84
255	8	49,89
511	9	55,93
1023	10	61,96
2047	11	67,98
4095	12	74,01
8191	13	80,03
16383	14	86,05
32767	15	92,01
65535	16	98,01

Число уровней квантования может быть от 3 до 65535 с числом разрядов в выборках от 2 до 16. В **табл.8.2** приведены возможное число уровней квантования и соответствующие им число разрядов и SNR. Допустимое значение SNR определяется заданной скоростью цифрового потока от 32 до 192 кбит/с, чем скорость меньше, тем меньше может быть SNR.

При компрессировании с использованием психоакустической модели квантуются только сигналы, уровень которых в полосах превышает порог слышимости, определяемый как абсолютным порогом, так и порогом слышимости, связанным с частотной и временной маскировкой.

В формате MP3 число используемых уровней квантования разное для различных диапазонов частот. В диапазоне от 0 до 2250 Гц используется квантование с числом уровней от 3 до 32767 с выборками от 2 до 15 бит. Выше 2250 Гц число уровней понижается, и в верхнем участке звукового диапазона от 16500 до 19500 Гц оно равно 3 или 5. На частотах выше 19500 Гц квантование вообще не производится (**табл.8.3**).

Необходимость использования на низких частотах квантования с большим числом уровней, связано с тем, что в формате MP3 на этих частотах абсолютный порог слышимости не вычисляется из-за низкой разрешающей способности БПФ. В диапазоне 2250...7500 Гц наиболее сильно проявляется эффект частотной маскировки, поэтому динамический диапазон слышимых сигналов понижается и требуется меньшее число уровней квантования. Выше частоты 7500 Гц достаточно резко поднимается абсолютный порог слышимости, динамический диапазон слышимых сигналов быстро уменьшается до 10...20 дБ, поэтому вполне достаточно 3 или 5 уровней квантования.

В формате MP3 предусмотрена возможность компрессирования звука без использования эффекта порога слышимости, в этом случае во всем звуковом диапазоне число уровней квантования равно 65535 (16 разрядное кодирование). При этом компрессирование осуществляется только благодаря субполосному кодированию, оптимизации квантования и упаковки цифровых данных

Табл. 8.3. Изменение числа уровней квантования			
Номер субполос	Диапазон, Гц	Число уровней квантования	Число бит
0...2	0...2250	3...32767 (65355)	2...15
3...10	2250...7500	3...8191 (65355)	2...13
11...22	7500...16500	3...31 (65355)	2...5
23...26	16500...19500	3,5 (65355)	2...2,5
27...31	19500...24000	0	

4.2. Измерение частотной характеристики цифровых трактов


Измерение частотной характеристики кодеков формате MP3 возможно только с помощью программного пакета Sound Forge 9c. Для этого используется тестовый сигнал белого шума в формате Wave с равномерной спектральной плотностью от 20 до 20000 Гц, и уровнем от минус 6 до минус 80 дБ.

5. Порядок выполнения лабораторной работы

5.1. Подготовка к измерениям

- Скопируйте с рабочего стола файл «Отчет LR-8 AT» на свою флешку, дайте ему свое имя в виде: «Группа»_ «Фамилия (работавшего на клавиатуре)»_ «Номер ЛР»_ «Дата».
- Откройте этот файл и сверните его окно.

В файл этого отчета вы будете копировать все экспериментальные графики, и заносить в таблицы отчета результаты измерений сразу в процессе их получения. В день выполнения лабораторной работы отчет в электронном виде необходимо сдать преподавателю, только тогда работа считается выполненной.

- Кнопкой с пиктограммой  откройте окно программы **Sound Forge 9c (SF)**, содержащее окно стерео сигналограммы и окно спектров левого и правого звуковых каналов. Не меняйте размеры этого окна, они согласованы с Вашим отчетом

- Нажмите кнопку с пиктограммой  и в появившемся меню

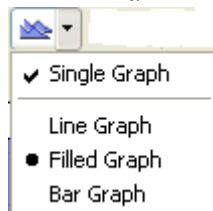



Рис.8.6. Меню анализатора спектра

выберите пункты, указанные на рис. 8.6. При этом будет строится только график спектра левого канала.

- Нажмите кнопку с пиктограммой  . По этой команде откроется окно предустановок (рис.8.7), их следует установить такими же как на рисунке. Обязательно используется линейная шкала частот

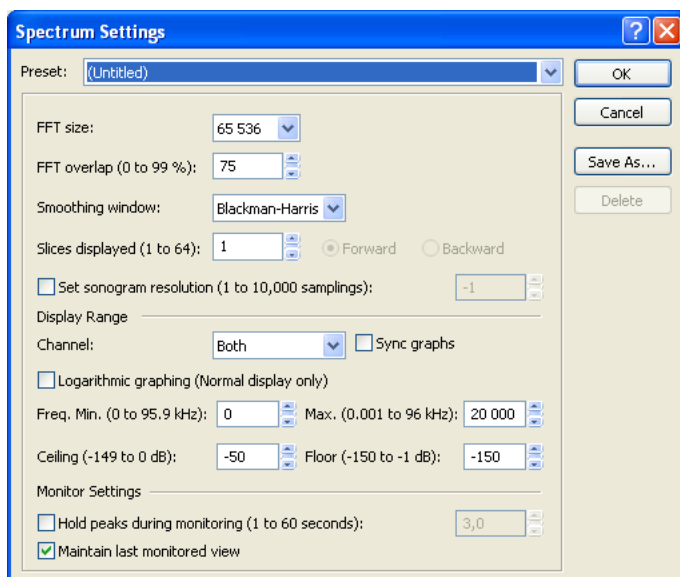




Рис.8.7. Окно предустановок анализатора спектра

Графики спектров строятся в диапазоне 20...20 кГц с уровнем от минус –50 до минус 150 дБ , FFT size = 65536.

- Кнопкой с пиктограммой  откройте окно программы **HyperSnap-DX**, предназначенной для вырезки и копирования графических изображений.
- Кнопкой с пиктограммой  откройте окно программы **Total Commander**, и на диске *C* в корневом каталоге откройте папку «Мои документы».Из нее скопируйте папку **“Lame 3.98.4”** в корневой каталог диска *C*.
- Кроме программных файлов в папке **“Lame 3.98.4”**. находятся также 5 файлов с расширением **.wav** . Эти файлы содержат тестовые сигналы белого шума длительностью 60 секунд каждый с уровнями: - 6, -20, -40, -60 и –80 дБ без компрессирования, частота дискретизации 48 кГц, число разрядов 16. Файлы имеют названия: *T* «уровень сигнала (в дБ)».

