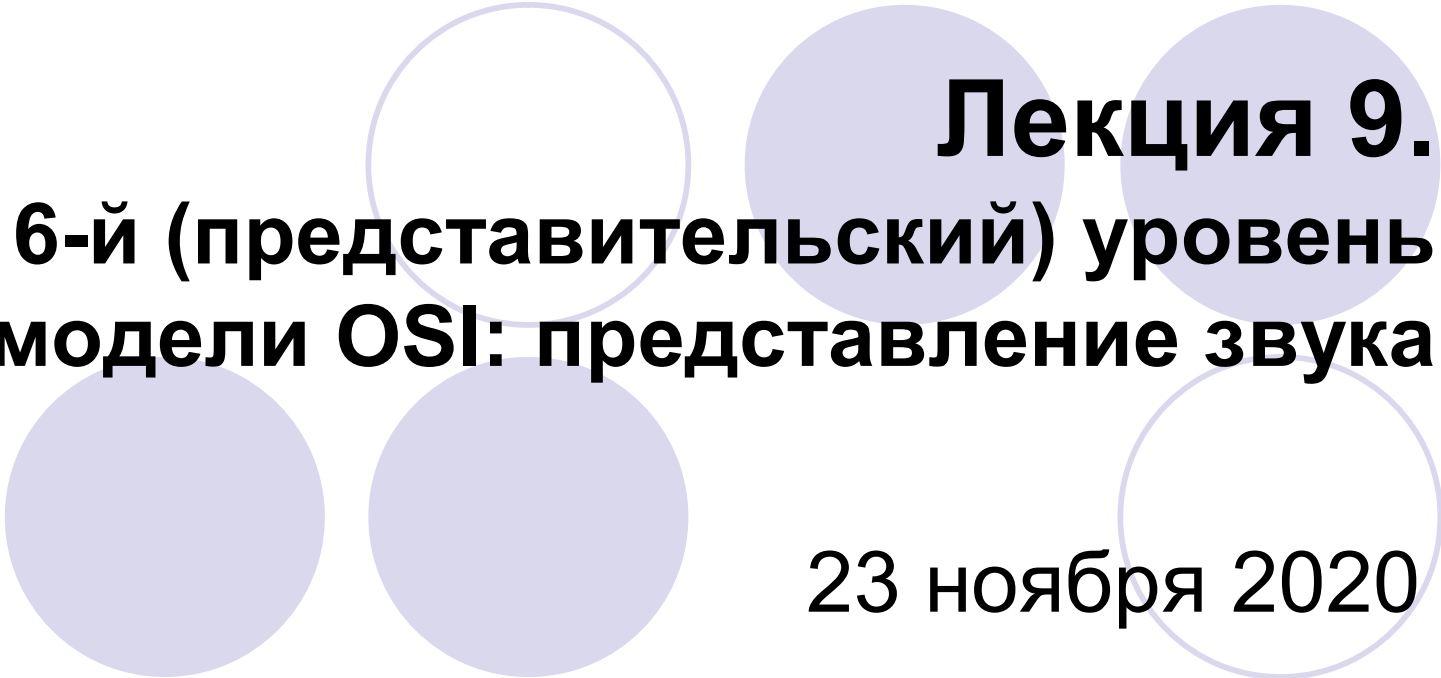


КПИ им. Игоря Сикорского, каф. микроэлектроники.

М. Р. Домбругов. Информатика-1.

Персональные компьютеры и основы сетевых технологий



Лекция 9.

6-й (представительский) уровень модели OSI: представление звука

23 ноября 2020

7-уровневая модель OSI

Тип данных	Уровень (layer)	Функции
Данные	7. Прикладной (application)	Доступ к сетевым службам
	6. Представительский (presentation)	Представление и шифрование данных
	5. Сеансовый (session)	Управление сеансом связи
Сегменты	4. Транспортный (transport)	Прямая связь между конечными пунктами и надежность
Пакеты	3. Сетевой (network)	Определение маршрута и логическая адресация
Кадры	2. Канальный (data link)	Физическая адресация
Биты	1. Физический (physical)	Работа со средой передачи, сигналами и двоичными данными

6-й уровень модели OSI

Представление и шифрование данных

- Представление текста
- - " - чисел
- **Представление звука**
- - " - изображений
- - " - видео
- Оптимальное кодирование и архиваторы
- Шифрование данных
- ...

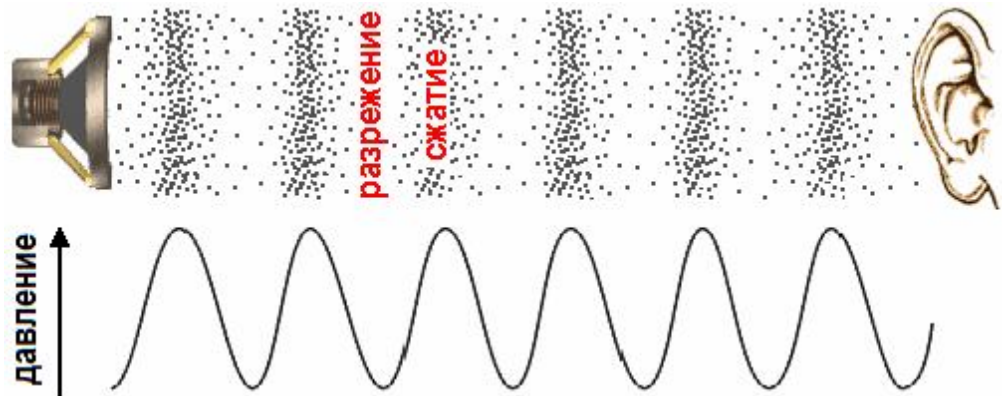
- Спектр периодического сигнала
- Природа и характеристики звука
- Цифровая звукозапись
- Оцифровка с уплотнением
- Музыкальный строй
- MIDI интерфейс

Природа звука

Звук – это волнообразные сгущения и разрежения воздуха (или другой среды), **воспринимаемые ухом человека**.

Механическое движение в упругой среде вызывает волнообразное колебание её плотности: частицы вещества колеблются относительно своего первоначального положения, это проявление упругости вещества.

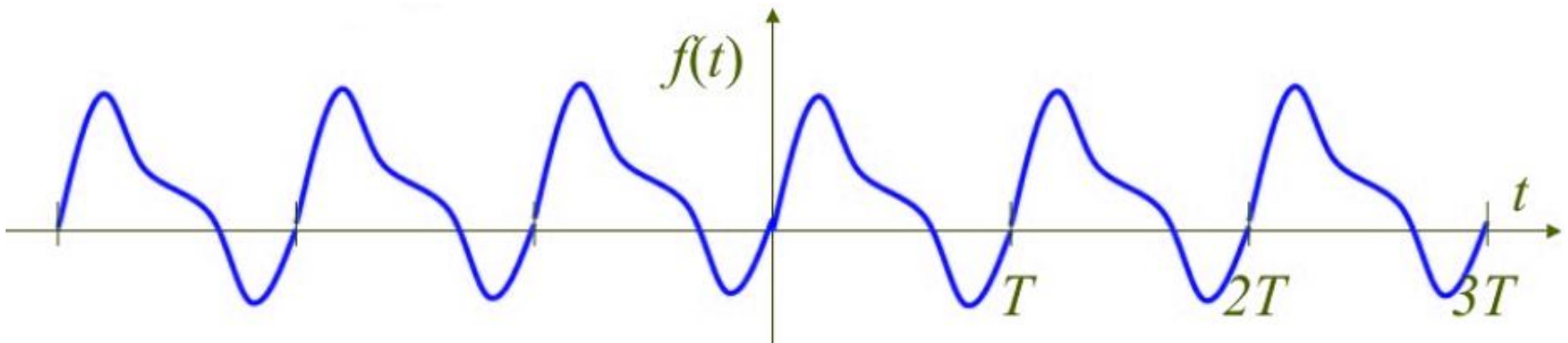
Звуковые волны – **продольные**, в них смещение частиц среды происходит в направлении распространения волны.



