

Предположим, что при  $n=0$ ,  $x(0)=0$ ,  $\hat{x}(0)=1$ ,  $d(0)=1$ ,  $\Delta(0)=\Delta_{\min}=1$  и  $\hat{d}(0)=1$ . Изобразить  $x(n)$  и  $\hat{x}(n)$  для  $0 \leq n \leq 25$ .

5.16. Рассмотрим два кодера, представленных на рис. 3.5.5а и б. В каждом кодере используется двухразрядный квантователь с характеристикой рис. 3.5.5в.

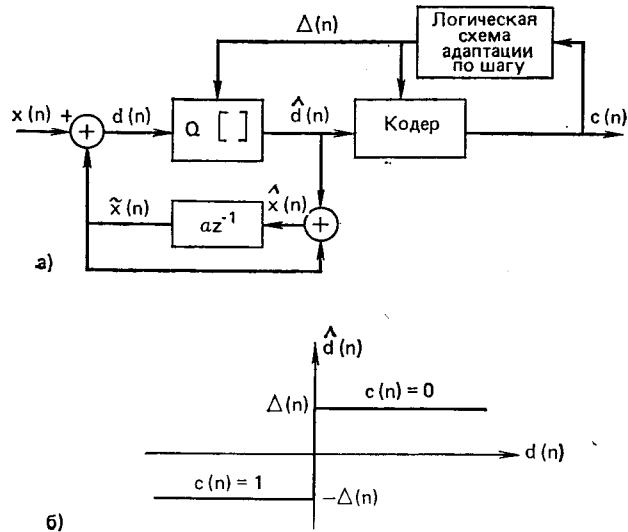


Рис. 3.5.4

Рассмотрим случай свободного канала, т. е. случай, когда есть шум малого уровня. Для простоты положим  $x(n)=0,1 \cos(\pi n/4)$ .

- Для  $0 \leq n \leq 20$  построить  $\hat{x}(n)$  для обоих кодеров.
- Для какого кодера шум незанятого канала будет более неприятным и почему?

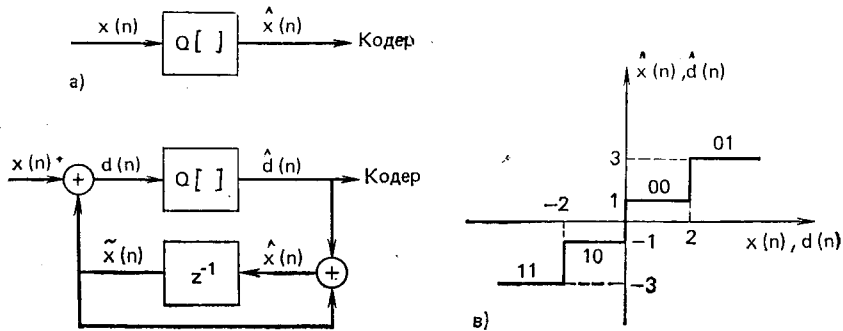


Рис. 3.5.5

5.17. Рассмотрим преобразование ИКМ-АРИКМ (см. рис. 5.46). Сигнал ИКМ  $y(n)$  можно представить в виде  $y(n)=x(n)+e_1(n)$ , где  $x(n)=x_a(nT)$  и  $e_1(n)$  — погрешность квантования при ИКМ представлении. Квантованный АРИКМ сигнал можно представить в виде  $\hat{y}(n)=y(n)+e_2(n)$ , где  $e_2(n)$  — погрешность квантования АРИКМ.

- Полагая погрешности квантования некоррелированными, показать, что полное отношение сигнал/шум  $SNR = \sigma_x^2 / (\sigma_{e_1}^2 + \sigma_{e_2}^2)$ .
- Показать, что отношение сигнал/шум можно записать в форме  $SNR = SNR_1 / [1 + (1 + SNR_1) / SNR_2]$ , где  $SNR_1 = \sigma_x^2 / \sigma_{e_1}^2$  и  $SNR_2 = \sigma_y^2 / \sigma_{e_2}^2$ .

## 6

# Кратковременный анализ Фурье

## 6.0. Введение

Представление сигналов или других функций с помощью сумм синусоид и комплексных экспонент часто приводит к удобным решениям задач науки и техники и помогает понять физику явления глубже, чем это возможно другими методами. Такое представление, часто называемое Фурье-представлением, полезно при обработке сигналов по двум причинам. Во-первых, в линейных системах легко определить отклик на суперпозицию синусоид или комплексных экспонент. Во-вторых, Фурье-представление часто выявляет такие свойства сигнала, которые в первоначальном виде скрыты или по крайней мере не очевидны.

Исследования и техника передачи речи представляют собой области, в которых по традиции преобразование Фурье играло ведущую роль. Чтобы понять, почему это так, полезно вспомнить, что модель образования стационарного речевого сигнала состоит просто из линейной системы, возбуждаемой либо периодически, либо случайно. Спектр на выходе такой модели равен произведению частотной характеристики голосового тракта и спектра возбуждения. Следует ожидать поэтому, что спектр на выходе должен отражать свойства как возбуждения, так и частотной характеристики голосового тракта. Однако речевой сигнал гораздо сложнее, чем просто продолжительный гласный или фрикативный звук. Поэтому стандартное Фурье-представление, вполне пригодное для периодических, импульсных или стационарных случайных сигналов, неприменимо к речевому сигналу, характеристики которого значительно меняются во времени. Однако мы уже имели возможность убедиться в том, что принцип кратковременного анализа оказывается полезным при обработке речи. Например, такие изменяющиеся во времени характеристики, как энергия, переходы через нуль и корреляция, можно считать постоянными на интервалах времени около 10—30 мс. В этой главе будет показано, что аналогичным образом и спектральные характеристики речи можно считать изменяющимися во времени сравнительно медленно.

Для изучения спектральных характеристик речевого сигнала удобно ввести формализм, связанный с понятием кратковременного преобразования Фурье. Введем кратковременное преобразование Фурье и процедуру синтеза, основанную на нем. При этом удобно рассматривать анализ Фурье, как преобразование сигнала в гребенке фильтров. Такой подход позволяет лучше понять как теоретические, так и практические (вычислительные) аспекты кратковременного анализа. Будут рассмотрены также и другие вычислительные методы, основанные на быстром алгоритме вычисления дискретного преобразования Фурье (алгоритмы БПФ). И, наконец, подробно изучив теорию и способ вычисления кратковременного преобразования Фурье, рассмотрим его применение в задачах анализа — синтеза речи (вокодеры), визуального отображения спектра и в таких важных задачах обработки речи, как формантный анализ и выделение основного тона.

<sup>1</sup> Перевод данной главы выполнен при участии канд. физ.-мат. наук А. Ю. Шевердяева. (Прим. ред.)

## 6.1. Определения и свойства

Потребность в спектральном представлении, отображающем меняющиеся во времени свойства речевых сигналов, побуждает нас ввести представление Фурье, зависящее от времени. Подходящим определением зависящего от времени преобразования Фурье будет следующее:

$$X_n(e^{i\omega}) = \sum_{m=-\infty}^{\infty} \omega(n-m)x(m)e^{-i\omega m}. \quad (6.1)$$

В (6.1)  $\omega(n-m)$  представляет собой действительную последовательность временного окна. Этой последовательностью выделяется часть входного сигнала в момент времени  $n$ . Ясно, что зависящее от времени преобразование Фурье представляет собой функцию двух переменных: времени  $n$ , которое предполагается дискретным, и частоты  $\omega$ , предполагаемой здесь непрерывной. Другая форма для (6.1) получается при замене переменной суммирования, что дает

$$\begin{aligned} X_n(e^{i\omega}) &= \sum_{m=-\infty}^{\infty} \omega(m)x(n-m)e^{-i\omega(n-m)} = \\ &= e^{-i\omega n} \sum_{m=-\infty}^{\infty} x(n-m)\omega(m)e^{i\omega m}. \end{aligned} \quad (6.2)$$

Определив

$$\tilde{X}_n(e^{i\omega}) = \sum_{m=-\infty}^{\infty} x(n-m)\omega(m)(e^{i\omega m}), \quad (6.3)$$

получим для  $X_n(e^{i\omega})$  выражение

$$X_n(e^{i\omega}) = e^{-i\omega n} \tilde{X}_n(e^{i\omega}). \quad (6.4)$$

Приведенные соотношения допускают две различные интерпретации. Во-первых, зафиксировав  $n$ , видим, что  $X_n(e^{i\omega})$  представляет собой обычное преобразование Фурье для последовательности  $\omega(n-m)x(m)$ ,  $-\infty < m < \infty$ . Поэтому для фиксированного  $n$   $X_n(e^{i\omega})$  обладает свойствами обычного преобразования Фурье. Вторая интерпретация получается, если зафиксировать  $\omega$  и рассматривать  $X_n(e^{i\omega})$  как функцию времени  $n$ . При таком подходе видно, что и (6.1) и (6.3) представляют собой свертки. Это позволяет рассматривать зависящее от времени представление Фурье с помощью линейной фильтрации. Как мы увидим ниже, обе интерпретации позволяют глубже понять существо дела, поэтому представляется целесообразным подробно изучить преобразование Фурье с обеих точек зрения.

### 6.1.1. Интерпретация преобразования Фурье

Рассмотрим  $X_n(e^{i\omega})$  как обычное преобразование Фурье последовательности  $\omega(n-m)x(m)$ ,  $-\infty < m < \infty$  при фиксированном  $n$ . Зависящее от времени преобразование Фурье представляет собой функцию индекса времени  $n$ , принимающего все целые значения, так что окно  $\omega(n-m)$  «скользит» вдоль последовательности  $x(m)$ . Это иллюстрируется на рис. 6.1, где  $x(m)$  и  $\omega(n-m)$  показаны как функции  $m$  для нескольких значений  $n$ . (И сигнал, и окно изображены для удобства непрерывными функциями, хотя определены они только для целых значений  $m$  и  $n-m$ .)

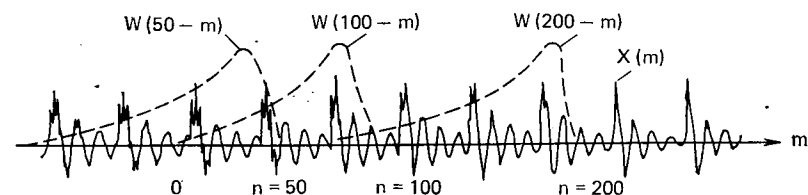


Рис. 6.1.  $x(m)$  и  $\omega(n-m)$  для нескольких значений  $n$

Условия существования зависящего от времени преобразования Фурье легко получить, вспомнив, что достаточным условием существования обычного преобразования Фурье служит абсолютная суммируемость последовательности. Потребуем, чтобы последовательность  $x(m)\omega(n-m)$  была абсолютно суммируемой при всех значениях  $n$ . Очевидно, что это выполняется, если, как это часто бывает,  $\omega(n-m)$  имеет конечную длительность.

Как и в случае обычного преобразования Фурье сигналов с дискретным временем, зависящее от времени преобразование Фурье периодически по  $\omega$  с периодом  $2\pi$ , в чем легко убедиться, подставив  $\omega+2\pi$  в (6.1). Отметим также, что это преобразование Фурье можно выразить как функцию частоты различными способами. Если, например,  $\omega = \Omega T$ , где  $T$  — период дискретизации, приводящей к последовательности  $x(m)$ , то  $\Omega$  будет непрерывной частотой в радианах. При подстановке  $\omega = 2\pi f$  или  $\omega = 2\pi fT$  можно выразить преобразование Фурье как функцию нормированной частоты  $f$  и соответственно как функцию обычной непрерывной частоты  $F$  (в герцах). Нам представится возможность воспользоваться разными частотными переменными в формулах и на рисунках; по мере знакомства с подобными простыми соотношениями возможностей для путаницы будет все меньше. То обстоятельство, что  $X_n(e^{i\omega})$  обладает при заданном  $n$  свойствами обычного преобразования Фурье, позволяет легко доказать, что входная последовательность  $x(m)$  может быть полностью восстановлена по ее зависящему от времени преобразованию Фурье. Вспомним, что, как было отмечено выше,  $X_n(e^{i\omega})$  представляет собой обычное преобразование

Фурье для  $\omega(n-m)x(m)$ :

$$\omega(n-m)x(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_n(e^{i\omega}) e^{i\omega m} d\omega. \quad (6.5)$$

Заметим, что интегрировать в (6.5) можно по любому интервалу длиной  $2\pi$  (например, от 0 до  $2\pi$ ), поскольку интегрируемая функция периодична с периодом  $2\pi$ . Если теперь  $\omega(0) \neq 0$ , из (6.5) при  $m=n$  можно получить

$$x(n) = \frac{1}{2\pi \omega(0)} \int_{-\pi}^{\pi} X_n(e^{i\omega}) e^{i\omega n} d\omega. \quad (6.6)$$

Следовательно, при слабом ограничении на  $\omega(0)$  и если  $X_n(e^{i\omega})$  известна при  $\omega$  в интервале, перекрывающем полный период, последовательность  $x(n)$  точно восстанавливается по значениям  $X_n(e^{i\omega})$ . Это важный теоретический результат. Он важен и в приложениях, если на окно наложены некоторые дополнительные ограничения.

Связь с кратковременной автокорреляционной функцией, определенной в гл. 4, представляет собой еще одно важное свойство  $X_n(e^{i\omega})$ . Рассматривая  $X_n(e^{i\omega})$  как обычное преобразование Фурье для  $\omega(n-m)x(m)$  при каждом  $n$ , легко понять, что

$$S_n(e^{i\omega}) = |X_n(e^{i\omega})|^2 = X_n(e^{i\omega}) X_n^*(e^{i\omega}) \quad (6.7)$$

есть преобразование Фурье для

$$R_n(k) = \sum_{m=-\infty}^{\infty} \omega(n-m)x(m)\omega(n-k-m)x(m+k). \quad (6.8)$$

Таким образом, соотношения (6.7) и (6.8) связывают кратковременное спектральное представление с кратковременной корреляцией, введенной в гл. 4.

Кратковременное преобразование Фурье<sup>1</sup> можно представить многими способами. Особенно простой вид оно принимает, если выражено через действительную и мнимую части<sup>2</sup>:

$$X_n(e^{i\omega}) = a_n(\omega) - i b_n(\omega). \quad (6.9)$$

Можно показать, что  $x(m)$  и  $\omega(n-m)$  удовлетворяют определенным условиям симметрии и периодичности, когда  $a_n(\omega)$  и  $b_n(\omega)$  действительны (см. задачу 6.1). Другое представление  $X_n(e^{i\omega})$  через амплитуду и фазу:

$$X_n(e^{i\omega}) = |X_n(e^{i\omega})| e^{i\theta_n(\omega)}. \quad (6.10)$$

<sup>1</sup> Авторы для преобразования (6.1) применяют три термина: short-time, time-dependent, time-varying. С целью устранения возможной путаницы далее сохраняется одно название — кратковременное преобразование Фурье (Прим. ред.)

<sup>2</sup> Обратите внимание на то, что  $a_n(\omega)$  есть действительная часть  $X_n(e^{i\omega})$ , а  $b_n(\omega)$  — мнимая, взятая со знаком «минус». Последнее сделано для удобства.

Величины  $|X_n(e^{i\omega})|$  и  $\theta_n(\omega)$  легко связать с  $a_n(\omega)$  и  $b_n(\omega)$  (см. задачу 6.3). Другие задачи в конце этой главы указывают на дополнительные свойства функций  $a_n(\omega)$ ,  $b_n(\omega)$  и  $X_n(e^{i\omega})$ .

До сих пор единственная роль окна  $\omega(n-m)$  состояла в том, чтобы выделить для анализа часть последовательности  $x(m)$ . Форма временного окна оказывает существенное влияние на характер кратковременного преобразования Фурье. Рассмотрим это влияние последовательности  $\omega(n-m)$ . Если представить себе  $X_n(e^{i\omega})$  как обычное преобразование Фурье последовательности  $\omega(n-m)x(m)$  и если предположить, что обычные преобразования Фурье

$$X(e^{i\omega}) = \sum_{m=-\infty}^{\infty} x(m) e^{-i\omega m} \quad (6.11)$$

и

$$W(e^{i\omega}) = \sum_{m=-\infty}^{\infty} \omega(m) e^{-i\omega m} \quad (6.12)$$

существуют, то обычное преобразование Фурье для  $\omega(n-m)x(m)$  (при фиксированном  $n$ ) представляет собой свертку преобразования  $\omega(n-m)$  и  $x(m)$ . Поскольку при фиксированном  $n$  преобразование Фурье  $\omega(n-m)$  есть  $W(e^{i\omega}) e^{-i\omega n}$ , имеем

$$X_n(e^{i\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(e^{-i\theta}) e^{-i\theta n} X(e^{i(\omega-\theta)}) d\theta. \quad (6.13)$$

Заменив в (6.13)  $\theta$  на  $-\theta$ , получим

$$X_n(e^{i\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(e^{i\theta}) e^{i\theta n} X(e^{i(\omega+\theta)}) d\theta. \quad (6.14)$$

Таким образом, преобразование Фурье последовательности  $x(m)$ ,  $-\infty < m < \infty$ , свертывается с преобразованием Фурье сдвинутой последовательности окна. Этот результат нуждается в оговорке, поскольку, строго говоря, обычного преобразования Фурье речевого сигнала не существует. Соотношение (6.14) станет полезным, если вспомнить, что смысл использования окна состоит в том, чтобы выделить некоторый конечный сегмент речевого сигнала в окрестности отсчета  $n$  и устранить остальную часть сигнала. И действительно, типичное окно таково, что  $\omega(n-m) = 0$  для  $m$ , не принадлежащих конечному интервалу вокруг  $n$ . Имея в виду конечный результат, вполне оправдано считать, что свойства последовательности определяются той ее частью, которая попадает в окно. Если, например, речевой сигнал в пределах окна соответствует гласному или другому вокализованному звуку, то можно рассматривать результирующую последовательность  $x(m)\omega(n-m)$ , как соответствующую периодически продолженному вокализованному звуку. Аналогичным образом, если речь в пределах окна не вокализована, можно считать, что характеристики невокализованного

сигнала сохраняются и вне окна. Равным образом можно считать, что сигнал вне окна равен нулю. Эта интерпретация пригодится при анализе переходных звуков, таких, как взрывные.

Следовательно, (6.14) имеет смысл как в случае, когда мы считаем, что  $X(e^{i\theta})$  представляет собой преобразование Фурье сигнала, свойства которого сохраняются и вне окна, так и в случае, когда  $X(e^{i\theta})$  соответствует сигналу, обращающемуся вне окна в нуль. Таким образом, кратковременное преобразование Фурье можно интерпретировать как сглаженное преобразование Фурье для части сигнала, попавшего в окно.

