

# **ОПТИМИЗАЦИЯ ПРОГРАММЫ ДЛЯ ОПРЕДЕЛЕНИЯ ЧАСТОТЫ АУДИОСИГНАЛА**

Ивонин Владимир Алексеевич, студент группы 01321-ДБ

# Введение

**Цель:** изучить и реализовать алгоритм быстрого преобразования Фурье и выяснить, как избежать ошибочного определения фундаментальной частоты.

## Задачи:

- Изучить алгоритм БПФ;
- Реализовать алгоритм БПФ;
- Изучить теорию восприятия высоты тона;
- Применить теорию восприятия высоты тона к предмету исследований.

# Преобразование Фурье

- $$S(k) = \sum_{n=0}^{N-1} s(n) e^{-i\frac{2\pi}{N}nk}, \quad k=0, \dots, N-1, \quad (1)$$

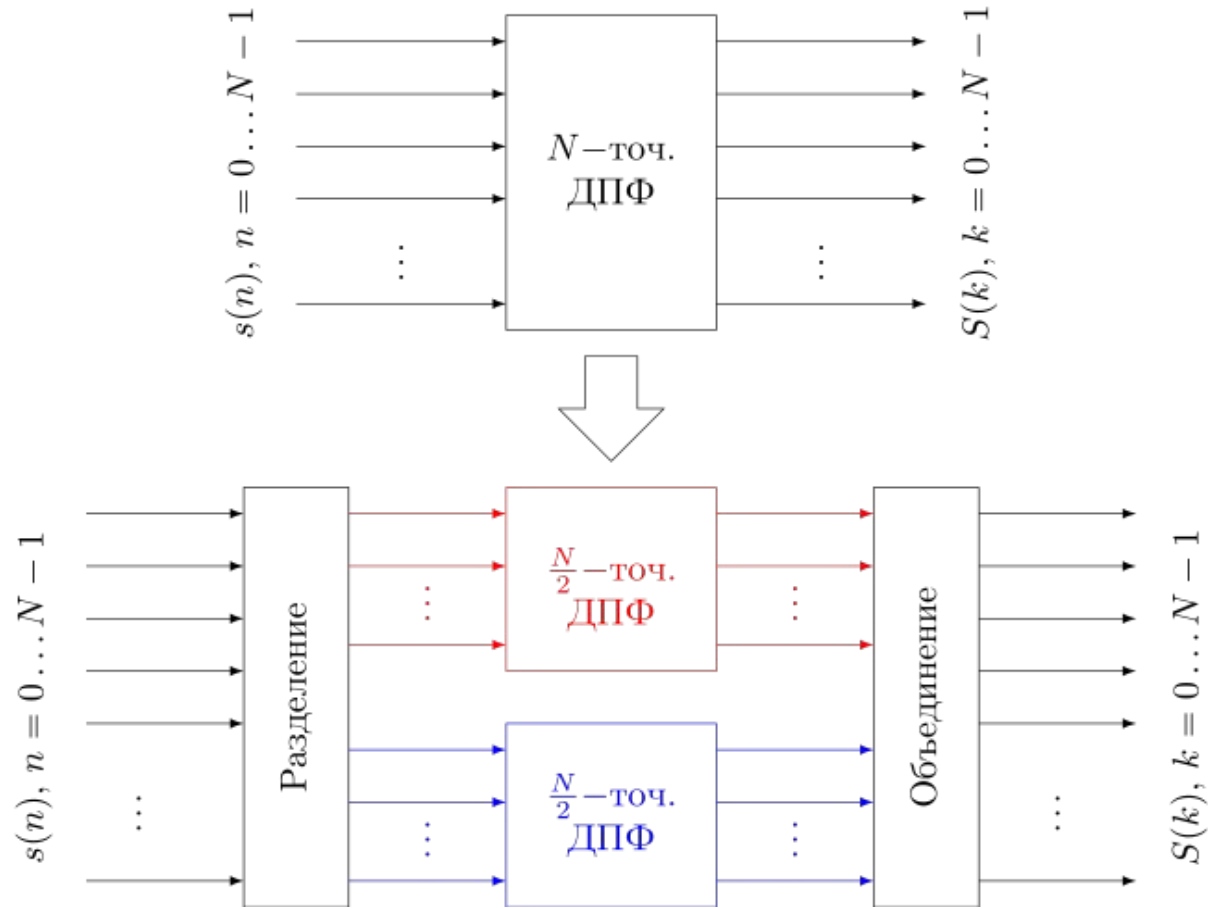
$$s(n) = \frac{1}{N} \sum_{k=0}^{N-1} S(k) e^{i\frac{2\pi}{N}nk}, \quad n=0, \dots, N-1, \quad (2)$$

где  $S(k)$  – отсчёты спектра отсчётов сигнала  $s(n)$ ,  $N$  – количество отсчётов,  $n$  – индекс временных отсчётов сигнала, а  $k$  – индекс спектральных отсчётов ДПФ.

(1) – прямое ДПФ, (2) – обратное ДПФ.

Вычислительная сложность алгоритма ДПФ составляет  $N^2$  операций комплексного умножения и сложения

# Замена $N$ -точечного ДПФ двумя $0,5N$ -точечными ДПФ



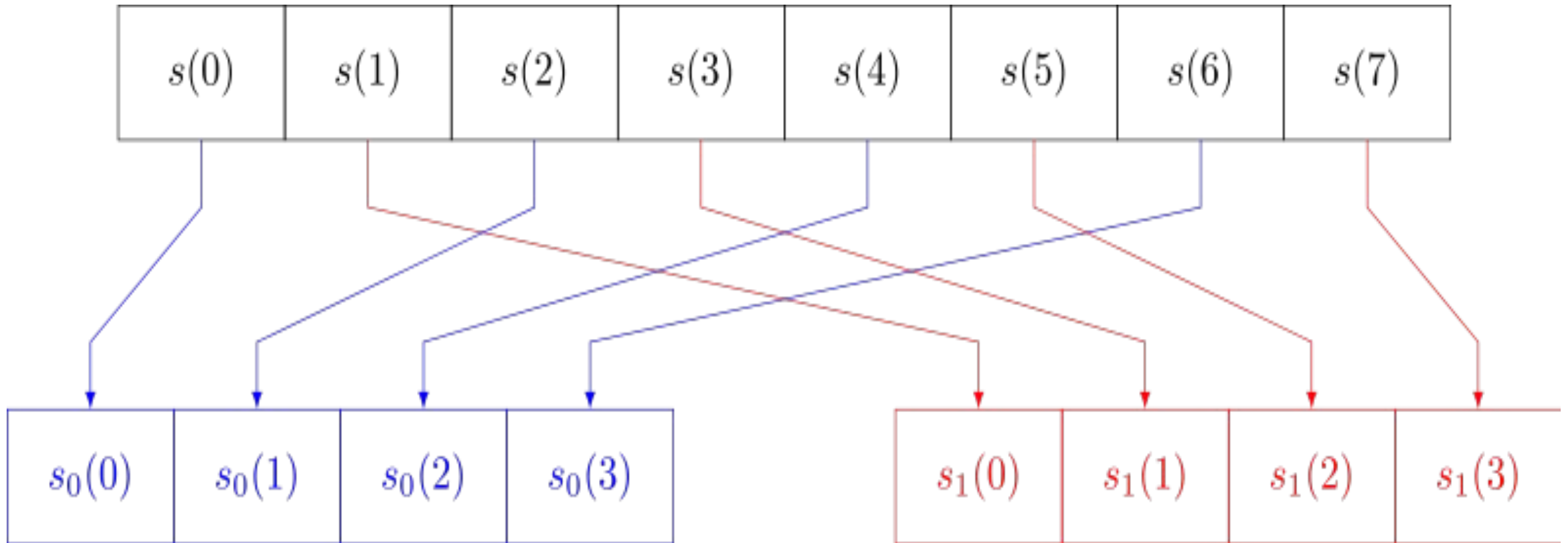
# БПФ

- Алгоритмы БПФ, использующие выборки длиной  $N=2^L$ , называются «алгоритмами БПФ по основанию 2»
- количество операций умножения будет  $\frac{N}{2} \log_2 N = \frac{NL}{2}$  раз.

# Быстрое преобразование Фурье с прореживанием по времени

- Введём обозначение в (1):  $W_N^k = e^{-i\frac{2\pi}{N}k}$ ,  $k=0, \dots, N-1$ , (3)  
где  $W_N^k$  – поворачивающий коэффициент,  $k$  – степень, а не индекс
- Прореживание по времени заключается в разделении исходной последовательности  $s(n)$ ,  $n=0, \dots, N-1$ , на две последовательности половинной длительности  $s_0(m)$  и  $s_1(m)$ ,  $m=0, \dots, 0,5N-1$ , таким образом, что  $s_0(m)=s(2m)$ , а  $s_1(m)=s(2m+1)$ . Последовательность  $s_0(m)$  содержит отсчёты с чётными индексами, а  $s_1(m)$  — с нечётными

# Прореживание по времени для $N=8$



- $$S(k) = S_0(k) + W_N^k S_1(k), \quad k=0, \dots, 0,5N-1, \quad (4)$$

$$S\left(k + \frac{N}{2}\right) = S_0(k) - W_N^k S_1(k), \quad k=0, \dots, 0,5N-1, \quad (5)$$

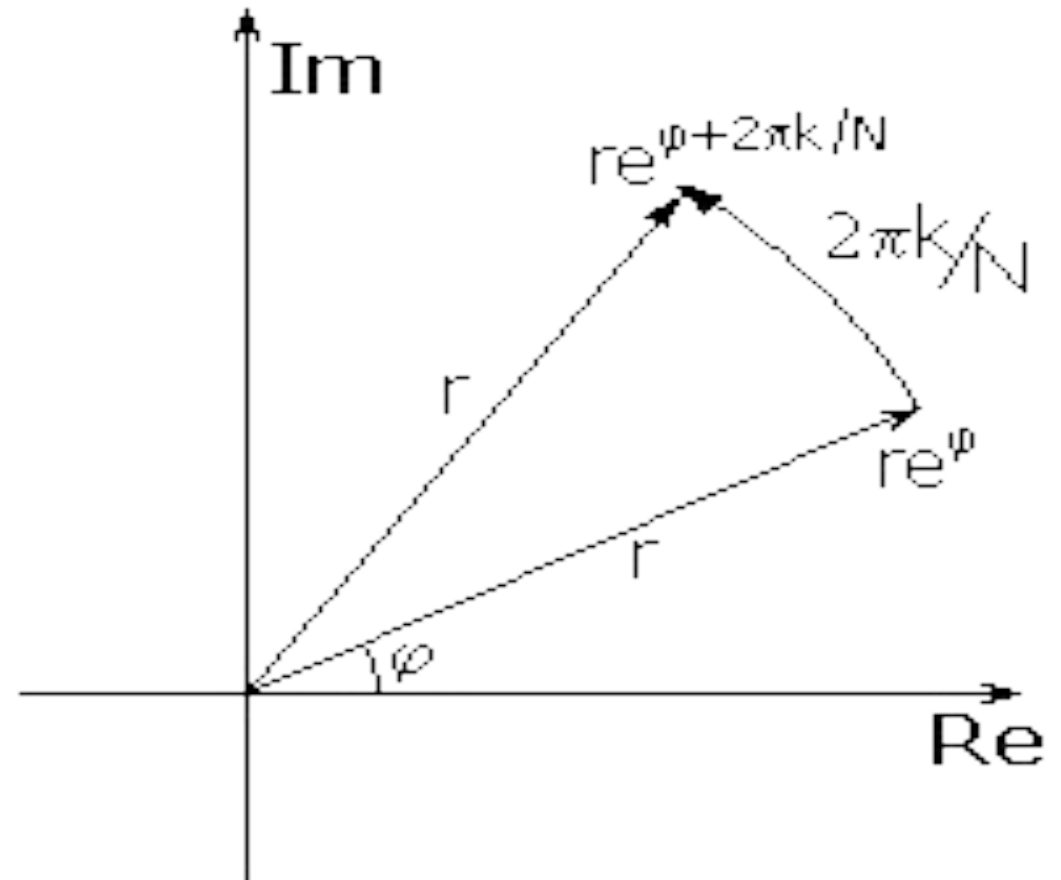
где  $S_0(k) = \sum_{m=0}^{\frac{N}{2}-1} s(2m) W_{\frac{N}{2}}^{mk}$ , а  $S_1(k) = \sum_{m=0}^{\frac{N}{2}-1} s(2m+1) W_{\frac{N}{2}}^{mk}$  –  $0,5N$ -точечные ДПФ прореженных последовательностей  $s_0(m)$  и  $s_1(m)$ ,  $m=0, \dots, 0,5N-1$ , соответственно.