(В отличие от **поперечных** волн, напр. волны на воде от брошенного камня)

Спектр периодического сигнала

Любая периодическая функция может быть разложена на сумму гармонических (синусоидальных) составляющих (ряд Фурье).



$$f(t) = A_0 + \sum_{n=1}^{\infty} \left[A_n \sin \left(2\pi n \frac{t}{T} + \varphi_n \right) \right]$$

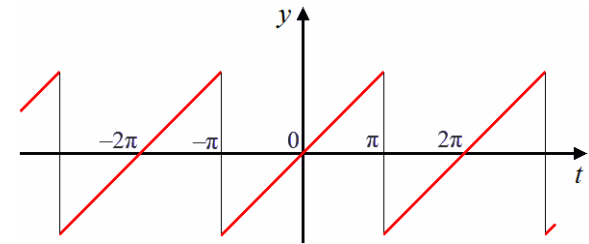
Здесь $f = 1/T$ – основная частота сигнала (1-я гармоника),
все составляющие суммы – гармоники с кратными частотами nf

Спектр периодического сигнала

Пример:

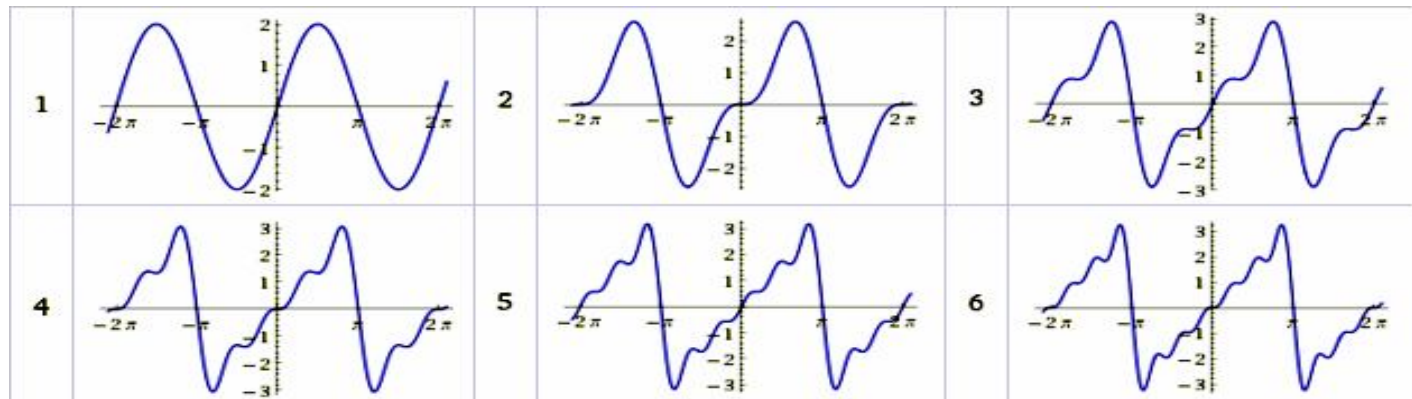
периодическая пилообразная функция, определяемая на одном периоде как

$$f(t) = t; \quad -\pi \leq t \leq \pi$$



Разложение Фурье:
$$f(t) \sim 2 \left(\sin t - \frac{1}{2} \sin 2t + \frac{1}{3} \sin 3t - \frac{1}{4} \sin 4t + \dots \right)$$

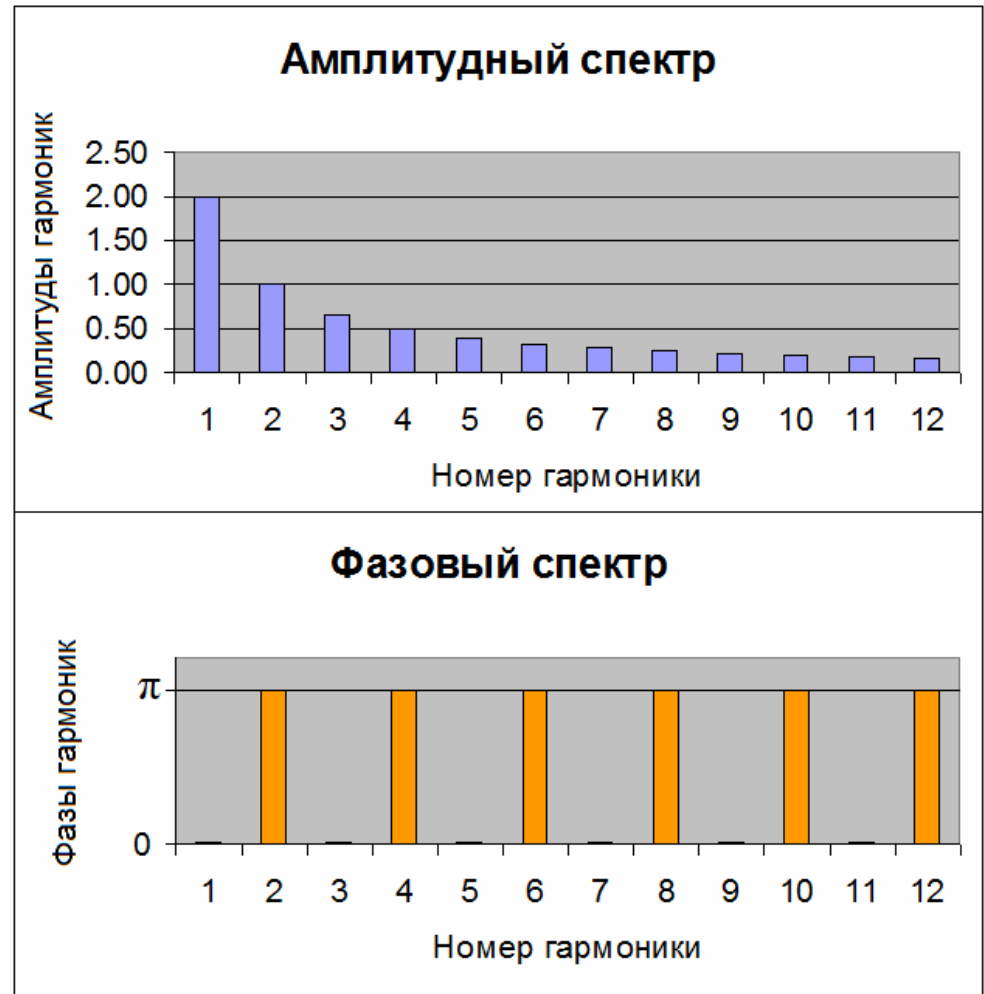
Суммы
нескольких
первых
гармоник:



Спектр периодического сигнала

Совокупность амплитуд гармоник A_n наз. **амплитудным спектром**, а их фаз φ_n – **фазовым спектром**.

Ухо чувствительно к амплитудному спектру и распознает отдельные частоты, но не чувствует различий в фазовом спектре.



Характеристики звука

Основные физические характеристики звука, воспринимаемые субъективно:

- **высота** – определяется основной частотой звуковой волны;
- **громкость** (интенсивность) – амплитудой волны;
- **тембр** – спектральным составом волны (соотношением между амплитудами отдельных гармоник).

Характеристики звука

Высота звука определяется основной частотой звуковой волны. Чем больше частота, тем выше звук. Условно считается, что ухо человека воспринимает колебания частотой **от 16 Гц до 20 кГц**.

Колебания частотой выше 20 кГц относятся к области **ультразвуков**, ниже 16 Гц – **инфразвуков**, которые воспринимаются всем телом как вибрации.

Основную звуковую информацию человеческий мозг получает в области частот до 4 кГц. Слышимость частот выше 4 кГц, как дополнение к основным частотам, создает у человека ощущение более качественного звучания.

У многих животных верхняя граница слуха выше, чем у человека. Напр., у собак – 38 кГц, кошек – 70 кГц, летучих мышей – 100 кГц.

Характеристики звука

Громкость звука определяется амплитудой звуковой волны.

