

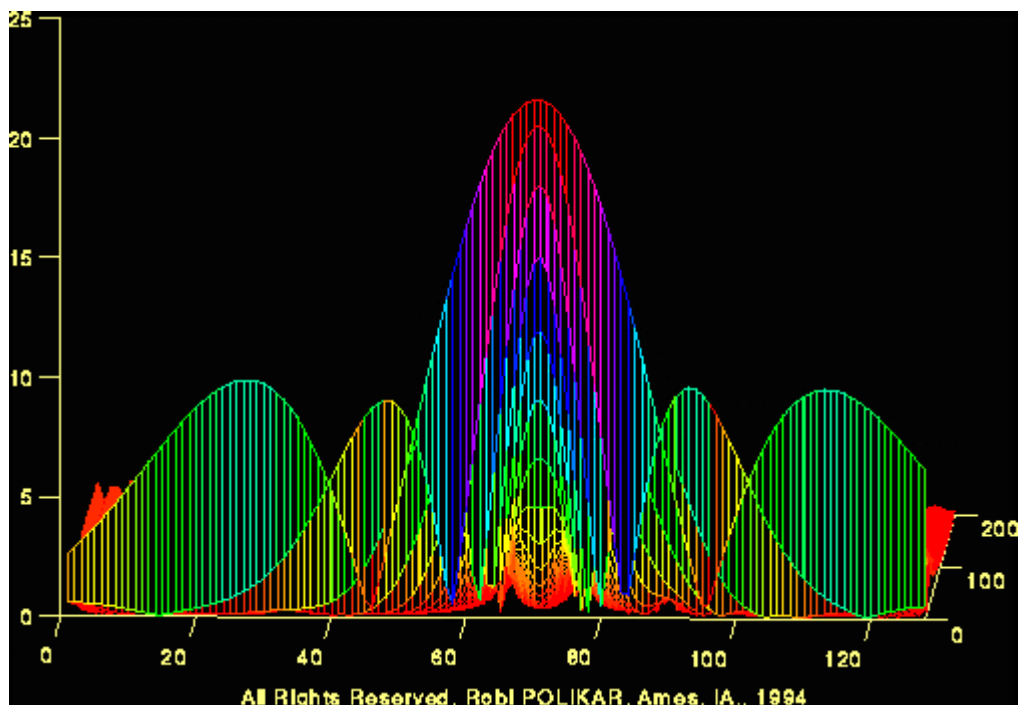
ВВЕДЕНИЕ В ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ

Robi Polikar, Iowa State University

Автор перевода: Грибунин В.Г.

Электронная версия книги подготовлена фирмой АВТЭКС Санкт-Петербург,
<http://www.autex.spb.ru>, E-mail: info@autex.spb.ru

ВВЕДЕНИЕ



Представляем Вам этот учебник по вейвлет-преобразованию. Большинство книг и статей по вейвлетам написаны математиками и для математиков, тогда как количество литературы для новичков в этой области весьма ограничено. Этим и вызвано написание учебника.

Этот учебник адресован в первую очередь вам, инженерам. Поэтому, многие теоретические выкладки здесь опущены. Да и вряд ли они оказались бы вам полезными. Тем не менее, для желающих более глубоко познать теорию вейвлет-преобразования даны все необходимые ссылки.

Учебник не предполагает наличие каких-либо предварительных знаний по вейвлетам. Если это не так, то следующий материал может показаться Вам тривиальным.

ЧАСТЬ I

ПРЕОБРА... ЧТО?

Итак, во-первых, для чего нам нужно преобразование? И что такое преобразование?

Математические преобразования применяются к сигналу для того, чтобы получить о нем какую-то дополнительную информацию, недоступную в исходном виде. В дальнейшем изложении сигнал во временной области будет называться «исходным», а преобразованный сигнал - трансформантой.

Среди многих известных преобразований сигналов наиболее популярным является преобразование Фурье.

Большинство сигналов, встречающихся на практике, представлены во **временной области**, то есть сигнал есть функция времени. Таким образом, при отображении сигнала на графике одной из координат (независимой) является ось времени, а другой координатой (зависимой) – ось амплитуд. Таким образом мы получаем **амплитудно-временное представление** сигнала. Для большинства приложений обработки сигнала это представление не является наилучшим. Во многих случаях наиболее значимая информация скрыта в частотной области сигнала. **Частотный спектр** есть совокупность частотных (спектральных) компонент, он отображает наличие тех или иных частот в сигнале.

Как известно, частота измеряется в Герцах [Гц], или в числе периодов в секунду. На рисунке ниже для примера представлены три синусоиды: 3Гц, 10Гц и 50Гц. Сравните их.

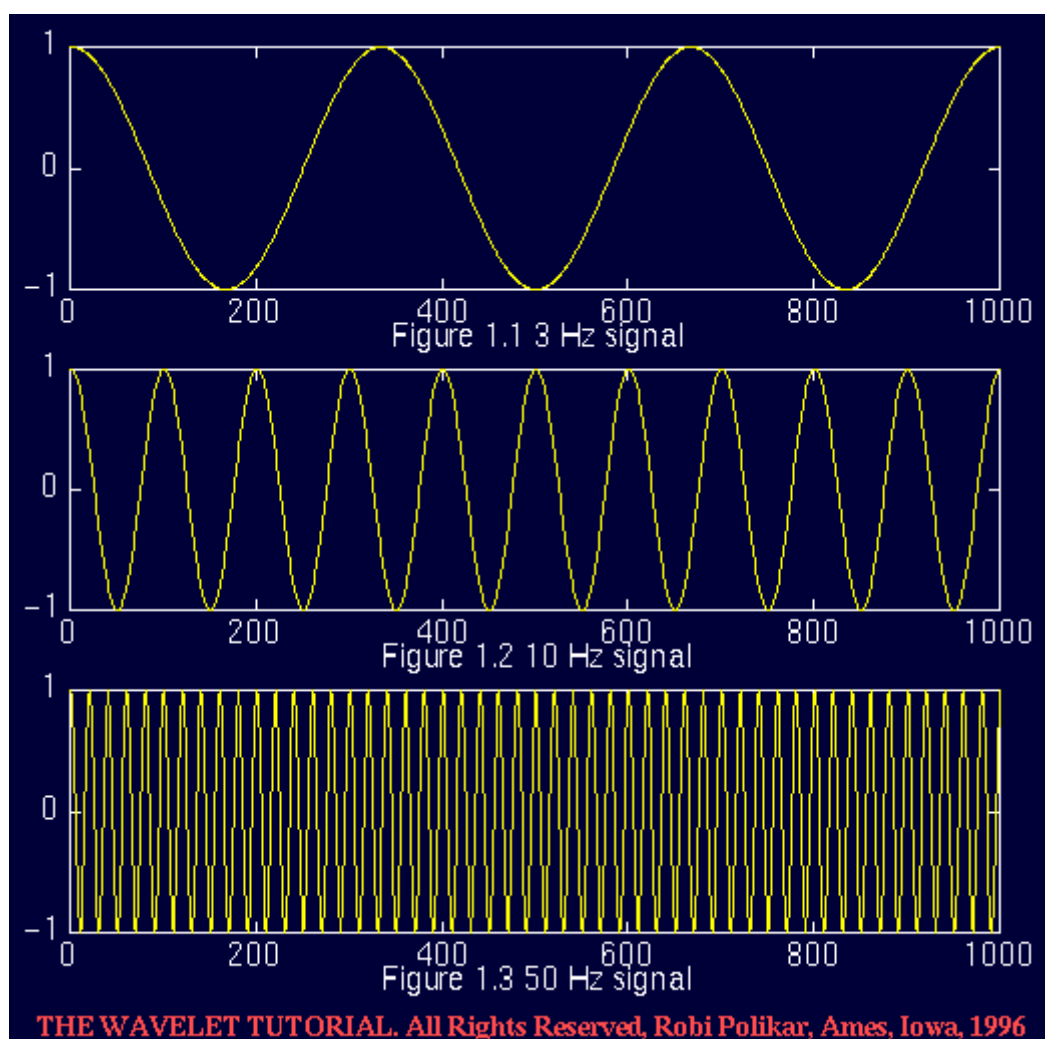


Рис.1-1, 1-2, 1-3.

А каким же образом мы измеряем, находим частоту сигнала? Способ называется ПРЕОБРАЗОВАНИЕ ФУРЬЕ (ПФ). В результате ПФ сигнала, заданного во временной области, мы получаем его спектральное представление. Вместо значений времени на оси абсцисс графика сигнала будет теперь отложены значения частоты, а ось ординат будет отображать амплитуду той ли иной частоты в сигнале.

Частотная ось начинается с нуля и простирается до бесконечности. Например, если взять ПФ сигнала электрического тока в наших розетках, то мы получим единственный импульс на частоте 50Гц, а остальные значения амплитуд будут равны нулю. Реальные сигналы обычно состоят из множества частот и редко имеют подобные простые графики. Ниже показано ПФ синусоиды с частотой 50Гц:

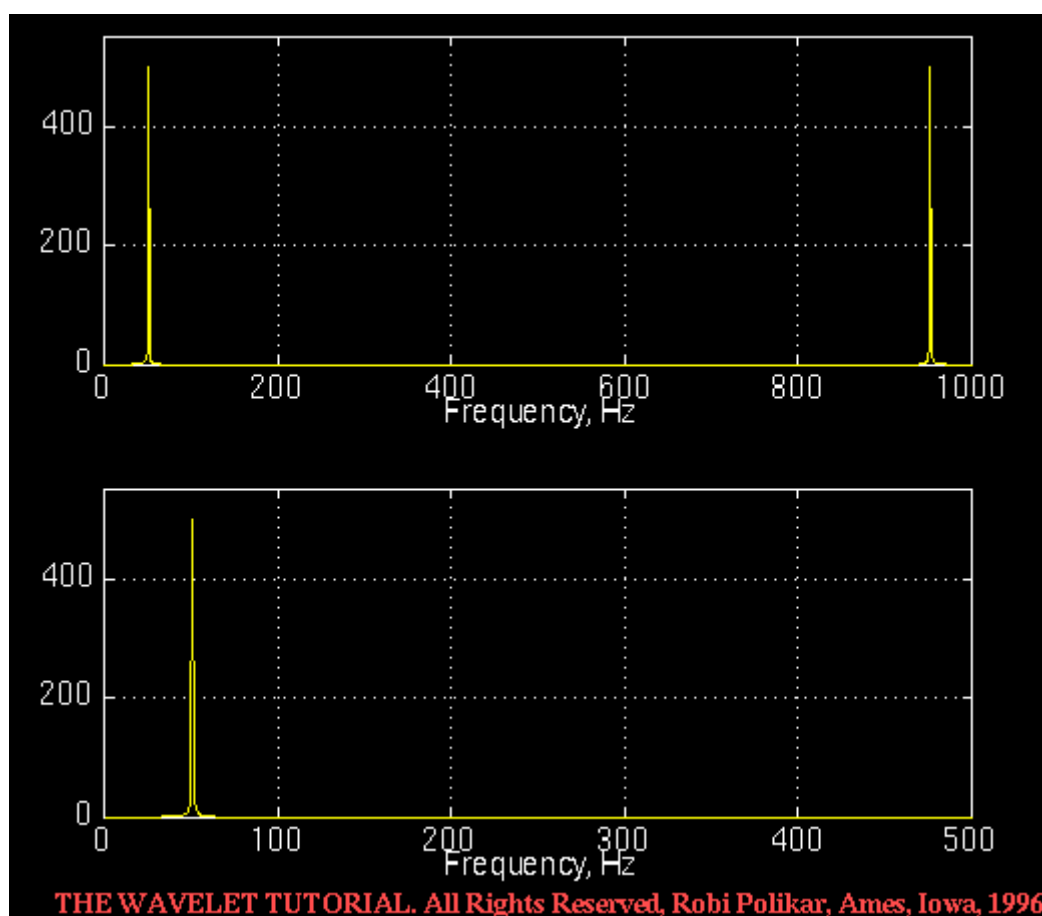


Рис.1-4. ПФ синусоиды 50Гц, показанной на Рис.1-3.

Посмотрите внимательнее на Рис.1-4. Как видно, нижний график представляет собой половину верхнего. Причина заключается в том, что ПФ вещественного сигнала всегда симметрично. Поэтому для амплитудного спектра достаточно отображать половину коэффициентов: остальные все равно не несут информацию. Так мы и будем показывать в дальнейшем. (Заметим, что оставшаяся часть информации сосредоточена в фазе преобразования, число коэффициентов которой также половина от исходного числа).

ДЛЯ ЧЕГО НАМ НУЖНА ЧАСТОТНАЯ ИНФОРМАЦИЯ?

Зачастую информация, не заметная во временном представлении сигнала, проявляется в его частотном представлении.

Рассмотрим в качестве примера биологический сигнал, например, электрокардиограмму (ЭКГ). Типичный вид ЭКГ хорошо известен кардиологам. Любое значительное отклонение от него рассматривается как патология.

Эта патология, однако, не всегда может быть заметна во временном представлении сигнала. Поэтому в последних моделях электрокардиографов для анализа используется и частотная область сигнала. Решение о патологии выносится только с использованием информации частотной области.

Кроме ПФ существует и много других часто применяемых преобразований сигнала. Примерами являются преобразование Гильберта, оконное ПФ, распределение Вигнера, преобразование Уолша, вейвлет-преобразование и многие другие. Для каждого преобразования можно указать наиболее подходящую область применения, достоинства и недостатки, и вейвлет-преобразование (ВП) не является в этом смысле исключением.

Для лучшего понимания потребности в ВП рассмотрим подробнее ПФ. ПФ (также, как и ВП) является обратимым образованием, то есть из его коэффициентов посредством обратного преобразования может быть получен исходный сигнал. Однако только одно из представлений доступно для нас в каждый момент времени: частотную информацию нельзя извлечь из временной, а временную - из частотной. Возникает естественный вопрос: возможно ли получить совместное частотно-временное представление сигнала?

Как будет показано, ответ зависит от конкретного приложения и от природы сигнала. Напомним, что ПФ дает частотную информацию, содержащуюся в сигнале, то есть говорит нам о том, каково содержание каждой частоты в сигнале. Однако в какой момент времени возникла та или иная частота, когда она закончилась - на эти вопросы ответ получить не удастся. Впрочем, эта информация и не требуется, если сигнал – стационарный.

Давайте поподробнее обсудим концепцию стационарности, так как она несомненно одна из наиболее важных при анализе сигнала. Стационарными называются сигналы, частотное наполнение которых не меняется во времени. Поэтому при частотном анализе таких сигналов и не требуется временная информация - все частоты присутствуют в сигнале на протяжении всего времени.

Например, следующий сигнал

$$x(t) = \cos(2\pi 10t) + \cos(2\pi 25t) + \cos(2\pi 50t) + \cos(2\pi 100t)$$

является стационарным, так как содержащиеся в нем частоты 10, 25, 50 и 100 Гц не меняются во времени. Этот сигнал изображен ниже:

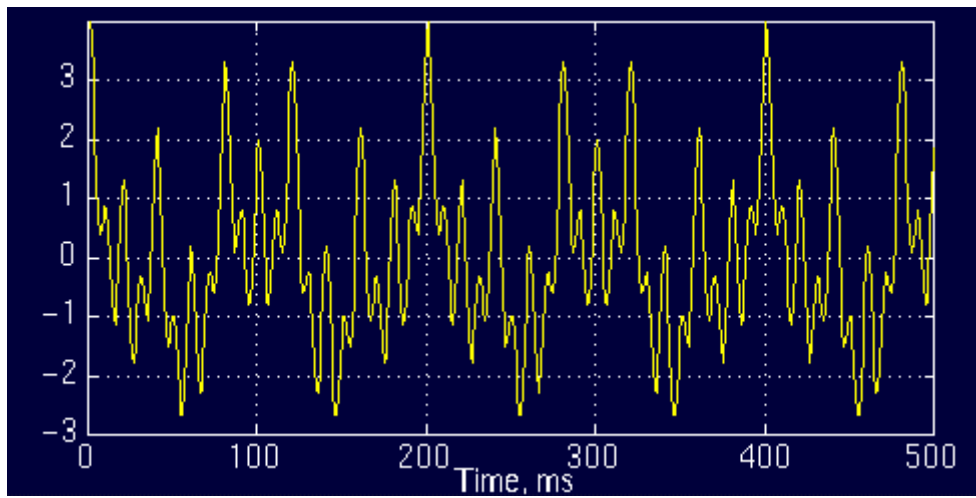


Рис.1-5.

А здесь показано его ПФ:

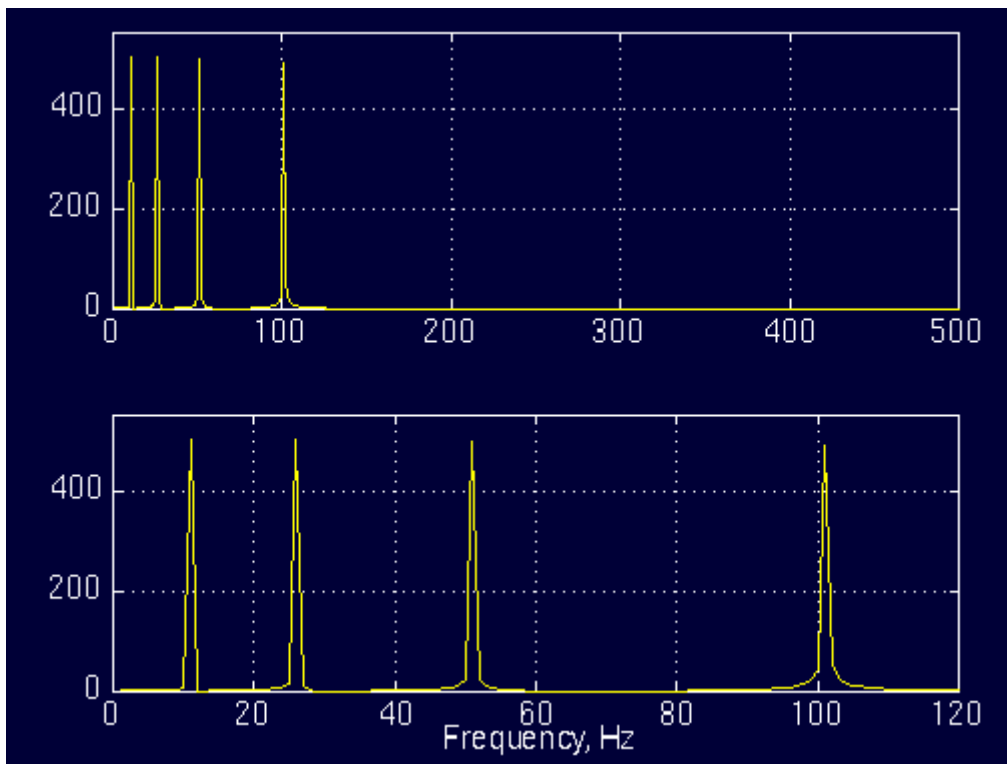


Рис.1-6.

На верхнем графике Рис.1-6 изображен частотный спектр сигнала, показанного на Рис.1-5. На нижнем графике изображена его увеличенная копия - только интересующий нас диапазон частот. Заметьте, что четыре частотные компоненты соответствуют частотам 10, 25, 50 и 100 Гц.

Сигнал, изображенный на Рис.1-7 нестационарный. Его частота непрерывно изменяется во времени. Такой сигнал называется сигналом с линейной частотной модуляцией (ЛЧМ). В зарубежной литературе он называется «chirp»-сигналом.

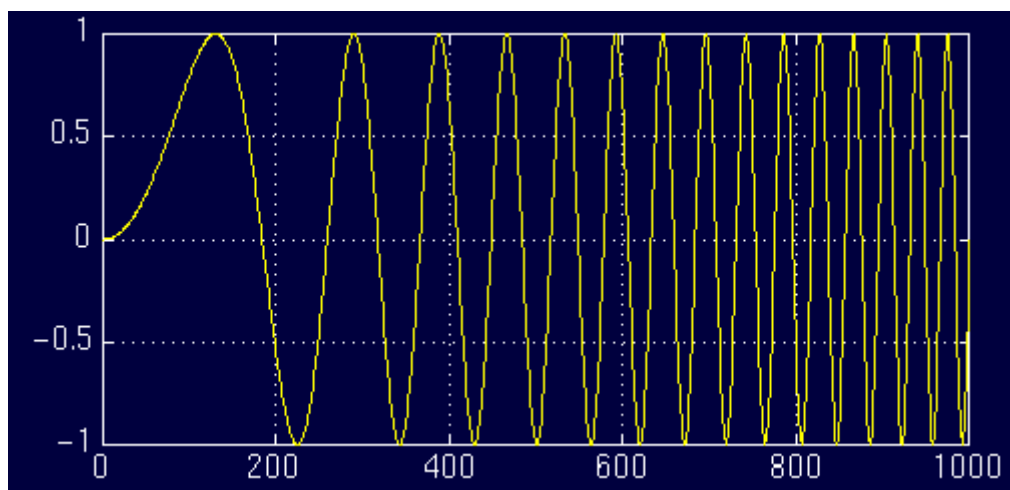


Рис.1-7.