5.2. Исследование частотных характеристик кодека MP3“Lame 3.98.4” в режиме CBR при разной скорости цифрового потока

Конвертирование тестовых сигналов из формата Wave в формат MP3 CBR

5.2.1. Выбираем для конвертирования файла **T6.wav** (или T20...T80) в формат **MP3 CBR** со скоростью цифрового потока 256 кбит/с

- Курсором активируйте программу **“Lame 3.98.4”**. По этой команде внизу окна **Total Commander** появится надпись **c:\Lame 3.98.4>**
- Вслед за этой надписью в командной строке наберите клавиатурой текст: **lame -h -b 256 T6.wav T6CBR256.mp3** (рис. 8.8) и дайте команду **Enter**. В этом тексте буква **h** означает высокое качество кодирования, за буквой **b** следует задаваемая скорость цифрового потока в килобитах
- По команде **Enter** через 10...20 с в папке **Lame 3.98.4** появится сконвертированный файл с таким названием: **T6CBR256.mp3** (белый шум в формате MP3 CBR с уровнем минус 6 дБ)

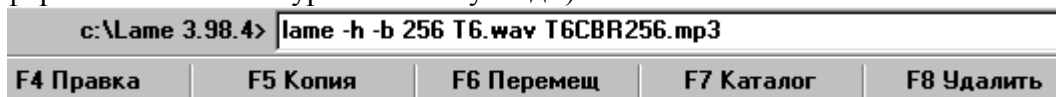


Рис. 8.8 Команды программы Lame в режиме CBR (-h -b скорость имя.wav имя .mp3)

5.2.2. Повторяем операции по п.5.2.1. и конвертируем файл T6.wav в формат **MP3 CBR** со скоростью цифрового потока 32, 64, 128 и затем 256 кбит/с

Измерение частотных характеристик кодека MP3 CBR при разных скоростях цифрового потока

5.2.2.

- Курсором активизируйте окно программы **Sound Forge 9c**
- В меню **File** программы **SF** нажмите кнопку **Open**. По этой команде на экране ПК появится окно **Открыть**, содержащее файлы папки **“Lame 3.98.4”** (рис.8.9). В этом окне курсором установите **Тип файлов MP3 audio**, и выберите в ней тест

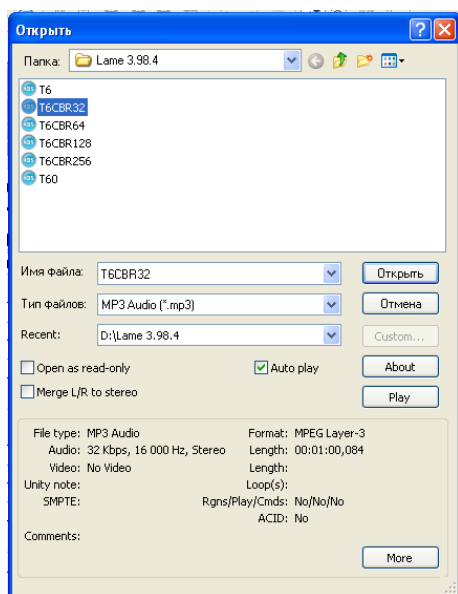


Рис.8.9 Окно программы SF, открывающее содержание папки программы Lame

T6CBR32.mp3, а затем нажмите кнопку **Открыть** (Open). По этой команде в окне **SF** появится сигналограмма тестового сигнала (рис.8.10)

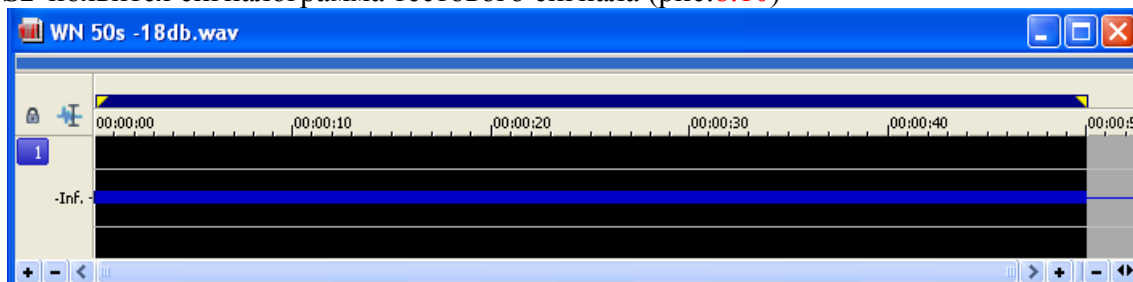




Рис.8.10. Выделенная фонограмма тестового сигнала

- Курсором выделите всю фонограмму от начала до конца, при этом ее фон становится черным
- В окне спектрального анализа нажмите кнопку с пиктограммой . Через 10...60 с в окне появится график спектра левого канала тестового сигнала. График этого спектра сохраните в ячейке памяти 1. Для этого последовательно нажмите кнопки **Set** и **1**.

5.2.3.

- Повторите измерения по п.5.2.2. для тестовых сигналов со скоростями 64,128, и 256 кбит/с, сохраняя их частотные характеристики последовательно в ячейках памяти 2, 3 и 4.
- При этих измерениях может автоматически измениться масштаб частотной шкалы, поэтому по окончании измерений кнопкой с пиктограммой  откройте окно предустановок, и установите верхнюю границу спектра 20000 Гц
- Скопируйте графики спектров левого канала(рис.8.11) в Ваш отчет.

Для этого активизируйте окно программы **HyperSnap-DX** и выберите в его меню **Capture** команду **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу графика спектра нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол графика спектра и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** в меню **File** выберите **Save As**. Установите путь копирования на Вашу флешкарту, дайте название файлу и сохраните его в своем отчете.

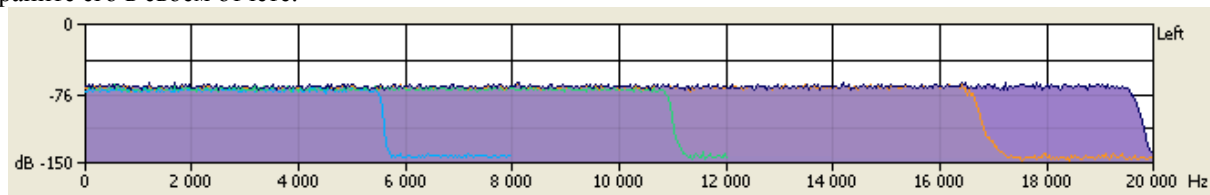


Рис.8.11 Графики частотных характеристик кодека MP3 CBR при разных скоростях цифрового потока (256, 128, 64 и 32 кбит/с.)

