



# Цифровое представление звука

## Лекция 8

# Природа звука

- звуки возникают благодаря преобразованию энергии в колебания воздуха или какой-то другой упругой среды
- звук, воспринимаемый человеком имеет частотный спектр в диапазоне от 20 Гц до 20 кГц

# Форма звукового сигнала

- музыка



- речь



- шумы



# Стереофония

- оцифровка звука
  - дискретизация
  - квантование
  - кодирование



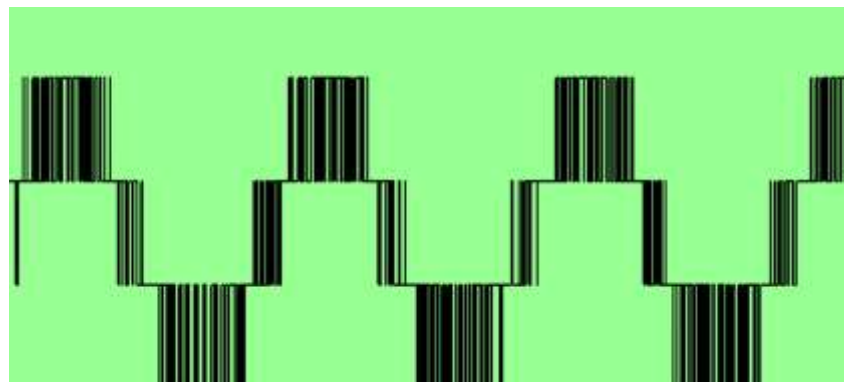
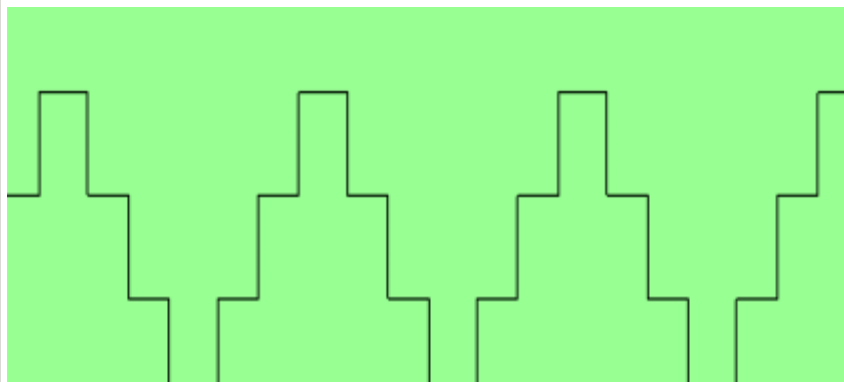
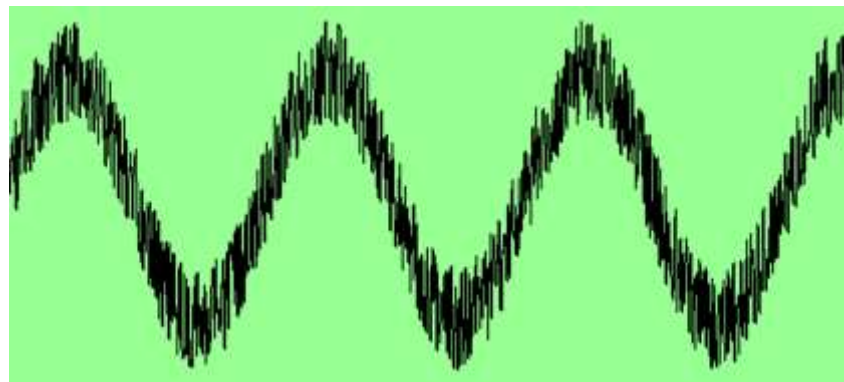
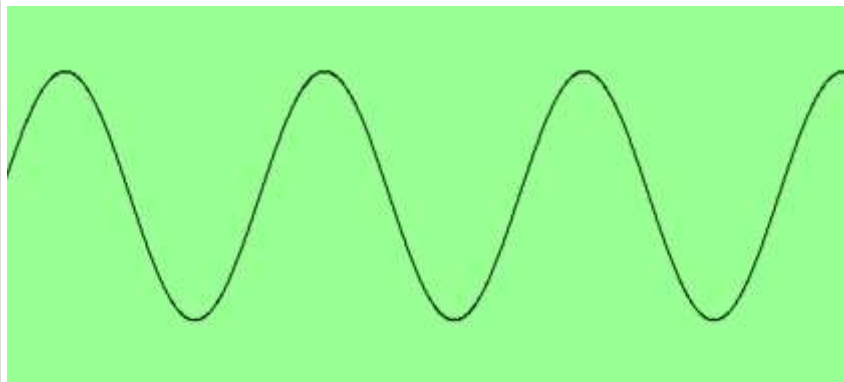
# Дискретизация звука