Рассмотрим еще один пример. На Рис.1-8 показан сигнал, состоящий из четырех различных частот, встречающихся на четырех различных интервалах и, следовательно, являющийся нестационарным. В интервале времени от 0 до 300мс частота сигнала 100Гц, от 300 до 600мс – 50Гц, от 600 до 800мс – 25Гц и на последнем интервале – 10Гц.

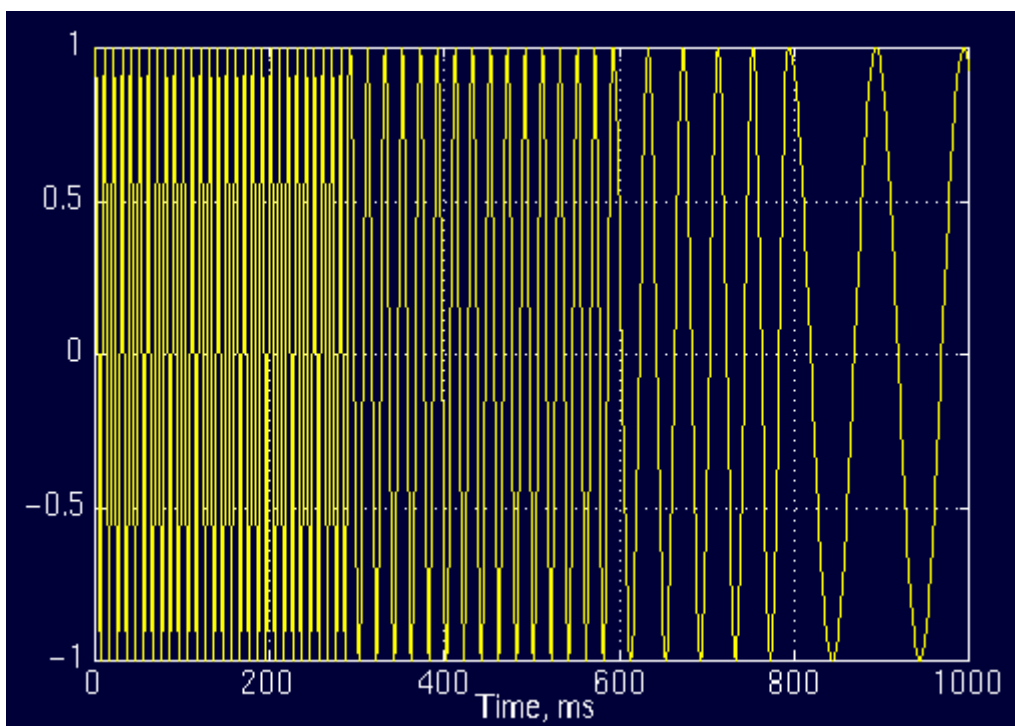


Рис.1-8.

А вот спектр (ПФ) этого сигнала:

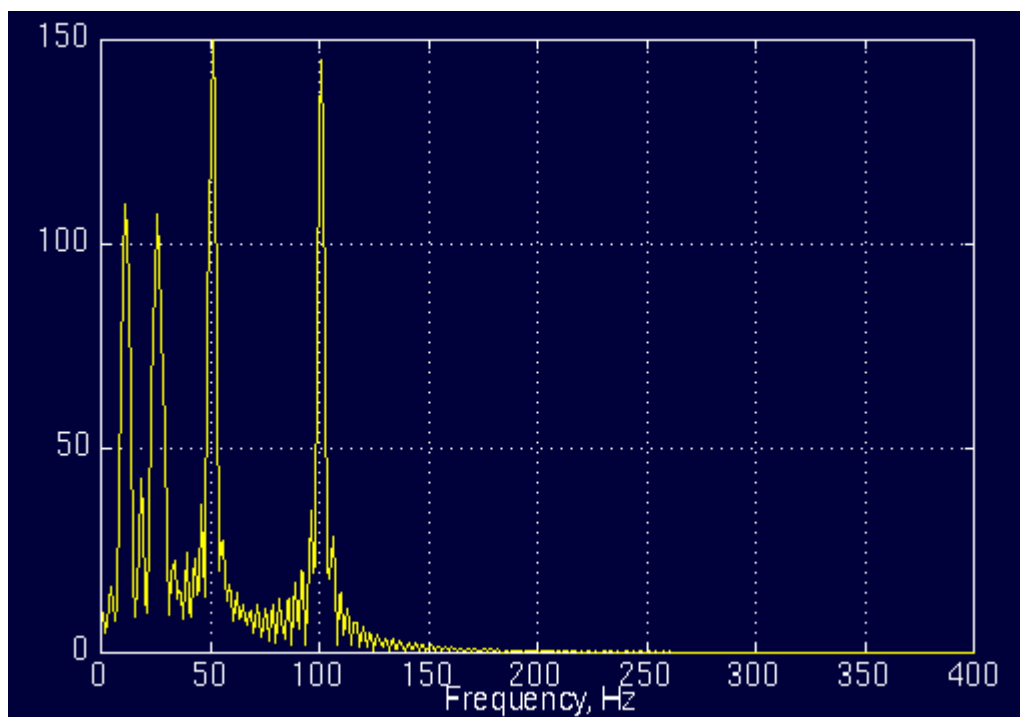


Рис.1-9.

Как видно из рисунка, все четыре частотные компоненты четко видны. «Ложные» частоты имеют небольшую величину в силу того, что частота сигнала изменяется редко. Заметьте, что амплитуды высокочастотных компонент больше, чем низкочастотных. Это связано с тем, что их длительность больше (300 против 200).

Не считая небольших «ложных» частот, все кажется верным. ПФ имеет четыре пика, соответствующие четырем частотам, присутствующим в сигнале... правильно...

НЕТ!

Впрочем, не совсем уж неправильно, но и не совсем правильно... И вот почему:

Для первого сигнала, показанного на Рис.1-5, рассмотрим следующий вопрос:

В какой момент времени (или хотя бы интервал) возникла та или иная частота?

Ответ:

Они существуют в течение всего времени! Напомним, что в стационарных сигналах все частотные компоненты присутствуют в течение всего времени. То есть 10, 50, 100Гц имеются на всем временном интервале.

Теперь рассмотрим тот же вопрос для нестационарного сигнала, показанного на Рис.1-7 или Рис.1-8.

В какое время существуют различные частоты?

Ясно, что не во все время. Однако, сравнив спектры Рис.1-6 и Рис.1-9 мы не обнаружим особой разницы. На обоих графиках видны четыре частотные составляющие 10, 25, 50 и 100Гц. Кроме небольших "ложных" частот и неодинаковости амплитуд пиков, других различий между спектрами нет, хотя они соответствуют различным сигналам во временной области. Каким же образом спектры двух столь разных сигналов оказались похожи? Таково свойство ПФ, которое позволяет увидеть частотное наполнение сигналов, но не позволяет определить, в какой момент времени существует та или иная частота. Поэтому, ПФ непригодно для анализа нестационарных сигналов, за одним исключением:

ПФ может использоваться для анализа нестационарных сигналов, если нас интересует лишь частотная информация, а время существования спектральных составляющих неважно. В противном случае надо искать более подходящий метод анализа.

В случае если требуется временная локализация спектральных компонент, необходимо обратиться к ЧАСТОТНО-ВРЕМЕННОМУ ПРЕДСТАВЛЕНИЮ сигнала.

ЭФФЕКТИВНОЕ РЕШЕНИЕ: ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ

Вейвлет-преобразование (ВП) и относится как раз к этому типу преобразований. Оно обеспечивает частотно-временное представление сигналов. (Существуют и другие преобразования сигналов, также выполняющие эту задачу, такие как оконное ПФ, преобразование Вигнера и др.)

Принципы ВП – это отдельная история. ВП будет объясняться после оконного ПФ (ОПФ). ВП было разработано в некотором роде как альтернатива ОПФ. ОПФ будет подробно освещено во второй части этого учебника. Пока же ограничимся лишь замечанием, что ВП было разработано для преодоления некоторых проблем ОПФ, связанных с плохим разрешением, как это поясняется в части II.

Пока же говоря о ВП будем иметь в виду следующее. Пропустим сигнал через два фильтра – низкочастотный и высокочастотный (фильтры соединены параллельно). Фильтры ВП должны удовлетворять жестким ограничениям – условию допустимости. Повторим эту процедуру для выхода низкочастотного фильтра, оставив выход высокочастотного фильтра неизменным.

Так, если исходный сигнал содержал частоты до 1000Гц, то после первого этапа получим два сигнала 0-500Гц и 500-1000Гц, после второго этапа - три сигнала 0-250Гц, 250-500Гц и 500-1000Гц. и т.д. Эта операция называется декомпозицией.

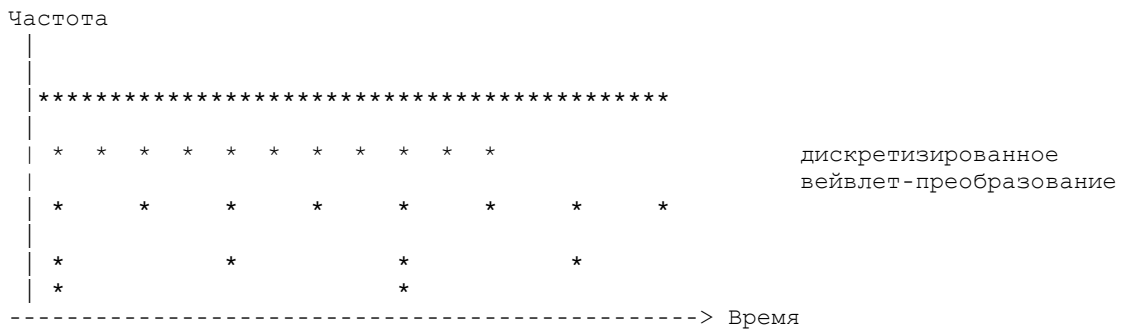
Декомпозиция продолжается какое-то число раз. В конечном счете, мы получаем множество субсигналов, которое представляет наш исходный сигнал. Каждый субсигнал соответствует како-то субполосе частот. Можно построить трехмерный график, отложив по одной оси время, по второй - частоту и по третьей - амплитуду. Таким образом мы можем увидеть, какие частоты присутствуют в каждый интервал времени. (Существует закон, называемый «принципом неопределенности Гейзенберга», который гласит, что нельзя одновременно измерить частоту и время с произвольно высокой точностью). Поэтому мы можем лишь говорить об интервале времени и о наблюдающейся в нем частотной полосе.

Это и есть, собственно, проблема разрешения анализа, решение которой невозможно в рамках оконного ПФ. ОПФ дает фиксированное разрешение на всех частотах, тогда как при ВП разрешение изменяется:

На высоких частотах лучше разрешение по времени, а на низких - по частоте. Это означает, что для высокочастотной компоненты мы можем точнее указать ее временную позицию, а для низкочастотной - ее значение частоты. Посмотрите на следующий график:



Этот график можно интерпретировать следующим образом: Верхняя строка показывает, что на высоких частотах мы имеем в распоряжении больше отсчетов, что соответствует меньшим интервалам времени. Нижняя строка соответствует низкочастотной компоненте, и мы видим, что в ней содержится меньше точек. Следовательно, временное разрешение для низкочастотных компонент сигнала хуже.



В случае дискретного времени разрешение по времени ведет себя также как и в предыдущем случае. Однако теперь и разрешение по частоте меняется от уровня к уровню. На низких частотах лучше разрешение по частоте, чем на высоких частотах. Отметим также увеличение расстояния между частотными точками с увеличением частоты.

Ниже приведены несколько примеров непрерывного ВП. Пусть сигнал состоит из двух синусоид, существующих в разное время:

Вначале в сигнале присутствует НЧ синусоида, затем - ВЧ.

Непрерывное ВП сигнала на Рис.1-10 показано на Рис.1-11:

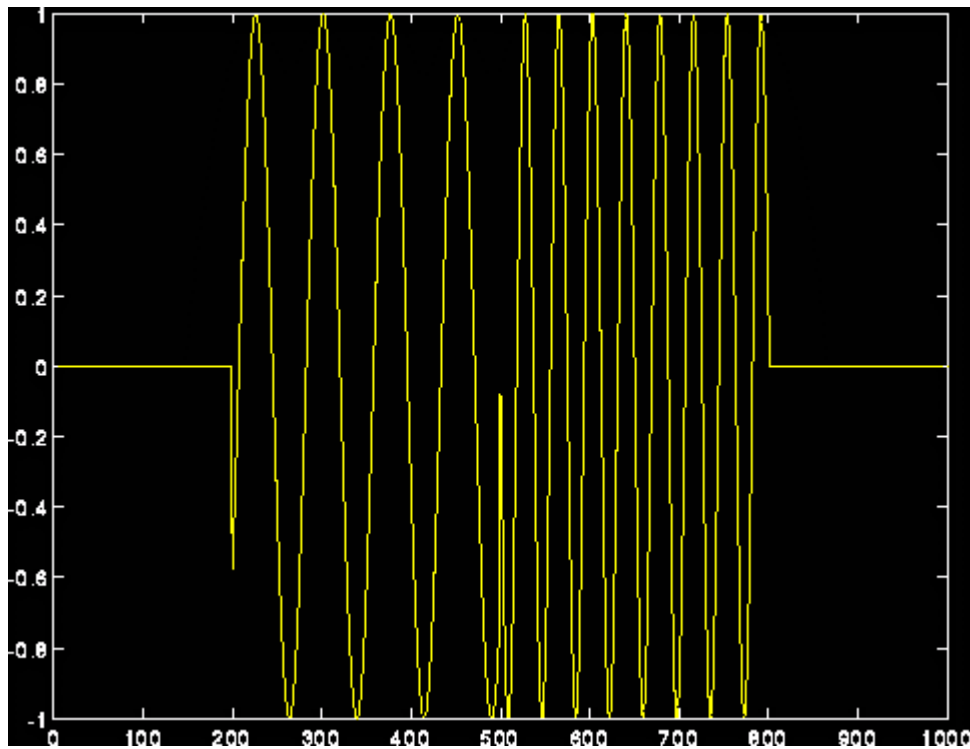


Рис.1-10.

SAME TRANSFORM, ROTATED -250 DEGREES, LOOKING FROM 45 DEG ABOVE

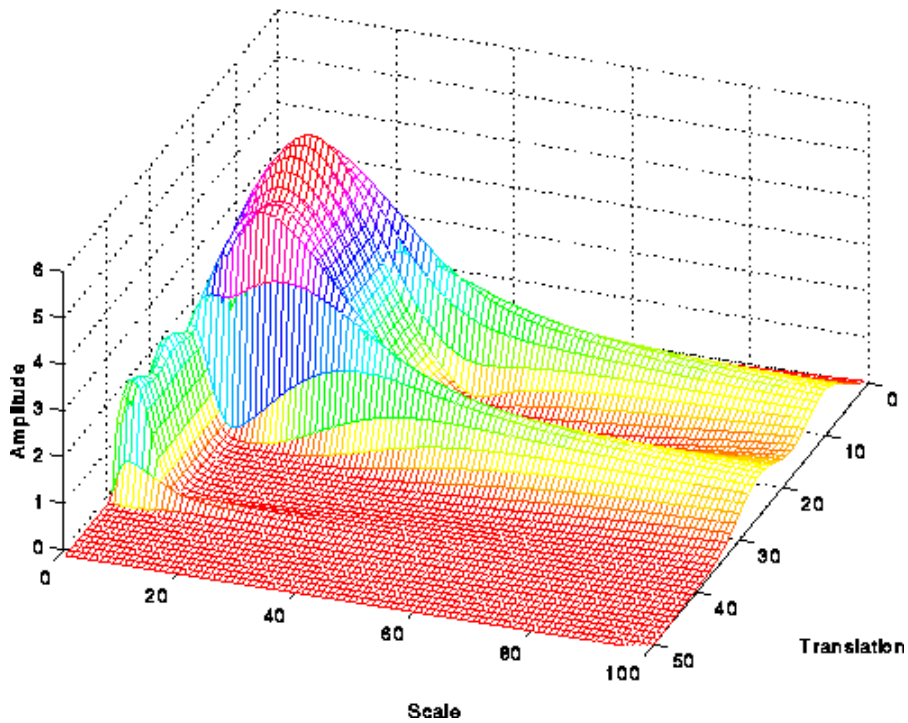


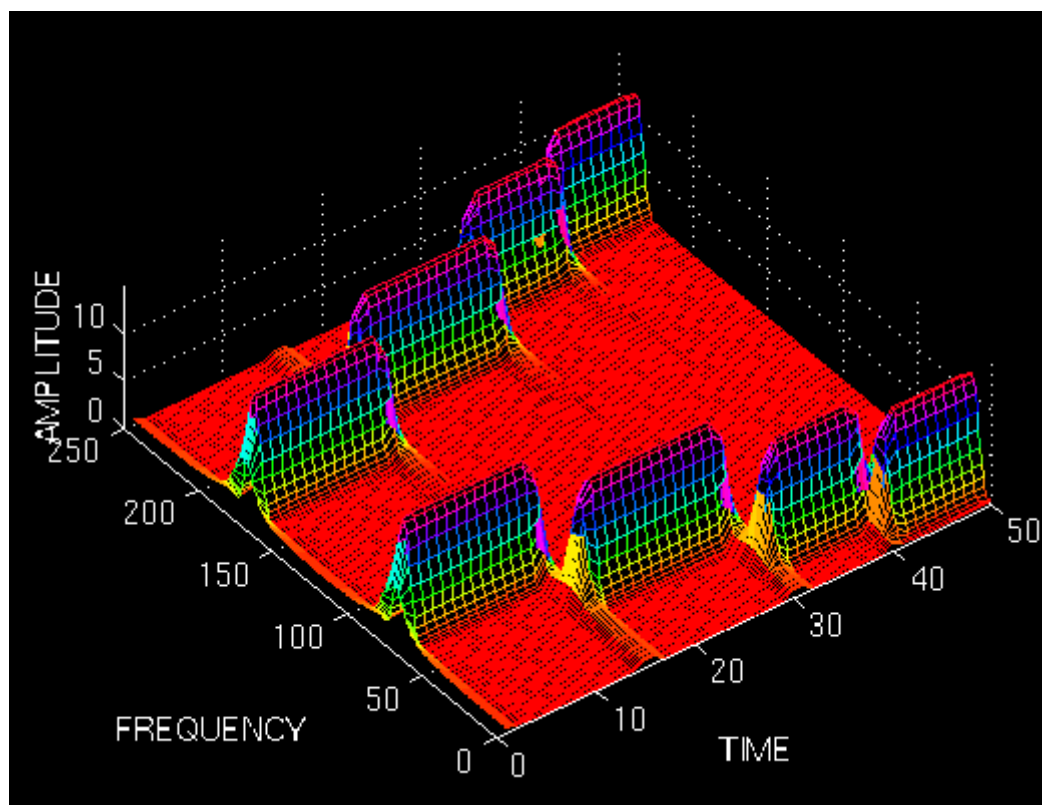
Рис.1-11.

Заметьте, что вместо оси частот на этих графиках обозначена ось масштаба. Концепция масштаба станет более понятной из объяснений последующих разделов. Пока же скажем лишь, что масштаб является понятием обратным частоте. То есть высокие шкалы соответствуют низким частотам, а низкие - высоким. Следовательно, малый пик на графике соответствует ВЧ компоненте сигнала, а большой пик - НЧ компоненте.

Возможно, вы сбиты с толку частотным разрешением сигнала, представленного на графике, так как кажется, что на высоких частотах имеется хорошее разрешение. Однако, на самом деле, это не что иное, как хорошее разрешение по масштабу. Следовательно, на ВЧ по частоте разрешение как раз таки плохое. Подробнее эти вопросы будут рассмотрены в части II и части III.

Итак, в первой части учебника мы дали краткое введение в обработку сигналов, преобразование Фурье и вейвлет-преобразование.

ЧАСТЬ II



ФУНДАМЕНТАЛЬНЫЕ ОСНОВЫ

Давайте кратко вспомним, что мы прошли в первой части. Итак, для анализа нестационарных сигналов предпочтительнее применять вейвлет-преобразование (ВП). Я написал, что преобразование Фурье (ПФ) не подходит для анализа нестационарных сигналов, и привел несколько тому примеров. Для быстрого повторения рассмотрим следующий пример.

Предположим, у нас имеется два различных сигнала. Также предположим, что их спектральные характеристики идентичны. Может ли быть такое? Как показано в примерах части I - может. В частности в случае, когда в одном из сигналов частоты присутствуют на протяжении всего интервала наблюдения, а в другом - эти же частоты встречаются поочередно во времени. Хотя сигналы полностью различны, их ПФ (амплитуда) полностью одинаково! Отсюда следует неэффективность применения ПФ для анализа нестационарных сигналов.

