

# БИОИНСПИРИРОВАННЫЙ И ПСИХОАКУСТИЧЕСКИ МОТИВИРОВАННЫЙ АНАЛИЗ ЗВУКОВЫХ СИГНАЛОВ

*к.т.н. Вашкевич М. И., д.т.н. Азаров И.С.*



*Белорусский государственный университет  
информатики и радиоэлектроники (БГУИР)*

*Кафедра электронных вычислительных средств*



# Цель доклада

Дать представление о существующих методах **биоинспирированного** и **психоакустически мотивированного** анализа звуковых сигналов, а также обозначить их области применения.

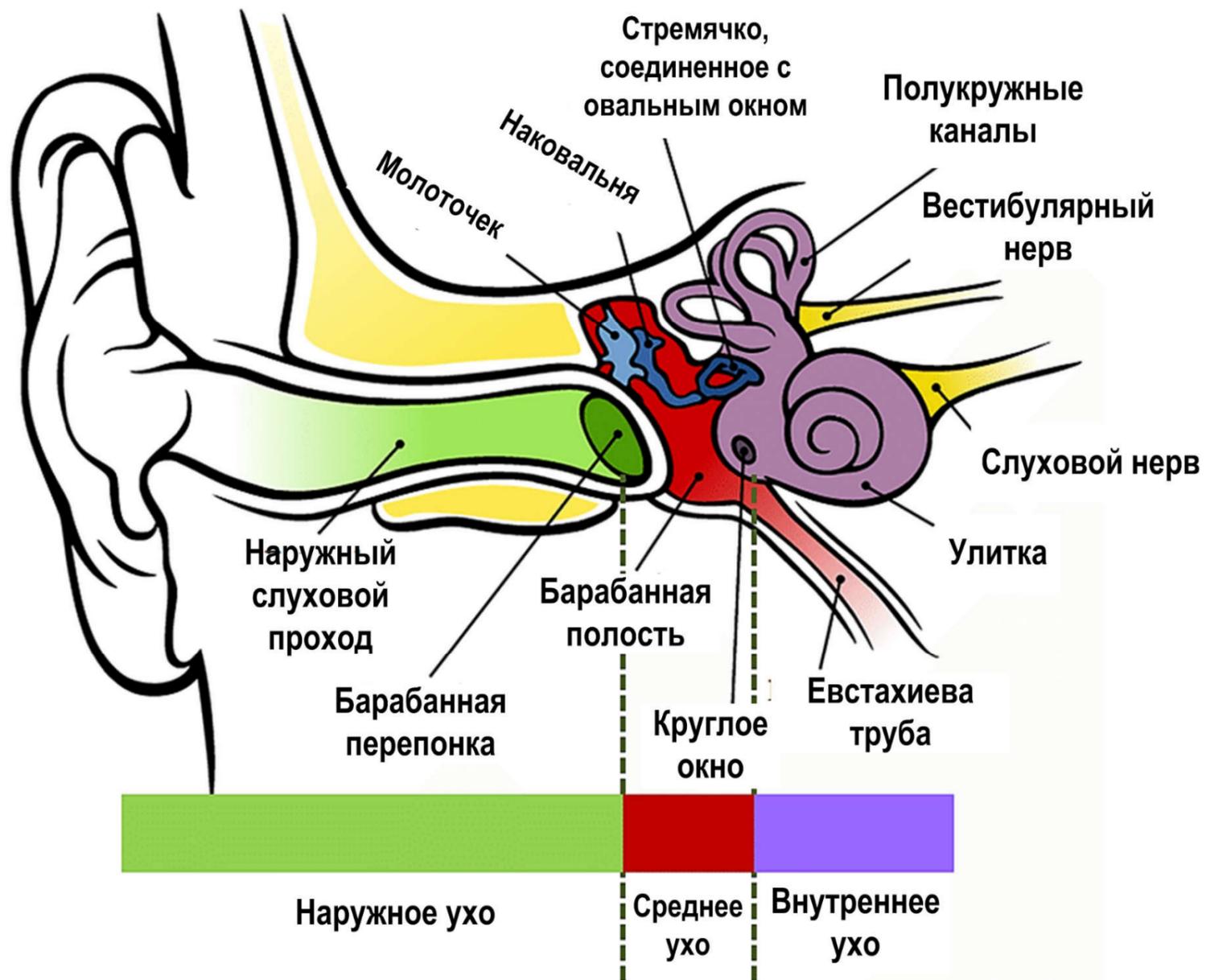
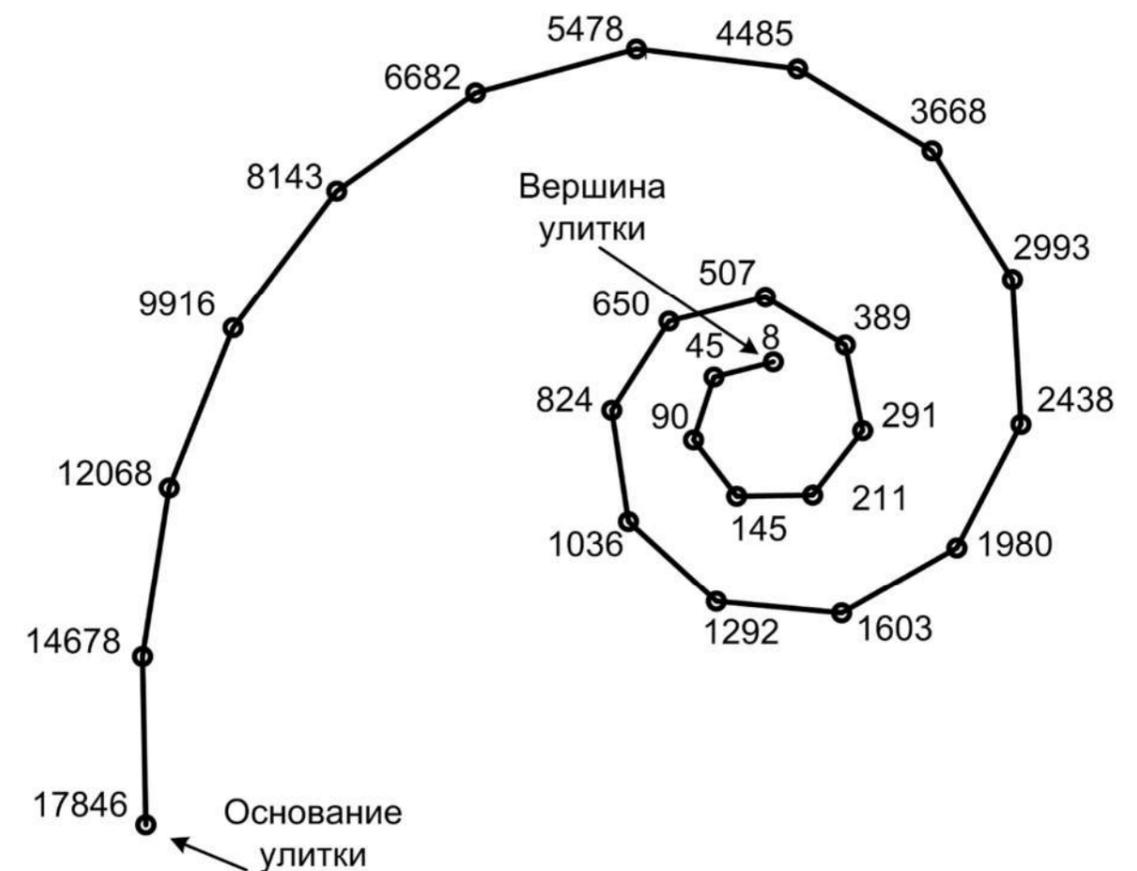


Диаграмма базилярной мембраны. Показаны позиции максимального смещения при стимуляции синусоидами с различной частотой (в Гц).



# Две группы методов

**1) Биоинспирированные методы** – решают задачу моделирования работы органов слуховой системы. В основе лежат параметрические выражения, подогнанные под имеющиеся физиологические данные.

*Примеры физиологических данных:*

- характеристики возбуждения слухового нерва;
- измеренное смещение базилярной мембраны;

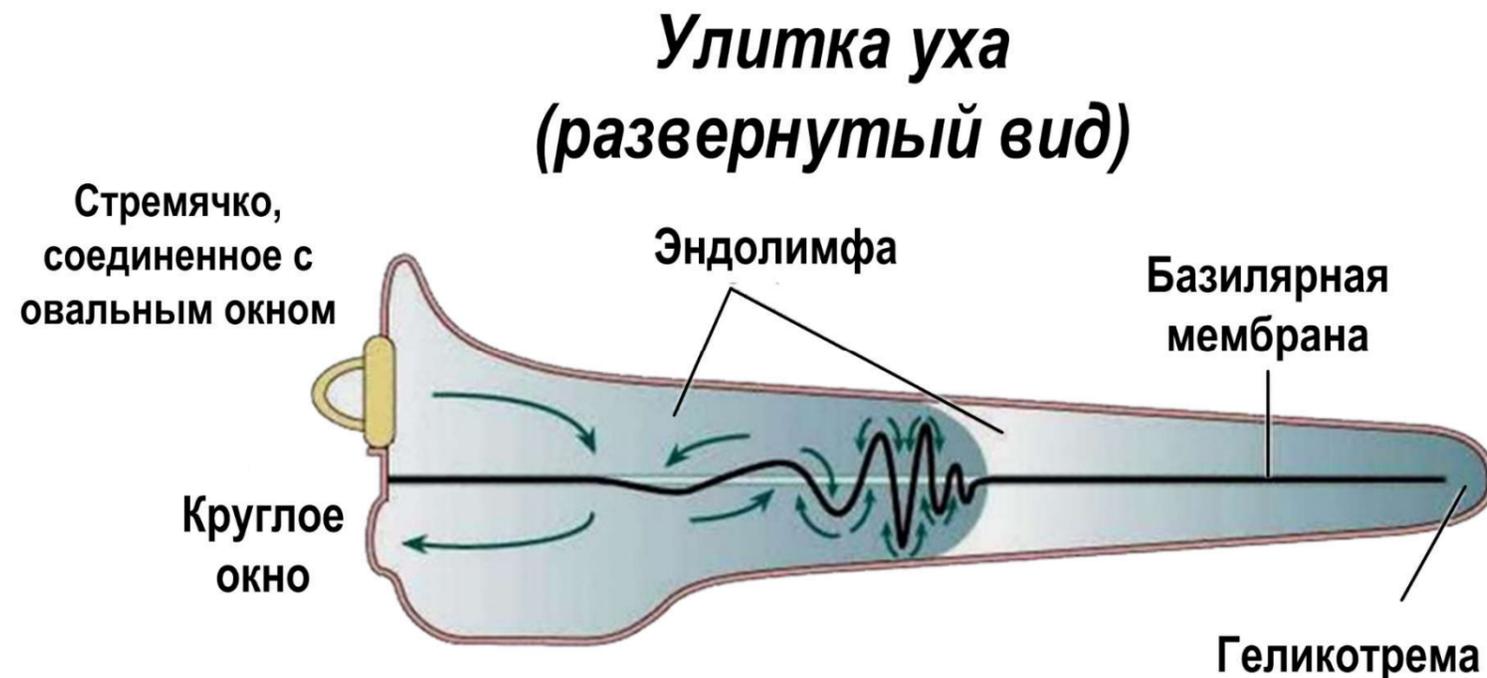
**2) Психоакустически мотивированные методы** – основаны на закономерностях, выявленных в процессе изучения слуховых впечатлений, вызываемых звуками.

*Примеры психоакустических закономерностей:*

- наличие критических частот;
- измерение высоты и громкости звука;

# Слуховые фильтры

На базилярной мембране имеется ~3000 внутренних волосковых клеток, т.е. каждой клетке отвечает свой *частотный диапазон*.



