



MSU GRAPHICS & MEDIA LAB

Лектор: Лукин Алексей Сергеевич

# Спектральный анализ

## Быстрая свертка

## Фильтрация

# План



- **Спектральный анализ**
  - ▶ Мощность сигнала, теорема Парсевала
  - ▶ Спектрограммы, усреднение во времени
  - ▶ Частотно-временное разрешение
- **Быстрая свертка**
  - ▶ Свертка в частотной области
  - ▶ Разбиение ядра
- **Фильтрация**
  - ▶ Основные понятия
  - ▶ Фильтры для изображений, их частотные характеристики
  - ▶ Применения: эквалайзеры, реверберация

# Мощность и амплитуда



MSU GRAPHICS & MEDIA LAB

- Мощность пропорциональна квадрату амплитуды

$$P \approx \frac{1}{N} \sum_{i=1}^N x[i]^2$$

- RMS (среднеквадратичное значение) сигнала:

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x[i]^2}$$

(пропорционально амплитуде)

- Окна при вычислении RMS
- Вычисление RMS в реальном времени:

$$RMS^2[n] = (1 - \alpha)RMS^2[n-1] + \alpha x[n]^2$$

(экспоненциальное окно)

# Теорема Парсеваля



MSU GRAPHICS & MEDIA LAB

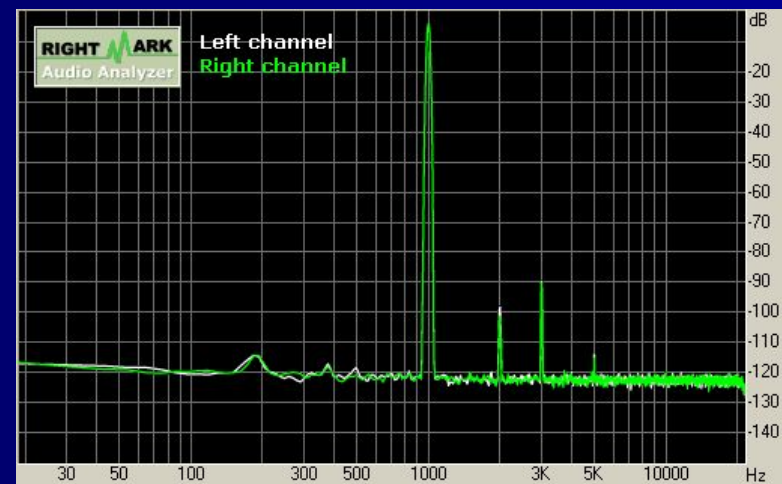
- Энергия сигнала равна энергии спектра  
(возможно, с точностью до множителя,  
зависящего от нормализации  
в формулах преобразования Фурье)
- Применение теоремы для оценки энергии  
различных частотных составляющих сигнала
- Пример:

вычисление THD

(total harmonic distortion)

$$THD = \frac{RMS_{distortion}}{RMS_{total}} \cdot 100\%$$

$$\sum_i x[i]^2 = \sum_k |X[k]|^2$$

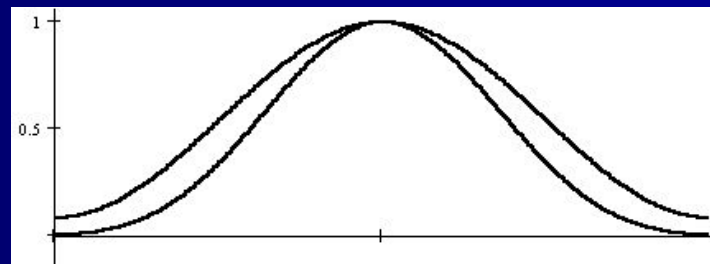


# Спектральный анализ



- Как вычислить и отобразить спектр сигнала?
  1. Взять нужный отрезок сигнала длины  $2^m$ ; если нужный отрезок короче – дополнить его нулями.
  2. Если нужно – умножить сигнал на весовое окно, плавно спадающее к краям. Обычно это улучшает свойства спектра.
  3. Вычислить FFT.
  4. Перевести комплексные коэффициенты в полярную форму: получить амплитуды.
  5. Отобразить график зависимости амплитуды от частоты.

Примеры весовых окон



# Свойства ДПФ



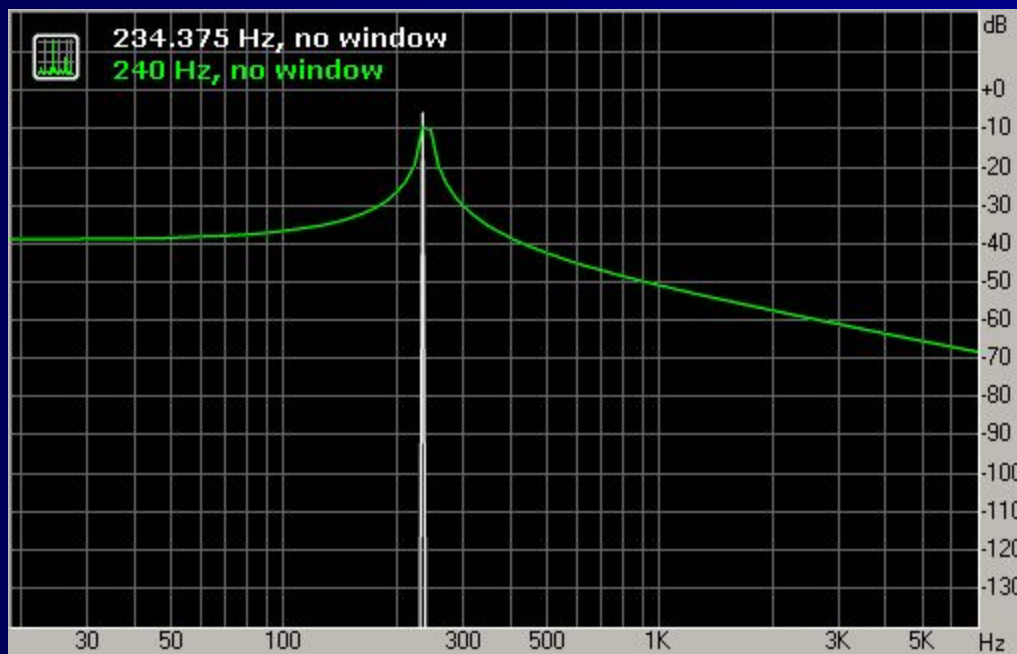
$$X_k = \sum_{n=0}^{N-1} x[n] \exp\left(-\frac{2\pi i}{N} kn\right) \quad x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X_k \exp\left(\frac{2\pi i}{N} kn\right)$$

- Линейность
- Комплексное ДПФ от вещественного сигнала обладает симметрией относительно коэффициента  $k = N/2$ . Таким образом, только половина коэффициентов являются значимыми.
- Периодическое продолжение сигнала
- Циклический сдвиг отсчётов сигнала не изменяет амплитудного спектра, а меняет только фазовый

# Размытие спектра



- Как выглядит спектр синусоиды, если её частота совпадает/не совпадает с базисными частотами ДПФ?

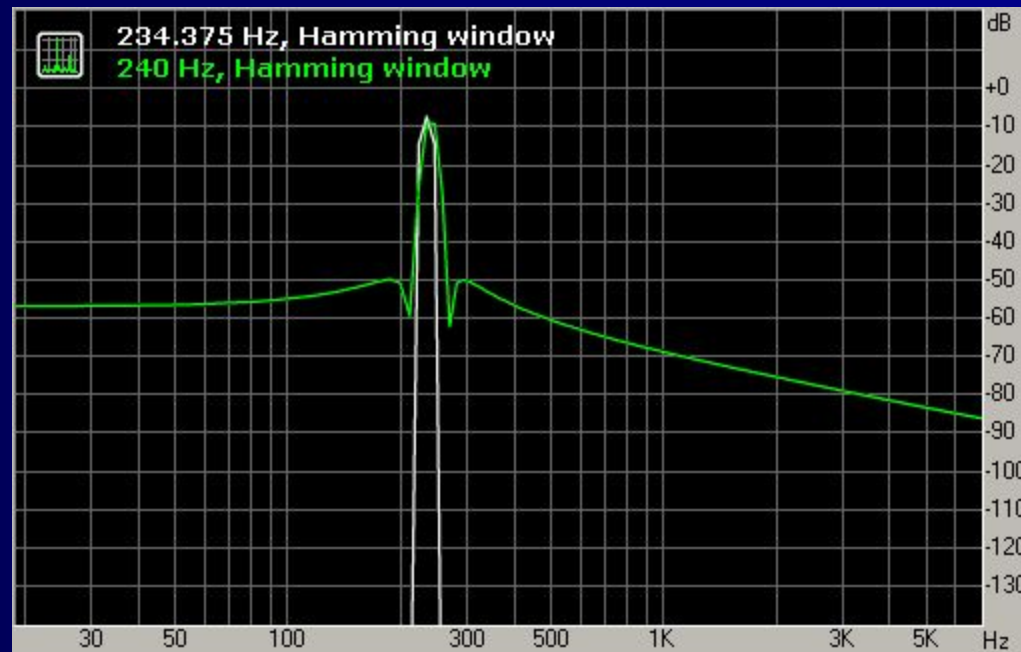


нет окна (прямоугольное окно)

