

Московский Государственный Университет  
имени М.В. Ломоносова



Факультет Вычислительной Математики и Кибернетики  
Кафедра Математических Методов Прогнозирования

ДИПЛОМНАЯ РАБОТА СТУДЕНТА 517 ГРУППЫ  
«Реализация спектрально-аналитического  
метода для описания звуковых сигналов»

Выполнил студент 5 курса  
517 группы

*Двойнев Александр Игоревич*

Научный руководитель:  
д.т.н., профессор

*Дедус Флоренц Фёдорович*

*Заведующий кафедрой  
Математических Методов  
Прогнозирования  
профессор, д. ф.-м. н.*

\_\_\_\_\_ *Ю. И. Журавлёв*

*К защите допускаю*

« \_\_\_\_\_ » \_\_\_\_\_ 2010 г.

*К защите рекомендую*

« \_\_\_\_\_ » \_\_\_\_\_ 2010 г.

Москва, 2010

# Содержание

<b>1</b>	<b>Введение</b>	<b>3</b>
<b>2</b>	<b>Общие сведения об ортогональных полиномах</b>	<b>5</b>
<b>3</b>	<b>Аппроксимация звуковых сигналов</b>	<b>6</b>
3.1	Реализация аппроксимации звукового сигнала с помощью дискретного косинусного преобразования . . . . .	7
3.2	Реализация аппроксимации звукового сигнала с помощью функций Майкнера . . . . .	7
3.2.1	Выбор семейства ортогональных полиномов . . . . .	7
3.2.2	Полиномы и функции Майкнера. Общие сведения.	8
3.2.3	Параметры регрессионной модели . . . . .	10
3.2.4	Рост погрешности ортонормированности при вычислении старших функций Майкнера . . . . .	12
3.2.5	Оптимизация модели . . . . .	14
3.3	Программная реализация . . . . .	17
<b>4</b>	<b>Сравнение работы ДКП и функций Майкнера</b>	<b>17</b>
4.1	Преимущества функций Майкнера . . . . .	18
4.1.1	Эффект Гиббса . . . . .	18
4.1.2	Описание затухающих сигналов . . . . .	19
4.2	Недостатки функций Майкнера . . . . .	20
4.2.1	Скорость сжатия и восстановления . . . . .	20
4.2.2	Гетероскедастичность ошибок . . . . .	20
4.3	Сравнение на реальных данных . . . . .	22
<b>5</b>	<b>Заключение</b>	<b>23</b>

# 1 Введение

Большинство современных аудио форматов используют дискретное косинусное преобразование (ДКП) для сжатия звуковых данных с потерями. Однако аппроксимация звукового сигнала рядами Фурье обладает рядом недостатков. Одним из главных недостатков является эффект «пре-эхо», который появляется из-за эффекта Гиббса. Также при описании некоторых звуковых сигналов недостатком дискретного косинусного преобразования является периодичность косинусов: дискретное косинусное преобразование идеально подходит для описания периодичных сигналов, но звуковые сигналы редко являются такими.

Классические ортогональные полиномы дискретного аргумента намного меньше подвержены эффекту Гиббса, чем дискретное косинусное преобразование. Также они не являются периодичными и теоретически должны лучше описывать непериодичные сигналы, чем косинусы. Основная идея данной работы состоит в том, что, если окажется, что ортогональные полиномы описывают звуковые сигналы лучше косинусов, то, если в аудио-кодеке в блоке сжатия с потерями заменить дискретное косинусное преобразование на аппроксимацию с помощью ортогональных полиномов, то кодек будет работать лучше.

В последнее время проводятся исследования по поиску наилучшей альтернативы ДКП [5], но вопрос всё еще остается открытым. Нельзя утверждать наперед, что полиномы будут описывать звук лучше косинусов. В данной работе установлено, что полифоническую музыку лучше аппроксимировать с помощью дискретного косинусного преобразования, но человеческую речь - с помощью ортогональных функций Майкснера.

## Цели работы:

- Изучение недостатков современных аудио-кодеков.
- Анализ аппроксимативных свойств функций Майкснера: параметры, погрешность.
- Сравнительная оценка описания звуковых сигналов с помощью дискретного косинусного преобразования и функций Майкснера.

## План работы:

1. Введение
2. Ознакомление с обобщенным спектрально-аналитическим методом
3. Составление алгоритма аппроксимации звуковых сигналов с помощью функций Майкснера. Оптимизация параметров базиса.
4. Разработка и реализация соответствующего алгоритма аппроксимации звуковых сигналов с помощью функций Майкснера.
5. Реализация алгоритма аппроксимации звуковых сигналов с помощью дискретного косинусного преобразования.
6. Сравнение аппроксимативных свойств функций Майкснера и дискретного косинусного преобразования
7. Заключение
8. Приложение. Программа, определяющая оптимальные параметры аппроксимации и число гармоник для описания данного звукового сигнала и заранее заданного порога качества, реализованная в среде MatLab 7.6.

## 2 Общие сведения об ортогональных полиномах

Пусть  $f(x)$  - непрерывная функция вещественного переменного. Функция  $f(x)$ ,  $x \in [a, b]$  принадлежит пространству  $L^2[a, b]$ , если она интегрируема с квадратом:

$$\int_a^b f^2(t) d\alpha(t) < \infty, \quad (1)$$

причем правая граница отрезка  $[a, b]$  не обязательно конечна.

Интеграл в формуле (1) является интегралом Стильеса. Рассмотрим пространство  $L_d^2[a, b]$  - множество функций *дискретного* аргумента, интегрируемых с квадратом на отрезке  $[a, b]$ . В нём  $\alpha(t)$  является функцией скачков. Пусть  $\alpha(t)$  имеет скачки величины  $j_i$  в точках  $t = t_i$ , а  $j'(t_i) = j_i$ . Тогда интеграл Стильеса в условии (1) сводится к сумме:

$$\sum_i f^2(t_i) j(t_i) < \infty. \quad (2)$$

Далее будем рассматривать только функции из пространства  $L_d^2[a, b]$ . Отметим, что функция скачков  $\alpha(t)$  положительна, причем  $\sum_i j(t_i)$  конечна, а точки  $t_i$  являются целыми и заданными на равномерной сетке на  $[a, b]$ .

Зададим в  $L_d^2[a, b]$  скалярное произведение:

$$(\varphi_1(t), \varphi_2(t)) = \int_a^b \varphi_1(t) \varphi_2(t) d\alpha(t) = \sum_i \varphi_1(t_i) \varphi_2(t_i) j(t_i). \quad (3)$$

Система функций  $\{\varphi_i(t)\}_{i=0}^N$  является ортонормированной, если

$$(\varphi_n, \varphi_m) = \begin{cases} 0, & \text{если } m \neq n; \\ 1, & \text{если } n = m. \end{cases} \quad (4)$$

Возьмем в  $L_d^2[a, b]$  базисную систему ортонормированных функций  $\{\varphi_i(t)\}$ . Разложим функцию  $f(t) \in L_d^2[a, b]$  по этому базису:

$$f(t) = \sum_n C_n \varphi_n(t). \quad (5)$$

Слагаемые суммы из формулы (5) будем называть *гармониками*. Найдем коэффициенты  $C_n$ :

$$\begin{aligned}
(f(t), \varphi(t)) &= \sum_{i=a}^b j(t_i) \varphi_n(t) \cdot \sum_k C_k \varphi_k(t) = \\
&= \sum_{i=a}^b j(t_i) \varphi_n(t) \varphi_n(t) C_n = C_n
\end{aligned} \tag{6}$$

Для аппроксимации разобьем формулу (5) на две суммы:

$$f(t) = \sum_{i=0}^k C_i \varphi_i(t) + \sum_{i=k+1}^n C_i \varphi_i(t) = \tilde{f}_n(t) + \varepsilon(t), \tag{7}$$

где  $\tilde{f}_n(t)$  - кривая регрессии с  $n$  гармониками,  $\varepsilon(t) = f(t) - \tilde{f}_n(t)$  - ошибка аппроксимации.