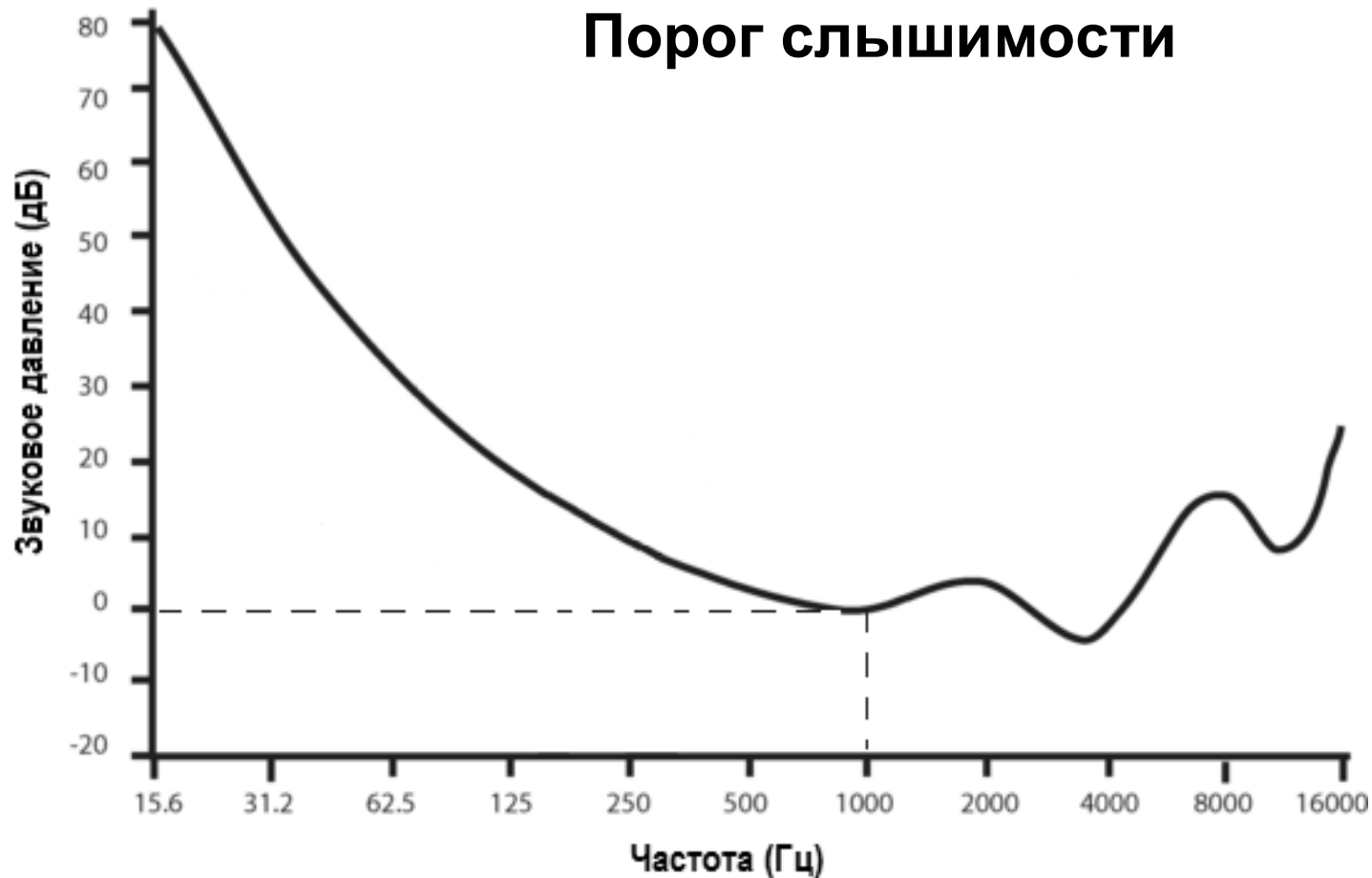
Может характеризоваться

- либо **интенсивностью**, т.е. мощностью, переносимой звуковой волной в направлении распространения [Вт/м²],
- либо **звуковым давлением** – амплитудой переменного избыточного давления, которое возникает в упругой среде при прохождении через неё звуковой волны [Па].

Интенсивность $\sim (\text{звуковое давление})^2$.

Громкость измеряется в децибелах по отношению к **опорному давлению** 20 мкПа, соответствующему **порогу слышимости** синусоидальной звуковой волны частотой 1 кГц.

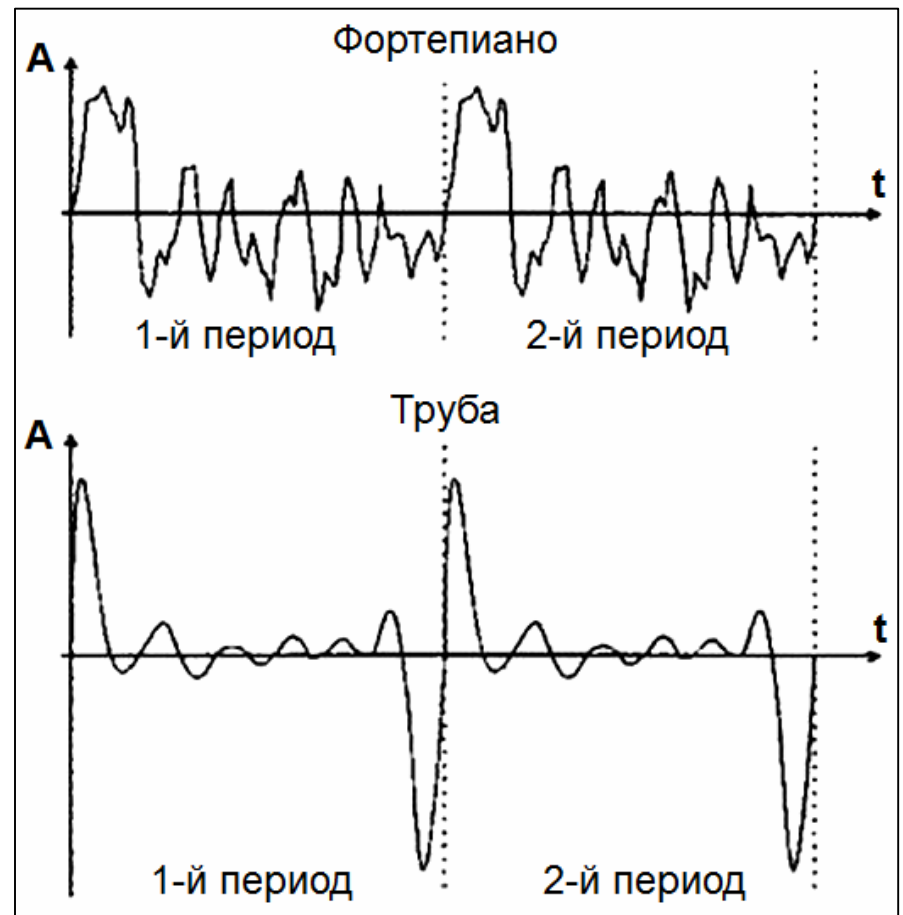
Характеристики звука



Характеристики звука

Тембр звука зависит от соотношения амплитуд составляющих его спектра – **основного тона** (1-й гармоники) и **обертонов** (высших гармоник).

Сигналограммы звучания фортепиано и трубы



Характеристики звука

Анализ спектра мозг выполняет при помощи **базиллярной мембраны**, которая помещается в улитке внутреннего уха.

Это структура переменной ширины (от 0,6 мм до 0,1 мм), жесткости, массы в разных своих точках по длине (ок. 35 мм).

Звуковой сигнал, приходящий в ухо, заставляет мембрану резонировать в разных частях. Максимальные колебания мембраны происходят в той ее части, которая соответствует частоте приходящей звуковой волны.

Вдоль мембраны расположены **волосковые клетки**. Те из них, что оказались в резонирующей области, начинают колебаться, возбуждая тем самым нервные окончания, которые передают информацию о колебаниях в мозг.

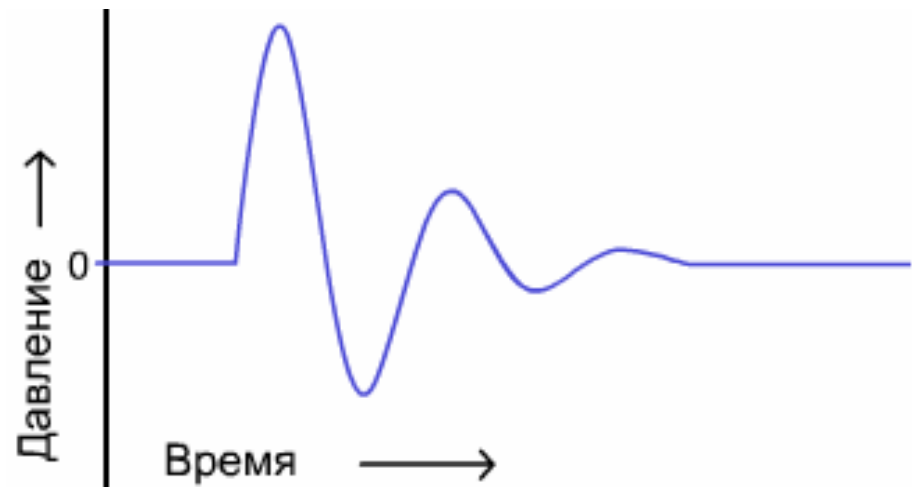
Анализируя эти нервные импульсы мозг интерпретирует их как звук той или иной высоты.