- по теореме отсчетов минимальная частота дискретизации должна быть в два раза выше максимальной частоты спектра, т.е.  $>40\text{кГц}$  для слышимых звуков
  - аудио компакт диски - CD:  $44.1\text{кГц}$ 
    - для Интернет аудио можно использовать  $22.05\text{кГц}$
    - для речи достаточно  $11.025\text{кГц}$
  - цифровые аудиоленты DAT:  $48\text{кГц}$ 
    - при переносе записи с CD или DAT может потребоваться передискретизация

# Квантование звука

- качество CD – 16 бит, что дает 65536 уровней квантования
  - минимально приемлемое качество при 8 битах (голосовая связь)
- ошибка квантования – разность между исходными и квантованными значениями отсчетов
  - добавление псевдослучайного шума делает ошибку квантования случайной (*Dithering*)

# Ошибка квантования



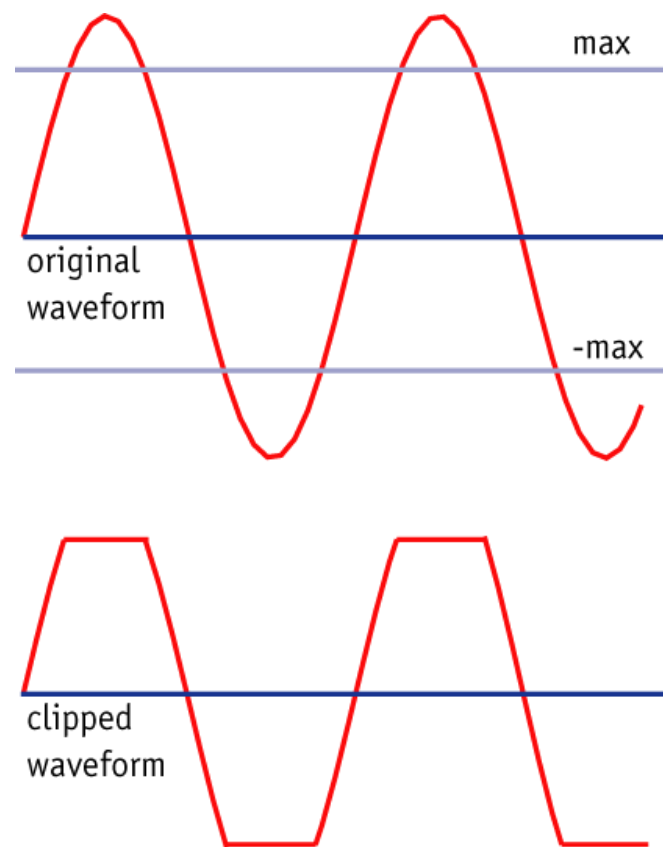
# Размер звукового файла

- частота дискретизации  $r$  в Гц
- размер выборки  $s$  в битах
- за секунду оцифрованный звук требует  $rs/8$  байт памяти
  - для CD :  $r = 44100$ ,  $s = 16$ , для каждой секунды требуется 86 килобайт ( $k=1024$ ), а за минуту около 5 Мбайт (mono)