- С помощью курсора измерьте на графиках верхние границы частотных характеристик, и результаты занесите в табл.8.4
 - В меню **File** программы **SF** нажмите кнопку **Open**. По этой команде на экране ПК появится окно, содержащее файлы папки “**Lame 3.98.4**”. Нажимая курсором последовательно на файлы с разными скоростями снимите показания для них
- показания и частоты дискретизации на выходе кодера f_s^* , занесите результаты в табл.8.4. Рассчитайте скорость входного цифрового потока v тестового сигнала белого шума по формуле

$$v = q \cdot f_s \cdot n = 16 \cdot 48000 \cdot 2 = 1536 \text{ кбит/с}$$

где q – число разрядов, f_s – частота дискретизации входного цифрового потока,

n – число звуковых каналов

Табл.8.4. Характеристики кодека MP3 в режиме CBR					
Файл	Уровень сигнала, дБ	Полоса частот, кГц	Скорость, кбит/с	f_s^* , кГц	K
T6CBR256	-6				
T6CBR128	-6				
T6CBR64	-6				
T6CBR32	-6				
T40CBR256	-40				
T60CBR256	-60				
T80CBR256	-60				

- Для данных табл.8.3 рассчитайте коэффициент компрессирования K как отношение скорости v к скорости цифрового потока MP3 кодера, результаты занесите в табл.8.4
- Сотрите все графики спектров из памяти

Для этого нажмите кнопку с пиктограммой 

5.3. Исследование частотных характеристик кодека MP3“Lame 3.98.4” в режиме CBR при разных уровнях тестовых сигнала

Конвертирование тестовых сигналов из формата Wave в формат MP3 CBR

5.3.1. Выбираем скорость цифрового потока 256 кбит/с (можно любую другую из стандартного ряда)


- Курсором активируйте программу “**Lame 3.98.4**”. По этой команде внизу окна **Total Commander** появится надпись **c:\Lame 3.98.4>**
- После этой надписи в командной строке наберите клавиатурой текст, в котором после буквы **-b** указывается скорость цифрового потока:
lame -h -b 256 T6.wav T6CBR256.mp3 (рис.8.12) и дайте команду **Enter**
- По этой команде через 10...20 в папке **Lame 3.98.4** появится сконвертированный файл с таким названием: **T6CBR256.mp3** (белый шум в формате MP3 CBR с уровнем минус 6 дБ)

5.3.2. Повторите операции по п.5.3.1. с конвертированием файлов T40.wav ... T80.wav в формат **MP3 CBR** при скорости цифрового потока 256 кбит/с .

Измерение частотных характеристик кодека MP3 CBR при разных уровнях тестовых сигналов

5.3.3

- Курсором активизируйте окно программы **Sound Forge 9c**
- В меню **File** программы **SF** нажмите кнопку **Open**. По этой команде на экране ПК появится окно **Открыть** , содержащее файлы папки “**Lame 3.98.4**” (рис. 8.7). В этом окне курсором установите **Тип файлов MP3 audio** и выберите в ней тест **T6CBR256.mp3**, а затем нажмите кнопку **Открыть** (Open). По этой команде в окне **SF** появится сигналограмма тестового сигнала
- Курсором выделите всю фонограмму от начала до конца, при этом ее фон становится черным

- В окне спектрального анализа нажмите кнопку с пиктограммой . Через 10...60 с в этом окне появится график спектра левого канала тестового сигнала. График этого спектра сохраните в ячейке памяти 1. Для этого последовательно нажмите кнопки **Set** и 1.

5.3.4.

- Повторите измерения по п.5.3.3. для тестовых сигналов с уровнем минус 40, 60 и 80 дБ при скоростях 256 кбит/с. Сохраните их частотные характеристики последовательно в ячейках памяти 2, 3 и 4.
- Скопируйте графики спектров левого канала (рис.8.12) в Ваш отчет.

Для этого активизируйте окно программы **HyperSnap-DX** и выберите в его меню **Capture** команду **Region**. По этой команде на экране появится перекрестие двух перпендикулярных линий. Курсором установите перекрестие на нижнем левом углу графика спектра нажмите левую клавишу мышки. Затем переместите курсор в правый верхний угол графика спектра и снова нажмите левую клавишу мышки (этим вы скопировали изображение гистограммы в оперативную память). Далее в окне программы **HyperSnap-DX** в меню **File** выберите **Save As**. Установите путь копирования на Вашу флешкарту, дайте название файлу и сохраните его в своем отчете.

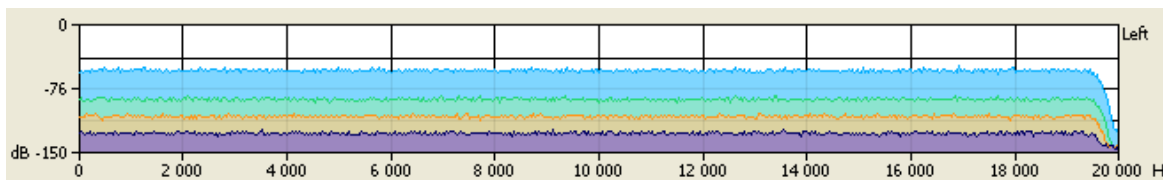


Рис.8.12. Графики частотных характеристик кодера MP3 CBR при разных уровнях тестового сигнала: минус 6, 40, 60 и 80 дБ при скорости цифрового потока 256 кбит/с.

- С помощью курсора измерьте на графиках верхние границы частотных характеристик, и результаты занесите в [табл.8.4](#).
- В меню **File** программы **SF** нажмите кнопку **Open**. По этой команде на экране ПК появится окно, содержащее файлы папки **“Lame 3.98.4”**. Нажимая курсором последовательно на файлы T6CBR256.mp3... T8CBR256.mp3 снимите показания их выходной частоты дискретизации и занесите результаты в [табл.8.4](#).
- Сотрите все графики спектров из памяти

Для этого нажмите кнопку с пиктограммой

5.4.Исследование кодера MP3 в режимах работы : CBR, ABR и VBR

Конвертирование тестовых сигналов из формата Wave в форматы MP3 CBR, ABR и VBR

5.3.1.

- Выбираем скорость цифрового потока 128 кбит/с (можно любую другую из стандартного ряда) и уровень тестовых сигналов минус 20 дБ (может быть задан и любой другой уровень)
- Курсором активизируйте программу **“Lame 3.98.4”**. По этой команде внизу окна **Total Commander** появится надпись **c:\Lame 3.98.4>**
- В режиме **CBR** в длинном окошке наберите клавиатурой текст, в котором после буквы **-b** указывается скорость цифрового потока:
lame -h -b 128 T20.wav T20CBR128.mp3 и дайте команду **Enter**
- По этой команде через 10...20 в папке **Lame 3.98.4** появится файл с таким названием: **T20CBR128.mp3** (белый шум в формате MP3 CBR с уровнем минус 6 дБ)
- В режиме **ABR** команда конвертирования имеет вид:
lame --abr 128 T20.wav T20ABR128.mp3
- В режиме **VBR** команда конвертирования имеет вид:
lame -h V (0 или 9) T20.wav T20VBRV(0 или 9).mp3.