В этом смысле становятся важными свойства преобразования Фурье окна  $W(e^{i\theta})$ . Из (6.14) ясно, что для достоверного отражения свойств  $X(e^{i\omega})$  в  $X_n(e^{i\omega})$ , необходимо чтобы функция  $W(e^{i\theta})$  носила импульсный характер по сравнению с  $X(e^{i\omega})$ . В гл. 4 уже обсуждалось свойство прямоугольного окна и окна Хемминга. Было показано, что ширина главного лепестка  $W(e^{i\theta})$  обратно пропорциональна ширине окна, в то время как уровень боковых лепестков от ширины окна по существу не зависит.

Эффекты, связанные с использованием окон при спектральном анализе речи, иллюстрируются рис. 6.2—6.5. Часть *a*) каждого из этих рисунков показывает сигнал  $x(m)w(n-m)$ , взвешенный окном Хемминга, часть *б*) представляет собой логарифм амплитуды  $X_n(e^{i\omega})$  (в децибелах), часть *в*) — сигнал, взвешенный прямоугольным окном, и часть *г*) — логарифм амплитуды соответствующего спектра. На рис. 6.2 представлены результаты для окна длительностью 500 отсчетов (50 мс при частоте дискретизации 10 кГц) и сегмента вокализованной речи. Отчетливо прослеживается периодичность сигнала на рис. 6.2*a* (временная диаграмма) и 6.2*б*, где основная частота и ее гармоники проявляются в кратковременном преобразовании Фурье как узкие пики, разнесенные равномерно по частоте. Из рис. 6.2*б* видно также, что спектр состоит из пика первой форманты в районе 300—400 Гц, и широкого пика около 2200 Гц, соответствующего второй и третьей формантам. Заметен также пик четвертой форманты — около 3800 Гц. И, наконец, видно, что спектр имеет тенденцию спадать на высоких частотах, что характерно для спектра импульсов возбуждения. Сравнение спектров на рис. 6.2*б* (окно Хемминга) и 6.2*г* (прямоугольное окно), выявляет их большое сходство по гармоникам основного тона, формантной структуре и форме спектра в целом. Видны и различия в спектрах, из них наиболее отчетливо — большая острота гармоник основного тона на рис. 6.2*г*, что вызвано лучшим частотным разрешением прямоугольного окна в сравнении с окном Хемминга той же ширины. Другое различие в спектрах вызвано тем, что сравнительно большие боковые лепестки прямоугольного окна дают «рваный», или «зашумленный» спектр. Этот эффект связан с тем, что боковые лепестки смежных гармоник взаимодействуют в интервалах между гармониками, иногда усиливаясь, иногда уничтожая друг друга, в целом производя впечатление довольно беспорядочного характера спектра между гармониками. Нежела-

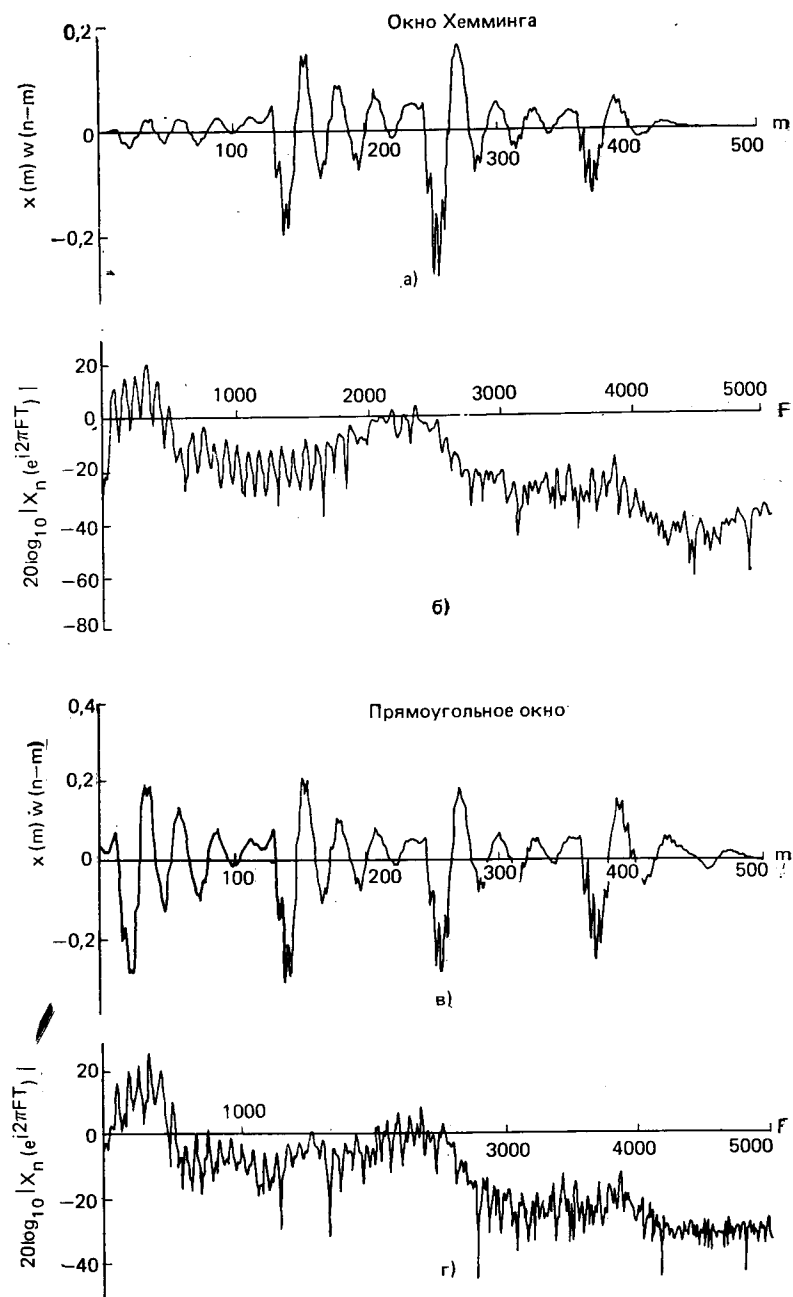


Рис. 6.2. Спектральный анализ вокализованной речи с окном длительностью 50 мс: *a*, *б*) — окно Хемминга; *в*, *г*) — прямоугольное окно (на рис. *a* и *в* показан сигнал, а на рис. *б* и *г* — спектры)

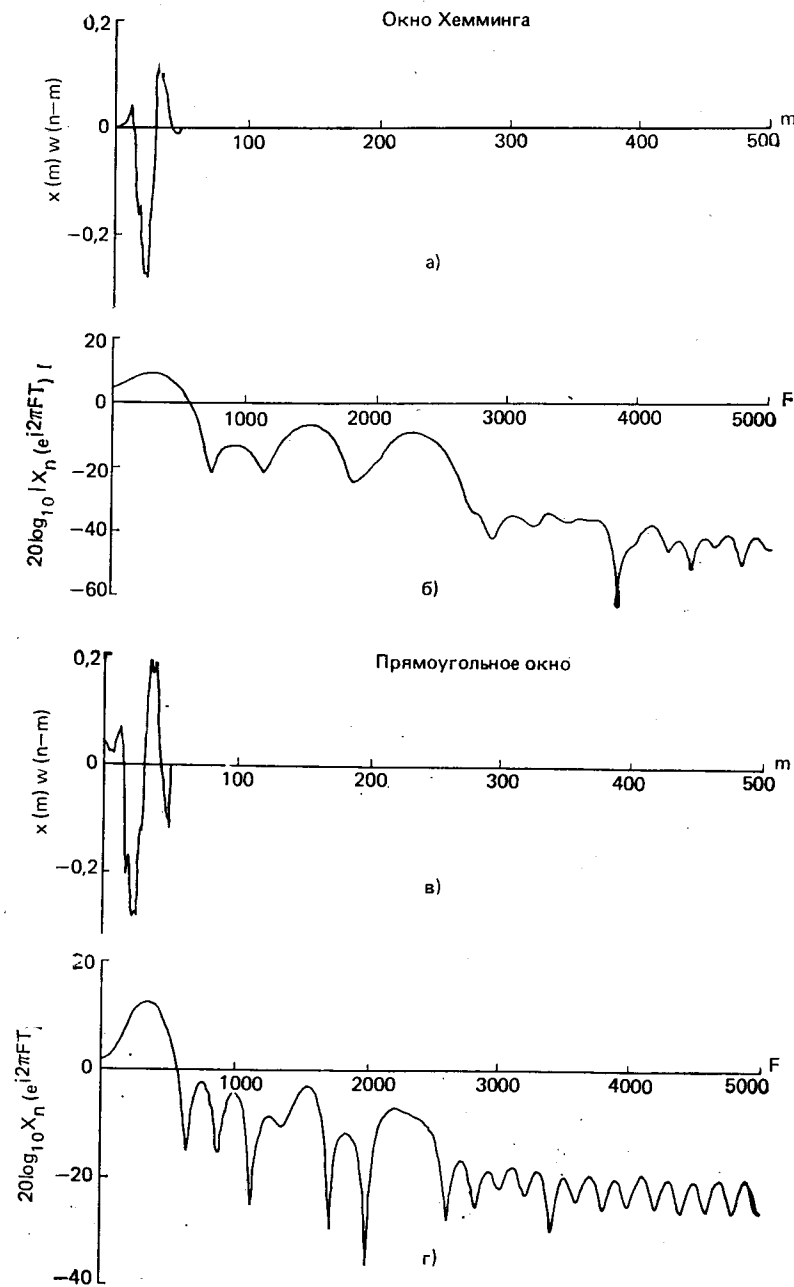


Рис. 6.3. Спектральный анализ вокализованной речи с окном длительностью 5 мс:  
 а, б) — окно Хемминга; в, г) — прямоугольное окно (на рис. а и в показан сигнал, а на рис. б и г — спектры)

тельное «рассеивание» спектра между смежными гармониками сводит на нет выгоды, связанные с узостью основного лепестка прямоугольного окна. Поэтому такие окна редко используются в спектральном анализе речи.

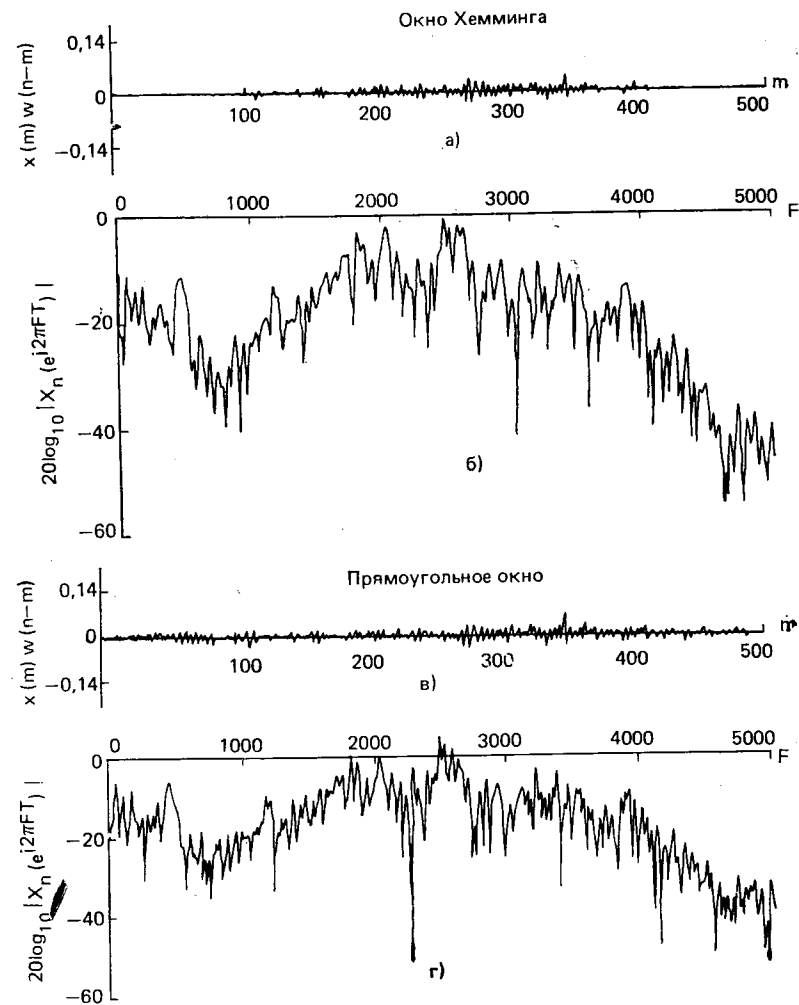


Рис. 6.4. Спектральный анализ невокализованной речи с окном длительностью 50 мс:  
 а, б) — окно Хемминга; в, г) — прямоугольное окно (на рис. а и в показан сигнал, а на рис. б и г — спектры)

На рис. 6.3 аналогичным образом проведено сравнение для сегмента вокализованной речи длительностью 50 отсчетов (5 мс). Для столь узкого окна ни временная последовательность  $x(m) \times w(n-m)$  (рис. 6.3а, в), ни спектр сигнала (рис. 6.3б, г) не вы-

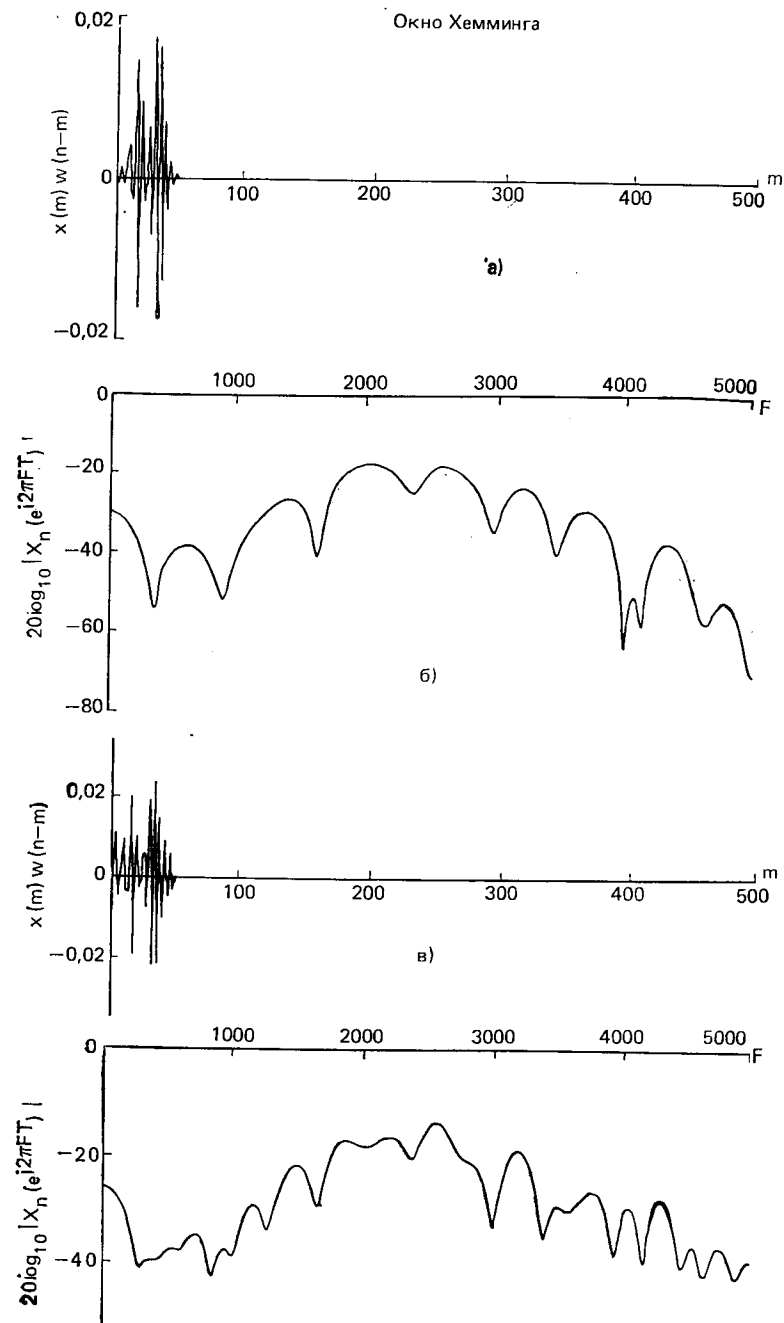


Рис. 6.5. Спектральный анализ невокализованной речи с окном длительностью 5 мс:  
 а, б) — окно Хемминга; в, г) — прямоугольное окно (на рис. а и в показан сигнал, а на рис. б и г — спектры)

являют периодичности сигнала. В сравнении с рис. 6.2 спектр на рис. 6.3 имеет только несколько довольно широких пиков вблизи 400, 1400 и 2200 Гц, соответствующих первым трем формантам отрезка речи, попавшего в окно. Сравнение спектров снова выявляет большее частотное разрешение прямоугольного окна.

Рисунки 6.4 и 6.5 иллюстрируют влияние окна для сегментов невокализованной речи (соответствующих фрикативу /sh/) длиной 500 отсчетов (рис. 6.4) и длиной 50 отсчетов (рис. 6.5). На этих рисунках видно, что спектр имеет медленно меняющийся тренд по частоте, на который наложена последовательность острых пиков и провалов. «Рваный» характер спектра (для обоих окон) связан со случайной природой невокализованной речи. И, наконец, видно, что окно Хемминга дает несколько более гладкий спектр в сравнении с прямоугольным окном.

Примеры, приведенные на рис. 6.2 и 6.5, хорошо иллюстрируют зависимость между длительностью окна и свойствами кратковременного преобразования Фурье: разрешение по частоте обратно пропорционально ширине окна. Вспомнив, что задача окна состоит в том, чтобы выделить подлежащий анализу интервал времени, сохранив, однако, характеристики колебания без существенных изменений, мы убеждаемся в необходимости компромисса. На рис. 6.2, например, видно, что частоты формант явно меняются на интервале, равном 50 мс. Необходим более короткий интервал анализа, чтобы выявить эти изменения. Окна шириной 5 мс, размещенные в начале и в конце 50-миллисекундного интервала, явно дадут другие кратковременные преобразования Фурье. Следовательно, хорошее разрешение по времени требует узкого окна, а хорошее разрешение по частоте — широкого. Позже, изучая приложения, мы приведем примеры использования обоих типов окон.

Итак, интерпретация кратковременного преобразования Фурье как обычного преобразования Фурье выделенного окном сегмента речевого сигнала позволяет глубже понять свойства кратковременного представления Фурье и роль самого окна.