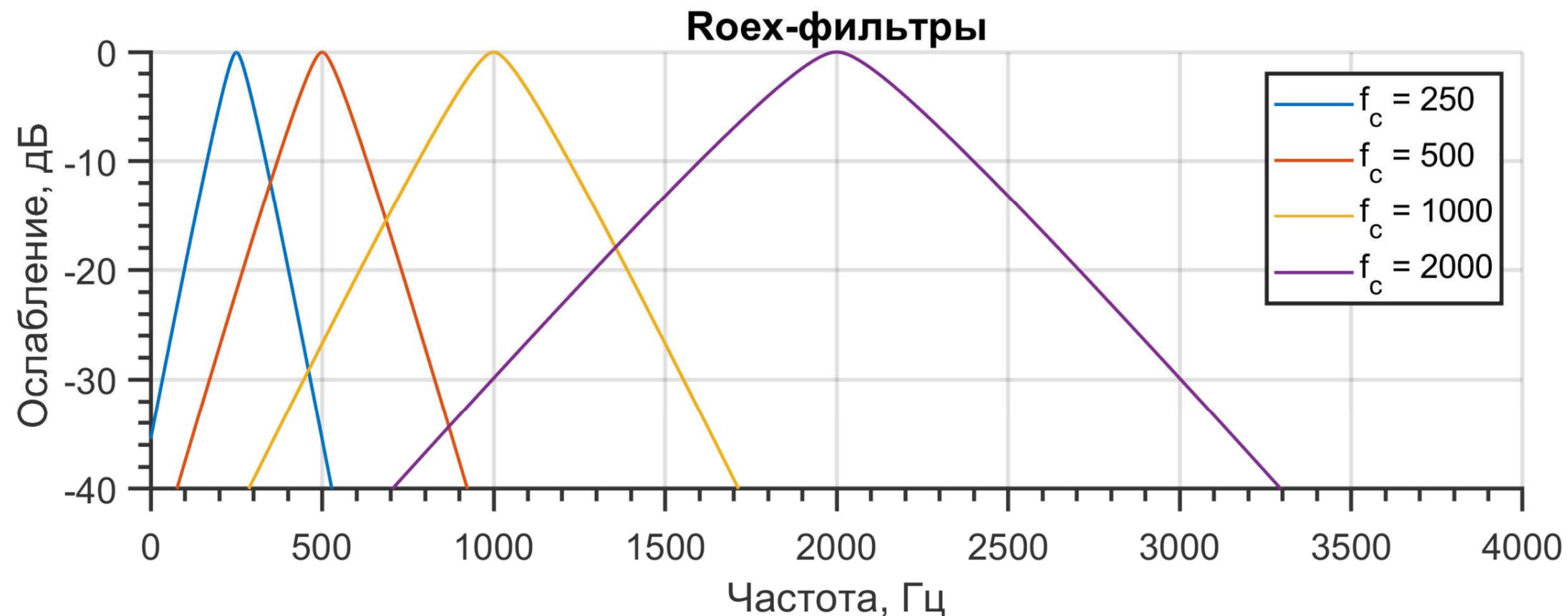
Механическое движение базилярной мембраны улитки уха моделируется при помощи **банка цифровых фильтров**. Такие фильтры называются **кохлеарными** или **слуховыми**.

# БМ слуховых фильтров: роех-фильтры

**Роех-фильтры** (англ. **round exponential**) – описание слухового фильтра в частотной области с помощью скругленной экспоненциальной функции:

$$|H(f)|^2 = (1 - r)(1 + pg)e^{-pg} + r, \quad g = |f - f_c|/f_c.$$

где  $g$  – нормализованное отклонение частоты  $f$  от центральной частоты фильтра  $f_c$ .

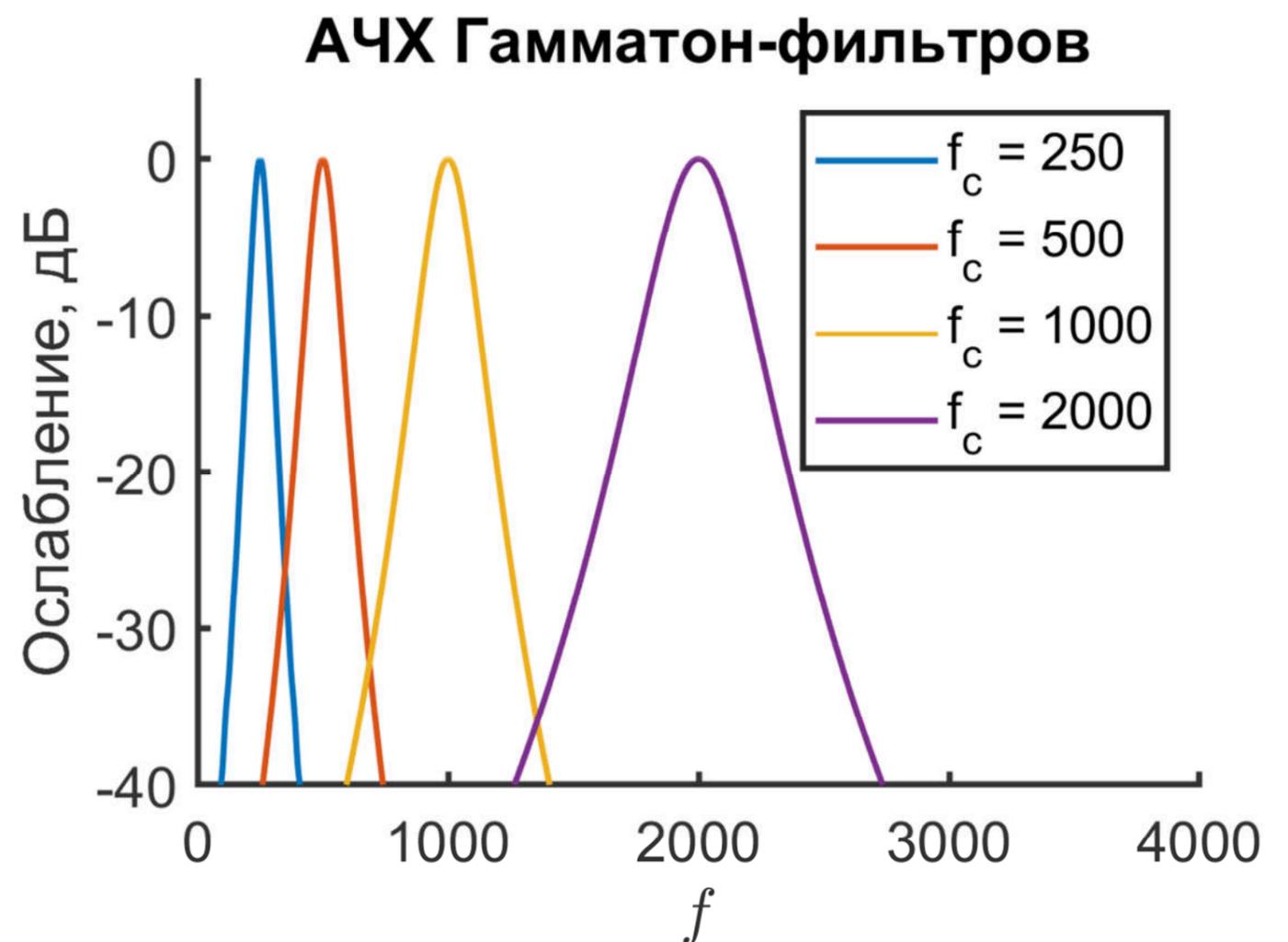
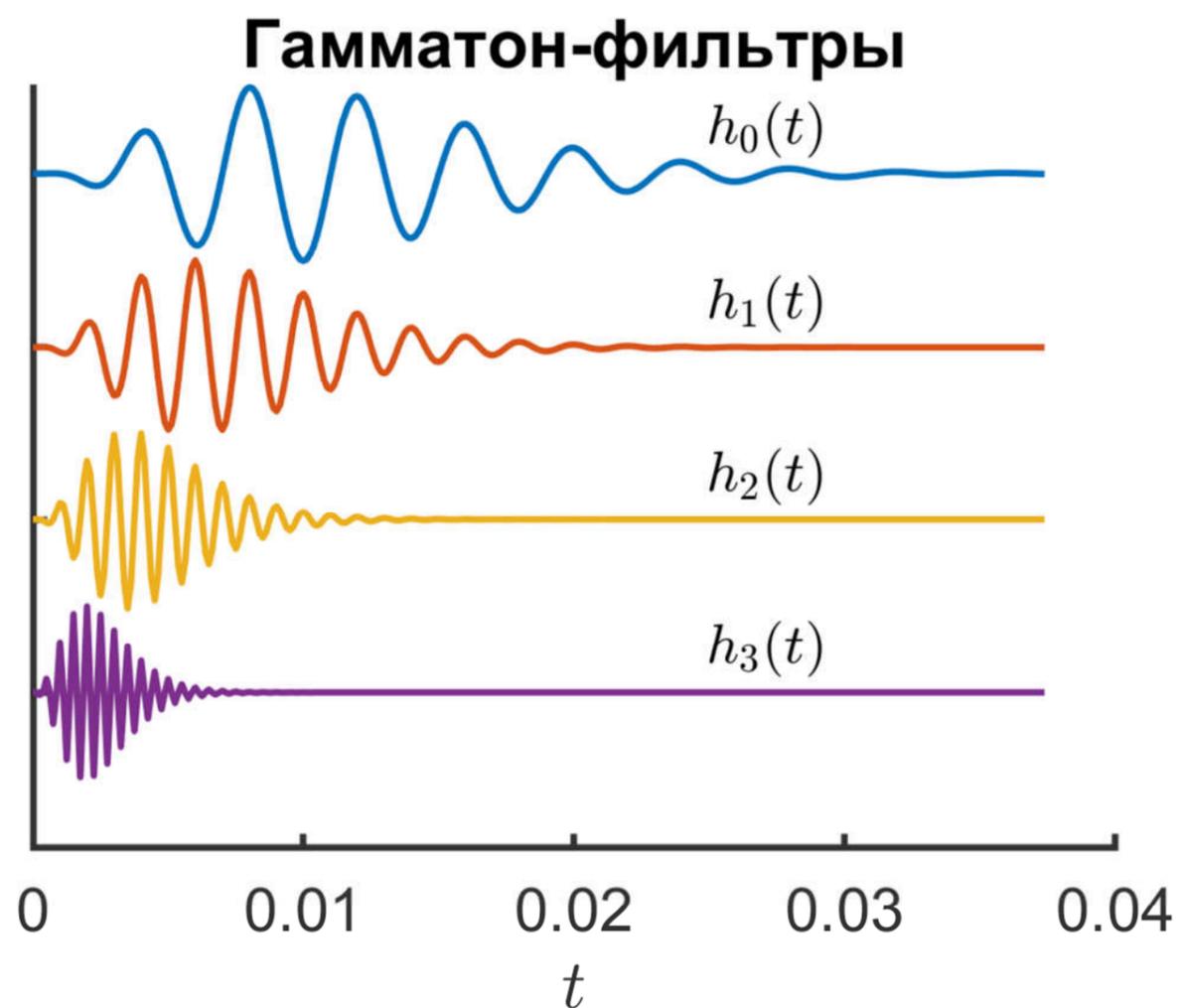


# БМ слуховых фильтров: гамматон-фильтры

**Гамматон-фильтры** (англ. **gammatone**) – описание слухового фильтра во временной области:

$$h(t) = t^{(l-1)} e^{-2\pi b \text{ERB}(f_c)} \cos 2\pi f_c t, \quad t > 0.$$

где  $b$  – параметр, определяющий ширину полосы,  $l$  – порядок.