Но в чем причина этого? Другими словами, почему два разных сигнала имеют одинаковый вид ПФ? КАК, ВООБЩЕ ГОВОРЯ, РАБОТАЕТ ПРЕОБРАЗОВАНИЕ ФУРЬЕ? Краеугольный камень цифровой обработки сигналов:

ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Я не буду рассматривать детали ПФ по двум причинам:

1. Это слишком обширная область, чтобы рассматривать здесь.
2. Да и в любом случае не ПФ находится в фокусе внимания учебника.

Однако мне хотелось бы упомянуть несколько важных свойств ПФ по двум причинам:

1. Знание ПФ необходимо для понимания того, как работает ВП.
2. ПФ было на протяжении многих лет важнейшим средством ЦОС (и я думаю, что так и останется в дальнейшем).

В XIX веке (в 1822 г., если быть точным) французский математик Дж. Фурье показал, что любая периодическая функция может быть выражена в виде бесконечной суммы периодических комплексных экспоненциальных функций. Через много лет эти идеи были применены для непериодических функций, а затем и для сигналов дискретного времени (периодических и непериодических). После такого обобщения потенциальная область применения ПФ стала значительно большей. В 1965 был разработан алгоритм быстрого вычисления ПФ (БПФ), что сделало ПФ еще более популярным.

Теперь рассмотрим, как ПФ работает:

ПФ декомпозирует сигнал на комплексные экспоненциальные функции различных частот. Процесс декомпозиции задается двумя равенствами:

$$X(f) = \int_{-\infty}^{\infty} x(t) \cdot e^{-2j\pi ft} dt \quad (1)$$

$$x(t) = \int_{-\infty}^{\infty} X(f) \cdot e^{-2j\pi ft} df \quad (2)$$

Рис.2-1.

В вышеприведенном равенстве, **t** обозначает время, **f** обозначает частоту, а **x** обозначает сам сигнал. Обозначение **x** используется для сигнала во временной области, а **X** - для сигнала в частотной области. Эти соглашения позволят различить две области представления сигнала. Равенство (1) называется **преобразованием Фурье x(t)**, а равенство (2) называется **обратным преобразованием Фурье X(f)**, в результате чего получается **x(t)**.

Рассмотрим внимательнее равенство (1):

Сигнал $x(t)$, умножается на экспоненту, имеющую некоторую частоту f , и затем интегрируется по ВСЕМУ ВРЕМЕНИ!!! (Ключевые слова здесь "все время", как будет показано ниже).

Отметим, что экспоненциальный член в (1) может быть также представлен как:

$$\cos(2\pi ft) + j \sin(2\pi ft) \quad (3)$$

Вышеприведенное выражение имеет вещественную часть (косинус частоты f) и мнимую (синус частоты f). Так что на самом деле мы умножаем исходный сигнал на комплексное выражение, состоящее из синусов и косинусов частоты f . Затем мы интегрируем это произведение. Другими словами, мы складываем все точки произведения. Если результат интегрирования имеет большое значение, то частота f существенно присутствует в сигнале $x(t)$. Если значение интеграла мало, то амплитуда частоты f в сигнале мала. Если интеграл равен нулю, то частота f отсутствует в сигнале.

Отметим, что интегрирование в выражении (1) осуществляется по времени. Однако, левая часть выражения (1) есть функция частоты. Поэтому, интеграл в (1) вычисляется для каждого значения f .

ВАЖНО(!) Интеграл берется от минус бесконечности до плюс бесконечности, то есть по всей временной оси. Поэтому, время существования той или иной частоты неважно: ее вклад будет все равно одинаковым. Именно поэтому преобразование Фурье не подходит для анализа сигналов с изменяющейся частотой, т.е. для нестационарных сигналов. Перед применением ПФ важно знать, стационарный сигнал или нет для оценки релевантности результатов.

Теперь становится более понятным пример из первой части. Повторим его еще раз. Взгляните на следующий рисунок, на котором представлен сигнал:

$$x(t) = \cos(2\pi 10t) + \cos(2\pi 25t) + \cos(2\pi 50t) + \cos(2\pi 100t)$$

т.е. состоящий из четырех частотных компонент 5, 10, 20 и 50Гц.

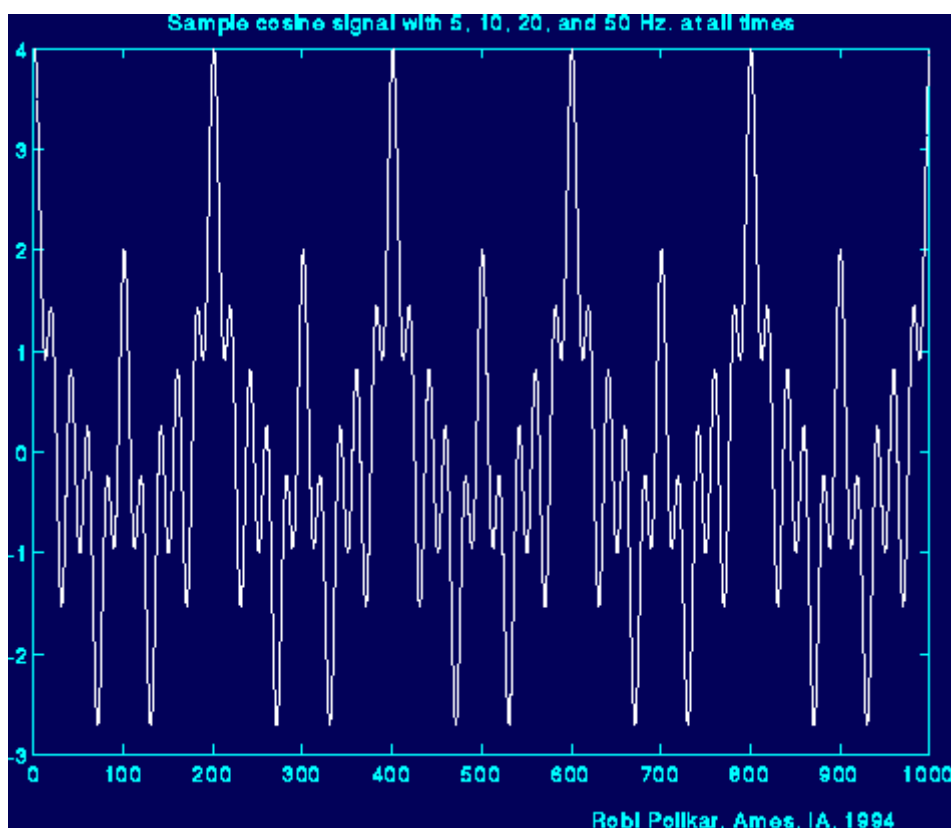


Рис.2-2.

На Рис.2-3 показано ПФ этого сигнала. На графике частотная ось «обрезана», но теоретически она простирается до бесконечности (для непрерывного ПФ). В данном случае мы вычисляем дискретное ПФ, в случае которого частотная ось простирается, по крайней мере, до удвоенной частоты дискретизации сигнала, а результат преобразования симметричен. Впрочем, для нас это неважно.

На графике отметим четыре пика, соответствующим четырем частотам.

На графике Рис.2-4 показан сигнал, в котором те же частоты встречаются в разные моменты времени, «по очереди».

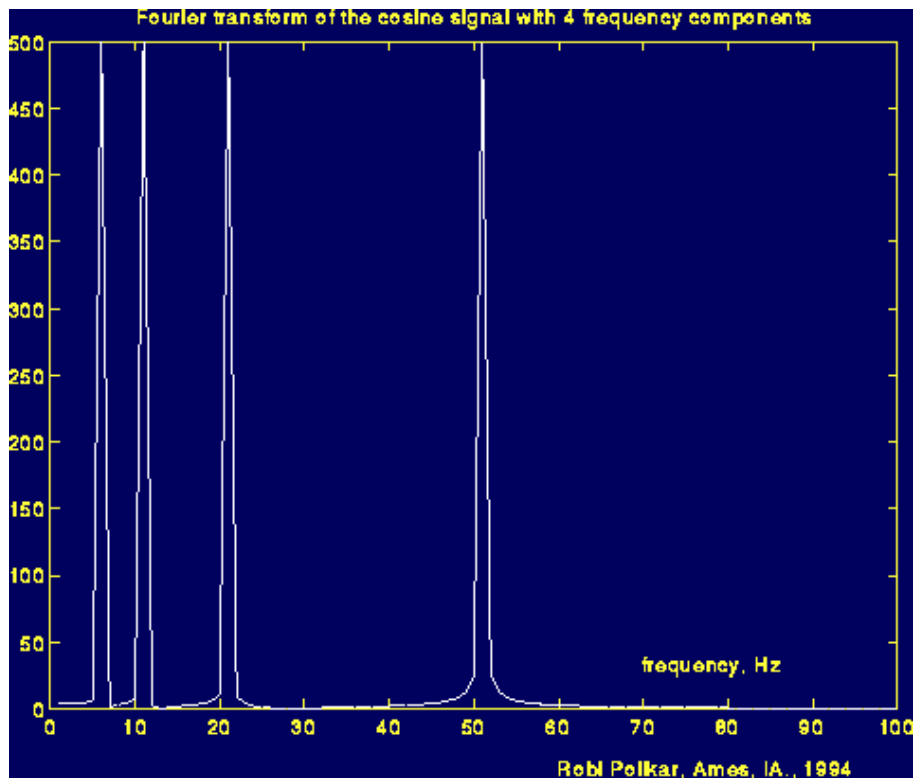


Рис.2-3.

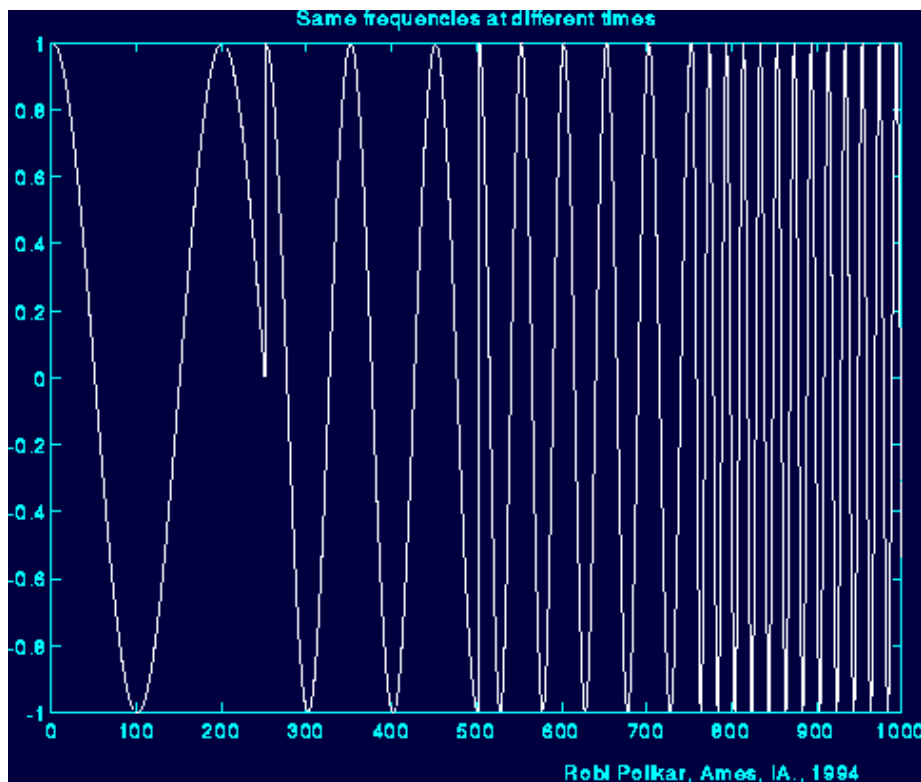


Рис.2-4.

А вот ПФ этого сигнала:

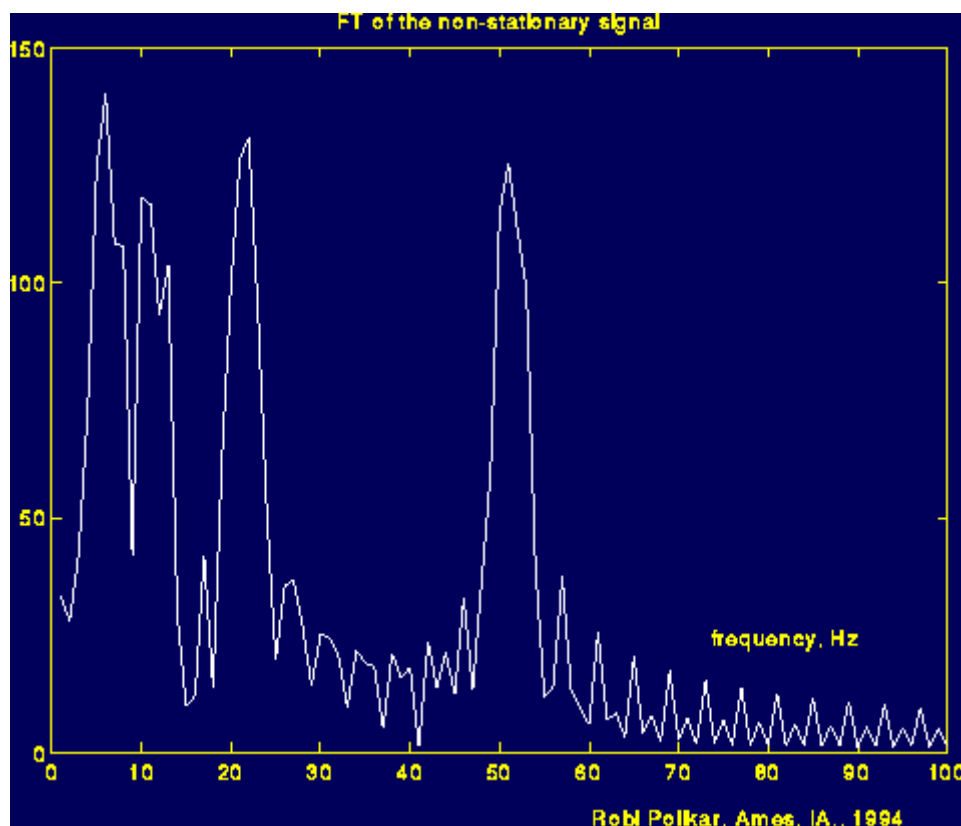


Рис.2-5.

Что ж, графики ПФ двух различных сигналов весьма похожи. На последнем видны еще и дополнительные значения частот, но они невелики, так как изменения частоты сигнала происходит редко. Применяя методы фильтрации, можно добиться полного подавления этих дополнительных частот. Таким образом, ПФ не может обеспечить различие двух этих сигналов.

Для решения этой задачи необходимо обратиться к частотно-временным преобразованиям. Одним из них является вейвлет-преобразование. Но оно было открыто сравнительно недавно. А как же раньше исследователи решали задачу анализа нестационарных сигналов?

Они использовали... **ЛИНЕЙНЫЕ ЧАСТОТНО-ВРЕМЕННЫЕ ПРЕДСТАВЛЕНИЯ.**

ОКОННОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

А что если предположить, что нестационарный сигнал кусочно-стационарен? Имеем ли мы право поступать таким образом?

Ответ: да.

Посмотрите на третий рисунок, представленный выше. Сигнал стационарен на каждом 250мс интервале.

Возможно, у вас возник вопрос:

А что если участок стационарности очень мал?

Что ж, тогда будем использовать окно, достаточно узкое для того, чтобы сигнал внутри него выглядел стационарным.

Такой подход получил название оконного (или кратковременного) преобразования Фурье (ОПФ).

При ОПФ сигнал делится на отрезки («окна»), в пределах которых его можно считать стационарным. Для этого к сигналу применяется оконная функция w , ширина которой должна быть равной ширине окна.

Пусть ширина оконной функции T сек. Тогда в момент времени $t=0$ она перекрывается с $T/2$ сек сигнала. Оконная функция и сигнал перемножаются. Если оконная функция прямоугольная и единичной высоты, то сигнал не изменяется. В противном случае он взвешивается с оконной функцией. Затем произведение подвергается преобразованию Фурье.

В результате мы получаем ПФ первых $T/2$ сек исходного сигнала. Если этот отрезок стационарен, как мы и предполагали, то полученный результат преобразования корректно отображает частотное наполнение первых $T/2$ сек сигнала.

Следующим шагом является сдвиг оконной функции на некоторую величину t_1 сек. Сдвинутая функция вновь умножается с сигналом, выполняется ПФ произведения. Эта процедура повторяется до достижения конца исходного сигнала. Все вышесказанное об ОПФ можно записать в следующем виде:

$$STFT_x^{(\omega)}(t', f) = \int_t [x(t) \cdot w^*(t-t')] \cdot e^{-j2\pi ft} dt \quad (4)$$

Рис.2-6.

Взгляните внимательно на это выражение. $x(t)$ - исходный сигнал, $w(t)$ - оконная функция, * означает комплексное сопряжение. Как видно из выражения, ОПФ есть не что иное, как ПФ сигнала, умноженного на оконную функцию.

Для каждого t' и f вычисляется свой коэффициент ОПФ. Следующий рисунок поможет понять вам это лучше:

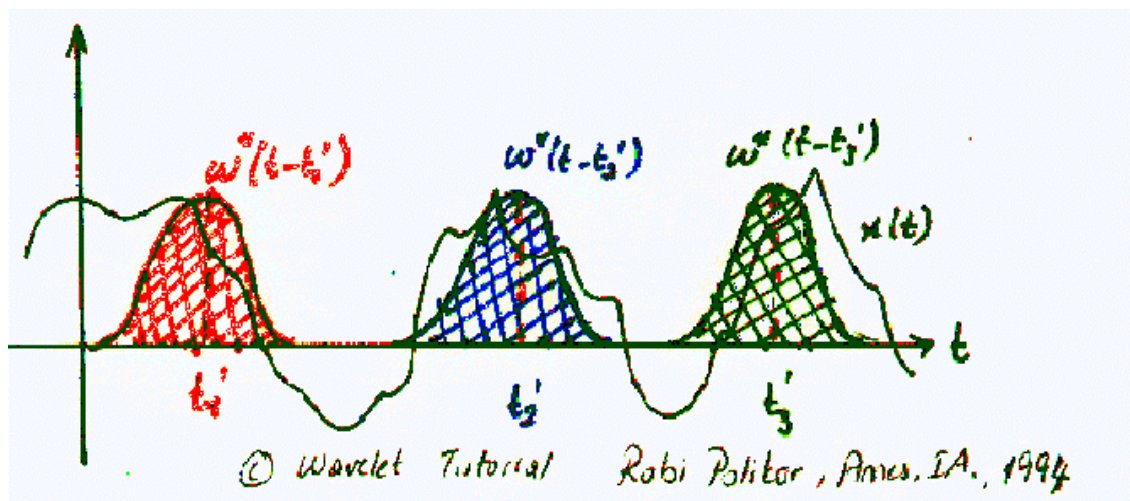


Рис.2-7.

Здесь колоколообразные гауссовские функции являются оконными функциями. Красным цветом показана функция в $t=t_1'$, синим - в $t=t_2'$ и зеленым - в $t=t_3'$. Это соответствует трем разным ПФ для трех моментов времени. Таким образом мы получаем истинное **частотно-временное преобразование** (ЧВП) сигнала.

Вероятно, наилучшим способом объяснения является разбор примера. Во-первых, так как наше преобразование является функцией как времени, так и частоты (в отличие от ПФ, которое зависит только от частоты), то оно является двумерным (а с учетом амплитуды, то и трехмерным). Пусть дан нестационарный сигнал, например, показанный на Рис.2-8:

В этом сигнале в разные моменты времени присутствуют различные частотные компоненты: от 0 до 250 мс - 300Гц, и, далее 200, 100 и 50Гц. Посмотрите на ОПФ этого нестационарного сигнала на Рис.2-9:

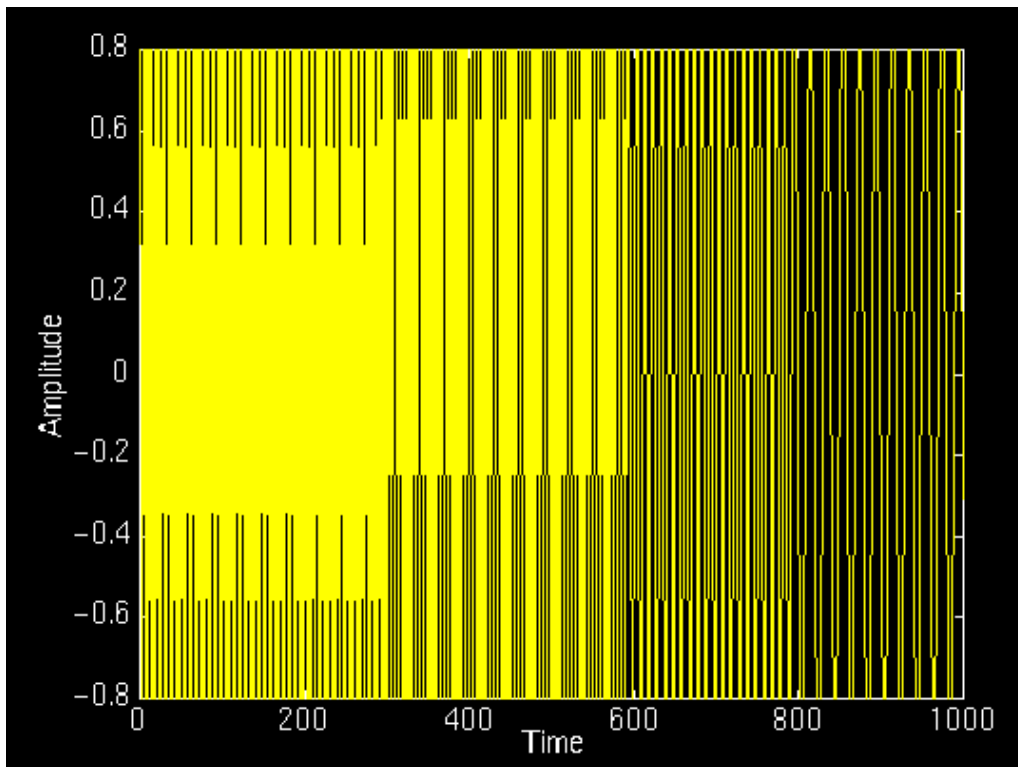


Рис.2-8.