### 6.1.2. Интерпретация посредством линейной фильтрации

Как ясно из (6.1), для каждого  $\omega$   $X_n(e^{j\omega})$  представляет собой свертку последовательности  $w(n)$  с последовательностью  $x(n)e^{-j\omega n}$ . Поэтому для фиксированных  $\omega$  можно представлять себе  $X_n(e^{j\omega})$  как выход системы, изображенной на рис. 6.6а, где  $w(n)$  играет роль импульсной характеристики линейной системы, инвариантной относительно сдвига. На рис. 6.6а вход и выход линейной системы комплексные. Выразим  $X_n(e^{j\omega})$  в виде

$$X_n(e^{j\omega}) = a_n(\omega) - i b_n(\omega). \quad (6.15)$$

Операции, необходимые для расчета  $a_n(\omega)$  и  $b_n(\omega)$ , показаны на рис. 6.6б, где имеются только действительные последовательности.

Чтобы понять, как в системе на рис. 6.6а формируется кратковременное преобразование Фурье на частоте  $\omega$ , полезно снова

предположить, что существует обычное преобразование Фурье  $x(n)$ . Обозначим преобразование Фурье  $x(n)$  через  $X(e^{i\theta})$ , чтобы избежать путаницы с частотными переменными. (Напомним, что теперь  $\omega$  — фиксированное значение частоты в радианах.) В результате модуляции преобразование Фурье входа линейного фильтра равно  $X(e^{i(\theta+\omega)})$ . Поэтому спектр  $x(n)$  на частоте  $\omega$  сдвигается

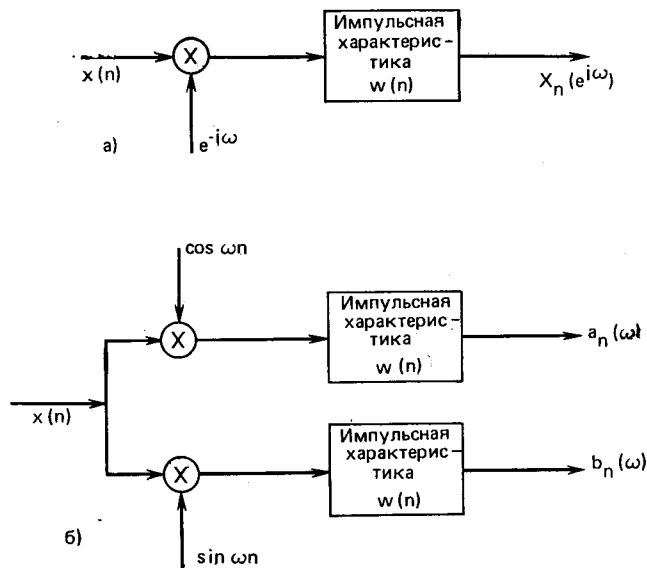


Рис. 6.6. Интерпретация кратковременного анализа как линейной фильтрации: а) — комплексные операции; б) — действительные операции

к нулю. Поскольку преобразование Фурье последовательности, поступающей с выхода фильтра, есть  $X(e^{i(\theta+\omega)})W(e^{i\theta})$ , то в случае фильтра нижних частот с очень узкой полосой пропускания выходная последовательность фильтра существенным образом зависит от  $X(e^{i\omega})$ . Поэтому, как следует из предшествующей интерпретации,  $W(e^{i\theta})$  должна быть отличной от нуля только в очень узком интервале вокруг нулевой частоты и быть по возможности меньшей вне этого интервала. Отметим интересный факт: правая часть (6.14) представляет собой преобразование Фурье  $W(e^{i\theta})X(e^{i(\theta+\omega)})$ .

Еще одна интерпретация для  $X_n(e^{i\omega})$  через линейные фильтры получается из рассмотрения (6.2). Как видно из рис. 6.7а,  $X_n(e^{i\omega})$  можно представить себе как результат модуляции выходным сигналом комплексного полосового фильтра с импульсной характеристикой  $w(n)e^{i\omega n}$ . Если преобразование Фурье  $W(e^{i\theta})$  имеет вид характеристики фильтра нижних частот, то фильтр, изображенный на рис. 6.7а, будет полосовым фильтром с центральной частотой  $\omega$ . На рис. 6.7б приведена та же система, что и на рис. 6.7а, но используются только действительные величины.

Сравнение рис. 6.6б и 6.7б показывает, что в случае, когда необходимы  $a_n(\omega)$  и  $b_n(\omega)$ , реализация системы рис. 6.6б проще. Если, однако, требуется получить только  $|X_n(e^{i\omega})|$ , то проще реализация с полосовым фильтром. Это становится понятным, если заметить, что из (6.4) и (6.9) получается

$$|X_n(e^{i\omega})| = [a_n^2(\omega) + b_n^2(\omega)]^{1/2} = \quad (6.16a)$$

$$= |\tilde{X}(e^{i\omega})| = [\tilde{a}_n^2(\omega) + \tilde{b}_n^2(\omega)]^{1/2}. \quad (6.16б)$$

Рисунок 6.8 иллюстрирует выражение (6.16а), а рис. 6.8б — (6.16б). Система, изображенная на рис. 6.8б, будет, вообще говоря, проще.

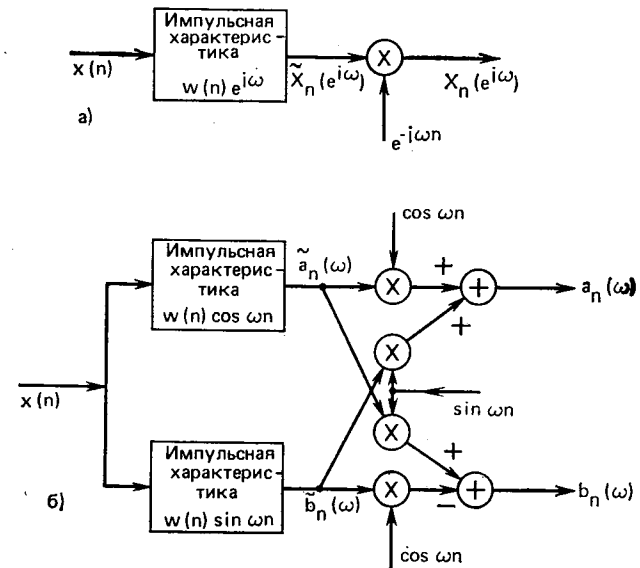


Рис. 6.7. Другая интерпретация кратковременного спектрального анализа как линейной фильтрации:

а) — комплексные операции; б) — действительные операции

Рассматривая  $X_n(e^{i\omega})$  при фиксированном  $\omega$  как выход системы, изображенной на рис. 6.6 и 6.7, можно воспользоваться нашими знаниями линейных систем, чтобы уяснить свойства кратковременного представления Фурье. Полезно, например, вспомнить, что импульсная характеристика линейной системы с дискретным временем, инвариантной относительно сдвига, может быть либо конечной (КИХ), либо бесконечной (БИХ) длительности. Аналогичным образом можно определить два класса окон для кратковременного анализа Фурье. Вспомним также, что линейная инвариантная к сдвигу система может быть или не быть физически реализуемой в зависимости от того, равна ли нулю ее импульсная характеристика при  $n < 0$ . Точно так же можно классифицировать окна на физиче-

ски реализуемые или нереализуемые. Для физически реализуемого окна

$$w(n) = 0, \quad n < 0, \quad (6.17a)$$

или, что эквивалентно,  $w(n-m) = 0, \quad n < m.$  (6.17б)

Окно Хемминга и прямоугольное окно являются окнами конечной ширины. При подходящем выборе начала отсчета во времени

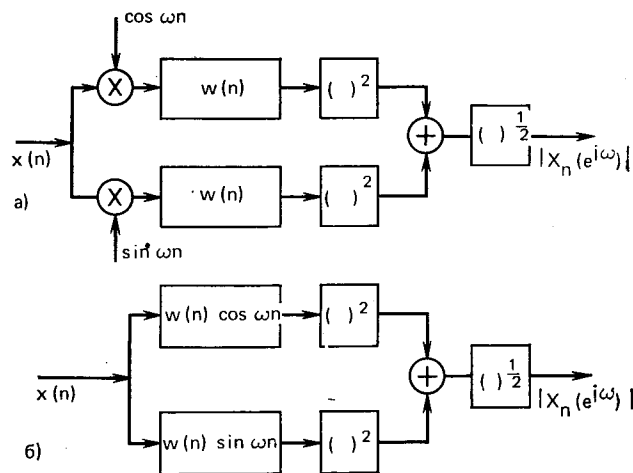


Рис. 6.8. Две схемы вычисления амплитудного кратковременного спектра: а) с фильтрами низкой частоты; б) с полосовыми фильтрами

их можно задать и как физически реализуемые. Как мы увидим ниже, такие окна пригодны как для реализации, основанной на схемах, изображенных на рис. 6.6 и 6.7, так и при реализации, основанной на дискретном преобразовании Фурье. Окна бесконечной ширины также оказываются полезными, особенно если  $X_n(e^{iω})$  вычисляется с помощью линейной фильтрации (рис. 6.6 и 6.7). В подобных случаях можно получить рекуррентные формулы для функции  $X_n(e^{iω})$  через ее значения в предыдущие моменты времени (см. задачу 6.6).

### 6.1.3. Частоты дискретизации $X_n(e^{iω})$ по времени и частоте<sup>1</sup>

Кратковременное преобразование Фурье есть двумерное представление одномерного сигнала  $x(n)$ . Иначе говоря,  $X_n(e^{iω})$  представляет собой функцию и времени  $n$  и круговой частоты  $ω$ . При цифровой реализации системы кратковременного анализа Фурье одним из основных оказывается вопрос о выборе частот дискретизации  $X_n(e^{iω})$  по времени и частоте так, чтобы избежать наложения и получить представление для  $X_n(e^{iω})$ , по которому можно было бы точно восстановить  $x(n)$ .

<sup>1</sup> Содержание оставшейся части § 6.1 основано на работах [1—4].

Вопрос этот ни в коей мере не тривиален и требует тщательного учета всех факторов, связанных с вычислением  $X_n(e^{iω})$ , для того, чтобы получить правильные значения частот дискретизации по времени и частоте. Как будет показано ниже, выбор частот дискретизации усложняется тем, что и в частотной и во временной областях можно использовать частоты дискретизации, меньшие теоретически минимальной, и все же точно восстановить  $x(n)$  по кратковременному представлению (при наличии наложения частот).

Такого рода представления (с пониженной частотой дискретизации) в действительности весьма полезны для приложений, в которых важен лишь кратковременный анализ Фурье (например, для оценки спектра, формантного анализа и анализа основного тона, для получения спектрограмм речи в дискретном времени и т. п.), а также применительно к вокодерам, в которых важнее всего минимизация общей скорости передачи в системе. В тех случаях, когда требуется получить кратковременное преобразование Фурье и далее произвести некоторые операции над ним (например, линейно или нелинейно отфильтровать), а затем синтезировать сигнал, важно, чтобы наложения не возникали ни в частотной, ни во временной областях.

Рассмотрим требования к частоте дискретизации, начиная с временной области. В этом случае наша интуиция может опираться на приведенную в предыдущем разделе интерпретацию посредством линейной фильтрации. Там было показано, что для фиксированного значения  $ω$   $X_n(e^{iω})$  будет выходом фильтра с импульсной характеристикой  $w(n)$ . Обозначим преобразование Фурье  $w(n)$  через  $W(e^{iω})$ . Для большинства разумно выбранных окон функция  $W(e^{iω})$  обладает свойствами частотной характеристики фильтра нижних частот. Обозначим эффективную полосу анализирующего фильтра через  $B$  Гц<sup>1</sup>. Таким образом, у  $X_n(e^{iω})$  такая же полоса, как и у окна, а следовательно, в соответствии с теоремой отсчетов  $X_n(e^{iω})$  должна быть дискретизирована, по крайней мере, с частотой  $2B$  отсч./с для того, чтобы избежать наложений. В качестве примера рассмотрим окно Хемминга:

$$w(n) = \begin{cases} 0,54 - 0,46 \cos(2\pi n/(L-1)), & 0 \leq n \leq L-1; \\ 0, & \text{в противном случае,} \end{cases} \quad (6.18)$$

Ширина полосы  $W(e^{iω})$ , выраженная через непрерывную частоту, приближенно равна

$$B = 2F_s/L, \quad (6.19)$$

где  $F_s$  — частота дискретизации сигнала  $x(n)$  и, следовательно, требуемая частота дискретизации  $X_n(e^{iω})$  во временной области равна  $2B = 4F_s/L$  отсч./с. Таким образом, для  $L=100$ ,  $F_s=$

<sup>1</sup> Отметим, что здесь может возникнуть путаница с частотными переменными. Напомним, что когда мы рассматриваем  $X_n(e^{iω})$  как функцию времени,  $ω$  — число (фиксировано). Однако эта переменная  $ω$  используется также для обозначения переменной частоты в  $X_n(e^{iω})$ .



$= 10\,000$  Гц получим  $B=200$  Гц, и  $X_n(e^{i\omega})$  должна вычисляться 400 раз/с, т. е. через каждые 25 отсчетов.

Ввиду того что функция  $X_n(e^{i\omega})$  периодична по  $\omega$  с периодом  $2\pi$ , дискретизация необходима только в интервале длиной  $2\pi$ . Воспользуемся интерпретацией  $X_n(e^{i\omega})$  через преобразование Фурье, чтобы найти подходящий конечный набор частот  $\omega_k = 2\pi k/N$ ;  $k = 0, 1, \dots, N-1$ , в которых необходимо задать  $X_n(e^{i\omega})$  для однозначного восстановления  $x(n)$ . Если окно ограничено во времени и если  $X_n(e^{i\omega})$  рассматривать как преобразование Фурье, то и обратное преобразование окажется ограниченным во времени. В таком случае теорема отсчетов требует, чтобы мы дискретизовали  $X_n(e^{i\omega})$  в частотной области с частотой, по крайней мере вдвое превышающей «ширину временной полосы». Поскольку обратным преобразованием Фурье  $X_n(e^{i\omega})$  будет сигнал  $x(m)\omega(n-m)$  и этот сигнал имеет продолжительность  $L$  отсчетов (опять из-за конечной ширины окна), то по теореме отсчетов  $X_n(e^{i\omega})$  необходимо дискретизовать (по частоте) на частотах

$$\omega_k = 2\pi k/L, \quad k = 0, 1, \dots, L-1, \quad (6.20)$$

чтобы точно восстановить  $x(n)$  по  $X_n(e^{i\omega_k})$  (см. задачу 6.8). Таким образом, в примере с окном Хемминга шириной  $L=100$  отсчетов требуется, чтобы  $X_n(e^{i\omega})$  вычислялось по крайней мере для 100 равномерно распределенных на единичной окружности частот.

Основываясь на приведенном обсуждении, можно определить полное число отсчетов, подлежащее вычислению каждую секунду для того, чтобы избежать наложений и получить представление исходного сигнала. Минимальная частота дискретизации  $X_n(e^{i\omega_k})$  во временной области равна  $2B$ , где  $B$  — ширина полосы частот окна, а минимальное число отсчетов в частотной области равно  $L$  — ширине окна во времени. Таким образом, «полная частота дискретизации» для  $X_n(e^{i\omega_k})$  равна

$$SR = 2BL \text{ отсч./с.} \quad (6.21)$$

Для большинства практически используемых окон  $B$  можно представить как кратное от  $(F_s/L)$ , где  $F_s$  — частота дискретизации  $x(n)$ , т. е.

$$B = CF_s/L \text{ Гц.} \quad (6.22)$$

Здесь  $C$  — константа пропорциональности. Следовательно, (6.21) можно переписать в виде

$$SR = 2CF_s \text{ отсч./с.} \quad (6.23)$$

Соотношение  $SR$  и  $F_s$  поэтому равно

$$SR/F_s = 2C. \quad (6.24)$$

Величина  $2C$  служит коэффициентом «излишка дискретизации» кратковременного анализа в сравнении с обычным дискретизированным представлением  $x(n)$ .

Из примеров видно, что  $2C=4$ , если  $\omega(n)$  — окно Хемминга, и  $2C=2$ , если  $\omega(n)$  — прямоугольное окно (и если ширина полосы

определена по первому нулю  $W(e^{i\omega})$ ). Следовательно, для кратковременного спектрального представления  $x(n)$  требуется примерно в 2—4 раза больше отсчетов, чем для описания самого сигнала. Взамен, однако, получено весьма гибкое представление сигнала, позволяющее производить целый ряд манипуляций как в частотной, так и во временной областях. Хотя рассчитанные частоты дискретизации и представляют собой теоретически минимальные, все же существуют специальные случаи, в которых можно дискретизовать  $X_n(e^{i\omega_k})$  с пониженной частотой во временной или частотной области и для которых сохраняется возможность точного восстановления  $x(n)$  без ошибок наложения. Такие случаи важны при реализации систем с минимальной памятью (скоростью передачи), например систем анализа — синтеза, визуального отображения спектра и т. п. Далее в этой главе мы рассмотрим, как спроектировать и реализовать такие системы. Вначале, однако, опишем два различных способа восстановления  $x(n)$  по дискретизированной  $X_n(e^{i\omega_k})$ , а затем влияние модификаций  $X_n(e^{i\omega_k})$  на результирующий сигнал.

#### 6.1.4. Кратковременный синтез методом суммирования выходов гребенки фильтров

Первый метод синтеза связан с интерпретацией кратковременного анализа как преобразования в гребенке фильтров. Ранее было показано, что для любой частоты  $\omega_k$   $X_n(e^{i\omega_k})$  есть узкополосное представление сигнала в полосе с центральной частотой  $\omega_k$ . Из (6.1) и (6.2) можно получить для  $X_n(e^{i\omega_k})$  выражения

$$X_n(e^{i\omega_k}) = \sum_{m=-\infty}^{\infty} \omega_k(n-m)x(m)e^{-i\omega_k m} \quad (6.25)$$

или

$$X_n(e^{i\omega_k}) = e^{-i\omega_k n} \sum_{m=-\infty}^{\infty} x(n-m)\omega_k(m)e^{i\omega_k m}, \quad (6.26)$$

где  $\omega_k(m)$  — окно, используемое на частоте  $\omega_k$ . Определив

$$h_k(n) = \omega_k(n)e^{i\omega_k n}, \quad (6.27)$$

перепишем (6.26):

$$X_n(e^{i\omega_k}) = e^{-i\omega_k n} \sum_{m=-\infty}^{\infty} x(n-m)h_k(m). \quad (6.28)$$

Поскольку окно  $\omega_k(n)$  обладает свойствами фильтра нижних частот, (6.28) можно интерпретировать как полосовую фильтрацию с импульсной характеристикой  $h_k(n)$  и последующей модуляцией комплексной экспонентой  $e^{-i\omega_k n}$ , как это показано на рис. 6.7. Определив

$$y_k(n) = X_n(e^{i\omega_k})e^{i\omega_k n}, \quad (6.29)$$

получим из (6.28) следующее выражение:

$$y_k(n) = \sum_{m=-\infty}^{\infty} x(n-m) h_k(m). \quad (6.30)$$

Таким образом,  $y_k(n)$  есть просто выходная последовательность полосового фильтра с импульсной характеристикой  $h_k(n)$  (6.27). Операции, определенные (6.28) и (6.29), изображены на рис. 6.9а. Поскольку (6.25) и (6.28) эквивалентны, то в (6.29) можно использовать любое из выражений для  $X_n(e^{i\omega_k})$ . В обоих случаях система, связывающая  $x(n)$  с  $y_k(n)$ , представляет собой полосовой фильтр с импульсной характеристикой  $h_k(n)$ . Это показано на рис. 6.9, причем рис. 6.9а отображает (6.28) и (6.29), а рис. 6.9б— (6.25) и (6.29). На рис. 6.9в показан эквивалентный полосовой фильтр.