Расчет характеристик кодера в различных режимах работы

- Активизируйте окно **Total Commander** с папкой Lame и из нее занесите в [табл.8.5](#) данные по размеру сконвертированных в разделе 5.3.1 файлов.
- Исходя из длительности всех файлов 60 с рассчитайте скорости цифровых потоков в кбит/с, данные занесите в [табл.8.5](#)
- Рассчитайте коэффициенты компрессии для всех файлов [табл.8.4](#), сравните результаты
- В меню **File** программы **SF** нажмите кнопку **Open**. По этой команде на экране ПК появится окно, содержащее файлы папки **“Lame 3.98.4”**. Нажимая курсором последовательно на файлы MP3 с уровнем тестового сигнала минус 20 дБ снимите показания для них частоты дискретизации на выходе кодера f_s^* , занесите результаты в [табл.8.5](#).

Табл.8.5. Характеристики кодека MP3 в режимах CBR, ABR и VBR					
Файл	Уровень сигнала, дБ	Размер, кбайт	Скорость, кбит/с	f_s^* , кГц	K
T20CBR128	-20				
T20ABR128	-20				
T20VBRV0	-20				
T20VBRV9	-20				

- Рассчитайте скорость входного цифрового потока v тестового сигнала белого шума по формуле

$$v = q \cdot f_s \cdot n = 16 \cdot 48000 \cdot 2 = 1536 \text{ кбит/с}$$

где q – число разрядов, f_s – частота дискретизации входного цифрового потока, n – число звуковых каналов

- Рассчитайте коэффициенты компрессии для всех форматов файлов, сравните результаты

6. Контрольные вопросы

1. Какие скорости цифровых потоков возможны в формате MP3
2. Какие частоты дискретизации допустимы на входе и выходе кодера MP3
3. В каких пределах может осуществляться компрессирование
4. Как зависит коэффициент компрессирования от скорости цифрового потока
5. Чем отличаются режимы кодирования CBR, ABR и VBR, какой из них обеспечивает больший коэффициент компрессирования
6. В каких пределах может изменять коэффициент компрессирования в режиме VBR
7. Как рассчитывается коэффициент компрессирования
8. Как меняется частотная характеристика кодека MP3 в зависимости от скорости цифрового потока
9. Зависит ли частотная характеристика кодека MP3 от уровня сигнала
10. Как и почему меняется частота дискретизации выходного сигнала кодера MP3 при изменении скорости цифрового потока

7. Содержание отчета

- 7.1. Титульный лист
- 7.2. Цель работы
- 7.3. Задание к лабораторной работе
- 7.4. Функциональная схема измерений
- 7.5. Результаты измерений в виде графиков
- 7.6. Таблицы с результатами измерений и расчетов
- 7.7. Выводы по результатам измерений

8. Литература

Ю.А. Ковалгин, Э.И. Вологдин Цифровое кодирование звуковых сигналов. СПб, Корона Принт, 2004.

Программные пакеты Sound Forge 9c и Audition 3

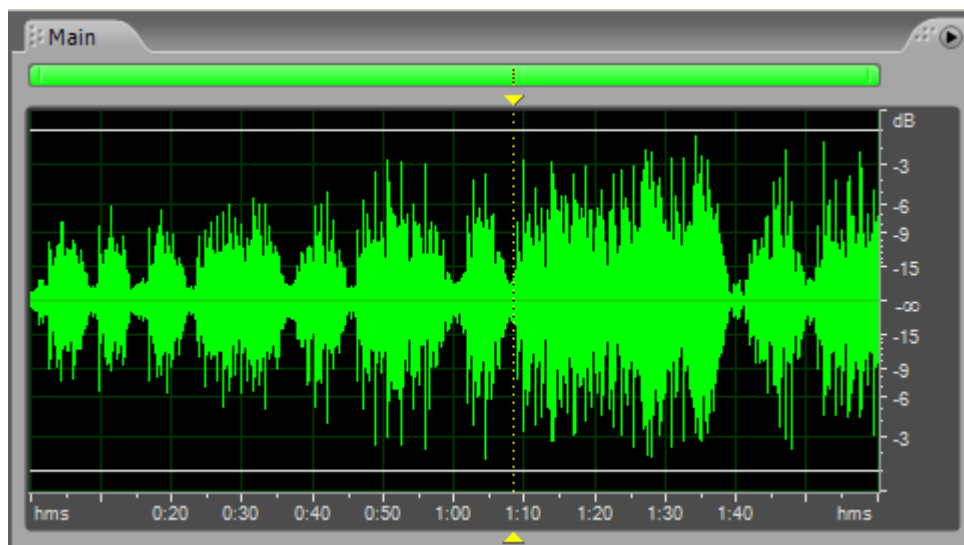
В настоящее время разработано громадное число программных продуктов, предназначенных для записи, редактирования и монтажа фонограмм с использованием звуковой карты и персонального компьютера. Среди них наибольшей популярностью пользуются Sound Forge 9c и Audition 3. В табл.1.. и 2.. приведены основные функциональные возможности этих пакетов.

Табл. 1..... Сравнительные характеристики пакетов программ двух звуковых редакторов			
	Функции	Sound Forge 9c	Audition 3
Запись	Число звуковых каналов (запись)	1,2,...8 (форматы 2.1; 5.1; 7.1)	1,2
	Число разрядов	8,16,24, 32 и 32, 64 (float)	8,16,32 (float)
	Частота дискретизации	8...192 кГц	
	Форматы записи	Wave Microsoft, Atrac Audio , ADPCM, Dolby Digital AC-3, MPEG 1-layer 1, MPEG 1- layer 2, MP3, ... всего 28	A/mu Law, Dialjgic ADPCM, DVI/IMA ADPCM, Microsoft ADPCM, MP3... всего 21
	Dithering	Triangular, Half Triangular, Half Rectangular, Rectangular, Gaussian	Triangular (Shaped) Gaussian (Shaped)
	Noise Shaping	High Pass Contour, Equal Loudness Contour	A, B, C1, C2, C3, D, E, E2
	Вертикальная шкала сигналограммы	Логарифмическая FS	Число уровней квантования, проценты, децибелы (шкала FS)
Ана	FFT size	128...65536	32...65536
	Оконные функции	6 (Hannig, Blackman....)	
	Диапазон уровней	0... минус 150 dB FS	0... минус 192 dB FS
	Частотные шкалы	Линейная, логарифмическая	
Синтез	Шум	Белый, розовый, коричневый	
	Форма волны сигнала	Синус, треугольник, прямоугольник, скользящий тон с линейным или логарифмическим законами изменения частоты	Синус и 5 гармоник, треугольник, прямоугольник
	Число сигналов, макс	4	1
	Длительность сигналов макс.	0.001... 600 с	100 с