Характеристики звука

Кратковременные звуковые всплески не могут заставить базилярную мембрану резонировать на нужной частоте, а значит, мозг не получит информацию о высоте тона.

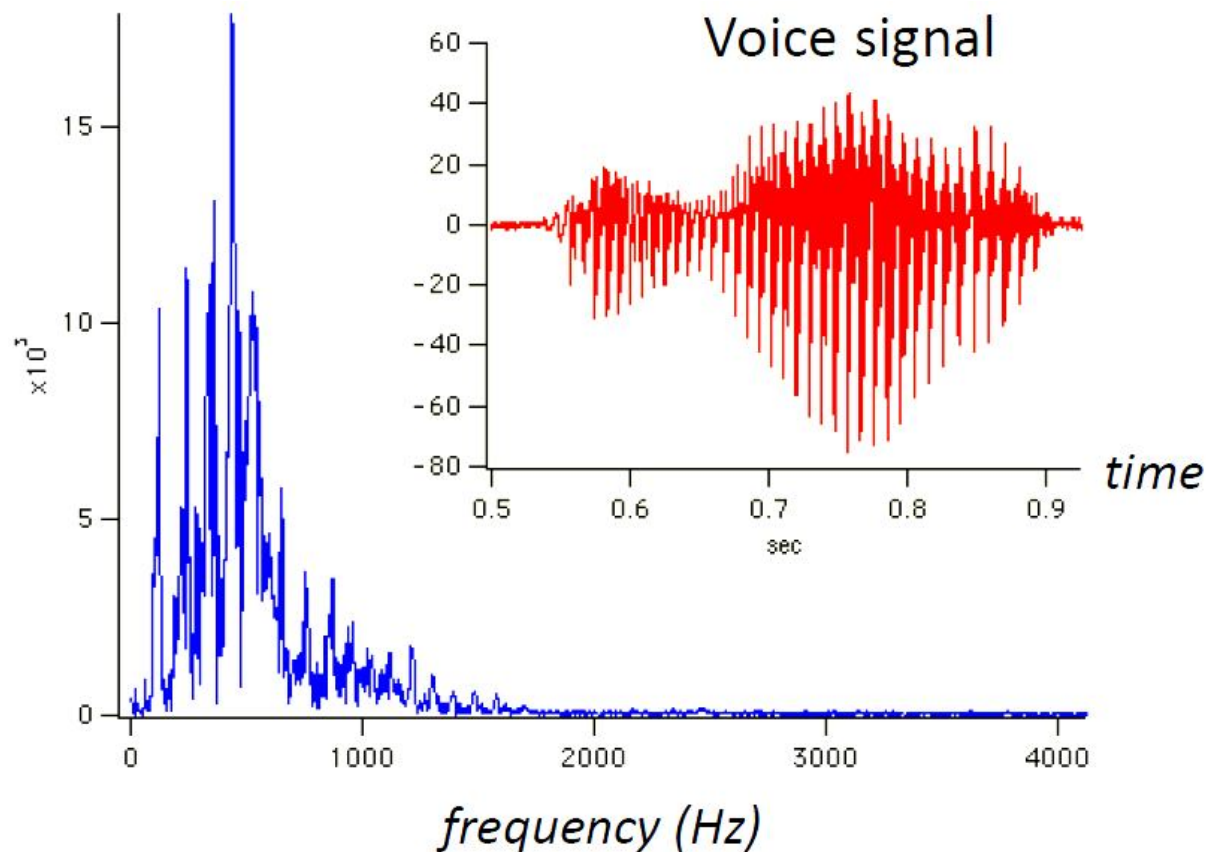
Поэтому очень кратковременное звучание (менее 20...30 периодов основного тона) любой частоты покажется на слух просто резким щелчком.

Это явление наз.
инерционностью слуха.



