



Современные системы цифрового телевидения

Старт 2-клик
Стоп - 1 клик

Лекция 5

Цифровое представление и сжатие звуковых сигналов

ФИО преподавателя: Смирнов

Александр Витальевич

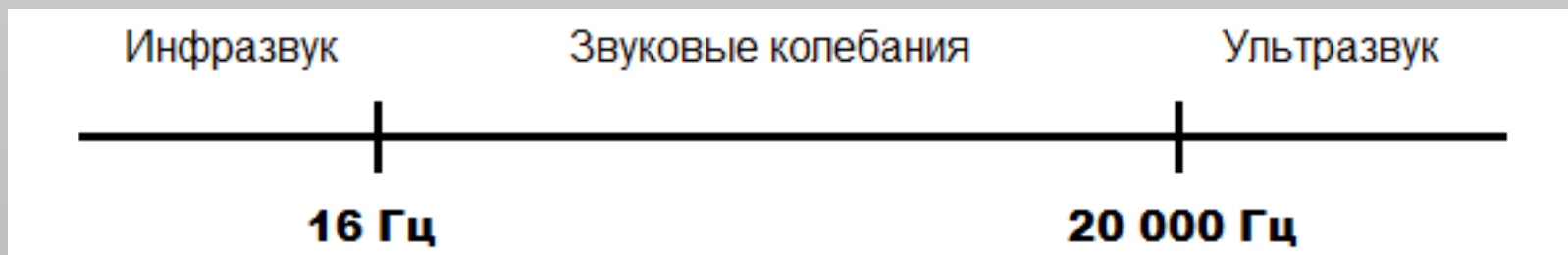
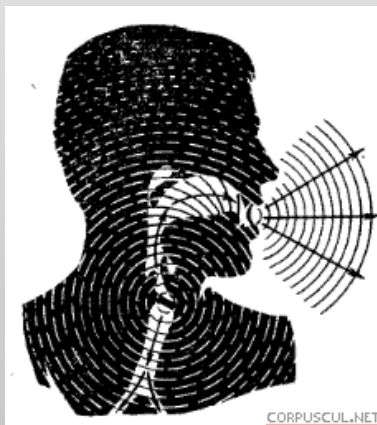
e-mail: av_smirnov@mirea.ru





Природа звука

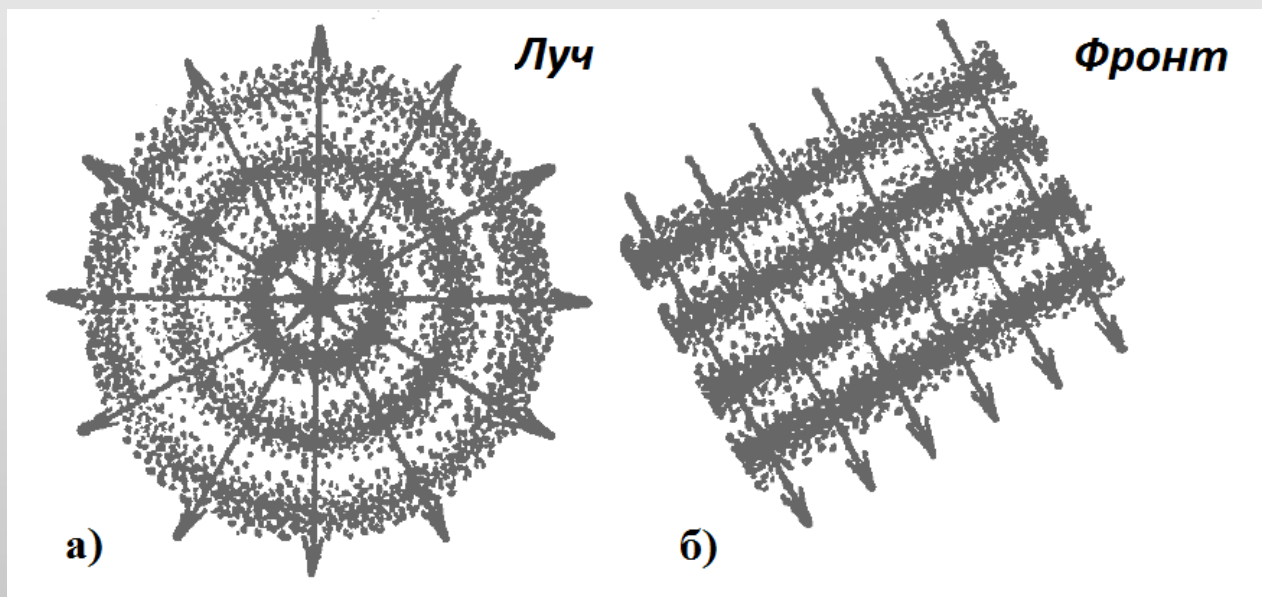
Звук - это механические колебания в различных средах, воспринимаемые слухом.



Скорость звука в воздухе $c_{зв} \approx 340$ м/с.



Звуковые волны



В звуковой волне чередуются сгущения и разрежения воздуха. Изменения давления в волне относительно постоянного атмосферного давления наз. звуковым давлением. Оно составляет малую долю атмосферного давления, равного в норме 100000 Па.



Интенсивность звука

Для плоской звуковой волны скорость колебательного движения частиц воздуха определяется формулой

$$v = p / \rho c_{зв},$$

где p - звуковое давление (Па),

ρ - плотность воздуха (кг/м³).

Интенсивность звука - количество энергии, проходящее за 1 секунду через единицу площади.

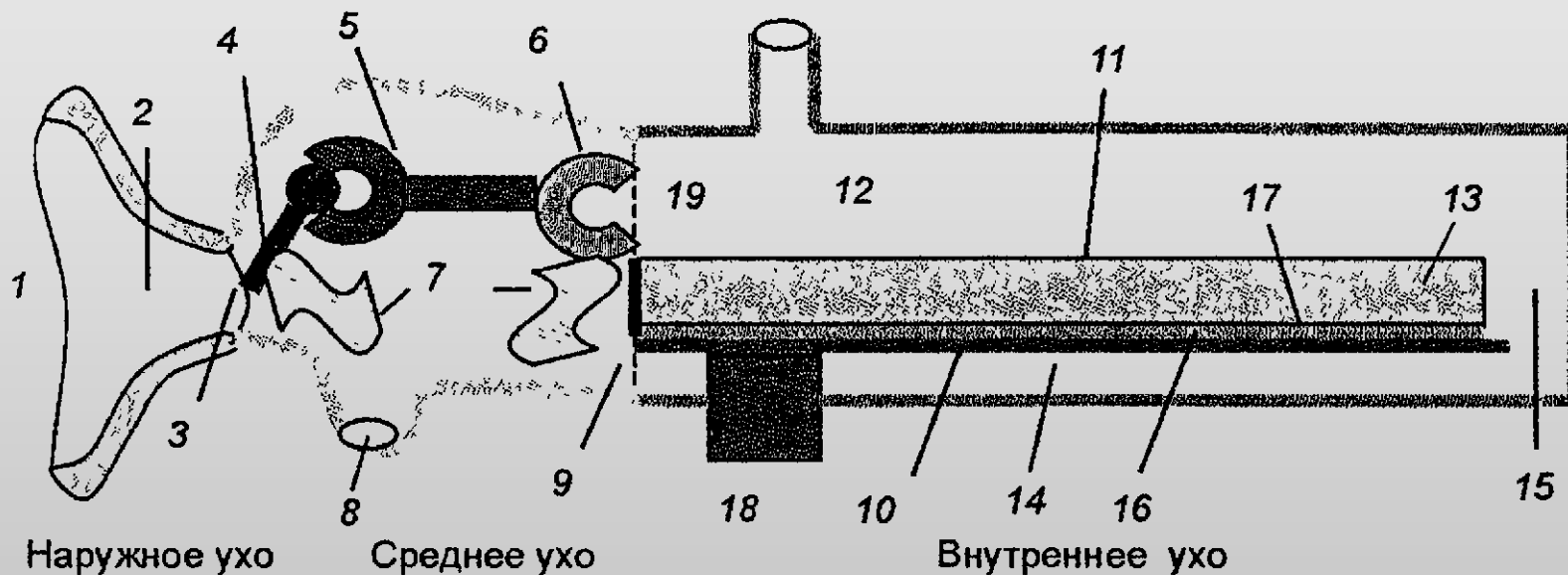
$$J = p_{эф} v_{эф} = p_{эф}^2 / \rho c_{зв} \quad (\text{Вт} / \text{м}^2),$$

где $p_{эф}$ и $v_{эф}$ - действующие значения звукового давления и скорости.