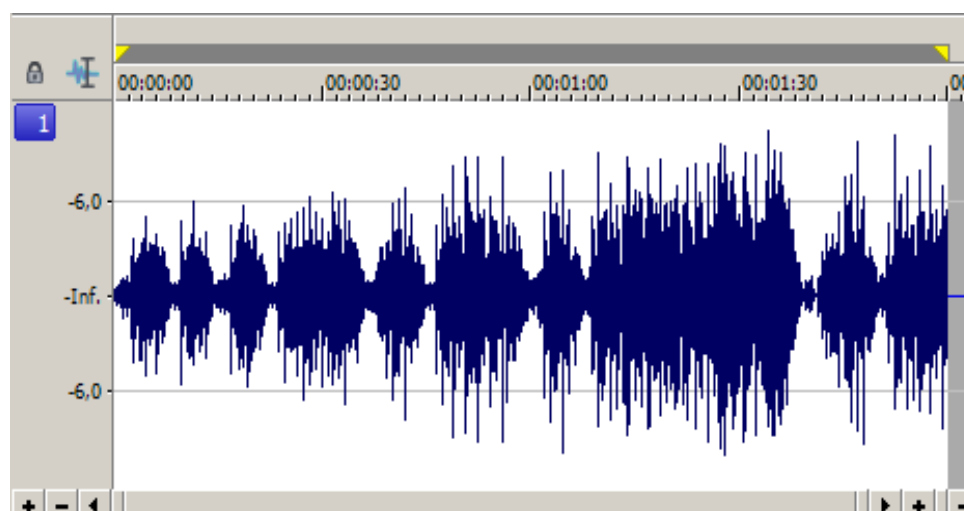
Число звуковых каналов, записываемых одновременно, в программе Audition 3 только 3, тогда как программа Sound Forge 9c позволяет осуществлять многоканальную запись в форматах 2.1, квадро, 5.1 и 7.1. Пределы изменения частоты дискретизации в обоих программах одинаковые от 8 до 192 кГц. Предпочтительно, что в Sound Forge 9c возможна 24 разрядная запись (не с плавающей запятой), позволяющая осуществлять реквантование 24 бит в 16 бит с применением технологий Dithering и Noise Shaping.

Программа Sound Forge 9c позволяет работать с большим числом форматов записи. Это очень важно, так как на этой базе возможна постановка лабораторных работ по исследованию как перцептуальных систем компрессирования, каких как Atrak, Dolby Digital AC-3, MPEG-1 с layers 1, 2 и 3, так и статистических систем кодирования с предсказанием, например, ADPCM. В программе Audition 3 из стандартов MPEG можно использовать только MP3Pro, но зато широко представлены форматы записи на основе ADPCM, можно осуществлять запись с компандированием по законам μ и A.

Возможности реализации технологии Dithering в программе Sound Forge 9c по программным продуктам значительно шире, но в программе Audition 3 предоставляется возможность пользователю самому выбирать уровень шума в битах. В отношении технологии Noise Shaping возможности выбора в программе значительно больше, но, к



a)



b)

Рис. 1 Сигналограммы одного и того же музыкального фрагмента в редакторах Audition 3 (a) и Sound Forge 9c (b)

сожалению, не раскрываются особенности фильтров, поэтому пользователю придется выбирать алгоритм методом проб и ошибок.

Несомненно, в программе Audition 3 более удобное для пользования окно сигналограммы, в котором вертикальная ось может быть децибелах, процентах и представлять число уровней квантования, особенно удобна шкала в децибелах с делениями от минус 3 до до минус 15 дБ (рис.), тогда как шкале только одно деление 6 дБ. Не маловажно, что пользователь может выбирать цветовую палитру, что не предусмотрено в редакторе Sound Forge 9c. .

Характеристики анализатора спектра в обоих редакторах почти одинаковые, однако, возможности программы Audition 3 значительно шире, так ее динамический диапазон простирается от 0 до минус 192 дБ FS, на 40 дБ шире. Частотные шкалы совершенно разные (рис.2). В редакторе Audition 3 привычная логарифмическая по основанию десять шкала. Неудобно, что фон только черный, при печати графики плохо видны. В

редакторе Sound Forge 9c очень сильно выделен диапазон от 0 до 500 Гц, поэтому частотные характеристики по форме отличаются очень сильно, непривычно.

Статистика звуковых сигналов в редакторе Sound Forge 9c крайне скудная, тогда как в редакторе Audition 3 она очень детальная, рассчитываются даже гистограммы с заданным временем усреднения. На [рис.3](#) приведены статистические данные одного и того же музыкального отрывка, иллюстрирующие как они изменяются при изменении

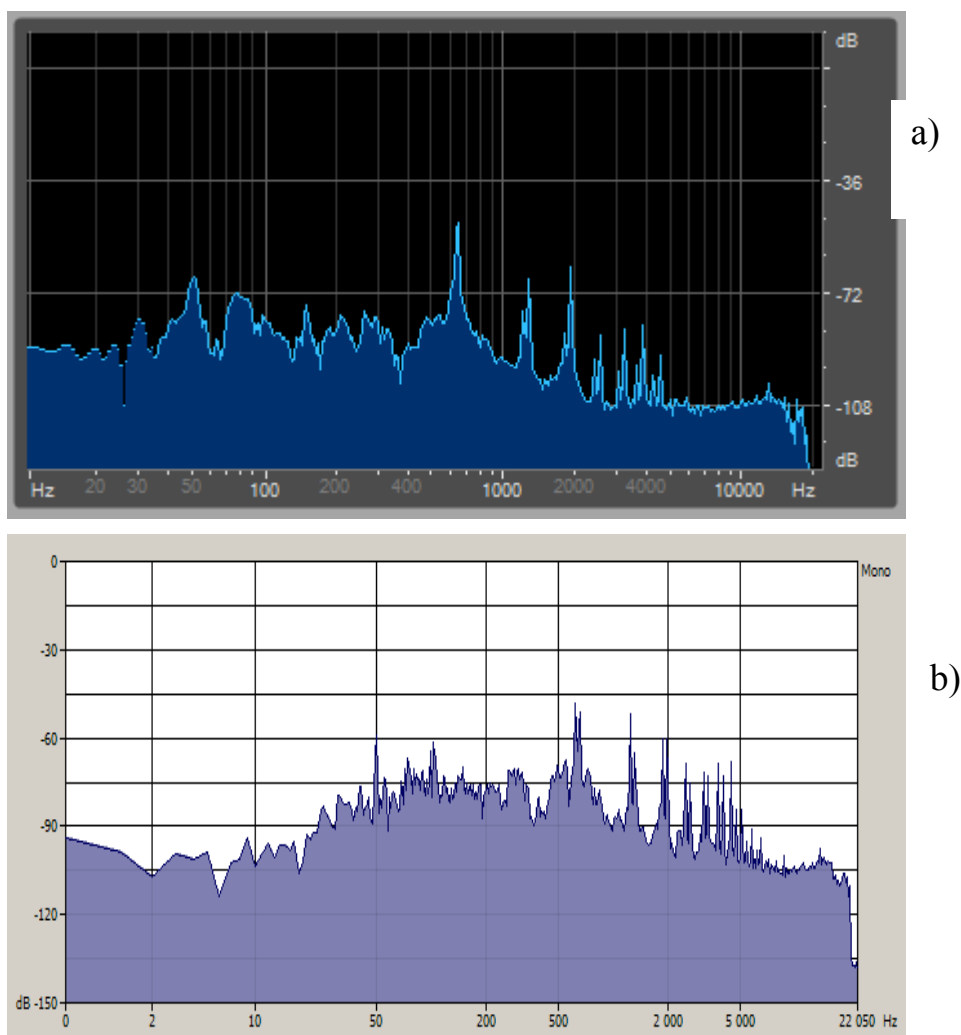


Рис. 2. Спектры одного и того же музыкального фрагмента в редакторах Audition 3 (a) и Sound Forge 9c (b)