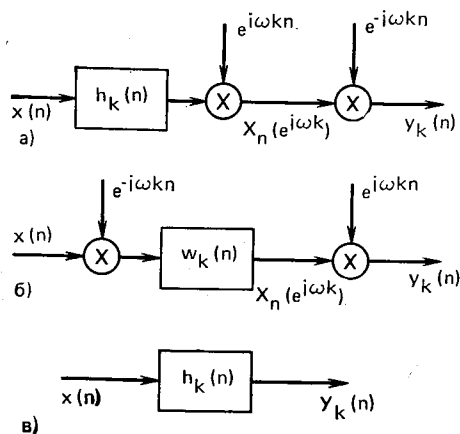


Рис. 6.9. Метод построения одного канала синтеза в рамках линейной фильтрации

$W_k(e^{i\omega})$  такого фильтра приведена на рис. 6.10а<sup>1</sup>. Частотной характеристикой соответствующего комплексного полосового фильтра с импульсной характеристикой  $h_k(n) = w_k(n) e^{i\omega_k n}$  будет тогда

$$H_k(e^{i\omega}) = W_k(e^{i(\omega - \omega_k)}), \quad (6.31)$$

как это показано на рис. 6.10б. Заметим, что центральная частота равна  $\omega_k$ , а ширина полосы равна  $2\omega_{pk}$ .

Рассмотрим набор из  $N$  полосовых фильтров с центральными частотами, равномерно разнесенными так, что перекрывается весь диапазон основных частот:

$$\omega_k = 2\pi k/N, \quad k = 0, 1, \dots, N-1. \quad (6.32)$$

Допустим также, что окно одинаково для всех каналов:

$$w_k(n) = w(n), \quad k = 0, 1, \dots, N-1. \quad (6.33)$$

<sup>1)</sup> Здесь  $\omega$  — частотная переменная.

Если теперь рассмотреть все полосовые фильтры вместе, считая, что вход одинаков, а выходы суммируются, как это показано на рис. 6.11, то общей частотной характеристикой, связывающей  $y(n)$  с  $x(n)$ , будет

$$\tilde{H}(e^{i\omega}) = \sum_{k=0}^{N-1} H_k(e^{i\omega}) = \sum_{k=0}^{N-1} W(e^{i(\omega - \omega_k)}). \quad (6.34)$$

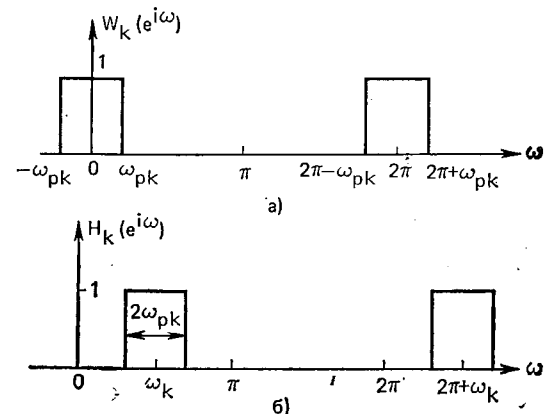


Рис. 6.10 Частотные характеристики: а) идеального ФНЧ; б) идеального полосового фильтра

Если  $W(e^{i\omega_k})$  должным образом дискретизована по частоте (т. е. если  $N \geq L$ , где  $L$  — длительность окна), то, как можно показать,

$$\frac{1}{N} \sum_{k=0}^{N-1} W(e^{i(\omega - \omega_k)}) = w(0) \quad (6.35)$$

для всех  $\omega$ .

Равенство (6.35) выводится следующим образом. Обратным преобразованием Фурье для  $W(e^{i\omega})$  служит окно  $w(n)$ . Если  $W(e^{i\omega})$  дискретизована по частоте с  $N$  равномерно разнесенными частотами, то обратное преобразование Фурье дискретизованной  $W(e^{i\omega_k})$  имеет вид

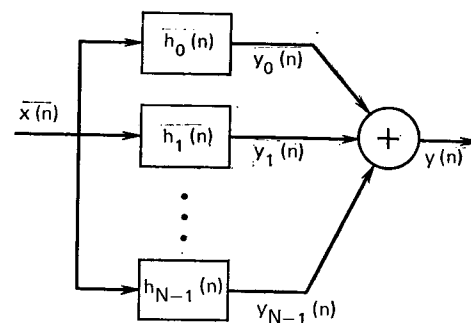


Рис. 6.11. Эквивалентная линейная система, связывающая  $y_k(n)$  и  $y(n)$  с  $x(n)$

$$\frac{1}{N} \sum_{k=0}^{N-1} W(e^{i\omega_k}) e^{i\omega_k n} = \sum_{r=-\infty}^{\infty} w(n+rN), \quad (6.36)$$

т. е. получено представление  $w(n)$  с наложениями (см. задачу 6.8). Если  $w(n)$  имеет продолжительность в  $L$  отсчетов, то

$$w(n) = 0, \quad n < 0, \quad n \geq L. \quad (6.37)$$

и наложений за счет дискретизации  $W(e^{i\omega})$  по частоте нет. В этом случае, вычислив (6.36) для  $n=0$ , получим

$$\frac{1}{N} \sum_{k=0}^{N-1} W(e^{i\omega_k}) = w(0). \quad (6.38)$$

Теперь легко получить (6.35), заметив, что  $W(e^{i(\omega-\omega_k)})$  представляет собой равномерно дискретизованную  $W(e^{i\omega})$ , вычисленную в точке  $\omega-\omega_k$  вместо  $\omega_k$ . В соответствии с теоремой отсчетов годится любой набор из  $N$  равномерно разнесенных отсчетов. Поэтому (6.35) следует из (6.38) и теоремы отсчетов.

Из (6.38) и (6.35) видно, что для системы в целом импульсная характеристика

$$\tilde{h}(n) = \sum_{k=0}^{N-1} h_k(n) = \sum_{k=0}^{N-1} w(n) e^{i\omega_k n} = N w(0) \delta(n) \quad (6.39)$$

просто совпадает с масштабированным единичным отсчетом  $Nw(0)\delta(n)$ , а следовательно, общим выходом  $y(n)$  будет  $Nw(0) \times X(n)$ .

Таким образом, в методе суммирования выходов гребенки фильтров восстановленный сигнал формируется как

$$y(n) = \sum_{k=0}^{N-1} y_k(n) = \sum_{k=0}^{N-1} X_n(e^{i\omega_k}) e^{i\omega_k n}, \quad (6.40)$$

и мы показали, что, если  $X_n(e^{i\omega_k})$  дискретизована должным образом по частоте,  $y(n) = Nw(0)x(n)$  независимо от конкретной формы окна. Операции анализа и синтеза, подразумеваемые в (6.40), изображены на рис. 6.12, причем фильтры на рисунке — полосовые.

Мы только что установили весьма важный результат: при условии, что  $w(n)$  имеет конечную длительность  $L$ , последовательность  $x(n)$  можно точно восстановить по кратковременному преобразованию Фурье, дискретизованному и по времени, и по частоте. Можно также показать, что если  $W(e^{i\omega})$  строго ограничена по частоте, то аналогично  $x(n)$  может быть точно восстановлена по  $X_n(e^{i\omega_k})$ . В действительности, существует много способов точного восстановления  $x(n)$  по кратковременному преобразованию.

Для того чтобы избежать наложений, нужно вычислить  $X_n(e^{i\omega_k})$ , по крайней мере на  $L$  равномерно разнесенных частотах, где  $L$  — длительность окна. Ширина полосы окна длительностью в  $L$  отсчетов лежит, вообще говоря, между  $2\pi/L$  (для прямоугольного окна) и  $4\pi/L$  (для окна Хемминга). Поскольку частоты анализа равны  $2\pi k/L$ , эффективные полосы полосовых фильтров пере-

крываются. Как указывалось ранее, существует способ, при котором можно вычислять  $X_n(e^{i\omega_k})$  в неперекрывающихся полосах и тем не менее восстановить  $x(n)$  точно (по крайней мере теоретически).

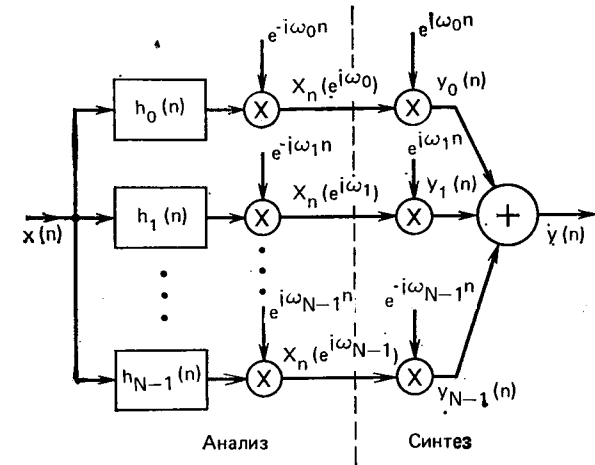


Рис. 6.12. Операции анализа и синтеза для процедуры кратковременного спектрального анализа

Чтобы показать это, допустим, что ширина окна для всех полос равна  $L$  отсчетам и что такое одно и то же окно используется для  $N$  равноразнесенных полос частот с частотами анализа:

$$\omega_k = 2\pi k/N, \quad k = 0, 1, \dots, N-1, \quad (6.41a)$$

где  $N$  может быть меньше  $L$ . Допустим также, что  $w(n)$  представляет собой идеальный фильтр нижних частот с частотой среза

$$\omega_p = \pi/N. \quad (6.41b)$$

Эта ситуация показана на рис. 6.13, где приведена общая характеристика для  $N=6$  равноразнесенных идеальных фильтров. В этом случае (6.39) переходит в

$$\tilde{h}(n) = \sum_{k=0}^{N-1} w(n) e^{i\omega_k n} = w(n) \sum_{k=0}^{N-1} e^{i\omega_k n}. \quad (6.42)$$

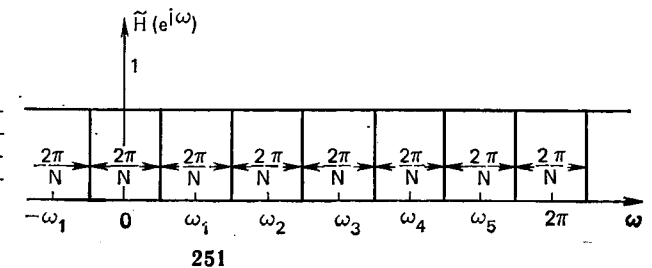


Рис. 6.13. Общая частотная характеристика для шести равноразнесенных идеальных фильтров

Определив

$$p(n) = \sum_{k=0}^{N-1} e^{i\omega_k n} = \sum_{k=0}^{N-1} e^{i2\pi kn/N}, \quad (6.43)$$

перепишем  $\tilde{h}(n)$  в виде

$$\tilde{h}(n) = w(n) p(n). \quad (6.44)$$

Легко видеть, что последовательность  $p(n)$  периодична с периодом  $N$ . В действительности можно показать (см. задачу 6.7), что  $p(n)$  есть периодическая последовательность импульсов с амплитудой  $N$ :

$$p(n) = N \sum_{r=-\infty}^{\infty} \delta(n - rN). \quad (6.45)$$

Поэтому для  $\tilde{h}(n)$  имеем

$$\tilde{h}(n) = N \sum_{r=-\infty}^{\infty} w(rN) \delta(n - rN). \quad (6.46)$$

Следовательно, общая импульсная характеристика представляет собой просто последовательность окна, дискретизованную через интервалы длиной в  $N$  отсчетов. Это показано на рис. 6.14: на рис. 6.14а — последовательность  $p(n)$ , на рис. 6.14б — импульсная характеристика идеального фильтра нижних частот с частотой среза  $\pi/N$ , т. е.

$$w(n) = \sin(\pi n/N)/\pi n. \quad (6.47)$$

Сравнивая рис. 6.14а и б можно понять, что произведение  $\tilde{h}(n) = p(n)w(n)$  равно нулю всюду, кроме точки  $n=0$ , где оно равно единице. Поэтому общая импульсная характеристика равна

$$\tilde{h}(n) = \delta(n). \quad (6.48)$$

И хотя при этом фильтр нижних частот предполагался идеальным, характер взаимодействия  $p(n)$  и  $w(n)$ , при котором образуется общая характеристика, подсказывает множество возможностей для выбора  $w(n)$  так, чтобы можно было восстановить сигнал по дискретизованному кратковременному преобразованию. Заметим, что если  $w(n)$  имеет конечную ширину  $L < N$  и физически реализуемо, то общая импульсная характеристика будет такой, как в (6.39), что и подтверждает наши рассуждения в предыдущем разделе. На рис. 6.14в приведен пример для этого случая. С другой стороны, можно использовать физически реализуемое окно, с шириной, большей  $N$ , если при этом  $w(n)$  обладает следующими свойствами:

$$w(n) = \begin{cases} 1/N, & n = r_0 N; \\ 0, & n = rN \begin{cases} r \neq r_0 \\ r = 0, \pm 1, \pm 2, \dots \end{cases} \end{cases} \quad (6.49a)$$

Тогда

$$\tilde{h}(n) = p(n)w(n) = \delta(n - r_0 N). \quad (6.49б)$$

На рис. 6.14 приведен пример окна конечной длительности, для которого  $r_0=2$ . На самом деле ясно, что для того, чтобы можно было по  $X_n(e^{i\omega})$  восстановить возможно с задержкой  $x(n)$ , вовсе не обязательно, чтобы  $w(n)$  было ограничено во времени или по частоте. Все, что необходимо — это, чтобы для  $w(n)$  (6.49а) выполнялось. На рис. 6.14д приведен пример окна бесконечной длительности с подходящими свойствами.

Смысл (6.49б) заключается в том, что общая частотная характеристика системы анализ — синтез (рис 6.12), должна иметь

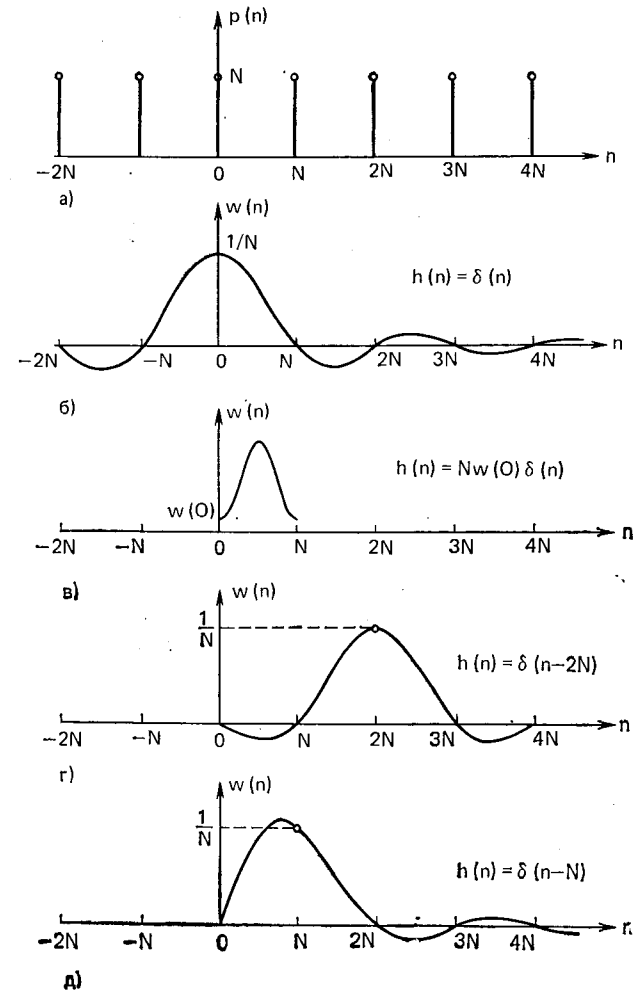


Рис. 6.14. Типичные последовательности  $p(n)$  и  $w(n)$  гребенки фильтров

плоскую амплитудно-частотную характеристику и линейную фазо-частотную, соответствующую задержке на  $r_0 N$  отсчетов, т. е.

$$\tilde{H}(e^{i\omega}) = e^{-i\omega r_0 N}. \quad (6.50)$$

Отсюда следует, что выходом системы анализ — синтез служит

$$y(n) = x(n - r_0 N). \quad (6.51)$$

Таким образом, если не считать задержки в  $r_0$  отсчетов, выход системы кратковременного анализа и синтеза Фурье представляет собой точную копию входной последовательности. Мы показали, следовательно, что возможно точное восстановление входного сигнала при числе частотных каналов, меньшем, чем необходимо по теореме отсчетов, и с помощью физически реализуемого окна, позволяющего реализовать анализ на полосовых фильтрах или фильтрах нижних частот. Возникает поэтому практически важный вопрос о том, насколько точно проектируемый фильтр будет аппроксимировать характеристики, приведенные на рис. 6.14. Этот вопрос рассмотрен в § 6.2.

Прежде чем перейти к другому методу синтеза по кратковременному спектру, нам следует обсудить способ практической реализации (6.40) (уравнение синтеза), поскольку было показано, что достаточно вычислить  $X_n(e^{i\omega_k})$  с частотой, определяемой шириной полосы окна. Допустим, что сигнал вычисляется в  $k$ -м канале только для каждого  $D_k$ -го отсчета входного сигнала. (Для равноразнесенных каналов  $D_k = D$  независимо от  $k$ .) Предположив, что  $X_n(e^{i\omega_k})$  вычисляется с частотой отсчетов входного сигнала (хотя фактически этого может и не быть), можно переделать рис. 6.9а и б, включив прореживатель на входе анализатора и интерполятор на выходе синтезатора, так, как это показано на рис. 6.15. Этим отображается утверждение о том, что  $X_n(e^{i\omega_k})$  можно дискретизовать с частотой  $F_s/D_k$ . Как уже отмечалось в гл. 2, прореживание заключается просто в том, что выкидываются  $D_k - 1$

отсчетов из каждых  $D_k$ , или, что то же самое,  $X_n(e^{i\omega_k})$  вычисляется через каждые  $D_k$  отсчетов. Интерполяция реализуется заполнением нулями  $D_k - 1$  отсчетов между значениями  $X_n(e^{i\omega_k})$ , полученными с уменьшенной частотой дискретизации и последующей фильтрацией в фильтре нижних частот.