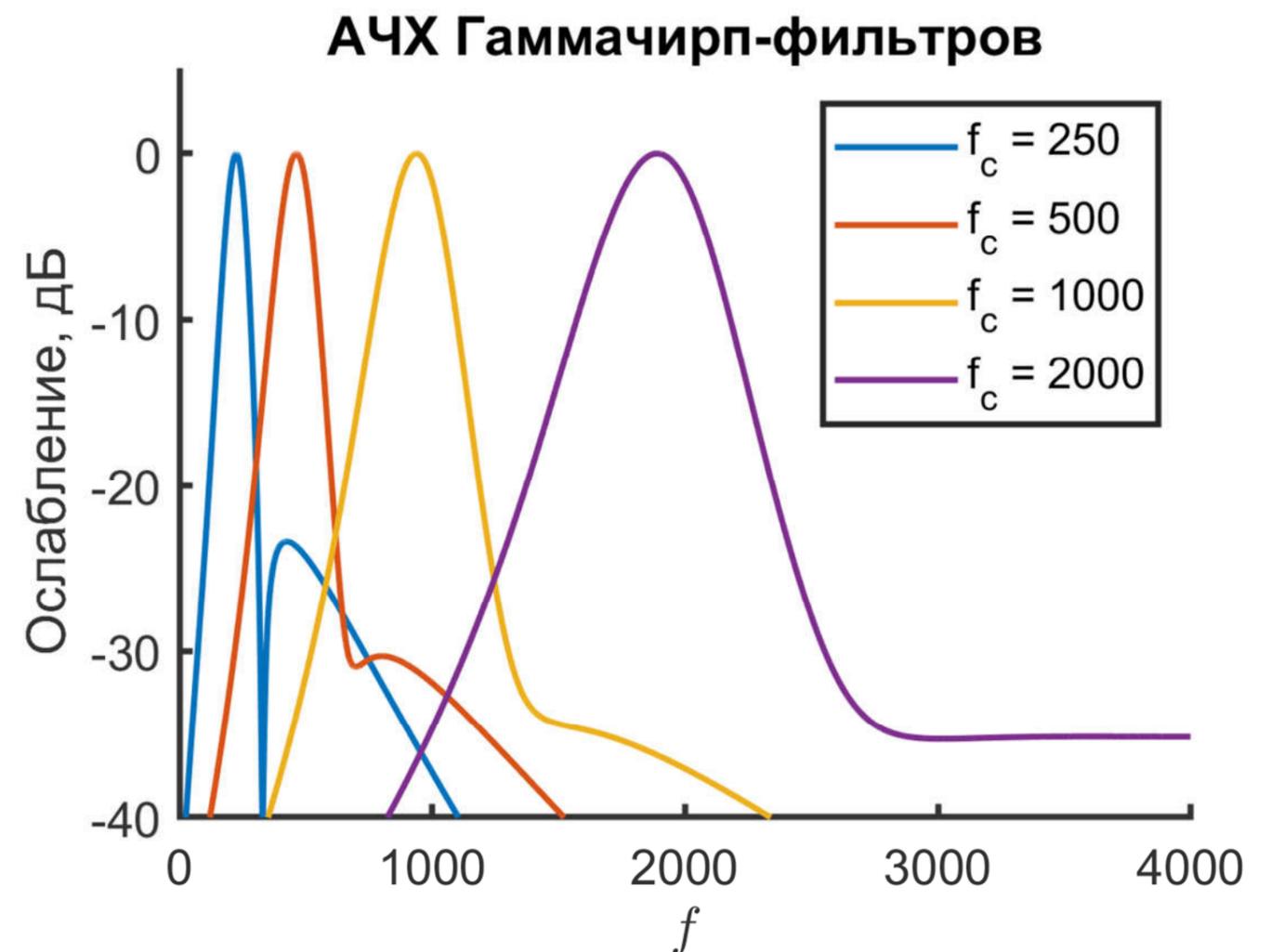
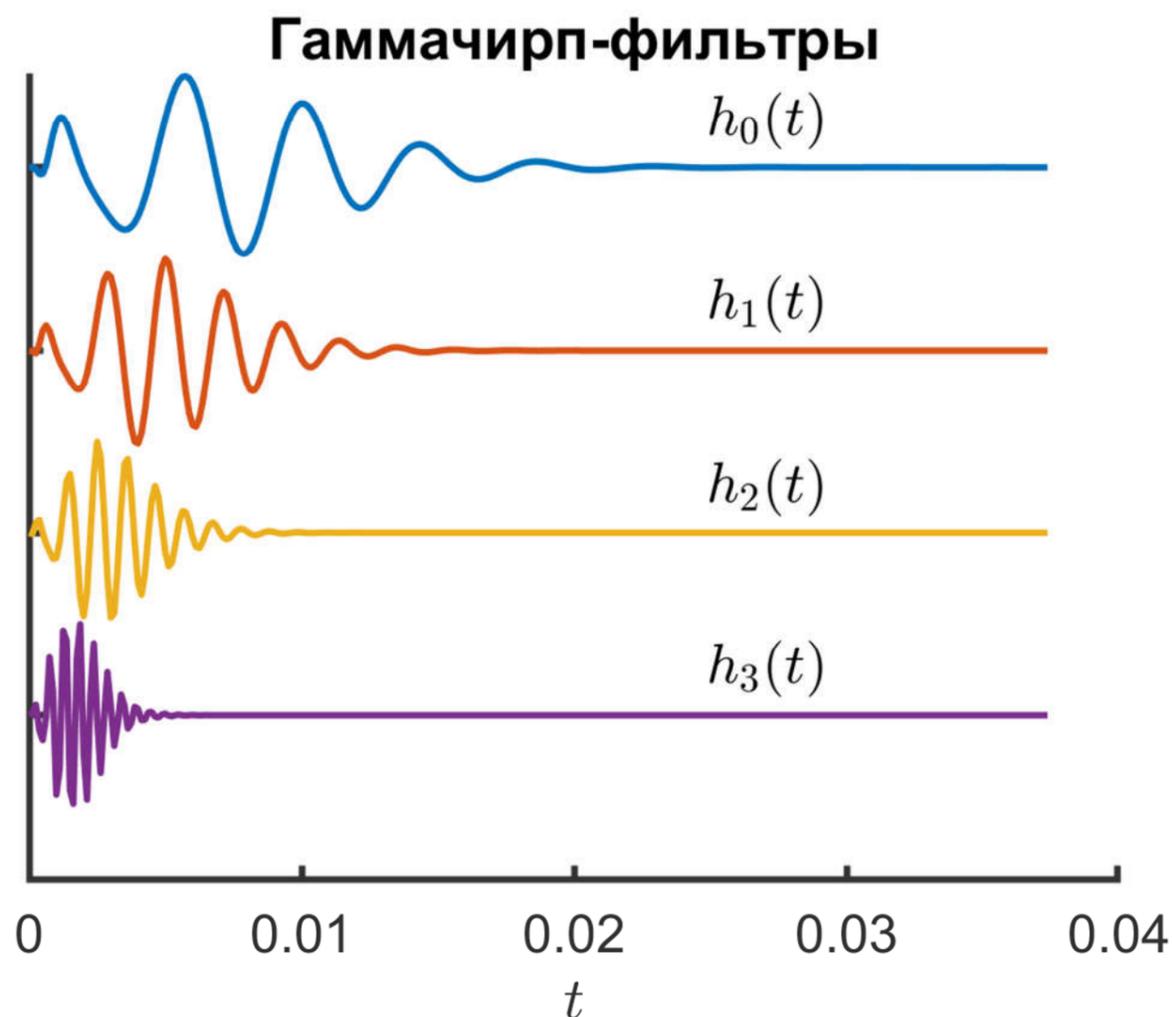


# БМ слуховых фильтров: гаммачирп-фильтры

**Гаммачирп-фильтры** (англ. **gammachirp**) – уточненное описание слухового фильтра во временной области:

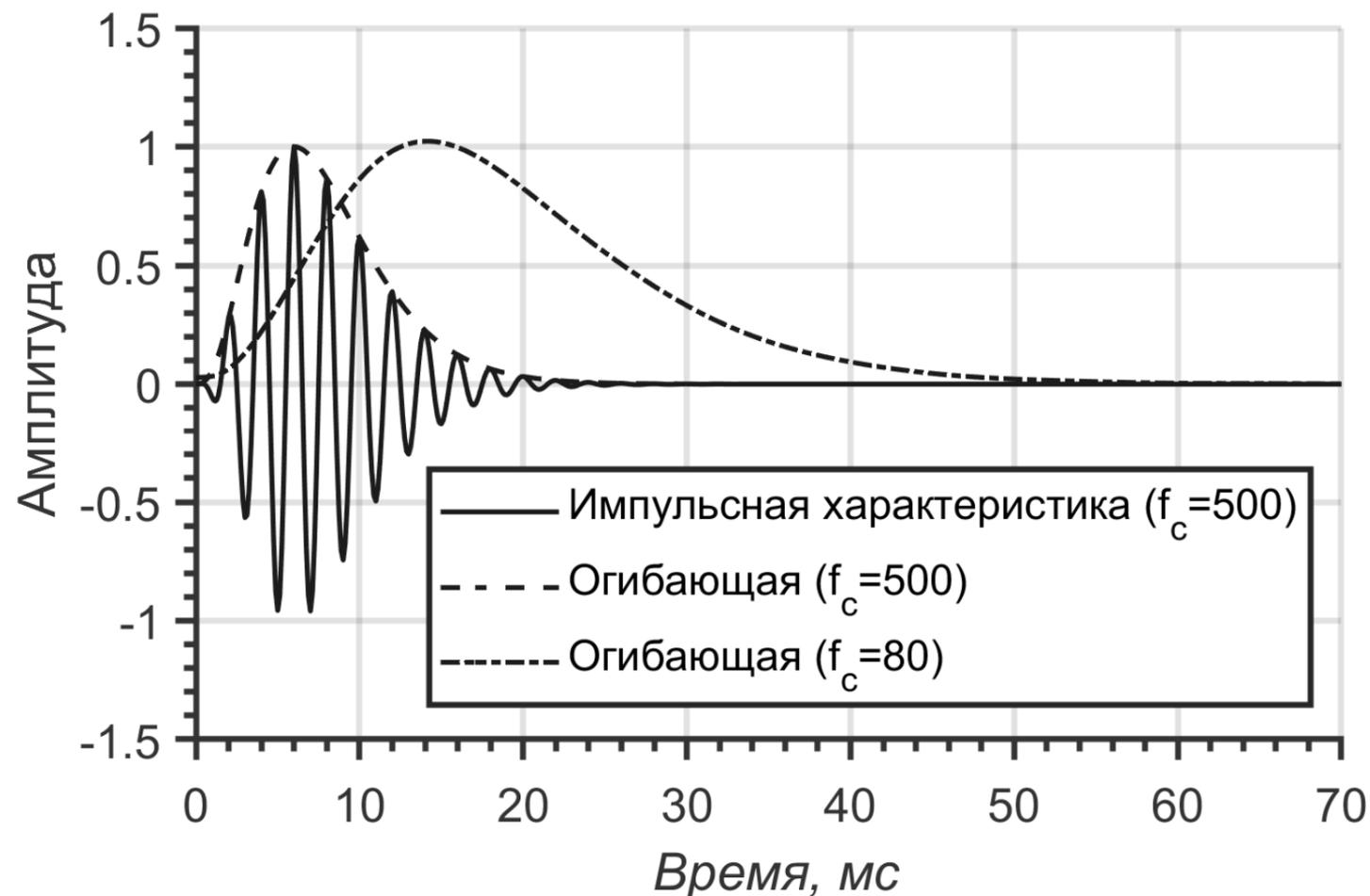
$$h(t) = t^{(l-1)} e^{-2\pi b \text{ERB}(f_r)} \cos(2\pi f_r t + c \ln t + \phi), \quad t > 0.$$

где  $f_r$  – пиковая частота,  $c$  – дополнительный параметр.



# Практическая реализация слуховых фильтров

Импульсные характеристики слуховых фильтров дискретизируют и ограничивают умножением на окно.