времени интегрирования с 5 до 500 мс. Нужно обратить внимание, что расчетное значение Average value программе Sound Forge 9c слишком мало. На [рис.4...](#) представлены гистограммы распределения, рассчитываемые только в редакторе Audition 3. Они позволяют видеть как меняется динамический диапазон

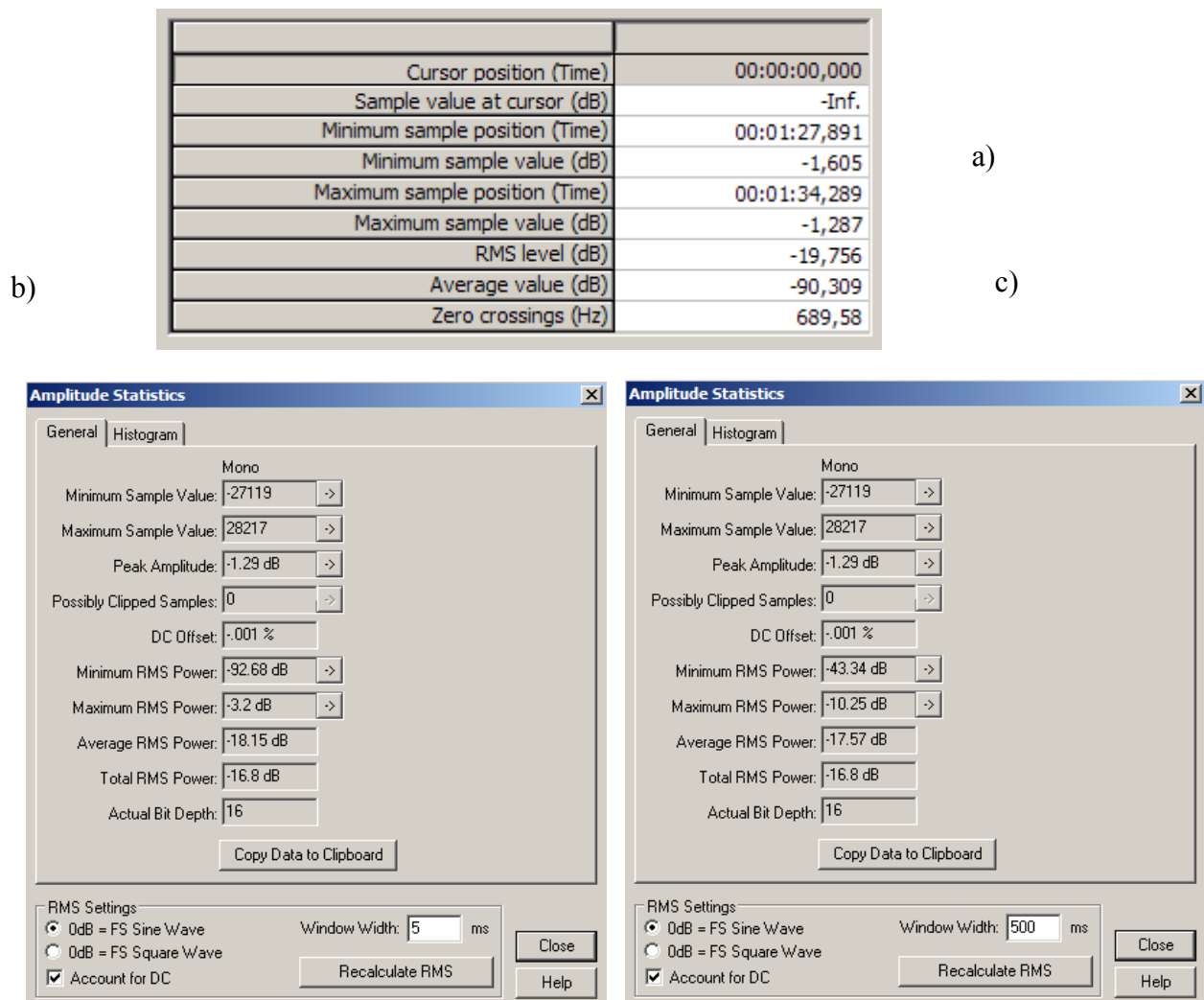


Рис 3..... Статистика одного и того же музыкального отрывка а) редактор Sound Forge, б,с) редактор Audition 3 с временем интеграции 5 и 500 мс.

Функции звуковых редакторов, связанные с обработкой фонограмм и звуковыми эффектами отличаются мало, особенно применительно к постановке лабораторных работ. В режиме пост обработки программа Sound Forge 9с позволяет одновременно обрабатывать до 32 каналов, а программа Audition 3 - только 6. В качестве примера на [рис. 5](#). приведены спектры шестиканального звука в формате 5.1.

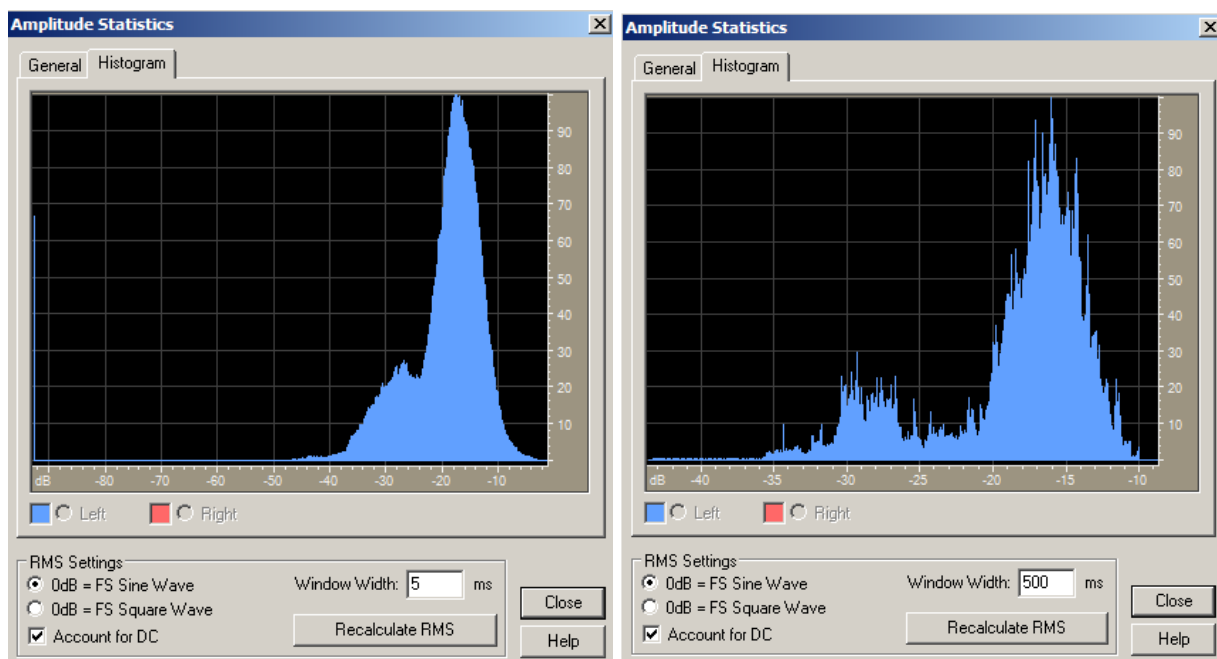


Рис. 4.... Гистограммы одного итого же музыкального отрывка при времени интегрирования 5 и 500 мс