- Таким образом, прореживание по времени можно считать алгоритмом разделения последовательности на две последовательности половинной длительности. Первая половина ДПФ – сумма ДПФ  $S_0(k)$  «чётной» последовательности  $s_0(m)$  и ДПФ  $S_1(k)$  «нечётной» последовательности  $s_1(m)$ , умноженного на поворотные коэффициенты  $W_N^k$ .

# Этимология названия поворачивающих коэффициентов

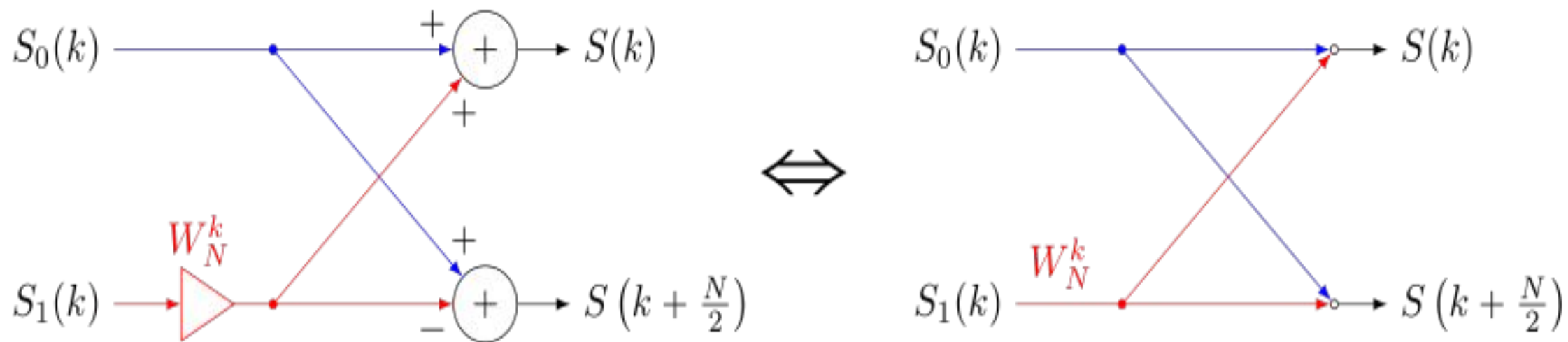
- $z_1 z_2 = |z_1| e^{i\varphi_1} |z_2| e^{i\varphi_2} = (|z_1| |z_2|) e^{i(\varphi_1 + \varphi_2)}$



# Свойства поворачивающих коэффициентов

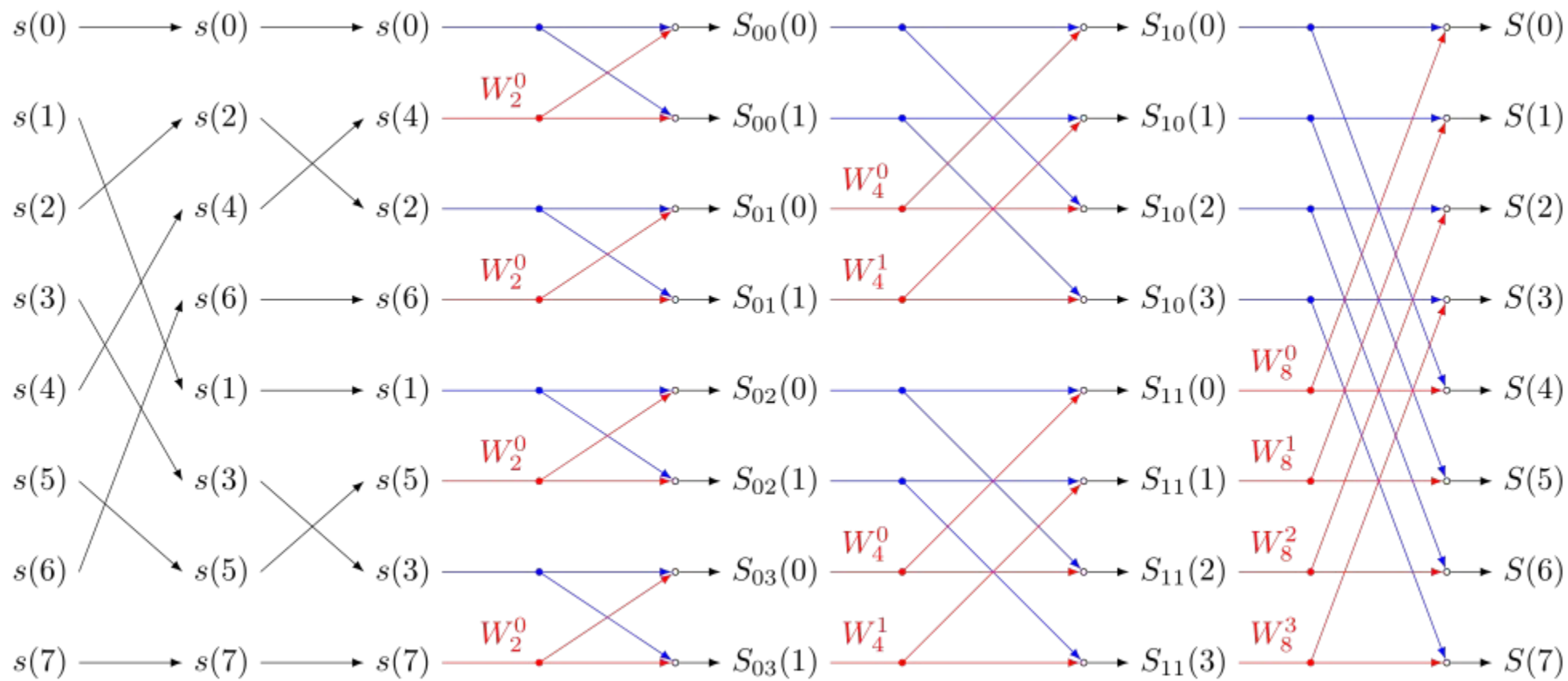
- $W_N^{(k+lN)(n+mN)} = W_N^k$
- $W_N^k = -W_N^{k-\frac{N}{2}}$
- $W_N = W_{NL}^L$

# Граф «бабочка»



- Такую процедуру объединения можно применить для расчёта каждого из  $0,5N$ -точечных ДПФ  $S_0(k)$  и  $S_1(k)$  через два  $0,25N$ -точечных ДПФ. Тогда для  $N=2^L$  можно произвести  $L-1$  этап деления последовательности на «чётную» и «нечётную» и после этого производить объединение спектра за  $L$  этапов. В результате получается полный граф алгоритма БПФ.

# Полный граф алгоритма БПФ по основанию 2 с прореживанием по времени для $N=8$



- На первом этапе отсчёты входного сигнала переставляются местами, и исходная последовательность делится на «чётную» и «нечётную» последовательности. Потом «чётная» и «нечётная» последовательности в свою очередь делятся на «чётную» и «нечётную» последовательности. Данная процедура называется двоично-инверсной перестановкой, так как можно выполнить перенумерацию отсчётов, переписав номер отсчёта в двоичной системе счисления в обратном направлении.

# Основные идеи алгоритма БПФ

- Разделить сумму (1) из  $N$  слагаемых на две суммы по  $N/2$  слагаемых и вычислить их по отдельности. Для вычисления каждой из подсумм нужно их тоже разделить на две и т.д.
- Повторно использовать уже вычисленные слагаемые



# Пример деления последовательности $s(n)$ для $N=16$

$x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8 x_9 x_{10} x_{11} x_{12} x_{13} x_{14} x_{15}$   
 $x_0 x_2 x_4 x_6 x_8 x_{10} x_{12} x_{14} | x_1 x_3 x_5 x_7 x_9 x_{11} x_{13} x_{15}$   
 $x_0 x_4 x_8 x_{12} | x_2 x_6 x_{10} x_{14} | x_1 x_5 x_9 x_{13} | x_3 x_7 x_{11} x_{15}$   
 $x_0 x_8 | x_4 x_{12} | x_2 x_{10} | x_6 x_{14} | x_1 x_9 | x_5 x_{13} | x_3 x_{11} | x_7 x_{15}$

До всех перестановок		После всех перестановок	
Десятичная с/с	Двоичная с/с	Десятичная с/с	Двоичная с/с
0	0000	0	0000
8	1000	1	0001
4	0100	2	0010
12	1100	3	0011

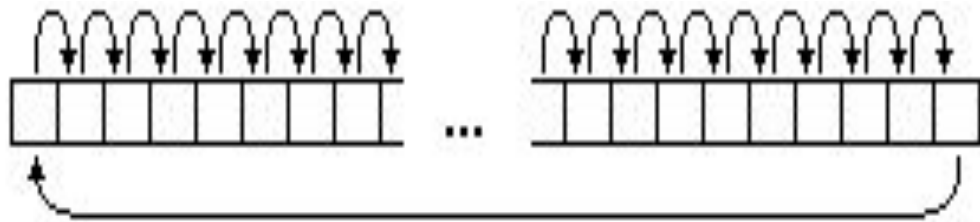
- На первом шаге чётные элементы с номером  $n$  переместились в позицию  $n/2$ , а нечётные – в позицию  $N/2+(n-1)/2$ , где  $n=0, 1, \dots, N-1$ . Таким образом, новая позиция вычисляется из старой позиции с помощью функции

$$\text{ror}(n, N) = \left[ \frac{n}{2} \right] + N\{n/2\},$$

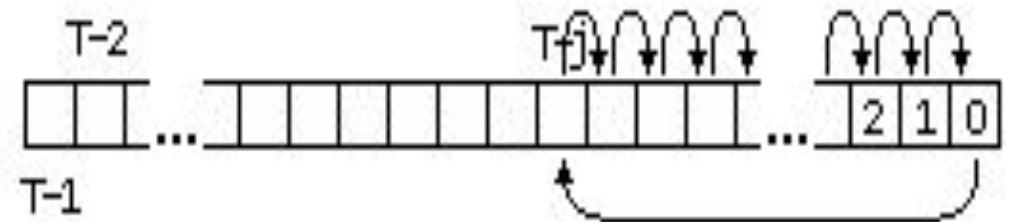
где  $[n/2]$  – целая часть от деления  $n$  на 2, а  $\{n/2\}$  – дробная.

