



Лекция 5. Основы записи, синтеза и воспроизведения звука.

1. Понятие звука
2. Кодирование звука
3. Кодирование оцифрованного звука перед его записью на носитель
4. Полный цикл преобразования звука
5. Методы кодирования
6. Программное обеспечение для работы со звуком



1. Понятие звука

Звук — физическое явление, представляющее собой распространение в виде **упругих волн** механических колебаний в твёрдой, жидкой или газообразной среде.



1. Понятие звука

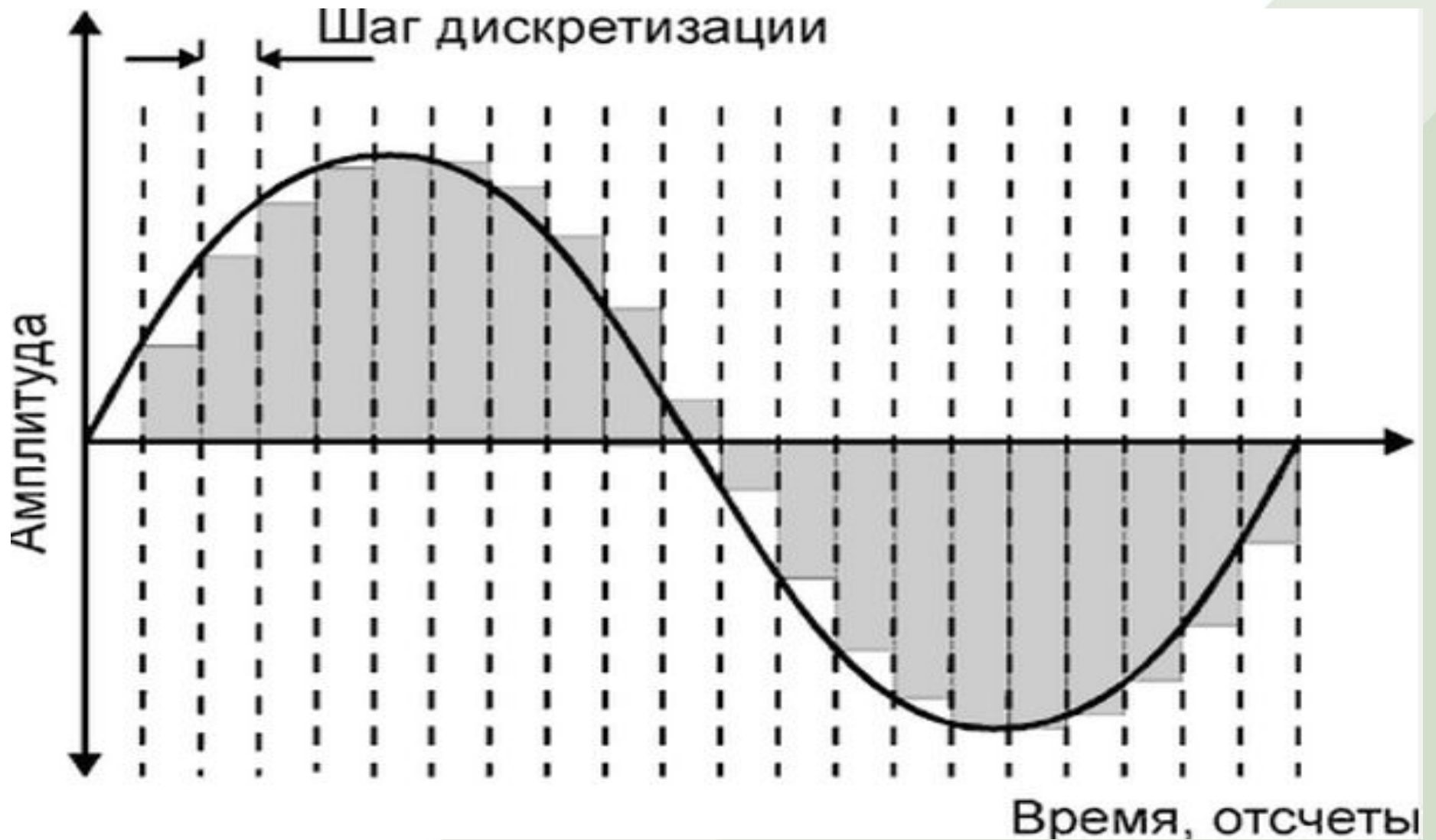
Как и любая волна, звук характеризуется амплитудой и спектром частот.