# Весовые окна



- Весовые окна уменьшают эффект размытия спектра, подавляя боковые лепестки за счёт расширения главного лепестка



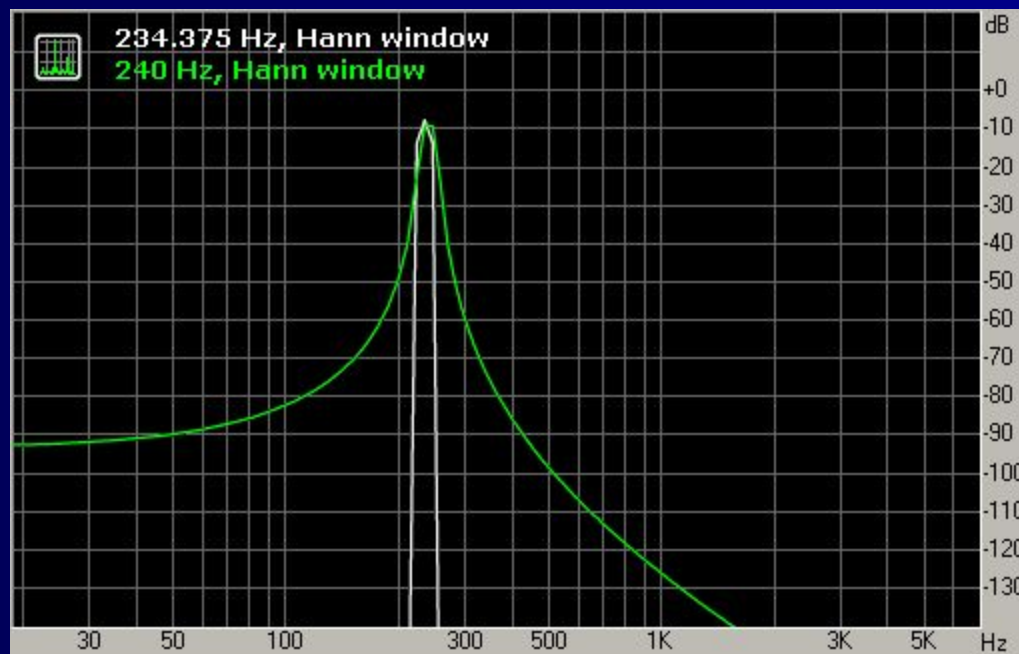
окно Хэмминга (Hamming)



# Весовые окна



- Весовые окна уменьшают эффект размытия спектра, подавляя боковые лепестки за счёт расширения главного лепестка

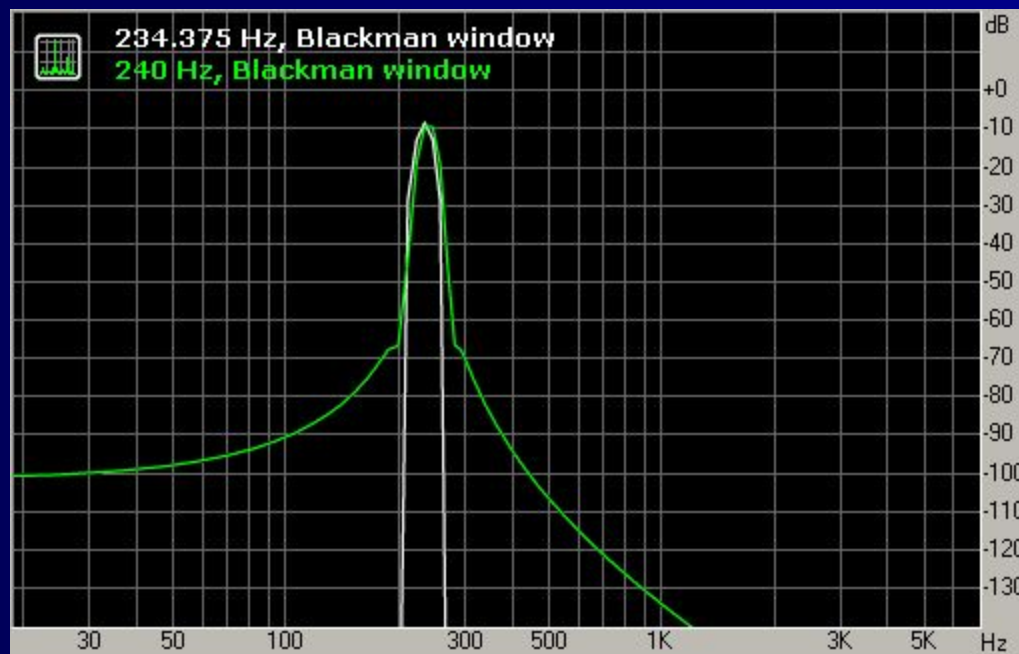


окно Хана (Hann)

# Весовые окна



- Весовые окна уменьшают эффект размытия спектра, подавляя боковые лепестки за счёт расширения главного лепестка

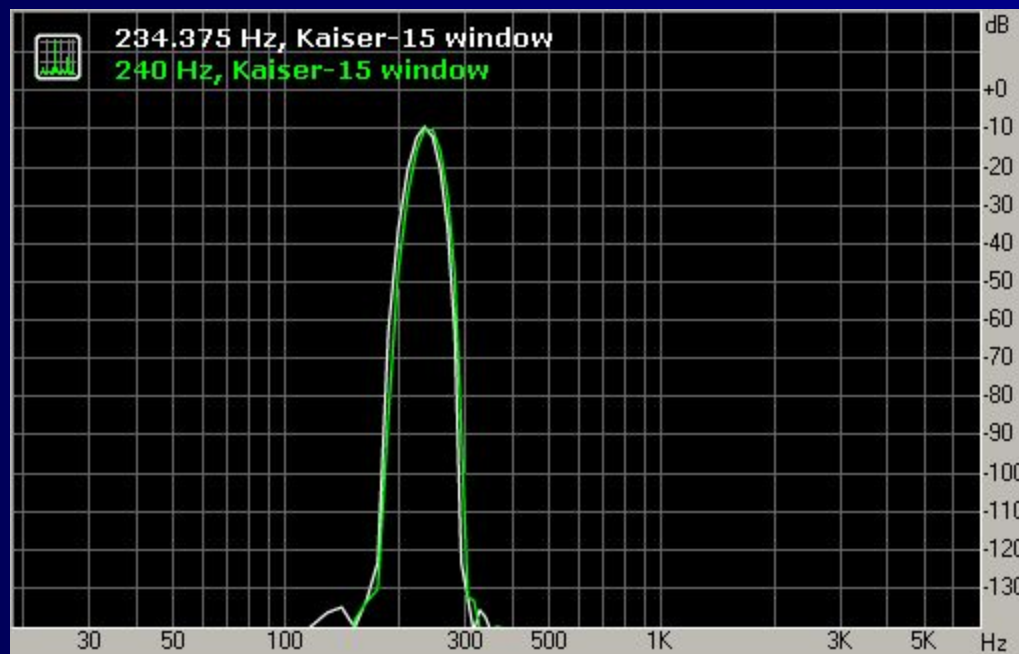


окно Блэкмана (Blackman)

# Весовые окна

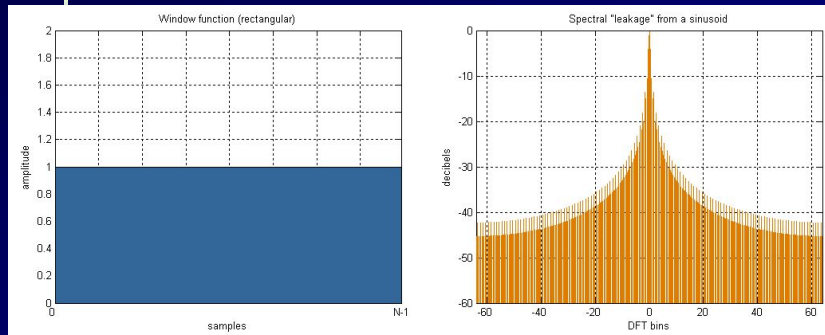


- Весовые окна уменьшают эффект размытия спектра, подавляя боковые лепестки за счёт расширения главного лепестка