Заметим, что благодаря ортогональности системы функций  $\{\varphi_i\}$  разложение по коэффициентам  $C_i$  является «жестким», то есть добавление новых гармоник для повышения точности аппроксимации не будет влиять на ранее вычисленные коэффициенты.

При аппроксимации классические ортогональные полиномы отвечают методу наименьших квадратов:

$$\sum_t \left( (f(t) - \tilde{f}_n(t))^2 \right) = \sum_t \varepsilon^2(t) \rightarrow \min_n \tag{8}$$

Если в системе ортонормированных полиномов функция скачков  $j(t)$  не единичная, то такой системе обычно сопоставляют систему функций  $\{\psi_i(t)\}$ :

$$\psi_i(t) = \sqrt{j(t)} \varphi_i(t) \tag{9}$$

Нетрудно видеть, что получившаяся система является ортонормированной с единичной функцией скачков.

### 3 Аппроксимация звуковых сигналов

Любой звуковой сигнал можно рассматривать как функцию зависимости амплитуды колебания от времени. Именно в таком виде звуковые данные хранятся в формате **\*.wav**.

Пусть  $f(t)$  - звуковой сигнал, где  $t = 0, \dots, T - 1$  - время, а значения  $f$  - амплитуда. Аппроксимируем функцию с помощью семейства ортонормированных функций  $\{\varphi_i(t)\}$ :

$$f(t) \approx \tilde{f}_n(t) = \sum_{i=0}^n C_i \varphi_i(t) \quad (10)$$

В качестве  $\{\varphi_i(t)\}$  рассмотрим семейства ортогональных косинусов и функции Майкснера и сравним их аппроксимативные свойства.

### 3.1 Реализация аппроксимации звукового сигнала с помощью дискретного косинусного преобразования

В большинстве современных аудио-кодеков, в частности в наиболее распространенном сегодня mp3, используется дискретное косинусное преобразование [9]. Поэтому аппроксимация с его помощью было взято как эталонное.

коэффициенты ДКП находят по формулам:

$$C_0 = \frac{1}{\sqrt{2T}} \sum_{t=0}^{T-1} f(t) \quad (11)$$

$$C_k = \sqrt{\frac{2}{T}} \sum_{t=0}^{T-1} f(t) \cos\left(\frac{2t+1}{2T}k\pi\right), \quad k \in [1, n-1] \quad (12)$$

Восстановление сигнала проводится по формуле:

$$\tilde{f}_n(t) = \sqrt{\frac{2}{T}} \sum_{k=0}^{n-1} C_k \cos\left(\frac{2t+1}{2T}k\pi\right), \quad (13)$$

где  $n$  - число гармоник. Чем больше  $n$ , тем точнее аппроксимация. При  $n = T$  получаем интерполяцию:  $f(t) = \tilde{f}_T(t)$ .

### 3.2 Реализация аппроксимации звукового сигнала с помощью функций Майкснера

#### 3.2.1 Выбор семейства ортогональных полиномов

Для того, чтобы полиномы аппроксимировали звуковой сигнал не хуже, чем ДКП, нужно предъявить к ним следующие требования:

- Полиномы должны быть дискретного аргумента, чтобы не проигрывать ДКП в скорости за счёт использования равномерной сетки представления.

- Полиномы должны быть затухающими, так как звуковой сигнал часто имеет затухающий вид.

Этим требованиям удовлетворяют следующие семейства полиномов:

1. функции Майкснера  $M_n^{\beta,c}(t)$  при всех значениях  $\beta$  и  $c$ ;
2. Функции Кравчука  $K_n^{p,q}(t)$  при некоторых значениях  $p$  и  $q$ ;
3. функции Хана  $H_n^{\alpha,\beta}(t)$  при некоторых значениях  $\alpha$  и  $\beta$ ;
4. функции Шарлье  $C_n(t, m)$  при некоторых значениях  $m$ .

Однако, функции Кравчука имеют периодический вид [1], поэтому их аппроксимативные свойства будут мало отличаться от дискретного косинусного преобразования. Функции Хана затухают не при всех параметрах. В итоге из четырех семейств было выбрано семейство функций Майкснера. Во-первых, функции Майкснера являются затухающими и непериодичными. Во-вторых, для функций Майкснера существует рекурсивная формула, считающая сразу нормированные полиномы и позволяющая вычислять полиномы степени 1000 и выше.

### 3.2.2 Полиномы и функции Майкснера. Общие сведения.

Для полиномов Майкснера используется обозначение  $M_n^{\beta,c}(t)$ , где  $t \in [0, \infty)$  - дискретный аргумент на равномерной сетке. Так как реальные сигналы рассматриваются на конечной сетке, на практике полубесконечный интервал ортогональности заменяется конечным интервалом  $[0, T]$ . Параметры:  $\beta > 0$ ,  $c \in (0, 1)$ .

Полиномы определяются дискретным аналогом формулы Родрига:

$$M_n^{\beta,c}(t) = \frac{t! \Gamma(\beta)}{\Gamma(\beta + t)} c^{-n-t} \cdot \Delta^n \left[ c^t \frac{\Gamma(\beta + t)}{(t - n)! \Gamma(\beta)} \right] \quad (14)$$

Для увеличения скорости и точности значения полиномов лучше считать по рекуррентной формуле:

$$M_{n+1}^{\beta,c} = \frac{1}{c} (tc - t + n + nc + c\beta) M_n^{\beta,c}(t) - \frac{1}{c} n(n - 1 + \beta) M_{n-1}^{\beta,c}(t). \quad (15)$$

Функция скачков:

$$j(t) = \frac{c^t (\beta)_t}{t!} \quad (16)$$



где  $(\beta)_t = \beta(\beta + 1) \dots (\beta + t - 1)$ ;  $(\beta)_0 = 1$  - символ Похгаммера.  
 Норма полиномов Майкснера:

$$\|M_n^{\beta,c}\| = \sqrt{\frac{n!(\beta)_t}{c^n(1-c)^\beta}} \quad (17)$$

Заметим, что при вычислении нормы при  $n > 170$ , оперируя типом `double`, мы получим ошибку переполнения регистра. Это происходит из-за наличия факториала в формуле, так как  $170! \approx 7 \cdot 10^{306}$ , а верхний предел машинного представления числа `double` равен  $10^{308}$  знаков. Для вычисления полиномов Майкснера степени больше 170 существует нормированная рекуррентная формула [6], которая сразу считает нормированные полиномы  $\overline{M}_n^{\beta,c}(t)$ :

$$\overline{M}_{n+1}^{\beta,c}(t) = (tc - t + n + nc + c\beta) \sqrt{\frac{1}{c(n+1)(\beta+n)}} \overline{M}_n^{\beta,c}(t) - \quad (18)$$

$$- \sqrt{\frac{(n-1+\beta)n}{(n+1)(\beta+n)}} \overline{M}_{n-1}^{\beta,c}(t) \quad (19)$$

Функции Майкснера  $m_n^{\beta,c}(t)$  получаются из полиномов по формуле (9):

$$m_n^{\beta,c} = M_n^{\beta,c}(t) \cdot \sqrt{\frac{c^t(\beta)_t}{t!}}. \quad (20)$$

Функция скачков для системы функций Майкснера единичная.

Вычисление коэффициентов разложения функции  $f(t)$  по функциям Майкснера выполняется по формуле:

$$E_k = \sum_{t=0}^T f(t) \cdot m_n^{\beta,c} \quad (21)$$

Кривая регрессии вычисляется по формуле:

$$\widetilde{f}_n^{\beta,c}(t) = \sum_{k=0}^n E_k \cdot m_k^{\beta,c}(t). \quad (22)$$

### 3.2.3 Параметры регрессионной модели

Итак, мы получили регрессионную модель  $m_n^{\beta,c}$ . Как видно из обозначения, у модели 3 параметра: параметры функций  $\beta$  и  $c$ , и число гармоник  $n$ . Рассмотрим каждый параметр отдельно.

**Параметр  $\beta$ .** Как видно из рисунков (1) и (2), чем больше параметр  $\beta$ , тем дальше нули функций смещены вправо и тем больше расстояние между ними.