Обычный человек способен слышать звуковые колебания в диапазоне частот от 16-20 Гц до 15-20 кГц. Звук ниже диапазона слышимости человека называют инфразвуком; выше: до 1 ГГц - ультразвуком, от 1 ГГц - гиперзвуком.

Громкость звука сложным образом зависит от эффективного звукового давления, частоты и формы колебаний, а высота звука - не только от частоты, но и от величины звукового давления.



1. Понятие звука





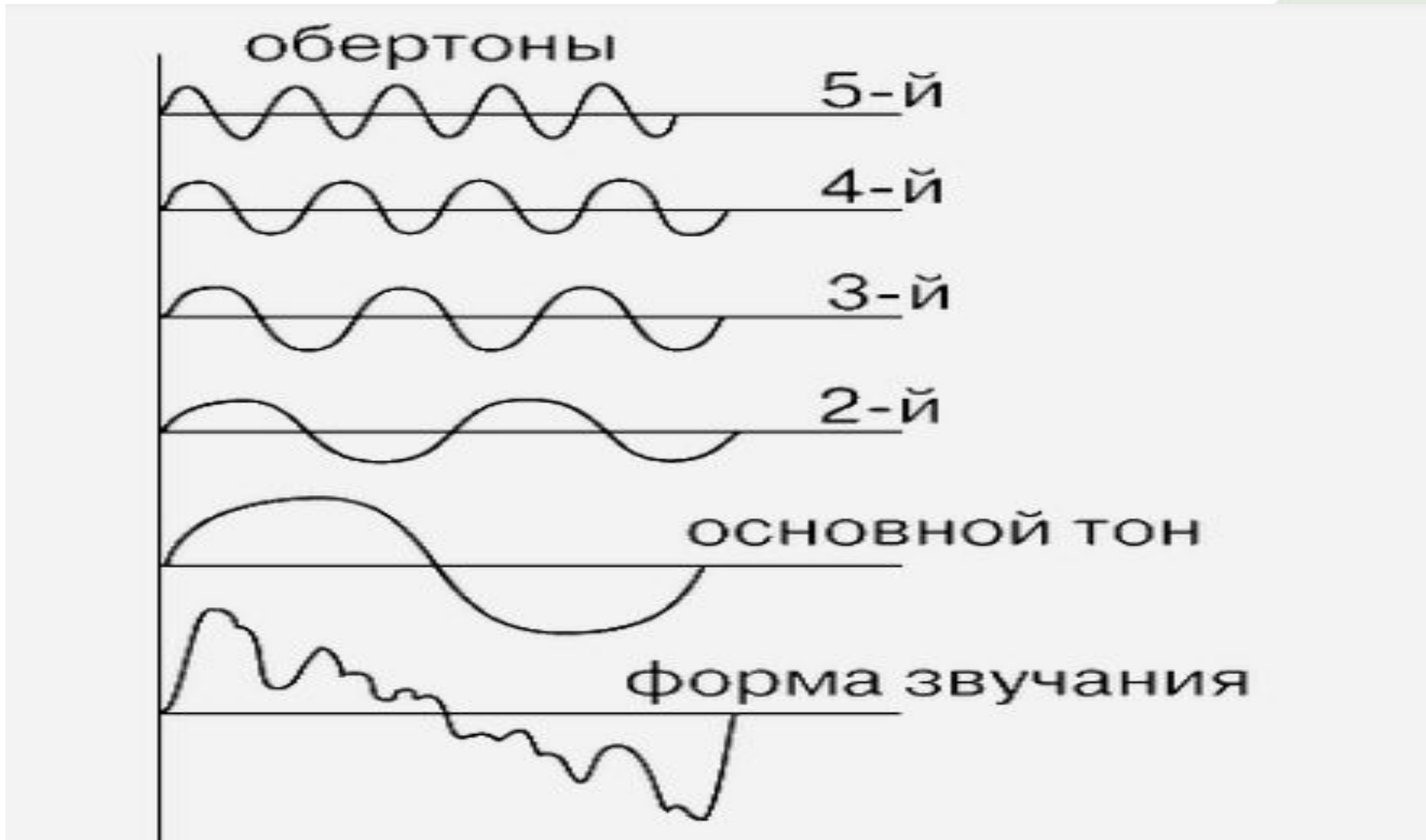
1. Понятие звука

В графическом представлении простейшую форму имеют звуковые колебания чистых тонов типа создаваемых камертоном. Им соответствуют синусоидальные кривые. Но большинство реальных звучаний имеет неправильную форму, которая однозначно характеризует звучание. Всякое звучание может быть разложено на чистые тона разных частот, которые состоят из:

- **основного тона**, с низшей частотой определяет высоту ноты;
- **обертонов (гармоник)**.