Структура органа слуха

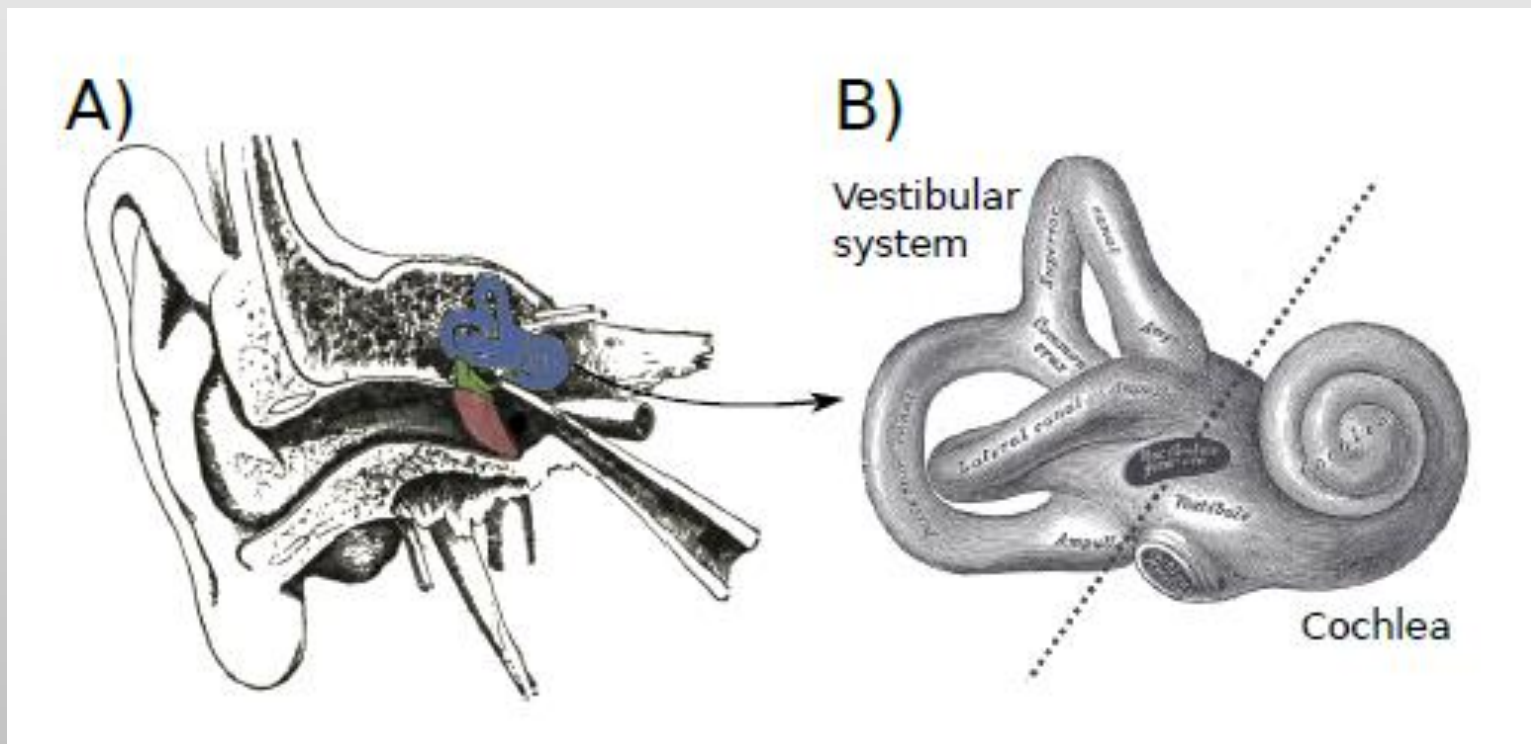
Длина улитки примерно 35 мм



3. Барабанная перепонка 4. Молоточек 5. Наковальня 6. Стремя 7. Мышцы
8. Евстахиева труба 9. Окно улитки 10, 11. Мембраны 12, 13, 14 Каналы,
заполненные лимфой 15. Геликоте́рма - соединяет каналы в вершине улитки
16. Кортиев орган – около 22 тысяч волосковых клеток-рецепторов
17. Покровная мембрана 18. Слуховой нерв 19. Окно
Разные частоты воспринимают рецепторы разной длины.



Реальный вид органа слуха и улитки

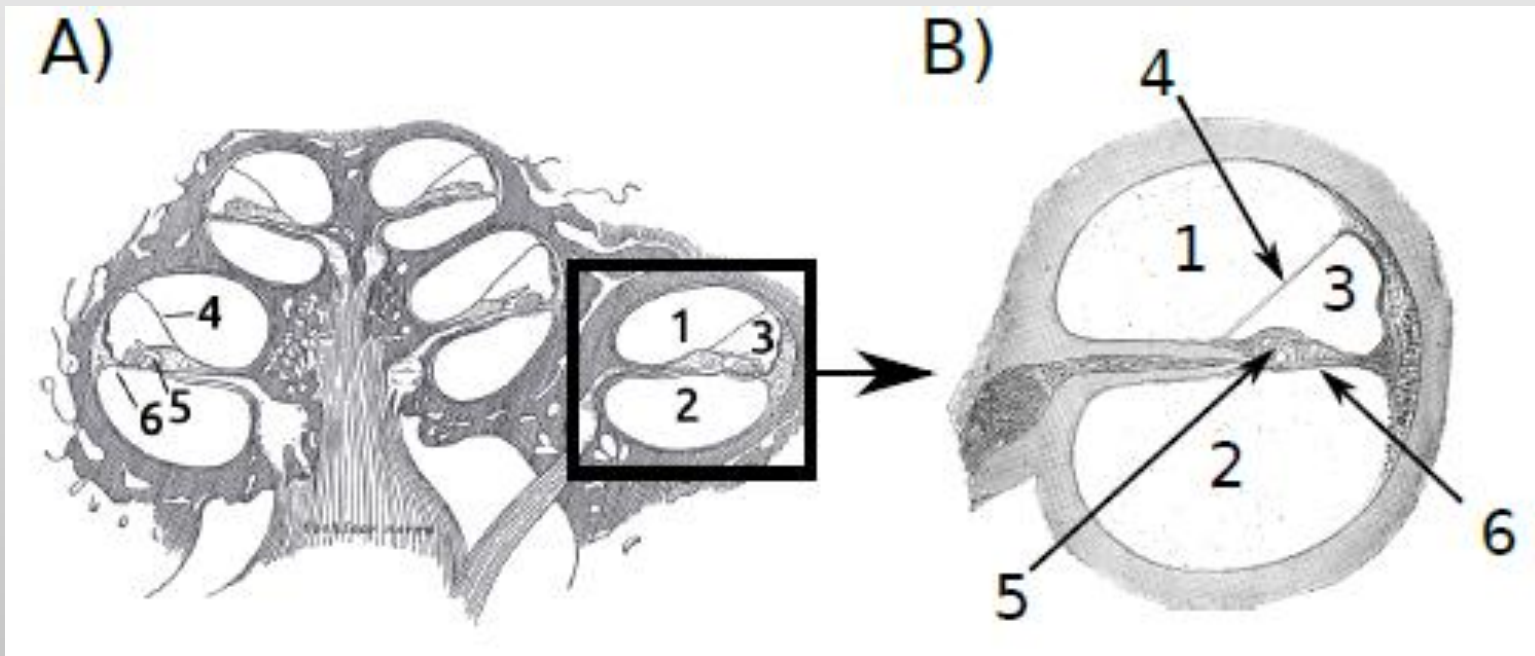


A) Орган слуха: барабанная перепонка (красная), косточки (зеленые), внутреннее ухо (синее)