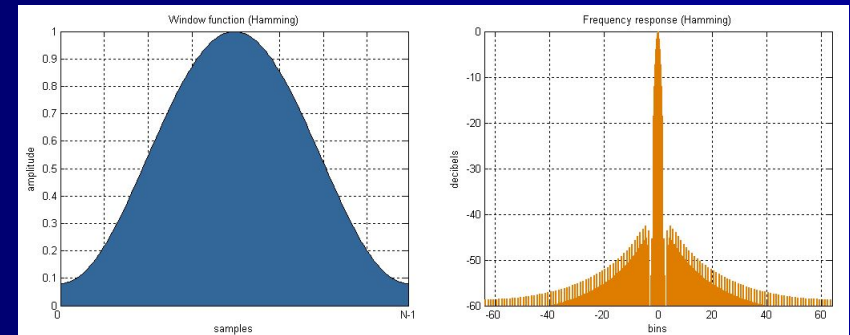


окно Кайзера (Kaiser),  $\beta = 15$

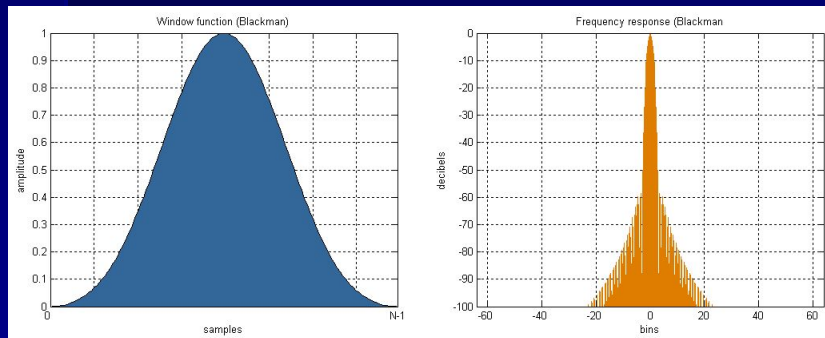
# Весовые окна



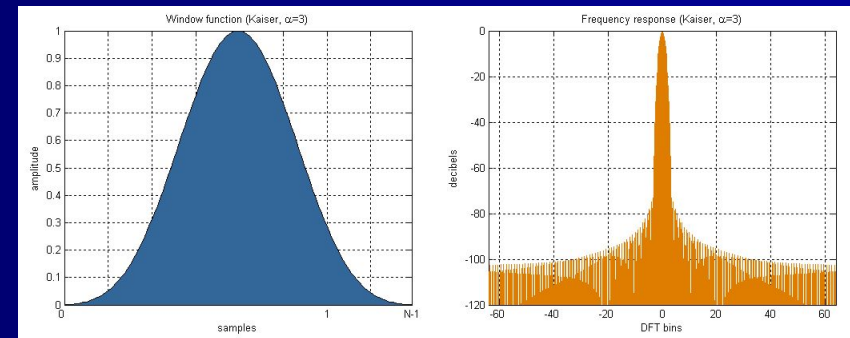
Прямоугольное (нет окна)



Hamming



Blackman



Kaiser

# Спектральный анализ

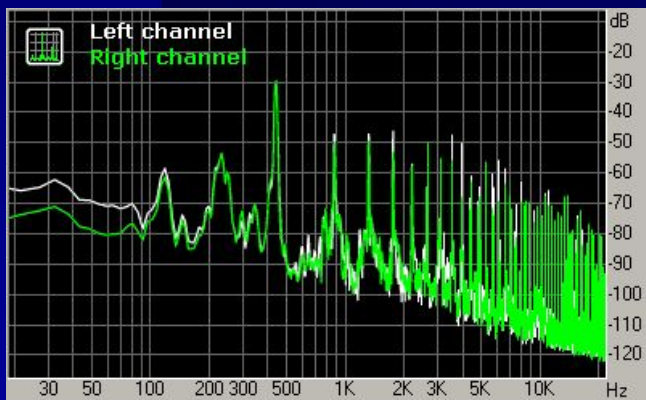


- Отображение спектра звука: спектрограмма
  - ▶ Спектрограмма – график зависимости амплитуды от частоты
  - ▶ Низкие частоты – слева, высокие – справа
  - ▶ Часто применяется логарифмический масштаб частот и амплитуд: “log-log-спектрограмма”
  - ▶ Временное и частотное разрешение спектрограммы

Децибелы:

$$D = 20 \lg \frac{A_1}{A_0}$$

$A_1$  – амплитуда измеряемого сигнала,  
 $A_0$  – амплитуда сигнала, принятого за  
начало отсчета (0 дБ)



Разница на 6 дБ – разница по амплитуде в 2 раза,  
разница на 12 дБ – разница по амплитуде в 4 раза.

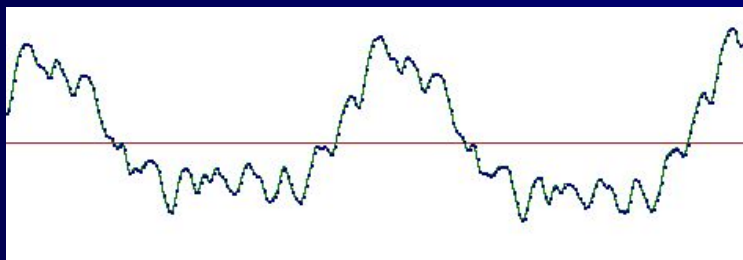
Часто за 0 дБ принимается либо самый тихий слышимый звук, либо самый мощный звук, который может воспроизвести аудиоустройство.

# Спектральный анализ

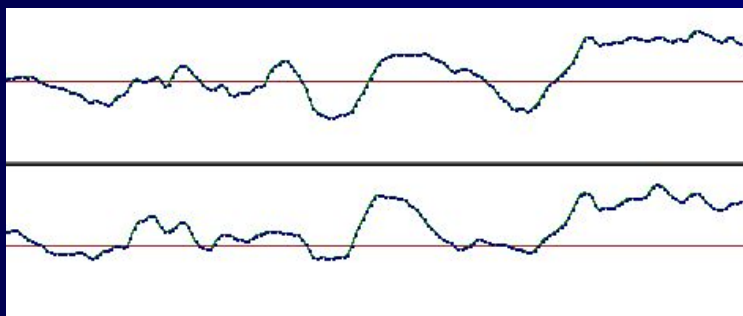
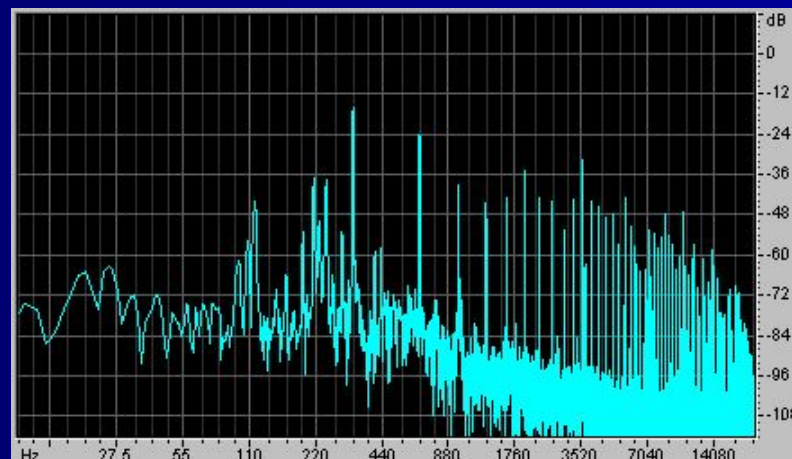