1. Понятие звука





1. Понятие звука

Полнота звучания обеспечивается при наличии всех обертонов, а для их воспроизведения необходимо, чтобы не искажались соотношения между уровнями основного тона и обертонов, т.е. частотная характеристика воспроизводящей системы должна быть линейной во всем диапазоне слышимых частот.



2. Кодирование звука

В основе **кодирования звука** с использованием ПК лежит процесс преобразования колебаний воздуха в колебания **электрического тока** и последующая **дискретизация аналогового электрического сигнала**.

Кодирование и воспроизведение звуковой информации осуществляется с помощью специальных программ (**редактор звукозаписи**). Качество воспроизведения закодированного звука зависит от **частоты дискретизации** и её разрешения (глубины кодирования звука - количество уровней)



2. Кодирование звука

Цифровой звук — это аналоговый звуковой сигнал, представленный посредством дискретных численных значений его амплитуды.

Оцифровка звука — технология поделенным временным шагом и последующей записи полученных значений в численном виде.

Другое название оцифровки звука — **аналогово-цифровое преобразование** звука.

Оцифровка звука включает в себя два процесса:

- процесс дискретизации (осуществление выборки) сигнала по времени
- процесс квантования по амплитуде.



2. Кодирование звука

Процесс дискретизации по времени — процесс получения значений сигнала, который преобразуется, с определенным временным шагом — *шагом дискретизации*. Количество замеров величины сигнала, осуществляемых в одну секунду, называют *частотой дискретизации* или *частотой выборки*, или *частотой семплирования* (от англ. «sampling» — «выборка»). Чем меньше шаг дискретизации, тем выше частота дискретизации и тем более точное представление о сигнале будет получено.



2. Кодирование звука

Теорема Котельникова (в зарубежной литературе встречается как теорема Шеннона).

Аналоговый сигнал с ограниченным спектром точно описуем дискретной последовательностью значений его амплитуды, если эти значения берутся с частотой, как минимум вдвое превышающей наивысшую частоту спектра сигнала.

То есть, аналоговый сигнал, в котором находится частота спектра равна F_m , может быть точно представлен последовательностью дискретных значений амплитуды, если для частоты дискретизации

F_d выполняется:

$$F_d > 2F_m.$$



2. Кодирование звука

На практике это означает, что для того, чтобы оцифрованный сигнал содержал информацию о всем диапазоне слышимых частот исходного аналогового сигнала (0 - 20 кГц) необходимо, чтобы выбранное значение частоты дискретизации составляло не менее 40 кГц.

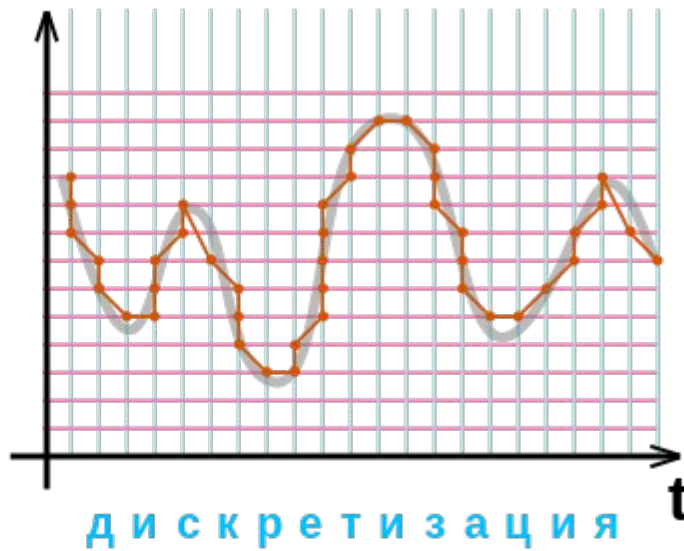
Количество замеров амплитуды в секунду называют *частотой дискретизации* (в случае, если шаг дискретизации постоянен). Основная трудность оцифровки заключается в невозможности записать измеренные значения сигнала с идеальной точностью.