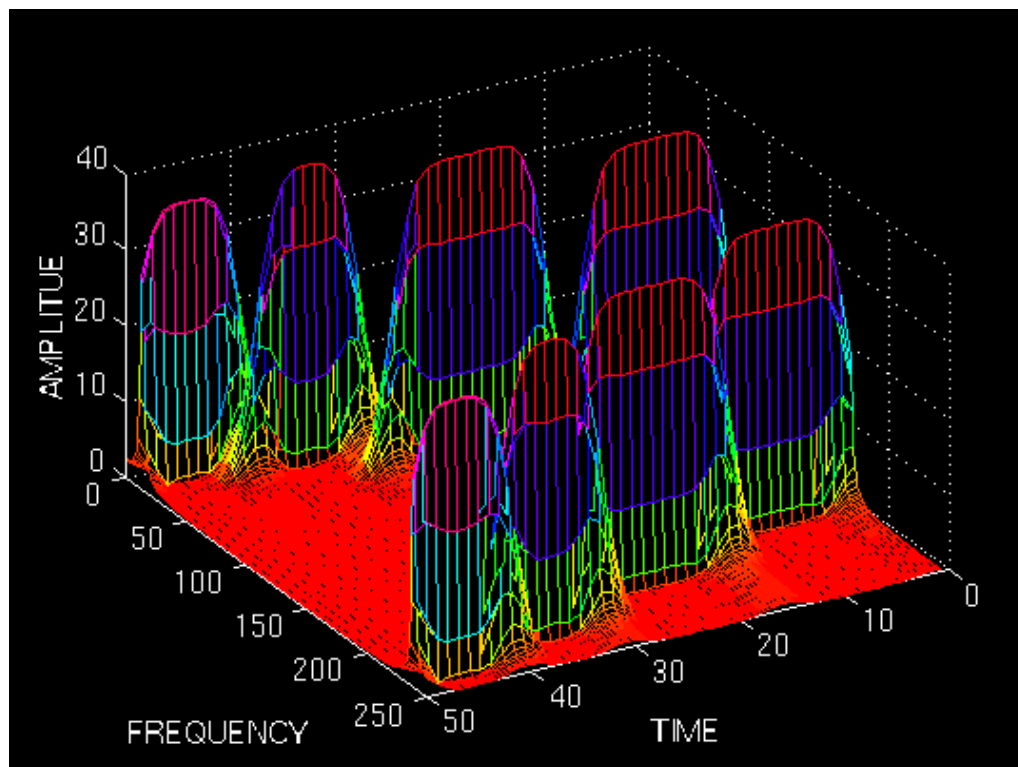


Рис.2-9.

Как мы и говорили, это трехмерный график. По осям «х» и «у» отложены время и частота, соответственно. Не обращайте внимания на числа, которые проставлены по осям: они получены после нормализации, которую мы здесь не рассматриваем. Давайте посмотрим на огибающую частотно-временного представления.

Во-первых, заметьте, что график симметричен относительно оси частот. Вспомните, что ПФ вещественного сигнала всегда симметрично, хотя мы и не показывали вторую половинку преобразования. Так как ОПФ основывается на ПФ, оно также симметрично. Говорят, что симметричная часть связана с отрицательными частотами. Это понятие мы не будем здесь рассматривать, да оно и неважно в контексте повествования.

Что действительно важно, так это наличие на графике четырех пиков, соответствующих четырем частотным компонентам. Заметьте, что в отличие от ПФ, эти пики локализованы в различных временных интервалах, как это имело место и в исходном сигнале.

Так что теперь мы имеем истинное частотно-временное представление сигнала. Мы не только знаем, какие частотные компоненты присутствуют в сигнале, но и в какой момент времени они встречаются. Это великолепно!!! Не так ли?

Не совсем так!

Вы можете спросить, если ОПФ дает частотно-временное представление сигнала, то для чего же нам вейвлет-преобразование? Присущий ОПФ недостаток не виден из рассмотренного примера. Этот пример как раз и был подобран для четкой и ясной демонстрации возможностей ОПФ.

Проблемы ОПФ имеют свои корни в явлении, которое называется **принципом неопределенности Гейзенберга**. Этот принцип в применении к ЧВП гласит, что невозможно получить произвольно точное частотно-временное представление сигнала, то есть нельзя определить для какого-то момента времени, какие спектральные компоненты присутствуют в сигнале. Единственное, что мы **можем** знать, так это **временные интервалы**, в течение которых в сигнале существуют **полосы частот**. Эта проблема называется проблемой разрешения.

Проблема ОПФ связана с **шириной** используемой оконной функции. Эта ширина называется еще **носителем** функции. Если окно достаточно узкое, то говорят о **компактном носителе**. Как мы увидим в дальнейшем, эта терминология особенно широко используется в теории вейвлет-преобразования.

Вот что происходит:

Вспомним, что при ПФ не существует проблемы разрешения в частотной области, аналогично нет проблем с разрешением во времени при временном представлении сигнала, так как мы точно знаем значение сигнала в каждый

момент времени. Однако, временная информация при ПФ, также как и частотная, при обычном представлении сигнала отсутствует. Оптимальное частотное разрешение достигается при ПФ за счет того, что окно, используемое в данном случае, есть ядро ПФ, функция $\exp\{j\omega t\}$, которая простирается от минус бесконечности до плюс бесконечности. При ОПФ окно имеет конечную длину, накрывает только часть сигнала, поэтому и частотное разрешение ухудшается. Под ухудшением я понимаю то, что мы теперь не знаем точно присутствующих в сигнале частот, а только полосы частот:

Вы можете спросить, а почему бы нам не сделать длину окна при ОПФ бесконечной, как при ПФ? Да, но тогда мы потеряем всю временную информацию, получив ПФ вместо ОПФ.

Итак, чем уже окно, тем лучше временное разрешение, но хуже частотное. И наоборот. Кроме того, чем уже окно, тем более справедливыми становятся наши предположения о стационарности сигнала в пределах окна.

Для того чтобы наблюдать эти эффекты, обратимся к примерам: Рассмотрим четыре окна разной ширины, вычислим с ними ОПФ и посмотрим что получится:

В качестве оконной функции будем использовать функцию Гаусса вида:

$$w(t) = e^{\frac{-at^2}{2}}$$

где **a** определяет ширину окна, а **t** - время. На следующем рисунке показаны четыре окна различной ширины, определяемой значением **a**. Не обращайте внимания на числовые значения **a**, так как функцию определяет также и временной интервал, на котором она вычисляется.

Рассмотренный ранее пример был посчитан при значении $a=0.001$. Теперь рассмотрим ОПФ того же сигнала при другом значении ширины окна.

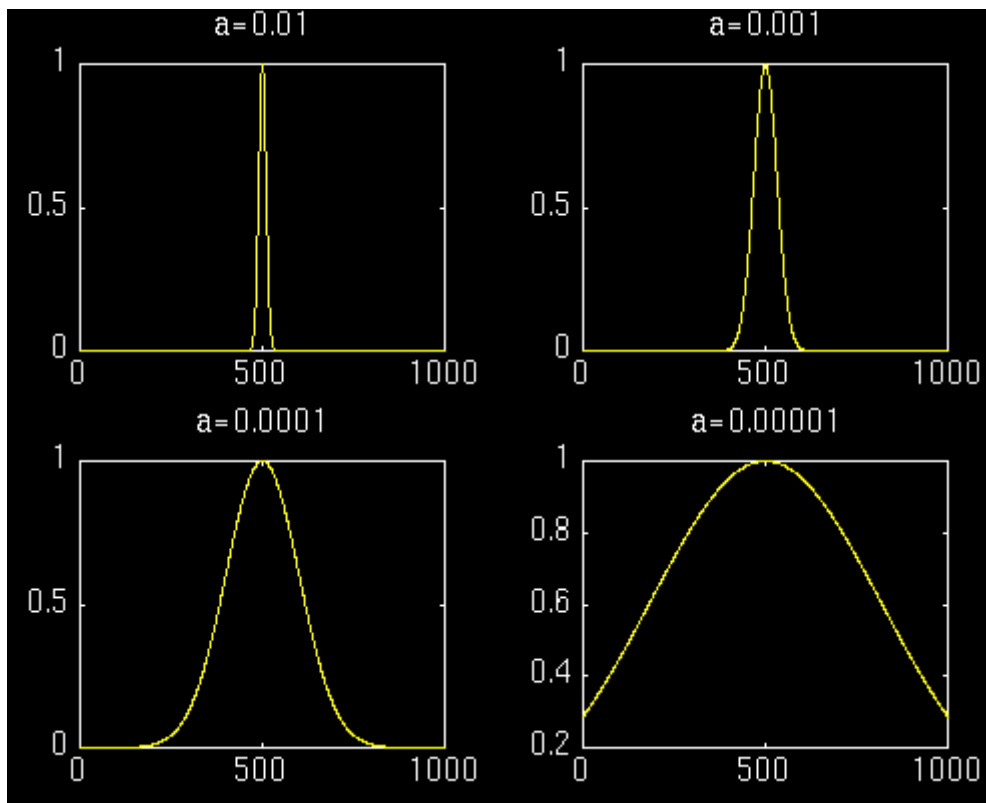


Рис.2-10.

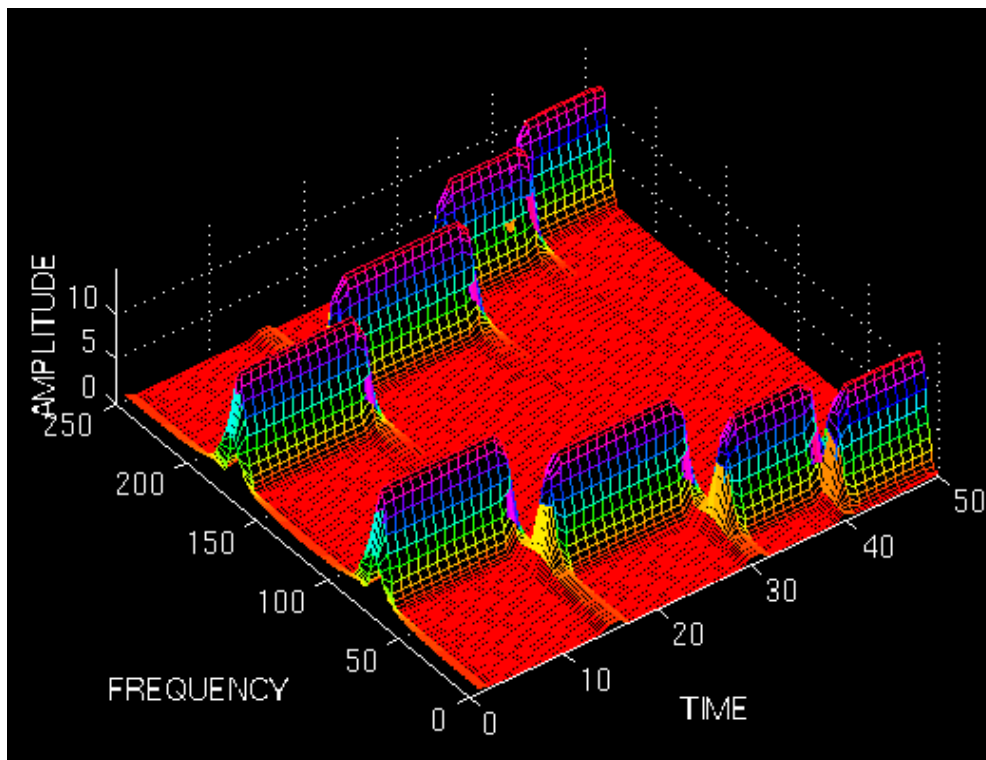


Рис.2-11.

Для начала используем первое, самое узкое окно. Мы можем ожидать хорошее разрешение по времени, но плохое по частоте (Рис.2-11).

Отметим, что четыре пика, показанных на рисунке, хорошо разделены по времени. Также отметим, что в частотной области каждый пик покрывает диапазон частот, а не одну какую-то частоту. Теперь увеличим ширину окна и посмотрим на следующий рисунок.

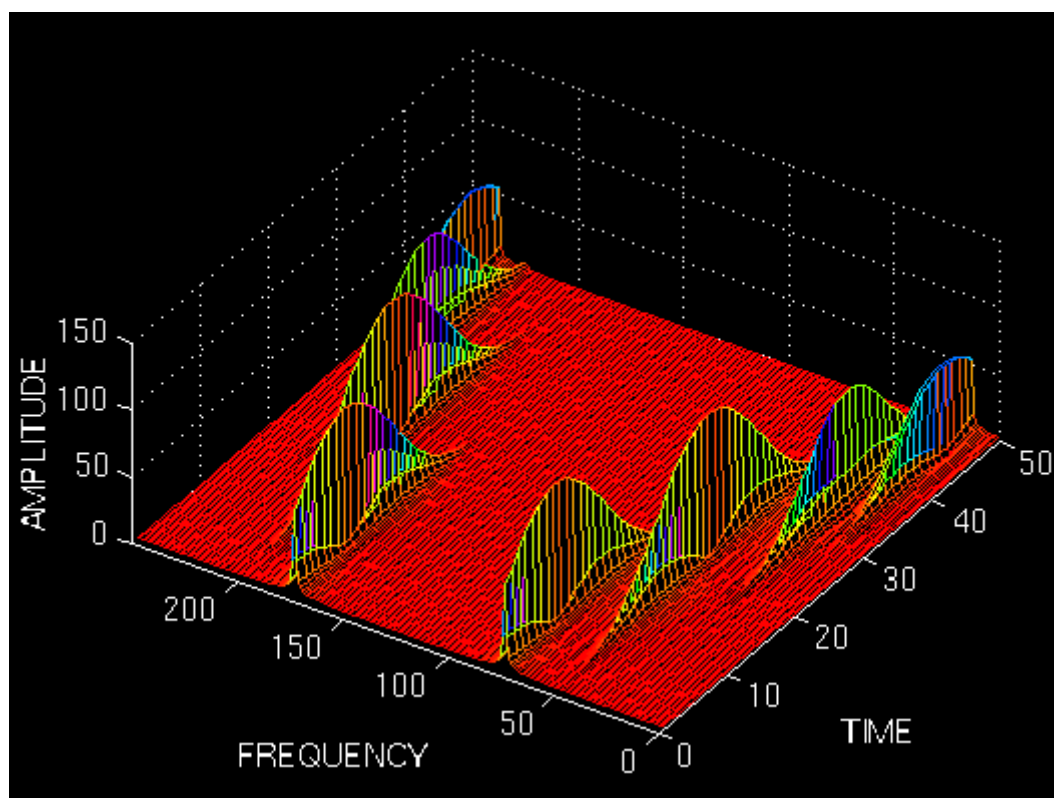


Рис.2-12.

Как видно из рисунка, пики теперь не столь хорошо разделены по времени. Однако частотное разрешение улучшилось. Увеличим еще ширину окна:

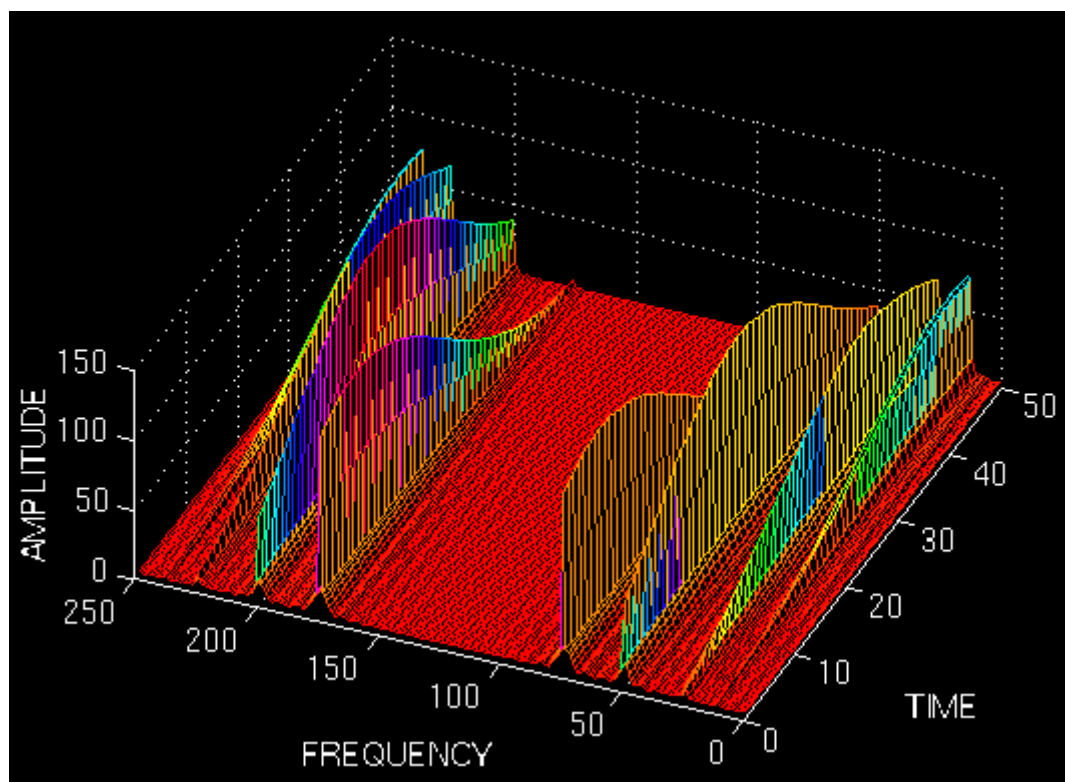


Рис.2-13.

Как и ожидалось, временное разрешение ухудшилось донельзя.

Приведенные примеры были призваны показать проблему разрешения, присущую ОПФ. Поэтому при применении ОПФ всегда возникают вопросы: какой вид окна использовать? Узкое окно обеспечивает лучшее временное разрешение, а широкое – лучшее частотное. Проблема состоит в том, что приходится выбирать окно «раз и навсегда», то есть для анализа всего сигнала, тогда как разные его участки могут требовать применения разных окон. Если сигнал состоит из далеко отстоящих друг от друга частотных компонент, то можно пожертвовать спектральным разрешением в пользу временного и наоборот.

Вейвлет-преобразование решает в какой-то степени эту проблему разрешения, как мы увидим в следующем разделе.

ЧАСТЬ III

КРАТНОМАСШТАБНЫЙ АНАЛИЗ

Несмотря на то, что проблема разрешения имеет физический характер и не может быть преодолена, существует возможность анализа сигнала при помощи альтернативного подхода, имя которому - кратномасштабный анализ (КМА). КМА, как видно из названия, анализирует сигнал на различных частотах и различном разрешении одновременно. Каждая спектральная компонента не анализируется отдельно, как это было в случае с ОПФ.

КМА позволяет получить хорошее разрешение по времени (плохое по частоте) на высоких частотах и хорошее разрешение по частоте (плохое по времени) на низких частотах. Этот подход становится особенно эффективным, когда сигнал имеет высокочастотные компоненты короткой длительности и протяженные низкочастотные компоненты. К счастью, именно такие сигналы и встречаются чаще всего на практике. Например, такой сигнал показан ниже. Он имеет сравнительно низкочастотную компоненту на протяжении всего сигнала и относительно высокую - на коротком интервале в середине сигнала.