MSU GRAPHICS & MEDIA LAB

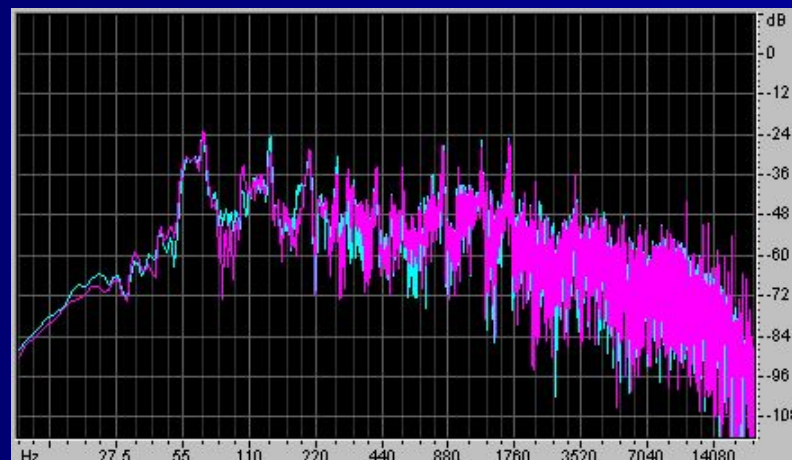
- Примеры звуков и их спектров



Нота на гитаре



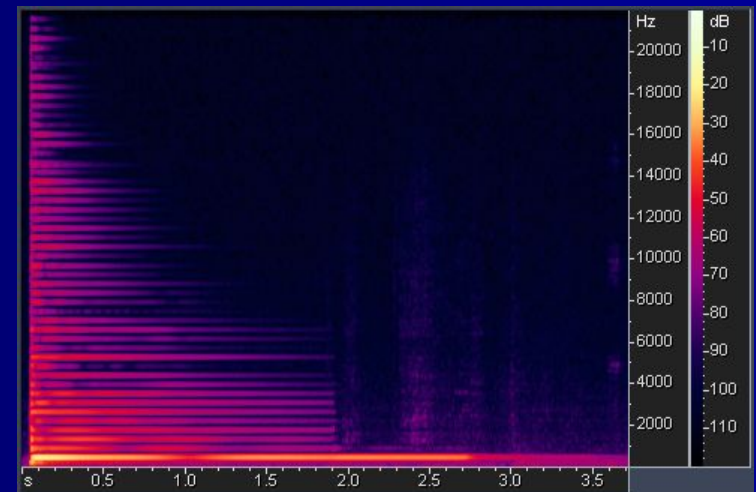
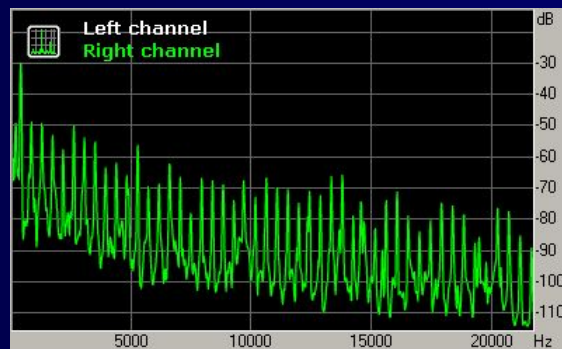
Песня (стерео запись)



# Спектральный анализ

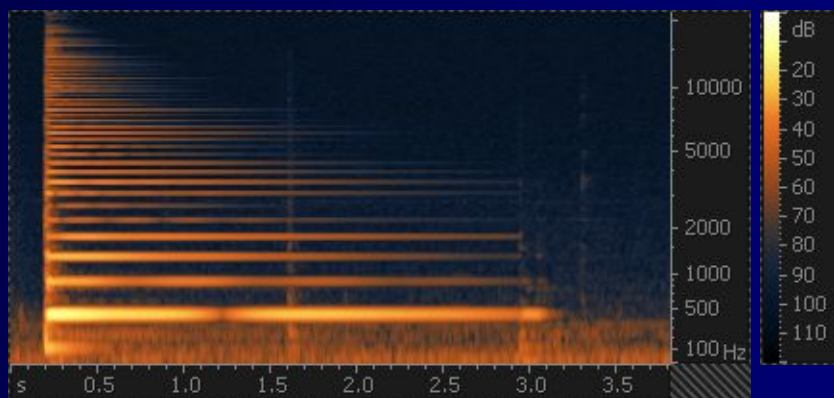
- Отображение спектра звука: спектрограмма (сонограмма)
  - ▶ Спектрограмма – график зависимости амплитуды от частоты и от времени, показывает изменение спектра во времени
  - ▶ Short Time Fourier Transform (STFT)

$$STFT[n, \omega] = \sum_{m=-\infty}^{+\infty} x[n+m] \cdot w[m] \cdot e^{-i\omega m}$$

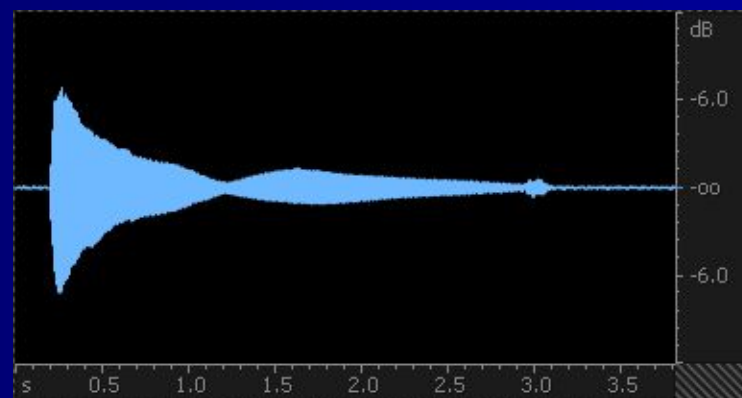


# Спектральный анализ

- Отображение спектра звука: спектрограмма (сонограмма)
  - ▶ Спектрограмма – график зависимости амплитуды от частоты и от времени, показывает изменение спектра во времени
  - ▶ Низкие частоты – снизу, высокие – сверху
  - ▶ Время идет справа налево
  - ▶ Амплитуда – яркость или цвет
  - ▶ Частотное и временное разрешение
  - ▶ Short Time Fourier Transform (STFT)



Спектрограмма

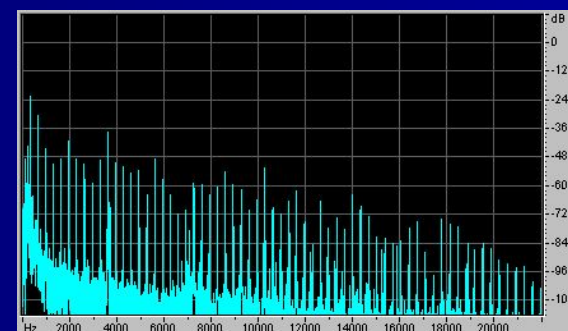
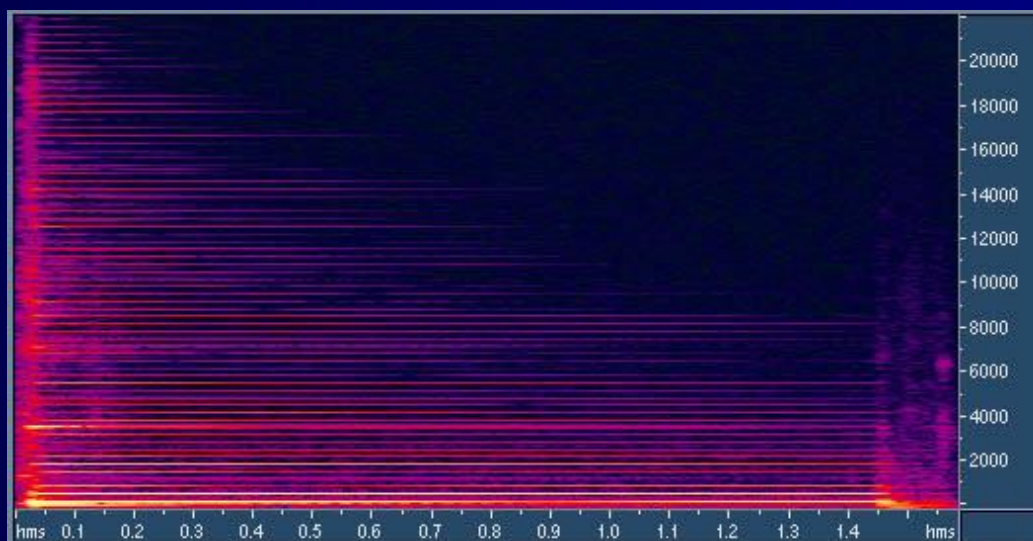
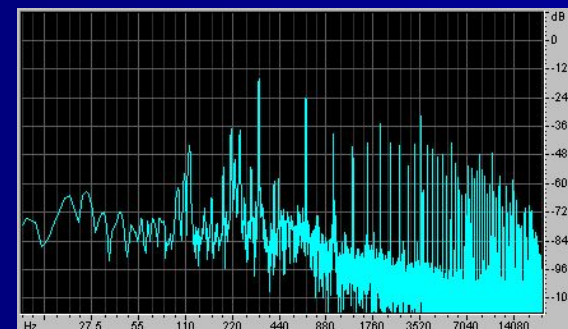
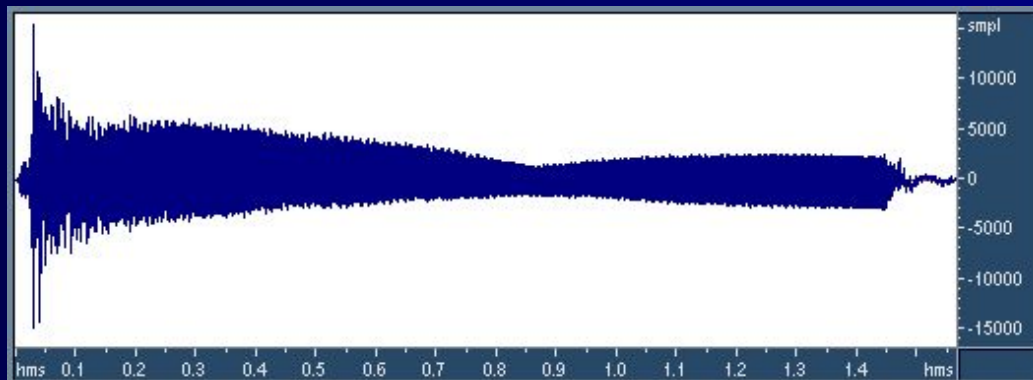