**Параметр  $c$ .** Как видно из рисунков (1) и (3), чем больше параметр  $c$ , тем медленнее затухают функции.

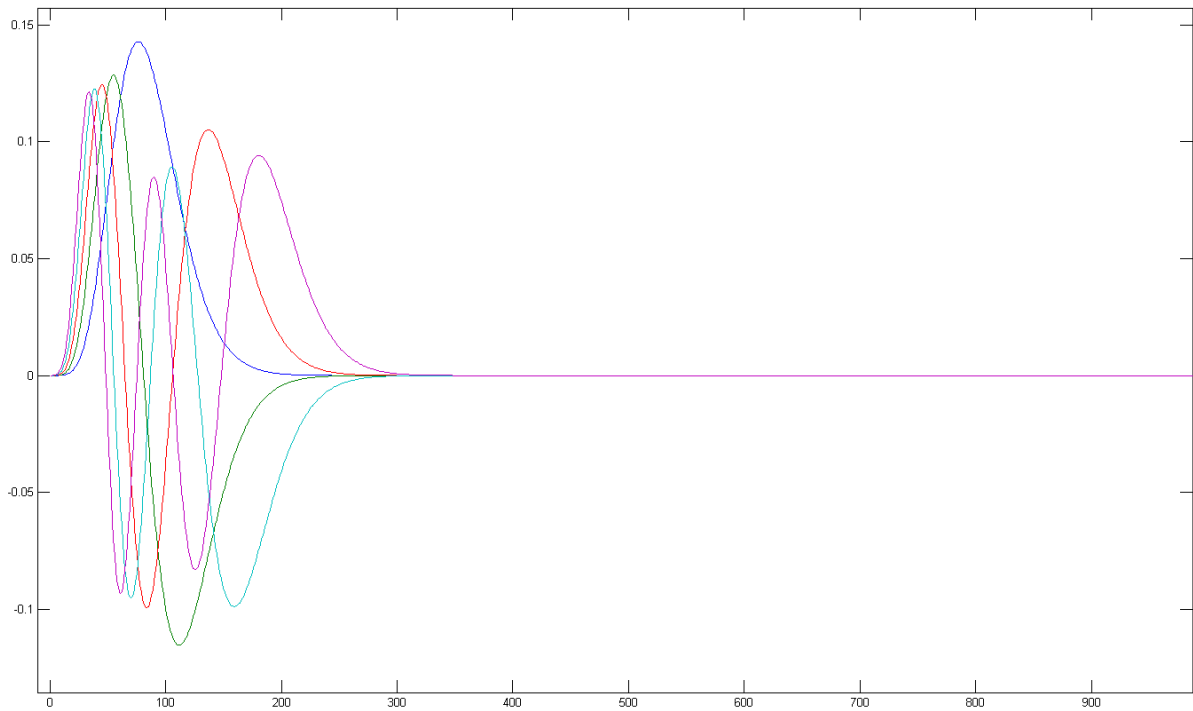


Рис. 1:  $\beta=20$ ,  $c=0.8$

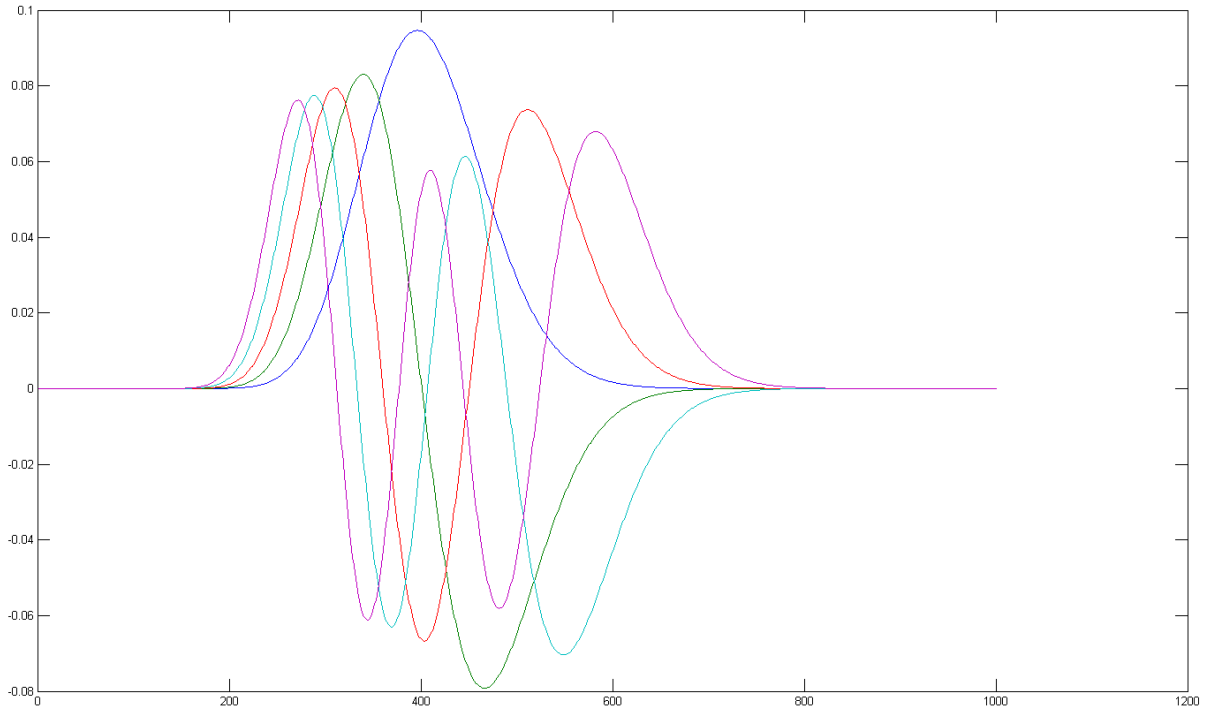


Рис. 2:  $\beta=100$ ,  $c=0.8$

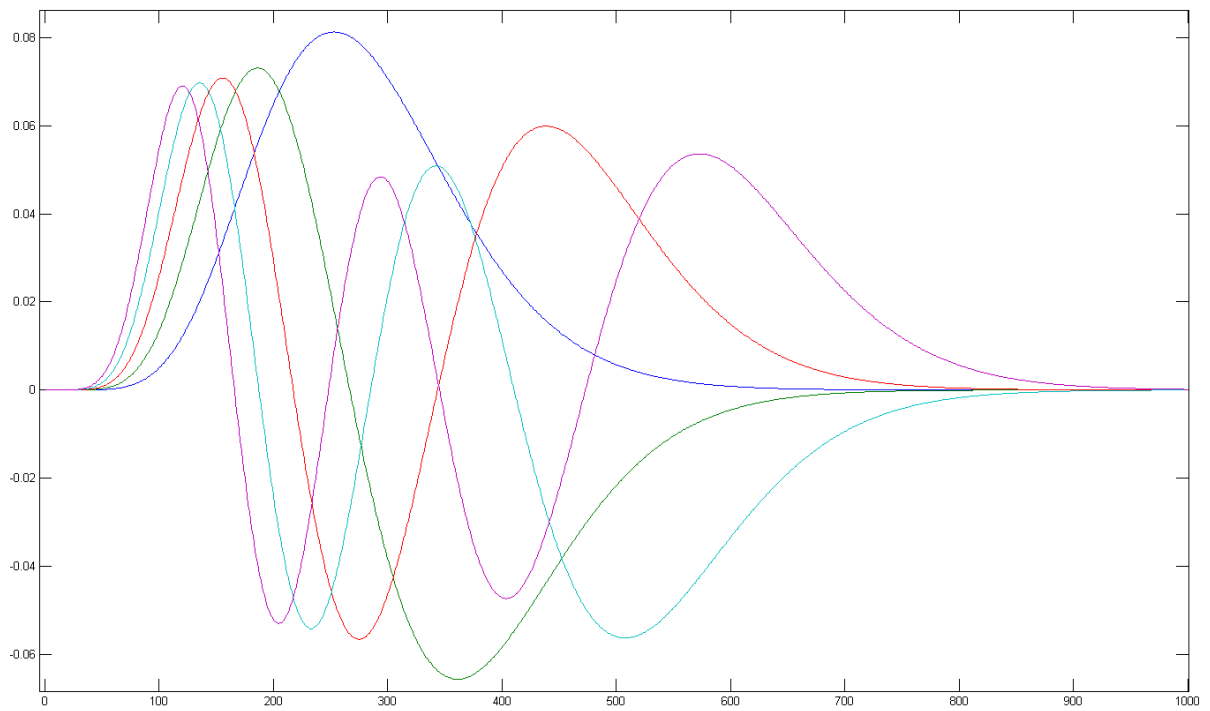


Рис. 3:  $\beta=20$ ,  $c=0.93$

### 3.2.4 Рост погрешности ортонормированности при вычислении старших функций Майкснера

Из-за ошибок округления ортонормированность функций Майкснера падает с ростом степени. Также ортонормированность сильно зависит от параметров  $\beta$  и  $c$ . Рассмотрим две модели:  $m_n^{260, 0.81}$  и  $m_n^{360, 0.25}$ .