### 6.1.5. Кратковременный синтез методом суммирования с наложением

Альтернативный метод восстановления  $x(n)$  по кратковременному спектру основан на интерпретации кратковременного спектра посредством обычного преобразования Фурье. Поскольку  $X_n(e^{i\omega_k})$  можно рассматривать как обычное дискретное преобразование Фурье последовательности

$$y_n(m) = x(m) \omega(n - m), \quad (6.52)$$

то, следовательно, можно восстановить  $x(m)$ , вычислив обратное дискретное преобразование Фурье от  $X_n(e^{i\omega_k})$  и разделив затем на окно (в предположении, что оно не равно нулю при всех  $m$ ). Этим способом можно вычислить  $L$  значений сигнала  $x(m)$ , где  $L$  — длительность окна. Затем окно сдвигается на  $L$  отсчетов и процесс повторяется. Из рассуждений, приведенных в 6.1.3, можно понять, что в этой процедуре используется представление  $X_n(e^{i\omega_k})$  с пониженной частотой дискретизации и, следовательно, оно весьма чувствительно к ошибкам наложения. Поэтому, несмотря на то, что такая процедура обоснована, она не нашла сколько-нибудь широкого применения там, где важно восстановить исходный (или преобразованный) сигнал. В этом разделе приведена более устойчивая процедура, сходная с методом вычисления периодической свертки на основе дискретного преобразования Фурье.

Допустим, что кратковременное преобразование дискретизируется с периодом, равным  $R$  отсчетов, во временной области, т. е. пусть  $Y_r(e^{i\omega_k}) = X_{rR}(e^{i\omega_k})$  где  $r$  — целое и  $0 \leq k \leq N-1$ . Метод суммирования с наложением основан на соотношении

$$y(n) = \sum_{r=-\infty}^{\infty} \left[ \frac{1}{N} \sum_{k=0}^{N-1} Y_r(e^{i\omega_k}) e^{i\omega_k n} \right]. \quad (6.53)$$

Иначе говоря, сигнал получается вычислением для каждого значения  $r$  обратного преобразования  $Y_r(e^{i\omega_k})$ , что дает последовательность

$$y_k(m) = x(m) \omega(rR - m), \quad -\infty < m < \infty. \quad (6.54)$$

Затем получают сигнал в момент  $n$ , суммируя значения всех последовательностей  $y_r(m)$ , перекрывающихся в этот момент. Таким образом,

$$y(n) = \sum_{r=-\infty}^{\infty} y_r(n) = x(n) \sum_{r=-\infty}^{\infty} \omega(rR - n). \quad (6.55)$$

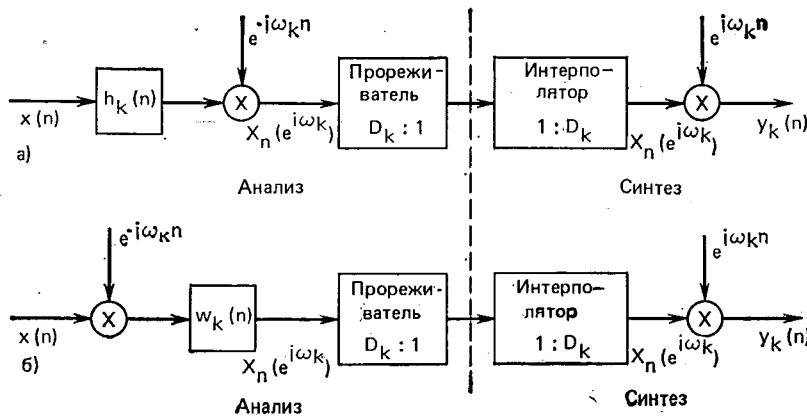


Рис. 6.15. Реализация одного канала кратковременного спектрального анализа с гребенкой фильтров

Легко показать (см. задачу 6.8), что если преобразование Фурье от  $w(n)$  ограничено по частоте и если  $X_n(e^{i\omega_k})$  дискретизирована надлежащим образом по частоте, т. е. если  $R$  достаточно мала, чтобы избежать наложения<sup>1</sup> во времени, то

$$\sum_{r=-\infty}^{\infty} w(rR-n) \approx W(e^{i0})/R \quad (6.56)$$

независимо от  $n$ .

Поэтому (6.55) переходит в

$$y(n) = x(n) W(e^{i0})/R. \quad (6.57)$$

Это показывает, что правило синтеза (6.53) приводит к точному восстановлению  $x(n)$  с точностью до множителя при суммировании перекрывающихся во времени сегментов сигнала.

На рис. 6.16 и 6.17 показано, как метод суммирования с наложением реализуется для  $L$ -точечного окна Хемминга с  $R=L/4$ . На рис. 6.16 приведена блок-схема алгоритма в предположении, что  $x(n)=0$  при  $n<0$ . Поскольку для окна Хемминга требуется перекрытие по времени 4:1, то для того, чтобы получить правильные начальные условия, начало первого анализируемого сегмента принято равным  $L/4$  (рис. 6.17). Используемое окно (предполагаемое отличным от нуля при  $0 \leq n \leq L-1$ ) дает сигнал  $y_r(m) = w(rR-m)x(m)$ , отличный от нуля для  $rR-L+1 \leq m \leq rR$ . Последовательность длиной  $L$  дополняется нулями

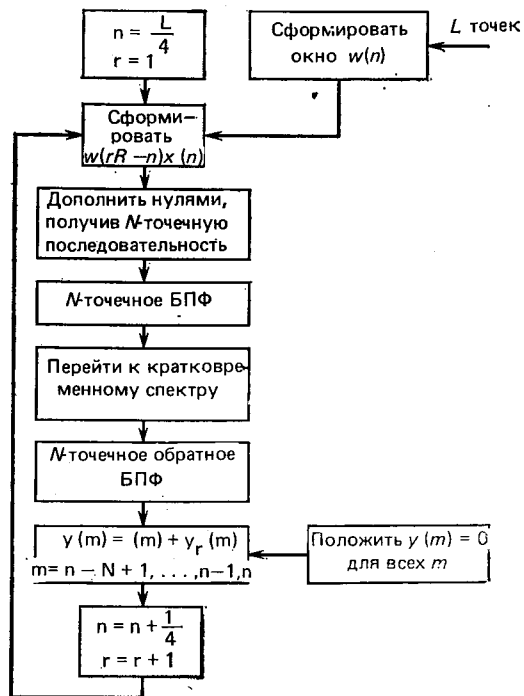


Рис. 6.16. Схема метода суммирования с наложением

для получения нужного изменения кратковременного спектра (см. следующий раздел), и затем к результирующей последовательности применяется  $N$ -точечное БПФ, дающее  $Y_r(e^{i\omega_k})$ .

Для восстановления сигнала в момент  $n$  можно воспользоваться выражением (6.53). На рис. 6.17 показаны операции, необхо-

димые для вычисления (6.53) при  $0 \leq n \leq R-1$ . Заметим, что  $y(n)$  представляется суммой четырех членов

$$y(n) = x(n)w(R-n) + x(n)w(2R-n) + x(n)w(3R-n) + x(n)w(4R-n). \quad (6.58)$$

При  $R \leq n \leq 2R-1$  член  $x(n)w(R-n)$  следует заменить членом  $x(n)w(5R-n)$  и т. д.

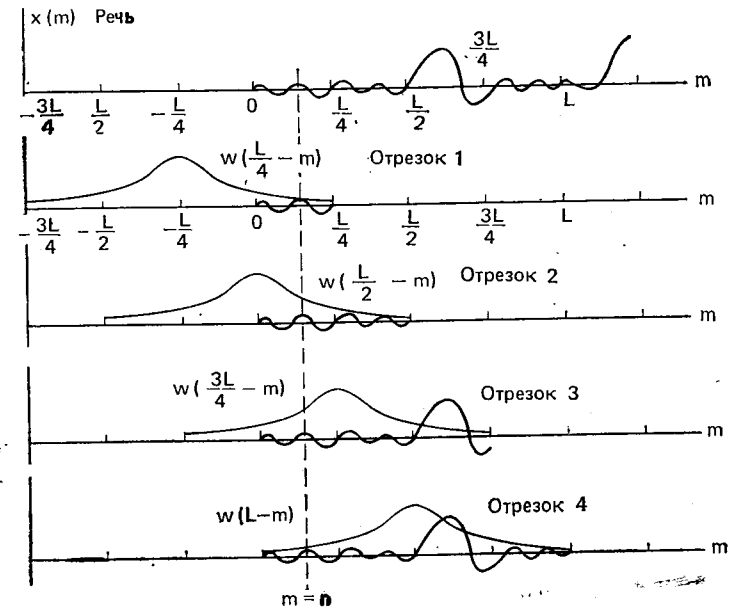


Рис. 6.17. Восстановление  $w(n)$  с  $L$ -точечным окном Хемминга

Метод суммирования выходов гребенки фильтров и метод суммирования с наложением по сути дела дуальны: один связан с дискретизацией по частоте, а другой — с дискретизацией во времени. Метод суммирования выходов гребенки фильтров требует, чтобы при дискретизации по частоте преобразование окна удовлетворяло соотношению

$$\sum_{k=0}^{N-1} W(e^{i(\omega-\omega_k)}) = w(0), \quad (6.59a)$$

в то время как в методе суммирования с наложением требуется, чтобы при дискретизации по времени окно удовлетворяло равенству

$$\sum_{r=-\infty}^{\infty} w(rR-n) = W(e^{i0})/R. \quad (6.59b)$$

Дуальность (6.59a и б) очевидна.

<sup>1</sup> Для  $L$ -точечного окна Хемминга  $R \leq L/4$ .

Для того чтобы сопоставить эти два метода восстановления сигнала по кратковременному преобразованию Фурье, ниже рассматривается влияние изменений кратковременного спектра на результат синтеза.

### 6.1.6. Влияние преобразований кратковременного спектра на синтез

Итак, имеется два метода восстановления сигнала по его кратковременному спектру. Оба метода позволяют точно восстановить исходный сигнал с точностью до множителя, если кратковременный спектр должным образом дискретизован как по частоте, так и по времени. Иногда требуется изменить кратковременный спектр для того, чтобы произвести постоянную или динамичную (т. е. переменную во времени) фильтрацию анализируемого сигнала. Рассмотрим влияние фиксированных или переменных во времени преобразований кратковременного спектра на результат синтеза.

**Метод суммирования выходов гребенки фильтров (СГФ).** Представим фиксированное преобразование кратковременного спектра в виде

$$\hat{X}_n(e^{i\omega_k}) = X_n(e^{i\omega_k}) P(e^{i\omega_k}), \quad (6.60)$$

где  $P(e^{i\omega_k})$  — весовая частотная функция кратковременного спектра. Предположим, что существует обратное дискретное преобразование Фурье от  $P(e^{i\omega_k})$ , и обозначим эту последовательность через  $p(n)$ :

$$p(n) = \frac{1}{N} \sum_{k=0}^{N-1} P(e^{i\omega_k}) e^{i\omega_k n}, \quad (6.61)$$

где  $N$  — количество частот, на которых вычисляется  $P(e^{i\omega_k})$ . Сигнал, восстановленный методом СГФ, получается подстановкой (6.60) в (6.40):

$$\begin{aligned} \hat{y}(n) &= \sum_{k=0}^{N-1} X_n(e^{i\omega_k}) P(e^{i\omega_k}) e^{i\omega_k n} = \sum_{k=0}^{N-1} \left[ \sum_{m=-\infty}^{\infty} \omega(n-m) x(m) e^{-i\omega_k m} \right] P(e^{i\omega_k}) e^{i\omega_k n} = \sum_{m=-\infty}^{\infty} \omega(n-m) x(m) \times \\ &\times \sum_{k=0}^{N-1} P(e^{i\omega_k}) e^{i\omega_k(n-m)} = \sum_{m=-\infty}^{\infty} \omega(n-m) x(m) N p(n-m) = \\ &= N x(n) * [\omega(n) p(n)]. \end{aligned} \quad (6.62)$$

Следовательно, эффект постоянного во времени преобразования с  $P(e^{i\omega_k})$  сводится к свертке сигнала  $x(n)$  с произведением окна  $\omega(n)$  и периодической последовательности  $p(n)$ . Преобразование кратковременного спектра вида (6.60) применяется тогда, когда

требуется линейно отфильтровать  $x(n)$ . Можно, например, потребовать, чтобы

$$\omega(n) p(n) = h_p(n) \quad (6.63)$$

была импульсной характеристикой линейного фильтра. Последовательность  $p(n)$  представляет собой периодическую последовательность, и если длительность  $\omega(n)$  не превосходит  $N$ , то структура  $h_p(n)$  окажется повторяющейся. В следующем разделе будет показано, что преобразования такого вида возникают при построении гребенок БИХ-фильтров. Таким образом, в методе суммирования выходов гребенки фильтров окно существенно влияет на фиксированное преобразование спектра. Справедливо лишь приближенное равенство

$$h_p(n) \approx p(n), \quad 0 \leq n \leq N-1, \quad (6.64)$$

да и то только в случае, когда  $p(n)$  сильно сконцентрирована или когда используется прямоугольное окно.

Для переменного во времени преобразования запишем  $\hat{X}_n(e^{i\omega_k})$  в виде

$$\hat{X}_n(e^{i\omega_k}) = X_n(e^{i\omega_k}) P_n(e^{i\omega_k}) \quad (6.65)$$

и определим переменную во времени импульсную характеристику  $p_n(m)$  так:

$$p_n(m) = \frac{1}{N} \sum_{k=0}^{N-1} P_n(e^{i\omega_k}) e^{i\omega_k m}. \quad (6.66)$$

Поступая как и раньше, получим

$$\begin{aligned} \hat{y}(n) &= \sum_{k=0}^{N-1} X_n(e^{i\omega_k}) P_n(e^{i\omega_k}) e^{i\omega_k n} = \sum_{k=0}^{N-1} e^{-i\omega_k n} \sum_{m=-\infty}^{\infty} x(n-m) \omega(n-m) e^{i\omega_k m} P_n(e^{i\omega_k}) e^{i\omega_k n} = \sum_{m=-\infty}^{\infty} x(n-m) \omega(n-m) \times \\ &\times \sum_{k=0}^{N-1} P_n(e^{i\omega_k}) e^{i\omega_k m} = \sum_{m=-\infty}^{\infty} x(n-m) \omega(n-m) N p_n(m) = N \sum_{m=-\infty}^{\infty} x(n-m) \omega(n-m) [p_n(m) \omega(n-m)]. \end{aligned} \quad (6.67)$$

Равенства (6.67) снова показывают, что в методе СГФ временная характеристика преобразования спектра взвешивается окном до свертывания с  $x(n)$ .

Итак, в методе суммирования выходов гребенки фильтров влияние преобразования спектра (постоянного или переменного во времени) сводится к свертыванию исходного сигнала с взвешенной временной характеристикой преобразования.

Метод суммирования с наложением (СН). Воспользовавшись (6.60), с помощью (6.53) можно выразить восстановленный сигнал в виде

$$\begin{aligned} \hat{y}(n) &= \sum_{r=-\infty}^{\infty} \frac{1}{N} \sum_{k=0}^{N-1} Y_r(e^{i\omega_k}) P(e^{i\omega_k}) e^{i\omega_k n} = \frac{1}{N} \sum_{r=-\infty}^{\infty} \sum_{k=0}^{N-1} \sum_{l=-\infty}^{\infty} x(l) \omega \times \\ &\times (rR-l) e^{-i\omega_k l} P(e^{i\omega_k}) e^{i\omega_k n} = \frac{1}{N} \sum_{l=-\infty}^{\infty} x(l) \left[ \sum_{k=0}^{N-1} P(e^{i\omega_k}) e^{i\omega_k(n-l)} \right] \times \\ &\times \left[ \sum_{r=-\infty}^{\infty} \omega(rR-l) \right] = \sum_{l=-\infty}^{\infty} x(l) p(n-l) W(e^{i0})/R \end{aligned} \quad (6.68)$$

или

$$\hat{y}(n) = (1/R) W(e^{i0}) [x(n) * p(n)]. \quad (6.69)$$

Соотношение (6.69) показывает, что  $\hat{y}(n)$  представляет собой свертку исходного сигнала с временной характеристикой преобразования спектра, т. е. в этом методе окно не изменяет последовательности  $p(n)$ <sup>1</sup>. (Читателю следует отдавать себе отчет в том, что схему, изображенную на рис. 6.16, следует видоизменить должным образом — дополнить нулями сигнал, — чтобы избежать наложений при реализации операций анализа и синтеза посредством БПФ).

Для случая переменного преобразования получим

$$\hat{y}(n) = \sum_{r=-\infty}^{\infty} \frac{1}{N} \left[ \sum_{k=0}^{N-1} Y_r(e^{i\omega_k}) P_r(e^{i\omega_k}) \right] e^{i\omega_k n}, \quad (6.70)$$

что можно преобразовать к виду

$$\hat{y}(n) = \frac{1}{N} \sum_{l=-\infty}^{\infty} x(l) \sum_{r=-\infty}^{\infty} \omega(rR-l) \left[ \sum_{k=0}^{N-1} P_r(e^{i\omega_k}) e^{i\omega_k(n-l)} \right]. \quad (6.71)$$

Из (6.66) получим

$$\hat{y}(n) = \sum_{l=-\infty}^{\infty} x(l) \sum_{r=-\infty}^{\infty} \omega(rR-l) p_r(n-l). \quad (6.72)$$

Положив  $q=n-l$  или  $l=n-q$ , получим для (6.72)

$$\hat{y}(n) = \sum_{q=-\infty}^{\infty} x(n-q) \sum_{r=-\infty}^{\infty} p_r(q) \omega(rR-n+q). \quad (6.73)$$

Если определить  $\hat{p}$  как

$$\hat{p}(n-q, q) = \hat{p}(m, q) = \sum_{r=-\infty}^{\infty} p_r(q) \omega(rR-m), \quad (6.74)$$

<sup>1</sup> Из-за вычисления выхода блоками по  $N$  отсчетов в (6.68),  $p(n)$  не периодична и имеет максимальную длительность, равную  $N$  отсчетам.