2. Кодирование звука

Кодирование сигнала в цифровой вид
(высокое качество)

К В А Н Т О В А Н И Е



Кодирование сигнала в цифровой вид
(низкое качество)

К В А Н Т О В А Н И Е





2. Кодирование звука

Линейное (однородное) квантование амплитуды

Отведём для записи одного значения амплитуды сигнала в памяти компьютера N бит. Значит, с помощью одного N -битного слова можно описать 2^N разных положений. Пусть амплитуда оцифровываемого сигнала колеблется в пределах от -1 до 1 некоторых условных единиц. Представим этот диапазон изменения амплитуды — динамический диапазон сигнала — в виде $2^N - 1$ равных промежутков, разделив его на 2^N уровней — квантов. Теперь, для записи каждого отдельного значения амплитуды, его необходимо округлить до ближайшего уровня квантования. Этот процесс носит название квантования по амплитуде.



2. Кодирование звука

Линейное (однородное) квантование амплитуды

Квантование по амплитуде - процесс замены реальных значений амплитуды сигнала значениями, приближенными с некоторой точностью. Каждый из 2^N возможных уровней называется **уровнем квантования**, а расстояние между двумя ближайшими уровнями квантования называется **шагом квантования**. Если амплитудная шкала разбита на уровни линейно, квантование называют линейным (однородным).

Точность округления зависит от выбранного количества (2^N) уровней квантования, которое, в свою очередь, зависит от количества бит (N), отведенных для записи значения амплитуды. Число N называют **разрядностью квантования** (подразумеваемая количество разрядов, то есть бит, в каждом слове), а полученные в результате округления значений амплитуды числа - *отсчетами или семплами* (от англ. « sample» - «замер»).



2. Кодирование звука

Линейное (однородное) квантование амплитуды

Принимается, что погрешности квантования, являющиеся результатом квантования с разрядностью 16 бит, остаются для слушателя почти незаметными. Этот способ оцифровки сигнала — дискретизация сигнала во времени в совокупности с методом однородного квантования — называется **импульсно-кодовой модуляцией, ИКМ** (англ. Pulse Code Modulation — PCM).

Оцифрованный сигнал в виде набора последовательных значений амплитуды уже можно сохранить в памяти компьютера. В случае, когда записываются абсолютные значения амплитуды, такой *формат записи* называется **PCM** (Pulse Code Modulation). Стандартный аудио компакт-диск (**CD-DA**), применяющийся с начала 80-х годов 20-го столетия, хранит информацию в формате PCM с частотой дискретизации 44.1 кГц и разрядностью квантования 16 бит.



2. Кодирование звука

Другие способы оцифровки

Способ неоднородного квантования предусматривает разбиение амплитудной шкалы на уровни по логарифмическому закону. Такой способ квантования называют **логарифмическим квантованием**. При использовании логарифмической амплитудной шкалы, в области слабой амплитуды оказывается большее число уровней квантования, чем в области сильной амплитуды (при этом, общее число уровней квантования остается таким же, как и в случае однородного квантования). Аналогово-цифровое преобразование, основанное на применении метода неоднородного квантования, называется **неоднородной импульсно-кодовой модуляцией** — неоднородной ИКМ (Nonuniform PCM).



2. Кодирование звука

Другие способы оцифровки

Альтернативным способом аналогово-цифрового преобразования является **разностная импульсно-кодовая модуляция** — разностная ИКМ (англ. « Differential PCM» — DPCM). В случае разностной ИКМ квантованию подвергают не саму амплитуду, а относительные значения величины амплитуды. В полной аналогии с ИКМ, разностная ИКМ может сочетаться с использованием как однородного, так и неоднородного методов квантования.



3. Кодирование оцифрованного звука перед его записью на носитель

Для хранения цифрового звука существует много различных способов. Оцифрованный звук является набором значений амплитуды сигнала, взятых через определенные промежутки времени.

Блок оцифрованной аудио информации можно записать в файл без изменений, то есть последовательностью чисел - значений амплитуды. В этом случае существуют два способа хранения информации.

- Первый - **PCM** (Pulse Code Modulation - импульсно-кодовая модуляция) - способ цифрового кодирования сигнала при помощи записи абсолютных значений амплитуд (в таком виде записаны данные на всех аудио CD).
- Второй - **ADPCM** (Adaptive Delta PCM - адаптивная относительная импульсно-кодовая модуляция) – запись значений сигнала не в абсолютных, а в относительных изменениях амплитуд (приращениях).