На графиках (4) и (5) изображены графики зависимости нормы функций от степени. На графике (4) видно, что после  $n = 230$  норма начинает отклоняться от единицы. Это происходит из-за ошибки округления. У функций с параметрами  $\beta = 260$  и  $c = 0.81$ , как видно из графика (5), норма не равна единице ни при каких  $n$ . Причиной тому может быть переполнение регистра при вычислении полиномов или функции скачков.

Для анализа ортогональности построим матрицу Грама. На рисунках (6) и (7) изображены соответствующие матрицы. При этом белому цвету соответствует значение 0, а черному - 0.3. Видно, что в первом случае ортогональность начинает нарушаться только при  $n > 230$ . Во втором же случае ортогональность нарушается даже при малых  $n$ . Причем чем ближе степени функций, тем больше ошибка ортогональности.

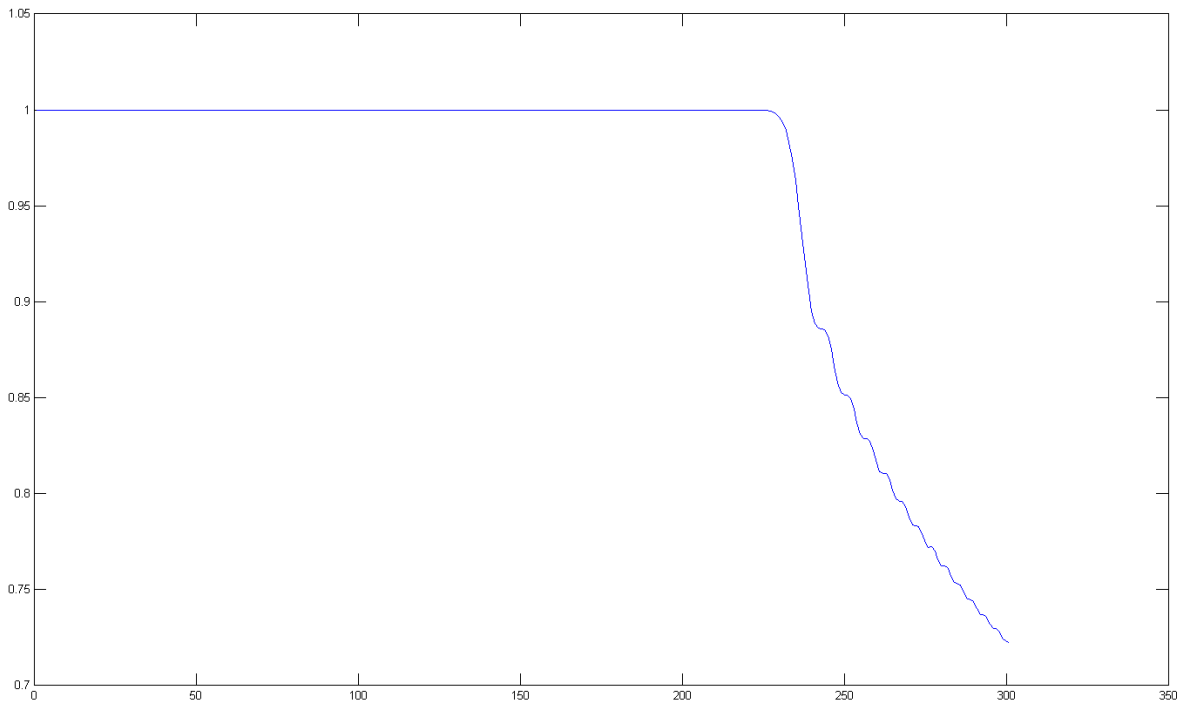


Рис. 4: Норма функций Майкснера при  $\beta=360, c=0.25$

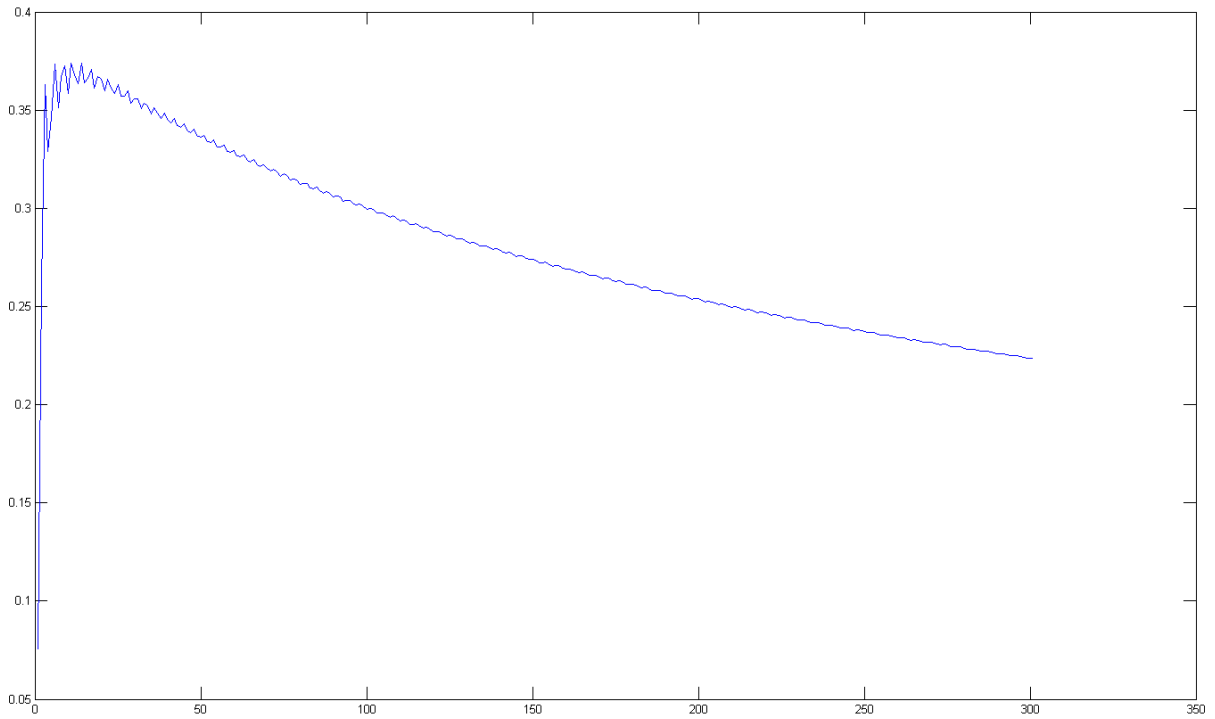


Рис. 5: Норма функций Майкснера при  $\beta=260$ ,  $c=0.81$

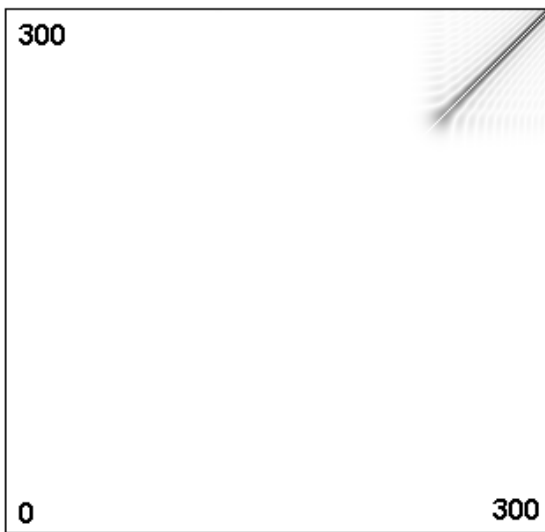


Рис. 6: Матрица Грама при  $\beta=360$ ,  $c=0.25$

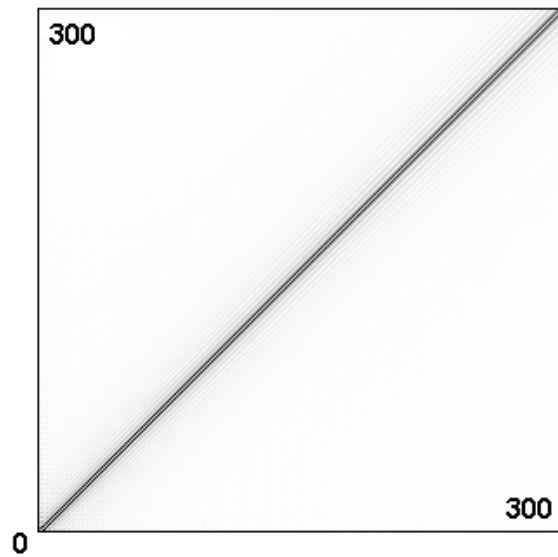


Рис. 7: Матрица Грама при  $\beta=260$ ,  $c=0.81$

### 3.2.5 Оптимизация модели

Так как функции Майкснера удовлетворяют МНК (8), в качестве функционала качества для оптимизации модели возьмем сумму квадратов ошибок:

$$Q(t, \beta, c, n) = \sum_t \left( f(t) - \widetilde{f}_n^{\beta, c}(t) \right)^2 = \sum_t \varepsilon^2(t) \rightarrow \min_{n, \beta, c}. \quad (23)$$

Как говорилось ранее, получившаяся модель регрессии оптимизируется по трём параметрам: параметрам функций Майкснера  $\beta$  и  $c$  и числу гармоник  $n$ .