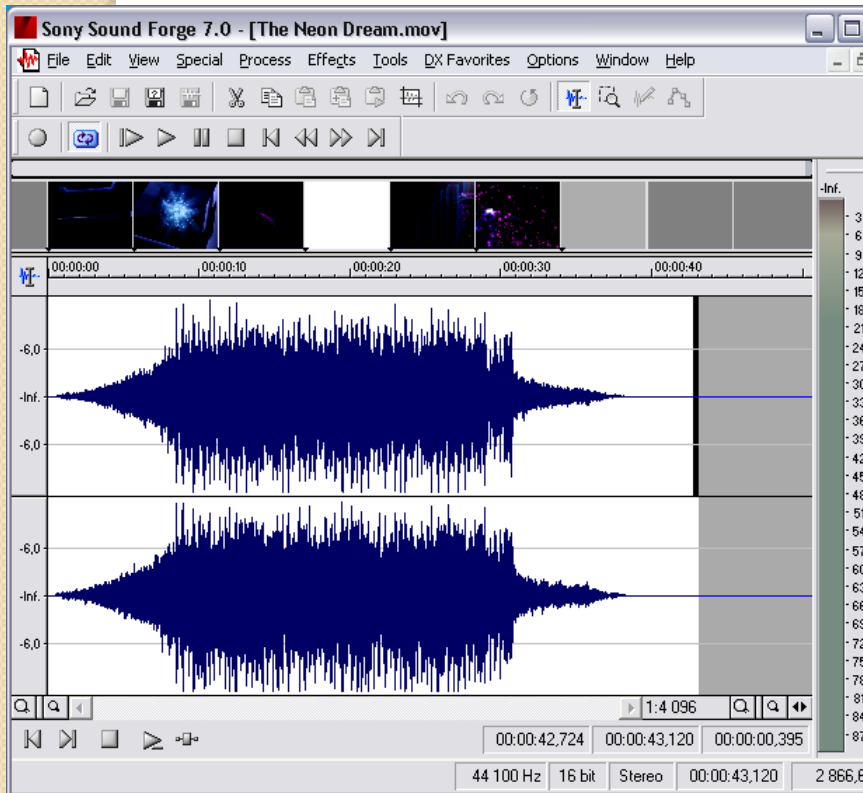
# Ограничение (Clipping)

- высокий уровень сигнала может быть ограничен
- низкий уровень сигнала сильно подвержен шуму



# Редактирование аудио

монтаж, комбинирование и  
переупорядочение звуковых клипов



# Методы синтеза

- таблично-волновой синтез
  - проигрывание нескольких сэмплов
  - дополнительная обработка звука
- синтез на основе частотной модуляции