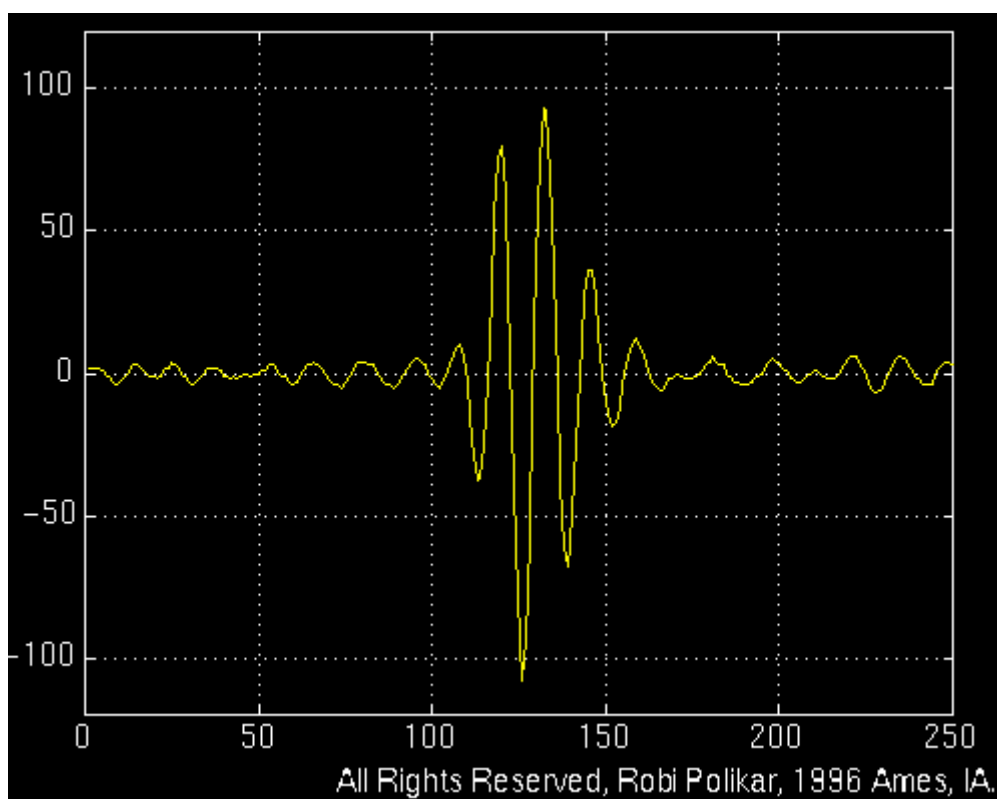


Рис.3-1.

НЕПРЕРЫВНОЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ

Непрерывный вейвлет-анализ (НВП) выполняется аналогично ОПФ, в том смысле, то сигнал перемножается с функцией (вейвлетом), также как и с оконной функцией при ОПФ, и преобразование выполняется отдельно для разных участков времени сигнала. Однако существует две существенные разности между ОПФ и НВП:

1. Не выполняется ПФ взвешенного с вейвлетом сигнала. Поэтому, единичный пик соответствует синусоиде, т.е. отрицательные частоты не вычисляются.
2. Ширина окна изменяется, так что преобразование вычисляется для каждой спектральной компоненты, что является наиболее важным свойством вейвлет-преобразования.

Непрерывное вейвлет-преобразование определяется следующим образом:

$$CWT_x^\psi(\tau, s) = \Psi_x^\psi(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \psi^* \left(\frac{t - \tau}{s} \right) dt$$

Уравнение 3-1.

Как видно из равенства, преобразованный сигнал есть функция двух переменных, **tau** и **s**, параметры сдвига и масштаба, соответственно. **psi(t)** – функция преобразования, называемая материнским вейвлетом. Этот термин обязан своему появлению двум важным свойствам вейвлет-преобразования, объясняемым ниже:

Слово вейвлет означает маленькая волна. Под маленькой понимается то, что эта функция (окно) имеет конечную ширину (компактный носитель). Слово «волна» отражает тот факт, что вейвлет-функция осциллирует. Термин «материнский» означает, что функции с различной шириной носителя, используемые в преобразовании, порождаются одной базовой функцией – материнским вейвлетом. То есть материнский вейвлет является прототипом для всех оконных функций.

Термин сдвиг используется здесь в том же смысле, что и при ПФ: он относится к местоположению окна, и окно движется вдоль сигнала. Этот термин относится, таким образом, к временной информации, присутствующей в результате преобразования. Однако при ВП мы не имеем частотного параметра, как это было при ОПФ. Вместо него здесь имеется параметр масштаба, который можно определить как величину, обратную частоте. Понятие масштаба более подробно объясняется в дальнейшем.

МАСШТАБ

Параметр масштаба в вейвлет-анализе имеет аналогию с масштабом географических карт. Большие значения масштаба соответствуют малому количеству деталей, глобальному представлению сигнала, а низкие значения масштаба позволяют различить детали. Аналогично, в терминах частоты, низкие частоты соответствуют глобальной информации о сигнале (которая содержится на всей его протяженности), а высокие частоты - детальной информации, скрытым особенностям, которые имеют обычно малую протяженность. На следующем рисунке приведены косинусные сигналы на различных масштабах.

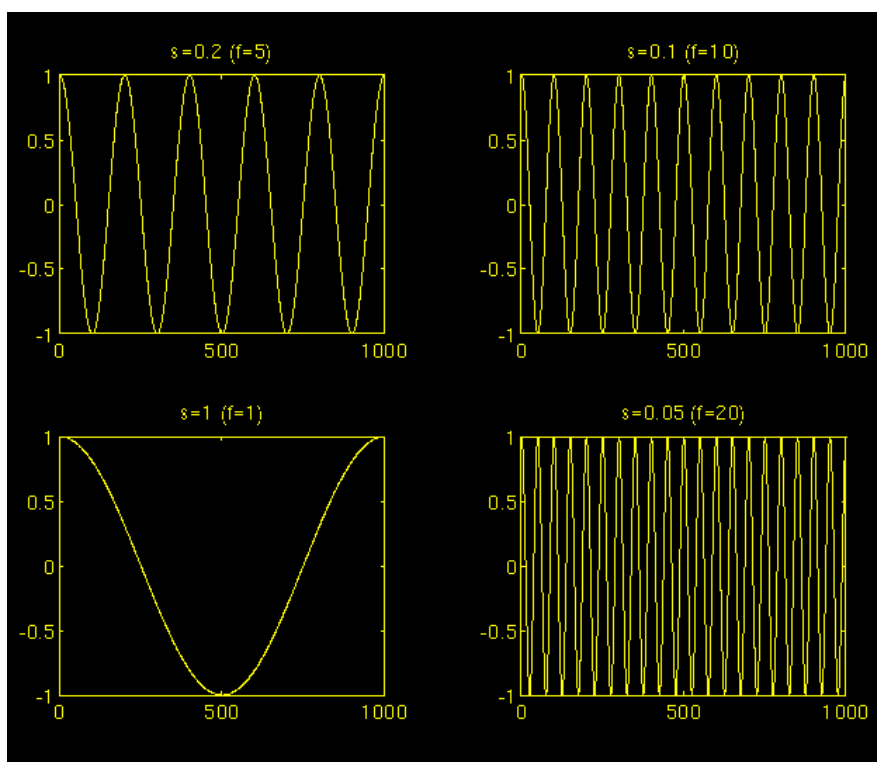


Рис.3-2.

Масштабирование, как математическая операция, расширяет или сжимает сигнал. Большие значения масштабов соответствуют расширениям сигнала, а малые значения - сжатым версиям. Все сигналы, показанные на рисунке, получены из одного и того же косинусного сигнала путем его сжатия и расширения. На приведенном выше рисунке $s=0.05$ - наименьший масштаб, а $s=1$ - наибольший.

Если $f(t)$ исходная функция, то $f(st)$ соответствует сжатой версии $f(t)$, если $s > 1$ и расширенной версии, если $s < 1$.

Однако в определении вейвлет-преобразования коэффициент масштаба стоит в знаменателе. Поэтому, $s > 1$ расширяет сигнал, а $s < 1$ сжимает его. Эти обозначения и будут использоваться в дальнейшем.

ВЫЧИСЛЕНИЕ НВП

В данном разделе приведена интерпретация вышеприведенного равенства. Пусть $x(t)$ - анализируемый сигнал. Выбирается материнский вейвлет, который будет прототипом для всех функций (окон), которые получаются из него путем сжатия (расширения). Существует несколько функций, применяющихся в качестве материнских вейвлетов. Двумя примерами являются вейвлеты Морле и Мексиканская шляпа, которые и используются в примерах раздела.

После выбора материнской функции вычисления начинаются с масштаба $s=1$. НВП вычисляется для всех значений s , меньших и больших «1». Однако полное преобразование обычно не требуется, так как реальные сигналы ограничены по полосе. Поэтому число масштабов может быть ограничено. В примерах этого раздела мы также используем ограниченное количество масштабов.

Процедура анализа стартует с масштаба $s=1$ и продолжается при увеличивающихся значениях s , т.е. анализ начинается с высоких частот и проводится в сторону низких частот. Первое значение s соответствует наиболее сжатому вейвлету. При увеличении значения s вейвлет расширяется.

Вейвлет помещается в начало сигнала, в точку, соответствующую времени $t=0$. Вейвлет-функция масштаба «1» умножается с сигналом и интегрируется на всем временном интервале. Интеграл умножается на константу $1/\sqrt{s}$ для нормализации, т.е. для того, чтобы сигнал на каждом масштабе имел бы одинаковую энергию.

Вейвлет масштаба $s=1$ затем сдвигается вправо на τ к точке $t=\tau$, и процедура повторяется. Получаем еще одно значение, соответствующее $t=\tau$, $s=1$ на частотно-временном плане.

Эта процедура повторяется до тех пор, пока вейвлет не достигнет конца сигнала. Таким образом получаем строку точек на масштабном-временном плане для масштаба $s=1$.

Теперь увеличим s на некоторое значение. Вообще говоря, так как преобразование непрерывное, то τ и s должны меняться непрерывно. При выполнении преобразования в компьютере мы вычисляем аппроксимацию, увеличивая оба параметра на некоторое малое значение. Тем самым мы осуществляем дискретизацию масштабного- временной плоскости.

Приведенная выше процедура повторяется для каждого значения s . При этом строка за строкой заполняется масштабном-временная плоскость. Так вычисляется НВП. Ниже рисунки иллюстрируют процесс преобразования шаг за шагом.

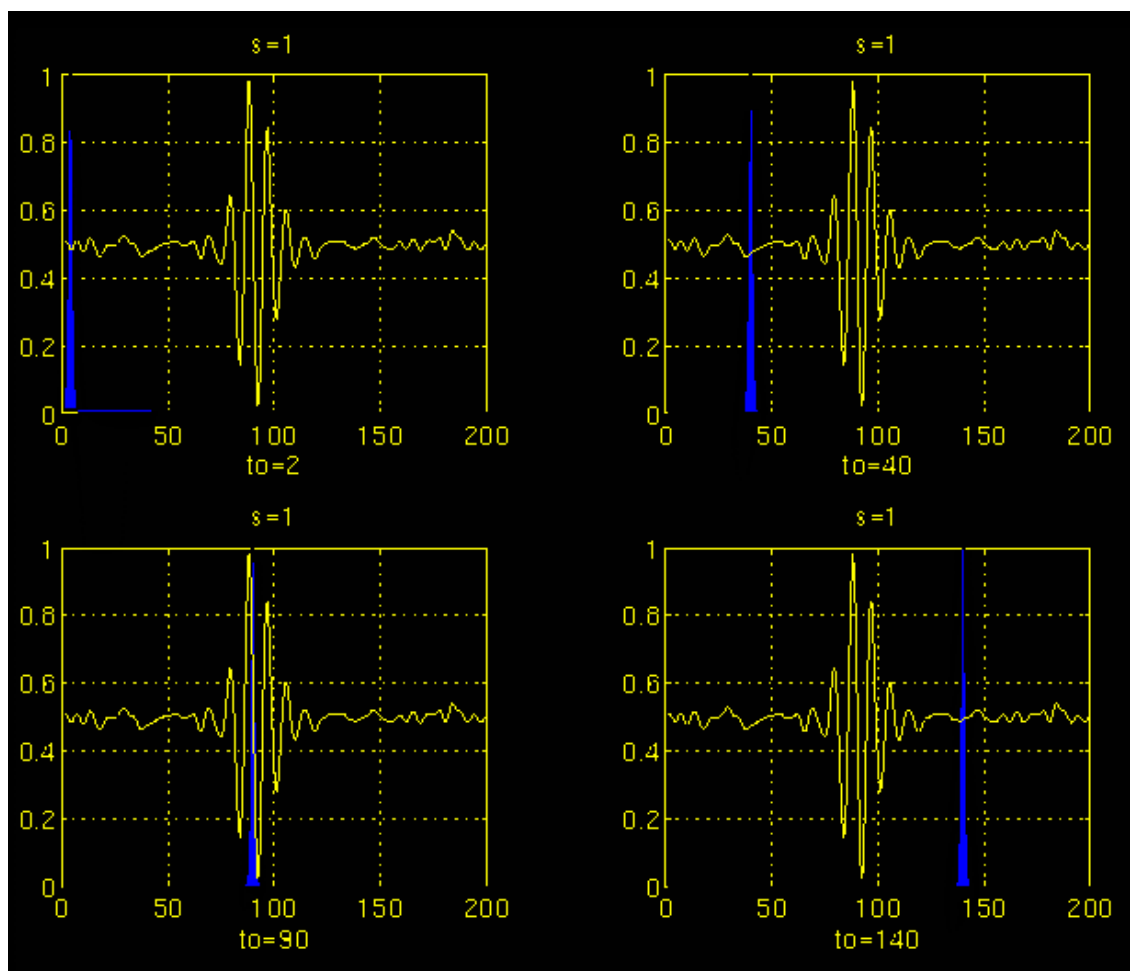


Рис.3-3.

На Рис.3-3 показаны сигнал и вейвлет-функция для четырех различных значений τ . В качестве сигнала используем усеченную версию сигнала, показанного на Рис.3-1. Значение масштаба, 1, соответствует наименьшему значению, или наибольшей частоте. Заметьте, сколь компактен носитель. Он должен быть так же узок, как и время жизни наивысшей частоты сигнала. На рисунке показаны четыре различных позиции вейвлет-функции в точках $\tau=2$, $\tau=40$, $\tau=90$ и $\tau=140$. В каждой позиции она умножается с сигналом. Произведение будет ненулевым лишь в случае, когда сигнал пересекается с носителем вейвлета, и нулевым – в остальных случаях. Сдвигам вейвлета по времени соответствует временная локализация сигнала, а сдвигам по масштабу – масштабная (частотная) локализация.

Если в сигнале присутствуют спектральные компоненты, соответствующие текущему значению s (которое в данном случае 1), то произведение вейвлета с сигналом в интервале, где эта спектральная компонента присутствует, дает относительно большое значение. В противном случае - произведение мало или

равно нулю. Сигнал, показанный на Рис.3-3, имеет спектральные компоненты, сравнимые с шириной окна при $s=1$ на интервале около $t=100$ мс.

НВП сигнала, показанного на Рис.3-3, дает большие значения для низких масштабов около времени $t=100$ мс и малые значения в остальных интервалах времени. Для высоких масштабов, напротив, НВП дает большие значения почти на всей длительности сигнала, так как низкие частоты присутствуют в нем все время.

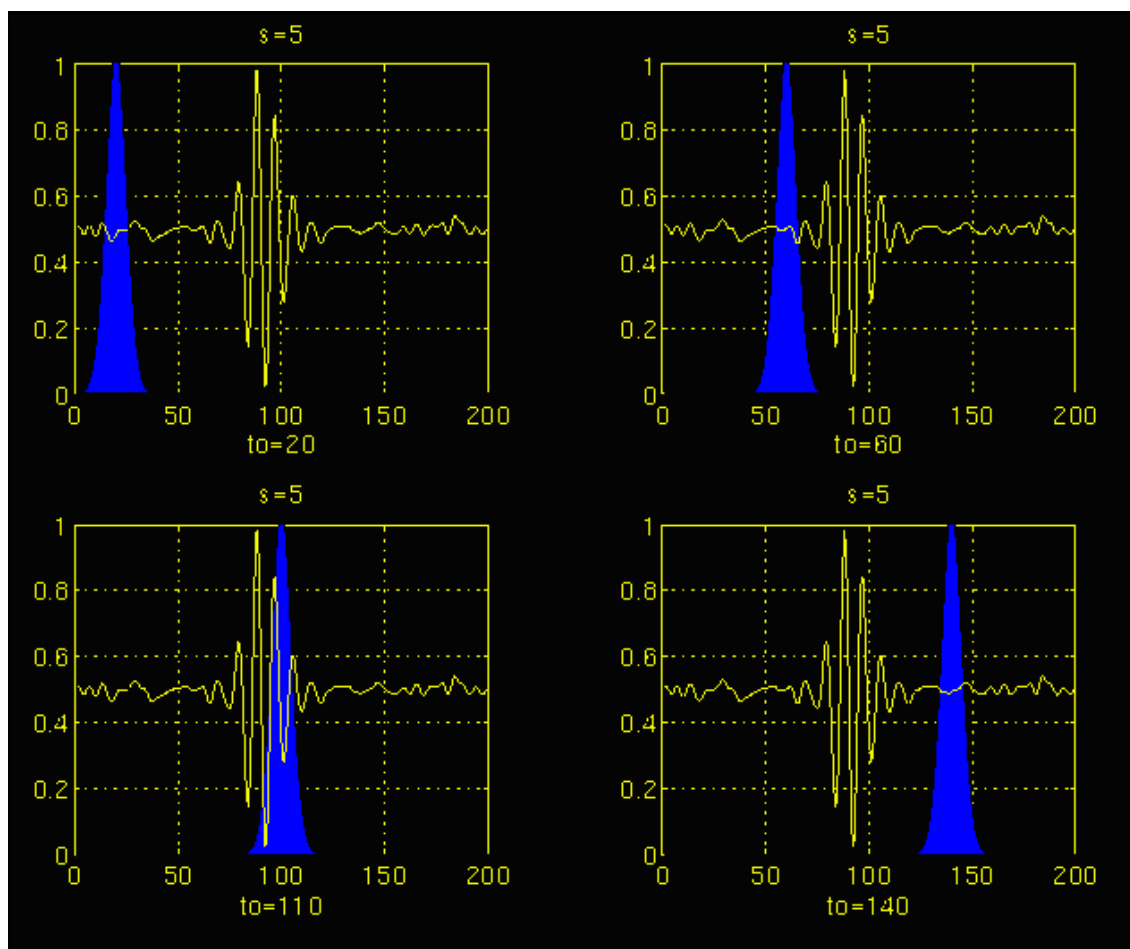


Рис.3-4.

На Рис.3-4,3-5 показан тот же процесс для масштабов $s=5$ и $s=20$, соответственно. Отметим, что ширина окна изменяется с увеличением масштаба. С увеличением ширины окна преобразование выделяет все более низкие частоты.

В конечном счете мы получаем точку на масштабно-временной плоскости для каждого значения масштаба и времени. Вычисления при фиксированном масштабе дают строку на плоскости, а вычисления при фиксированном времени - столбец.

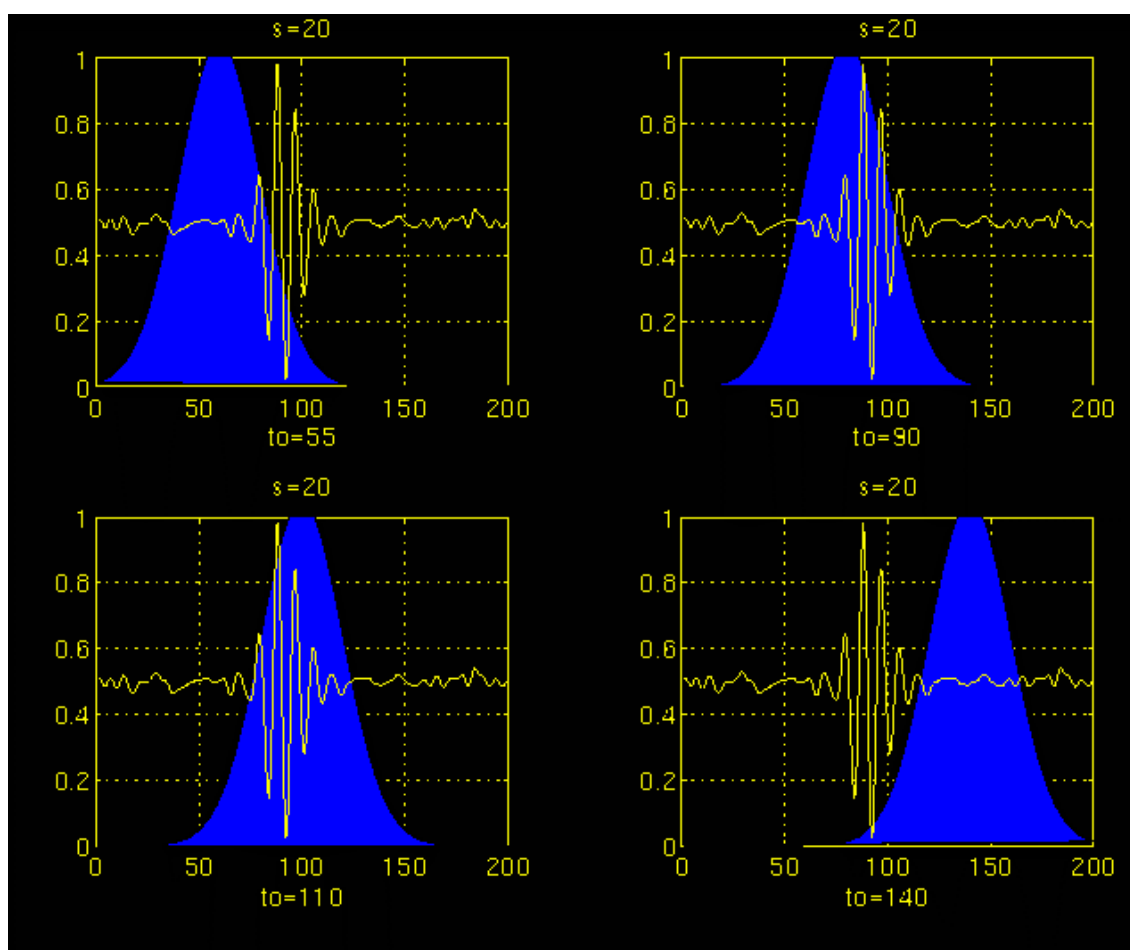


Рис.3-5.