**Верхняя граница параметра  $\beta$ .** В определении полиномов Майкснера параметр  $\beta$  может принимать любое положительное значение. Однако на практике следует ограничить параметр сверху. Во-первых, граница сверху облегчит оптимизацию модели по параметру. Во-вторых, если  $\beta$  больше длины сигнала, функции Майкснера не способны адекватно описать этот сигнал.

Рассмотрим конкретный пример. На рисунке (8) изображена функция  $f(t)$  заданная на отрезке  $t \in [0, 100]$  и функции Майкснера  $m_n^{100, 0.6}$ . Видно, что основную информацию функции несут на отрезке  $t \in [50, 300]$ . Как отмечалось ранее, параметр  $\beta$  отвечает за сдвиг нулей функций Майкснера вправо. Из-за слишком большого  $\beta$  функции на отрезке  $t \in [0, 50]$  почти нулевые, поэтому на этом отрезке функции не способны адекватно аппроксимировать функцию. Поэтому далее будем брать  $\beta \in (0, T)$ .

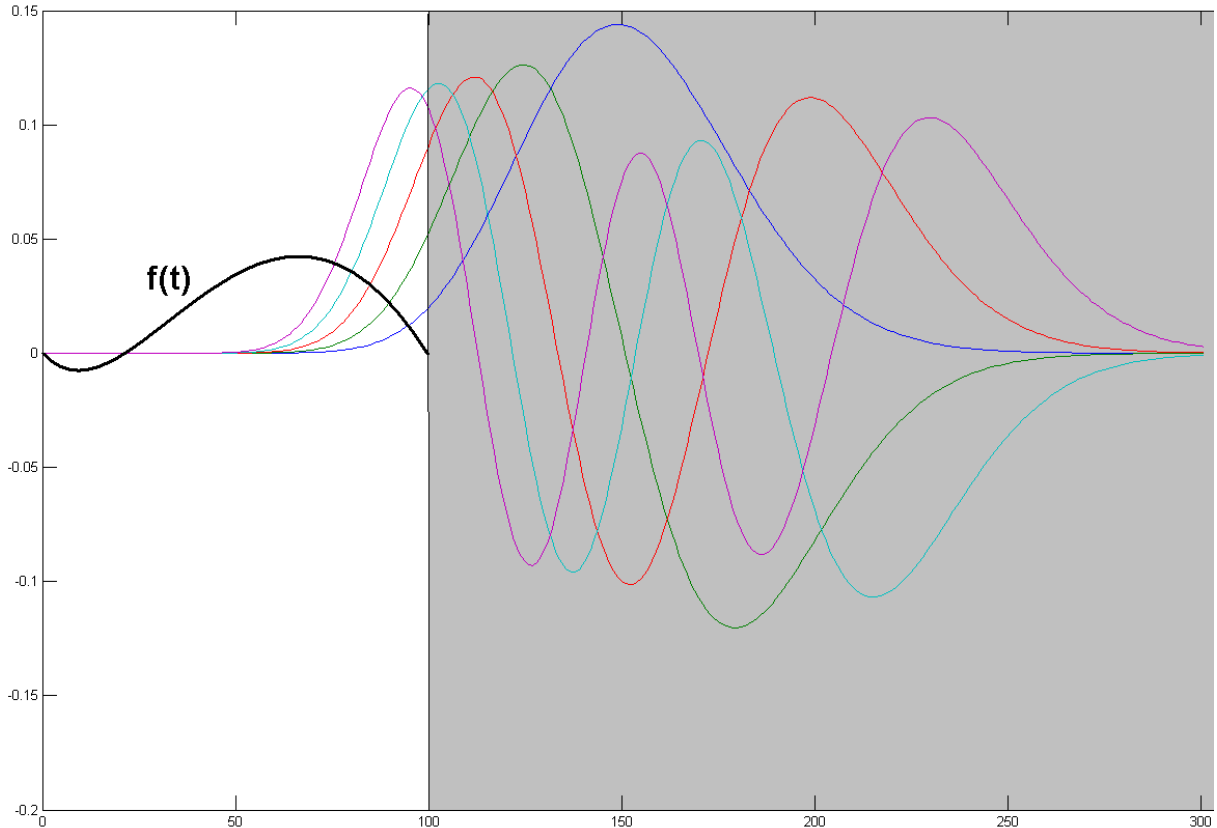


Рис. 8: Слишком большой сдвиг вправо

**Оптимизация по параметрам  $\beta$  и  $c$ .** Зафиксируем число гармоник  $n$  и рассмотрим зависимость функционала  $Q$  от  $\beta$  и  $c$ .

На рисунке (9) изображен пример графика зависимости функционала  $Q$  от  $\beta$  и  $c$ . Как и на этом рисунке, эта зависимость всегда имеет форму оврага. Такая унимодальность облегчает поиск минимума функционала  $Q$  по  $\beta$  и  $c$ , позволяя использовать градиентные методы поиска минимума. В частности, шаблонный Cross поиск [8], который находит минимум за логарифмическое время и который был использован в программной реализации метода.

**Оптимизация по  $n$ .** Очевидно, что функционал качества  $Q(t, \beta, c, n)$  монотонно убывает при росте числа гармоник  $n$ , и при  $n = T$  получится интерполяция, и в  $Q$  останутся только ошибки округления. Поэтому нужно задать пороговое значение  $\varepsilon_0$  и наращивать число гармоник, пока  $Q(t, \beta, c, n) > \varepsilon_0$ .