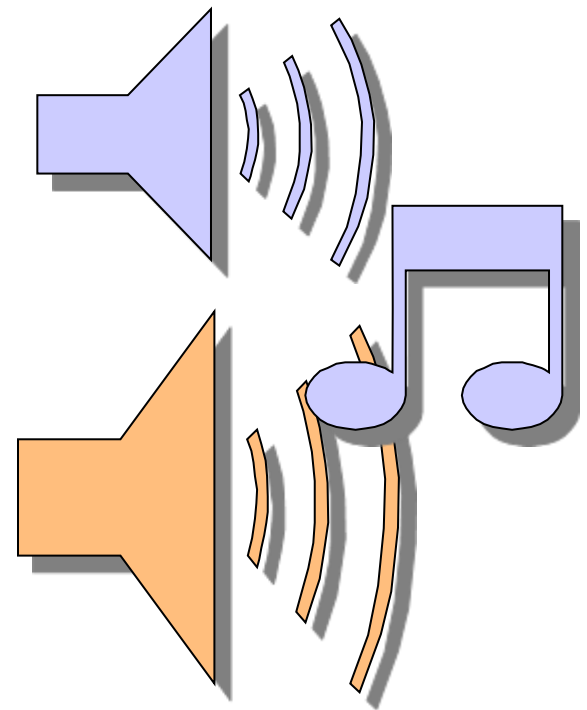
# MIDI-секвенсоры

используются для аранжировки музыкальных произведений

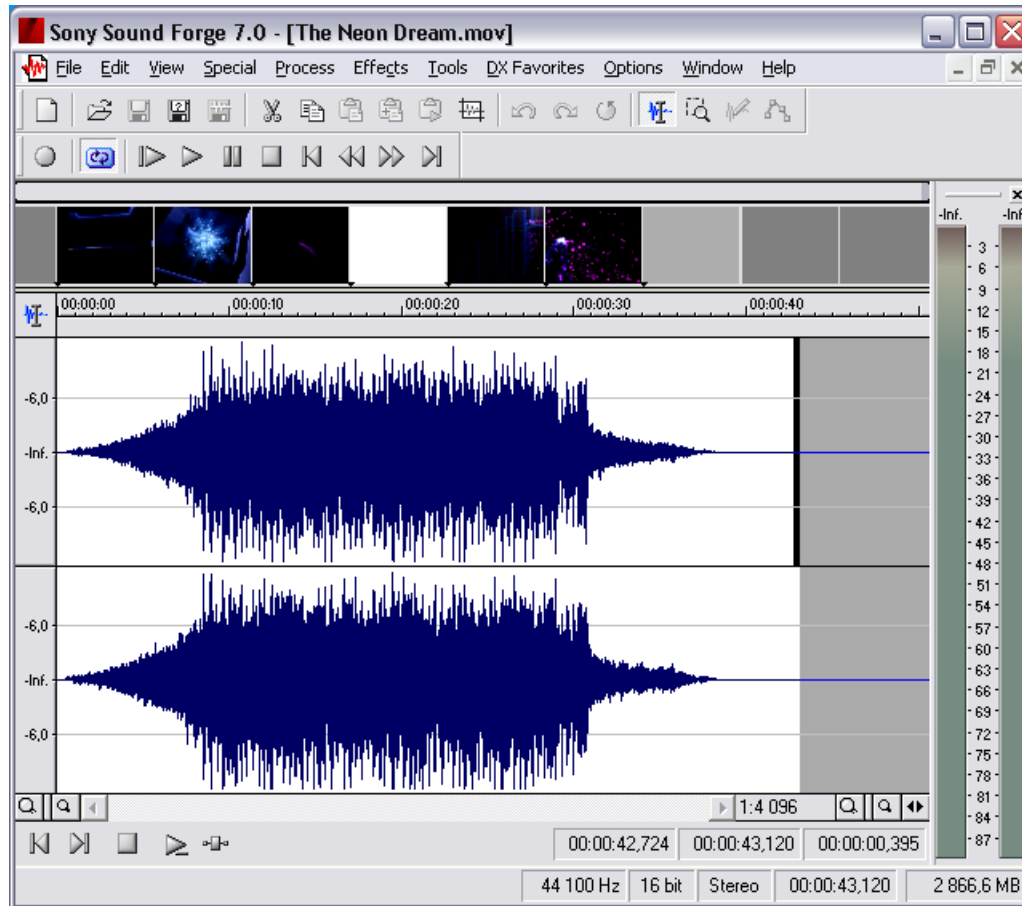
The screenshot displays the Band-in-a-Box for Windows software interface. The title bar reads "Band-in-a-Box for Windows [BLUES.SG6] Chorus 2/5 [Blues Shuffle Style] 1:49". The menu bar includes File, Edit, Styles, Opt., Play, Lyrics, Melody, Soloist, Audio, GM, User, Harmony, Not'n, and Help. The instrument selection bar shows "Piano" selected, with other options like Combo, Bass, Drums, Guitar, String, Melody, Soloist, and Thru. The instrument is set to "Rock Organ". The volume and pan controls are set to 90 and 0 respectively. The reverb and chorus settings are also visible. The keyboard view shows a sequence of notes. The control panel includes buttons for New, Open, Save, .MID, .WAV, Play, Stop, H, From, Melodist, Soloist, <, Juke, >, and Rec. The song information bar shows "C Blues (Pop)" and "ZZBLUSHF.STY". The Wizard panel includes buttons for Intro, RME, EDIT, C7, II, SONG TITLE, L, Pref, PG, and MIDI. The main sequencer area shows a grid of notes with a 2/5 time signature and a tempo of 140. The notes are: 1a C, 2 F, 3 C, 4 C, 5 F, 6 F, 7 C, 8 C, 9 G, 10 F, 11 C, 12 C, 13a C, 14 F, 15 C, 16 C, 17 C, 18 F, 19 C, 20 C, 21 C, 22 F, 23 C, 24 C, 25 C, 26 F, 27 C, 28 C, 29 C, 30 F, 31 C, 32 C.