3. Кодирование оцифрованного звука перед его записью на носитель

Можно сжать данные так, чтобы они занимали меньший объем памяти, чем в исходном состоянии. Существует два способа:

- 1. Кодирование данных без потерь (lossless coding).**
- 2. Кодирование данных с потерями (lossy coding).**



3. Кодирование оцифрованного звука перед его записью на носитель

Кодирование данных без потерь (lossless coding) - способ кодирования аудио, который позволяет осуществлять стопроцентное восстановление данных из сжатого потока. К нему прибегают в тех случаях, когда сохранение оригинального качества данных особо значимо. Существующие сегодня алгоритмы кодирования без потерь (например, Monkeys Audio) позволяют сократить занимаемый данными объем на 20-50%, но при этом обеспечить стопроцентное восстановление оригинальных данных из полученных после сжатия.



3. Кодирование оцифрованного звука перед его записью на носитель

Кодирование данных с потерями (lossy coding).

Цель метода - добиться схожести звучания восстановленного сигнала с оригиналом при как можно меньшем размере сжатого файла. Это достигается путем использования алгоритмов, «упрощающих» оригинальный сигнал (удаляющих из него «несущественные», неразличимые на слух детали). Это приводит к тому, что декодированный сигнал перестает быть идентичным оригиналу, а является лишь «похоже звучащим».



3. Кодирование оцифрованного звука перед его записью на носитель

Методов сжатия, а также программ, реализующих эти методы, существует много. В среднем, коэффициент сжатия, обеспечиваемый такими кодерами, находится в пределах 10-14 (раз). В основе всех lossy-кодеров лежит использование так называемой психоакустической модели.

Она занимается этим самым «упрощением» оригинального сигнала. Такое сжатие приводит к сильной потере качества, поскольку удалению могут подлежать не только незаметные, но и значимые детали звучания



3. Кодирование оцифрованного звука перед его записью на носитель

Кодер – программа (или устройство), реализующая определенный алгоритм кодирования данных (например, архиватор, или кодер MP 3), которая в качестве ввода принимает исходную информацию, а в качестве вывода возвращает закодированную информацию в определенном формате.

Декодер – программа (или устройство), реализующая обратное преобразование закодированного сигнала в декодированный.

Кодек (от англ. « codec » - « Coder / Decoder ») - программный или аппаратный блок, предназначенный для кодирования/декодирования данных.



3. Кодирование оцифрованного звука перед его записью на носитель

Наиболее распространённые кодеки

- MP3 – MPEG-1 Layer 3
- OGG – Ogg Vorbis
- WMA – Windows Media Audio
- MPC - MusePack
- AAC – MPEG-2/4 AAC (Advanced Audio Coding)
 - Стандарт MPEG-2 AAC
 - Стандарт MPEG-4 AAC



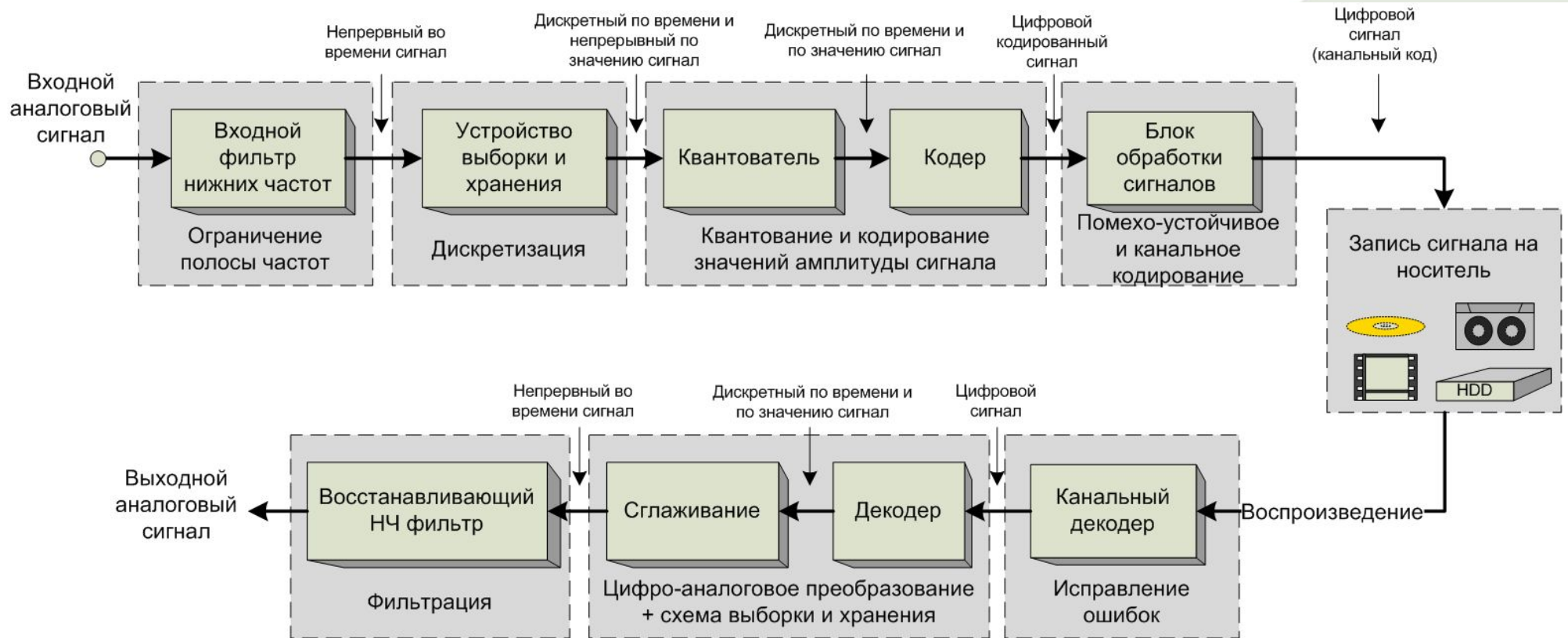
3. Кодирование оцифрованного звука перед его записью на носитель