B) В) внутреннее ухо: вестибулярная система и улитка



Вид сечения улитки

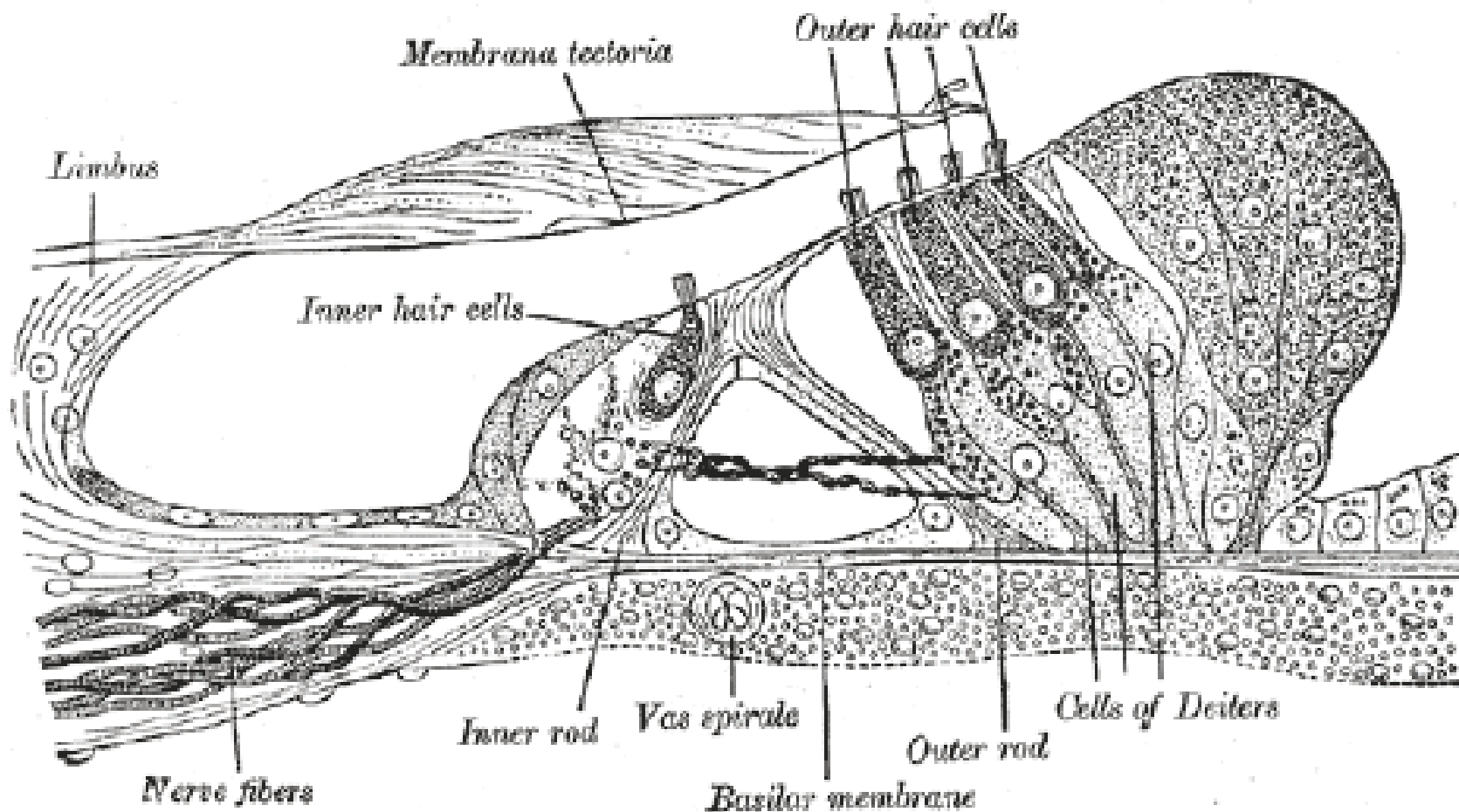


A) Сечение улитки B) увеличенное сечение каналов

1) Scala vestibuli; 2) Scala tympani; 3) Scala media; 4) Reissner membrane;
5) Corti's organ; 6) Basilar membrane.



Кортиев орган





Восприятие звука

Многие виды восприятия внешних воздействий, в том числе и слух, подчиняются закону Вебера - Фехнера:

$$\Delta X_{\text{пор}} = k X_{\text{фон}},$$

где $X_{\text{фон}}$ - фоновое значение воздействия, k - коэффициент, $\Delta X_{\text{пор}}$ - минимально различимое изменение воздействия.

Следовательно, на низком фоне различимы малые изменения, а на высоком фоне - только большие. Поэтому шкалы восприятия, как правило, логарифмические.



Восприятие частоты звука

Для восприятия частоты $k = 0,002 \dots 0,01$ (0,2 ... 1,0 %).

Ощущение частоты звука называется высотой тона.

Октава - изменение частоты в 2 раза.

Полутон - $1/12$ октавы - минимальный частотный интервал в музыке.

Звуковые колебания обычно содержат основной тон и его высшие гармоники. Тембр звука определяется числом и амплитудами гармоник.



Уровень интенсивности звука

Акустический уровень интенсивности:

$$N_J = 10 \lg(J/J_0) \quad (\text{дБ}),$$

где $J_0 = 10^{-12} \text{ Вт/м}^2$ соответствует порогу слышимости звука на частоте 1000 Гц.

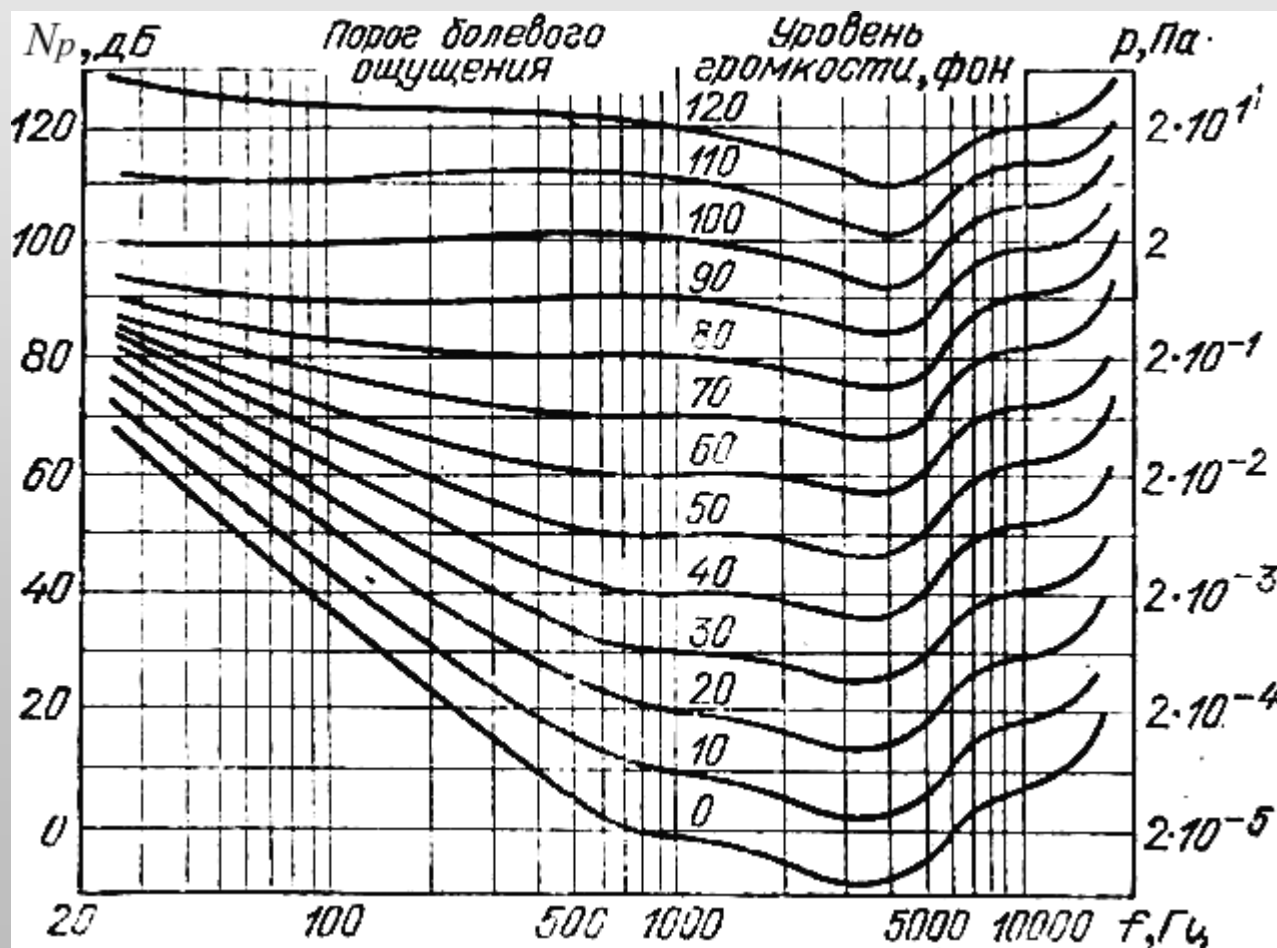
Акустический уровень звукового давления

$$N_p = 20 \lg(p/p_0) \quad (\text{дБ}),$$

где $p_0 = 2 \cdot 10^{-5} \text{ Па}$ соответствует порогу слышимости звука на частоте 1000 Гц.