# Циклический сдвиг вправо

Циклический сдвиг вправо при одном разбиении



Циклический сдвиг вправо при многократных разбиениях



# Оптимизация вычисления поворачивающих коэффициентов

Формула 
$$W_N^{k+1} = W_N W_N^k, \quad (6)$$

позволяет заменить операцию возведения в степень операцией умножения, если заранее вычислить  $W_N$  для  $N=2, 4, 8, \dots$

- $W_N$  не целое, т.к.  $W_N = e^{\frac{-i2\pi}{N}} = \cos \frac{2\pi}{N} - i \sin \frac{2\pi}{N}$ , но можно избежать множества обращений к относительно медленным функциям  $\sin(x)$  и  $\cos(x)$ :

$$W_N^{13} = \left( (W_N^2)^2 \right)^2 (W_N^2)^2 W_N$$

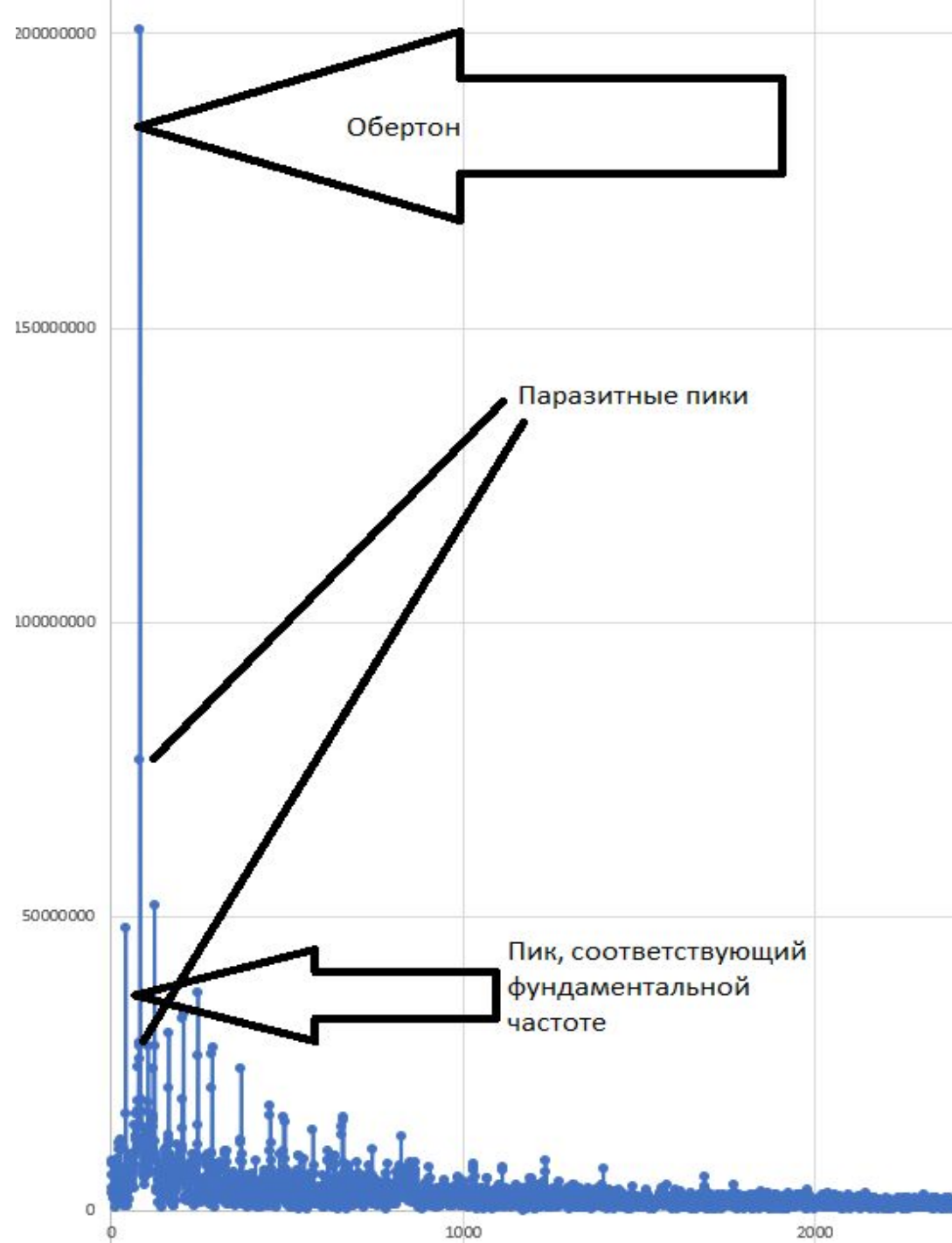
- В приведённом примере понадобилось всего 5 умножений, вместо 13. Для возведения в степень  $0,5N-1$  нужно  $\log_2 N$  умножений, вместо  $0,5N$ , что позволяет значительно уменьшить погрешность при вычислении поворачивающих коэффициентов.
- Можно вдвое сократить количество умножений на каждом шаге, если использовать уже рассчитанные результаты:  $W_N, W_N^2, W_N^4, W_N^8, \dots, W_N^{\frac{N}{2}}$ , для хранения которых нужно дополнительно использовать  $\log_2 N \max$  комплексных ячеек памяти

- $$W_N^{\frac{N}{2}-1} = W_N^{\frac{N}{4}} \cdot W_N^{\frac{N}{8}} \cdot \dots \cdot W_N^8 \cdot W_N^4 \cdot W_N^2 \cdot W_N$$
- Если какое-либо  $W_N^k$  не было вычислено предварительно, то анализируется двоичный вид числа  $k$ . Каждому единичному биту соответствует ровно один коэффициент. В общем случае единице в бите с номером  $b$  (младший бит имеет номер 0) соответствует коэффициент  $W_N^{2^b}$ , хранящийся в  $b$ -й ячейке упомянутого выше массива.

- Можно уменьшить количество умножений для вычисления  $W_N^k$  до одного на два цикла. Для этого нужно отвести  $0,5N$  комплексных ячеек для хранения всех предыдущих  $W_N^k$ . Алгоритм таков: нечётные элементы вычисляются по формуле (6), чётные – по формуле (7):

$$W_N^{2m} = W_{\frac{N}{2}}^m. \quad (7)$$

- То есть ничего не вычисляется, а используется одно из значений, вычисленных на предыдущем шаге, когда  $N$  было вдвое меньше.





# Основные задачи психоакустики

- понять, как слуховая система расшифровывает звуковой образ
- установить основные соответствия между физическими стимулами и слуховыми ощущениями
- выявить, какие именно параметры звукового сигнала являются наиболее значимыми для передачи семантической (смысловой) и эстетической (эмоциональной) информации.

# Периферическая часть слухового анализатора

- акустическая антенна, принимающая, локализирующая, фокусирующая и усиливающая звуковой сигнал;
- микрофон;
- частотный и временной анализатор;
- аналого-цифровой преобразователь, преобразующий аналоговый сигнал в двоичные нервные импульсы – электрические разряды.

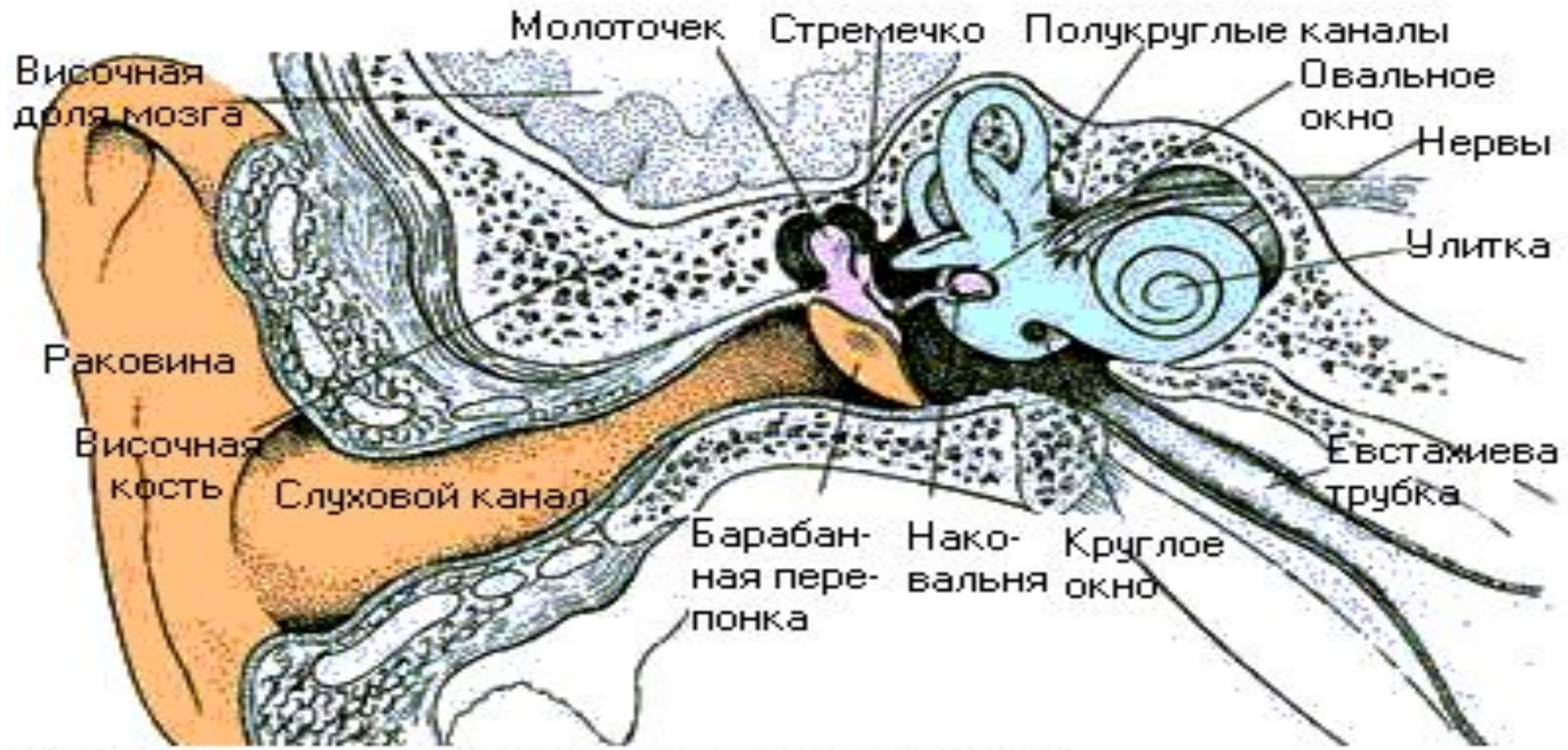
# Основные функции внешних ушей

- бинауральное (пространственное) восприятие
- локализация звукового источника
- усиление звуковой энергии

# Строение внешних ушей

- Уши и голова – компоненты внешней акустической антенны
- Слуховой канал представляет собой изогнутую цилиндрическую трубку длиной 22,5 мм, имеющую первую резонансную частоту порядка 2,6 кГц, поэтому в этой области частот он существенно усиливает звуковой сигнал
- Барабанная перепонка – тонкая плёнка толщиной 74 мкм, имеет вид конуса, обращённого остриём в сторону среднего уха. На низких частотах она движется, как поршень, на более высоких – на ней образуется сложная система узловых линий