Теперь рассмотрим пример и увидим, на что похож результат вейвлет-преобразования. Пусть имеем нестационарный сигнал, показанный на Рис.3-6. Сигнал аналогичен сигналу, приведенному в примере для ОПФ, за исключением частот. В данном случае сигнал состоит из частот 30, 20, 10 и 5Гц.

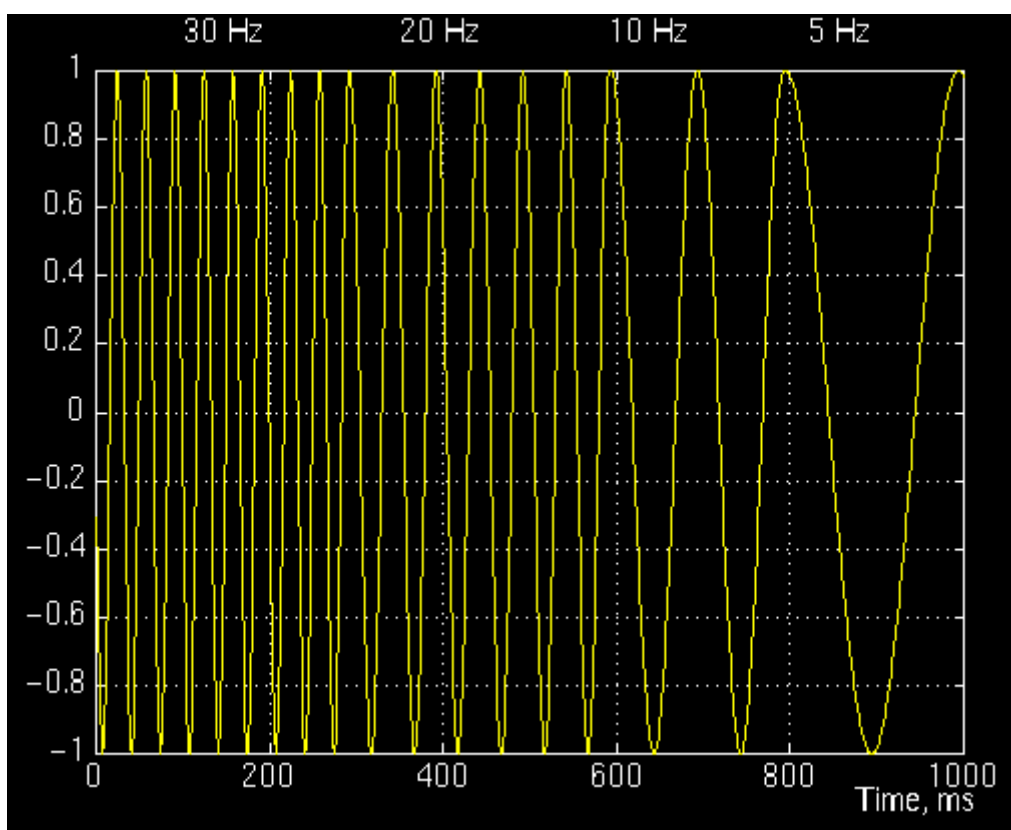


Рис.3-6.

На Рис.3-7 показано непрерывное вейвлет-преобразование этого сигнала. Заметим, что в качестве осей использованы сдвиги и масштаб, а не время и частота. Однако сдвиг тесно связан со временем, так как он показывает местоположение вейвлета по времени. Сдвиг материнского вейвлета может рассматриваться, как время, прошедшее с момента $t=0$. Масштаб же обратен частоте.

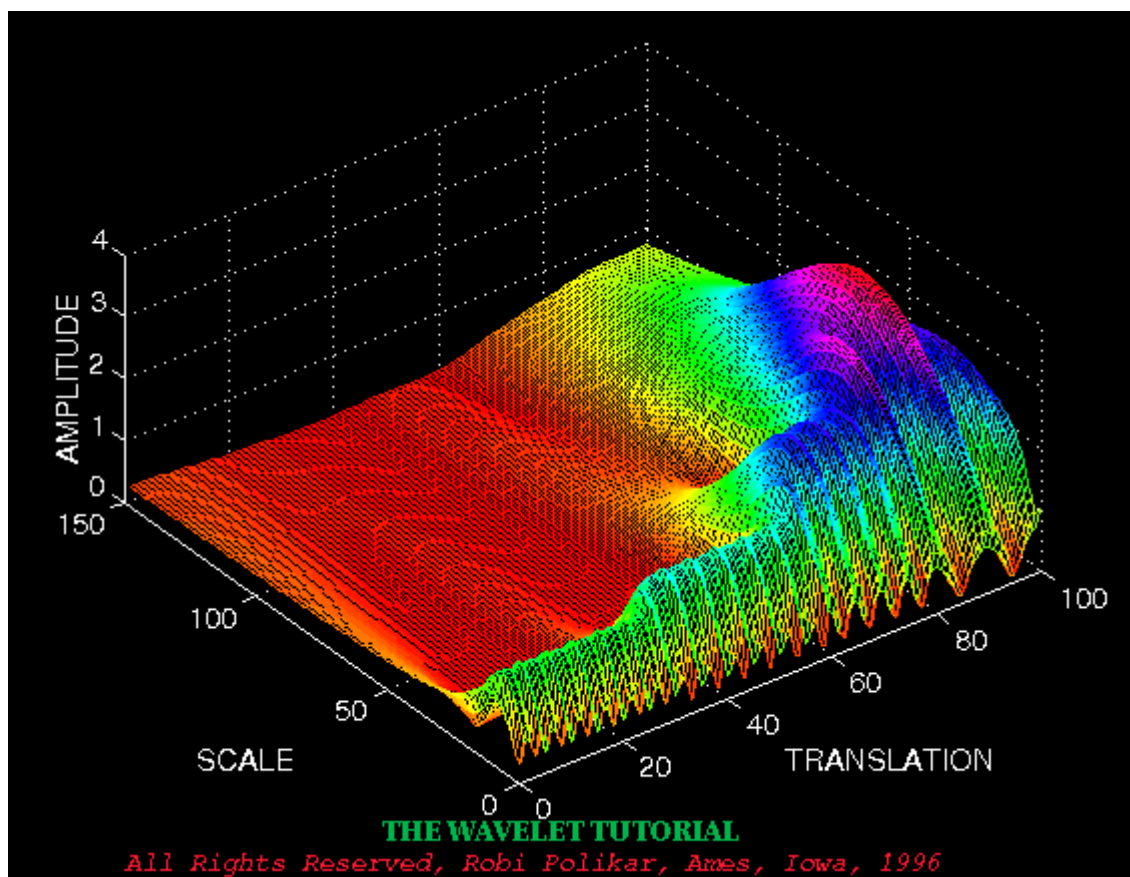


Рис.3-7.

Малые масштабы соответствуют высоким частотам. Поэтому, на Рис.3-7 часть графика, где масштабы близки нулю, соответствует высоким частотам. Верхняя частота анализируемого сигнала 30Гц и она появляется на самых малых масштабах при сдвигах от 0 до 30. Самая низкая частота сигнала – 5Гц появляется в конце оси сдвигов и на самых больших масштабах, как и ожидалось.

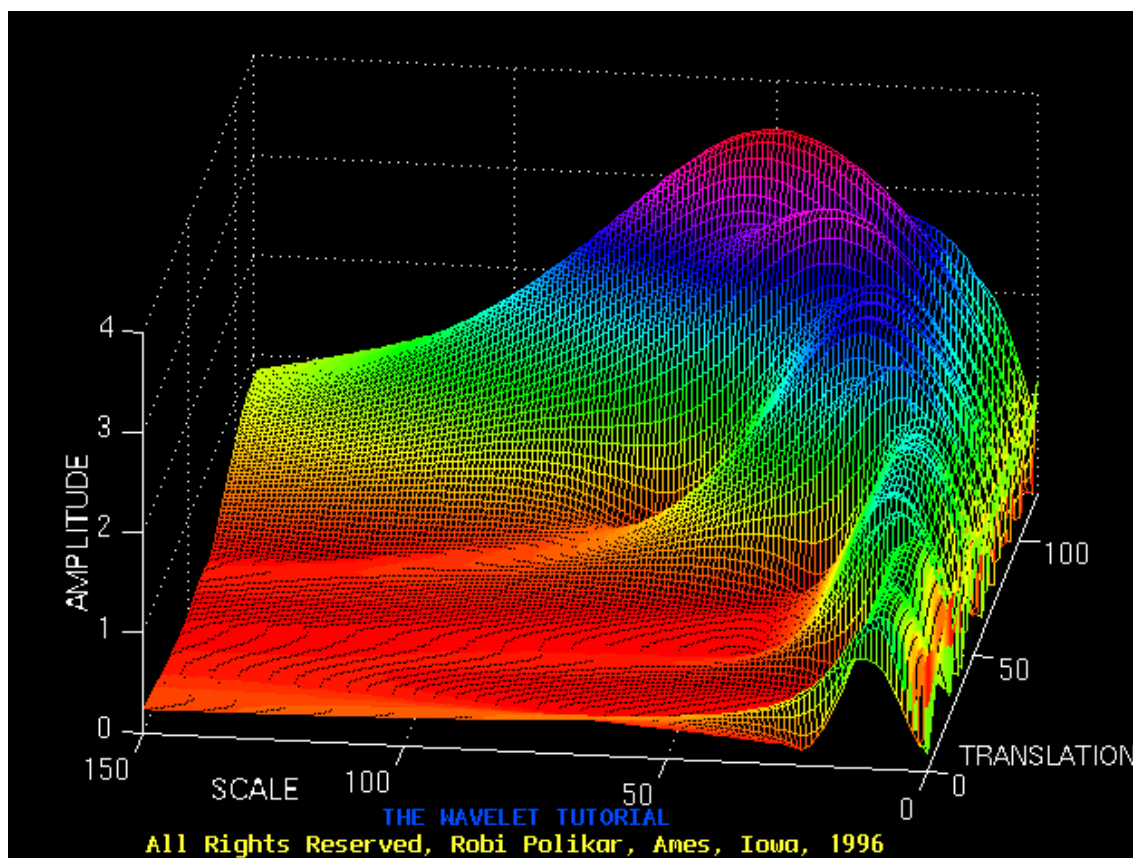


Рис.3-8.

Теперь рассмотрим свойства разрешения: в отличие от ОПФ, которое имеет фиксированное разрешение для всех времен и частот, НВП имеет хорошее разрешение по времени и плохое по частоте на высоких частотах. На нижних частотах - наоборот. На Рис.3-8 показана та же плоскость, что и на рис.3.7, но под другим углом обзора, чтобы лучше проиллюстрировать свойства разрешения. На Рис.3-8 на высоких частотах мы наблюдаем лучшее разрешение по масштабу, (т.е. масштабы узки, мы можем точно сказать, какой масштаб использован), что соответствует худшему разрешению по частоте. На нижних частотах (больших масштабах) мы наблюдаем обратную картину.

Оси графиков Рис.3-7 и 3-8 нормализованы. Вообще говоря, 100 точек оси сдвигов соответствуют 1000 мс, а 150 точек оси масштабов соответствуют полосе частот 40Гц. (Числа, которые стоят по осям, не соответствуют значениям частоты и времени, это лишь номера отсчетов при вычислении).

РАЗРЕШЕНИЕ ПО ВРЕМЕНИ И ПО ЧАСТОТЕ

В этом разделе мы глубже исследуем свойства разрешения вейвлет-преобразования. Вспомним, что проблема разрешения была главной причиной перехода нашего внимания от ОПФ к НВП.

Для объяснения разрешения при вейвлет-преобразовании обычно используется Рис.3-9. Каждый прямоугольник соответствует значению вейвлет-преобразования на частотно-временной плоскости. Площадь прямоугольников ненулевая, что означает то, что мы не можем точно вычислить какую-нибудь точку плоскости. Все точки, принадлежащие одному прямоугольнику, представляются одним значением вейвлет-преобразования.

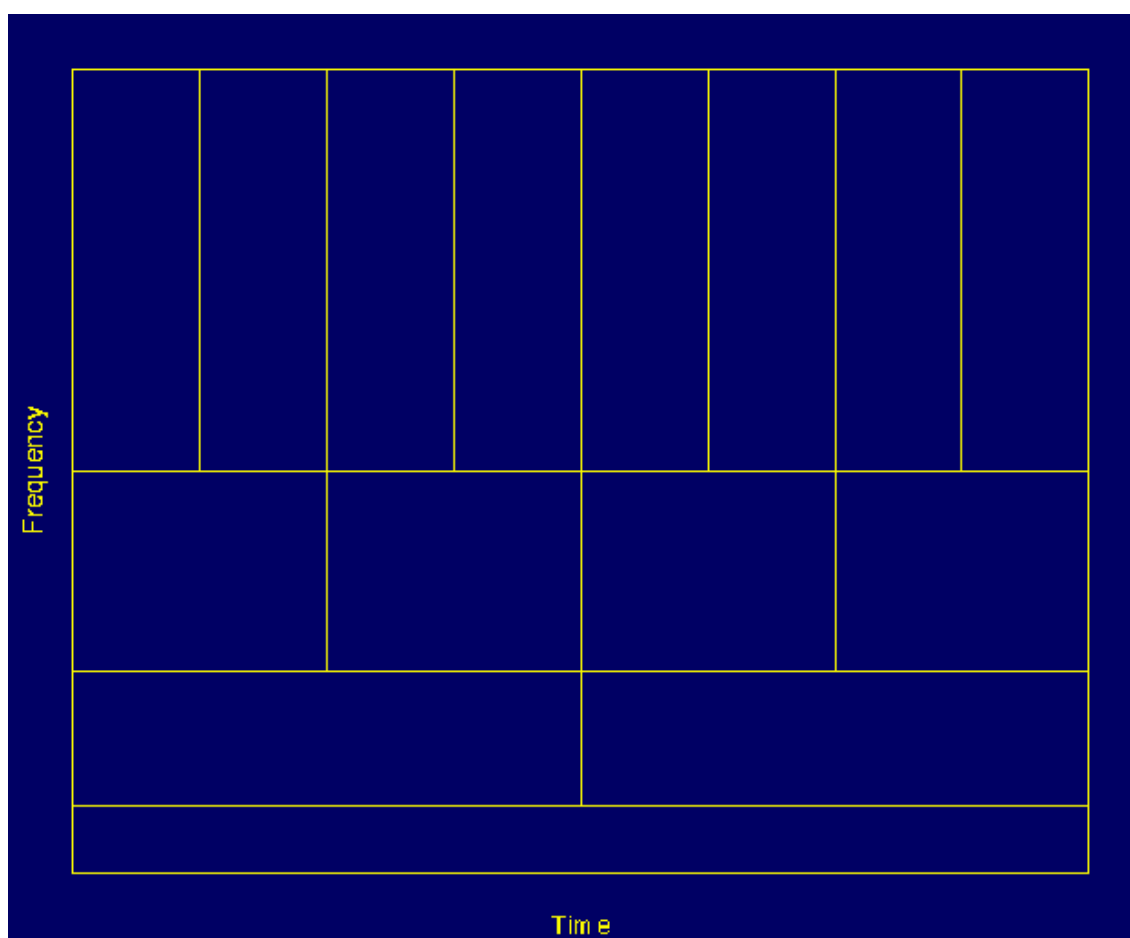


Рис.3-9.

Посмотрите внимательнее на Рис.3-9: прямоугольники разной ширины и высоты имеют одинаковую площадь. Каждый прямоугольник дает равный вклад в частотно-временную плоскость, но с различными долями частоты и времени. На нижних частотах высота прямоугольников меньше (что соответствует лучшему разрешению по частоте, так как меньше неопределенность насчет ее точного значения). Однако ширина прямоугольников больше (что соответствует худшему разрешению по времени). На высоких частотах разрешение по времени улучшается, а по частоте - ухудшается. В случае ОПФ ширина окна выбирается раз и навсегда для анализа всего сигнала. Поэтому частотно-временная плоскость ОПФ состоит из прямоугольников одинакового размера.

Площади прямоугольников ОПФ и вейвлет-преобразования равны и определяются принципом неопределенности Гейзенберга. Конечно, эта площадь зависит от используемой оконной функции при ОПФ или материнского вейвлета при вейвлет-преобразовании. Однако она ограничена снизу величиной $1/4 \pi$.

ТЕОРИЯ ВЕЙВЛЕТОВ: МАТЕМАТИЧЕСКИЙ ПОДХОД

В этом разделе рассмотрены основные идеи теории вейвлет-анализа, которые могут также рассматриваться как основа вообще любых методов обработки сигналов. Преобразование Фурье использует базисные функции для анализа и реконструкции функций. Любой вектор в векторном пространстве может быть записан в виде линейной комбинации базисных векторов этого пространства, т.е. путем умножения векторов на некоторые константы и затем взятия суммы произведений. Анализ сигнала заключается в оценивании этих констант (которые называются коэффициентами преобразования, или коэффициентами Фурье, вейвлет-коэффициентами и т.д.). Синтез, или реконструкция, заключается в вычислении линейной комбинации.

БАЗИСНЫЕ ВЕКТОРЫ

Обозначения: В большинстве выражений используются символы греческого алфавита. В тексте они записываются прямо по их именам, например, tau, psi, phi и т.д. Для обозначения прописных символов первая буква пишется прописной, например, Tau, Psi, Phi и т.д. Нижние индексы показаны знаком подчеркивания, τ , а верхние индексы – знаком \wedge . Все символы, записанные жирным шрифтом, означают векторы.

Базисом векторного пространства V называется множество линейно независимых векторов, таких, что любой вектор v из V может быть записан в виде их линейной комбинации. У векторного пространства может быть более одного базиса. Однако, число векторов в базисах одинаково, и это число называется размерностью векторного пространства. Например, в двумерном пространстве базис состоит из двух векторов.

$$v = \sum_k v^k b_k$$

Уравнение 3-2.

Выражение 3-2 показывает, как любой вектор v может быть записан в виде линейной комбинации базисных векторов b_k и соответствующих коэффициентов v^k .

От векторов можно легко перейти к функциям, заменив понятие базисных векторов b_k на базисные функции $\phi_k(t)$, и вектор v на функцию $f(t)$. Выражение 3-2 тогда принимает вид:

$$f(t) = \sum_k \mu_k \phi_k(t)$$

Уравнение 3-2а.

Для ПФ базисными функциями являются комплексные экспоненты (синусы и косинусы). Эти функции являются ортогональными, что упрощает осуществление реконструкции.

Пусть $f(t)$ и $g(t)$ - две функции в $L^2 [a,b]$. ($L^2 [a,b]$ означает множество функций, интегрируемых с квадратом на интервале $[a,b]$). Скалярное произведение этих двух функций определяется выражением 3-3:

$$\langle f(t), g(t) \rangle = \int_a^b f(t) g^*(t) dt$$

Уравнение 3-3.

В соответствии с приведенным выражением для скалярного произведения, НВП может рассматриваться как скалярное произведение анализируемого сигнала и базисных функций $\psi_{\tau,s}(t)$:

$$CWT_x^\psi(\tau, s) = \Psi_x^\psi(\tau, s) = \int x(t) \psi_{\tau,s}^*(t) dt$$

где,

$$\psi_{\tau,s} = \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-\tau}{s}\right)$$

Уравнения 3-4, 3-5.

Это определение НВП показывает, что вейвлет-анализ измеряет близость между базисными функциями (вейвлетами) и самим сигналом. Близость здесь понимается в смысле сходства частотного наполнения. Вейвлет-коэффициенты отражают близость сигнала к вейвлету данного масштаба.

СКАЛЯРНЫЕ ПРОИЗВЕДЕНИЯ, ОРТОГОНАЛЬНОСТЬ И ОРТОНОРМАЛЬНОСТЬ

Два вектора v, w называются ортогональными, если их скалярное произведение равно нулю:

$$\langle v, w \rangle = \sum_n v_n w_n^* = 0$$

Уравнение 3-6.