Кривые равной громкости



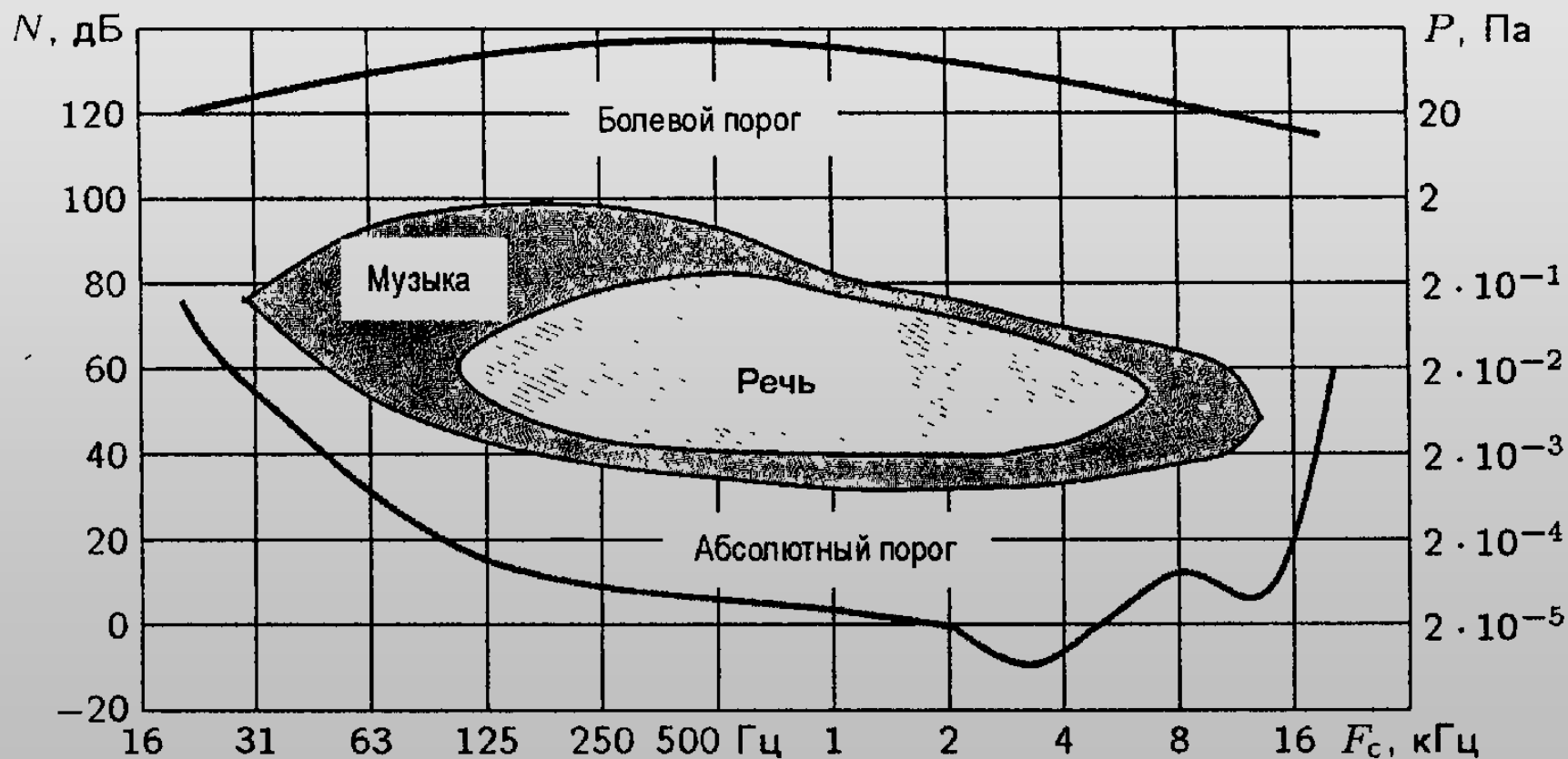


Уровень громкости

Уровень громкости измеряется в единицах, называемых *фон* и численно равен уровню давления или интенсивности звука с частотой 1000 Гц, громкость которого на слух равна громкости измеряемого звука.



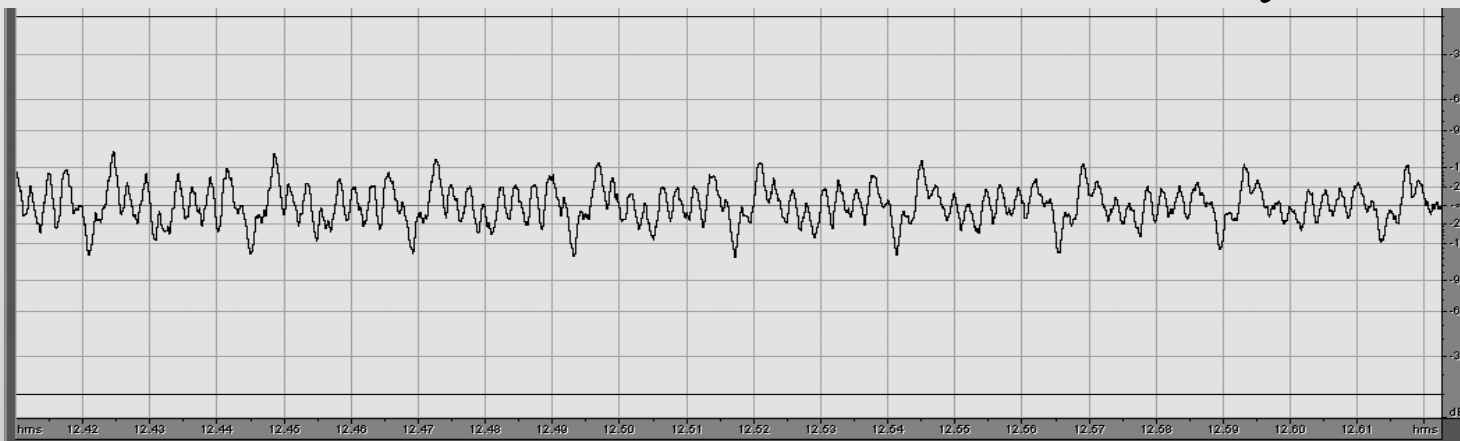
Диапазоны частот и громкостей



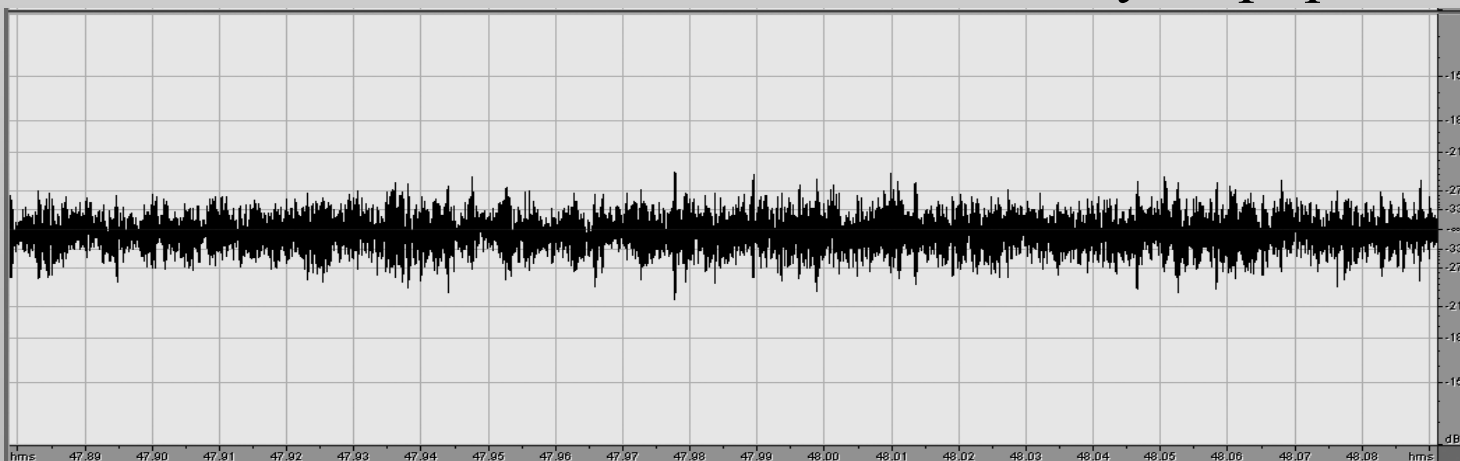


Форма сигнала (примеры)