Казалось бы, что благодаря «жесткости» разложения сигнала по ортогональным функциям для того, чтобы увеличить степень регрессии

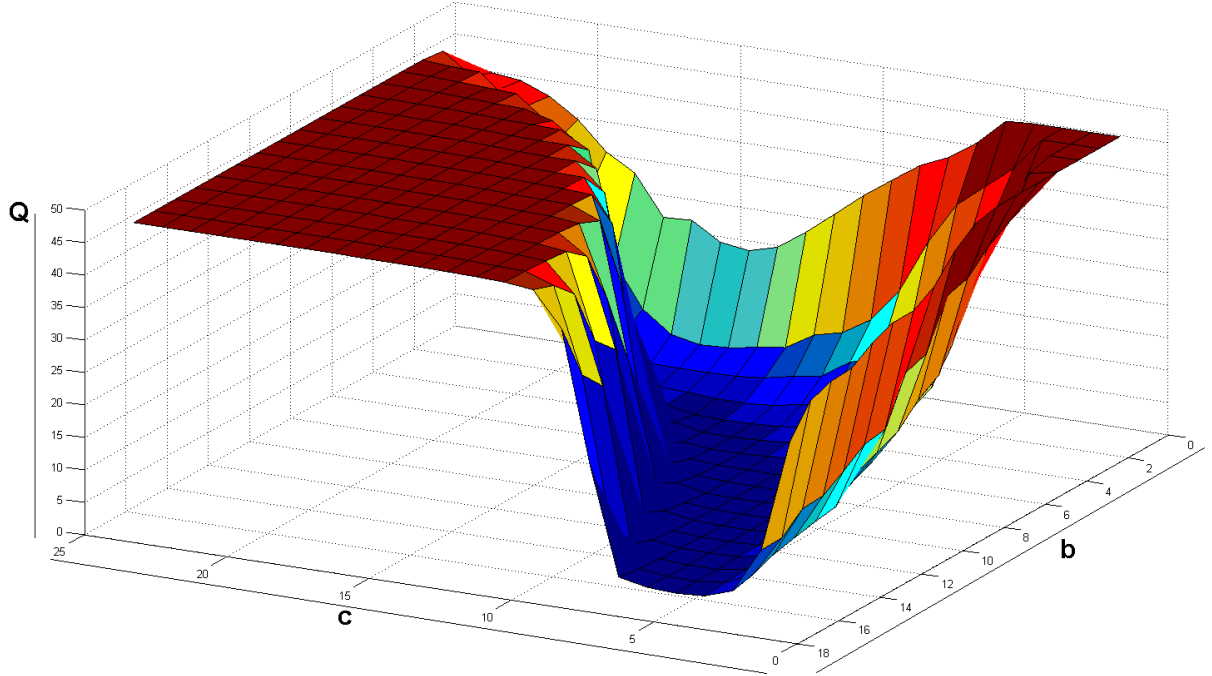


Рис. 9: График зависимости  $Q$  от  $\beta$  и  $c$

$\widetilde{f}_n^{\beta,c}(t)$ , достаточно добавить гармонику степени  $n + 1$ . Однако, наличие параметров  $\beta$  и  $c$  играет тут отрицательную роль, так как при увеличении степени мы можем перейти к другой системе функций с другими параметрами.

Рассмотрим пример. Пусть мы аппроксимировали сигнал  $f(t)$  и получили кривую регрессии  $\widetilde{f}_n^{\beta_1,c_1}(t)$  с оптимальными параметрами  $\beta_1$  и  $c_1$ . Прибавим к ней гармонику степени  $n + 1$  и получим  $\widetilde{f}_{n+1}^{\beta_1,c_1}(t)$ . Однако, для степени  $n + 1$  оптимальными могут быть другие параметры  $\beta_2$  и  $c_2$ , то есть модель  $\widetilde{f}_{n+1}^{\beta_2,c_2}(t)$  описывает сигнал лучше, чем  $\widetilde{f}_{n+1}^{\beta_1,c_1}(t)$  (что чаще встречается на практике). При этом к модели  $\widetilde{f}_n^{\beta_1,c_1}(t)$  нельзя прибавлять гармонику степени  $n + 1$  с параметрами  $\beta_2$  и  $c_2$ , так как функции Майкснера с разными параметрами не являются ортогональными.

Поэтому при вычислении модели следующей степени приходится вычислять все коэффициенты заново, что плохо сказывается на скорости оптимизации. В реальной реализации метода был использован двоичный поиск с логарифмической скоростью.



### 3.3 Программная реализация

На основе проведенного теоретического исследования на языке Matlab была написана программа, которая для сигнала  $f(x)$  и порогового значения  $\varepsilon_0$  вычисляет коэффициенты разложения сигнала по функциям Майкснера с оптимальными параметрами.

Структура программы:

1. Сегментирование сигнала
2. Разделение сигнала на отрезки, если тот слишком длинный
3. Нахождение оптимального числа гармоник  $n$  с помощью метода дихотомии.
4. Для каждого фиксированного  $n$  вычисление оптимальных параметров  $\beta$  и  $c$  с помощью двоичного Cross поиска.
5. Вычисление коэффициентов разложения сигнала по функциям Майкснера с оптимальными  $\beta$ ,  $c$  и  $n$ .

## 4 Сравнение работы ДКП и функций Майкснера

Пусть  $N_m$  - минимальное число гармоник, требуемое для описания сигнала  $f(t)$  с суммарной квадратичной ошибкой  $\varepsilon_0$  с помощью функций Майкснера, а  $N_c$  - с помощью дискретного косинусного преобразования. Для сравнения качества сжатия звукового сигнала функциями Майкснера и косинусами будем рассматривать коэффициент

$$K = \frac{N_c}{N_m}. \quad (24)$$

Если  $K > 1$ , то для описания сигнала функциями Майкснера нужна степень меньше, чем косинусам, и поэтому сжатие с помощью функций Майкснера лучше. Если  $K < 1$ , то сжатие с помощью дискретного косинусного преобразования лучше.

Здесь и далее под сжатием будем понимать преобразование набора значений сигнала в набор коэффициентов, под восстановлением - преобразование набора коэффициентов в набор значений приближенной функции.

## 4.1 Преимущества функций Майкснера

### 4.1.1 Эффект Гиббса

Одним из главных недостатков формата mp3 является так называемый «эффект преехо», который возникает из-за эффекта Гиббса. Эффектом Гиббса называются волнообразные колебания аппроксимирующей функции перед и после скачка сигнала.

К сожалению, эффект Гиббса невозможно и даже ослабить лишь увеличением числа гармоник. В этом случае просто возрастает частота волнообразных колебаний, но их относительная амплитуда меняется незначительно [3]. Эффект Гиббса можно устранить лишь интерполяцией.

Рассмотрим рисунок (10). На нём синим цветом изображен график звукового сигнала  $f(x)$  - щелчка пальцем. Красным цветом изображен график аппроксимация с помощью функций Майкснера, чёрным - с помощью ДКП с одинаковым числом гармоник. Из графика аппроксимации с помощью косинусов видно, что ей надо «разогнаться» перед скачком. Такие волнообразные колебания и вызывают «эффект преехо». Такого рода сигналы возникают чаще всего при описании с помощью ДКП речи.