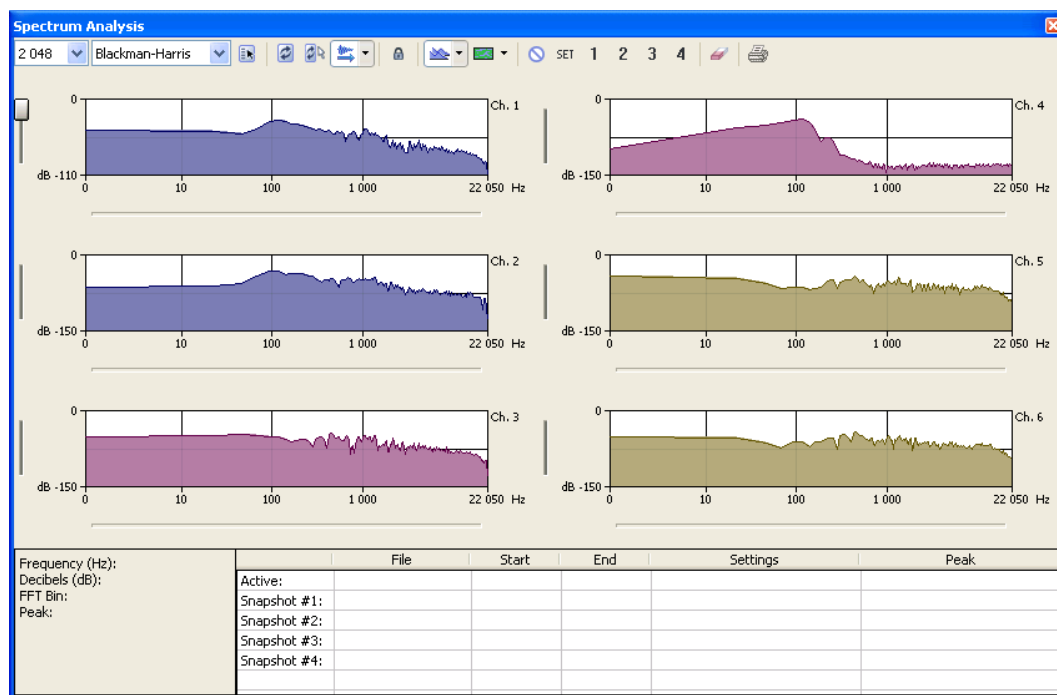


Рис.5. Спектр 6-канального звукового сигнала

Табл.2.... Сравнительные характеристики двух звуковых редакторов		
Функции и звуковые эффекты	Sound Forge 9c	Audition 3
Число звуковых каналов (обработка)	1...32	1...6
Динамическая обработка	Компрессоры: графический , многополосный , жесткий лимитер, максимайзер	Компрессоры: графический , многополосный, жесткий лимитер
Частотная обработка	Эквалайзеры: графический (10, 20 полос), параметрический и параграфический	Эквалайзеры: графический (10, 20 и 30 полос), параметрический , FFT фильтр, Notch фильтр
Задержка (Delay)	Задержка, эхо (Multi Tap, Simple)	Задержка, эхо, эхо камеры, Multy Tap Delay
Модуляция	Тремоло, вибрато, фленжер, Вау-вау, хорус	Хорус, фланжер, Sweeping Phaser
Реверберация	Полная реверберация	Сверточная, полная и студийная реверберация
Высота тона (Pitch)	Pitch Bender, Pitch Shifter	Pitch Bender, Pitch correction, Pitch Shifter
Психоакустическая обработка	Графический синтезатор нелинейных искажений	Графический синтезатор нелинейных искажений, гитарный процессор
Шумоподавление	Шумоподаватель, устрани- тель щелчков и Noise Gate	Адаптивные шумоподаватель и устранитель щелчков

Общий недостаток графических процессоров редактора Audition 3 это использование черного фона окон. При этом графики хорошо смотрятся на экране монитора и трудно различимы при печати. Для сравнения на рис. 6. приведены окна