**Основная проблема:**  
импульсные характеристики имеют большую длину.

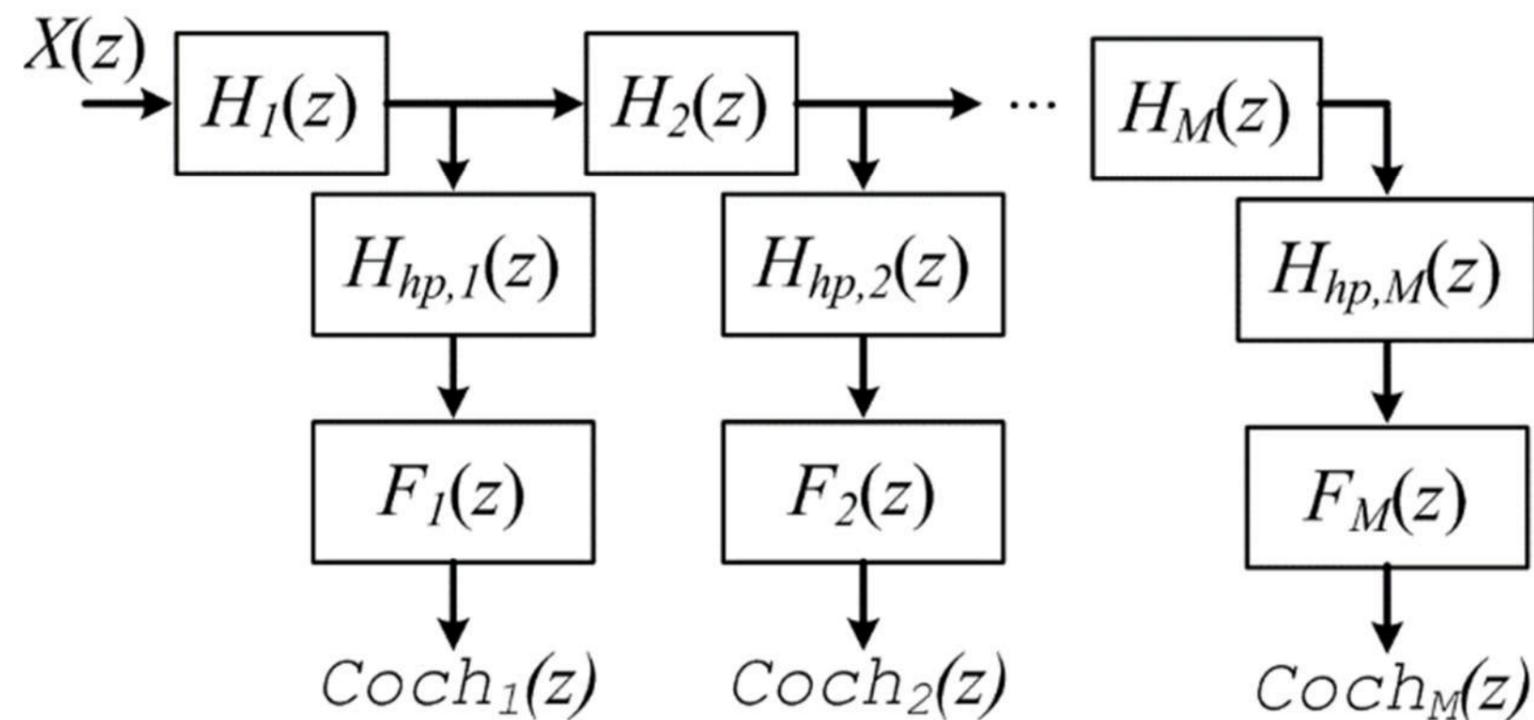
Пример ( $f_s = 8000$  кГц):

при  $f_c=500$  Гц длина импульсной характеристики **240 отсчетов**.

при  $f_c=80$  Гц длина импульсной характеристики **600 отсчетов**.

# Модель улитки уха во временной области (1)

Дж. Кейтс предложил моделировать распространение *бегущей волны* на мембране при помощи каскада цифровых БИХ-фильтров второго порядка.

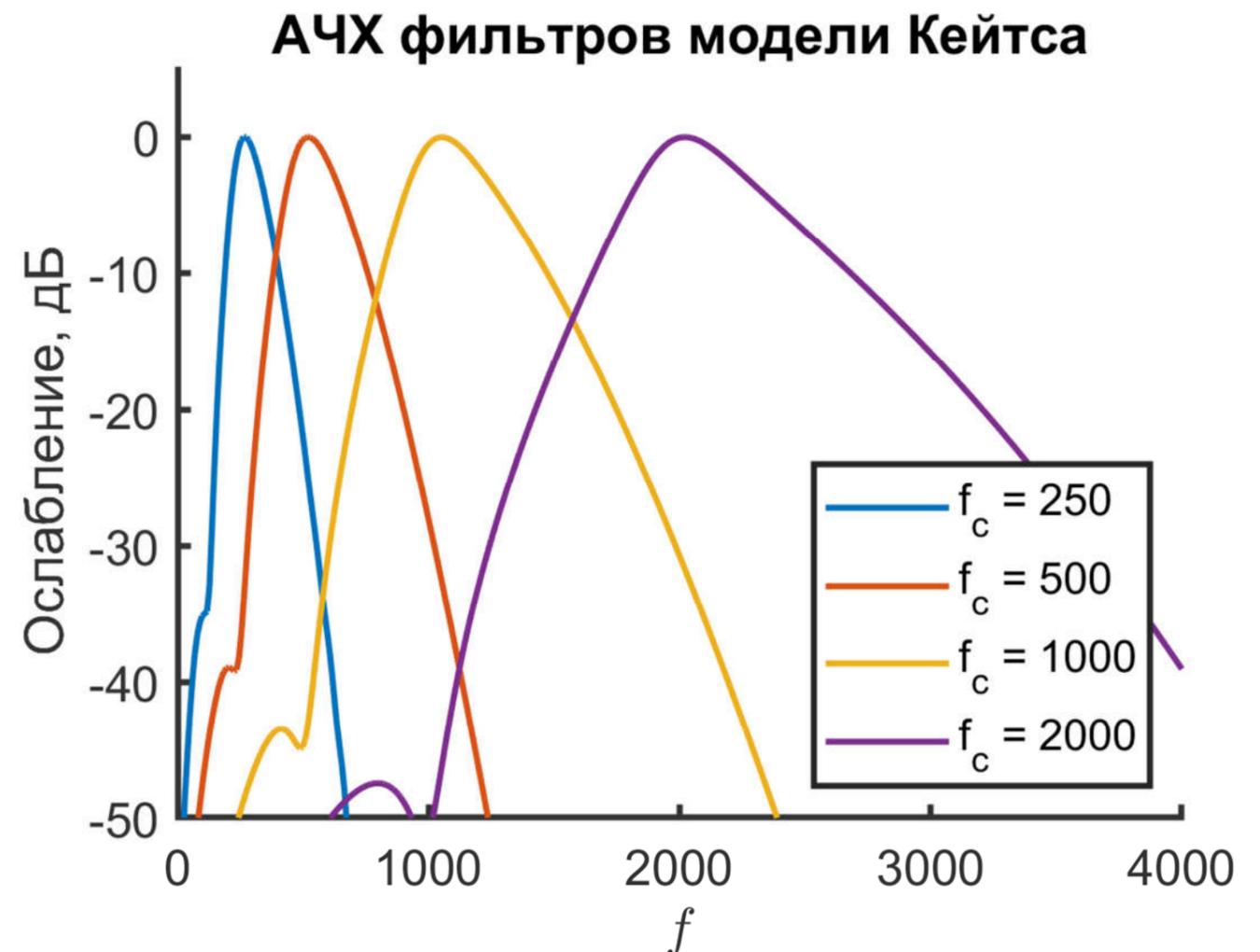
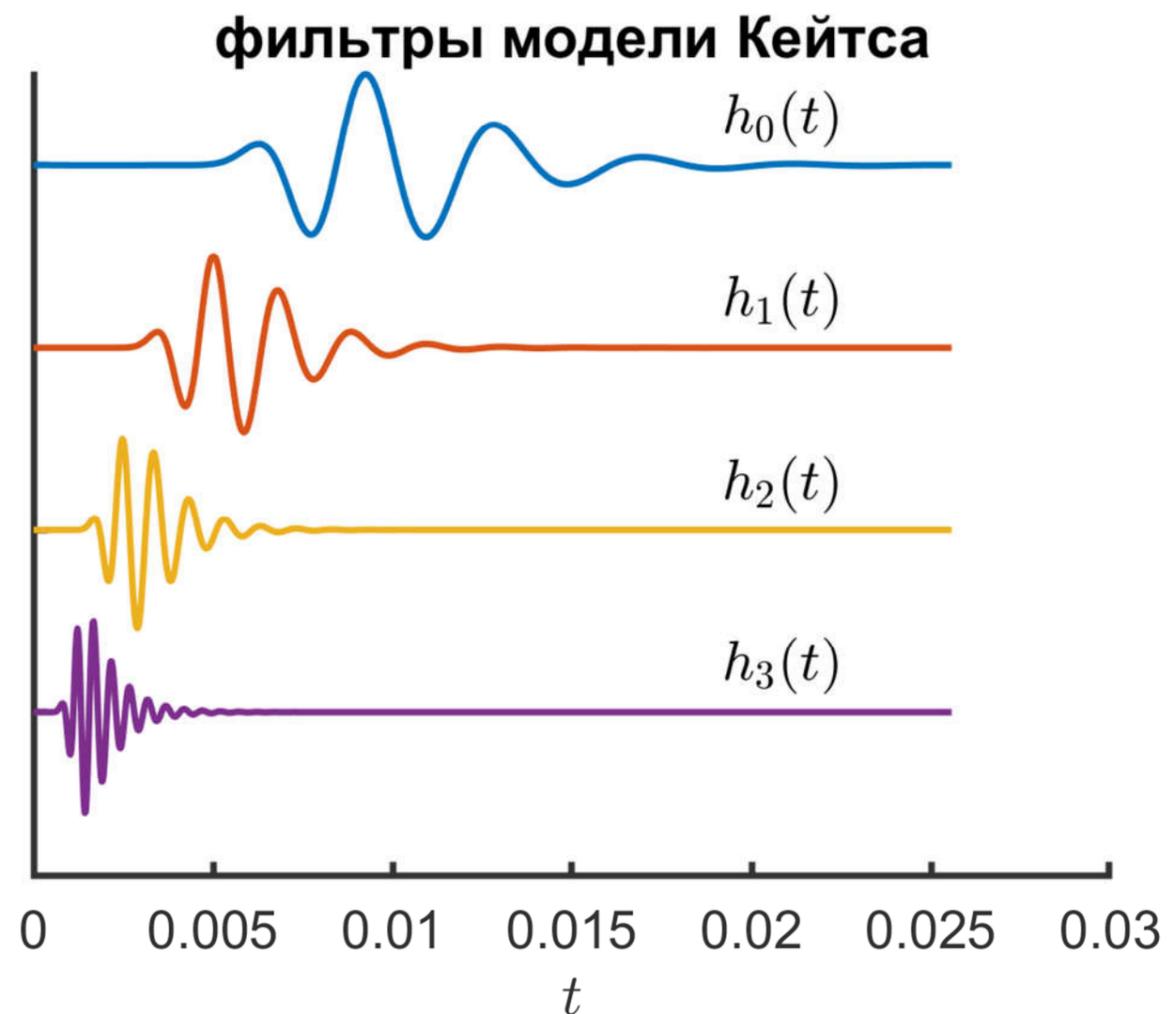


$H_{hp,k}(z)$  – однополосный фильтр, моделирующий преобразование «давление-скорость»;  $F_k(z)$  – режекторный фильтр;  $H_k(z)$  – отдельная секция фильтра бегущей волны, которая усиливает сигнал на частотах близких к резонансной частоте.

# Модель улитки уха во временной области (2)

Передаточная функция слуховых фильтров модели Кейтса имеет вид:

$$\text{Coch}_k(z) = H_{hp,k}(z)F_k(z) \prod_{i=1}^k H_k(z).$$



# Психоакустически-мотивированный анализ

Недостатки *биоинспирированных* методов анализа:

- ✓ большие вычислительные затраты;
- ✓ методы далеки от преобразований, которые обычно используются в ЦОС;
- ✓ отсутствие процедур восстановления сигнала.