соотношение (6.72) перейдет в

$$\hat{y}(n) = \sum_{q=-\infty}^{\infty} x(n-q) \hat{p}(n-q, q). \quad (6.75)$$

Интерпретация (6.74) такова: для  $q$ -го значения  $\hat{p}(m, q)$  представляет свертку  $p_r(q)$  и  $\omega(r)$ . Поэтому каждый из коэффициентов временной характеристики преобразования сглаживается окном (фильтруется фильтром нижних частот). Следовательно, в методе суммирования с наложением любое преобразование ограничивается по частоте окном, но соответствует обычной свертке. Это прямо противоположно методу суммирования выходов гребенки фильтров, в котором преобразование ограничивается во времени окном и может мгновенно меняться.

### 6.1.7. Аддитивное преобразование

Мы рассмотрели влияние неслучайного мультипликативного преобразования кратковременного спектра. Важно также понимать влияние аддитивного, не зависящего от сигнала (случайного) изменения кратковременного спектра, такого, какое может возникнуть при реализации анализа с конечной точностью вычисления (шумы округления) или при квантовании спектра, как это происходит в вокодерах. Запишем такое аддитивное преобразование спектра в виде

$$\hat{X}_n(e^{i\omega_k}) = X_n(e^{i\omega_k}) + E_n(e^{i\omega_k}), \quad (6.76)$$

где шумовая последовательность  $E_n(e^{i\omega_k})$  определена так:

$$e(n) = \sum_{k=0}^{N-1} E_n(e^{i\omega_k}) e^{i\omega_k n}. \quad (6.77)$$

В случае, когда  $e(n)$  представляет собой случайный шум, требуется статистическая модель для  $e(n)$  и  $E(e^{i\omega_k})$ . Приводимые ниже результаты не зависят от природы модели. В методе СГФ эффект аддитивного преобразования (6.76) сводится к

$$\hat{y}(n) = \sum_{k=0}^{N-1} [X_n(e^{i\omega_k}) + E_n(e^{i\omega_k})] e^{i\omega_k n}, \quad (6.78)$$

что, ввиду линейности, можно записать следующим образом:

$$\hat{y}(n) = y(n) + \sum_{k=0}^{N-1} E_n(e^{i\omega_k}) e^{i\omega_k n} \quad (6.79)$$

или

$$\hat{y}(n) = y(n) + e(n). \quad (6.80)$$

Следовательно, аддитивное преобразование спектра приводит к появлению аддитивной компоненты в восстановленном сигнале. Следует заметить, что «анализирующее» окно не оказывает прямого влияния на аддитивные члены при синтезе.



Для СН метода аддитивное преобразование (6.76) сводится к

$$\hat{y}(n) = \sum_{r=-\infty}^{\infty} \frac{1}{N} \sum_{k=0}^{N-1} (Y_r(e^{i\omega_k}) + E_r(e^{i\omega_k})) e^{i\omega_k n}, \quad (6.81)$$

что можно переписать в виде

$$\hat{y}(n) = y(n) = \sum_{r=-\infty}^{\infty} \left[ \frac{1}{N} \sum_{k=0}^{N-1} E_r(e^{i\omega_k}) e^{i\omega_k n} \right] = y(n) + \sum_{r=-\infty}^{\infty} e_r(n). \quad (6.82)$$

Следовательно, результат синтеза содержит большую аддитивную (шумовую) компоненту в случае СН метода по сравнению с методом СГФ из-за перекрытий между сегментами анализа. Для окна Хемминга с перекрытием 4 : 1 аддитивная добавка будет примерно в 4 раза больше при синтезе методом СН в сравнении с методом СГФ. Таким образом, метод СН оказывается более чувствительным к ошибкам вычислений, чем СГФ, и, следовательно, менее полезен для применения в вокодерах и т. п.

### 6.1.8. Обзор методов кратковременного анализа и синтеза речи

В этой главе было показано, что полезное определение кратковременного преобразования Фурье имеет вид

$$X_n(e^{i\omega}) = \sum_{m=-\infty}^{\infty} \omega(n-m) x(m) e^{-i\omega m},$$

где  $\omega(n)$  — анализирующее окно, выделяющее часть входного сигнала в момент  $n$ . Показано также, что  $X_n(e^{i\omega})$  можно интерпретировать посредством линейной фильтрации, как выход полосового фильтра с близкой к нулю полосой пропускания или, равным образом, как обычное преобразование Фурье последовательности  $\omega(n-m)x(m)$ .

Показано также, что можно определить частоты дискретизации по времени и частоте, основываясь на теореме отсчетов и рассматривая представление окна в частотной и временной областях. Требуемые частоты надлежащим образом дискретизированного кратковременного спектрального представления оказались в 2—4 раза выше, чем в эквивалентном представлении самого сигнала во временной области.

На основе двух интерпретаций кратковременного анализа были введены две различные процедуры синтеза. В первом методе, названном суммированием выходов гребенки фильтров, сигнал синтезируется так:

$$y(n) = \sum_{k=0}^{N-1} X_n(e^{i\omega_k}) e^{i\omega_k n},$$

т. е. выходной сигнал представляет собой сумму сигналов гребенки фильтров, смещенных к центральным частотам полос.

Второй метод синтеза, названный методом суммирования с наложением, приводит к следующей процедуре синтеза:

$$y(n) = \sum_{r=-\infty}^{\infty} \frac{1}{N} \sum_{k=0}^{N-1} Y_r(e^{i\omega_k}) e^{i\omega_k n},$$

где  $Y_r(e^{i\omega_k}) = X_{rR}(e^{i\omega_k})$ , т. е. взвешенные окном сегменты, разнесенные по времени на  $R$  отсчетов, суммируются, образуя восстановленный сигнал. Эти два метода синтеза обладают двойственными свойствами, дуальными как по отношению к синтезу, так и по реакции на преобразование кратковременного спектра.

На этом мы заканчиваем формальное обсуждение общих свойств методов кратковременного анализа и синтеза речи. В последующих параграфах внимание сконцентрировано на методах проектирования цифровых гребенок фильтров для случая пониженной частоты дискретизации, например для вокодера, и обсуждаются некоторые из многочисленных приложений теории кратковременного анализа и синтеза к обработке речи.

## 6.2. Проектирование гребенок цифровых фильтров

Ниже будут рассмотрены некоторые практические методы проектирования цифровых фильтров, используемых в методе суммирования выходов гребенки фильтров. Нашей целью будет проектирование гребенки фильтров с общей частотной характеристикой, хорошо аппроксимирующей идеальную — с плоской амплитудно-частотной и линейной фазо-частотной частями. Начнем с некоторых подробностей, общих для проектирования гребенок фильтров, независимо от типа используемых фильтров. Затем приведем примеры использования БИХ- и КИХ-фильтров.

### 6.2.1. Соображения практического характера

Из ранее изложенного теоретического анализа следует, что достигнуть точного воспроизведения входного сигнала на выходе гребенки равномерно разнесенных фильтров можно, когда нули фильтра нижних частот (или анализирующего окна  $\omega(n)$ ) равномерно разнесены с интервалом, равным  $N$  отсчетам. Казалось бы, что при проектировании такой гребенки нужно сначала выбрать количество фильтров, а следовательно, разнесение по частоте. Затем спроектировать фильтр нижних частот с надлежащим разрешением по частоте и разнесом нулей во временной области. К сожалению, возникает ряд соображений практического характера, усложняющих эту процедуру. Во-первых, часто желательно использовать неравномерное разнесение фильтров, поэтому данное выше доказательство того, что можно получить безупречную общую характеристику, не дает каких-либо указаний по поводу практического решения. Во-вторых, некоторая часть спектра на практике часто анализу не подвергается. Следовательно, необходимо рассмотреть, как повлияет отсутствие фильтра на общую характеристику. И, наконец, большинство процедур проектирования фильтров нижних частот не допускает введения одновременных ограни-

чений на частотную и временную характеристики. Поэтому может оказаться невозможным получить и требуемое разрешение по частоте и необходимое распределение нулей импульсной характеристики.

Для того чтобы представить в простом виде некоторые из таких специальных соображений, а также превосходящая разрешение некоторых из указанных трудностей, полезно рассмотреть несколько модифицированную структуру гребенки фильтров<sup>1</sup>, в которой в каждом из каналов (см. рис. 6.11 или равным образом рис. 6.12) комплексный сигнал умножается на комплексную величину, обозначенную через  $P_k = |P_k| e^{i\phi_k}$ . Для одного канала этот случай показан на рис. 6.18. Заметим, что поскольку каждый из каналов представляет собой линейную систему, умножение можно отнести как ко входу, так и к выходу. На рис. 6.18б и в комплексная кон-

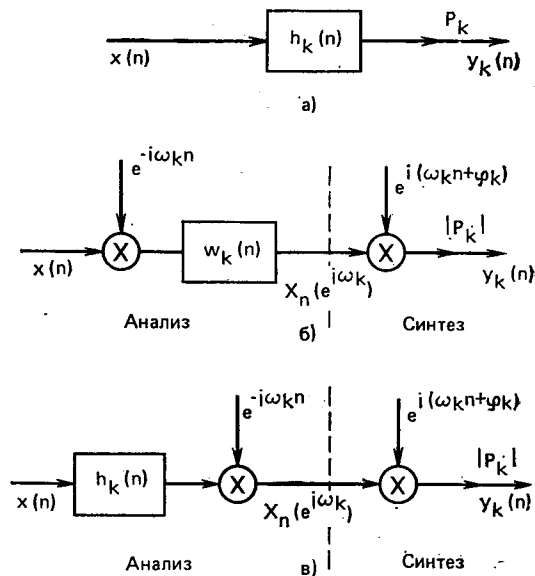


Рис. 6.18. Канал модифицированной гребенки фильтров: а) эквивалентный полосовой фильтр для процедуры анализ—синтез; б) реализация с модулятором и фильтром нижних частот; в) реализация с модулятором и полосовым фильтром

станта включена в часть канала, относящуюся к синтезу. В другом случае, когда умножение выполняется на входе, выходом блока анализа будет  $P_k X_n(e^{i\omega_k})$ . В каждом из этих случаев, рассмотрим общий выход системы  $N$  комплексных каналов, получим

$$y(n) = \sum_{k=0}^{N-1} P_k y_k(n) = \sum_{k=0}^{N-1} P_k X_n(e^{i\omega_k}) e^{i\omega_k n}. \quad (6.83)$$

<sup>1)</sup> Эта модификация относится к преобразованиям общего вида, рассмотренным в 6.1.6.

Общая импульсная характеристика системы имеет вид

$$\tilde{h}(n) = \sum_{k=0}^{N-1} P_k h_k(n) = \sum_{k=0}^{N-1} |P_k| w_k(n) e^{i(\omega_k n + \phi_k)}. \quad (6.84)$$

Таким образом, в терминах кратковременного анализа — синтеза Фурье  $X_n(e^{i\omega_k})$  взвешивается комплексной последовательностью  $\{P_k\}$ ,  $k=0, 1, \dots, N-1$ . В случае гребенки фильтров комплексные константы  $P_k$  позволяют изменять усиление и фазу отдельных фильтров гребенки.

Первый шаг в проектировании системы с гребенкой фильтров (или системы кратковременного анализа — синтеза Фурье) состоит в выборе частот анализа  $\{\omega_k\}$  для  $0 \leq k \leq N-1$ . При их выборе обычно руководствуются требованиями разрешения по частоте. Например, если требуется, чтобы основная частота и ее гармоники оказались разделенными, понадобятся близкие частоты анализа и полосовые или фильтры нижних частот с достаточно узкими полосами. Во многих случаях требуются равномерно разнесенные частоты анализа и фильтры с одинаковой шириной полосы. Однако иногда приходится использовать неравномерное распределение частот анализа. Так будет, например, в случае системы обработки речи, в которой используется ухудшение чувствительности слуха на высоких частотах. Обычно выбирают симметричный набор частот в диапазоне  $0 \leq \omega \leq 2\pi$ , т. е. так, что  $\omega_{N-k} = 2\pi - \omega_k$ . Это показано на рис. 6.19 для четных и нечетных  $N$ . Заметим, что для четных  $N$  имеется канал с центральной частотой  $\omega = \pi$ . Если к тому же  $w_k(n) = w_{N-k}(n)$ , то, как и ранее,

$$X_n(e^{i\omega_k}) = X_n^*(e^{i(2\pi - \omega_k)}) = X_n^*(e^{-i\omega_k}), \quad (6.85)$$

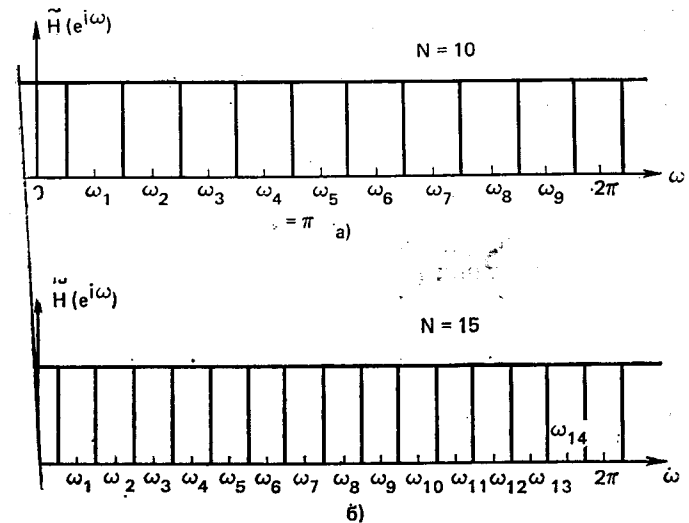


Рис. 6.19. Распределение частот анализа для гребенки фильтров с четным (а) и нечетным (б)  $N$

что приводит к упрощению гребенки фильтров, поскольку здесь необходимы только действительные полосовые фильтры.

Допустим, что  $\omega_k = 2\pi - \omega_{N-k}$ ,  $P_k = P_{N-k}^*$  и  $\omega_k(n) = \omega_{N-k}(n)$ . Тогда из (6.84) можно показать, что для четных  $N$

$$\tilde{h}(n) = P_0 \omega_0(n) + \sum_{k=1}^{\frac{N}{2}-1} 2|P_k| \omega_k(n) \cos(\omega_k n + \Phi_k) + P_{N/2} \omega_{N/2} \times (n) (-1)^n, \quad (6.86)$$

а для  $N$  нечетных

$$\tilde{h}(n) = P_0 \omega_0(n) + \sum_{k=1}^{\frac{N-1}{2}} 2|P_k| \omega_k(n) \cos(\omega_k n + \Phi_k). \quad (6.87)$$

Поэтому можно ограничиться гребенкой, состоящей из фильтров нижних частот  $P_0 \omega_0(n)$  и набора полосовых фильтров с действительными импульсными характеристиками:

$$h_k(n) = 2|P_k| \omega_k(n) \cos(\omega_k n + \Phi_k). \quad (6.88)$$

Если  $N$  четное, то потребуется дополнительный фильтр высоких частот с центральной частотой  $\omega = \pi$ ; тогда окажется перекрытым весь диапазон частот  $0 \leq \omega < 2\pi$ . Импульсная характеристика этого фильтра должна быть равной  $P_{N/2} \omega_{N/2}(n) (-1)^n$ .

За исключением того, что частоты анализа должны быть размещены симметрично на интервале  $0 \leq \omega < 2\pi$ , нет никаких других ограничений на частоты  $\{\omega_k\}$ . Коль скоро эти частоты выбраны, нам нужно найти соответствующий набор фильтров нижних частот или анализирующих окон  $\{\omega_k(n)\}$  с требуемым частотным разрешением и нужной общей характеристикой.

Часто необходимо вычислять  $X_n(e^{i\omega_k})$  только на частотах некоторого поддиапазона основной полосы  $0 \leq \omega < 2\pi$ . Например, часто не проводят анализа на частоте  $\omega_k = 0$ , поскольку эта часть речевого спектра не представляет интереса в большинстве систем обработки речи. Аналогичным образом опускают анализ на высоких частотах ( $\omega_k$ , близких к  $\pi$ ), поскольку в процессе ограничения спектра входного сигнала, предшествующего дискретизации, эта часть спектра сильно ослаблена и поэтому несет мало достоверной информации.

Чтобы оценить зависимость общей характеристики от пропуска каналов, вернемся к случаю разноразнесенных частот анализа. Предположив, что имеются идентичные анализирующие окна, получим из (6.84)

$$\tilde{h}(n) = \omega(n) \sum_{k=0}^{N-1} P_k e^{i2\pi kn/N}. \quad (6.89)$$

Следовательно, если определить

$$p(n) = \sum_{k=0}^{N-1} P_k e^{i2\pi kn/N}, \quad (6.90)$$

то

$$\tilde{h}(n) = \omega(n) p(n) \quad (6.91)$$

в соответствии с (6.89). Видно, что, как и раньше, последовательность  $p(n)$  периодична с периодом  $N$ . В действительности комплексные коэффициенты усиления в каналах играют роль коэффициентов ряда Фурье. Заметим, что в случае, когда  $P_k = 1$ ,  $0 \leq k \leq N-1$ , все каналы оказываются включенными и (6.89) становится идентичным (6.42). Эффект пропуска каналов весьма удобно наблюдать, положив соответствующие  $P_k$  равными нулю. Для того чтобы исключить, например, нулевую частоту, можно положить  $P_0 = 0$ . Чтобы, кроме того, исключить каналы с частотами выше  $\omega_M = 2\pi M/N$ , положим  $P_k = 0$  для  $k > M$ . В этом случае

$$p(n) = \sum_{k=1}^M e^{i\frac{2\pi}{N} kn} + \sum_{k=N-M}^{N-1} e^{i\frac{2\pi}{N} kn} = \sum_{k=1}^M \left[ e^{i\frac{2\pi}{N} kn} + e^{-i\frac{2\pi}{N} kn} \right]. \quad (6.92)$$

Это можно выразить в более компактной форме

$$p(n) = \frac{\sin\left[\frac{\pi}{N}(2M+1)n\right]}{\sin\left[\frac{\pi}{N}n\right]} - 1. \quad (6.93)$$

В качестве примера на рис. 6.20 изображена последовательность (6.93) для случая  $N=15$  и  $M=2$ . Ясно, что здесь  $p(n)$  периодична

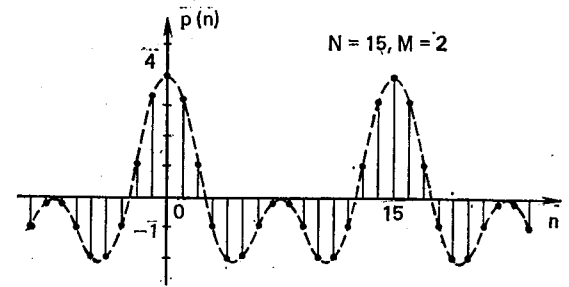


Рис. 6.20. График  $p(n)$  для гребенки фильтров с  $N=15$ ,  $M=2$

с периодом 15, но вместо одного единичного импульса через каждые 15 отсчетов  $p(n)$  теперь содержит импульсы с амплитудой и шириной, зависящими от  $N$  и  $M$ . Можно понять, что если канал на нулевой частоте будет включен, то член  $-1$  в правой части (6.93) исчезнет. Кроме того, для нечетного  $N$  при наличии всех каналов будет  $M = (N-1)/2$ , так что  $p(n)$  можно выразить в виде

$$p(n) = \frac{\sin(\pi n)}{\sin(\pi n/N)} = N \sum_{r=-\infty}^{\infty} \delta(n-rN). \quad (6.94)$$

Таким образом, только в случае, когда имеются все каналы, последовательность  $p(n)$  может быть такой, что окна  $w(n)$  с разнесенными на интервалы в  $N$  отсчетов нулями достаточно для выполнения равенства

$$\tilde{h}(n) = \delta(n - r_0 N). \quad (6.95)$$

Это, разумеется, оправдано, поскольку присутствует не весь частотный спектр. Разумно предположить, что в области включенных фильтров общая амплитудно-частотная характеристика будет плоской, а общая фазо-частотная — линейной. Выбрасывание канала на нулевой частоте и каналов на высших частотах эквивалентно полосовой фильтрации. Это подтверждается на последующих примерах.