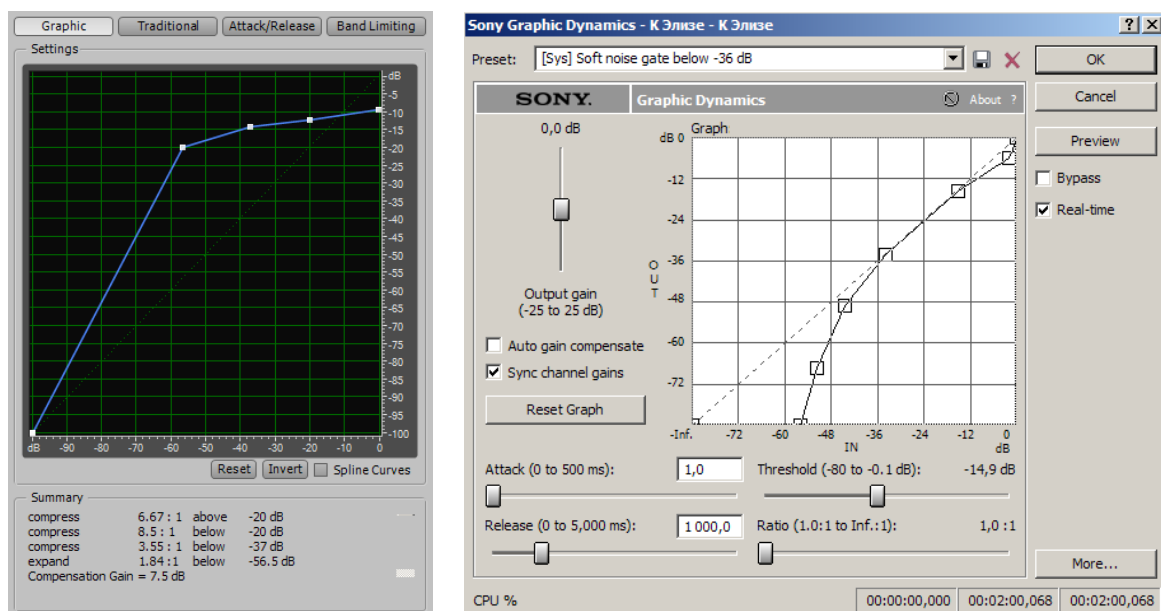
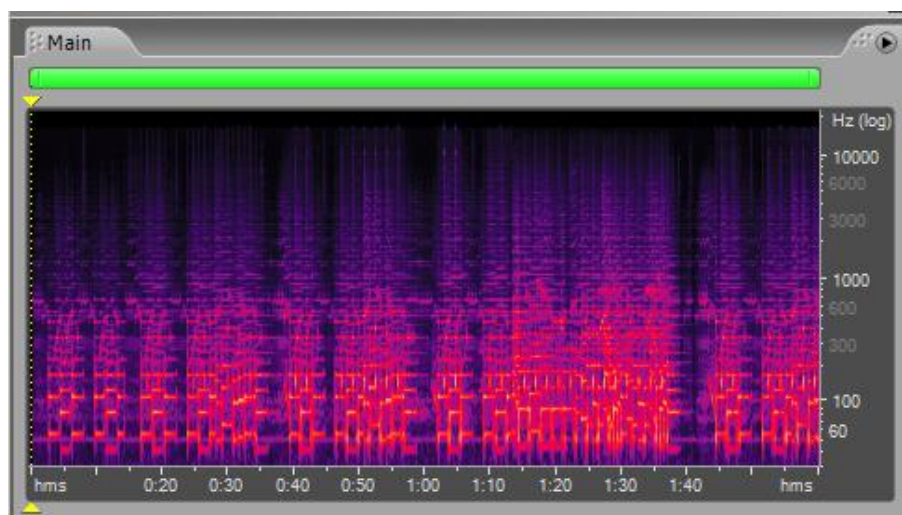
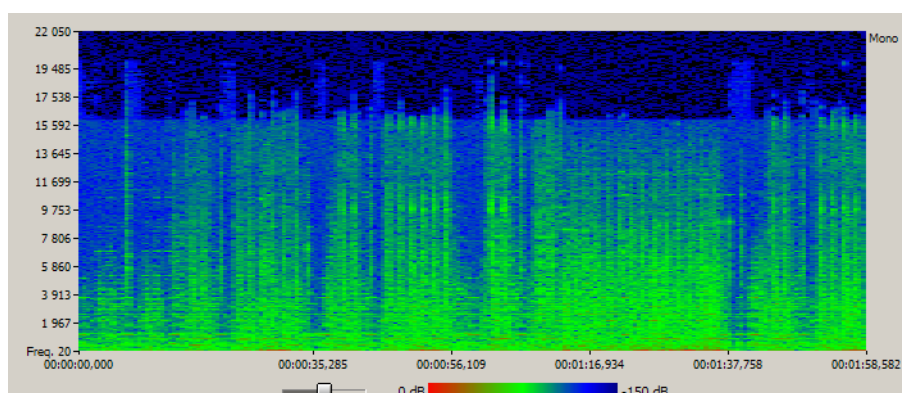


Рис. 6. Окна графических динамических процессоров редакторов Audition 3 (a) и Sound Forge (b) графических динамических процессоров двух редакторов. В редакторе Sound Forge 9c все функции управления сосредоточены в одном окне, в редакторе Audition 3 те же функции распределены по 4 окнам. Такие же незначительные отличия имеют место при сравнении и других процессоров.



a)



b)

Рис.7. Сонограммы музыки Бетховена. a) Audition 3 b) Sound Forge 9

В большинстве звуковых редакторов измеряется текущий спектр музыкальной программы, усредненный за достаточно длительный промежуток времени. Он

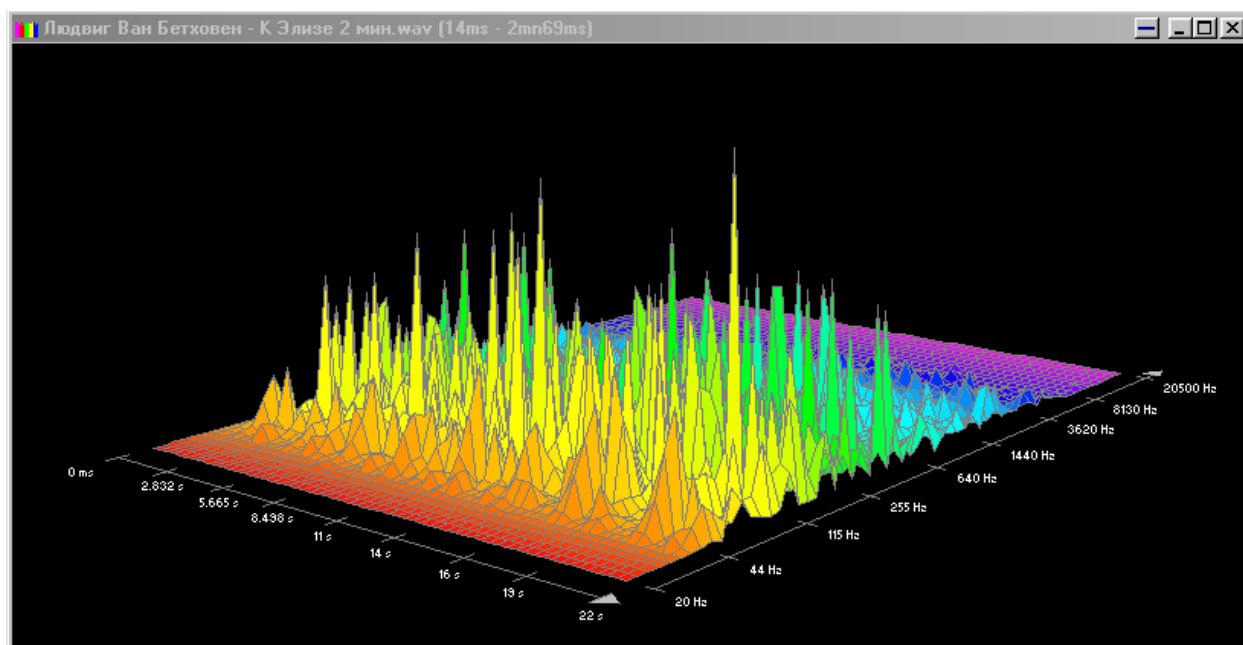


Рис.8. Трехмерный спектр из музыки Бетховена (отрывок 2 мин). Редактор WavLab

практически не связан с качеством звучания и используется как статистическая характеристика, вроде средней температуры больных по палате или всей больнице.

Практический интерес для звукорежиссеров представляет мгновенный спектр в виде сонограмм, приведенных на [рис.7](#). Как видно у одного и того же музыкального отрывка в разных редакторах даже близко не похожие сонограммы. Более наглядным представляется трехмерный спектр, [рис.8](#), но измерить его можно, например, с помощью звукового редактора WavLab.