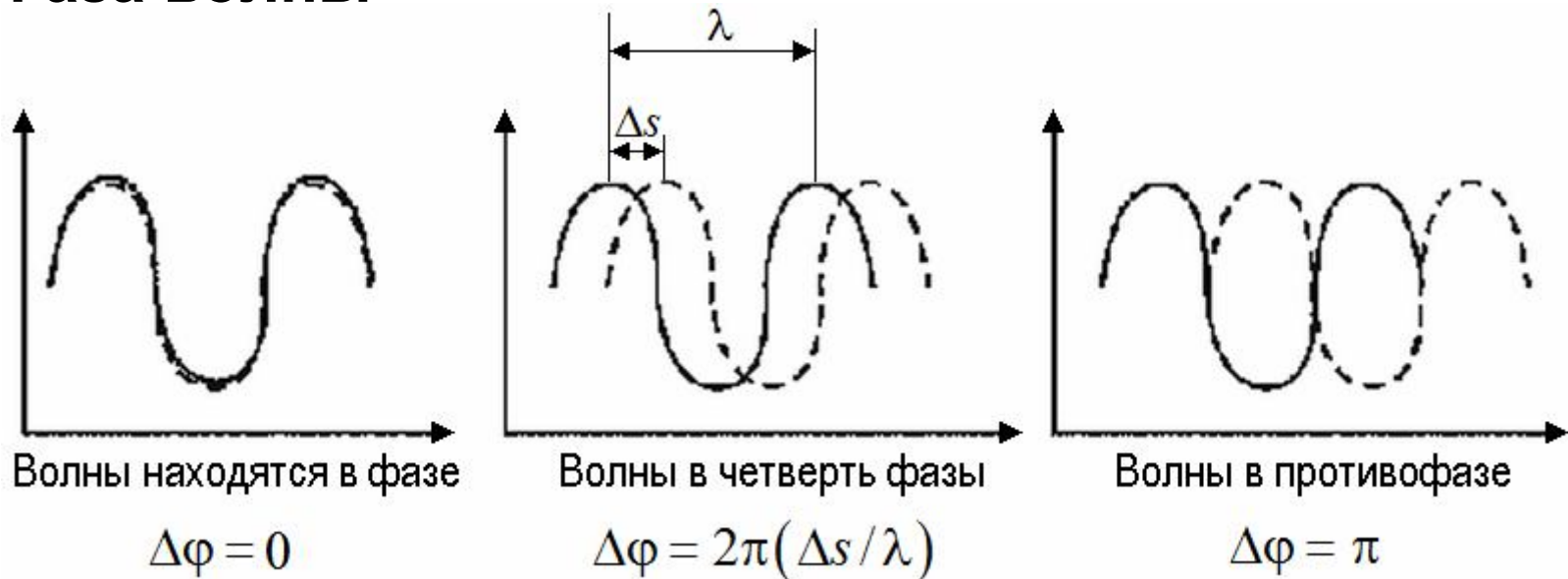
Характеристики звука

Типичные сигналограмма и спектрограмма голоса



Характеристики звука

Фаза волны

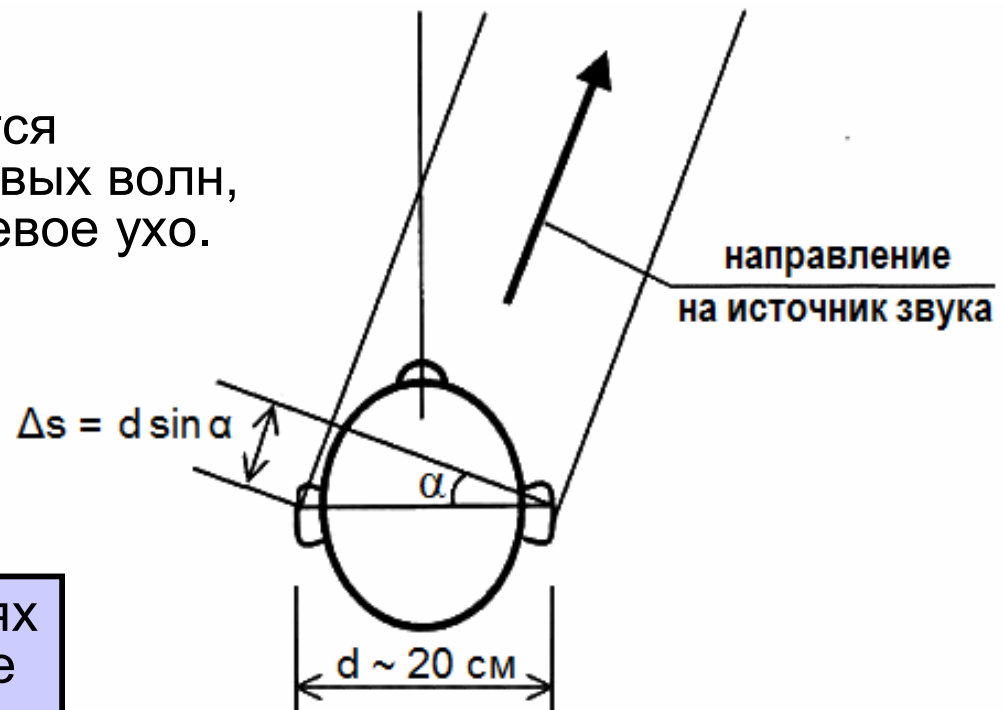


Хотя ухо не в состоянии сопоставить фазы **разных гармоник**, ухо и мозг, однако, в состоянии извлечь информацию о разности фаз сигналов **одной частоты** от правого и левого уха.

Характеристики звука

Частотам $f \sim 300 \dots 1000$ Гц
соответствуют длины волн
 $\lambda = v_{\text{зв}}/f \sim 113 \dots 34$ см.

На этих частотах становится
заметным сдвиг фаз звуковых волн,
попадающих в правое и левое ухо.



При нормальных условиях
скорость звука в воздухе
 $v_{\text{зв}} = 340$ м/с

Характеристики звука

Бинауральный (двуушный) эффект –

способность уха определять местоположение источника звука.

- Практически отсутствует на частотах ниже 300 Гц,
- На частотах от 300 до 1000 Гц становится заметным сдвиг фаз звуковых волн, попадающих в правое и левое ухо.
- При частотах более 1000 Гц сдвиг фаз становится очень небольшим и поиск направления осуществляется за счет сравнения силы звука, приходящего с разных сторон.

При неподвижном источнике звука слух способен определить направление движения звуковых волн по горизонтали не точнее 12° , а по вертикали – $17...20^\circ$

Характеристики звука

- <https://digitalmusicacademy.ru/lesson-sound-as-a-physical-phenomenon>
- <https://function-x.ru/rows5.html>
- https://ru.wikipedia.org/wiki/Ряд_Фурье
- https://ru.qaz.wiki/wiki/Basilar_membrane
- <http://www.midi.ru/doc/38.htm>

Цифровая звукозапись

Цифровая звукозапись – технология преобразования аналогового звука в цифровой для возможности его хранения, пересылки и воспроизведения.

- Вначале аналоговый сигнал преобразуется в цифровой при помощи **аналого-цифрового преобразователя, АЦП (Analog-to-digital converter, ADC)**.
- Полученные цифровые данные сохраняются на носителе.
- Для воспроизведения сделанной записи требуется обратное преобразование в аналоговый сигнал при помощи **цифро-аналогового преобразователя, ЦАП (Digital-to-analog converter, DAC)**.



Цифровая звукозапись

Импульсно-кодовая модуляция, ИКМ (Pulse code modulation, PCM)

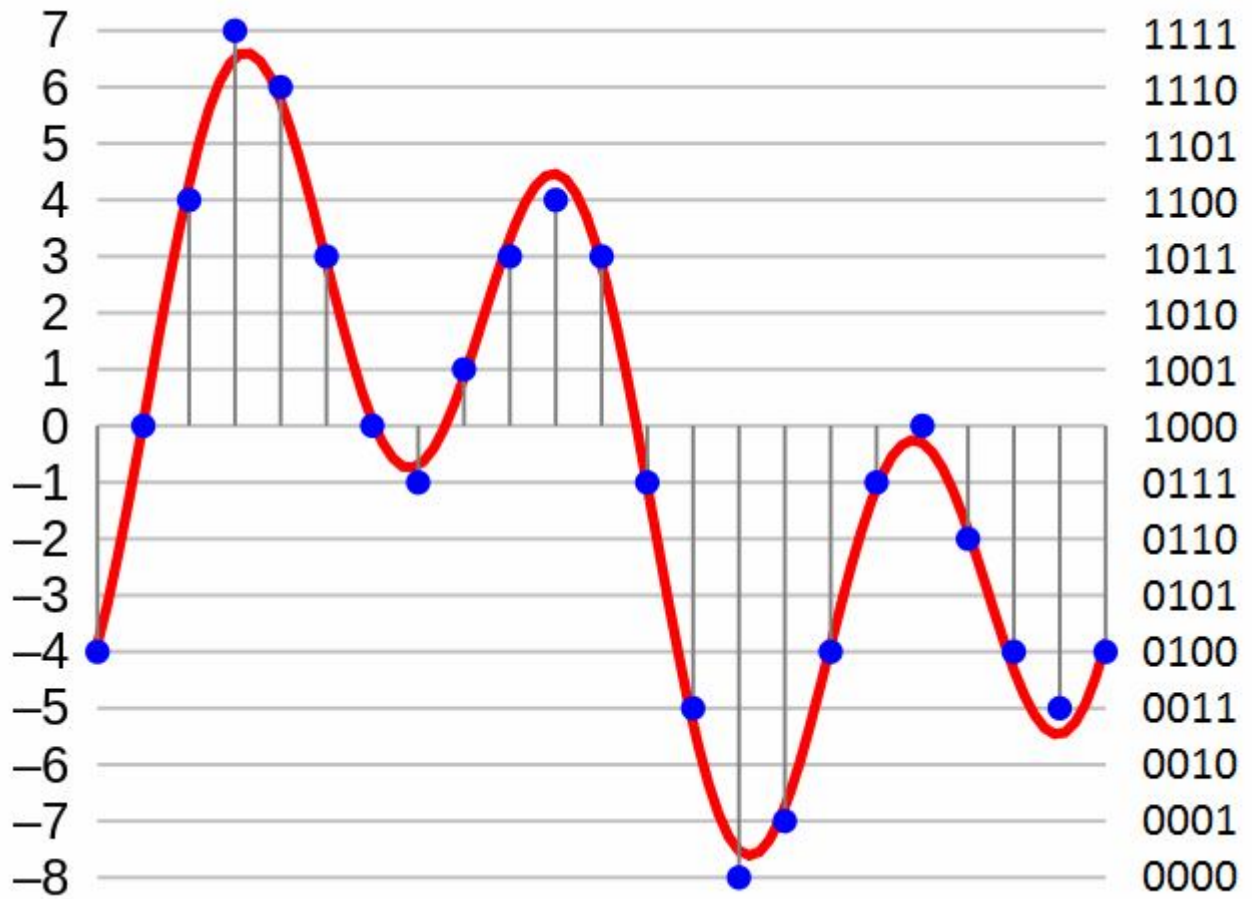
Кодек PCM преобразует сигнал в цифровой в три этапа:

- **дискретизация** (сэмплирование, *sampling*) – представление непрерывного сигнала в виде последовательного набора отдельных амплитуд («сэмплов»);
- **квантование** – разделение всего диапазона сигнала на заданное число уровней и округление каждой из его последовательных амплитуд до ближайшего уровня;
- **кодирование** – запись данных позиции и уровня амплитуды в цифровом виде.

Цифровая звукозапись

Звуковая волна
(красная) –
аналоговый
сигнал.