Сравнение звуковых форматов

Название формата	Квантование, бит	Частота дискретизации, кГц	Число каналов	Величина потока данных с диска, кбит/с	Степень сжатия/упаковки
CD	16	44,1	2	1411,2	1:1 без потерь
Dolby Digital (AC3)	16-24	48	6	до 640	~12:1 с потерями
DTS	20-24	48; 96	до 8	до 1536	~3:1 с потерями
DVD-Audio	16; 20; 24	44,1; 48; 88,2; 96	6	6912	2:1 без потерь
DVD-Audio	16; 20; 24	176,4; 192	2	4608	2:1 без потерь
MP3	плавающий	до 48	2	до 320	~11:1 с потерями
AAC	плавающий	до 96	до 48	до 529	с потерями
AAC+ (SBR)	плавающий	до 48	2	до 320	с потерями
Ogg Vorbis	до 32	до 192	до 255	до 1000	с потерями
WMA	до 24	до 96	до 8	до 768	2:1, есть версия без потерь



4. Полный цикл преобразования звука





4. Полный цикл преобразования звука

Помехоустойчивое и канальное кодирование

Помехоустойчивое кодирование позволяет при воспроизведении сигнала выявить и устранить (или снизить частоту их появления) ошибки чтения с носителя. Для этого при записи к сигналу полученному на выходе АЦП добавляется искусственная избыточность (контрольный бит), которая впоследствии помогает восстановить поврежденный отсчет. В устройствах записи звука обычно используется комбинация из двух или трех помехоустойчивых кодов. Для лучшей защиты от пакетных ошибок также применяется перемежение.



4. Полный цикл преобразования звука

Канальное кодирование служит для согласования цифровых сигналов с параметрами канала передачи (записи/воспроизведения). К полезному сигналу добавляются вспомогательные данные, которые облегчают последующее декодирование. Это могут быть сигналы **временного кода**, служебные сигналы, сигналы синхронизации. В устройствах воспроизведения цифровых сигналов канальный декодер выделяет из общего потока данных тактовые сигналы и преобразует поступивший канальный сигнал в цифровой поток данных. После коррекции ошибок сигнал поступает в ЦАП.



4. Полный цикл преобразования звука

Принцип действия ЦАП

Цифровой сигнал, полученный с декодера, преобразовывается в аналоговый. Это преобразование происходит следующим образом:

1. Декодер ЦАП преобразует последовательность чисел в дискретный квантованный сигнал
2. Путем сглаживания во временной области из дискретных отсчетов вырабатывается непрерывный во времени сигнал
3. Окончательное восстановление сигнала производится путем подавления побочных спектров в аналоговом фильтре нижних частот



4. Полный цикл преобразования звука

Параметры, влияющие на качество звука при его прохождении по полному циклу

Основными параметрами, влияющими на качество звука при этом являются:

- Разрядность АЦП и ЦАП.
- Частота дискретизации АЦП и ЦАП.
- Джиттер АЦП и ЦАП
- Передискретизация



4. Полный цикл преобразования звука

Джиттер (англ. *jitter* — дрожание) или **фазовое дрожание цифрового сигнала данных**

Передискретизация (англ. *resampling*) в обработке сигналов — изменение частоты дискретизации дискретного (чаще всего цифрового) сигнала



5. Методы кодирования

Частотная модуляция

Методы кодирования звука (электрический сигнал, поступающий с микрофона) основаны на том, что теоретически любой **сложный звук** можно разложить на **последовательность простейших гармонических сигналов разных частот**, каждый из которых представляет собой синусоиду, называемых спектром исходного сигнала.