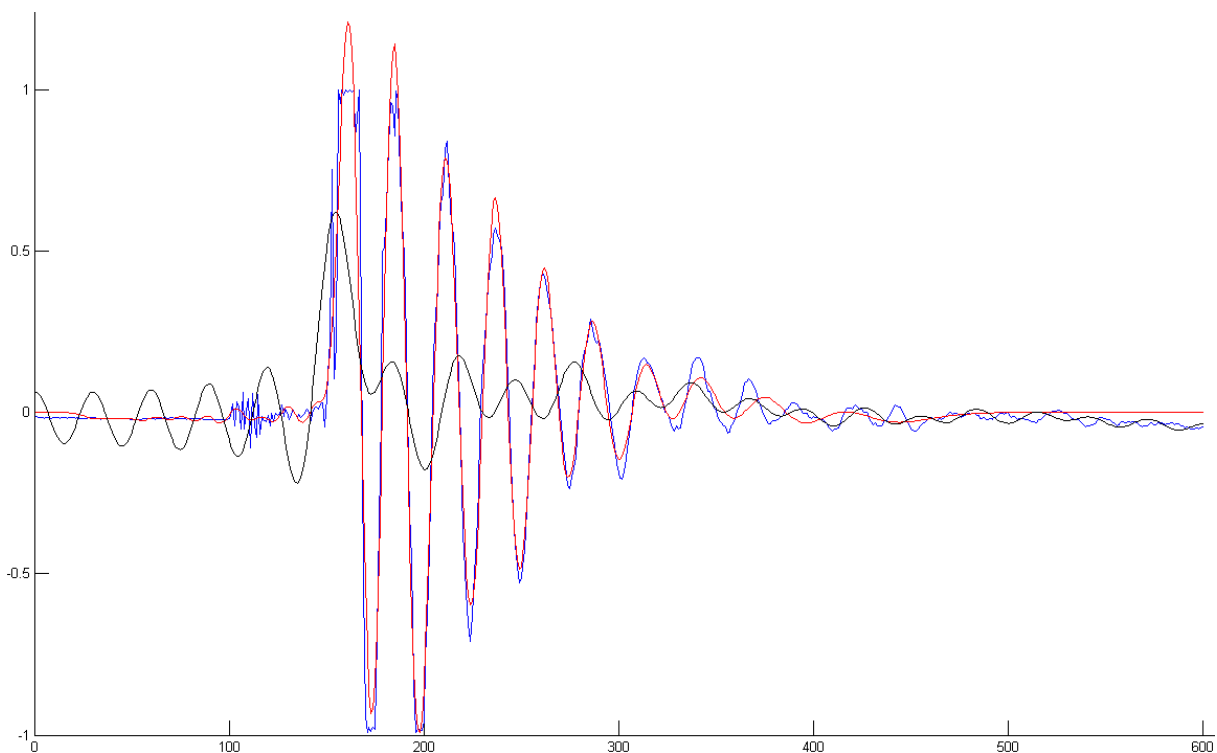


Рис. 10: Эффект Гиббса

При описании полифонической музыки эффект Гиббса выражен не так сильно.

Также из графика видно, что функции Майкснера не подвержены эффекту Гиббса. Следовательно, функции Майкснера лучше подходят для описания речи.

#### 4.1.2 Описание затухающих сигналов

Естественно предположить, что благодаря затухающему виду функции Майкснера должны лучше описывать затухающие сигналы. На рисунке (11) синим цветом изображен график затухающей функции, красным - аппроксимация с помощью функций Майкснера, чёрным - с помощью ДКП. Видно, что в начале функции Майкснера намного лучше описывают сигнал, а в конце - немного хуже. Правая часть графика показывает, что затухающий вид функций Майкснера может также являться недостатком. Подробнее это описывается в разделе 13.

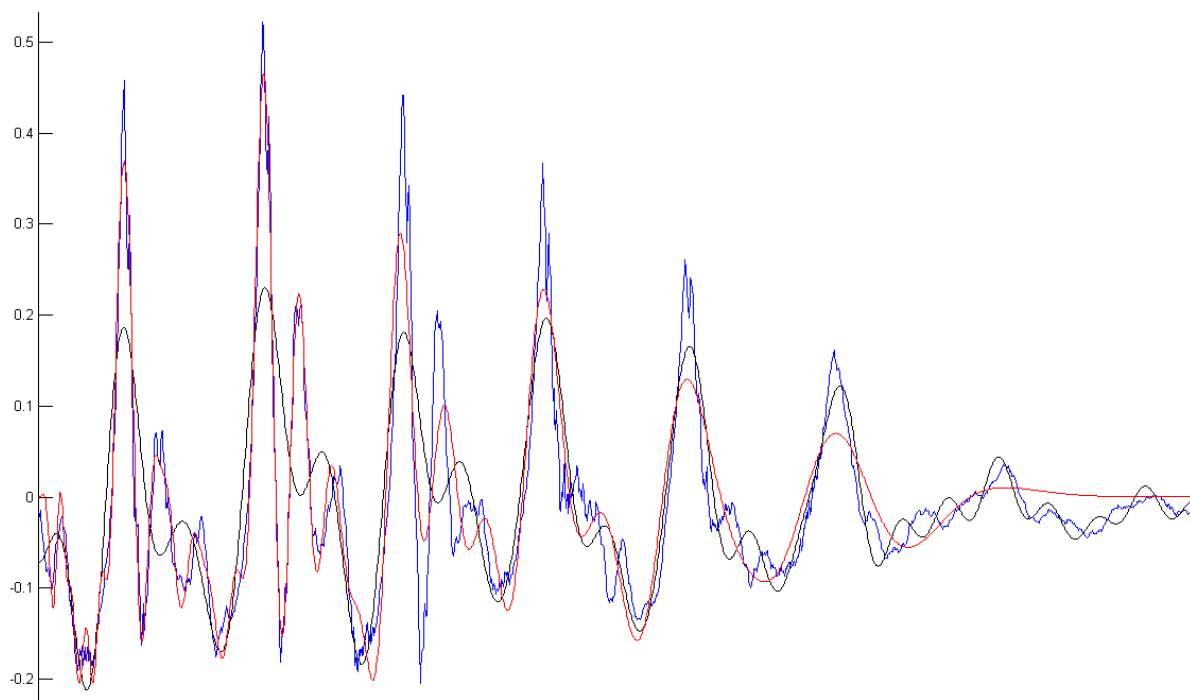


Рис. 11: Затухающая функция

## 4.2 Недостатки функций Майкснера

### 4.2.1 Скорость сжатия и восстановления

Дискретное косинусное преобразование является симметричным методом. То есть время сжатия и восстановления примерно равны.

Преобразование с помощью функций Майкснера не является симметричным. На этапе сжатия с помощью функций Майкснера требуется искать оптимальные параметры. К тому же, как говорилось в разделе 3.2.5, нельзя наращивать степень аппроксимации простым добавлением новых гармоник. На этапе восстановления просто перемножаются две матрицы.

В итоге на этапе сжатия функции Майкснера сильно проигрывают косинусам по времени. На этапе восстановления они проигрывают тоже, но не на много. Однако такой проигрыш во времени на этапе сжатия не является большим недостатком, так как скорость важна именно на этапе восстановления сигнала, так как именно восстановление сигнала происходит в аудиоплеерах. Сжатие сигнала, можно производить на суперкомпьютерах, так как алгоритм сжатия очень хорошо распараллеливается.

### 4.2.2 Гетероскедастичность ошибок

Как утверждалось в разделе 4.1.2, благодаря затухающему виду функции Майкснера отлично описывают затухающие сигналы. Однако в при описании сигналов с примерно постоянной амплитудой затухающий вид функций может играть отрицательную роль. На рисунке (12) синим цветом изображен пример такого сигнала, а красным - его аппроксимация функциями Майкснера. Видно, что в начале кривая регрессии почти полностью повторяет сигнал. Однако правую часть сигнала функции Майкснера описали намного хуже.

На рисунке (13) изображен график квадрата ошибки

$$\varepsilon^2(t) = \left( f(t) - \widetilde{f}_n^{\beta,c}(t) \right)^2. \quad (25)$$

Видно, что ошибка непостоянна и растет со временем. Этот эффект называется гетероскедастичностью [7]. Обычно незатухающий вид имеет полифоническая музыка.