# Структура периферической слуховой системы

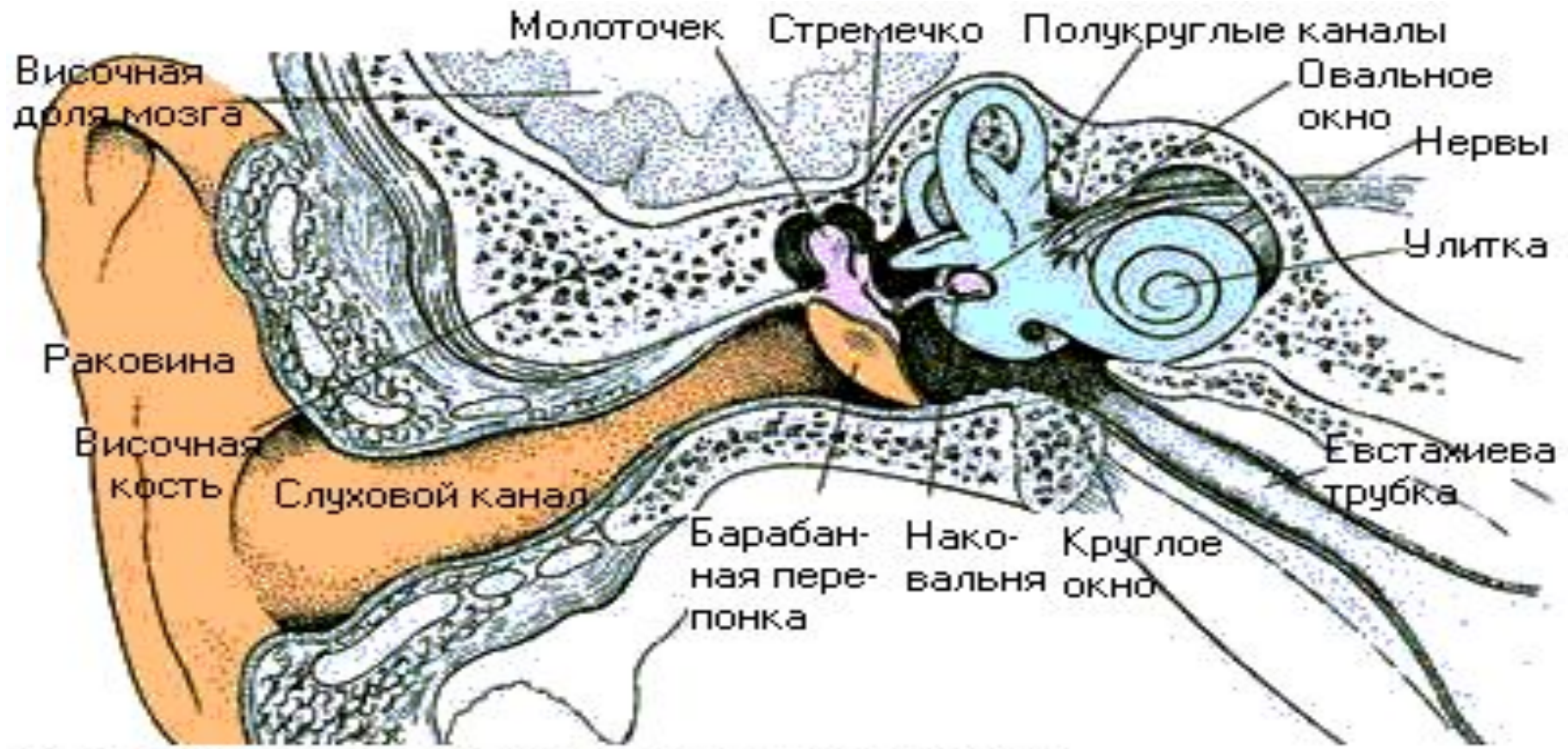


# Строение среднего уха

Это заполненная воздухом полость, соединённая с носоглоткой евстахиевой трубой для выравнивания атмосферного давления. Там находятся три маленькие слуховые косточки:

- молоточек
- наковальня
- стремечко

# Структура периферической слуховой системы



# Основные функции среднего уха

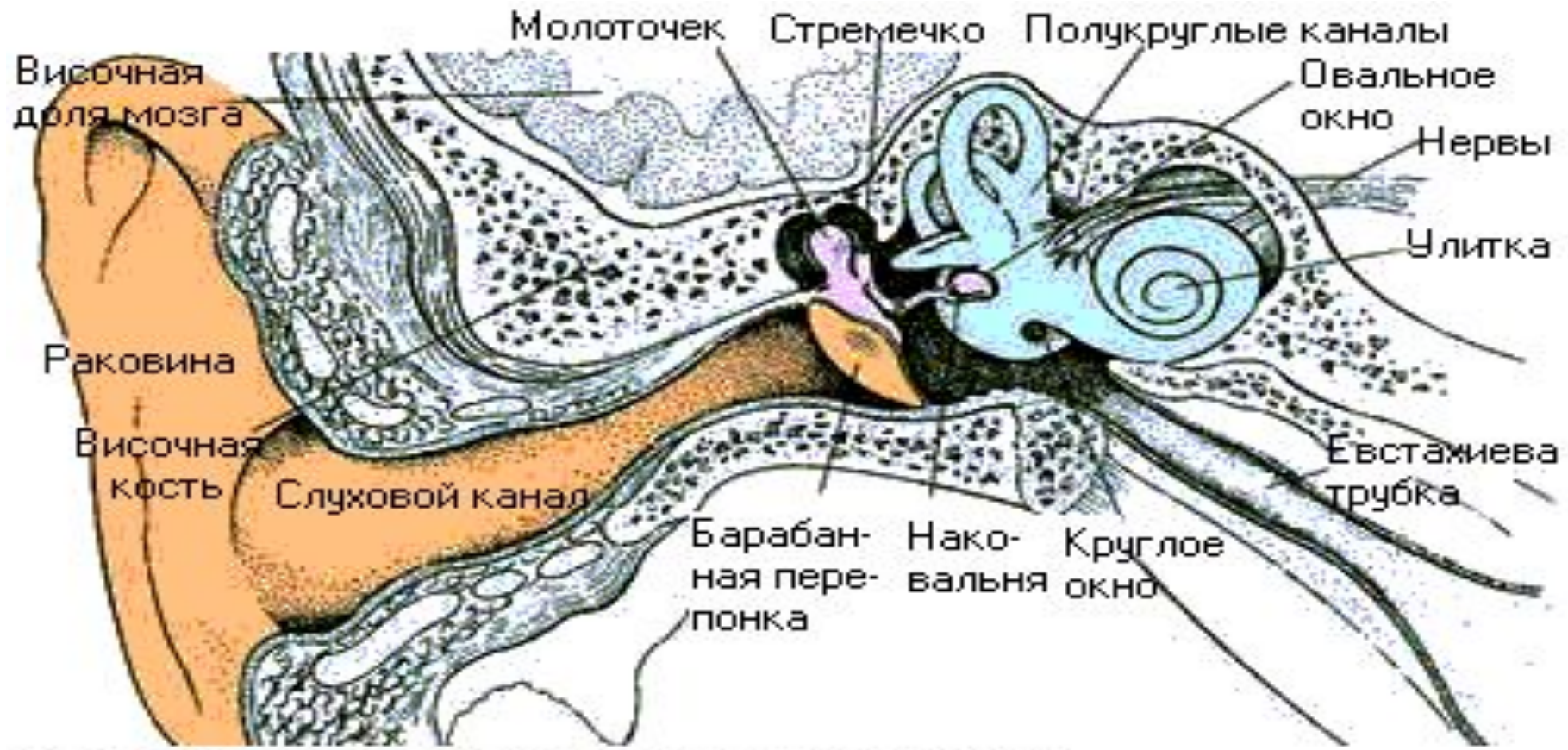
- согласование импеданса воздушной среды с жидкой средой улитки внутреннего уха
- защита от громких звуков (акустический рефлекс)
- усиление (рычаговый механизм), за счёт которого звуковое давление, передаваемое во внутреннее ухо, усиливается почти на 38 дБ по сравнению с тем, которое попадает на барабанную перепонку.



# Строение внутреннего уха

- орган равновесия (вестибулярный аппарат)
- улитка

# Структура периферической слуховой системы



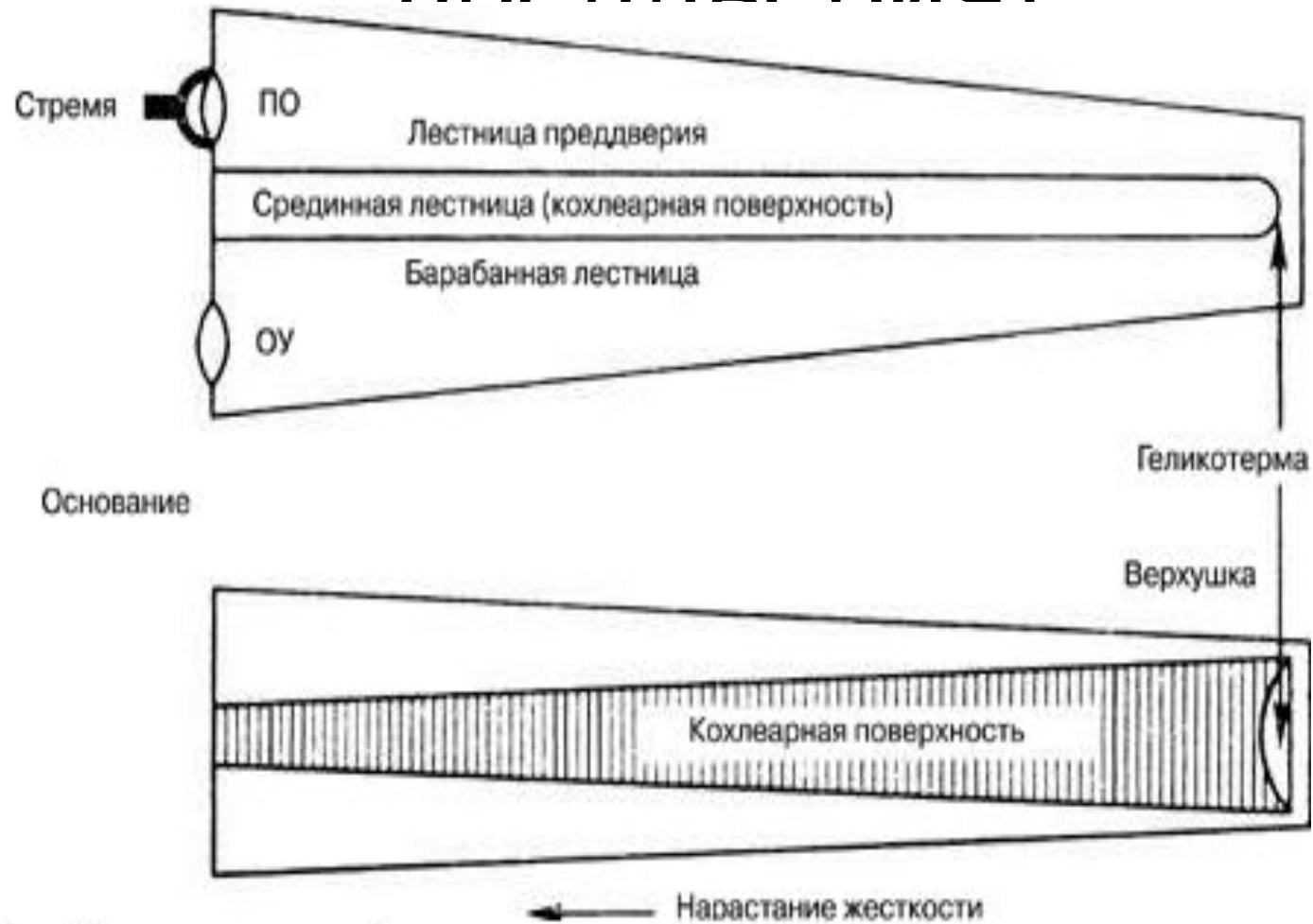
# Строение улитки

Это трубка переменного сечения, свёрнутая три раза, внутри по всей длине она разделена двумя мембранами на три полости:

- лестница преддверия
- срединная полость
- барабанная лестница

Сверху срединная полость закрыта мембраной Рейсснера, снизу – базилярной мембраной. Все полости заполнены жидкостью. Верхняя и нижняя полости соединены через отверстие у вершины улитки (геликотрему). В верхней полости находится овальное окно, через которое стремечко передаёт колебания во внутреннее ухо, в нижней полости находится круглое окно, выходящее обратно в среднее ухо.