*Психоакустически мотивированные* методы анализа:

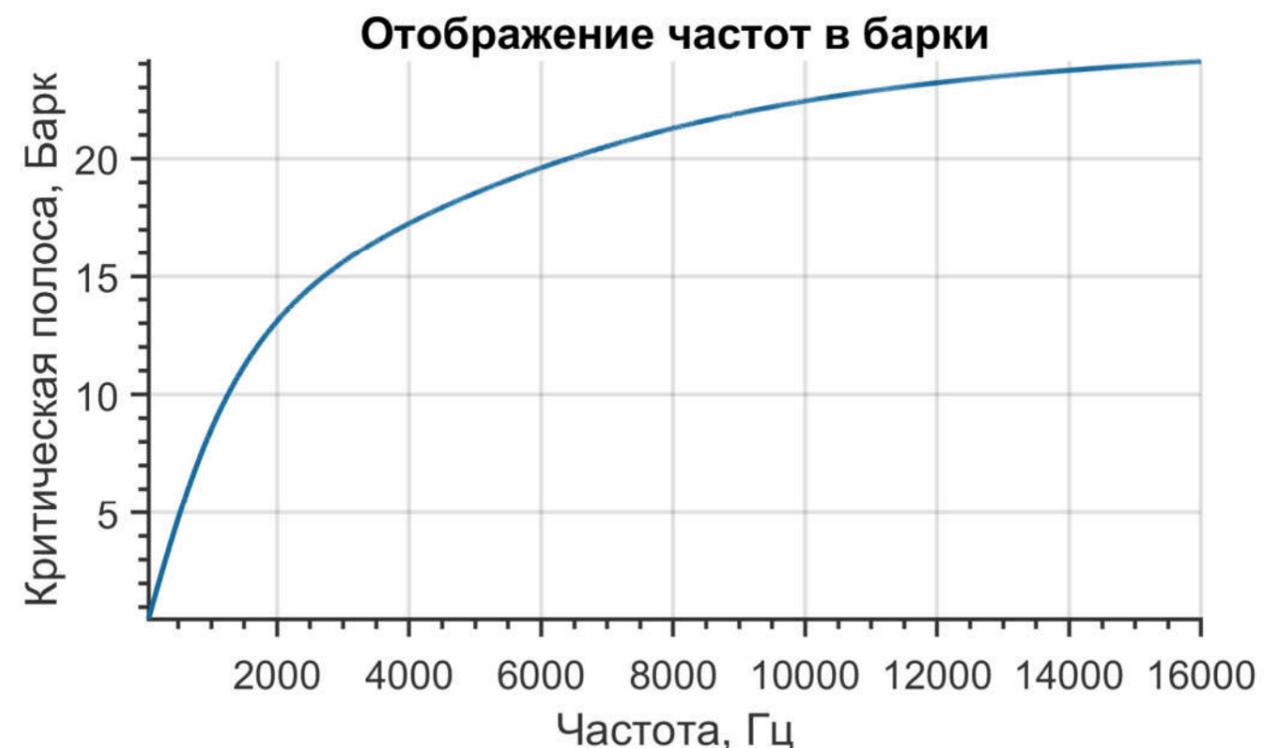
- ✓ основаны на результатах психоакустических экспериментов;
- ✓ моделируют частотное (иногда временное) разрешение слуха;
- ✓ базируются на хорошо изученных и эффективных преобразованиях ЦОС.

# Критические полосы частот

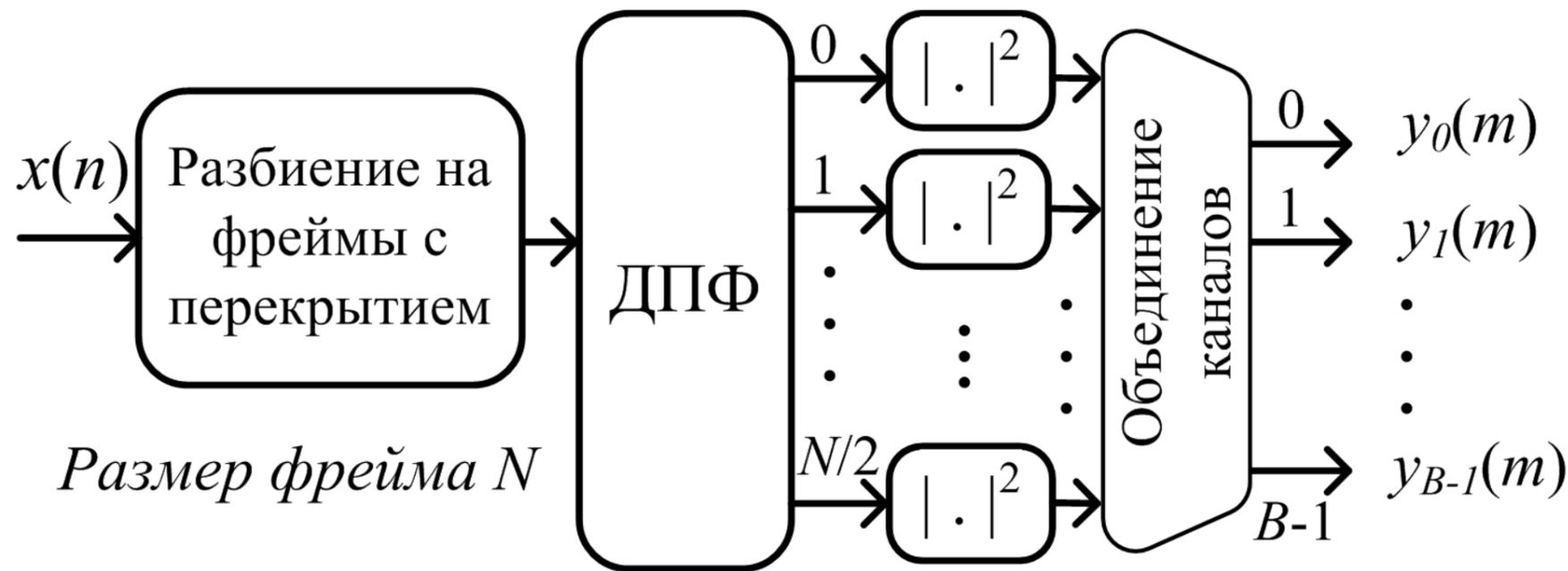
**Критические полосы** определяют частотные зоны внутри которых наблюдается взаимодействие звуковых энергий.

**Критическая полоса** не есть диапазон с фиксированными нижней и верхней границами: это частотный интервал («барк») с шириной 20% от средней частоты, который может быть обнаружен в любом месте слухового диапазона в зависимости от условий эксперимента. Слуховой диапазон покрывается 24 барками.

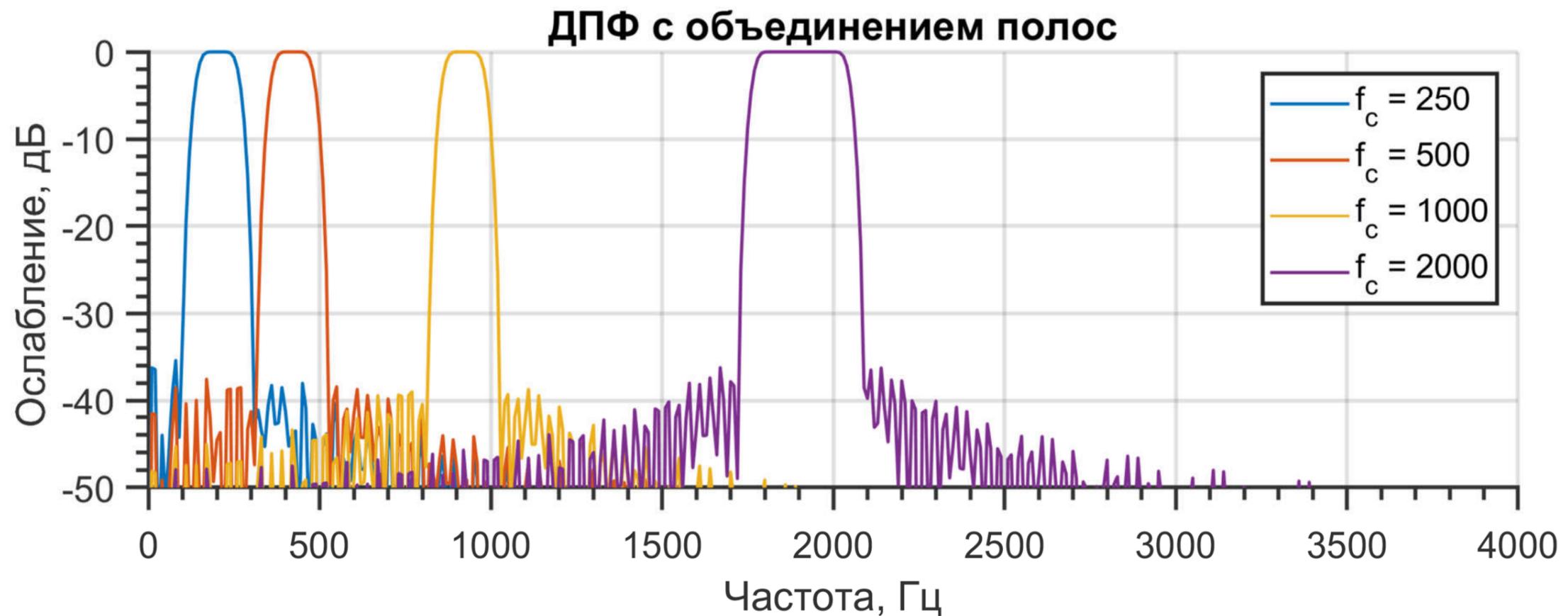
$$z(f) = 13 \operatorname{arctg} \left( \frac{7,6}{10^4 f} \right) + 3,5 \operatorname{arctg} \left( \left( \frac{f}{7500} \right)^2 \right) \text{ [барк]}$$



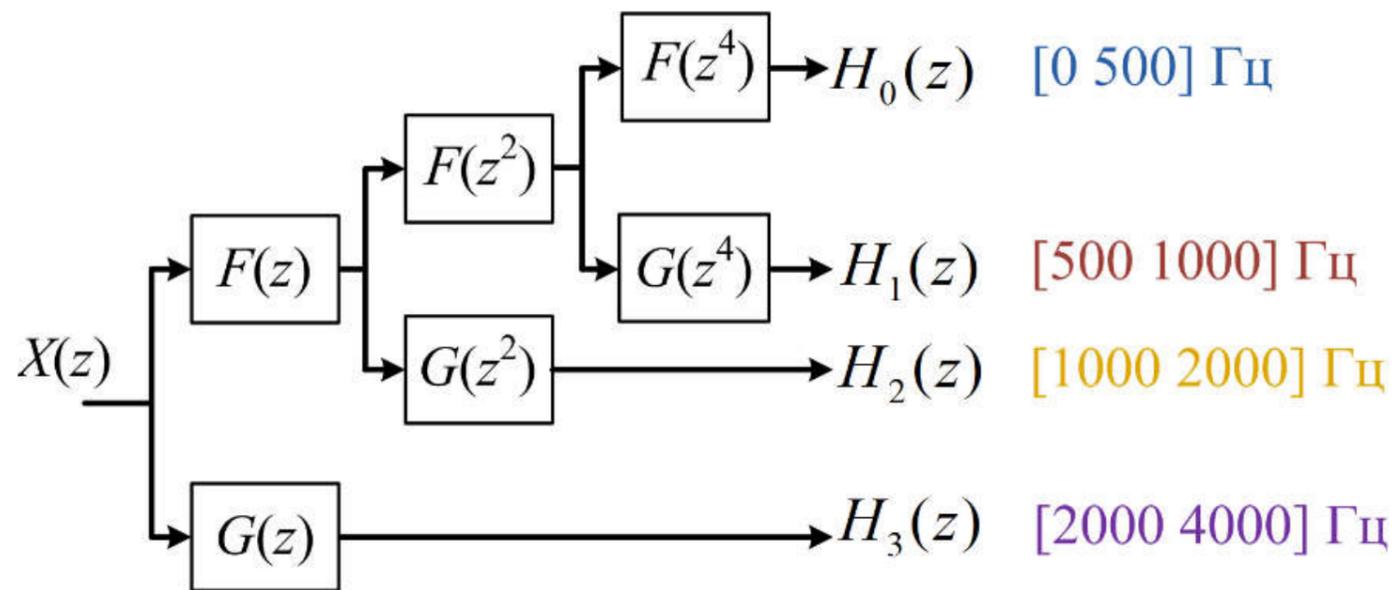
# ДПФ с объединением полос



- ✓ Фильтры определены в **частотной области**.
- ✓ Низкая вычислительная сложность.
- ✓ Не моделируется временное разрешение слуха.