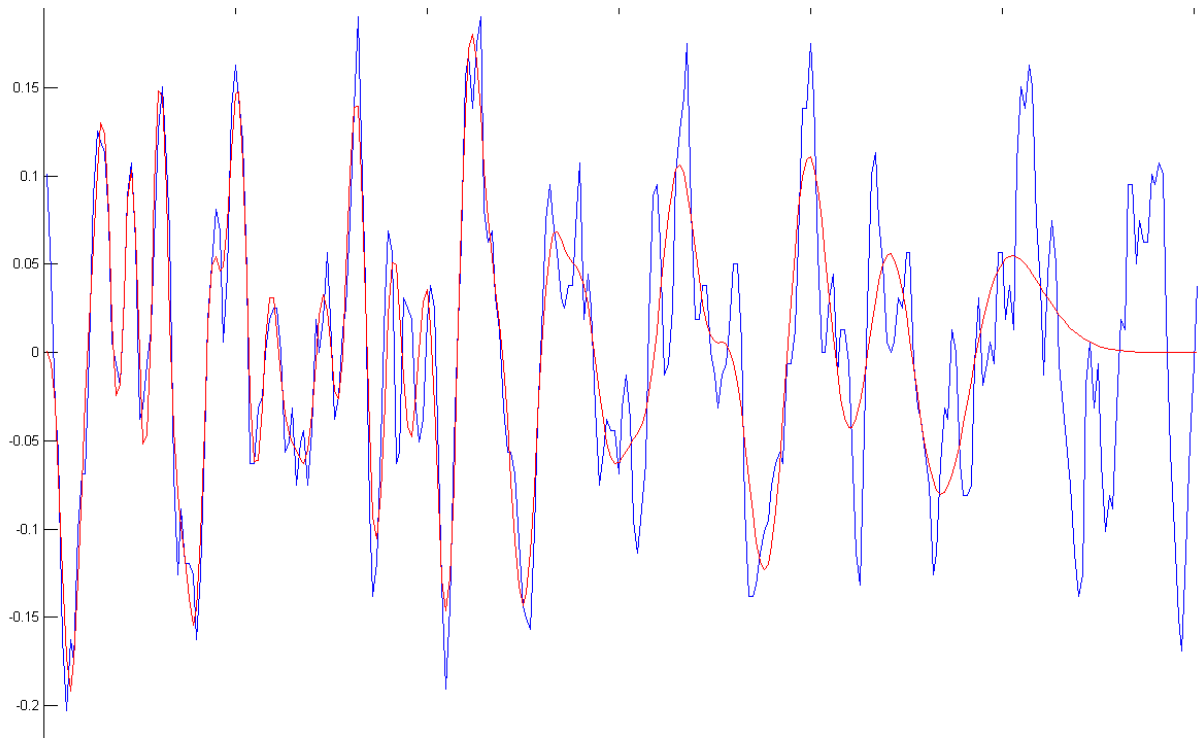


Рис. 12: Незатухающая функция

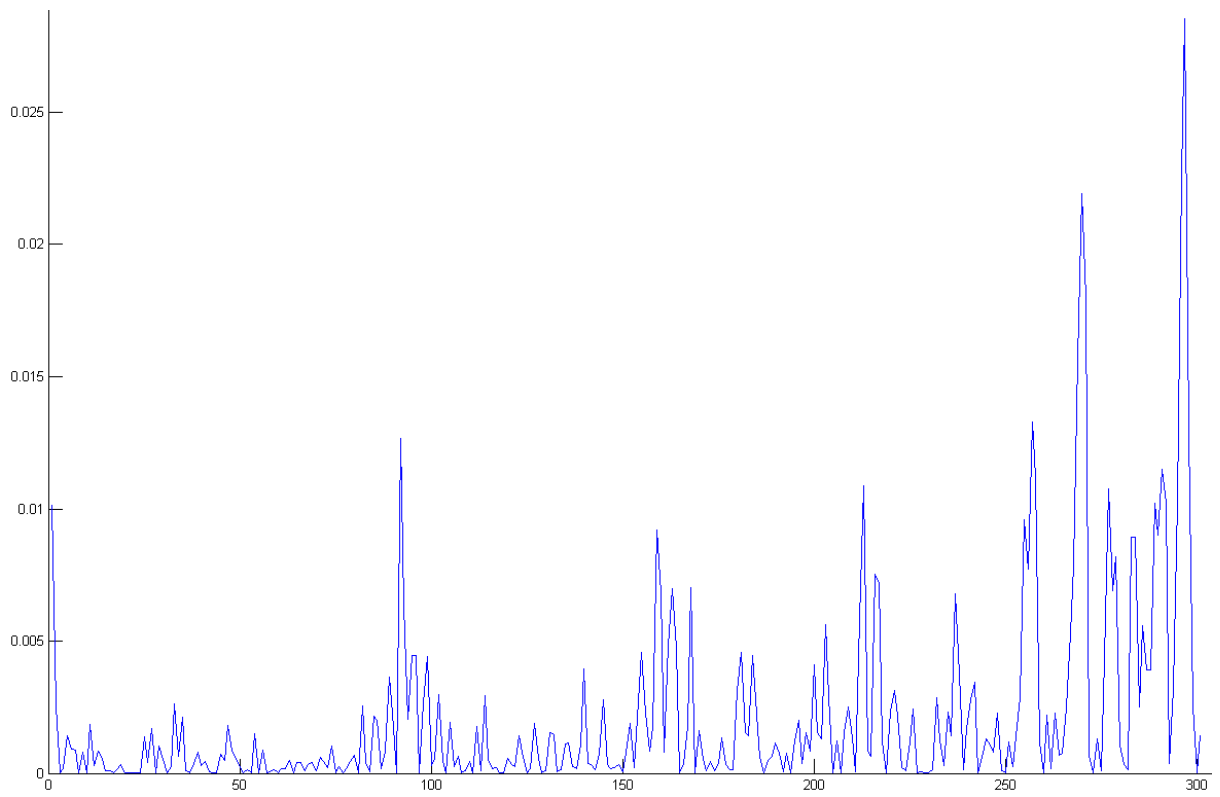


Рис. 13: Гетероскедастичность ошибки

### 4.3 Сравнение на реальных данных

Для сравнения качества сжатия было взято 5 музыкальных файлов и 5 аудиокниг, из каждого файла было случайным образом взято 10 отрезков длиной  $T = 1000$  и для каждого отрезка по формуле (24) был вычислен коэффициент  $K$ .

В итоге:

1. При сжатии речи среднее значение  $K = 1.15$
2. При сжатии музыки среднее значение  $K = 0.7$

Как и предполагалось, функции Майкснера лучше описывают речь, а косинусы лучше описывают полифоническую музыку.

## 5 Заключение

В ходе работы над данной темой были исследованы аппроксимативные свойства функций Майкснера и систем ортогональных косинусов. Было изучено влияние параметров функций Майкснера на их вид и способы подбора оптимальных параметров.

На основе теоретического исследования была написана программа на языке MatLab, которая для сигнала  $f(x)$  и порогового значения  $\varepsilon_0$  вычисляет коэффициенты разложения сигнала по функциям Майкснера с оптимальными параметрами. С помощью этой программы на модельных и реальных данных сделано сравнение аппроксимативных свойств ДКП и функций Майкснера.

Благодаря двум параметрам, позволяющим подстраивать вид функций под сигнал, модель регрессии с помощью функций Майкснера оказалась более гибкой. Однако, из-за необходимости поиска оптимальных параметров описание сигнала функциями Майкснера требует намного больше времени, чем дискретное косинусное преобразование.

Главное отличие функций Майкснера и косинусов - затухающий вид первых. Затухающий вид на разных реальных данных играл как положительную, так и отрицательную роль. Благодаря затухающему виду разложение по функциям Майкснера почти не подвержено эффекту Гиббса, поэтому с помощью функций Майкснера лучше описывалась речь. Однако, из-за затухающего вида функций Майкснера намного хуже описывается полифоническая музыка, амплитуда которой не имеет резких скачков.

## Список литературы

- [1] Ф. Ф. Дедус, Л. И. Куликова, А. Н. Панкратов, Р. К. Тетуев. Классические ортогональные базисы в задачах аналитического описания и обработки информационных сигналов. — М.: Изд. отдел ВМиК МГУ, 2004.
- [2] Д. Джексон. Ряды Фурье и ортогональные полиномы. — М.:Гос. изд. иностранной литературы, 1948.
- [3] А. В. Давыдов. Сигналы и линейные системы.
- [4] А. Ф. Никифоров, В. Б. Уваров. Специальные функции математической физики.
- [5] N. M Atakishyev, K. V. Wolf. Fractional Fourier-Kravchuk transform.
- [6] О. П. Давыдова. Разработка алгоритмов и программ аналитического описания сложных сигналов.
- [7] Л. И. Лопатников. Экономико-математический словарь.
- [8] Д. Ватолин. Методы сжатия Медиаданных
- [9] Joebert S. Jacaba. Audio compression using modified discrete cosine transform: the mp3 coding standart.