# Схематическое изображение развёрнутой улитки (оп – окно преддверия)



# Базиллярная мембрана

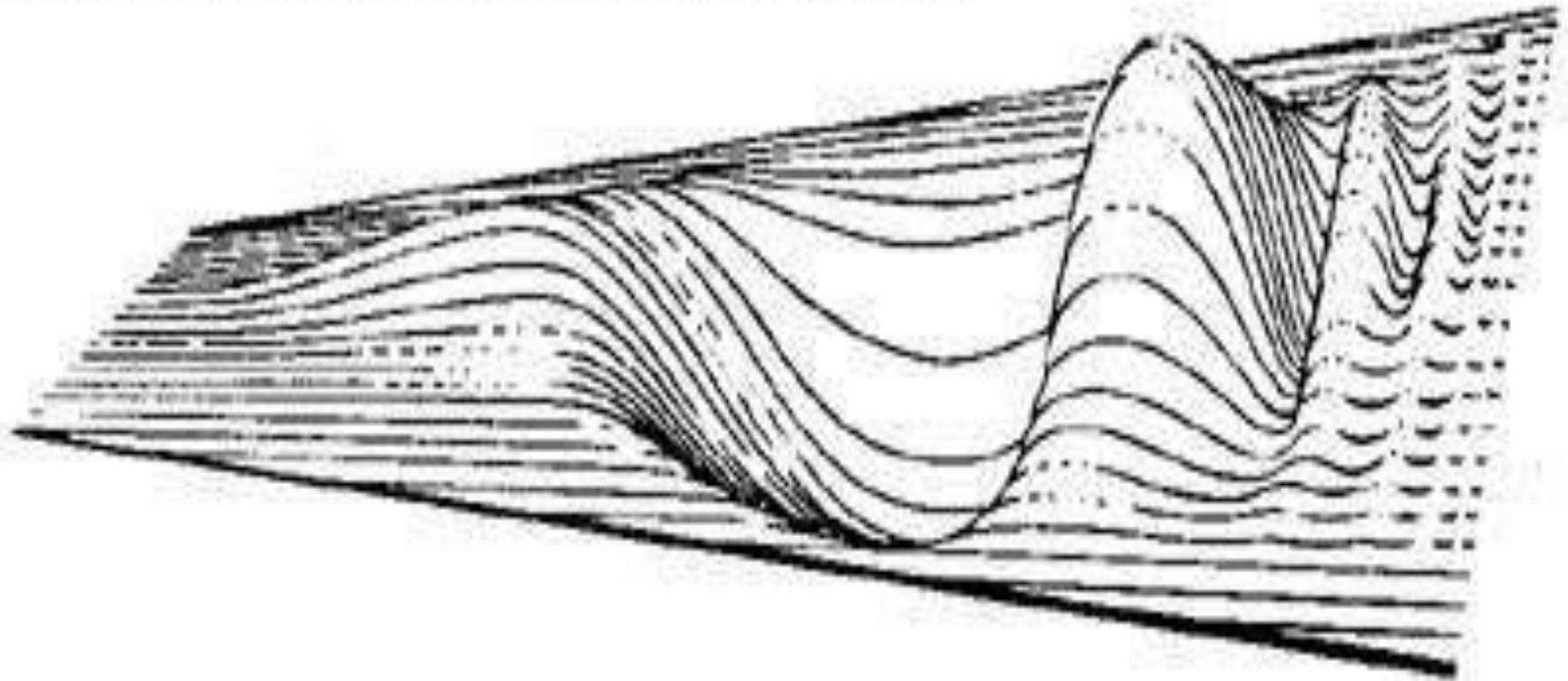
- Состоит из нескольких тысяч поперечных волокон: длина – 32 мм, ширина у стремечка – 0,05 мм (этот конец узкий, лёгкий и жесткий), у геликотремы ширина – 0,5 мм (этот конец толще и мягче). На внутренней стороне базилярной мембраны находится орган Корти, а в нём – специализированные слуховые рецепторы – волосковые клетки. В поперечном направлении орган Корти состоит из одного ряда внутренних волосковых клеток и трёх рядов наружных волосковых клеток. Между ними образуется тоннель. Волокна слухового нерва пересекают тоннель и контактируют с волосковыми клетками.

- Слуховой нерв представляет собой перекрученный ствол, сердцевина которого состоит из волокон, отходящих от верхушки улитки, а наружные слои – от нижних её участков. Войдя в ствол мозга, нейроны взаимодействуют с клетками различных уровней, поднимаясь к коре и перекрещиваясь по пути так, что слуховая информация от левого уха поступает в основном в правое полушарие, где происходит, главным образом, обработка эмоциональной информации, а от правого уха – в левое полушарие, где в основном обрабатывается смысловая информация. В коре основные зоны слуха находятся в височной области, между обоими полушариями имеется постоянное взаимодействие.

# Упрощённый механизм передачи звука

1. Звуковые волны проходят звуковой канал и возбуждают колебания барабанной перепонки
2. Эти колебания через систему косточек среднего уха передаются овальному окну, толкающему жидкость в верхнем отделе улитки, в ней возникает импульс давления, заставляющий жидкость переливаться из верхней половины в нижнюю через барабанную лестницу и геликотрему, и оказывает давление на перепонку круглого окна, вызывая при этом его смещение в сторону, противоположную движению стремечка
3. Движение жидкости вызывает колебания базилярной мембраны (бегущая волна).
4. Преобразование механических колебаний мембраны в дискретные электрические импульсы нервных волокон происходят в органе Корти. Когда базилярная мембрана вибрирует, реснички на волосковых клетках изгибаются, и это генерирует электрический потенциал, что вызывает поток электрических нервных импульсов, несущих всю необходимую информацию о поступившем звуковом сигнале в мозг для дальнейшей переработки и реагирования.