Музыка

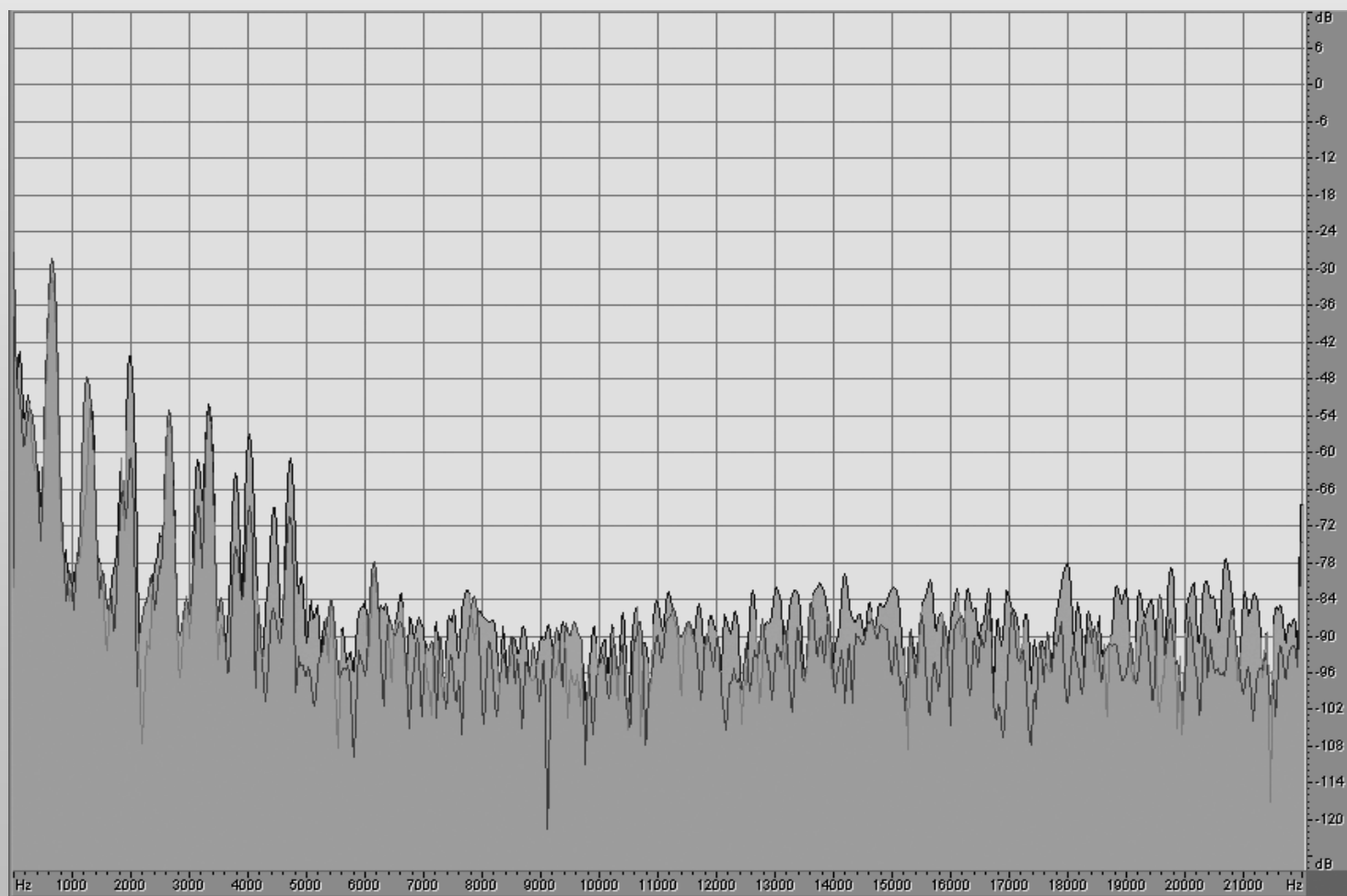


Шум природы



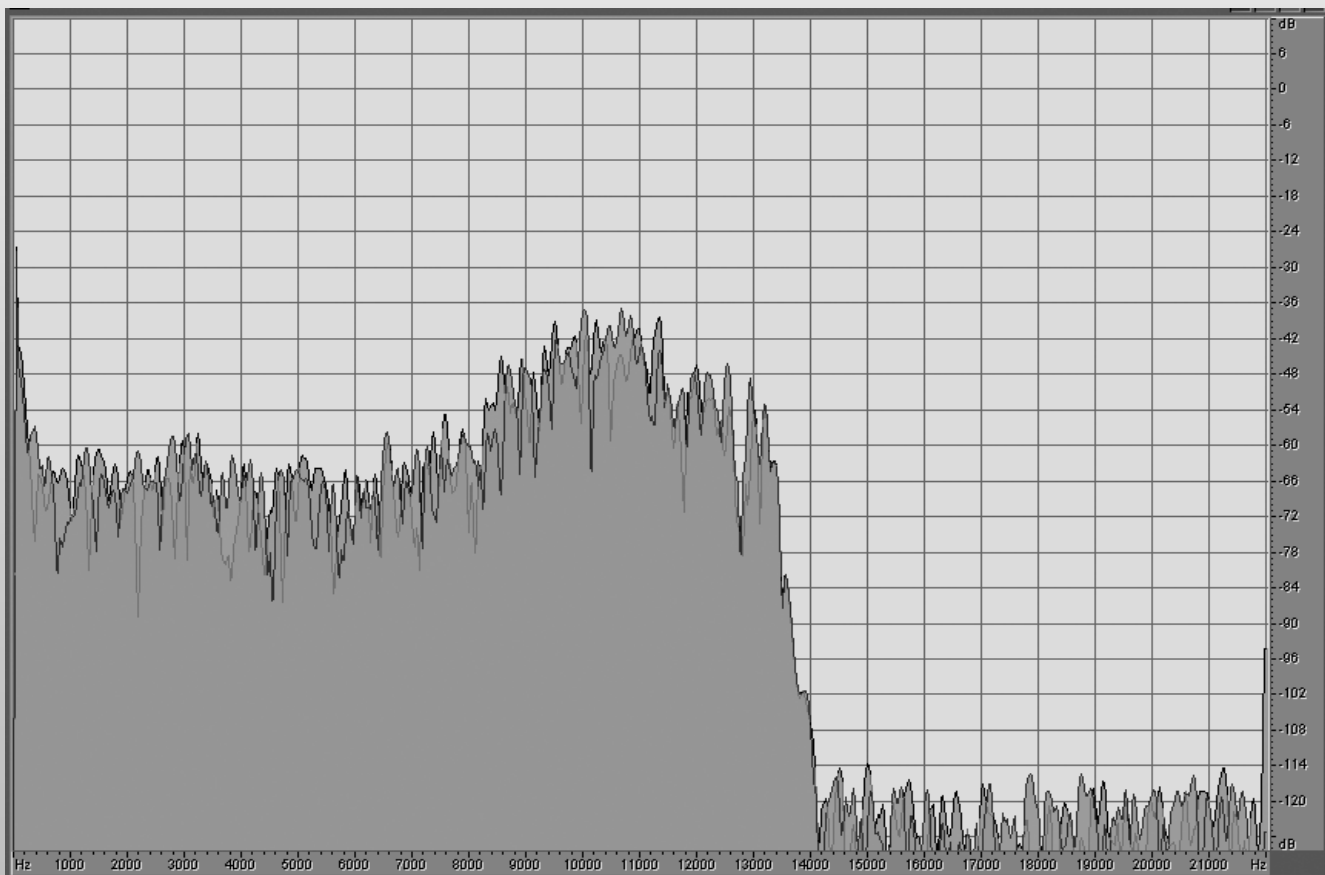


Спектр сигнала музыки





Спектр сигнала шума природы



Спектр ограничен до 12 кГц при оцифровке аудиозаписи



Уровень звукового сигнала

Уровень N_U сигнала звукового вещания определяется относительно номинальной величины напряжения сигнала $U_{\text{ном}} = 0,775 \text{ В}$.

$$N_U = 20 \lg \frac{U_c}{U_{\text{ном}}} \text{ дБ}$$

Динамический диапазон сигнала звукового вещания определяется как разность максимального и минимального уровней сигнала за время контроля.

$$\Delta N_U = N_{U_{\text{макс}}} - N_{U_{\text{мин}}}$$

$N_{U_{\text{макс}}}$ - уровень, который превышает не более чем в течении 1 - 2 % времени.

$N_{U_{\text{мин}}}$ превышает в течении 98 - 99% времени.



Параметры звуковых сигналов

Воспроизводимые диапазоны частот для разных классов аппаратуры:

Высший: 30 - 15000 Гц.

Первый: 50 - 10000 Гц.

Второй: 100 - 6300 Гц.

Диапазон частот в телефонной сети: 300 - 3400 Гц.

Динамические диапазоны некоторых источников звукового сигнала:

Речь диктора: 25 - 35 дБ.

Симфонический оркестр: 65 - 75 дБ.



Линейные искажения звука

Неравномерность АЧХ приводит к ухудшению воспроизведения некоторых диапазонов частот.
(например, высоких или низких частот)

Нелинейность ФЧХ может сказаться при воспроизведении стереофонического звука.



Нелинейные искажения звука

Нелинейные искажения создают высшие гармоники и приводят к появлению заметных искажений.

(типа хрипения, дребезжания и т.д.)

Коэффициент гармоник при практических измерениях:

$$K_{\Gamma} = \frac{\sqrt{U_2^2 + U_3^2}}{U_1},$$

Допустимый коэффициент гармоник
на частотах выше 200 Гц:

- высший класс 1,5%;
- первый класс 3-3,5%;
- второй класс 4-5%.



Защищенность от интегральной помехи

Интегральная помеха - это сумма всех помех, действующих в тракте ЗВ.