Дискретизация,
4-битное
квантование
(синие)
и **кодирование**
4-битным
словом –
цифровая запись



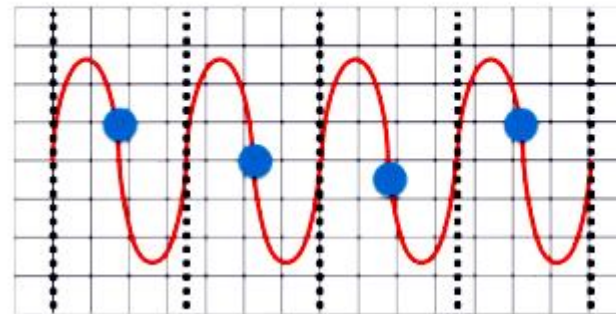
Цифровая звукозапись

Теорема Котельникова (Найквиста-Шеннона)

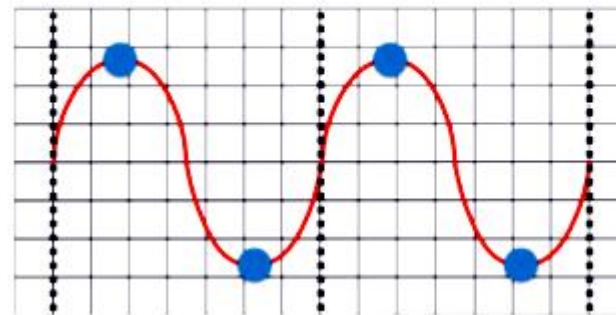
Любую функцию, состоящую из частот от 0 до f_{max} , можно непрерывно передавать с любой точностью при помощи чисел, следующих друг за другом с частотой $f_{дискр} = 2f_{max}$.

Т.е. частота дискретизации $f_{дискр}$ должна как минимум вдвое превышать частоту любой гармоники из спектра самой функции («два отсчета на период»).

ONE Digital Sample per Cycle



TWO+ Digital Samples per Cycle

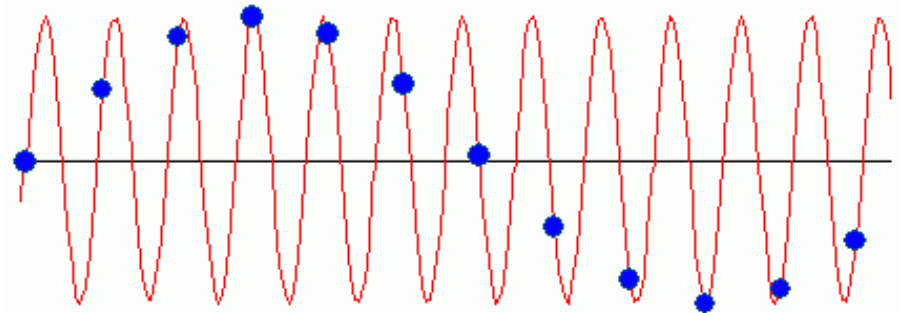


Цифровая звукозапись

Алиасинг

Паразитный эффект.

Если требование теоремы Котельникова не выполняется, и в исходном аналоговом сигнале присутствует помеха с частотой $f_{\text{помех}} > f_{\text{дискр}}/2$, то в сигнале после дискретизации появится паразитная низкочастотная гармоника.



Для предотвращения алиасинга аналоговый сигнал перед дискретизацией нужно обязательно **отфильтровать от высоких частот**, пропустив его через **фильтр низких частот (ФНЧ)** и удалив все частотные составляющие выше $f_{\text{дискр}}/2$.

Цифровая звукозапись

Частота дискретизации 44 100 Гц

Поскольку слышимый диапазон для человеческого до 20 кГц, то частота дискретизации должна быть выше 40 кГц.

Первое оборудование для цифровой звукозаписи создано на базе имеющегося оборудования для аналоговой видеозаписи, а носителем служила видеокассета с пленкой.

По ТВ-стандарту PAL кадры по 588 строк чередуются по 25 кадров в секунду, т.е. частота строчной развертки равна $588 \times 25 = 14700$ Гц.

Утроенная частота $3 \times 14700 = 44100$ Гц как раз подходит для цифровой звукозаписи. Это значит, что за время развертки аналоговой ТВ-строки записывается 3 звуковых отсчета.

Цифровая звукозапись

Разрядность АЦП 16 бит

Точное восстановление непрерывного сигнала по дискретным отсчетам требует точного значения каждой амплитуды сигнала, т.е. бесконечно большого **числа уровней квантования**.

На практике число уровней квантования устанавливают в 2^N , где N – разрядность АЦП.

Общее правило – «чем больше – тем лучше», но 16 бит (65536 уровней) практически достаточно.

Цифровая звукозапись

Компакт-диск (Compact Disc, CD)

Для записи **2-канального стереозвука на CD** используется $f_{\text{дискр}} = 44100$ Гц и 16-разрядный АЦП.

Формируемый цифровой поток («**битрейт**») определяется произведением $N \cdot f_{\text{дискр}}$ бит/с и для 2-канальной записи составит:
 $44100 \times 16 \times 2 = 1\,411\,200$ бит/с = 0.168 МБ/с = **10.1 МБ/мин.**

Первые студийные адаптеры, которые записывали звук в таком стандарте на видеокассеты – Sony PCM-1600 (1979).

Для тиражирования звукозаписей был разработан CD (1982). По легенде, его диаметр (120 мм) и плотность гравирования дорожек выбраны так, чтобы на CD помещалась целиком запись 9-й симфонии Бетховена (**74 мин или 750 МБ**).

Звуковой файл такого формата в Windows имеет расширение **.WAV**

Цифровая звукозапись

- https://ru.wikipedia.org/wiki/Цифровая_звукозапись
- https://ru.wikipedia.org/wiki/Цифровые_аудиоформаты
- https://ru.wikipedia.org/wiki/Кодирование_звуковой_информации
- [https://ru.wikipedia.org/wiki/Квантование_\(обработка_сигналов\)](https://ru.wikipedia.org/wiki/Квантование_(обработка_сигналов))
- <https://ru.wikipedia.org/wiki/%Алиасинг>
- <https://ru.wikipedia.org/wiki/Компакт-диск>

Сжатие аудиоданных

Сжатие (компрессия) аудиоданных –

уменьшение скорости цифрового потока данных за счет сокращения избыточности цифрового звукового сигнала.

Кодек PCM обеспечивает практически полную идентичность оригинальному аудио, но не очень экономичен, что выражается в очень больших объемах файлов.

Также он не подходит для потокового вещания, поскольку битрейт 1411.2 кбит/с для многих передающих каналов слишком велик.

Сжатие аудиоданных

Эффект маскирования –

неспособность человека в определенных случаях различать тихие звуки в присутствии более громких.

Эффекты слухового маскирования разделяются на две группы:

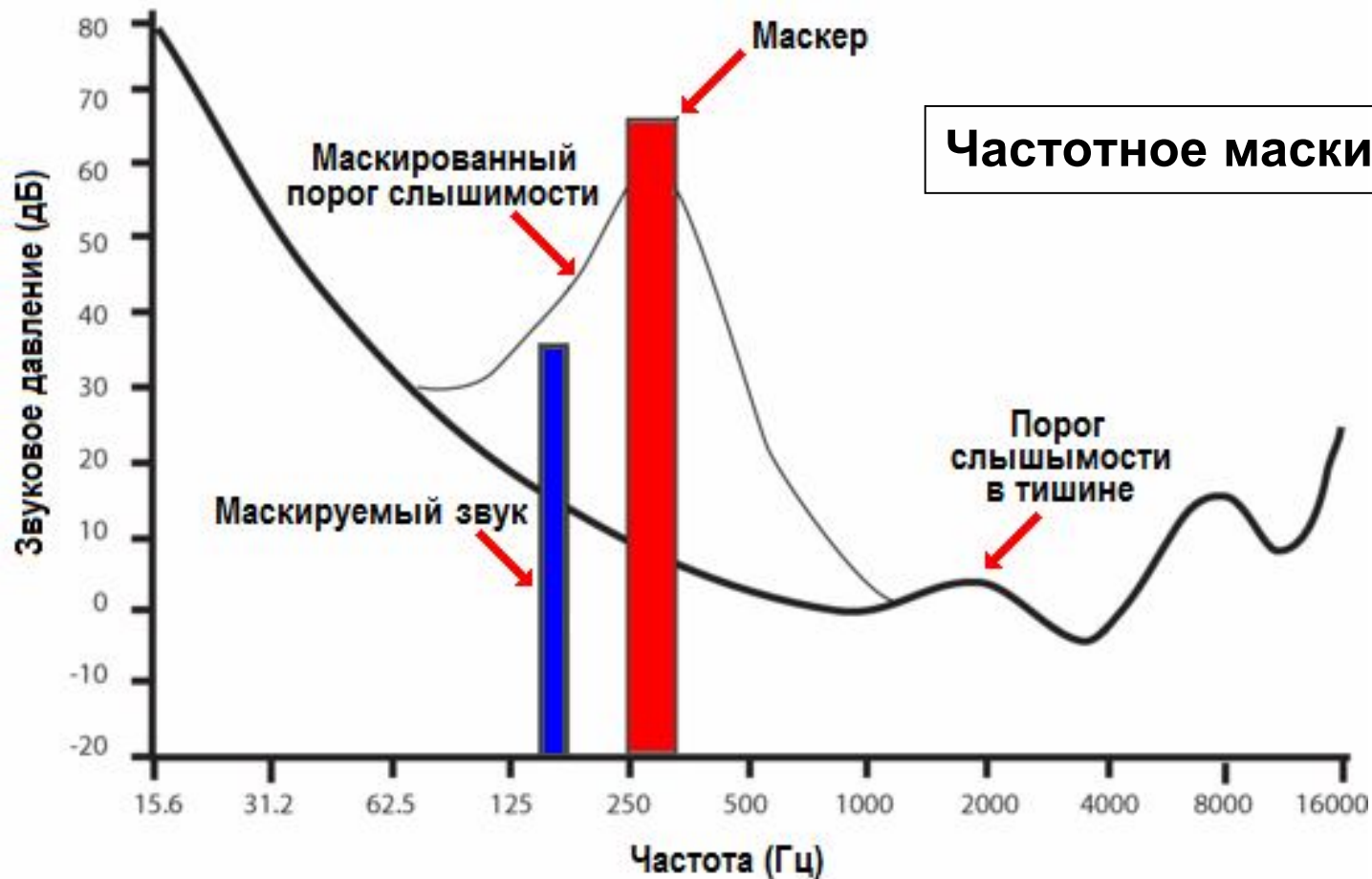
- **частотное** (одновременное) маскирование;
- **временное** (неодновременное) маскирование.