# Слуховые фильтры на основе вейвлетов



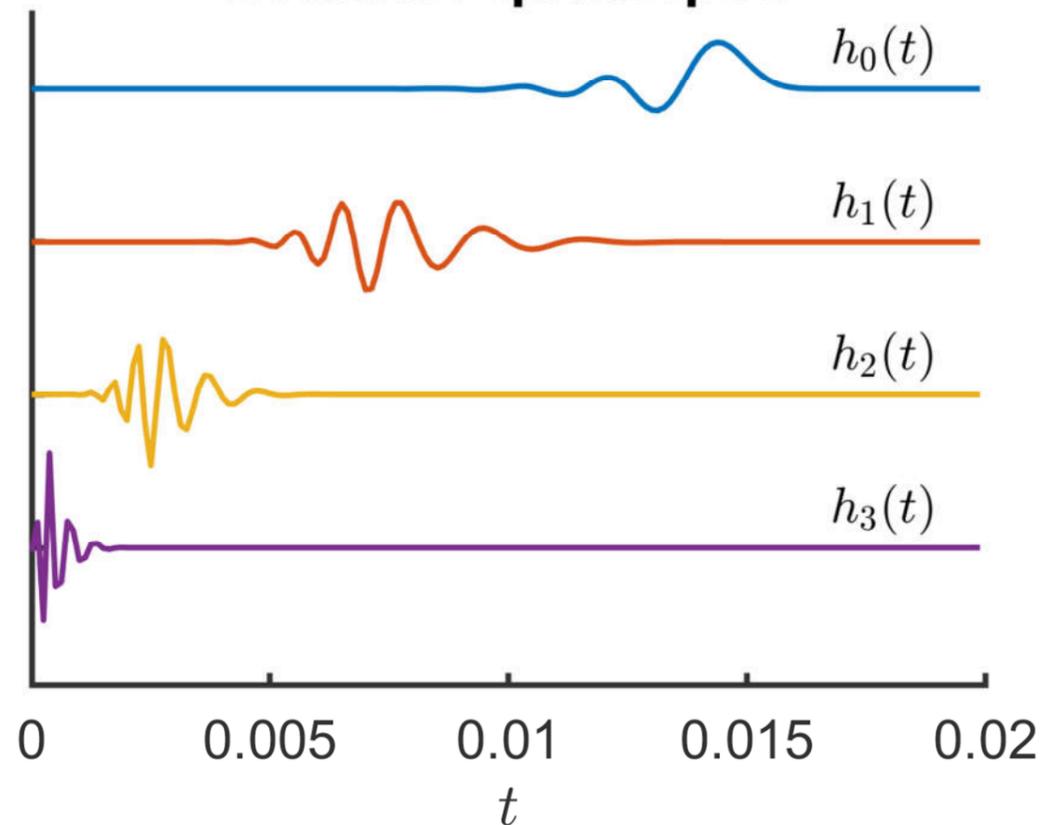
✓ Моделируется **временное** разрешение слуха.

✓ Относительно низкая вычислительная сложность.

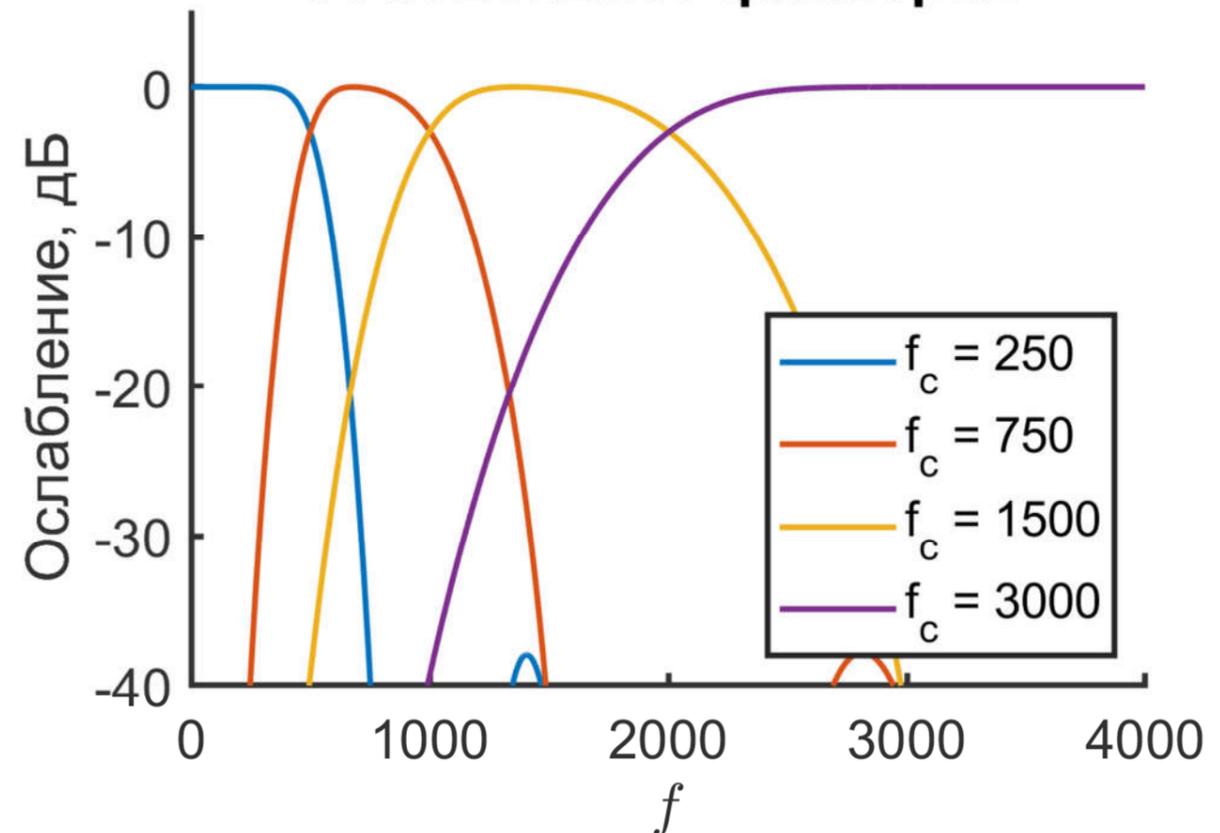
✓ Жесткая привязка частотного

разрешения к структуре банка фильтров.

Вейвлет-фильтров



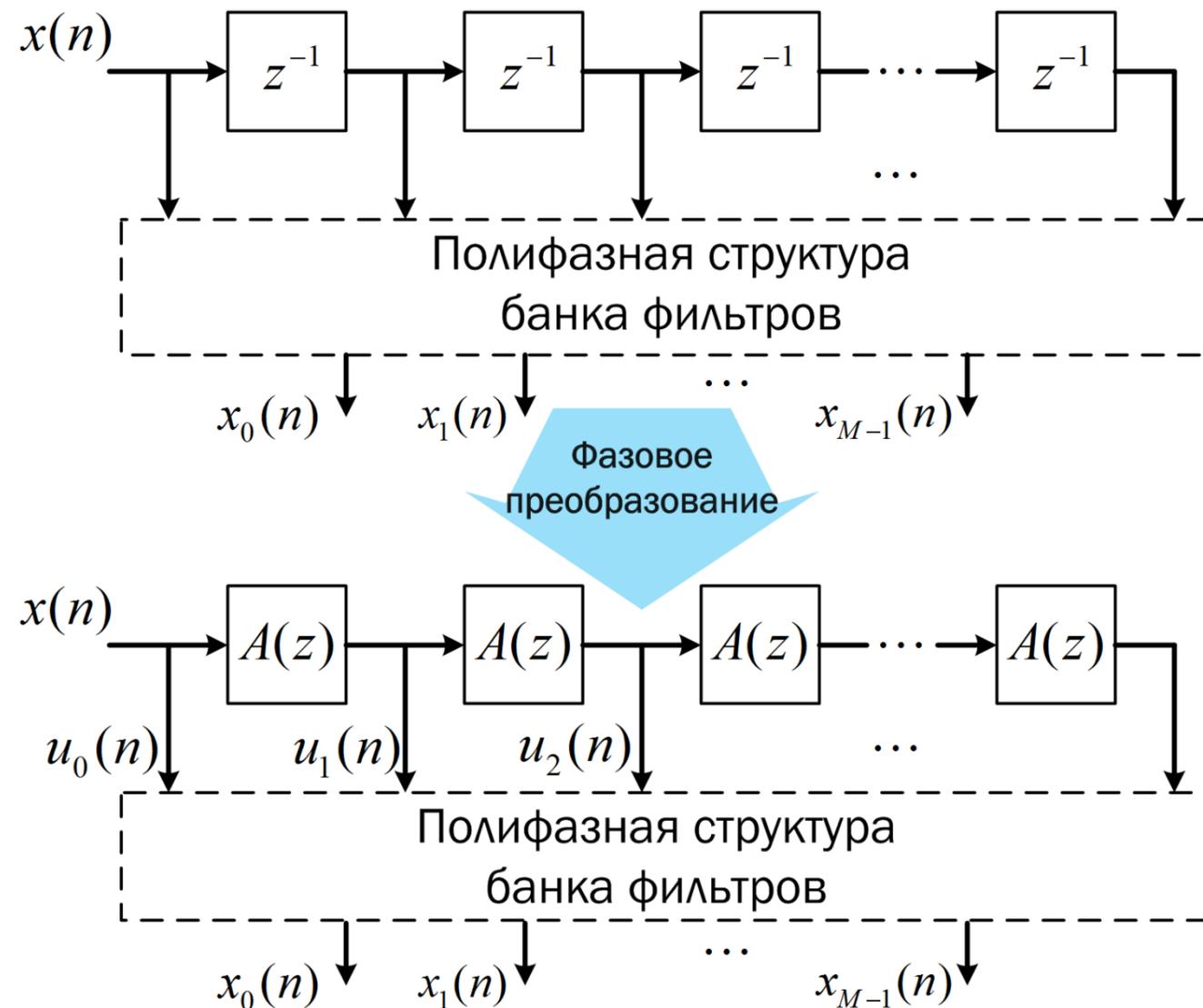
АЧХ вейвлет-фильтров



# Банк фильтров на основе фазового преобразования

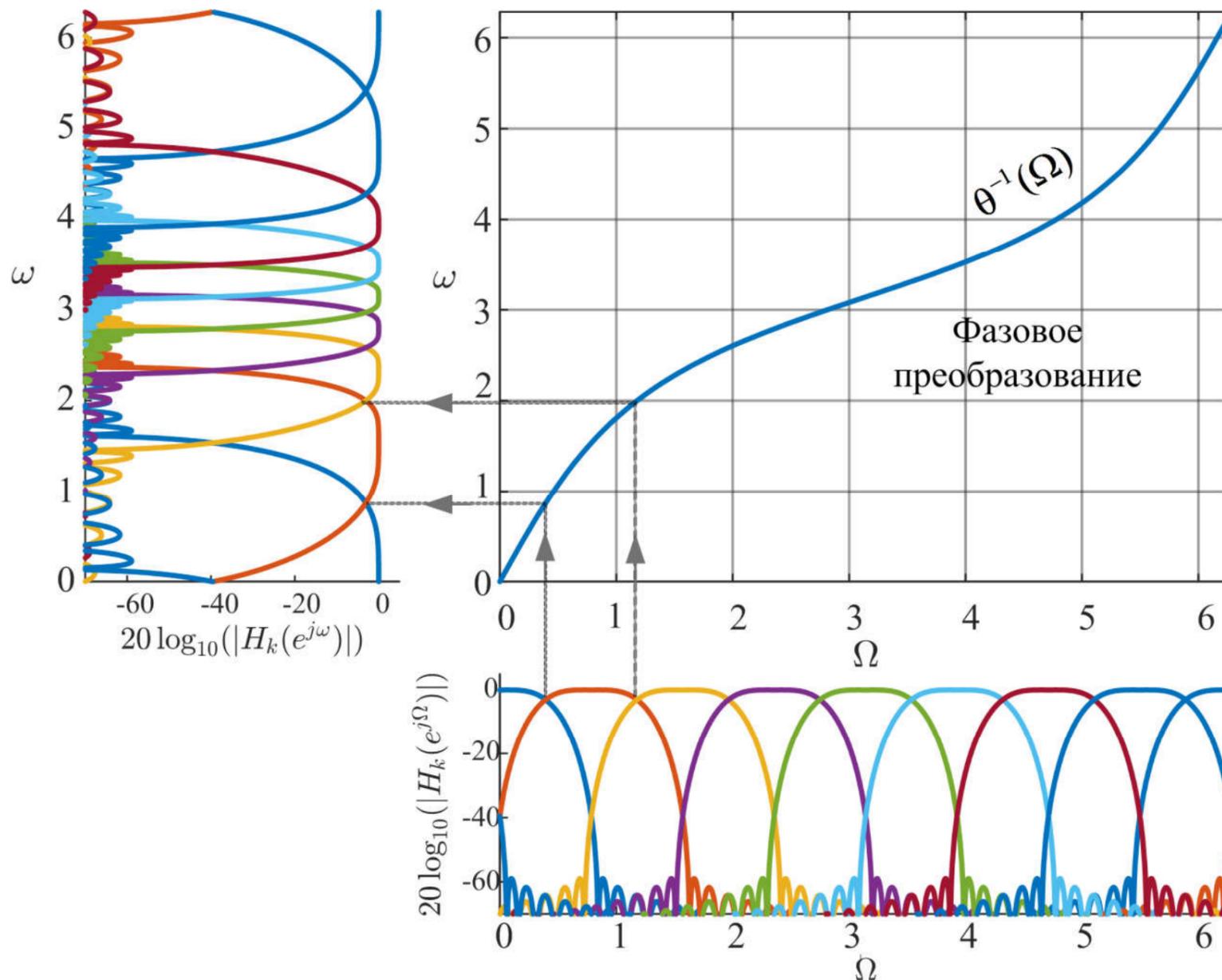
Фазовое преобразование позволяет трансформировать **равнополосный** банк фильтров в **неравнополосный**.