Осциллограмма (форма волны)



# Спектральный анализ

- Примеры звуков и их спектрограмм



Нота на гитаре

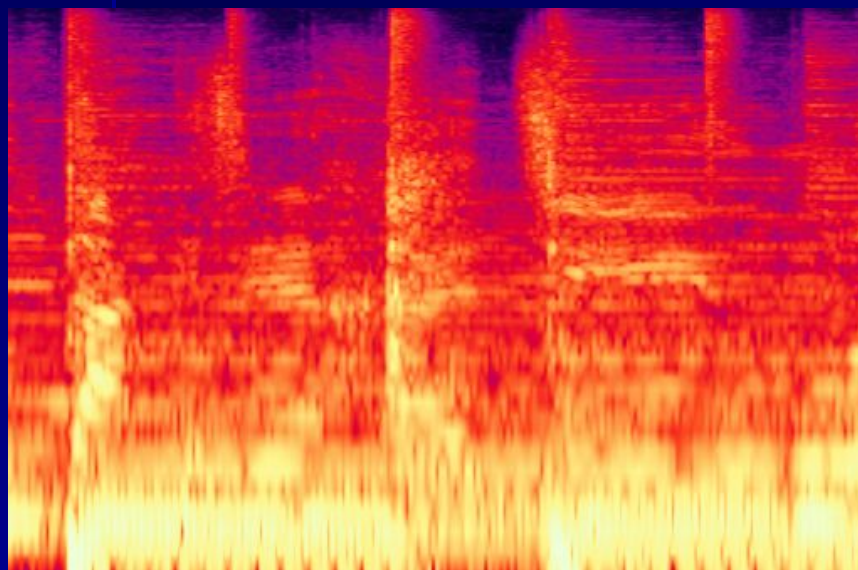
# Спектральный анализ



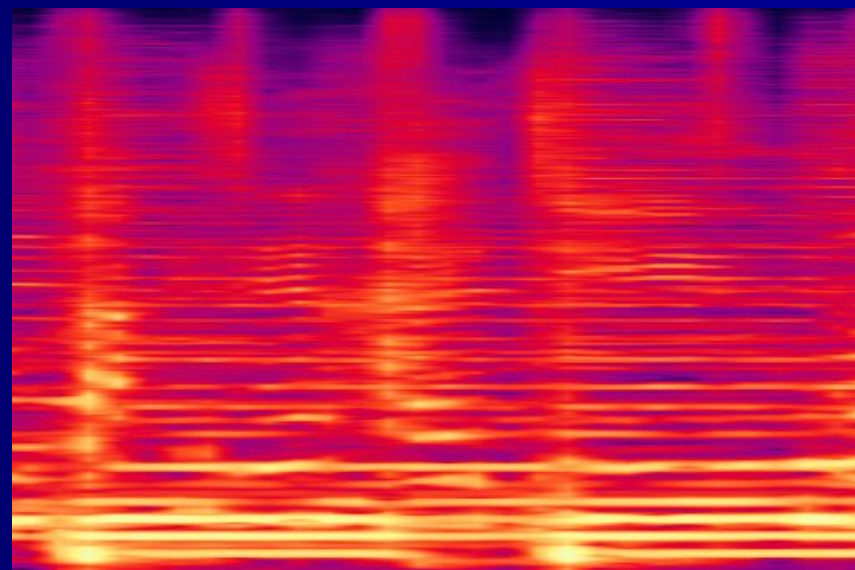
- Спектрограммы шума
- Типы шумов: белый (постоянная PSD), розовый (3 дБ/окт.), коричневый (6 дБ/окт.)
- Вариация и усреднение спектрограмм

# Построение спектрограмм

- Мел-шкала частот
- Недостаточное частотное разрешение на НЧ, недостаточное временное разрешение на ВЧ



Обычная спектрограмма,  
окно 512 точек

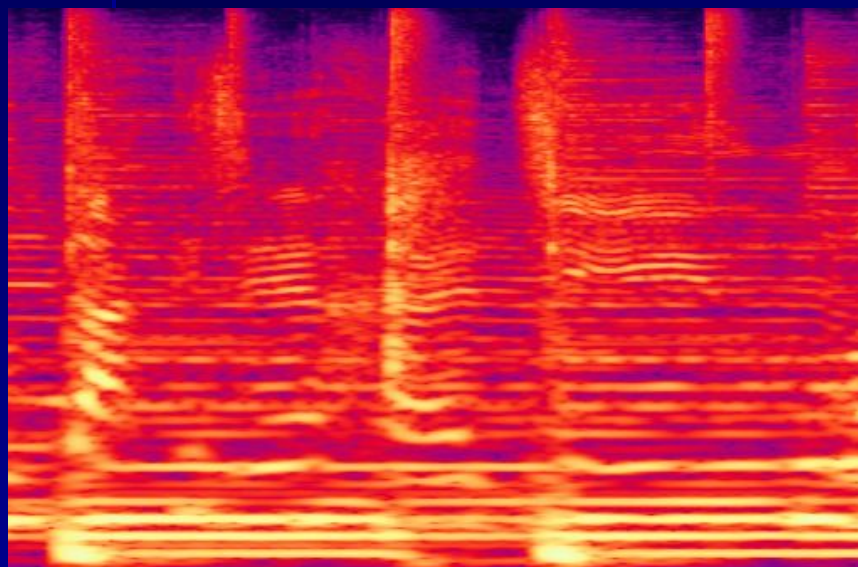


Обычная спектрограмма,  
окно 4096 точек

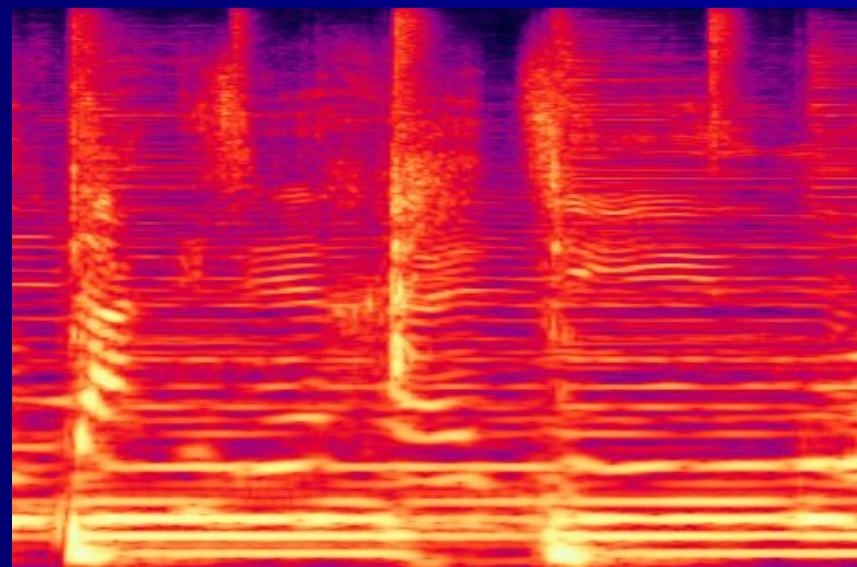
# Построение спектрограмм



- Адаптивное частотно-временное разрешение спектрограмм
- Критерий оптимальности: максимально компактная локализация энергии (минимальное размытие спектрограммы)



Адаптивная спектрограмма,  
разрешение зависит от частоты



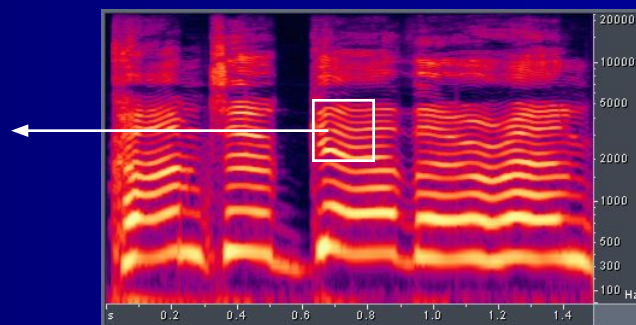
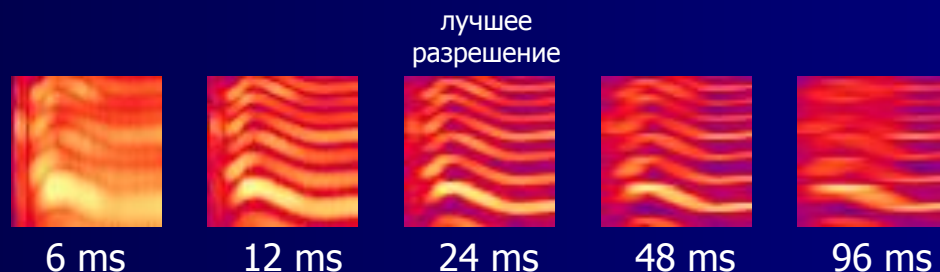
Адаптивная спектрограмма,  
оптимальное разрешение

# Построение спектрограмм

- Критерий оптимальности: максимальное компактирование (локализация) энергии

$$r_0 = \arg \min_r S_r \quad S_r = \frac{\text{norm}_{L1}(a)}{\text{norm}_{L2}(a) \cdot \sqrt{n}} = \frac{\sum_{i=1}^n a_{i,r}}{\sqrt{n \sum_{i=1}^n a_{i,r}^2}}$$

Здесь  $a_{i,r}$  – амплитуды STFT в блоке разрешения  $r$ ,  
 $S_r$  – мера размытости спектра для данного разрешения  $r$ ,  
 $r_0$  – лучшее разрешение, для которого размытость минимальна.