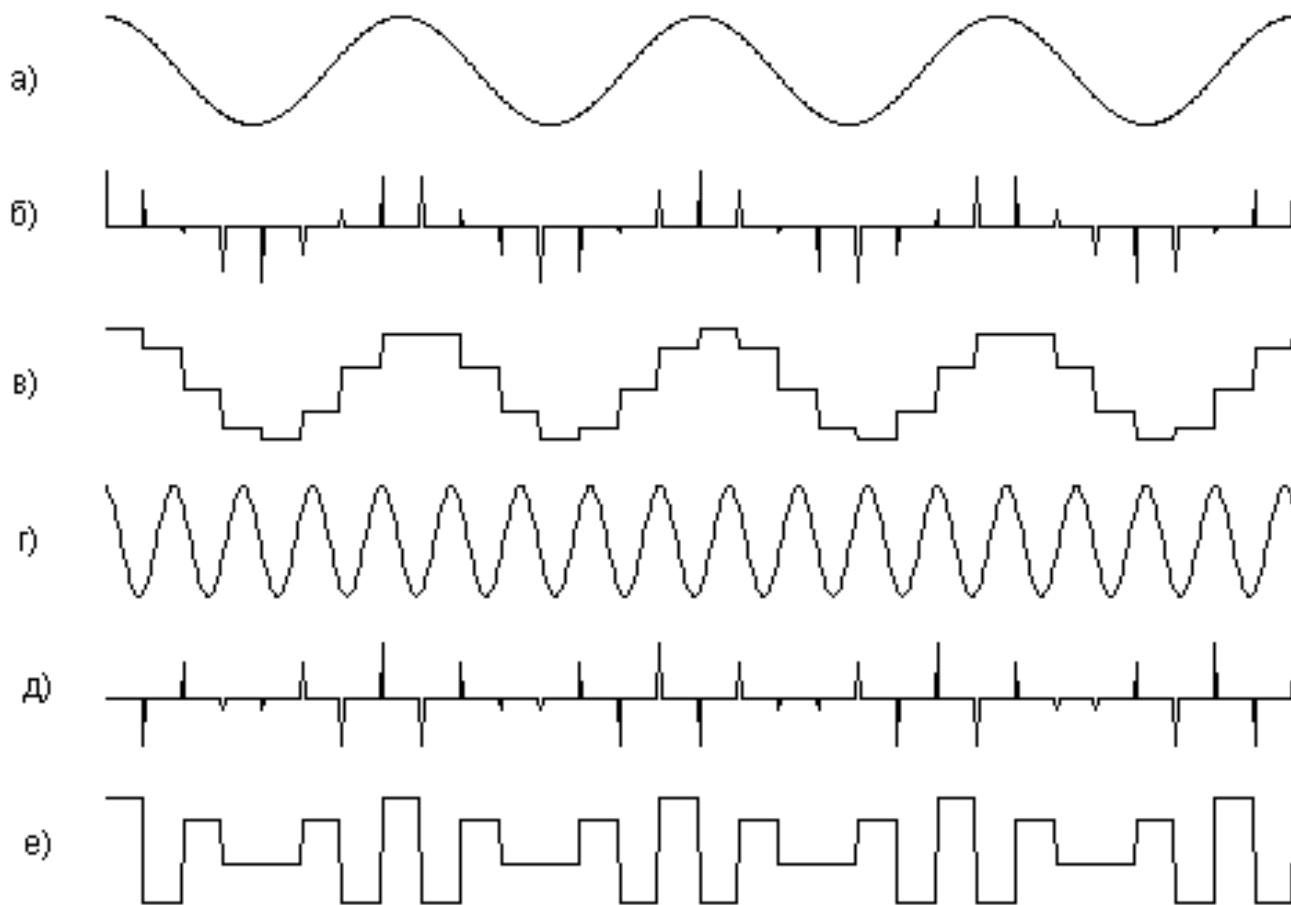
$$A_{\text{ш}} = 20 \lg (U_c / U_{\text{ш}}) \text{ дБ},$$

где U_c - номинальное действующее напряжение сигнала;

$U_{\text{ш}}$ - действующее напряжение шума.



Дискретизация сигнала





Условие Котельникова - Найквиста

Частота дискретизации f_d должна удовлетворять условию:

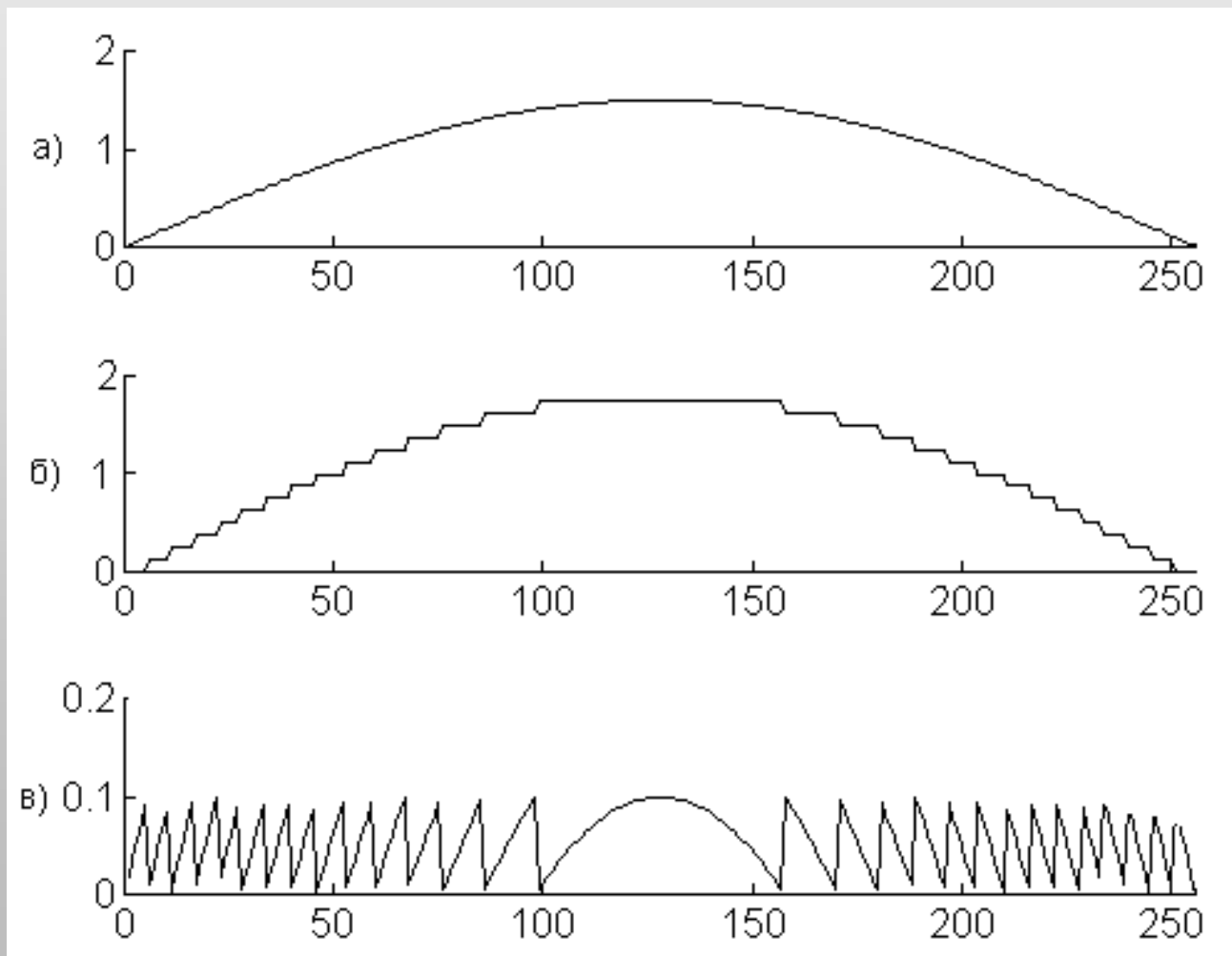
$$f_d > 2 f_v,$$

где f_v - верхняя граничная частота аналогового сигнала.

При нарушении этого условия в спектре сигнала появляются ложные составляющие, которые искажают его форму. Этот эффект называется *aliasing* ("элайзинг") - наложение.



Квантование сигнала





Шум квантования

$$\sigma_{\text{кв}} = \frac{h}{\sqrt{12}} \quad h - \text{шаг квантования}$$

Относительный уровень шума квантования в логарифмическом масштабе (обратная величина к $C/\text{Ш}_{\text{кв}}$)

$$\text{Ш}_{\text{кв}} = 20 \lg \left(\frac{h}{\sqrt{12}} \right) = 20 \lg \left(\frac{2^{-b}}{\sqrt{12}} \right) \approx -(6,02b + 10,8) \text{ дБ.}$$

b – число разрядов квантования.

При $b = 7$ получаем $\text{Ш}_{\text{кв}} \approx -53$ дБ (телефон).

При $b = 15$ получаем $\text{Ш}_{\text{кв}} \approx -101$ дБ (CD).



Цифровое представление звука

Для передачи речи достаточна частота дискретизации 8 кГц, число разрядов АЦП 8.

В звукозаписи и вещании применяют частоты дискретизации 44,1 и 48 кГц; 16-разрядное квантование по уровню.

В студийной аппаратуре может быть: 18-, 20-, 24- разрядное квантование; частоты дискретизации 96 и 192 кГц.

Это позволяет сохранить высшие гармоники сигнала, не воспринимаемые слухом, но влияющие на формирование общей звуковой картины.



Скорость цифрового потока звукового сигнала

$$V_B = f_d b \text{ (бит/с)},$$

где f_d - частота дискретизации,

b - число двоичных разрядов квантования.

При $f_d = 48$ кГц, $b = 16$ $V_B = 768$ кбит/с на канал,
 $\approx 1,5$ Мбит/с для стереозвука.



Динамический диапазон звукового сигнала

Максимальное $C/\mathcal{W}_{\text{кв}}$ при $b = 16$ около 100 дБ.