# Программы подготовки аудио

- *Sound Forge*
- *Adobe Audition*
- *WaveLab*
- *PowerTracks Pro*
- *Cakewalk Pro Audio*
- *Logic Audio Platinum*
- *Cubase SX3*



# Sony Sound Forge



представляет данные  
в двух видах (спектр и  
фонограмма)

# Adobe Audition

The screenshot displays the Adobe Audition interface for the file 'The Neon Dream.mp3'. The main window shows a waveform view with two channels of audio data, rendered in a bright cyan color. The vertical axis represents amplitude in decibels (dB), ranging from -12 to 0. The horizontal axis represents time in hours, minutes, and seconds (hms), with markers every 5 seconds from 0:05.0 to 0:40.0. The interface includes a menu bar (File, Edit, View, Effects, Generate, Favorites, Options, Window, Help), a toolbar with various editing tools, and a workspace set to 'Edit View (Default)'. On the left, there is a 'Main' panel with a list of effects categories: Multitrack Effects, Amplitude, Delay Effects, Filters, Restoration, Special, Time/Pitch, Generate, VST, and a group of 'Apply' options (Apply Invert (process), Apply Reverse (process), Apply Mute (process)). Below the effects list are buttons for 'Group By Category' and 'Group Real-Time Effects'. The bottom section of the interface contains a transport control panel with play, stop, and record buttons, a time display showing '0:00.000', a zoom control, and a 'Selection/View' panel with input fields for 'Begin' and 'Length' (both currently at 0:00.000). At the very bottom, a status bar provides technical details: 'Opened in 0.64 seconds', 'R: -29.1dB @ 0:00.887', '22050 • 16-bit • Stereo', '3.64 MB', '74.26 GB free', '251:08:05.04 free', and 'Alt Waveform'.

# Форматы аудиофайлов

<b>.AIFF</b>	<b>Audio Interchange Format</b>
<b>.WAV</b>	<b>Windows Audio</b>
<b>.AU/.SND</b>	<b>Sun Audio</b>
<b>.WMA</b>	<b>Windows Media Audio</b>
<b>.RA</b>	<b>Real Audio</b>
<b>.MIDI</b>	<b>MIDI Sound</b>
<b>.MP3</b>	<b>MPEG Audio</b>