Многие из стандартных методов проектирования фильтров не допускают введения одновременных ограничений на частотную и временную характеристики. Следовательно, может оказаться невозможным получить фильтр нижних частот с импульсной характеристикой, обращающейся в нуль через каждые  $N$  отсчетов. Чтобы уяснить влияние этого обстоятельства, рассмотрим рис. 6.21.

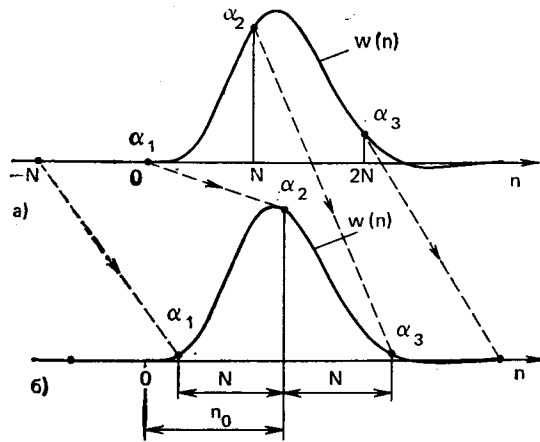


Рис. 6.21. Иллюстрация подстройки параметра  $n_0$ :  
 а) общая импульсная характеристика для  $n_0 = 0$ ; б)  $n_0$  выбрано так, чтобы минимизировать амплитудную и фазовую ошибки (пунктиром показано перемещение отдельных импульсов) [1]

Для удобства здесь принято, что задействованы все каналы так, что  $p(n)$  представляет собой последовательность единичных импульсов с периодом  $N$ . Огибающая последовательности  $w(n)$  показана как непрерывная кривая. Произведение  $p(n)$  и  $w(n)$  представлено отсчетами, обозначенными через  $\alpha_1, \alpha_2, \alpha_3$  и т. д. Ясно, что в этом случае для общей характеристики имеем приближенно

$$\tilde{h}(n) = \alpha_2 \delta(n - N) + \alpha_3 \delta(n - 2N), \quad (6.96)$$

если пренебречь импульсами в моменты  $3N, 4N$  и т. д. Общая частотная характеристика будет иметь вид

$$\begin{aligned} \tilde{H}(e^{i\omega}) &= \alpha_2 e^{-i\omega N} + \alpha_3 e^{-i\omega 2N} = \alpha_2 e^{-i\omega N} \left( 1 + \frac{\alpha_3}{\alpha_2} e^{-i\omega N} \right) = \\ &= \alpha_2 e^{-i\omega N} \tilde{G}(e^{i\omega}), \end{aligned} \quad (6.97)$$

где

$$\tilde{G}(e^{i\omega}) = 1 + (\alpha_3/\alpha_2) e^{-i\omega N} \quad (6.98)$$

задает отклонение общей частотной характеристики от  $\alpha_2 e^{-i\omega N}$ . Последняя соответствует точному воспроизведению входного сигнала с задержкой в  $N$  отсчетов и при масштабировании коэффициентом  $\alpha_2$ . Амплитуда и фаза  $\tilde{G}(e^{i\omega})$  задаются соответственно

$$|\tilde{G}(e^{i\omega})| = \left[ 1 + \left( \frac{\alpha_3}{\alpha_2} \right)^2 + 2 \left( \frac{\alpha_3}{\alpha_2} \right) \cos \omega N \right]^{1/2} \quad (6.99)$$

$$\arg[\tilde{G}(e^{i\omega})] = \text{tg}^{-1} \left[ \frac{-(\alpha_3/\alpha_2) \sin \omega N}{1 + (\alpha_3/\alpha_2) \cos \omega N} \right]. \quad (6.100)$$

Набросок этих функций для  $N=4$  приведен на рис. 6.22. Можно видеть, что модуль общей характеристики имеет множитель ошибки, который, вообще говоря, осциллирует с периодом  $2\pi/N$ .

Это вызвано разносом между фильтрами. Действительно, пики  $|\tilde{G}(e^{i\omega})|$  приходятся на частоты анализа  $\omega_k = (2\pi/N)k$ , а провалы лежат строго посередине между ними, т. е. в «перекрестных» точках фильтров. Величина пульсаций общей характеристики зависит от соотношения между  $\alpha_2$  и  $\alpha_3$ . И для фазы видно уклонение от линейной кривой; оно также периодически с периодом  $2\pi/N$ . Отметим, что амплитудная и фазовая ошибки исчезают при  $\alpha_3=0$  и максимальны при  $\alpha_3=\alpha_2$ .

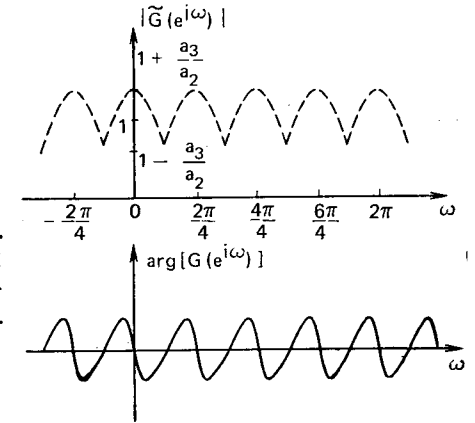


Рис. 6.22. Амплитуда и фаза гребенки фильтров

На рис. 6.21 предлагается несколько подходов к уменьшению этих погрешностей. Одна из возможностей состоит в использовании более узкого окна, что позволяет добиться уменьшения  $\alpha_3$  путем снижения  $w(n)$  при  $n=2N$ . Неприятным эффектом оказывается расширение преобразования Фурье  $w(n)$  и, следовательно, сокращение частотного разрешения в отдельных каналах. Вторая возможность заключается в том, чтобы увеличить  $N$ , сохранив  $w(n)$  без изменений. Третья возможность показана на рис. 6.21б. Если сдвинуть бесконечную последовательность импульсов относительно  $w(n)$ , то получим последовательность из трех импульсов вместо двух:

$$\tilde{h}(n) = \alpha'_1 \delta(n - n_0) + \alpha'_2 \delta(n - N - n_0) + \alpha'_3 \delta(n - 2N - n_0) \quad (6.101)$$

Тогда общей частотной характеристикой будет функция

$$\tilde{H}(e^{i\omega}) = e^{-i\omega(n_0+N)} [\alpha'_1 e^{i\omega N} + \alpha'_2 + \alpha'_3 e^{-i\omega N}]. \quad (6.102)$$

Если определить, как и раньше,

$$\tilde{G}'(e^{i\omega}) = \frac{\alpha'_1}{\alpha'_2} e^{i\omega N} + 1 + \frac{\alpha'_3}{\alpha'_2} e^{-i\omega N}, \quad (6.103)$$

то

$$\tilde{H}(e^{i\omega}) = \alpha'_2 e^{-i\omega(n_0+N)} \tilde{G}'(e^{i\omega}). \quad (6.104)$$

Ясно, что для  $\alpha'_1 = \alpha'_3$

$$\tilde{G}'(e^{i\omega}) = (2\alpha'_3/\alpha'_2) \cos \omega N + 1, \quad (6.105)$$

т. е. фазовая ошибка отсутствует. Амплитудная ошибка отложена для  $N=4$  (рис. 6.23). В первом случае максимум ошибки равен

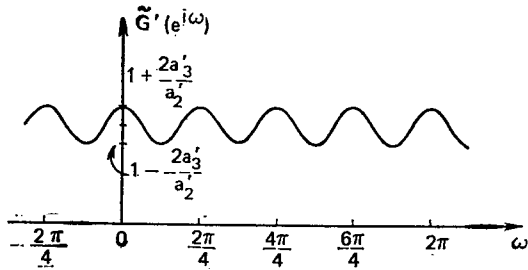


Рис. 6.23. Амплитуда гребенки фильтров, когда  $p(n)$  сдвинута относительно  $\omega(n)$

$2\alpha_3/\alpha_2$ . Во втором —  $4\alpha'_3/\alpha'_2$ . Следовательно, в дополнение к тому, что при  $\alpha'_1 = \alpha'_3$  отсутствует фазовая ошибка, при

$$2\alpha'_3/\alpha'_2 < \alpha_3/\alpha_2 \quad (6.106)$$

и амплитудная ошибка будет меньше.

Механизм сдвига  $p(n)$  относительно  $\omega(n)$  подсказывается (6.90). Чтобы сдвинуть бесконечную последовательность единичных импульсов на  $n_0$ , положим

$$P_k = e^{-12\pi k n_0/N}, \quad 0 \leq k \leq N-1, \quad (6.107)$$

тогда

$$p(n) = \sum_{k=0}^{N-1} e^{i2\pi k(n-n_0)/N} = N \sum_{r=-\infty}^{\infty} \delta(n-n_0-rN). \quad (6.108)$$

Если, положив некоторые из  $P_k$  равными нулю, мы пропускаем соответствующие каналы, то в результате вместо каждого импульса появится сдвинутая последовательность единичных импульсов. Подбирая фазы каналов в соответствии с (6.107), можно переместить  $p(n)$  относительно  $\omega(n)$ . Регулировку фазы можно производить так, как это показано на рис. 6.18. Рассмотрим далее некоторые примеры гребенок БИХ- и КИХ-фильтров.

## 6.2.2. Проектирование гребенок с БИХ-фильтрами

Проектирование гребенок БИХ-фильтров рассмотрено в [1], где показано, что общий принцип подбора фазы в отдельных каналах в соответствии с (6.107) может оказаться полезным для оптимизации.

Следующий пример также взят из [1]. Предполагается, что входная последовательность получается дискретизацией с частотой 10 000 отсч./с. Нужно спроектировать гребенку фильтров с равномерным разнесением в 100 Гц (по аналоговой частоте). Отсюда следует, что  $N=10\,000/100=100$  и что частоты анализа равны  $\omega_k = 2\pi k/100$ ,  $k=0, 1, \dots, M$ . Область анализа предполагается лежащей между 100 и 3000 Гц. Это означает, что потребуется 30 каналов, т. е.  $M=30$ . Плоской общей амплитудно-частотной и линейной фазо-частотной характеристик проще всего добиться, когда отдельные фильтры обладают такими характеристиками. По этой причине используются фильтры Бесселя с максимально-плоской характеристикой. Для этого примера цифровой фильтр нижних частот был получен из фильтра Бесселя шестого порядка методом инвариантной импульсной характеристики. На рис. 6.24 показаны свойства отклика базового для анализа фильтра. Отметим, что 30 мс соответствуют 300 отсчетам при частоте дискретизации 10 кГц. Отметим также, что фаза на рис. 6.24 в вполне линейна. Номинальная частота среза фильтра равна 60 Гц. Этот фильтр использовался в 30-канальной гребенке фильтров с разнесением между каналами в 100 Гц. Канал на нулевой частоте был опущен. Из (6.93) при  $M=30$  и  $N=100$  имеем

$$p(n) = \frac{\sin [0,61 \pi n]}{\sin [0,01 \pi n]} - 1. \quad (6.109)$$

Отметим, что  $p(n)$  периодична с периодом  $N=100$  отсчетов. Результирующая общая характеристика показана на рис. 6.25. На общей импульсной характеристике видно расширение импульсов, что и следовало ожидать, поскольку все каналы участвуют в формировании отклика. Видно также, что из-за того, что продолжительность  $\omega(n)$  превышает  $2N=200$  отсчетов, у  $\tilde{h}(n)$  появляются большие пики. Это приводит к значительным пульсациям в общих амплитудно-частотной и фазо-частотной характеристиках — около 4 дБ пульсаций по амплитуде и  $25^\circ$  для максимума фазовой ошибки.

Для того чтобы улучшить характеристики, не меняя  $\omega(n)$ , была введена задержка, достаточная для выравнивания амплитуды первого и третьего пиков. Это показано на рис. 6.26. Пик функции  $p(n)$  при  $n=0$  сдвинулся на 129 отсчетов вправо. На рис. 6.26б и в видно, что результирующие амплитудно-частотная и фазо-частотная характеристики существенно улучшились. Теперь значение пульсации по амплитуде стало около 0,8 дБ, а максимальная ошибка по фазе — всего лишь  $0,6^\circ$ .

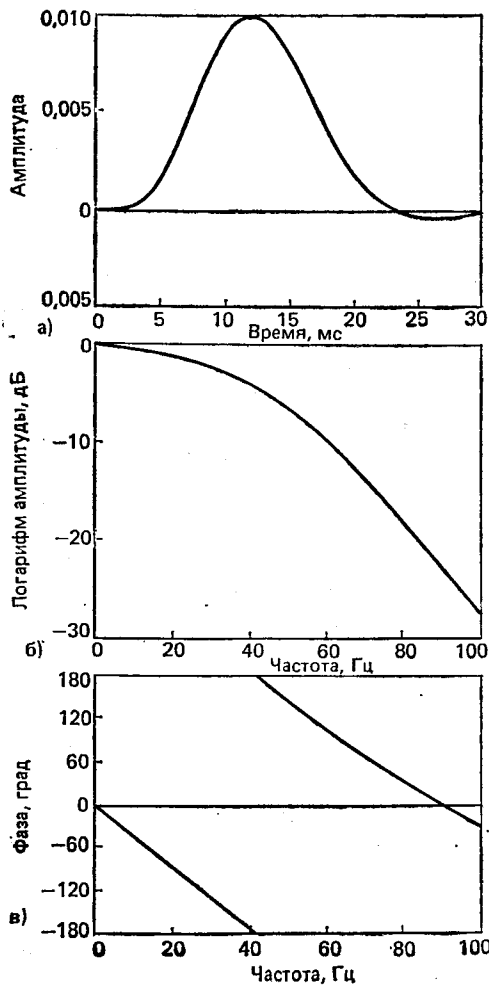


Рис. 6.24. Характеристики фильтра Бесселя 6-го порядка:  
 а) импульсная характеристика;  
 б) амплитудно-частотная характеристика;  
 в) фазо-частотная характеристика [1]

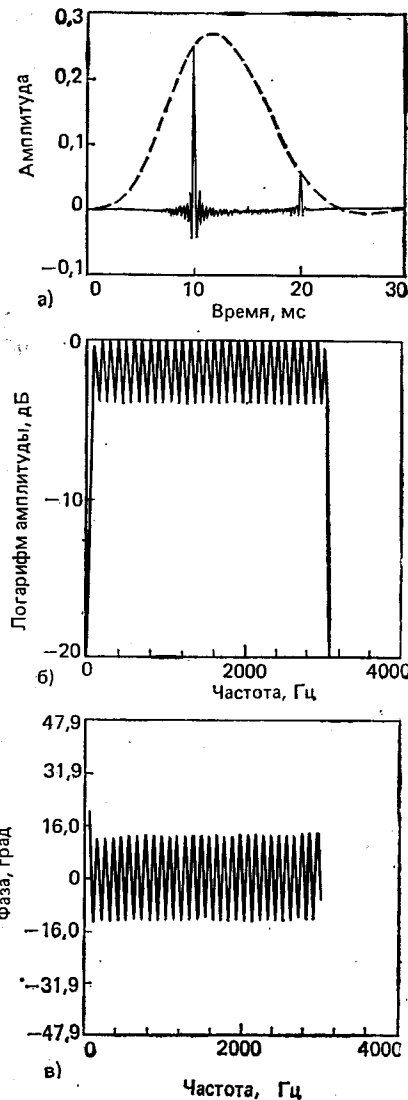


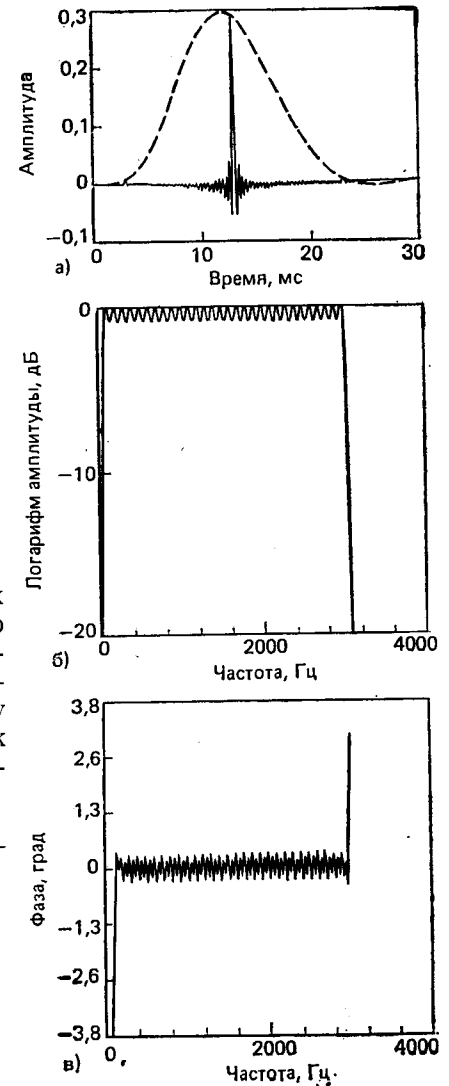
Рис. 6.25. Характеристики 30-канальной гребенки фильтров:  
 а) импульсная характеристика (пунктир — импульсная характеристика фильтра-прототипа нижних частот);  
 б) общая амплитудно-частотная характеристика;  
 в) общая фазо-частотная характеристика за вычетом задержки 10 мс [1]

Приведенный пример иллюстрирует следующую процедуру проб и ошибок, которая позволяет получать гребенки фильтров с хорошими общими характеристиками:

1. Определить разнос и количество фильтров, требуемых для перекрытия полосы анализа.
2. Спроектировать фильтр с нужной избирательностью в каждом канале. Здесь будет получено окно  $\omega(n)$ .
3. Вычислить  $\omega(n)$  и определить  $n_0$ , при котором  $\alpha'_1 = \alpha'_3$  (см. рис. 6.21).
4. Вычислить характеристику. Если она не удовлетворяет вас, изменить разнос, полосу частот и повторить расчет.