Ограничивающие диапазон факторы:

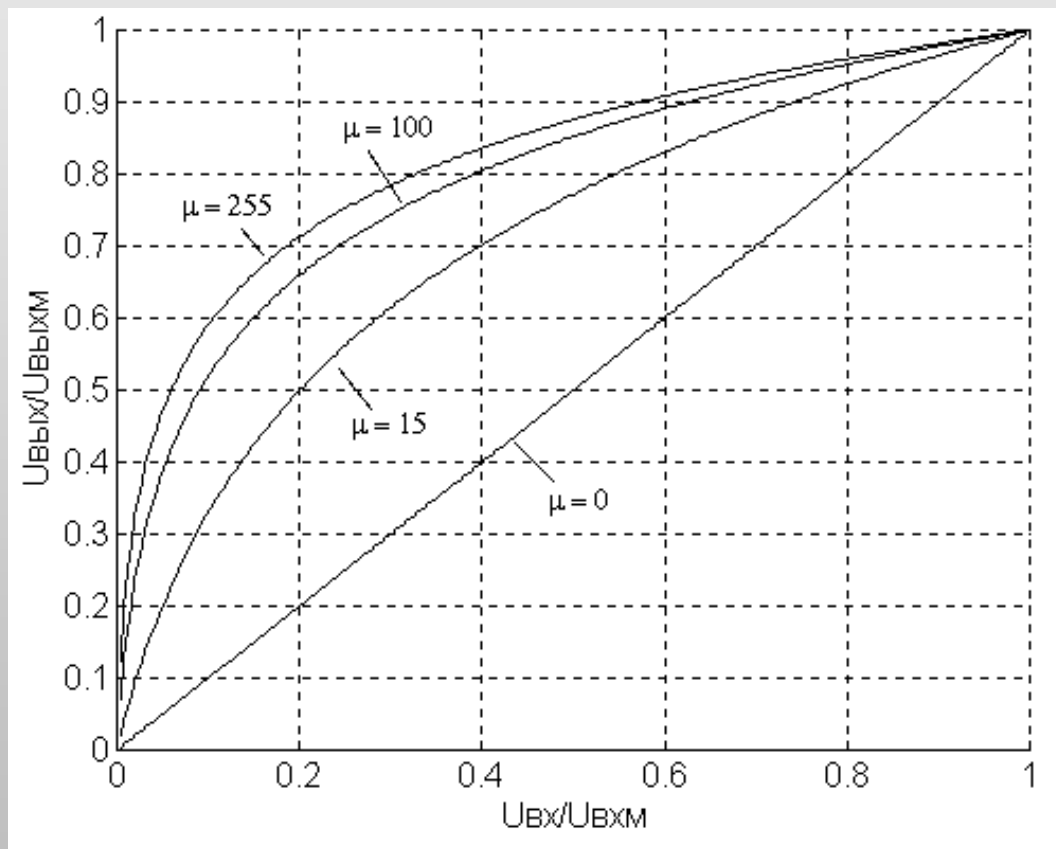
- учет пик-фактора звукового сигнала - 12 дБ;
- запас сверху, чтобы предотвратить ограничение, примерно 10 дБ;
- запас снизу, чтобы тихие звуки не тонули в шуме, примерно 20 дБ.

Реальный динамический диапазон 55 - 60 дБ.

Компандирование перед АЦП улучшает C/\mathcal{W} для тихих звуков, ухудшая C/\mathcal{W} для громких звуков, но на фоне громких звуков шум малозаметен.



Компандирование звукового сигнала



Для Hi-Fi $\mu = 15$;
для телефона
 $\mu = 100 - 255$.

10.8.20

$$u_{\text{ВЫХ}} = u_{\text{ВХ МАКС}} \frac{\ln(1 + \mu |u_{\text{ВХ}}| / u_{\text{ВХ МАКС}})}{\ln(1 + \mu)},$$

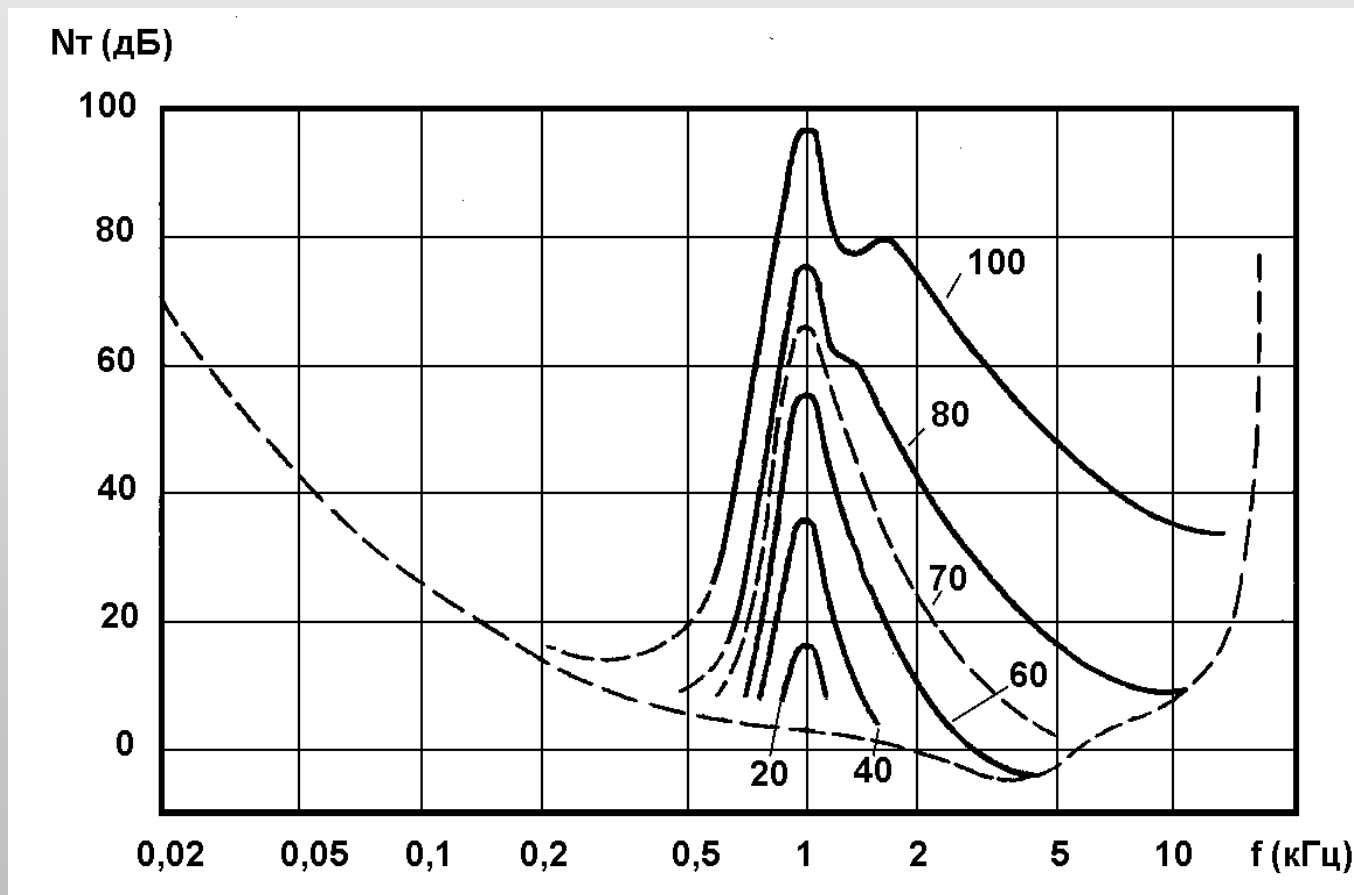


Принципы сжатия звука

1. Используются маскировка громкими звуками более тихих в близких частотных поддиапазонах и различная чувствительность слуха на разных частотах.
2. Сжатие основано на адаптивном распределении битов, то есть раздельном квантовании сигналов частотных поддиапазонов в зависимости от их слышимости.