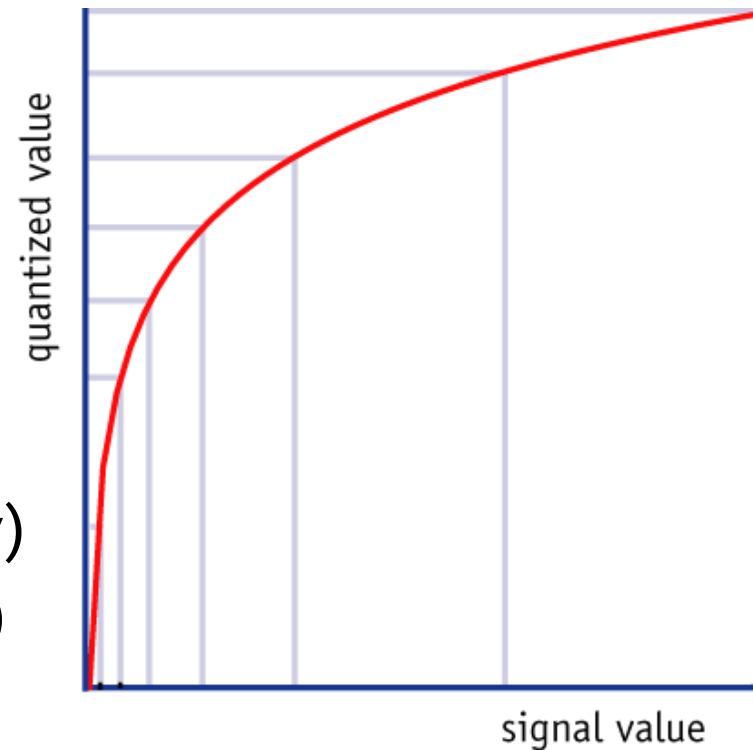


# Сжатие звука (Compression)

- сложная и непредсказуемая природа звуковых сигналов не позволяет сжимать их без потерь
  - стерео запись CD качества на 3 минуты требует 25 Мбайт
- удаление пауз с указанием их длительности
  - высокие частоты связаны с быстрыми изменениями звука

# Компандирование (Componding)

- нелинейное квантование (логарифмическая характеристика)
- меньшее число уровней квантования для всего диапазона
- $\mu$ -характеристика ( $\mu$ -law)
- A-характеристика (A-law)