# Свертка и фильтрация



## ■ Основные термины

- ▶ Свертка (*convolution*), фильтрация (*filtering*)
- ▶ Фильтр (*filter*), ядро фильтра (*kernel*)
- ▶ Импульсная, частотная и фазовая характеристики (*impulse, frequency, phase response*)

$$y[n] = h[n] * x[n] \qquad y[n] = \sum_{k=-\infty}^{+\infty} x[n-k] \cdot h[k]$$

## ■ Применения фильтрации

- ▶ Анти-алиасинг изображений, нахождение границ
- ▶ Звуковой эквалайзер
- ▶ Моделирование реверберации помещения

# Быстрая свертка



MSU GRAPHICS & MEDIA LAB

- Прямое вычисление:  $M \cdot N$  умножений ( $M$  – размер ядра свертки,  $N$  – длина сигнала)
- Теорема свертки: свертка\* во временной области эквивалентна умножению в частотной области, умножение во временной области эквивалентно свертке\* в частотной области.
- Алгоритм быстрой свертки:
  1. Вычислить спектры сигнала и ядра свертки (FFT)
  2. Перемножить эти спектры
  3. Вернуть полученный спектр во временную область (IFFT)
- Почему это быстрее? Потому что переход в частотную область и обратно быстрый: FFT

---

\* Речь идет о т.н. круговой свертке

# Быстрая свертка



MSU GRAPHICS & MEDIA LAB

- Как изменяется длина сигнала при свертке? Она увеличивается на длину ядра минус 1 (т.к. каждый входной отсчет превращается в ядро и они складываются с наложением)
- Значит, если взять сигнал длины  $N$ , ядро длины  $M$  и произвести свертку через FFT размера  $N$ , то результат свертки (длины  $N+M-1$ ) не поместится в результате IFFT (длины  $N$ ). Произойдет **круговая свертка** (заворачивание результата по времени).
- Следовательно, для предотвращения круговой свертки надо взять размер FFT как минимум  $N+M-1$



# Вычисление свертки



- Быстрая свертка через FFT  $IFFT(FFT(x(n)) \cdot FFT(h(n)))$ 
  - ▶ Неоптимальна по скорости, когда длины сигнала и ядра сильно различаются
  - ▶ Неприменима при потоковой обработке (*когда входной сигнал поступает в реальном времени и его длина неизвестна*)
- Секционная свертка
  - ▶ Используем свойство линейности и инвариантности к сдвигу: разбиваем сигнал на блоки и делаем быструю свертку блочно
  - ▶ Увеличение длины каждого блока на  $M-1 \rightarrow$  сложение результатов с перекрытием (*метод OLA: overlap-add*)

Подробности – см. в методичке

# Разбиение ядра



- Проблема секционной свертки: сигнал надо накапливать блоками, сравнимыми по размеру с длиной ядра свертки (*иначе – неэффективно*) → задержки при real-time
- Partitioned convolution (*свертка с разбиением ядра*)
  - ▶ Цель – уменьшение задержки
  - ▶ Недостаток – некоторое увеличение времени обработки

Секционная свертка:  $x(n) * h(n) = (x_1(n) + \dots + x_m(n) + \dots) * h(n)$

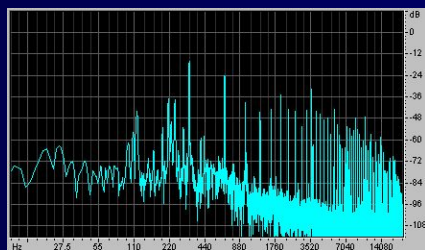
Разбиение ядра:  $x(n) * h(n) = (x_1(n) + \dots + x_m(n) + \dots) * (h_1(n) + \dots + h_k(n))$

(раскрываем скобки)

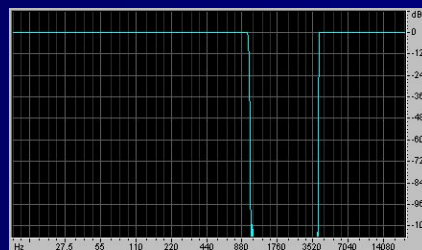
# Частотная характеристика



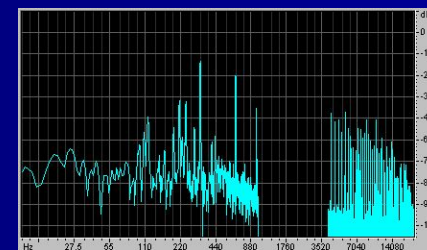
- Получение частотной характеристики по ядру
  1. Дополнение нулями
  2. FFT
  3. Берем модули комплексных коэффициентов
- Свертка = перемножение частотных характеристик



\*



=

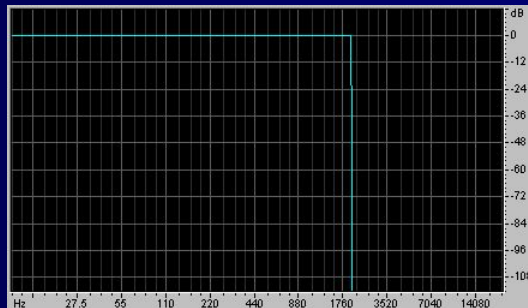


Перемножение амплитуд = сложение децибелов

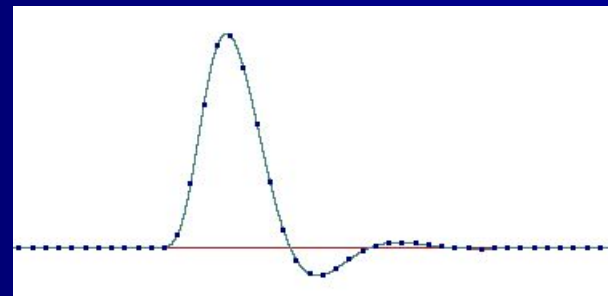
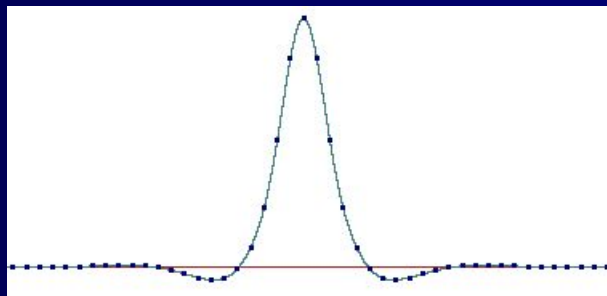
# Виды фильтров



- Идеальные и реальные фильтры, виды искажений



- FIR (*finite impulse response*) и IIR (*infinite i. r.*) фильтры
- Линейность ФЧХ («линейность фазы»)



# Построение фильтров



- Простейший НЧ-фильтр – windowed sinc

$$\frac{\sin(\alpha n)}{\alpha n} \cdot w(n)$$

- Метод спектральной инверсии\*
  - ▶ Вычитание фильтра из единичного импульса – вычитает его АЧХ из единичной
  - ▶ Инверсия каждого второго отсчета в фильтре – разворачивает АЧХ фильтра по частоте

---

\* Применим для фильтров с линейной ФЧХ

# Построение фильтров



- **Метод весового окна**
  - ▶ Построение фильтра с линейной фазой по произвольной заданной частотной характеристике
  - ▶ Частотная характеристика приближается с любым заданным уровнем точности
- **Более сложные методы**
  - ▶ Чуть меньшее число точек ядра при соблюдении той же частотной характеристики
  - ▶ Специальные требования (*к фазовой характеристике, к монотонности частотной характеристики и т.п.*)

# Построение фильтров



- **Метод весового окна** (*идея метода – взять обратное FFT от требуемой АЧХ*)
  1. Дана требуемая АЧХ (частотная характеристика)
  2. Произвольно выбираем размер FFT  $N$  и берем амплитуды коэффициентов Фурье по АЧХ ( $N/2+1$  точка)
  3. Записываем нули в фазы и делаем IFFT
  4. Циклический сдвиг ядра
  5. Применение весового окна
  6. Контроль реальной полученной АЧХ (*см. получение АЧХ по ядру фильтра*)
  7. Если АЧХ недостаточно соответствует требованиям – увеличиваем размер FFT и снова конструируем фильтр