Частотные характеристики маскировки

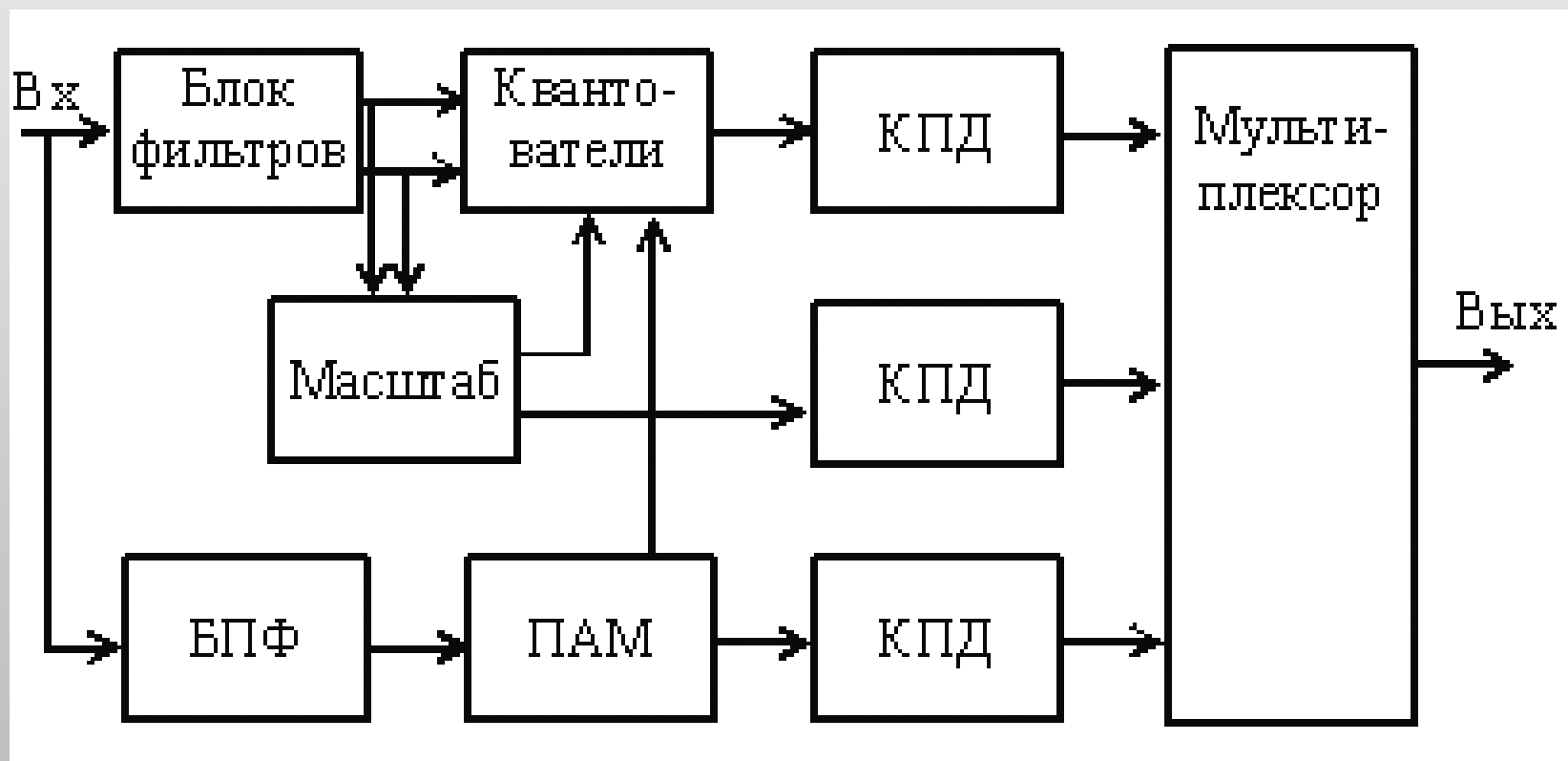


Показаны зависимость порога слышимости от частоты (пунктир) и области маскировки звуком с частотой 1 кГц.

10.8.20



Кодер звука MPEG





Звуковые сигналы в MPEG-2

Частоты дискретизации:

- 48,0; 44,1; 32,0; 24,0; 22,05 и 16 кГц

Число разрядов квантования: 16, 20

Каналы: L, R – базовый поток,
C, Ls, Rs, Sub – расширения.

Звуковой кадр: 384 отсчета на Layer 1
1152 отсчета на Layer 2 и 3



Операции при сжатии звука

1. Разложение на 32 равных частотных поддиапазона по 36 отсчетов в каждом (Блок фильтров).
2. Оценка уровня звука в каждом поддиапазоне (БПФ).
3. Определение масштабных множителей для групп по 12 отсчетов с целью приведения к единому динамическому диапазону (Масштаб).
4. Оценка слышимости звуков и заметности их искажений при сжатии. Распределение битов по поддиапазонам (ПАМ).
5. Квантование в поддиапазонах в соответствии с выделенным количеством битов (Квант).
6. Энтропийное кодирование данных. Формирование потока звуковых данных (Мп - мультиплексор).

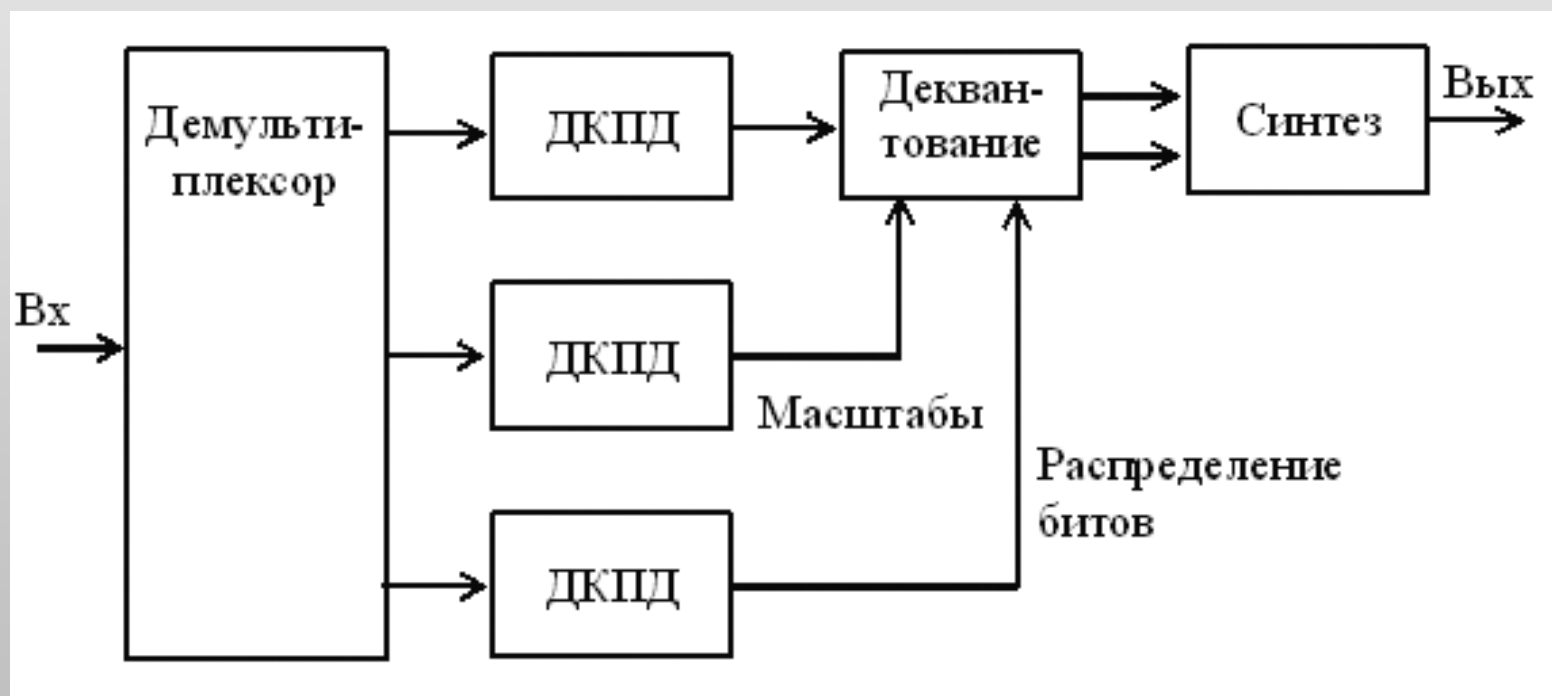


Особенности Layer 3 (MP3)

1. Каждый частотный поддиапазон с помощью Дискретного Косинусного Преобразования (ДКП) разделяется еще на 18 частотных составляющих. Это позволяет более точно учесть слышимость различных звуков и достичь большего сжатия при сохранении требуемого качества звука
2. Распределение битов выполняется в несколько проходов с проверкой получаемого качества звука после каждого прохода. Это также способствует улучшению результатов.



Декодер звука MPEG





Структура потока звуковых данных

Поток данных состоит из звуковых кадров (*Frame*).

Звуковой кадр начинается с заголовка:

- синхрослово (11111111111b или FFFh);
- параметры режима кодирования.

Затем идут данные каналов L и R – базовая часть звукового кадра, совместимая с MPEG-1.

Затем – расширения, содержащие данные других каналов объемного звука.

Звуковой кадр также содержит контрольные слова, позволяющие обнаруживать ошибки.



Сравнение уровней MPEG-2 аудио

Скорость потока до сжатия примерно 700 кбит/с на канал.
Приемлемое качество звука после сжатия обеспечивается при:

Layer 1 - 192 кбит/с на канал;

Layer 2 - 128 кбит/с на канал

(используется в радио- и ТВ-вещании);

Layer 3 - 64 кбит/с на канал

(используется для звукозаписи “MP3”);

AAC (Advanced Audio Coding) - 32 кбит/с на канал.



Advanced Audio Coding - AAC

Усовершенствованное кодирование звука.

Используется в цифровом ТВ вещании.

Расширен набор частот дискретизации: от 8 до 96 кГц.

Длина кодируемого отрезка 2048 или 256 отсчетов.

Как и в Layer 3 кодируются коэффициенты ДКП.

Усовершенствованы алгоритмы оценки заметности искажений при сжатии и распределения битов при квантовании.

Сжатие улучшено в ≈ 2 раза по сравнению с MP3.

Дальнейшее развитие - AAC Plus, сжатие до 48 кбит/с для стереозвука. Используется в DRM.



Спасибо за внимание!