Аналогично, две функции $f(t)$ и $g(t)$ называются ортогональными, если их скалярное произведение равно нулю:

$$\langle f(t), g(t) \rangle = \int_a^b f(t) g^*(t) dt$$

Уравнение 3-7.

Множество векторов $\{v_1, v_2, \dots, v_n\}$ называется ортонормальным, если все векторы попарно ортогональны и имеют единичную длину. Это может быть выражено как:

$$\langle v_m, v_n \rangle = \delta_{mn}$$

Уравнение 3-8.

Аналогично, множество функций $\{\phi_k(t)\}$, $k=1,2,3,\dots$, называется ортонормальным, если

$$\int_a^b \phi_k(t) \phi_l^*(t) dt = 0 \quad k \neq l$$

Уравнение 3-9.

и

$$\int_a^b |\phi_k(t)|^2 dt = 1$$

Уравнение 3-10.

или, эквивалентно:

$$\int_a^b \phi_k(t) \phi_l^*(t) dt = \delta_{kl}$$

Уравнение 3-11.

где δ_{kl} есть дельта-функция Кронекера, определяемая как:

$$\delta_{kl} = \begin{cases} 1 & \text{если } k = l \\ 0 & \text{если } k \neq l \end{cases}$$

Уравнение 3-12.

Как отмечалось выше, в пространстве может быть больше одного множества базисных функций (векторов). Среди них особую важность имеет ортонормальный базис функций (векторов) в силу его привлекательных свойств нахождения коэффициентов анализа. Ортонормальный базис позволяет вычислить эти коэффициенты очень простым образом.

Для ортонормального базиса коэффициенты μ_k могут быть найдены как

$$\mu_k = \langle f, \phi_k \rangle = \int f(t) \phi_k^*(t) dt$$

Уравнение 3-13.

и функция $f(t)$ может затем быть восстановлена из выражения 3-2а путем подстановки коэффициентов μ_k . Это дает

$$f(t) = \sum_k \mu_k \phi_k(t) = \sum_k \langle f, \phi_k \rangle \phi_k(t)$$

Уравнение 3-14.

Так как не всегда удастся использовать ортонормальный базис, зачастую используют биортогональный базис. Термин «биортогональный» означает два различных базиса, ортогональных друг другу. Между тем каждый из них не является ортогональным.

Иногда и биортогональные базисы применить не удастся. Тогда используют фреймы. Фреймы составляют важную часть теории вейвлетов и интересующиеся математикой могут обратиться к книге Добеши (переведена на русский язык).

Приведем некоторые примеры НВП, в том же порядке, как мы их приводили во втором разделе для ОПФ.

Познакомьтесь с двумя материнскими функциями, чаще всего применяющимися в НВП. Вейвлет Мексиканская шляпа определяется как вторая производная от функции Гаусса:

$$w(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{t^2}{2\sigma^2}}$$

Уравнение 3-15.

которая равна

$$\psi(t) = \frac{1}{\sqrt{2\pi\sigma^3}} \left(e^{-\frac{t^2}{2\sigma^2}} \cdot \left(\frac{t^2}{\sigma^2} - 1 \right) \right)$$

Уравнение 3-16.

Вейвлет Морле определяется как

$$w(t) = e^{iat} e^{-\frac{t^2}{2\sigma}}$$

Уравнение 3-16а.

где **a** - параметр модуляции, а **sigma** параметр масштаба, влияющий на ширину окна.

ПРИМЕРЫ

Все сигналы в примерах, приведенных ниже, взяты из реальной жизни (медицина). Первый сигнал относится к нормальному человеку, второй – к больному. Так как эти сигналы достаточно сложны, интерпретация их НВП является непростой задачей. Они приведены здесь для того, чтобы показать какие трудности возникают перед исследователями.

На Рис.3-11 показана характеристика нормального человека.

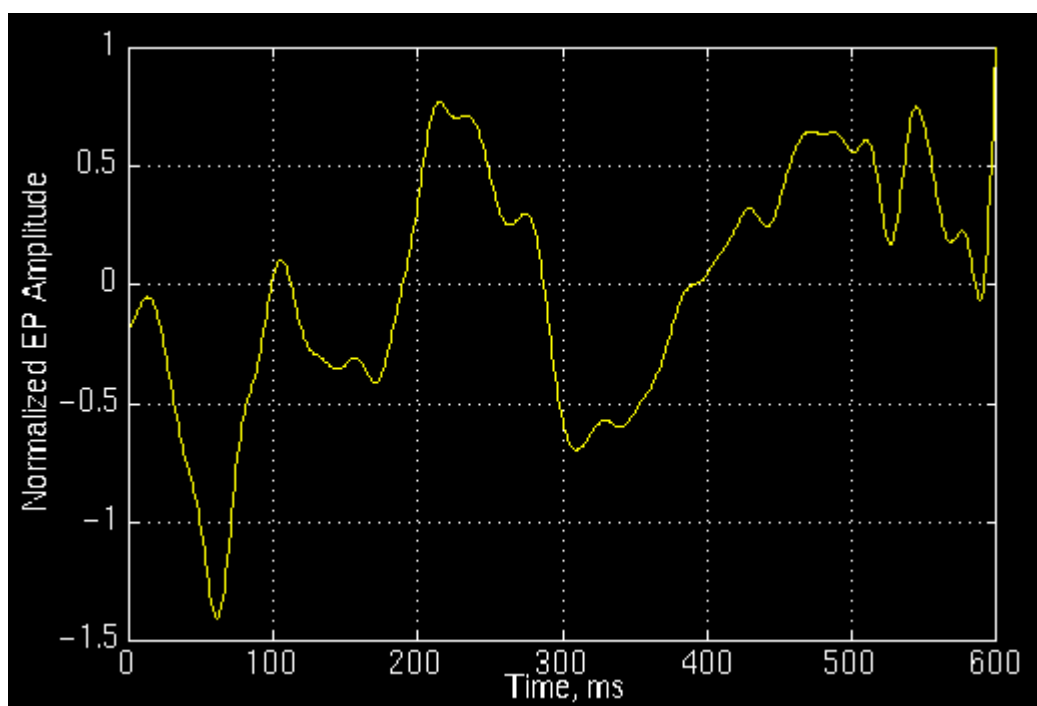


Рис.3-11.

Далее показано НВП этого сигнала. Не обращайте внимания на числа, стоящие на осях: они просто показывают, что НВП вычисляется для 350 сдвигов и 60 масштабов. Важно отметить, что нами вычислялось не истинное НВП, так как число сдвигов - конечно. Таким образом мы получили дискретизированную версию НВП, которая будет обсуждаться ниже. (Не путайте это преобразование с дискретным вейвлет-преобразованием, которое будет рассматриваться в четвертой части).

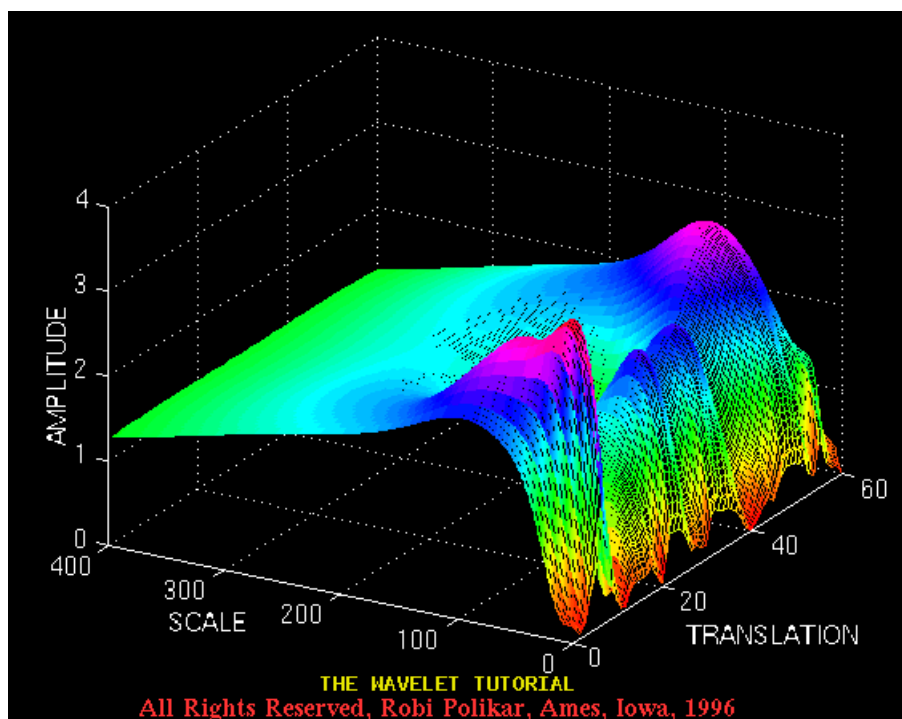


Рис.3-12.

На Рис.3-13 показано то же преобразование, но под несколько другим углом для лучшей визуализации.

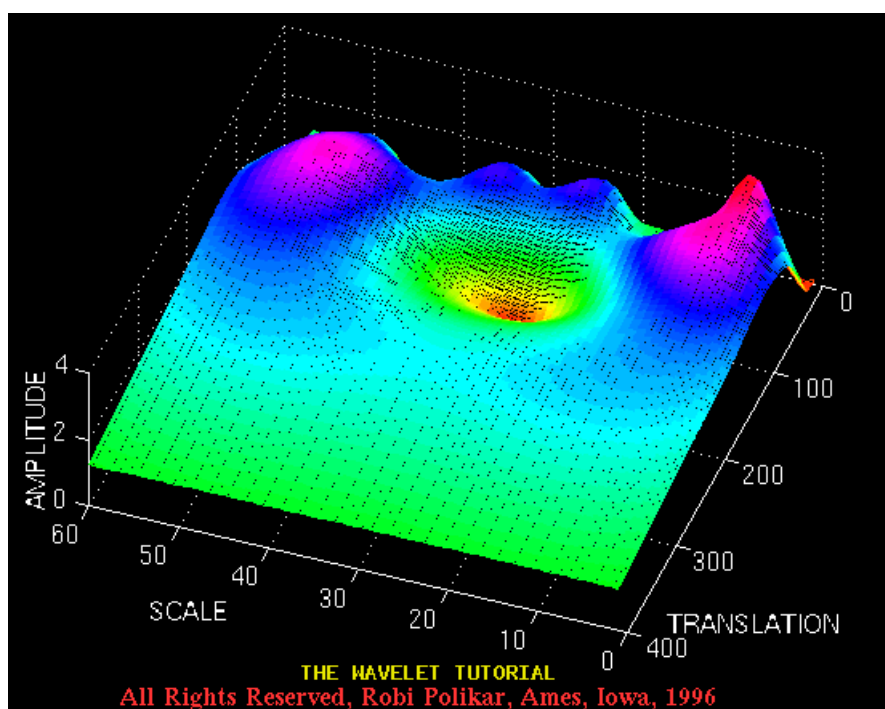


Рис.3-13.

На Рис.3-14 показана характеристика человека, больного болезнью Альцгеймера.

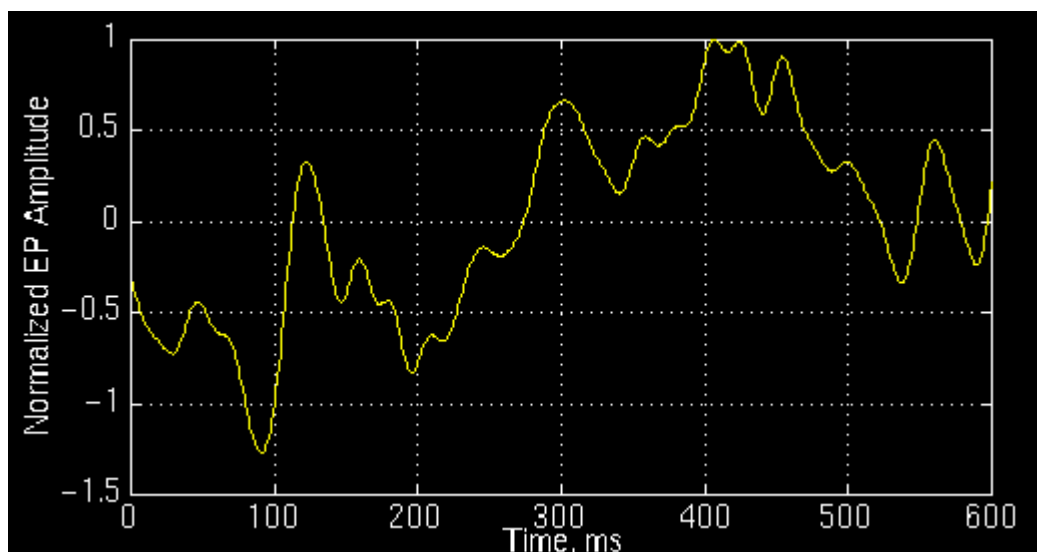


Рис.3-14.

На Рис.3-15 показано НВП этого сигнала:

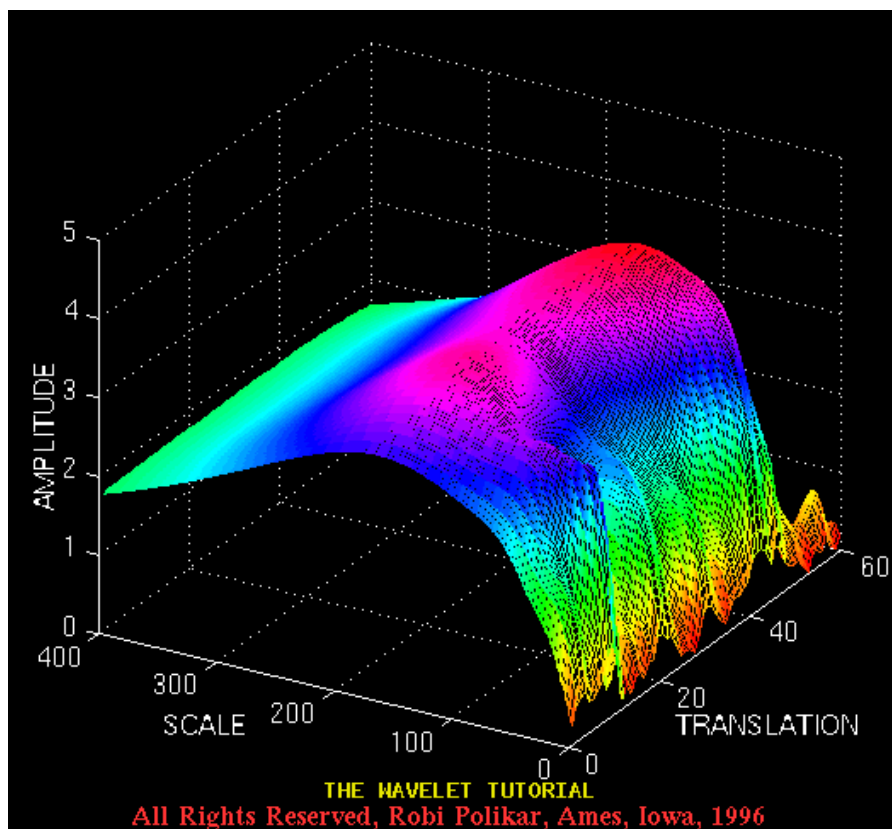


Рис.3-15.

и далее то же НВП под другим углом обзора:

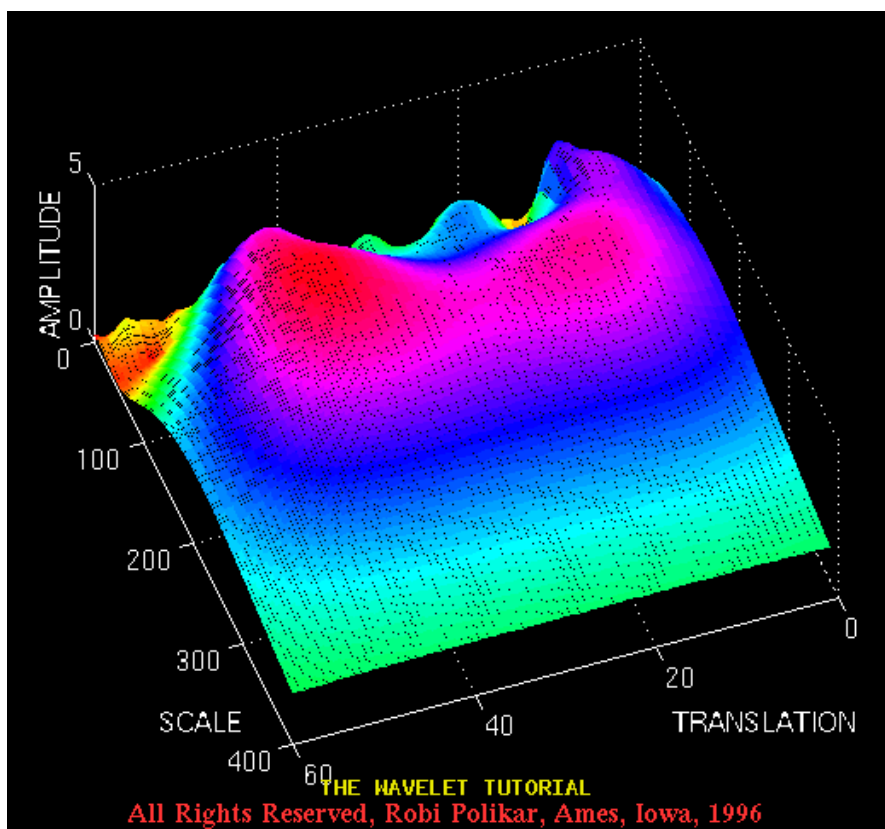


Рис.3-16.

ВЕЙВЛЕТ-СИНТЕЗ

Непрерывное вейвлет-преобразование является обратимым преобразованием, если выполняется равенство 3-18. По счастью, это требование не накладывает серьезных ограничений. НВП будет обратимым при выполнении 3-18, даже если базисные функции будут неортогональными.

Реконструкция выполняется по формуле:

$$x(t) = \frac{1}{C_{\psi}^2} \int_s \int_{\tau} \Psi_x^{\psi}(\tau, s) \frac{1}{s^2} \psi\left(\frac{t-\tau}{s}\right) d\tau ds$$

Уравнение 3-17.

где C_{ψ} - константа, зависящая от используемого вейвлета. Эта константа называется константой допустимости, и она должна удовлетворять условию допустимости:

©Вадим Грибунин, E-mail: wavelet@autex.spb.ru
 ©АВТЭКС Санкт-Петербург, <http://www.autex.spb.ru>, E-mail: info@autex.spb.ru

$$c_\psi = \left\{ 2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi \right\}^{1/2} < \infty$$

Уравнение 3-18.

где $\hat{\psi}(\xi)$ есть ПФ $\psi(t)$. Выражение 3-18 означает, что $\hat{\psi}(0) = 0$, или

$$\int \psi(t) dt = 0$$

Уравнение 3-19.

Как отмечалось выше, условие 3-19 не накладывает много ограничений, так как может быть найдено множество вейвлет-функций, удовлетворяющих 3-19. Из 3-19 видно, что вейвлет должен колебаться вокруг оси абсцисс.

ДИСКРЕТИЗАЦИЯ НЕПРЕРЫВНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ: РЯДЫ ВЕЙВЛЕТОВ

В настоящее время большинство вычислений выполняют компьютеры. Это означает, что ни ПФ, ни ОПФ, ни НВП не могут быть практически вычислены путем взятия интегралов и т.д. На практике преобразования дискретизируются, и интегралы заменяются суммами. Наиболее простым и естественным способом является равномерная дискретизация частотно-временной плоскости.

На высоких масштабах (нижних частотах) частота дискретизации может быть понижена, в соответствии с теоремой отсчетов. Другими словами, если требуется дискретизация масштабно-временной плоскости с частотой N_1 при масштабе s_1 , то при масштабе s_2 эта частота может быть понижена до N_2 , где $s_1 < s_2$ (т.е. $f_1 > f_2$) и $N_2 < N_1$. Взаимосвязь между N_1 и N_2 следующая:

$$N_2 = \frac{s_1}{s_2} N_1$$

Уравнение 3-20.

или

$$N_2 = \frac{f_1}{f_2} N_1$$

Уравнение 3-21.

Таким образом, на нижних частотах частота дискретизации может быть снижена, что уменьшает число вычислений.

Надо отметить, что если нас интересует только анализ сигнала (без последующего его синтеза), то дискретизация может выполняться с любой частотой, даже и не удовлетворяющей критерию Найквиста. Этот критерий существен, только если мы собираемся синтезировать непрерывный сигнал из его дискретных отсчетов. И здесь как раз важны базисные векторы, рассмотренные ранее.