# Бегущая волна на базилярной мембране

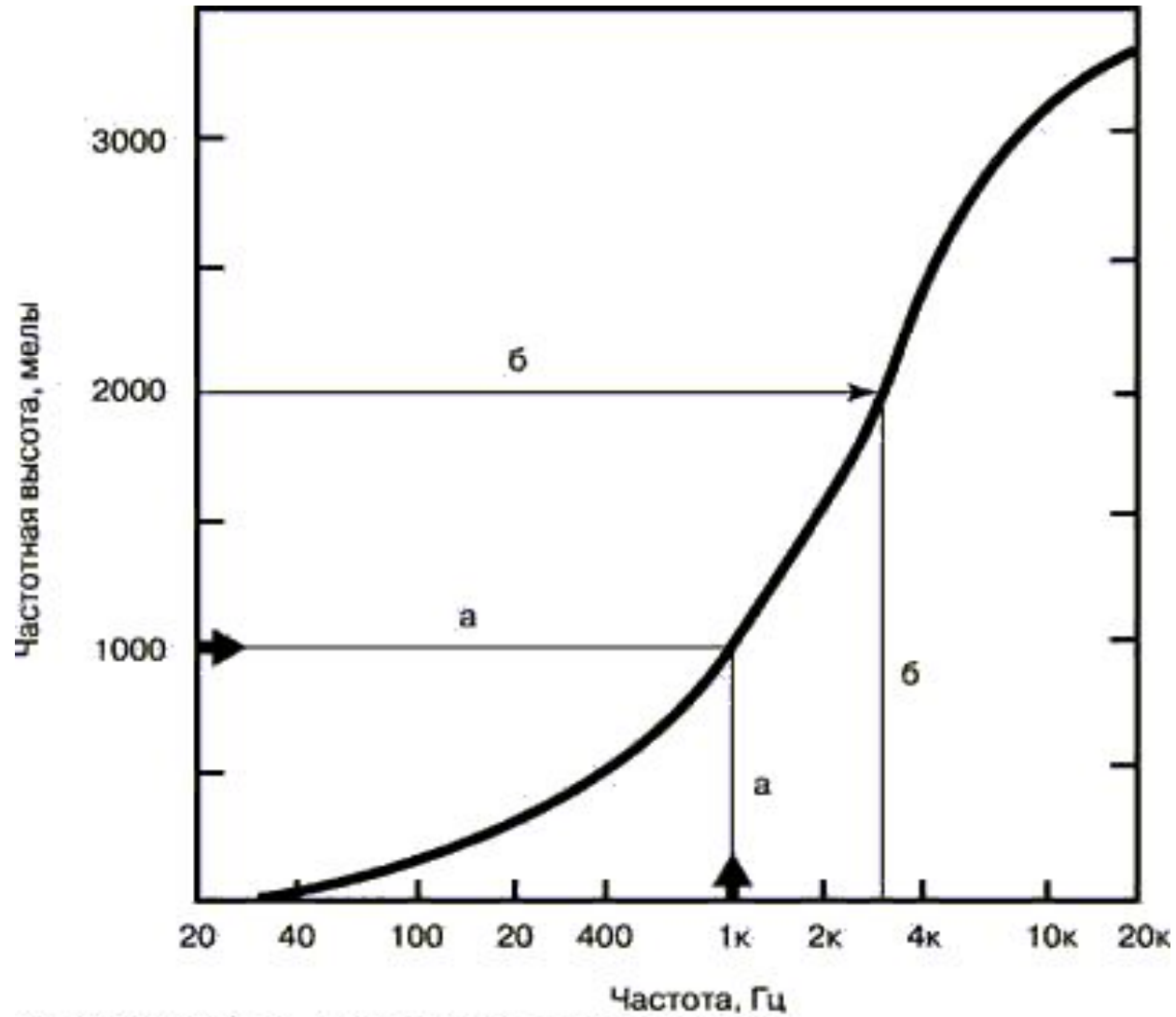




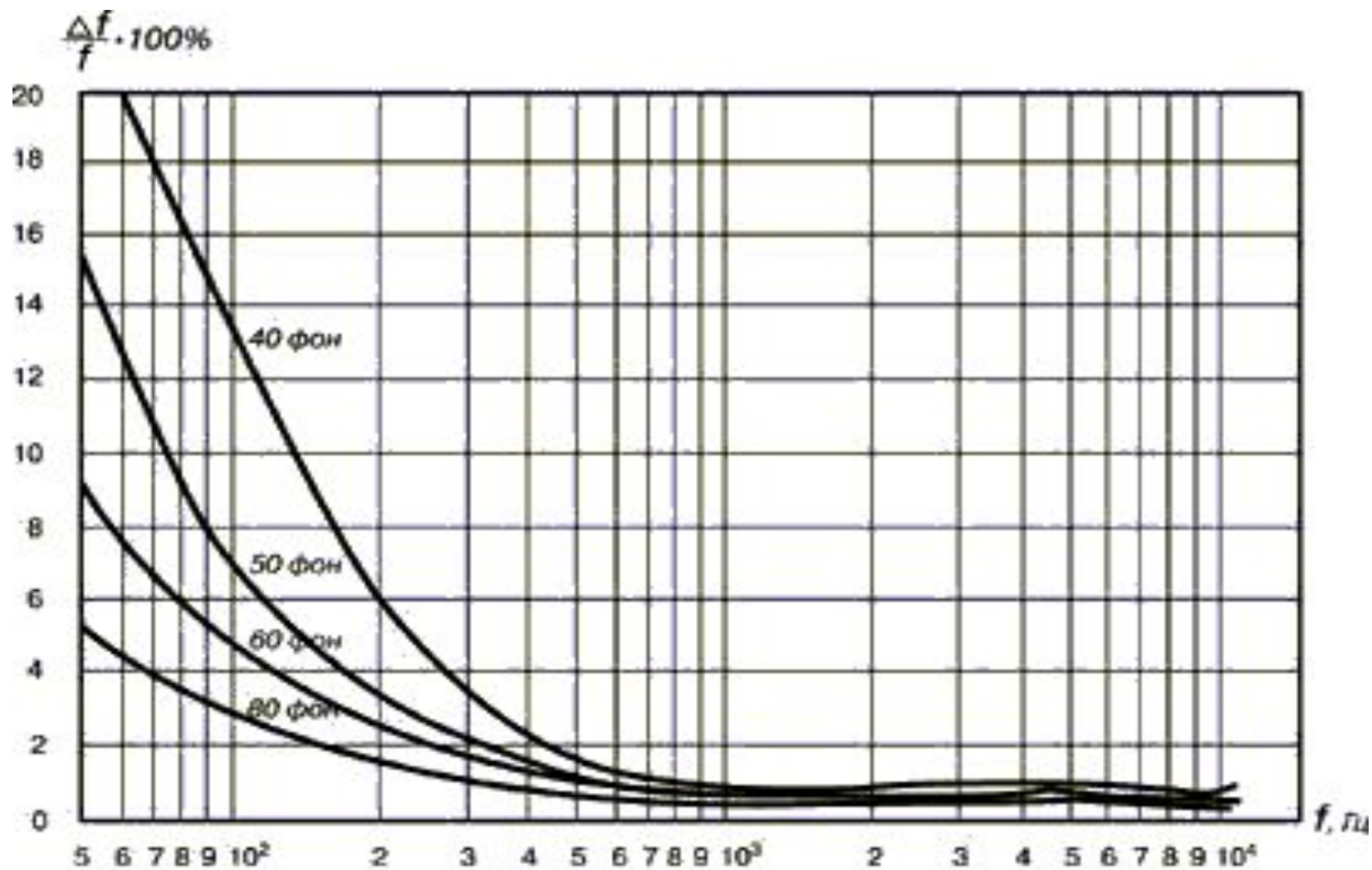
- 
- Слуховая система способна различать высоту звука только у периодических сигналов. Если это простое гармоническое колебание, например, синусоидальный сигнал от генератора, то период колебаний  $T$  определяет частоту  $f_0=1/T$ , поэтому определяющим параметром для различения высоты является частота сигнала
- Если это сложный звук, то высоту слуховая система может присвоить по его основному тону, но только если он имеет периодическую структуру, т.е. спектр его состоит из гармоник – обертонов – с частотами

$$f_n = (n + 1)f_0$$

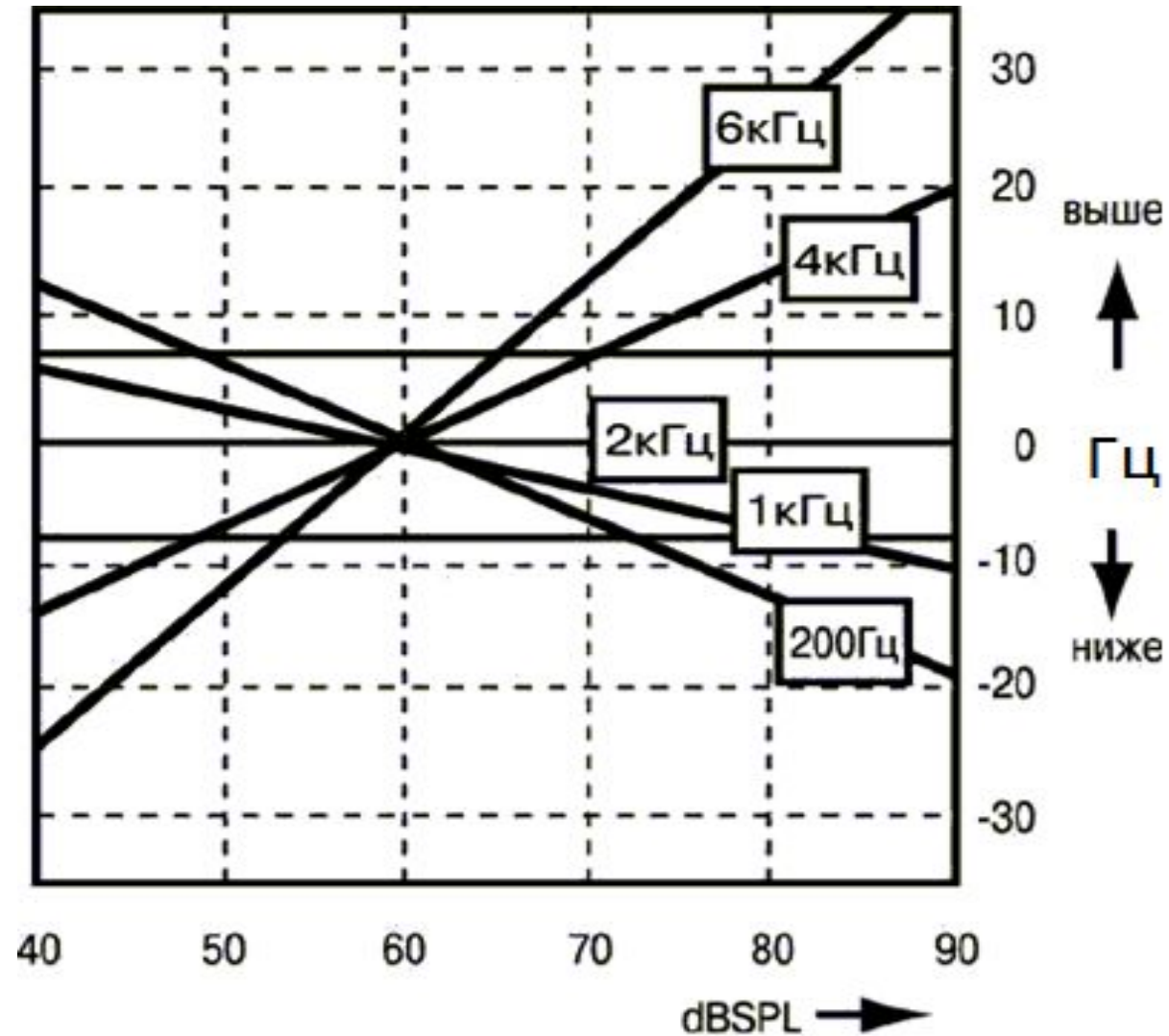
# Зависимость высоты тона от частоты звука



# Кривые частотной разрешающей способности слуха



# Зависимость высоты звука от его ИНТЕНСИВНОСТИ



- Ощущение высоты тона зависит и от его длительности: короткие звуки воспринимаются как сухой щелчок, но при удлинении звука щелчок начинает давать ощущение высоты тона. Время, требуемое для перехода от щелчка к тону, зависит от частоты: для низких частот требуется для распознавания высоты тона примерно 60 мс, для частот от 1 до 2 кГц – 15 мс. Для сложных звуков это время увеличивается, для звуков речи оно может составлять 20-30 мс.

- В музыке простые синусоидальные тоны практически не используются, каждый музыкальный тон имеет сложную структуру и состоит из основного тона и гармоник. Однако можно установить соответствие по высоте музыкального тона, например, ноты ля первой октавы и чистого синусоидального сигнала с частотой 440 Гц. Высоты этих двух звуков будут одинаковыми, но тембры – разными. Это свидетельствует о том, что для сложных периодических сигналов высота присваивается по частоте основного тона – именно он имеет частоту 440 Гц.

- До частоты примерно 5000 Гц увеличение высоты тона на октаву связано с удвоением частоты. Но выше частоты 5000 Гц это соответствие нарушается: чтобы получить ощущение увеличения высоты на октаву, надо увеличить соотношение частот почти в 10 раз. Это дало основание некоторым ученым предложить две размерности высоты тона: психофизическую – в мелах, пропорциональную в некоторых пределах логарифму частоты, установленную для чистых тонов, и музыкальную, соответствующую названию нот, которая может быть определена примерно до 5000 Гц.

# Теория места

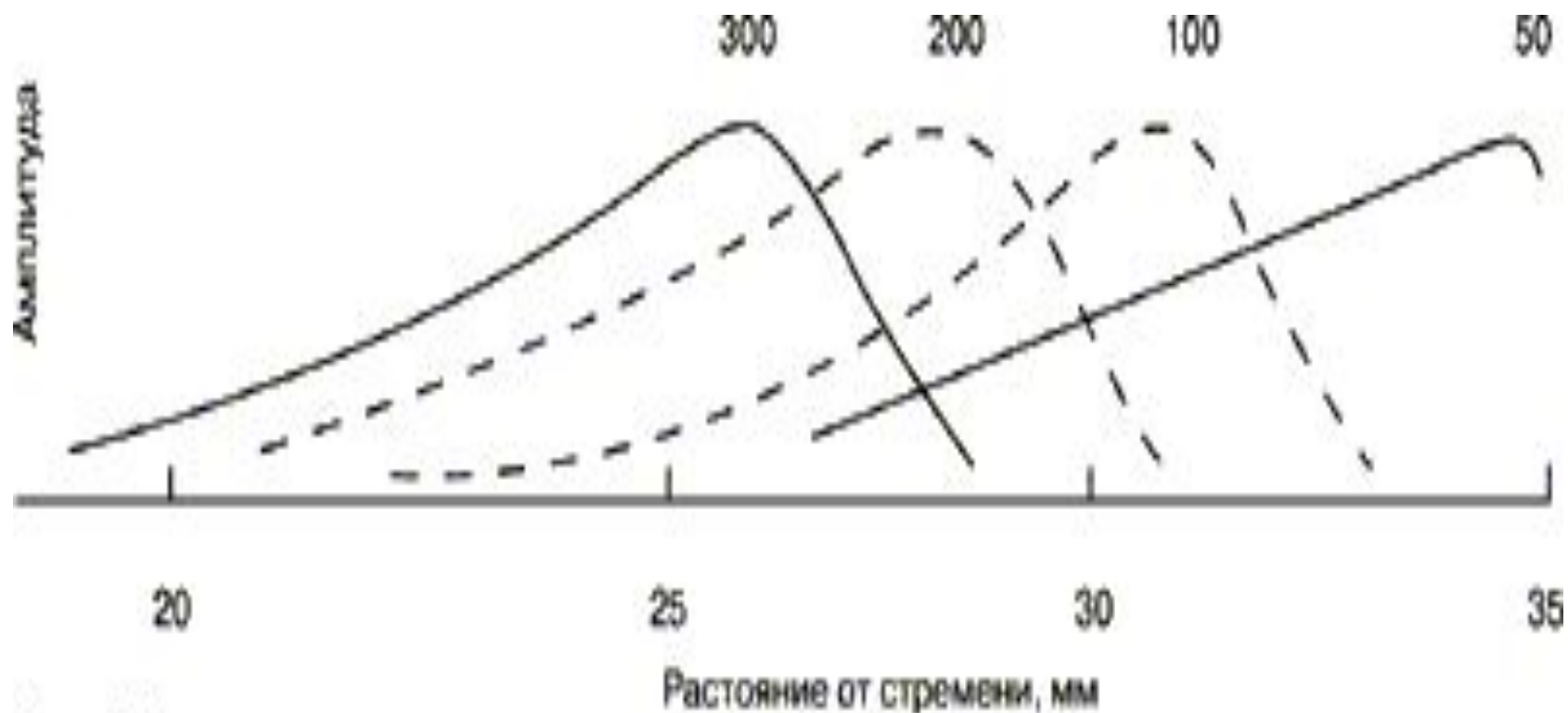
- Теория места при восприятии высоты основана на способности базилярной мембраны выполнять частотный анализ сложного звука, т.е. действовать как спектральный анализатор
- Базилярная мембрана организована тонотопически, т.е. каждый тон имеет свою топографию размещения
- Максимум смещения бегущей волны располагается в разных местах базилярной мембраны: низкие частоты имеют максимум смещения вблизи вершины мембраны, высокие – вблизи овального окна. Каждая частота имеет своё место максимума возбуждения на мембране