$$z^{-1} \rightarrow A(z) = \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}}, \quad A(e^{j\omega}) = e^{j\theta(\omega)}.$$



# Банк фильтров на основе фазового преобразования

Применение *фазового преобразования* к дискретной системе приводит к трансформации оси частот:



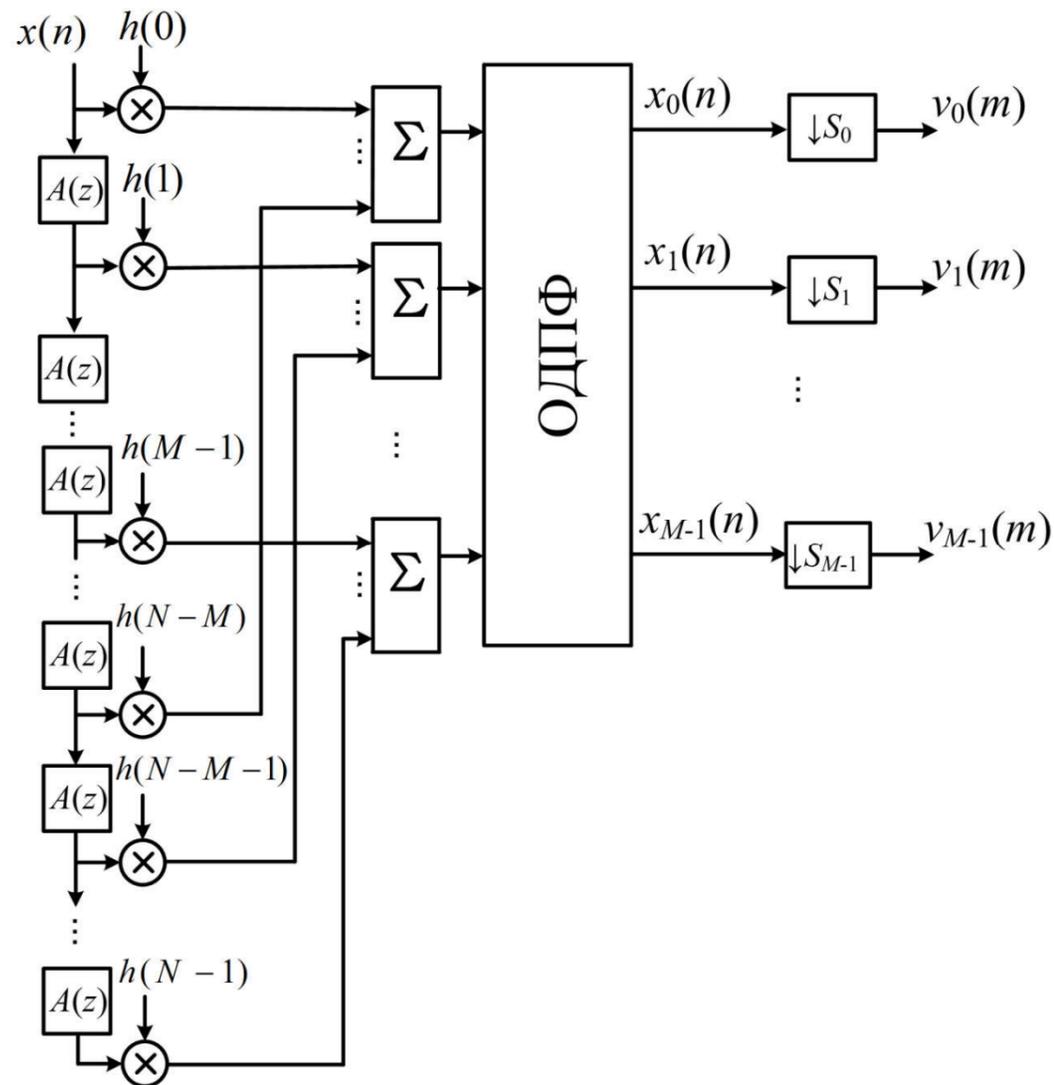
$$\theta^{-1}: \Omega \rightarrow \omega,$$

где

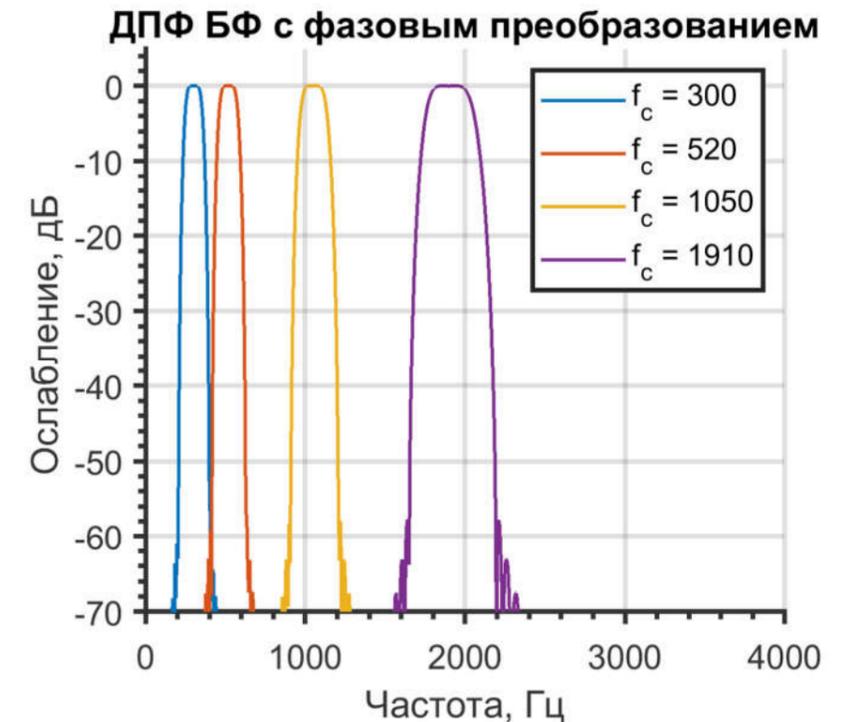
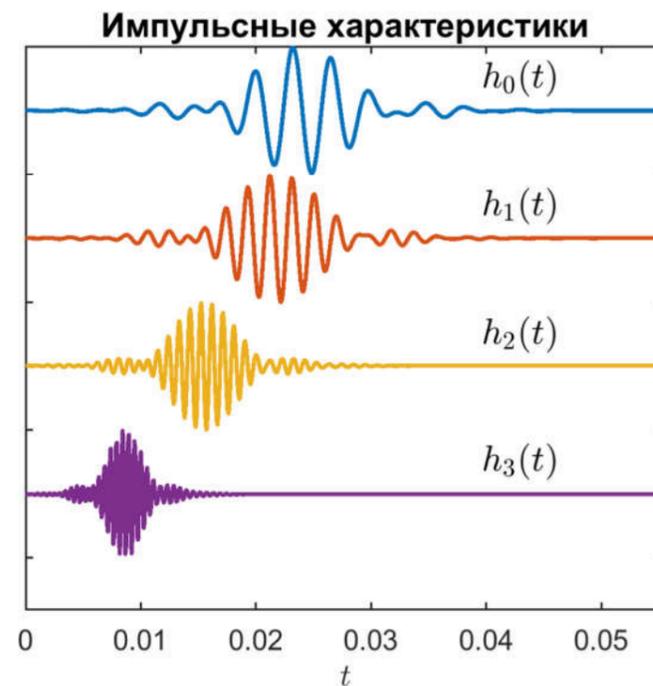
$$\theta(\omega) = \omega + \operatorname{arctg} \left( \frac{\alpha \sin \omega}{1 - \alpha \cos \omega} \right).$$

$\alpha$  – регулирует степень деформации частотной оси.

# Банк фильтров на основе фазового преобразования

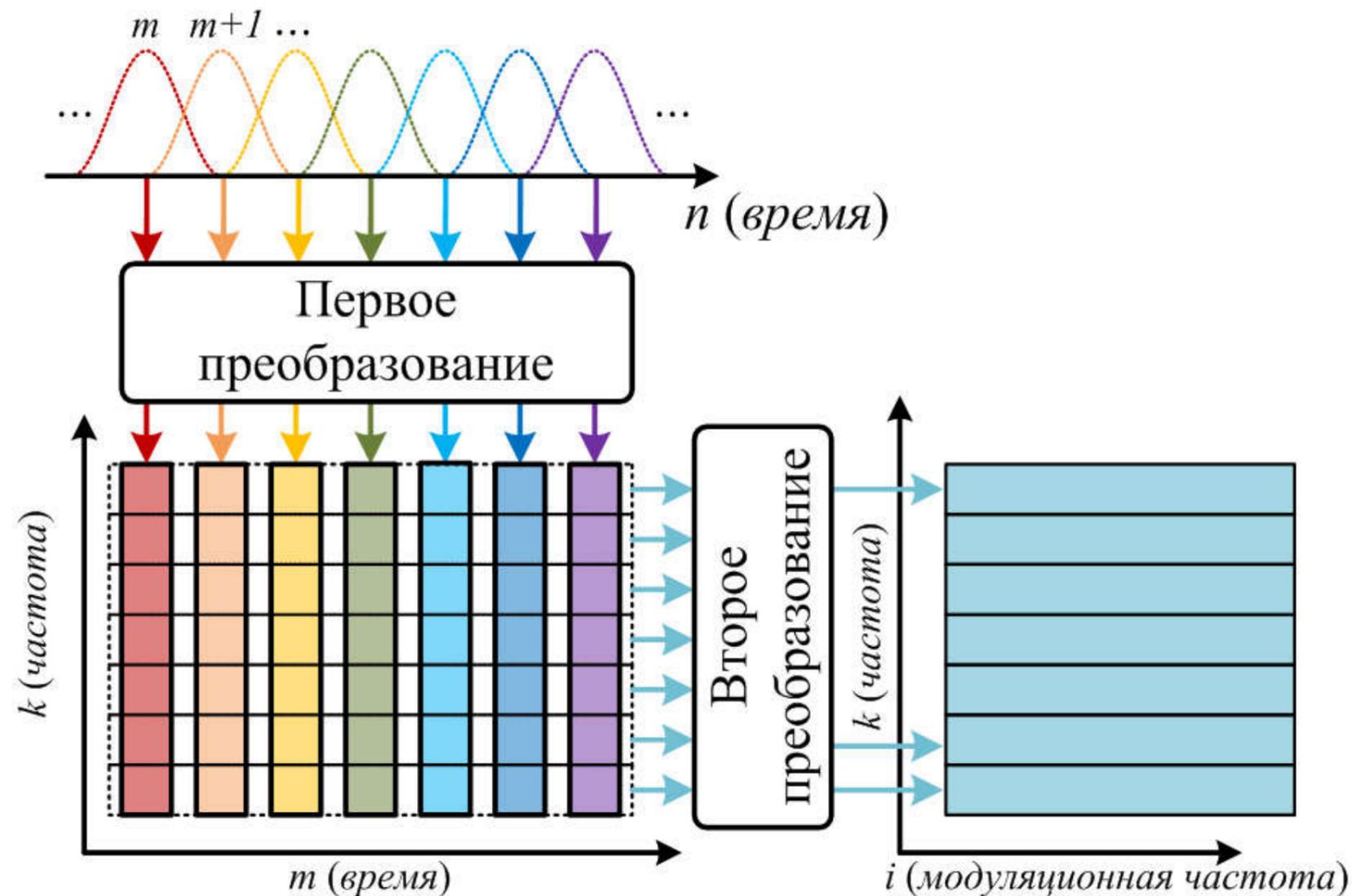


- ✓ Моделируется **частотное и временное** разрешение слуха.
- ✓ Относительно низкая вычислительная сложность.
- ✓ Нет жесткой привязки частотного разрешения к структуре банка фильтров.