Психоакустическое («perceptual») сжатие звука

(от «perception» - «восприятие», «ощущение» чего-либо») основывается на несовершенстве человеческого слуха при восприятии звука.

Неслышимые аудиосигналы несущественны для человеческого восприятия и поэтому могут быть удалены.

Сжатие аудиоданных



Сжатие аудиоданных

Маскирование в частотной области

связано с тем, что в присутствии больших звуковых амплитуд ухо нечувствительно к малым амплитудам близких частот.

Т. е., когда два сигнала одновременно находятся в ограниченной частотной области, то более слабый сигнал становится неслышимым на фоне более сильного.

Частотные спектры нужно определять каждые 15-20 мс.

Маскирование во временной области

изменяет порог слышимости одного сигнала в присутствии другого, когда маскирующий и маскируемый сигналы звучат не одновременно.

Тихий сигнал в присутствии громкого маскира становится неслышимым за 5...50 мс до включения маскира и становится вновь слышимым через 50...200 мс после его выключения.

Сжатие аудиоданных

Аудио MPEG1 level III

Наиболее распространенный в интернете формат сжатых аудиоданных, где сжатие обеспечивается, в основном, за счет устранения психофизической избыточности.

При кодировании задают требуемую величину битрейта, а алгоритмы сжатия добиваются требуемого значения уровня сжатия за счет потери качества.

Звуковой файл такого формата имеет расширение **.MP3**

При подаче на вход кодека MP3 входного цифрового сигнала стерео 16-бит РСМ, 44.1 кГц с битрейтом 1411.2 кбит/с на выходе можно получить битрейт 64...384 кбит/с.

128 кбит/с (уплотнение в 11 раз) в большинстве случаев не приводит к существенным потерям в качестве звука.

Сжатие аудиоданных

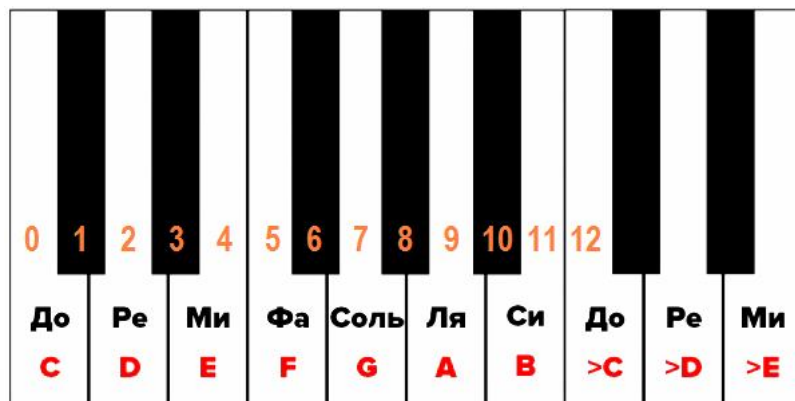
- https://ru.wikipedia.org/wiki/Сжатие_аудиоданных
- https://scinse.donntu.edu.ua/ks/kolesnik_a/library/files/4.pdf
- http://astro.tsu.ru/MT/text/2_3.html
- <https://studfile.net/preview/2690112/page:50/>
- <https://digitalmusicacademy.ru/lesson-digital-audio-compressing>
- http://mf.grsu.by/UchProc/livak/po/comprsite/theory_souvid.html

Музыкальный строй

Ноты – другой способ распространения музыкальных записей



Музыкальный строй



Октава =
12 полутонов



Почему ступеней именно 12?

Музыкальный строй

Музыкальный строй – соответствие ступеней музыкального звукоряда звукам определённой высоты.

Интервал – соотношение двух музыкальных звуков по их высоте. Определяется **отношениями частот** звуков.

Октава – музыкальный интервал, в котором соотношение частот между звуками составляет 2:1. Субъективно на слух октава воспринимается как базисный музыкальный интервал. Два звука, отстоящие на октаву, воспринимаются очень похожими друг на друга, хотя явно различаются по высоте.

Восходящая октава



Нисходящая октава



Музыкальный строй

Для получения звукоряда октаву делят на N равных интервалов, тогда частоты соседних ступеней будут различаться в

$\sqrt[N]{2}$ раз (**равномерно темперированный строй**).

Важное преимущество равномерной темперации – возможность **транспонирования** пьесы на произвольный интервал, т.е. ее **исполнения с систематическим сдвигом высот всех звуков** на заданный интервал.

Требование равенства соседних интервалов не столь очевидно, и данный подход был применен лишь с 18 века (И.С.Бах. Хорошо темперированный клавир. 1722).

Раньше применялись и другие звукоряды.
Свидетельство тому – белые и черные клавиши фортепиано.

Музыкальный строй

Созвучными (консонанс) считаются одновременно звучащие тоны, частоты которых соотносятся как малые целые числа. Помимо октавы (2:1) это квинта (3:2) и кварта (4:3).