- В зависимости от спектрального состава на базилярной мембране возбуждаются различные участки. Возбуждаются волосковые клетки, находящиеся на этом месте, и их электрическая активность сообщает мозгу, какие частоты присутствуют в спектре
- Таким образом, частота тона представлена в коде, основанном на том, нейроны каких участков активны, а каких – нет

- При действии синусоидального сигнала в слуховом нерве формируется «образец возбуждения» – скорость разрядов нейронов как функция места на базилярной мембране. При этом пик этого образца движется вдоль мембраны при изменении частоты. Для того чтобы слух различил два тона по высоте, необходимо, чтобы на базилярной мембране максимум смещения, соответствующий данным частотам, сместился всего на 52 мкм
- Таким образом, можно считать, что периферическая слуховая система содержит банк полосовых фильтров («слуховых фильтров») с перекрывающимися полосами. Их ширина свыше 1 кГц составляет примерно 10-17% от центральной частоты

# «Слуховые фильтры»



# Методы определения высоты

1. Локализовать место фундаментальной частоты и по нему определить высоту тона
2. Найти минимальную частотную разницу между соседними гармониками, равную фундаментальной частоте:

$$(n + 1)f_0 - nf_0 = nf_0 + f_0 - nf_0 = f_0, \quad (8)$$

где  $n = 1, 2, 3, \dots$  и принять её за основу при распознавании высоты

3. Найти общий наибольший сомножитель, который получается при делении всех гармоник на последовательные целые числа, и использовать его как базу для определения частоты.

# Метод 1

- Присутствие фундаментальной частоты является обязательным для определения высоты звука
- Но в 1940 году Шутен продемонстрировал, что ощущение высоты тона не изменится, если вырезать в музыкальном тоне фундаментальную частоту. Этот эффект получил название «феномен пропущенной фундаментальной» и доказал, что метод 1 не может служить единственной базой для определения высоты сложного тона, хотя он работает для большинства музыкальных звуков.

# Метод 2

- Для большинства музыкальных звуков соседние гармоники обычно присутствуют. Слуховая система, оценивая положение их максимумов на базилярной мембране, вычисляет частотную разницу между ними и по ней определяет высоту
- Но если подать звук, в котором присутствуют только нечётные гармоники  $f_0, 3f_0, 5f_0, 7f_0$ , например, 100 Гц, 300 Гц, 500 Гц, 700 Гц и др. Если фундаментальная частота есть в спектре, то слух определяет высоту по ней  $f_0=100$  Гц. Если её вырезать, то расстояние между гармониками останется равным  $2f_0$ , но слух продолжает определять высоту тона, равную фундаментальной  $f_0=100$  Гц.

# Метод 3

- Позволяет объяснить и пропущенную фундаментальную частоту, и наличие только нечётных гармоник, т.к. от отсутствия каких-либо гармоник общий наибольший сомножитель 100 Гц не меняется. Этот метод позволяет также объяснить восприятие слабого ощущения высоты тона у колоколов и других источников квазипериодических тонов.

# Выводы теории места

- Механизм места разворачивает данную гармонику, если критическая полоса её слухового фильтра, построенного на ней как на срединной частоте, достаточно узкая и соседние гармоники внутри этого фильтра не попадают
- Если гармоники находятся настолько близко по частоте друг от друга, что внутри одного слухового фильтра попадает несколько гармоник, то они не разворачиваются.
- Какой бы ни была фундаментальная частота, слуховой механизм разворачивает только первые 6-7 гармоник – именно они и являются определяющими при определении высоты звука



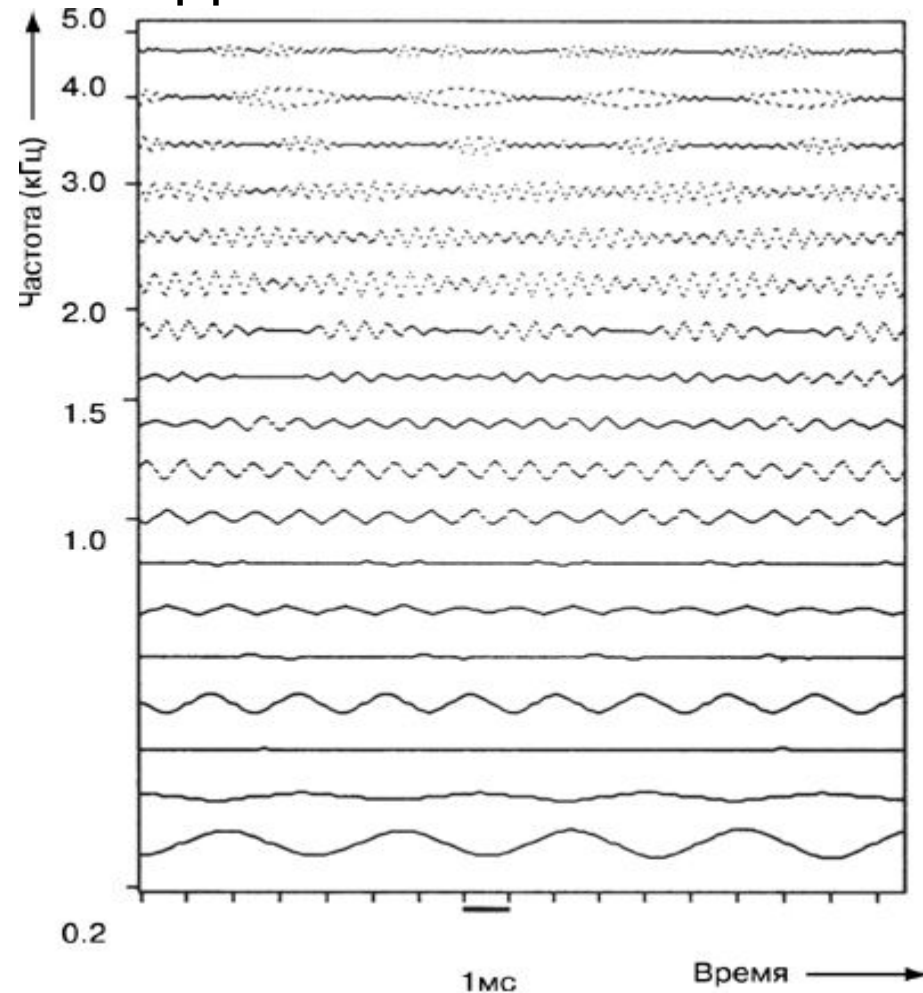
# Выводы теории места

- Теория места создаёт базис для понимания того, как можно определить высоту путём анализа гармонического ряда, но эта теория не может объяснить ряд проблем, например, очень высокая точность определения высоты звука для тонов, чьи частотные компоненты не разворачиваются (т.е. звуки с гармониками выше седьмой).

# Временная теория

- Эта теория использует синхронизацию разрядов нейронов органа Корти с фазой колебания базилярной мембраны (эффект запирающей фазы). При смещениях определённой точки мембраны в сторону расположения волосковых клеток в них возникает электрический потенциал, при смещении в противоположную сторону потенциал отсутствует. Благодаря фазовому запирающему времени между импульсами в любом отдельном волокне будет равно целому числу 1, 2, 3, ..., умноженному на период в основной звуковой волне

Основа временной теории – анализ формы волны в различных частях базилярной мембраны. Если рассматривать механизм частотного анализа на базилярной мембране как работу линейки фильтров различной ширины, то форма волны звукового сигнала, выходящего из этого набора фильтров, должна иметь такой вид:



- Волновая форма выходного сигнала для фильтра, центральная частота которого выше шестой, не синусоидальная, т.к. гармоники не разворачиваются индивидуально, демонстрируя, что частотный диапазон полосового фильтра шире, чем расстояния между ними. По меньшей мере две гармоники комбинируются на выходе этого фильтра. Известно, что если две частоты находятся достаточно близко друг от друга, между ними возникают биения, т.е. амплитудно модулированное колебание с частотой модуляции, равной разности частот. В данном случае, когда взаимодействуют две гармоники, этот период определяется фундаментальной частотой  $T=1/f_0$ . Таким образом, период всех волн, выходящих после фильтров с центральной частотой выше шестой гармоники и состоящих из соседних гармоник, будет одинаковым и равным  $1/f_0$ .

- Минимальное время между импульсами от различных мест на базилярной мембране определяется периодом волны, выходящей от соответствующего фильтра. Для мест, соответствующих частотам от основной до шестой гармоники, минимальное время равно периоду данной гармоники. Для мест, соответствующих более высоким гармоникам, промежутки между импульсами равны периоду огибающей, т.е. основному тону.
- Таким образом, выше шестой гармоники разряды нейронов синхронизированы с формой огибающей, и период разрядов совпадает с периодом для фундаментальной частоты. Иными словами, для всех гармоник периоды разрядов или равны, или отличаются в целое число раз от частоты основного тона.