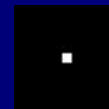
*Подробности – см. в методичке*

# Двумерные фильтры



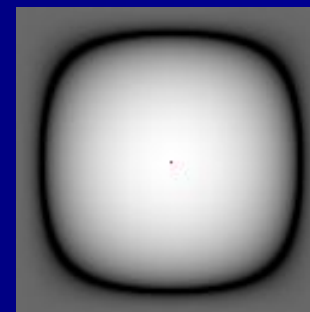
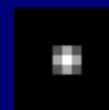
- Единичный импульс

$$\delta[m, n]$$



- Простейшее размытие

$$Ker[k, p] = \frac{1}{15} \cdot \begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{vmatrix}$$





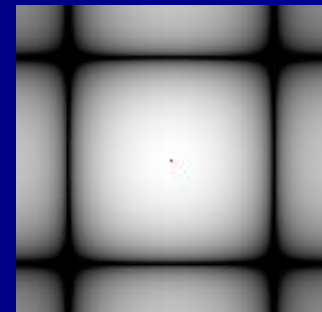
# Двумерные фильтры



MSU GRAPHICS & MEDIA LAB

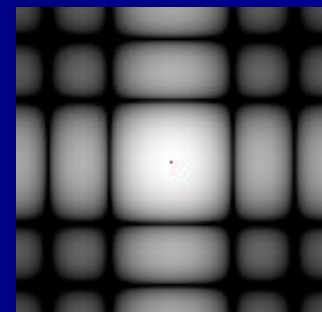
- Константное размытие 3x3

$$Ker[k, p] = \frac{1}{Sum}$$



- Константное размытие 5x5

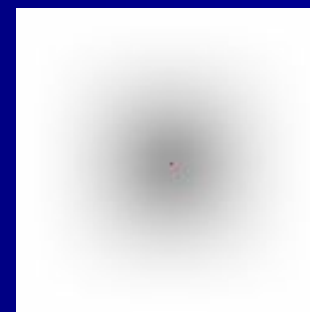
$$Ker[k, p] = \frac{1}{Sum}$$



# Двумерные фильтры

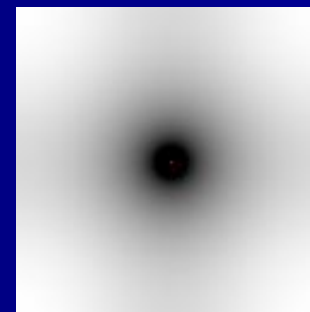
- Повышение четкости

$$\frac{1}{10} \cdot \begin{vmatrix} -1 & -2 & -1 \\ -2 & 22 & -2 \\ -1 & -2 & -1 \end{vmatrix}$$



- Выделение границ

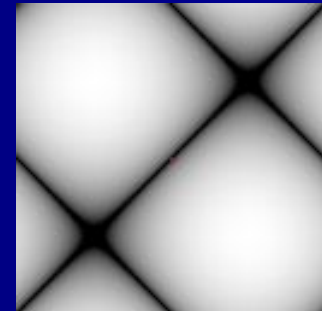
$$\begin{vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{vmatrix}$$



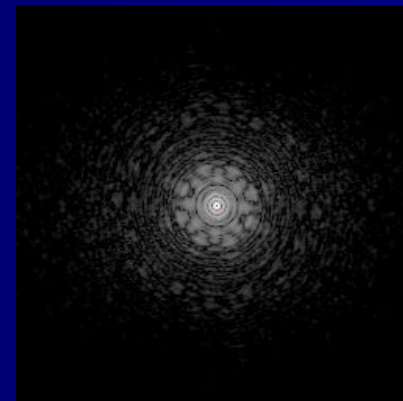
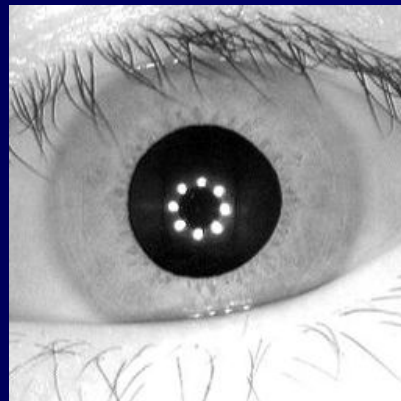
# Двумерные фильтры

- Тиснение

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$



- Пример спектра изображения



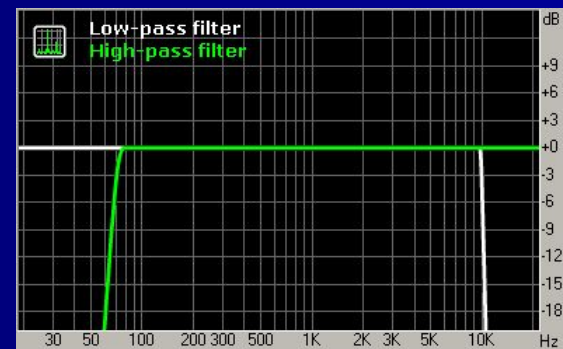
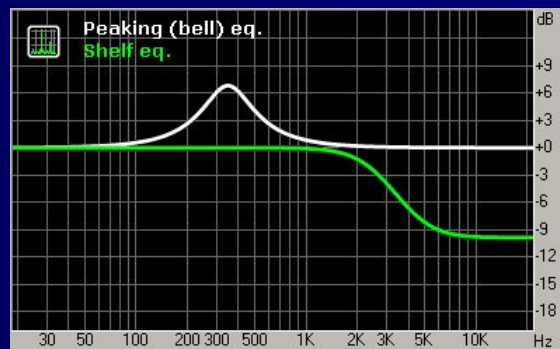
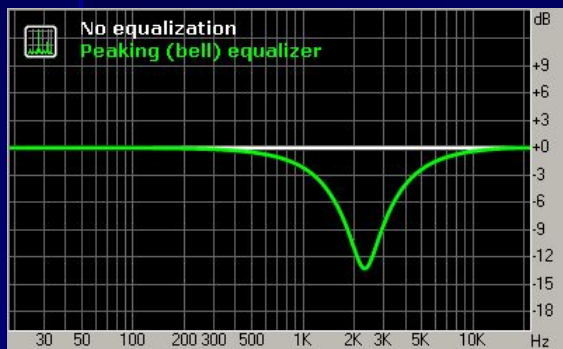
# Эквалайзеры



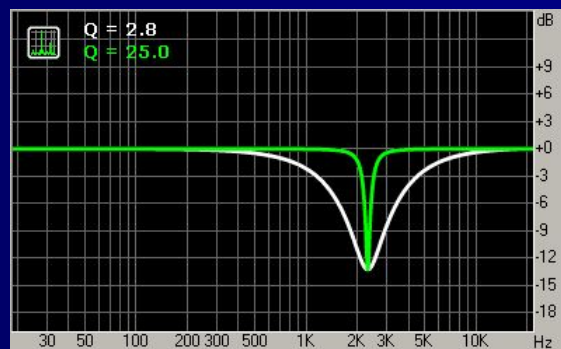
MSU GRAPHICS & MEDIA LAB

- Эквалайзер – устройство коррекции тембра сигнала, изменяющее амплитуды его частотных составляющих
- Изначально применялись для выравнивания АЧХ неидеального звукового тракта
- Вскоре стали использоваться и творчески, для создания нужных тембров или аккуратного совмещения инструментов в фонограмме

- Амплитудно-частотная характеристика (frequency response)



- Добротность ( $Q$ ) определяет ширину полосы воздействия



# Виды эквалайзеров



MSU GRAPHICS & MEDIA LAB

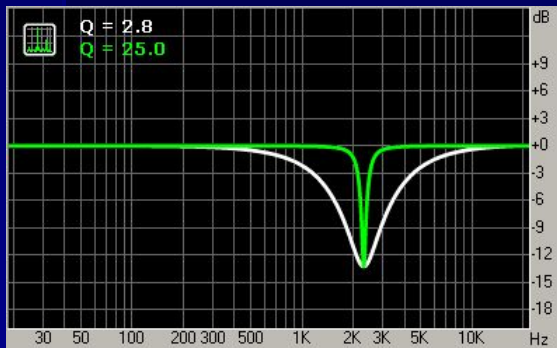
- По управлению АЧХ:
  - ▶ Параметрические: можно выбирать одну из имеющихся форм АЧХ и задавать ее параметры: центральную частоту, коэффициент усиления и добротность
  - ▶ Графические: пользователь «рисует» требуемую АЧХ непосредственно на дисплее или с помощью набора регуляторов усиления на различных частотах
  - ▶ Параграфические: гибрид предыдущих
- По принципу действия:
  - ▶ Аналоговые: состоят из конденсаторов, катушек индуктивности, операционных усилителей
  - ▶ Цифровые: используют FIR или IIR фильтры