Задачей кодирования звука является представление его в форме аналогового или цифрового сигнала, более удобного для передачи или хранения.

Реальные источники звука имеют ограниченную ширину спектра, поэтому для кодирования применяют такие методы преобразования, которые преобразуют исходный сигнал в такой, спектр которого наиболее подходит для передачи по выбранному каналу.



5. Методы кодирования

Частотная модуляция

Представление аналогового сигнала в виде другого аналогового сигнала обычно называется модуляцией, а представление в цифровом виде - кодировкой. Это разделение очень условно. Аналоговый сигнал может быть представлен в виде гармонического сигнала (т.е. синусоиды), параметры которого изменяются в зависимости от значения первоначального сигнала. В том случае, когда с изменением первоначального сигнала изменяется амплитуда синусоиды - мы имеем дело с амплитудной модуляцией (АМ).



5. Методы кодирования

Частотная модуляция

Если в зависимости от значения исходного сигнала изменяются частота или фаза синусоиды - мы имеем дело с частотной модуляцией (FM) или фазовой модуляцией (PM).

Амплитудная и частотная модуляция, например, широко используются для передачи звука **по радио**. Эти виды модуляции, конечно, не являются разложением исходного сигнала по гармоникам. Развитие цифровой техники и применение компьютерной обработки и хранения информации привело к широкому применению **импульсных методов модуляции** или кодирования (импульсно-кодовая модуляция, при которой значение исходного сигнала через определенные промежутки времени представляется в виде кода).



5. Методы кодирования

Частотная модуляция

Подавляющее большинство "компьютерного звука" является именно записью двоичного кода сигнала, полученного через небольшие равные промежутки времени, определяемые частотой дискретизации. Для хранения и передачи по каналам связи такой сигнал обычно подвергается сжатию (уменьшению объема путем отбрасывания ненужной или малозначимой информации).

Кроме импульсно-кодовой модуляции для кодирования звука применяют и другие виды цифровой модуляции (широко-импульсную, частотно импульсную и пр.). Эти виды модуляции иногда применяют для передачи звука по цифровым каналам, но в компьютерной технике их используют редко и в основном как промежуточные для некоторых видов цифровой фильтрации и создания звуковых эффектов.



5. Методы кодирования

Таблично-волновой синтез

Метод таблично-волнового (Wave-Table) синтеза лучше соответствует современному уровню развития техники. Если говорить упрощенно, то можно сказать, что в заранее подготовленных таблицах хранятся образцы звуков для множества различных музыкальных инструментов (хотя не только для них).

В технике такие образцы называют семплами. Числовые коды выражают тип инструмента, номер его модели, высоту тона, продолжительность и интенсивность звука, динамику его изменения, некоторые параметры среды, в которой происходит звучание, а также прочие параметры, характеризующие особенности звука. Поскольку в качестве образцов используются «реальные» звуки, то качество звука, полученного в результате синтеза, получается очень высоким и приближается к качеству звучания реальных музыкальных инструментов.