# Основа временной теории восприятия высоты тона

- Мозг определяет периодичность разрядов и по ним восстанавливает частоту основного тона. Восприятие музыкальной высоты связано с оценкой временной формы звукового сигнала (за счёт использования эффекта «фазового запирания»)

# Выводы временной теории

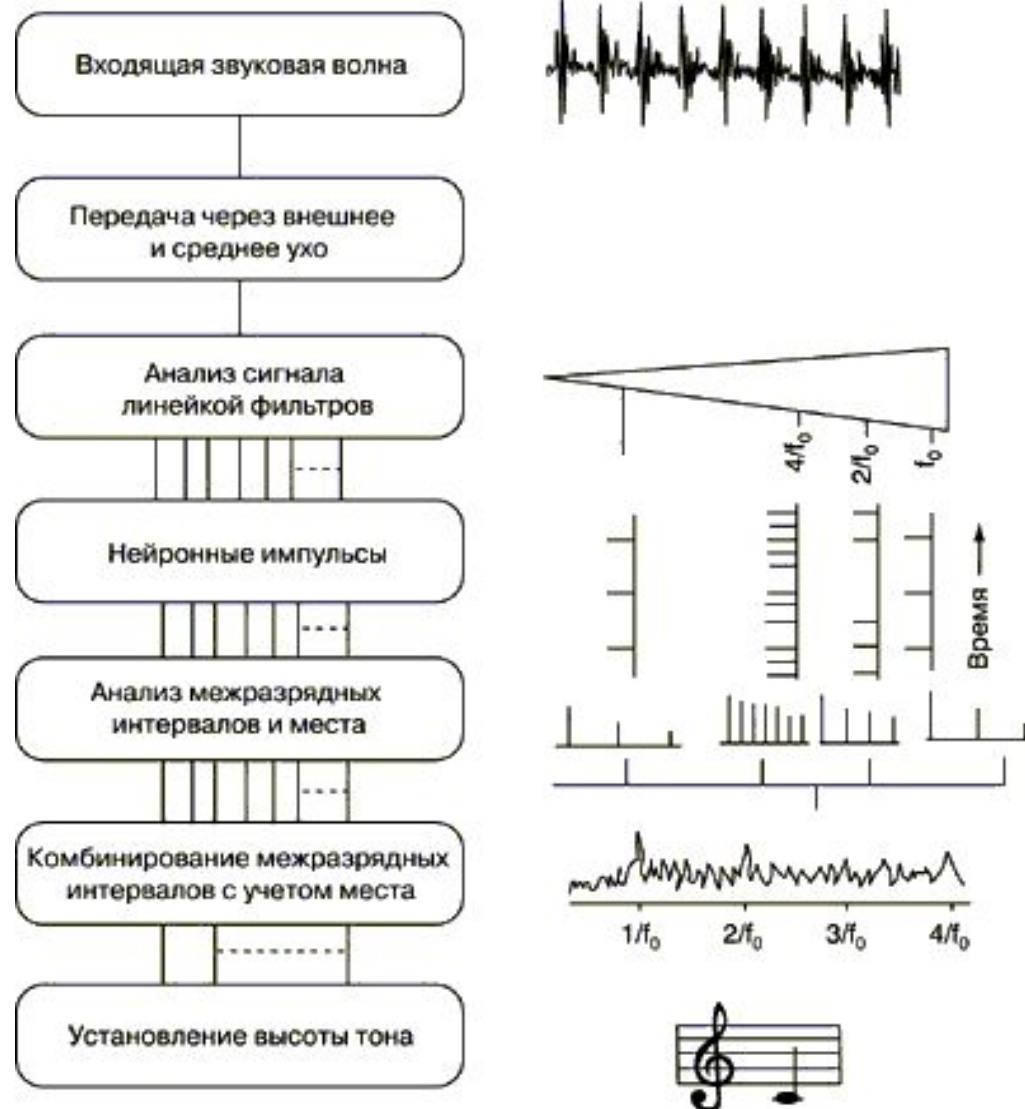
- Временная теория позволяет понять, как найти фундаментальную частоту на основе анализа временных интервалов между нервными импульсами от различных мест на базилярной мембране и по ней определить высоту тона.
- Но временная теория не объясняет восприятия высоты тона на частотах выше 5000 Гц, т.к. эффект фазового запираения не срабатывает на этих частотах. Вероятно, в этой области частот меняется механизм восприятия высоты тона.

# Современная теория восприятия высоты тона

- Согласно современным теориям, мозг принимает информацию от периферийной слуховой системы как за счёт индикации места (частотный анализ), так и за счёт информации о форме звуковой волны (временной анализ).



# Современная модель для восприятия ВЫСОТЫ ТОНА



# Анализ восприятия высоты музыкального тона с помощью предложенной модели позволил получить ряд интересных результатов:

1. Для музыкальных тонов с основной частотой от 100 до 400 Гц (с уровнем звукового давления не менее 50 дБ) основную роль в определении высоты тона играют первые пять-шесть гармоник (если их уровень превышает 10 дБ), т.е. те гармоники, которые разворачиваются слуховыми фильтрами
2. Звуковые сигналы, содержащие только очень высокие гармоники (свыше двадцатой), не вызывают ощущения высоты тона
3. Музыкальные сигналы, содержащие очень низкие частоты (с основной частотой ниже 50 Гц) вызывают ощущение высоты тона только по обертонам, т.к. такие низкие частоты не вызывают смещений базилярной мембраны – они на ней не размещаются, им не хватает места. При этом наиболее существенную роль играют пятые-шестые гармоники
4. Фундаментальная частота звука, если она выше 1000 Гц, является доминантной компонентой в определении высоты тона

5. Музыкальные звуки, содержащие только неразвёрнутые гармоники (свыше шестой), могут дать ощущение высоты тона по огибающей, при этом слух производит достаточно тонкую дифференциацию сдвига максимума огибающей, т. е. точно ощущает высоту.
6. Фазовые соотношения различных гармоник в музыкальном сигнале оказывают влияние на восприятие высоты, т.к. их изменение приводит к изменению структуры огибающей для высших неразвёрнутых гармоник. Для музыкальных сигналов, содержащих много низких и высоких гармоник, изменение фазовых соотношений может привести к улучшению чёткости восприятия высоты, не вызывая её сдвига (т.к. они не влияют на оценку низших развёрнутых гармоник). Для сигналов, содержащих в основном высокие гармоники, изменение их фазы может вызвать сдвиг высоты тона и изменение его чёткости, т.к. может привести к сдвигу пиков в огибающей, по которым и определяется высота тона.

# Обработка звуковой информации в мозге

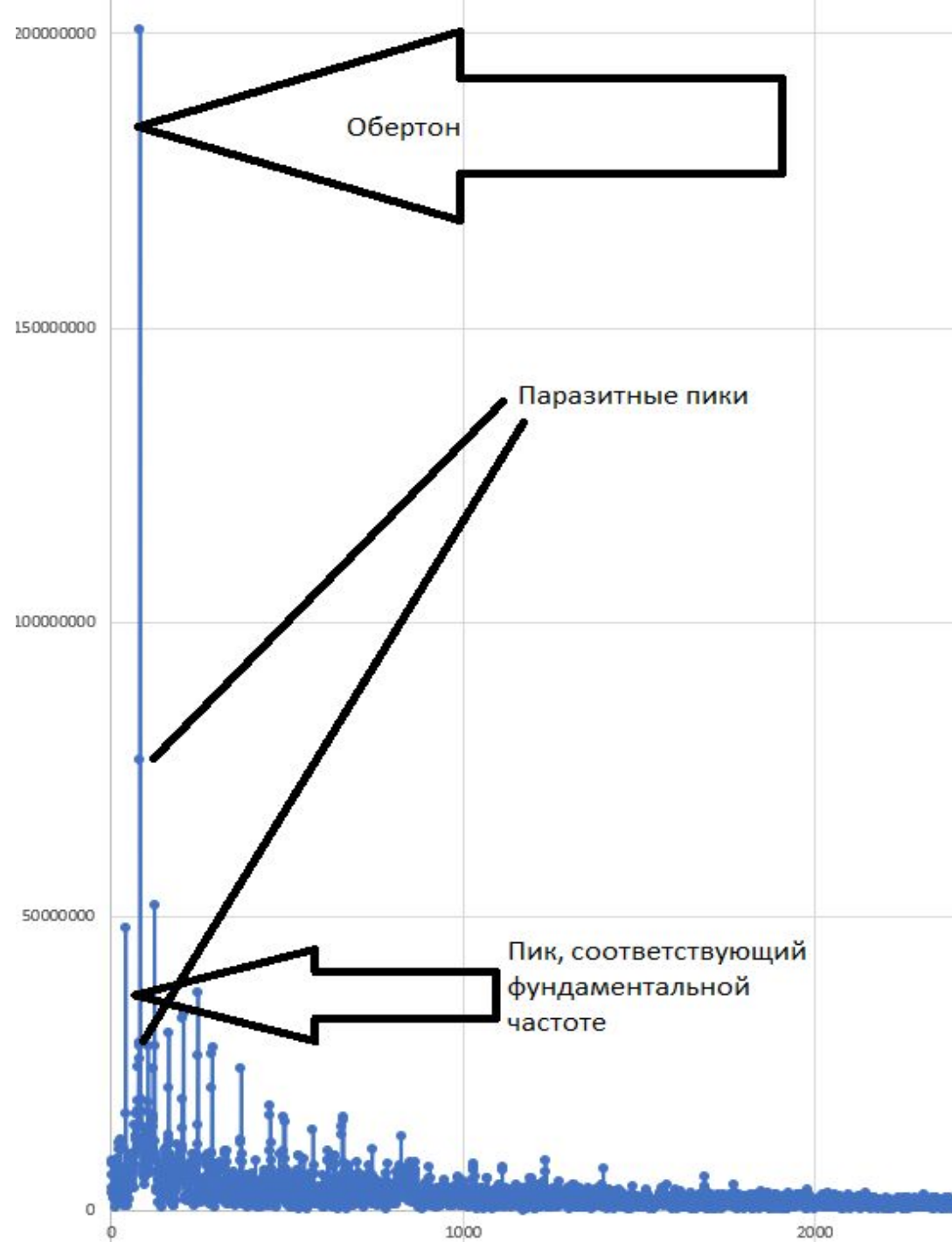
- В настоящее время принята гипотеза, что мозг, получив информацию от периферической слуховой системы о наличии компонент с кратными периодами в музыкальном звуке, группирует их и сравнивает с гармоническим шаблоном, в котором имеются все последовательные гармоники
- Если к данному сигналу подходят два шаблона с разными фундаментальными частотами, можно ожидать услышать или неопределённую высоту, или две высоты

# Применение теории восприятия высоты тона к программе

- В программе был применён 1 метод теории места, в результате чего возникает «феномен пропущенной фундаментальной»
- Было решено воспользоваться вторым методом теории места, т.е. находить два пика, соответствующих соседним гармоникам (двум обертонам или обертону и фундаментальной частоте), находить разность их частот и принимать её за фундаментальную частоту

# Возникшие проблемы

- Но возникла проблема: в амплитудном спектре пик с максимальной среди всех других амплитудой оказался как бы размазанным, то есть на расстоянии одного или совсем небольшого числа отсчётов есть пики, меньшие по амплитуде, чем обертон, но большие, чем все остальные, но расположенные практически в том же месте спектра. Это т.н. эффект размазывания, из-за которого воспользоваться вторым методом теории места невозможно, если его (эффект) не устранить, потому что вместо двух соседних обертонов определяется размазанный один и тот же.



# Заключение

- В ходе данной работы был изучен и реализован алгоритм быстрого преобразования Фурье по основанию два с прореживанием по времени, в результате чего скорость работы программы многократно возросла. Была изучена теория восприятия высоты тона, в результате чего была установлена причина ошибочного определения фундаментальной частоты и найдены методы избегания данной ошибки. Но при попытке применить эти методы была определена ещё одна нерешённая проблема: эффект размазывания.



# Вследствие чего задачами и перспективами дальнейших исследований являются

1. Изучить и устранить эффект размазывания;
2. Применить второй метод теории места;
3. Применить временную теорию, т.е. задействовать фазовый спектр;
4. Для снижения влияния помех реализовать оконное преобразование Фурье;
5. Реализовать анализ аудиосигнала непосредственно со звукопринимающего устройства в режиме реального времени, сделать тюнер.