# ФЧХ и ГВЗ

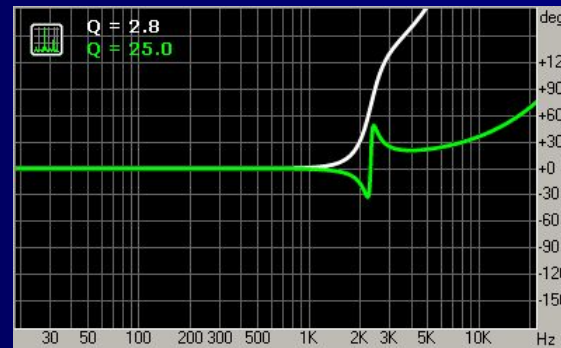


MSU GRAPHICS & MEDIA LAB

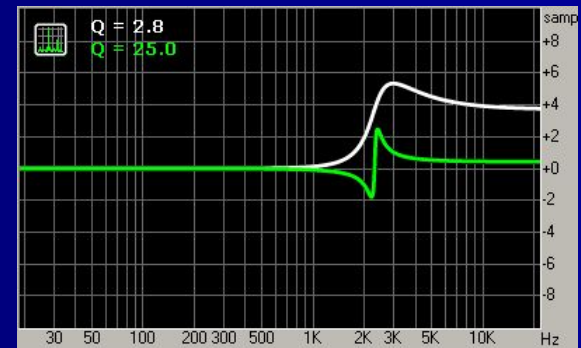
- Фазово-частотная характеристика (*phase response*)
  - ▶ Изменение фазы в зависимости от частоты
- Групповое время задержки (*group delay*)
  - ▶ Задержка различных частотных составляющих



АЧХ



ФЧХ



ГВЗ

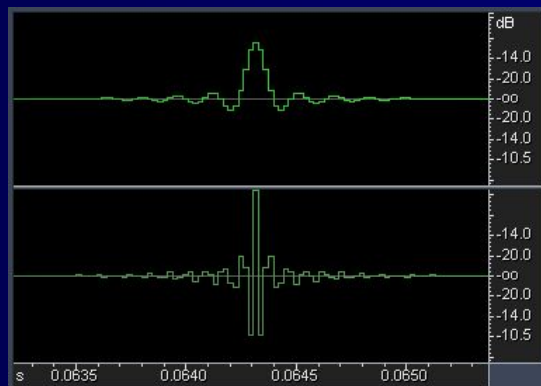
# Импульсная характеристика



- Конечная (FIR)
  - ▶ Легко достичь линейной ФЧХ
- Бесконечная (IIR)
  - ▶ Нелинейная ФЧХ
- Звон фильтров (ringing)

$$y_i = \sum_{-M \leq k \leq M} x_{i-k} h_k$$

$$y_i = \sum_{0 \leq k \leq N} b_k x_{i-k} + \sum_{1 \leq k \leq N} a_k y_{i-k}$$



Линейная ФЧХ



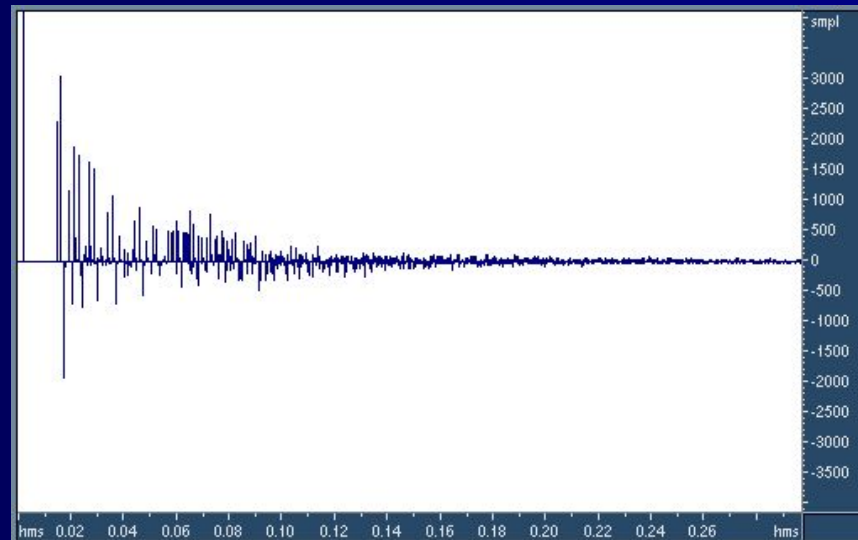
Нелинейная ФЧХ



# Измерение реверберации



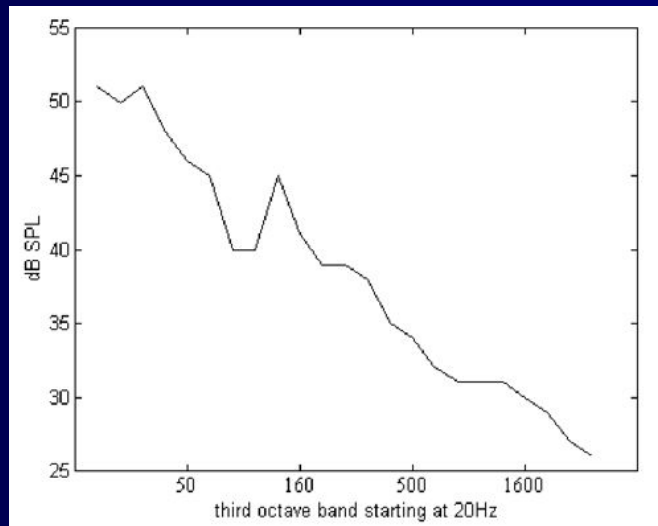
- Импульсный отклик помещения
- Реверберация: прямой звук, ранние и поздние отражения



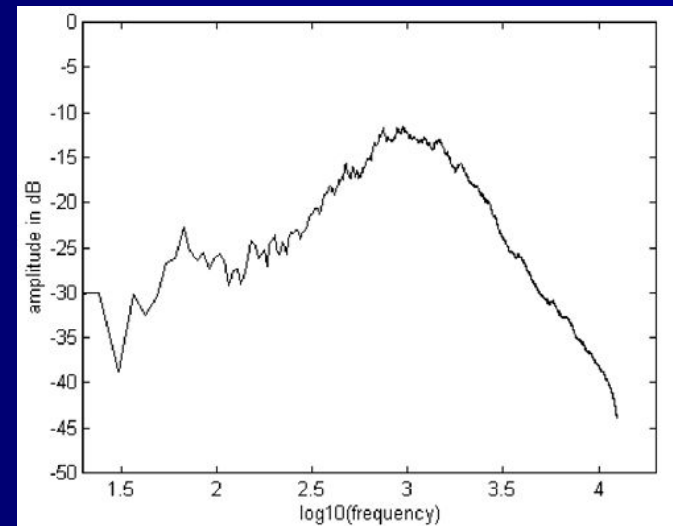
# Измерение реверберации



- Способы измерения импульсного отклика:
  - ▶ Единичный импульс (*плохое соотношение сигнал/шум*)
  - ▶ Звук стартового пистолета / лопающегося воздушного шарика (*неровный спектр, маленькая энергия на НЧ*)



Шум измеряемого зала



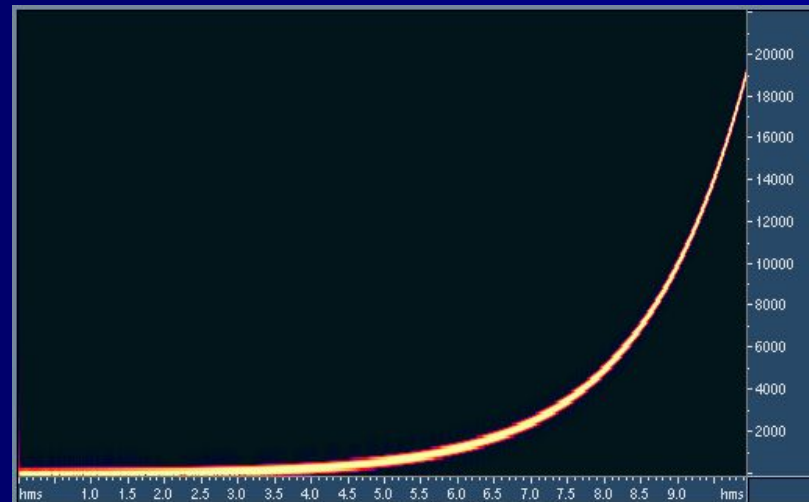
Стартовый пистолет

# Измерение реверберации



- Способы измерения импульсного отклика:
  - ▶ Белый шум, MLS-последовательности (*чувствительны к нелинейным искажениям*)
  - ▶ Скользящий тон (swept sine, «плавающий синус»)

- Деконволюция  
(обратная свертка)



Скользящий тон

# Реверберация



- Что почитать:
  - ▶ Дэвид Гризингер: <http://world.std.com/~griesngr/>
  - ▶ Ирина Алдошина: <http://www.625-net.ru/archive/z0803/aldo.htm>