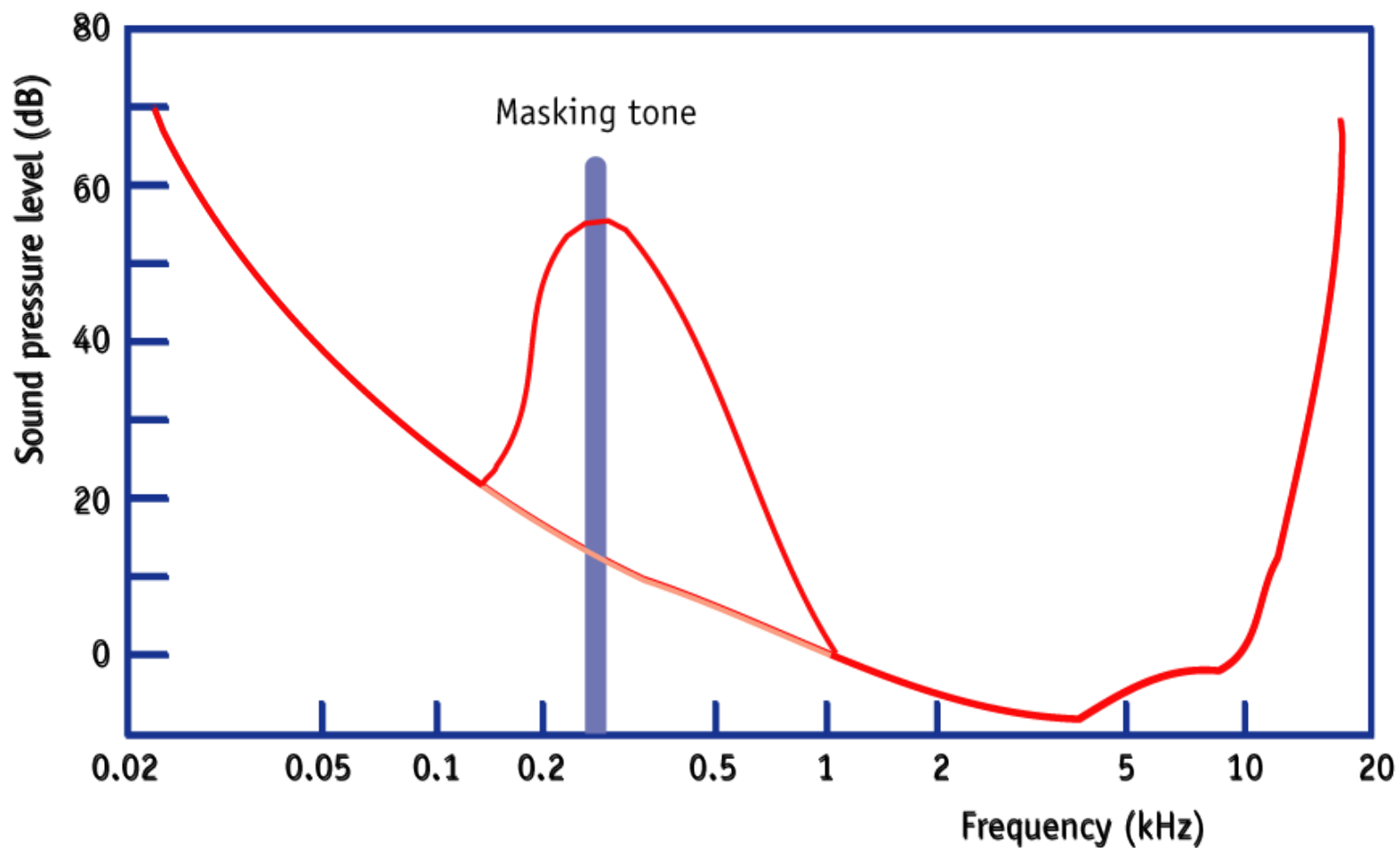
# Импульсно-кодовая модуляция

- Дифференциальная импульсно-кодовая модуляция (Differential Pulse Code Modulation – DPCM)
  - вычисляет предсказанное значение выборки
- Адаптивная ДИКМ (*Adaptive* Differential Pulse Code Modulation – ADPCM)
  - основана на записи разностей последовательных выборок

# Сжатие на основе восприятия

- использует *психо-акустическую модель* восприятия, имитирующую реакцию системы ухо-мозг на звук
- *порог слышимости* – звук, имеющий минимальный уровень
- *маскировка* – «затенение» звука другим с большим уровнем

# Порог слышимости



# Алгоритм сжатия

- разделение сигнала на частотные полосы (обычно 32)
- вычисление среднего уровня в каждой полосе с определением *порога маскировки*
- отбрасывание полосы при меньшем уровне сигнала
- квантование с меньшим числом бит для остальных полос

# Стандарт MP3

- *MPEG Audio, Layer 3*
- три уровня сжатия аудио в MPEG-1 (MPEG-2 идентичен по звуку)
- скорость потока 192Kb/s - Layer 1, 128 Kb/s - Layer 2, 64 Kb/s - Layer 3
- сжатие 10:1 при высоком качестве
- кодирование с переменной скоростью передачи (VBR)

# Интерфейс MIDI

- *Musical Instruments Digital Interface*
- Стандартный протокол связи между электронными инструментами синтезаторы (synthesizers), сэмплеры (samplers), драм-машинами (drum machines)
- Файлы MIDI можно внедрять в мультимедийную продукцию



# Спецификация General MIDI

- инструменты, контролируемые MIDI, имеют набор *голосов*
- в стандарте General MIDI определены 128 стандартных голосов
  - хотя GM задает только имена голосов, он не гарантирует, что для каждого имени будет воспроизводиться один и тот же звук на разных инструментах

# Преобразования MIDI

- MIDI записывает музыкальную структуру звука
  - преобразование MIDI в аудио
  - преобразование аудио в MIDI
  - музыку в исполнении обычных инструментов можно записывать в виде нот

**Цикл лекций подготовлен в 2011/2012 уч. году  
Кузнецовым Игорем Ростиславовичем,  
доцентом кафедры радиоэлектронных средств  
Санкт-Петербургского  
Государственного электротехнического  
университета «ЛЭТИ»**

Прочитан в дисциплине  
«Информационные технологии»