6. Программное обеспечение для работы со звучком

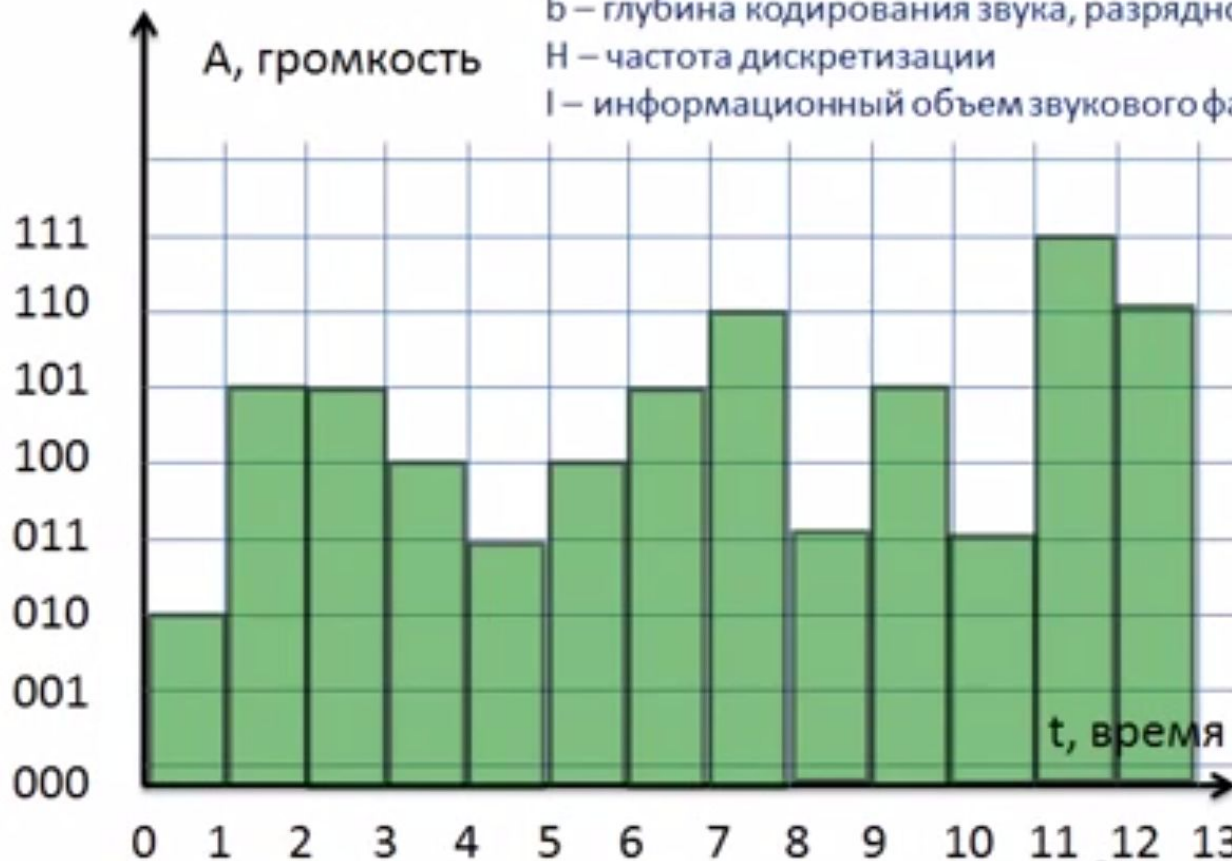
Программное обеспечение для работы со звуком можно разделить по функциональному назначению на следующие группы:

- **Плееры** - для воспроизведения звука;
- **Грabbеры** - программное обеспечение для «захвата» существующей звуковой информации и перекодировки в другой формат;
- **Устройства для создания и редактирования звука** - программное обеспечение для создания звука (с возможностями гармонизации, аранжировки, стилистической обработки, наложения голоса, добавления спецэффектов).



Информационный объем оцифрованного звука

N – количество уровней громкости
 b – глубина кодирования звука, разрядность квантования
 H – частота дискретизации
 I – информационный объем звукового файла



$$2^b = N$$

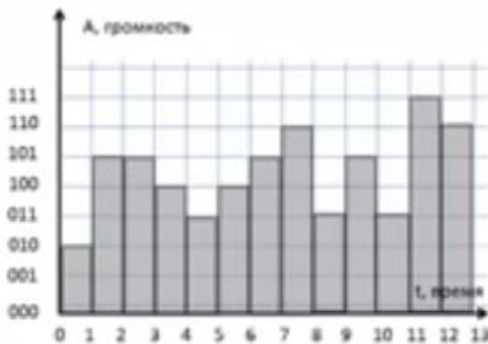
$$2^3 = 8$$

$$I = H \times t \times b$$

$$I = 13 \times 1 \times 3 = 39 \text{ (бит)}$$



Информационный объем оцифрованного звука



$$2^b = N$$

$$I = H \times t \times b$$

N – количество уровней громкости

b – глубина кодирования звука, разрядность квантования

H – частота дискретизации

I – информационный объем звукового файла

В течение 25 секунд производилась запись звука в компьютер при частоте дискретизации 22,050 кГц и разрядности квантования 16 бит. Определите информационный объем получившегося звукового файла.



Производится одноканальная (моно) цифровая звукозапись. Значение сигнала фиксируется 48 000 раз в секунду, для записи каждого значения используется 32 бит. Запись длится 4 минуты, её результаты записываются в файл, сжатия данных не производится. Какая из приведённых ниже величин наиболее близка к размеру полученного файла?

- 1) 44 Мбайт 2) 87 Мбайт 3) 125 Мбайт 4) 175 Мбайт