Эта процедура предназначена для равноразнесенных частот анализа. В [1] рассматривается также подход к проектированию неравноразнесенных гребенок, которые разбиваются на несколько подгребенок. Каждая из последних набрана из равноразнесенных фильтров. Этот подход довольно эффективен, однако лучшие результаты можно получить с помощью КИХ-фильтров. Поэтому вопросы проектирования гребенок с БИХ-фильтрами более рассматриваться не будут.

Рис. 6.26. Характеристики 30-канальной гребенки фильтров:  
 а) импульсная характеристика для  $n_0 = 129$ ;  
 б) общая амплитудно-частотная характеристика;  
 в) общая фазо-частотная характеристика за вычетом задержки 12,9 мс [1]



### 6.2.3. Проектирование гребенок с КИХ-фильтрами

Цифровые КИХ-фильтры весьма удобны для проектирования гребенок фильтров анализаторов речи. Во-первых, такие фильтры легко спроектировать со строго линейной фазо-частотной характеристикой, просто наложив ограничение

$$w(n) = w(L-1-n), \quad 0 \leq n \leq L-1 \quad (6.110)$$

для каждого фильтра<sup>1</sup>, где  $w(n)$  — импульсная характеристика фильтра, а  $L$  — ее длина в отсчетах. Это означает, что требование линейности фазы общей характеристики гребенки фильтров выполняется тривиально, если отдельные фильтры обладают одинаковыми линейными фазо-частотными характеристиками. Можно поэтому сконцентрировать внимание на достижении нужной частотной избирательности каждого из фильтров и требуемой плоской характеристики всей гребенки фильтров. Второе достоинство КИХ-фильтров состоит в том, что существует множество методов проектирования — от прямого метода взвешивания до метода последовательных приближений, допускающих гибкий подход к реализации сложных требований проектирования.

Представляется, что метод взвешивания имеет ряд преимуществ при проектировании полосовых фильтров или фильтров нижних частот класса КИХ, предназначенных для использования в гребенках. Этот метод иллюстрируется на рис. 6.27. Прежде чем

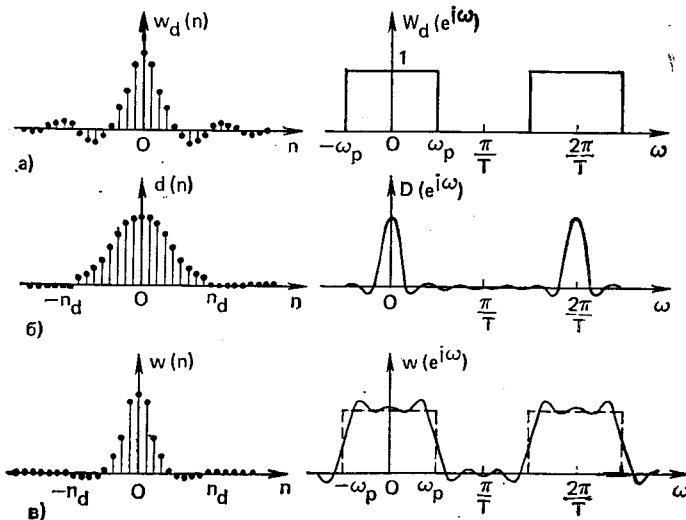


Рис. 6.27. Конструктивное окно для идеального фильтра нижних частот [2]

перейти к подробностям проектирования гребенки, рассмотрим метод взвешивания. Во-первых, идеальный фильтр нижних частот вида

$$W_d(e^{i\omega}) = \begin{cases} e^{-i\omega n_d}, & |\omega| \leq \omega_p; \\ 0, & \text{в противном случае} \end{cases} \quad (6.111)$$

<sup>1</sup> Для простоты предполагается, что импульсная характеристика каждого полосового фильтра имеет длительность  $L$ , хотя такое ограничение снимается тривиально, если добавить в каждый канал соответствующую задержку.

задается частотой среза  $\omega_p$ . Отметим, что для простоты на рисунке опущен член  $e^{-i\omega n_d}$  с линейной фазой, соответствующий задержке на  $n_d$  отсчетов. Требуемое значение  $n_d$  равно  $(L-1)/2$ . Это означает, что при четном  $L$  задержка соответствует нецелому числу отсчетов. Следовательно, идеальная импульсная характеристика задается выражением

$$w_d(n) = \frac{1}{2\pi} \int_{-\omega_p}^{\omega_p} e^{-i\omega n_d} e^{i\omega n} d\omega = \frac{\sin[\omega_p(n-n_d)]}{\pi[n-n_d]} \text{ для всех } n. \quad (6.112)$$

Эта характеристика имеет бесконечную длительность, и ее необходимо усечь, чтобы получить КИХ-фильтр. Это достигается вводом

$$w(n) = d(n-n_d)w_d(n), \quad (6.113)$$

где  $d(n)$  — конструктивное окно фильтра, а  $w(n)$  — импульсная характеристика результирующего фильтра нижних частот<sup>1</sup>. Ширина окна, обозначенная через  $L$ , может быть либо четным ( $L=2q$ ), либо нечетным ( $L=2q+1$ ) целым числом. На рис. 6.27 приведен случай нечетного  $L$ .

В частотной области результатом умножения импульсной характеристики фильтра нижних частот на конструктивное окно будет свертка идеальной частотной характеристики с преобразованием Фурье  $D(e^{i\omega})$  конструктивного окна:

$$W(e^{i\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_d(e^{i\theta}) D(e^{i(\omega-\theta)}) d\theta. \quad (6.114)$$

Результат этой свертки приведен на рис. 6.27в. Основные эффекты сводятся к появлению мягкого перехода от полосы пропускания к полосе задерживания и появлению пульсаций в полосе пропускания и полосе задерживания. Свойства такой аппроксимации отражены на рис. 6.28. Когда  $\omega_p$  больше ширины основного лепестка  $D(e^{i\omega})$ , справедливо следующее:

1. Переходная область  $\Delta\omega$  обратно пропорциональна  $L$ .

2. Функция  $W(e^{i\omega})$  весьма близка к нечетной в окрестности точки ( $\omega_p=0,5$ ).

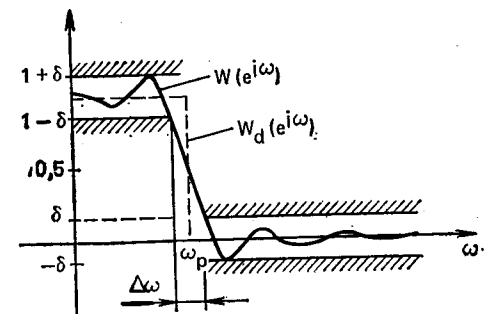


Рис. 6.28. Характеристика ФНЧ, спроектированного методом взвешивания [2]

<sup>1</sup> Здесь важно не запутаться в терминологии. Мы назвали  $d(n)$  конструктивным окном фильтра, а  $w(n)$  — импульсной характеристикой фильтра нижних частот;  $w(n)$  представляет собой также и окно кратковременного анализа Фурье.

3. Максимумы погрешностей аппроксимации в полосе пропускания и в полосе задерживания весьма близки.

4. Погрешность аппроксимации максимальна в окрестности  $\omega_p$  и убывает с изменением  $\omega$  в обе стороны от  $\omega_p$ .

Указанные выше свойства метода взвешивания выполняются для всех обычно применяемых окон. Однако Кайзер предложил семейство весьма гибких окон, близких к оптимальным, если они используются при проектировании фильтров [5]. Окно Кайзера равно:

$$d(n) = \begin{cases} \frac{I_0[\alpha \sqrt{1 - (n/n_d)^2}]}{I_0(\alpha)}, & |n| \leq n_d, \\ 0, & \text{в противном случае.} \end{cases} \quad (6.115)$$

Здесь  $n_d = (L-1)/2$ , а  $I_0[x]$  — модифицированная функция Бесселя нулевого порядка первого рода. Подбирая параметр  $\alpha$ , можно добиться компромисса между шириной переходной области и пиковой погрешностью аппроксимации. Больше того, Кайзер формализовал процедуру выбора окон, дав эмпирическую формулу

$$L = \frac{-20 \log_{10} \delta - 7,95}{14,36 \Delta f} + 1, \quad (6.116a)$$

где  $L$  — порядок фильтра;  $\delta$  — пиковая погрешность аппроксимации;  $\Delta f$  — нормированная ширина переходной области:

$$\Delta f = \Delta\omega/2\pi. \quad (6.116b)$$

При использовании этой формулы задают  $\delta$  и  $\Delta f$ , чем добиваются нужной частотной избирательности. Затем используют (6.116a) для вычисления  $L$ , а параметр  $\alpha$  можно подсчитать из уравнения [5]:

$$\left. \begin{aligned} \alpha &= 0,1102 (-20 \log_{10} \delta - 8,7), \text{ для } -20 \log_{10} \delta > 50; \\ \alpha &= 0,5842 (-20 \log_{10} \delta - \\ &- 21)^{0,4} + 0,07886 \times \\ &\times (-20 \log_{10} \delta - 21), \text{ для } 21 < -20 \log_{10} \delta < 50. \end{aligned} \right\} \quad (6.117)$$

На практике выбор  $\delta$  и  $\Delta f$  зависит от требований к полосовым фильтрам, из которых набрана гребенка.

Гребенка фильтров составлена из набора полосовых фильтров с импульсными характеристиками вида

$$h_k(n) = P_k \omega_k(n) e^{i\omega_k n}, \quad 0 \leq k \leq N-1, \quad (6.118)$$

где  $\omega_k(n)$  — импульсная характеристика фильтра нижних частот. На рис. 6.29 приведены три идеальных полосовых фильтра с абсолютно плоской характеристикой в области  $\omega_{min} \leq \omega \leq \omega_{max}$ . Проектируя фильтры нижних частот, мы выбираем набор частот анализа  $\{\omega_k\}$  и набор частот среза  $\omega_{pk}$  так, чтобы точно перекрыть полосу, как это показано на рис. 6.29. Идеальные полосовые фильтры затем аппроксимируются методом взвешивания.

Выбор пиковой погрешности аппроксимации связан с необходимостью для данного применения затуханием в полосе задерживания. Типичное значение  $-20 \log_{10} \delta$  оказывается, как правило, между 40 и 60 дБ. Значение  $\alpha$  вычисляется по (6.117). И, наконец, нормированное значение ширины переходной области  $\Delta f$  задается для

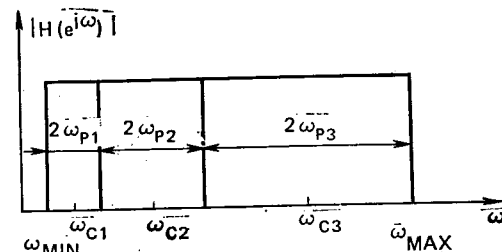


Рис. 6.29. Общая характеристика трех идеальных полосовых фильтров с разными полосами пропускания

того, чтобы по (6.116b) вычислить  $L$ . Выбор  $\Delta f$  или  $\Delta\omega$  опять-таки определяется нужной частотной избирательностью отдельных фильтров. Ясно, что ширина переходной области  $\Delta\omega_k$  не должна превышать  $2\omega_{pk}$ .

При разработке гребенки фильтров предполагается, что  $\Delta\omega$  одинакова для всех фильтров, следовательно, можно воспользоваться приведенным выше свойством 2. Если переходная область у всех фильтров одна и та же и если, кроме того, переходы симметричны в окрестности точки переходов, то можно ожидать, что сумма частотных характеристик окажется близкой к единице. Отсюда вытекает, что  $d(n)$  должна быть одинаковой для всех фильтров гребенки.

Чтобы уяснить эффект одинакового выбора  $d(n)$  для всех фильтров, рассмотрим общую частотную характеристику, которая, как следует из (6.118), равна

$$\tilde{H}(e^{i\omega}) = \sum_{k=0}^{N-1} P_k W_k(e^{i(\omega-\omega_k)}). \quad (6.119)$$

Теперь при одинаковых на всех частотах анализа конструктивных окнах можно записать

$$W_k(e^{i(\omega-\omega_k)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_{dk}(e^{i(\theta-\omega_k)}) D(e^{i(\omega-\theta)}) d\theta. \quad (6.120)$$

Подставив (6.120) в (6.119) и изменив порядок суммирования и интегрирования, получим

$$\tilde{H}(e^{i\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[ \sum_{k=0}^{N-1} P_k W_{dk}(e^{i(\theta-\omega_k)}) \right] D(e^{i(\omega-\theta)}) d\theta. \quad (6.121)$$

Определив

$$\tilde{H}_d(e^{i\omega}) = \sum_{k=0}^{N-1} P_k W_{dk}(e^{i(\omega-\omega_k)}), \quad (6.122)$$



можно переписать (6.121) в виде

$$\tilde{H}(e^{i\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{H}_d(e^{i\theta}) D(e^{i(\omega-\theta)}) d\theta. \quad (6.123)$$

Видно, что функция  $\tilde{H}_d(e^{i\omega})$  представляет собой требуемую общую характеристику. Если, например,  $P_k=1$  для  $0 \leq k \leq N-1$ , ширина полос и центральные частоты выбраны так, чтобы перекрыть всю полосу частот  $-\pi \leq \omega \leq \pi$ , то

$$\tilde{H}_d(e^{i\omega}) = e^{i\omega n_d}, \quad -\pi \leq \omega \leq \pi. \quad (6.124)$$

Подставив этот результат в (6.123), получим

$$\tilde{H}(e^{i\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i\theta n_d} D(e^{i(\omega-\theta)}) d\theta = d(n_d) e^{-i\omega n_d}. \quad (6.125)$$

Отсюда следует, что общая импульсная характеристика имеет вид

$$\tilde{h}(n) = d(n_d) \delta(n - n_d). \quad (6.126)$$

Следовательно, если имеется достаточное количество каналов для линейности фазо-частотной характеристики и если все фильтры проектируются с одинаковым окном, общая характеристика оказывается также идеальной. Иначе говоря, независимо от распределения центральных частот и полос общая характеристика оказывается идеальной при любом конструктивном окне. Итак, при использовании КИХ-фильтров теоретически можно получить точное воспроизведение входного сигнала при произвольном распределении центральных частот и полос.

Эффект пропуска части частотного диапазона  $-\pi \leq \omega \leq \pi$  можно выяснить, заметив, что (6.123) справедливо независимо от выбора  $P_k$  в (6.122). Следовательно, если из анализа исключаются как нижние, так и высшие частотные диапазоны (см. рис. 6.29), то  $P_0$  и  $P_k=0$  для  $k > M$ , где  $M$  — количество каналов. Общая характеристика будет иметь вид

$$\tilde{H}_d(e^{i\omega}) = e^{-i\omega n_d}, \quad \omega_{min} \leq |\omega| \leq \omega_{max}. \quad (6.127)$$

Таким образом, для заданного конструктивного окна  $d(n)$ , заданных  $\omega_{min}$  и  $\omega_{max}$  общая характеристика будет характеристикой полосового фильтра с переходной областью и пульсациями (в полосе пропускания и полосе задерживания), идентичными с каждым из фильтров каналов. Это происходит потому, что одно и то же конструктивное окно перемножается с каждой из отдельных импульсных характеристик. Следовательно, опять общая характеристика не зависит от числа и распределения отдельных полосовых фильтров.

Проектирование гребенок фильтров в соответствии с указанными принципами иллюстрируют следующие примеры.

**Пример 1.** Допустим, что входная частота дискретизации равна 9,6 кГц и требуется спроектировать гребенку из 15 равнозначенных фильтров, перекрывающую область от 200 до 3200 Гц. Частотой среза для всех фильтров нижних частот будет

$$F_p = \omega_p / 2\pi T = (3200 - 200) / 2 (15) = 100. \quad (6.128)$$

а центральные частоты равны

$$F = \omega_k / 2\pi T = 200k + 100, \quad k = 1, 2, \dots, 15. \quad (6.129)$$

При таком выборе центральных частот и полос 15 идеальных фильтров перекрывают интервал от 200 до 3200 Гц. Предположив, что требуется затухание на 60 дБ вне переходных областей каждого канала, найдем из (6.117), что  $\alpha = -5,65326$ . Поскольку для фильтров-прототипов частота среза равна 100 Гц, разумно выбрать в качестве самой широкой из областей перехода значение 200 Гц. Воспользовавшись этим значением и тем, что  $-20 \log_{10} \delta = 60$ , получим из (6.116а)  $L=175$  нижнюю оценку для  $L$ . Отметим, что, если бы были допустимы меньшие затухания, можно было бы получить меньшее  $L$  при том же  $\Delta f$ . Характеристики гребенки, спроектированной с указанными параметрами, приведены на рис. 6.30. В верхней части рисунка показаны характеристики отдельных

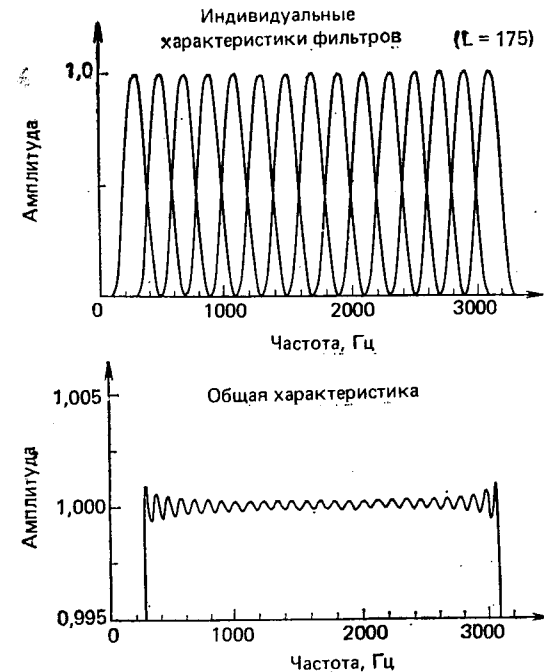


Рис. 6.30. Индивидуальная и общая характеристики гребенки из 15 полосовых фильтров с одинаковыми полосами пропускания при  $L=175$  [2]

полосовых фильтров. Спад в переходной области каждого фильтра компенсируется подъемом следующего. Отметим, что смежные каналы пересекаются на уровне 0,5. В нижней части рисунка — общая амплитудно-частотная характеристика гребенки. Фазо-частотная характеристика линейна с задержкой  $n_d = -(175-1)/2 = 87$  отсчетов. Ясно, что фильтры хорошо согласуются на краях частотного диапазона. Действительно, отклонение от единицы не превышает или равно пиковой погрешности аппроксимации  $\delta = 0,001$ , выбранной при проектировании фильтров нижних частот.