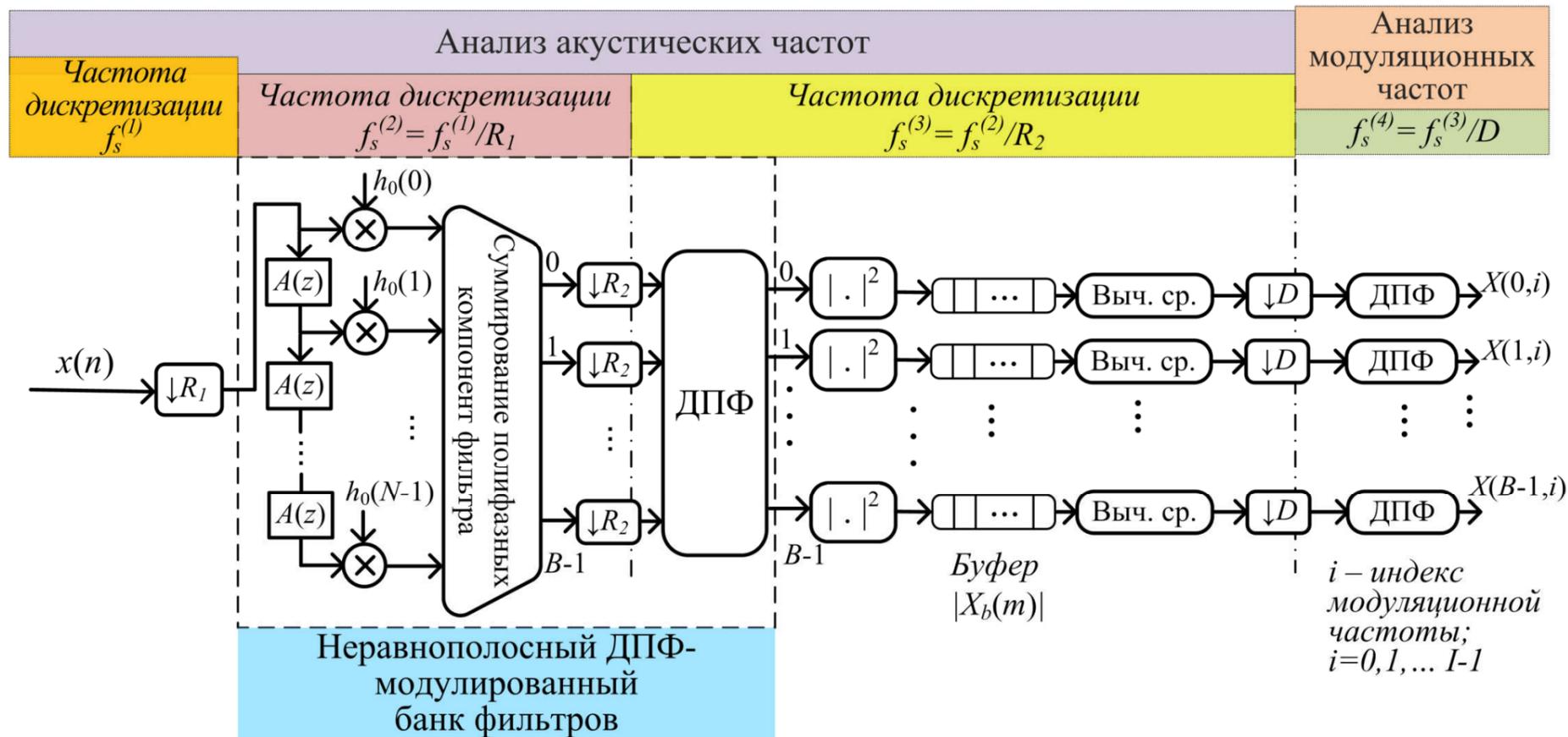
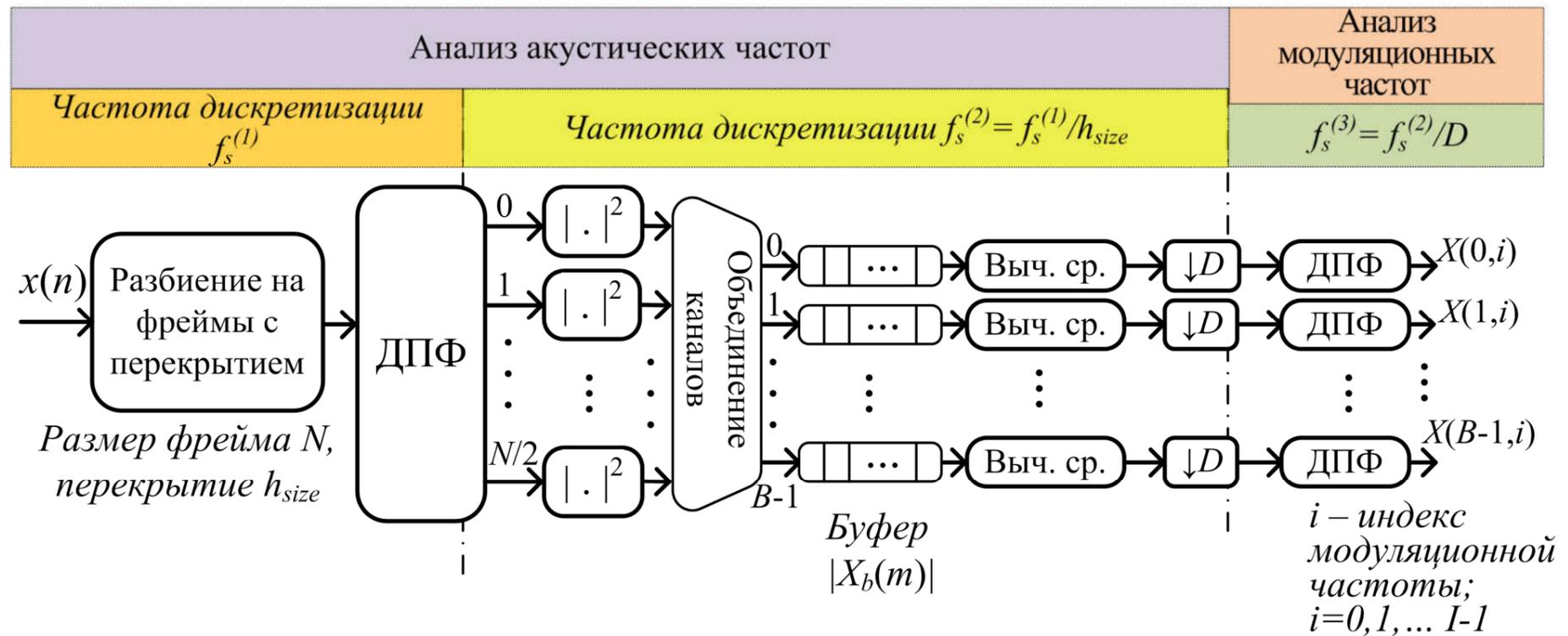


# Определение патологии голоса

Исследовался вопрос детектирования патологии патологии голосового аппарата на основе анализа *модуляционного спектра* речи в *критических* *полосах*.

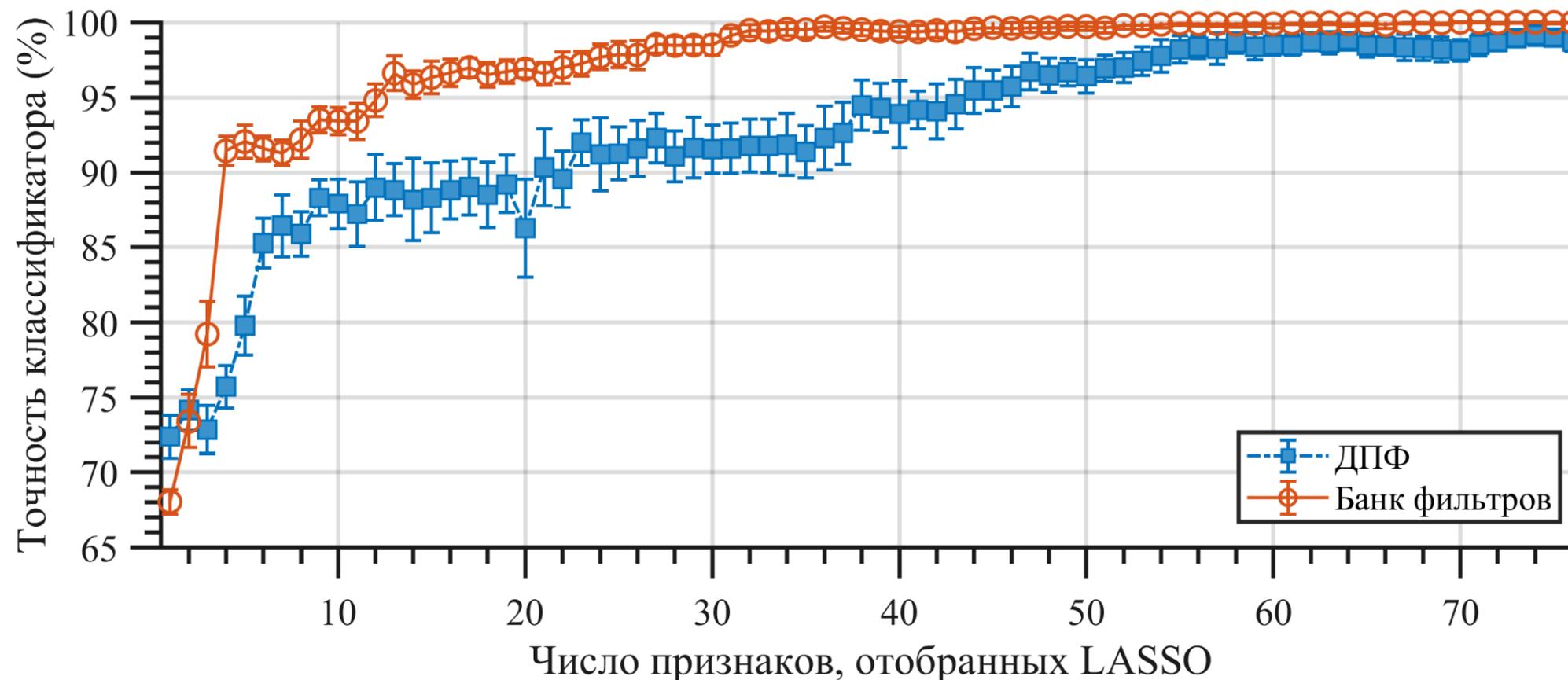


# Схемы анализа



# Результаты классификации

Задача: *выявление патологии гортани*. База голосов: 30 записей голосов с различными патологиями гортани + 30 здоровых голосов.



Вывод: *параметры выделяемые схемой на основе банка фильтров с фазовым преобразованием более эффективны.*

# Заключение

## 1) **Биоинспирированные методы** анализа:

- *гоех-фильтры;*
- *гамматон фильтры;*
- *гаммачирп фильтры;*
- модель Кейтса

## 2) **Психоакустически мотивированные методы** анализа:

- *ДПФ с объединением в полосах;*
- *банк фильтров на основе вейвлетов;*
- *банк фильтров на основе фазового преобразования;*

**3) Применение** психоакустически мотивированных методов анализа звукового сигнала к задаче определения патологии голоса.

**СПАСИБО ЗА ВНИМАНИЕ!**