Учитывая, что

$$\left(\sqrt[12]{2}\right)^5 = 1.335 \approx 4:3$$

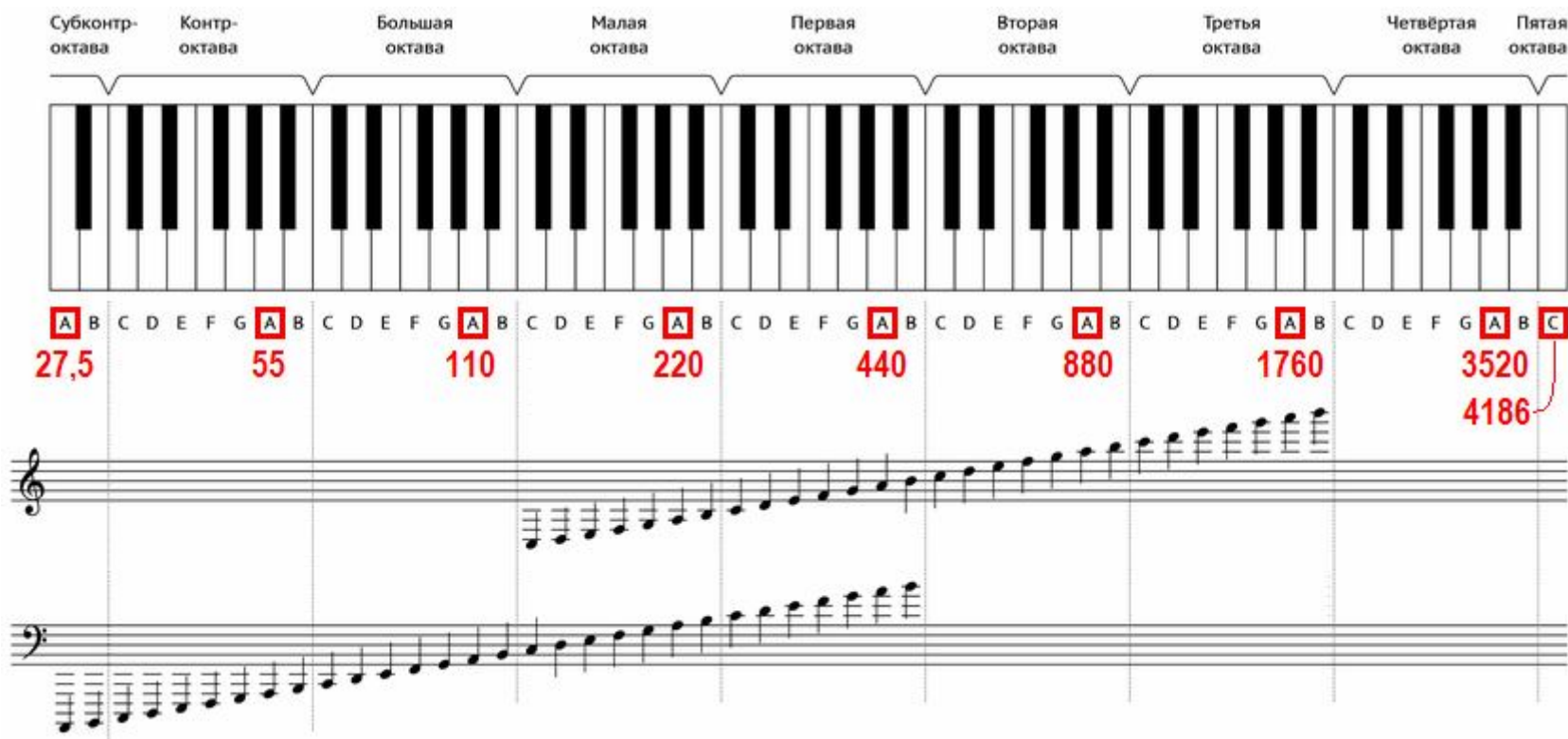
$$\left(\sqrt[12]{2}\right)^7 = 1.498 \approx 3:2$$

кварта и квинта хорошо вписываются в 12-ступеневой темперированный звукоряд.

Ля 1-й октавы = 440 Гц

Музыкальный строй

Стандартная клавиатура фортепиано (88 клавиш).
Частоты некоторых звуков (Гц) нанесены красным.



Музыкальный строй

- <https://habr.com/ru/post/471600/>
- https://ru.wikipedia.org/wiki/Равномерно_темперированный_строй
- [https://ru.wikipedia.org/wiki/Интервал_\(музыка\)](https://ru.wikipedia.org/wiki/Интервал_(музыка))
- https://ru.wikipedia.org/wiki/Консонанс_и_диссонанс

MIDI – Musical Instrument Digital Interface

MIDI – формат обмена данными между электронными музыкальными инструментами.

MIDI-файл (обычно с расширением **.MID**) содержит ссылки на ноты. Его данные – поток сообщений в реальном времени, каждое из которых является командой для музыкального инструмента.

Стандартизованы номера 128 программ («тембров» или «инструментов»), напр.:

- | | |
|------------------|----------------------|
| ● 1 - Фортепиано | ● 57 - Труба |
| ● 16 - Аккордеон | ● 74 - Флейта |
| ● 41 - Скрипка | ● 97 - Дождь |
| ● 44 - Контрабас | ● 110 - Волынка |
| ● 47 - Арфа | ● 127 - Аплодисменты |

MIDI – Musical Instrument Digital Interface

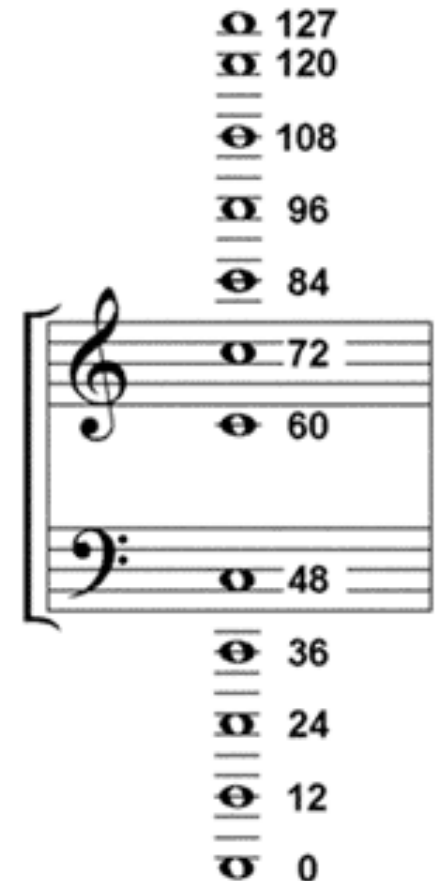
В MIDI используется 128 нот различной высоты (от 0 до 127).
Нота 69 – Ля первой октавы (частота 440 Гц).

Для перевода номера ноты в частоту используется формула:

$$f_N = 440 \cdot 2^{\frac{N-69}{12}},$$

где N – номер ноты, f_N – частота ноты в Гц.

Полный частотный диапазон превышает 10 октав (8...12543 Гц), что более чем достаточно.



MIDI – Musical Instrument Digital Interface

Стандарт предусматривает 16 независимых логических каналов, каждому каналу обычно назначается свой инструмент.

Воспроизводящее устройство или программа наз. **синтезатором (секвенсором) MIDI** и фактически является автоматическим музыкальным инструментом.

Первые звуковые карты с MIDI-синтезатором выпустила фирма Creative Technology (Сингапур) («**Sound Blaster**», 1989)

MIDI – Musical Instrument Digital Interface

Всякий синтезатор (в том числе и звуковая карта) содержит в своей памяти банк со звуками всех тембров.

Каждый тембр, как правило, состоит из **сэмплов** – аудиозаготовок, которые определяют звучание инструмента. Сэмпл может быть как синтезированным так и записанным при помощи реальных инструментов.

Когда MIDI-совместимое устройство получает сообщение, которое просит его воспроизвести какую-то ноту определённым тембром, оно воспроизводит сэмпл данного тембра, меняя его высоту в соответствии с высотой указанной ноты.

Реальные результаты воспроизведения зависят от занесенного в память синтезатора набора сэмплов.

MIDI – Musical Instrument Digital Interface

блок заголовка
длина блока - 6 байт
формат файла 0
один трек
96 тиков в четверти

Пример команд MIDI-файла (аккорд До-мажор)

На первом MIDI-канале – тембр 3 (электророяль),
на втором – 17 (электроорган), на третьем – 47 (арфа).

блок трека
длина блока - 59 байт
дельта-время 0, размер 4/4, один удар метронома на четверть,
8 тридцать вторых на 24 MIDI Clock
дельта-время 0, темп - 652174 микросекунды на четверть
канал 1, программа 3
канал 2, программа 17
канал 3, программа 47
взять ноту До малой октавы на канале 3
дельта-время 96 тиков, взять ноту Соль первой октавы на канале 2
взять ноту До первой октавы на канале 2 (Running Status)
дельта-время 96 тиков, взять ноту Ми второй октавы на канале 1
дельта-время 192 тика, снять ноту До малой октавы на канале 3
снять ноту Соль первой октавы на канале 2
снять ноту До первой октавы на канале 2 (Running Status)
снять ноту Ми второй октавы на канале 1
конец трека

$\bullet = 92$

Channel 1
Program 3

Channel 2
Program 17

Channel 3
Program 47

MIDI – Musical Instrument Digital Interface

Синтез
гитарного боя



MIDI – Musical Instrument Digital Interface

- <https://ru.wikipedia.org/wiki/MIDI>
- https://ru.wikipedia.org/wiki/General_MIDI
- <http://www.midi.ru/article.php>
- <http://www.midi.ru/doc/13.htm>
- <http://www.midi.ru/doc/35.htm>
- <http://el-izdanie.narod.ru/gl5/5-3.htm>
- <https://habr.com/ru/post/471600/>
- <http://www.audacity.ru/p9aa1.html>
- <http://www.muzoborudovanie.ru/articles/midi/midi5.php>