Как мы уже знаем, вейвлет $\psi(\tau, s)$, удовлетворяющий выражению 3-18, позволяет реконструировать сигнал в соответствии с уравнением 3-17. Однако, это истинно только для непрерывного преобразования. Вопрос: можем ли мы восстановить сигнал при дискретизированных значениях времени и масштаба? ДА, но при некоторых дополнительных условиях.

Параметр масштаба s дискретизируется на логарифмической сетке. Параметр времени дискретизируется потом в соответствии с параметром масштаба, т.е. на разных масштабах - разная частота дискретизации. Другими словами, дискретизация выполняется на диадической сетке, показанной на Рис.3-17:

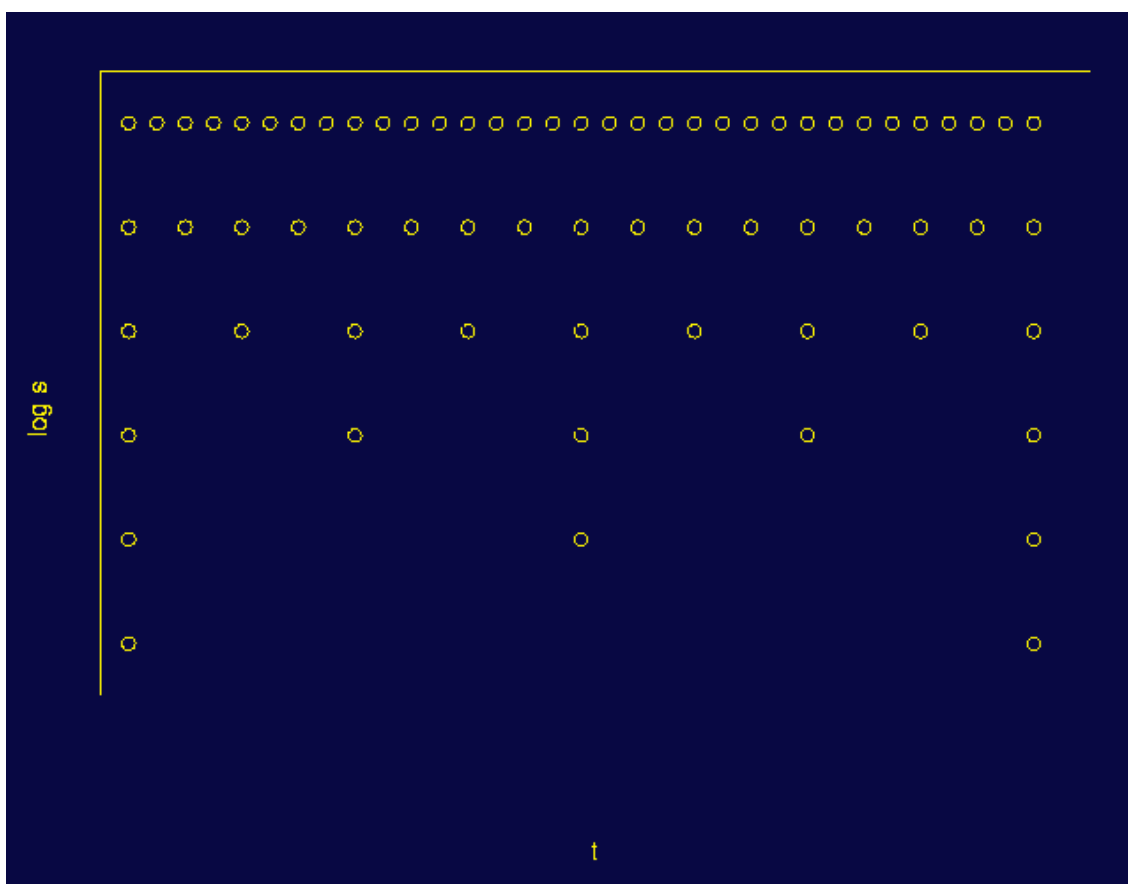


Рис.3-17.

Область, занимаемую графиком, можно рассматривать как всю масштабнo-временную плоскость. НВП сопоставляет значения с континуумом точек этой плоскости. Поэтому, число коэффициентов НВП бесконечно. Сначала рассмотрим дискретизацию вдоль оси масштаба. Берем точки в соответствии с логарифмическим масштабом, причем основание логарифма может быть произвольным. Обычно в качестве основания выбираю 2 и вычисляют токи при масштабах 2, 4, 8, 16, 32, 64,...и т.д. Затем соответствующим образом дискретизируется временная ось. При этом на каждом следующем масштабе частота дискретизации снижается вдвое.

Заметьте, что на самом низшем масштабе ($s=2$) выбираются только 32 точки временной оси (для частного случая Рис.3-17). При $s=4$ берется 16 точек и т.д.

Хотя обычно используется название масштабнo-временная плоскость, точнее было бы назвать ее масштабнo-сдвиговая плоскость, так как «время» в области преобразования в действительности соответствует сдвигам вейвлета по времени. Для вейвлет-рядов действительное время остается непрерывным.

Так же как существуют непрерывное преобразование Фурье, ряды Фурье и дискретное вейвлет-преобразование, имеется непрерывное вейвлет-преобразование, полудискретное вейвлет-преобразование (вейвлет-ряды) и дискретное вейвлет-преобразование.

Вышеприведенную дискретизацию масштаба можно записать как $s = s_0^j$, а дискретизацию сдвигов - $\tau = k.s_0^j.\tau_0$, где $s_0 > 1$ и $\tau_0 > 0$. Таким образом, дискретизация сдвигов зависит от дискретизации масштаба.

Непрерывная вейвлет-функция

$$\psi_{\tau,s} = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right)$$

Уравнение 3-22.

$$\psi_{j,k}(t) = s_0^{-j/2} \psi(s_0^{-j}t - k\tau_0)$$

Уравнение 3-23.

подставив $s = s_0^j$ и $\tau = k.s_0^j.\tau_0$.

Если $\{\psi_{j,k}\}$ образует ортонормальный базис, то ряды вейвлетов могут быть записаны в виде:

$$\Psi_x^{\psi_{j,k}} = \int x(t) \psi_{j,k}^*(t) dt$$

Уравнение 3-24.

или

$$x(t) = c_{\psi} \sum_j \sum_k \Psi_k^{\psi_{j,k}} \psi_{j,k}(t)$$

Уравнение 3-25.

Ряды вейвлетов требуют, чтобы $\{\psi_{j,k}\}$ были ортонормальными, биортонормальными или фреймом. Если $\{\psi_{j,k}\}$ не являются ортонормальными, то выражение 3-24 становится:

$$\Psi_x^{\psi_{j,k}} = \int x(t) \hat{\psi}_{j,k}^*(t) dt$$

Уравнение 3-26.

где $\{\psi_{j,k}^*(t)\}$ есть дуальный ортогональный базис, или фрейм. * Означает комплексное сопряжение. Если $\{\psi_{j,k}\}$ ортонормальны или биортонормальны, то преобразование будет безизбыточным, если они образуют фрейм - то избыточным (то есть число коэффициентов больше числа отсчетов исходного сигнала). Как уже отмечалось, фреймов гораздо «больше», чем ортогональных или биортонормальных базисов.

Для вейвлет-преобразования можно предложить следующую аналогию с человеческим зрением. При взгляде на объект глаз человека видит в первую очередь общие очертания, зависящие от расстояния, с которого смотрят. Это соответствует установке параметра масштаба s_0^{-j} . Когда приближаемся очень близко к объекту, j отрицательно и имеет большое значение (низкий масштаб, высокие частоты, анализ деталей в сигнале). Малые перемещения глаз (изменение угла обзора) соответствуют малым значениям $\tau = k.s_0^j.\tau_0$. При отрицательном и большом j это соответствует малым изменениям во времени, τ , (высокой частоте дискретизации) и большим изменениям в s_0^{-j} (низкий масштаб). Параметр масштаба может рассматриваться также как увеличение. Насколько мала может быть частота дискретизации при обеспечении реконструкции сигнала? Это главный вопрос, на который надо ответить для оптимизации всей процедуры. Для программирования удобнее всего было бы значение «2» для s_0 и «1» для τ . Конечно, при предельно низкой частоте дискретизации число доступных ортонормальных вейвлетов уменьшается.

Итак, примеры НВП, представленные в данной главе, есть ряды вейвлетов. Параметры выбираются в зависимости от сигнала. Так как реконструкция не требуется, частоты дискретизации были далеки от критических.

На этом мы заканчиваем часть III учебника. Я надеюсь, что вы получили общее представление о том, что такое вейвлет-преобразование. Нам осталось изучить еще один вид вейвлет-преобразования. Хотя дискретизированное вейвлет-преобразование может быть вычислено на компьютере, объем вычислений велик. Дискретное вейвлет-преобразование (ДВП), рассматриваемое в главе IV, имеет исключительно быстрый алгоритм вычисления.

ЧАСТЬ IV

ДЛЯ ЧЕГО НУЖНО ДИСКРЕТНОЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ?

Дискретизированная версия непрерывного вейвлет-преобразования, рассмотренная в части 3, требует большого числа вычислений. Кроме того, в результате получается избыточное количество коэффициентов, намного превосходящее число отсчетов исходного сигнала (если мы думаем о реконструкции). Дискретное вейвлет-преобразование (ДВП) обеспечивает достаточно информации как для анализа сигнала, так и для его синтеза, являясь вместе с тем экономным, как по числу операций, так и по требуемой памяти.

ДИСКРЕТНОЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ (ДВП)

Истоки ДВП восходят к 1976, когда Croiser, Esteban и Galand разработали метод декомпозиции дискретных сигналов. Crochiere, Weber и Flanagan в тот же год опубликовали аналогичную работу по кодированию речевых сигналов. Они назвали свой метод анализа **субполосным кодированием**. В 1983 Burt разработал похожий метод и назвал его **пирамидальным кодированием**. Позднее, в 1989, Vetterli и Le Gall разработали некоторые улучшения схемы субполосного кодирования, заключающиеся в устранении избыточности преобразования. Субполосное кодирование будет объяснено ниже.

СУБПОЛОСНОЕ КОДИРОВАНИЕ (СУБКОД) И КРАТНОМАСШТАБНЫЙ АНАЛИЗ (КМА)

Основная идея - та же, что и при НВП. Масштабно-временное представление сигнала получается с использованием методов цифровой фильтрации. Вспомним, что НВП измеряло корреляцию между вейвлетом на разных масштабах и сигналом. НВП вычислялось путем изменения масштаба окна анализа, сдвига окна во времени, умножения на сигнал, интегрирования по всей временной шкале. В дискретном случае для анализа сигнала на разных масштабах используются фильтры с различными частотами среза. Сигнал пропускается через древовидно соединенные ВЧ и НЧ фильтры.

Разрешение сигнала, являющееся мерой количества детальной информации в сигнале, изменяется за счет фильтрации сигнала, а масштаб изменяется за счет децимации и интерполяции. Децимация соответствует снижению частоты дискретизации, или удалению некоторых отсчетов из сигнала. Например, децимация в два раза означает, что из сигнала удаляется каждый второй отсчет.

Интерполяция соответствует увеличению частоты дискретизации сигнала путем добавлению новых отсчетов между существующими. Обычно в качестве новых отсчетов используется нуль. Например, интерполяция в два раза означает вставку нулей между каждыми отсчетами.

Коэффициенты ДВП обычно получаются из коэффициентов НВП путем дискретизации последних на диадической сетке $s_0 = 2$ and $t_0 = 1$, хотя это и не является обязательным. То есть $s=2^j$ и $t = k*2^j$, как описано в третьей части. Так как сигнал является функцией дискретного времени, то термины функция и последовательность применяются в дальнейшем взаимозаменяемо. Последовательность обозначается как $x[n]$, где n - целое.

Процедура начинается с пропускания сигнала (последовательности) через полуполосный цифровой низкочастотный фильтр с импульсной характеристикой $h[n]$. Фильтрация сигнала соответствует математической операции свертки сигнала и импульсной характеристики фильтра. Операция свертки для дискретного времени определяется как:

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

Уравнение 4-1.

Полуполосный НЧ фильтр обрезает все частоты, большие половины верхней частоты сигнала. Например, если верхняя частота=1000Гц, то все частоты, большие 500Гц будут удалены из сигнала.

Разберемся с единицей измерения частоты. Для дискретных сигналов частота обычно выражается в радианах. Тогда частота дискретизации равна 2π радиан. Поэтому, наибольшая частота, которая будет присутствовать в сигнале, равна π радиан, если сигнал был дискретизирован в соответствии с теоремой отсчетов. Тем не менее, мы иногда будем использовать и герцы.

После пропускания сигнала через НЧ полуполосный фильтр половина отсчетов может быть удалена из сигнала в соответствии с теоремой отсчетов, так как сигнал теперь имеет наивысшую частоту $\pi/2$ радиан, а не π . Удалим каждый второй отсчет, т.е. выполним децимацию в два раза. Масштаб сигнала получился удвоенным. НЧ-фильтрация удаляет ВЧ компоненты, но оставляет масштаб неизменным. Масштаб изменяет операция децимации. С другой стороны, разрешение связано с количеством информации, содержащейся в сигнале, поэтому оно меняется при фильтрации (после полуполосной - уменьшается в два раза). Заметим, что последующая децимация не изменяет разрешения, так как удаляемые отсчеты все равно являются избыточными и не несут информации.

На языке формул эта процедура запишется как:

$$y[n] = \sum_{k=-\infty}^{\infty} h[k] * x[2n - k]$$

Уравнение 4-2.

Теперь мы понимаем, как вычисляется ДВП: оно анализирует сигнал в различных частотных полосах с различным разрешением путем декомпозиции его на грубую аппроксимацию и детали. В ДВП работают два множества функций - масштабирующие функции и вейвлеты, соответствующие НЧ и ВЧ фильтрам. Один уровень ДВП можно записать следующим образом (сигнал пропускается через НЧ и ВЧ фильтры и затем децимируется в два раза):

$$y_{high}[k] = \sum_n x[n] \cdot g[2k - n]$$

$$y_{low}[k] = \sum_n x[n] \cdot h[2k - n]$$

Уравнение 4-3.

где $y_{high}[k]$ и $y_{low}[k]$ есть прореженные в два раза выходы ВЧ и НЧ фильтров, соответственно.

В результате первого шага ДВП временное разрешение уменьшается в два раза, так как лишь половина отсчетов характеризует весь сигнал (рассматриваем выход НЧ-фильтра). Однако частотное разрешение удваивается, так как сигнал занимает теперь половинную полосу частот и неопределенность уменьшается. Вышеприведенная процедура, известная как субполосное кодирование, повторяется далее. Выход НЧ фильтра подается на такую же схему обработки, а выход ВЧ фильтра считается вейвлет-коэффициентами. Этот процесс показан на Рис.4-1, где $x[n]$ -исходный сигнал, $h[n]$ и $g[n]$ - НЧ и ВЧ фильтры, соответственно. Полоса сигнала на каждом уровне обозначена здесь как «f».

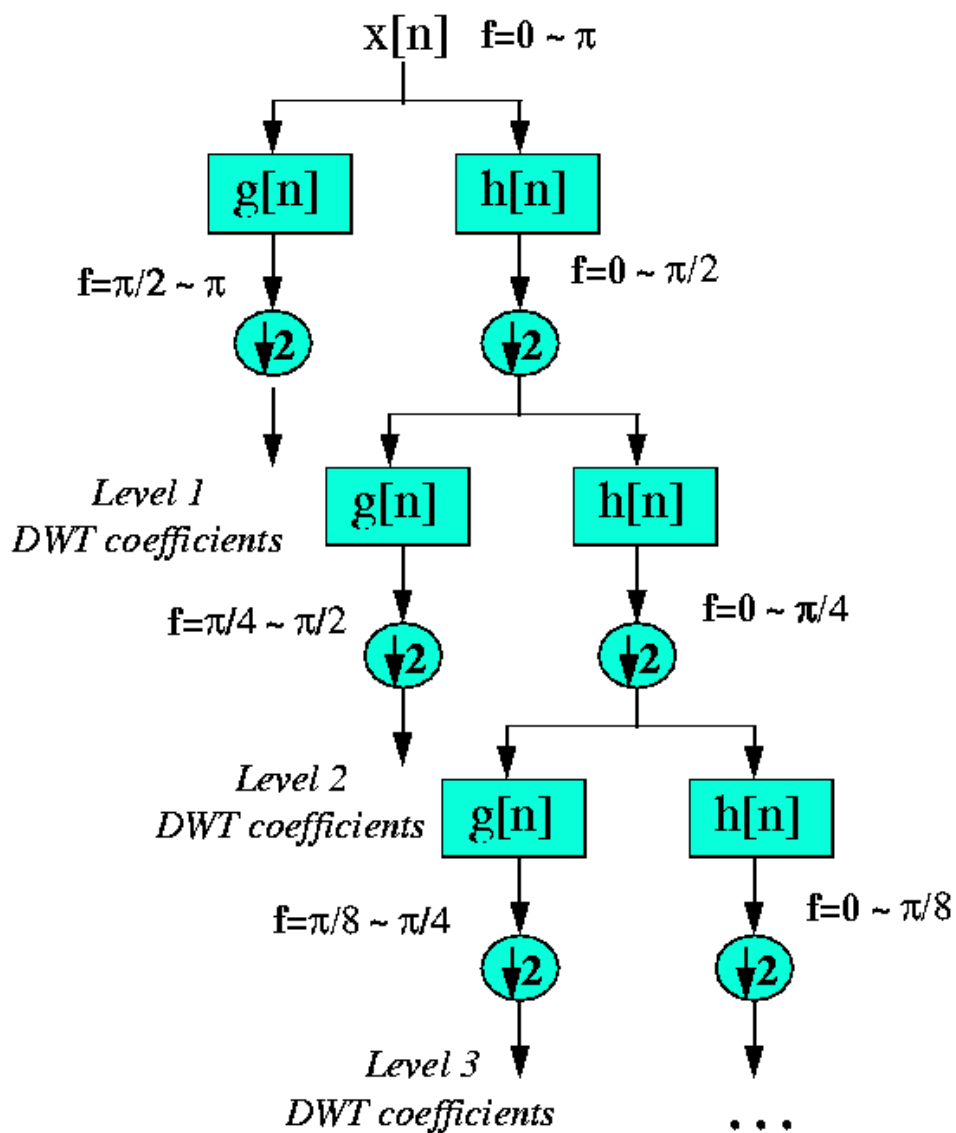


Рис.4-1. Алгоритм субполосного кодирования.

В качестве примера представьте, что сигнал $x[n]$ имеет 512 отсчетов и частотную полосу от 0 до π рад/с. После первого уровня декомпозиции на выходе ВЧ фильтра-дециматора имеется 256 отсчетов. Разрешение по времени уменьшилось в два раза, но частотный диапазон, занимаемой сигналом, теперь от $\pi/2$ до π рад/с (т.е. разрешение по частоте удвоилось). Эти 256 отсчетов есть вейвлет-коэффициенты первого уровня. Выход НЧ-фильтра занимает полосу частот от 0 до $\pi/2$ рад/с. Этот сигнал поступает опять на два фильтра и т.д.

Этот процесс может продолжаться до тех пор, пока не останутся 2 отсчета, т.е. для данной длины сигнала может быть 8 уровней декомпозиции. ДВП исходного сигнала получается затем путем объединения вейвлет-коэффициентов всех уровней, начиная с последнего (2 коэффициента). К ним присоединяются спереди отсчеты аппроксимации сигнала последнего уровня (2 отсчета). Таким образом, число коэффициентов преобразования равно числу отсчетов в исходном сигнале.

Наиболее значимые частоты исходного сигнала будут отображаться как большие амплитуды вейвлет-коэффициентов, «отвечающих» за соответствующий частотный диапазон. Отличие ДВП от преобразования Фурье заключается в том, что время появления частот в данном случае не утеряно. Однако временная локализация будет иметь разрешение, зависящее от уровня преобразования, на котором появляется частота. Если основная информация сигнала лежит в области высоких частот, как это часто случается, то временная локализация этих частот будет более точной, так как они характеризуются большим количеством отсчетов. На низких частотах, напротив, отсчетов мало. Поэтому, временное разрешение будет плохим, зато хорошим будет частотное разрешение (так как сигнал занимает узкую полосу частот). Такая схема анализа подходит для большинства сигналов, встречающихся на практике.

Малые значения вейвлет-коэффициентов означают низкую энергетическую плотность соответствующих частотных полос в сигнале. Эти коэффициенты могут быть отброшены без существенного искажения сигнала. Таким образом достигается уменьшение данных. На Рис.4-2 показан пример того, на что похоже ДВП и как получается уменьшение данных. На Рис.4-2а показан типичный сигнал, состоящий из 512 отсчетов и имеющий нормализованную амплитуду. На Рис.4-2б показано 8-уровневое ДВП этого сигнала. Последние 256 отсчетов этого графика соответствуют высшим частотам, имеющимся в сигнале и т.д. Отметим, что только первые 64 отсчета являются значимыми, т.е. несут важную (энергетически) информацию о сигнале. Поэтому, можно оставить только эти 64 отсчета, а остальные отбросить, получив, таким образом, сжатие данных.

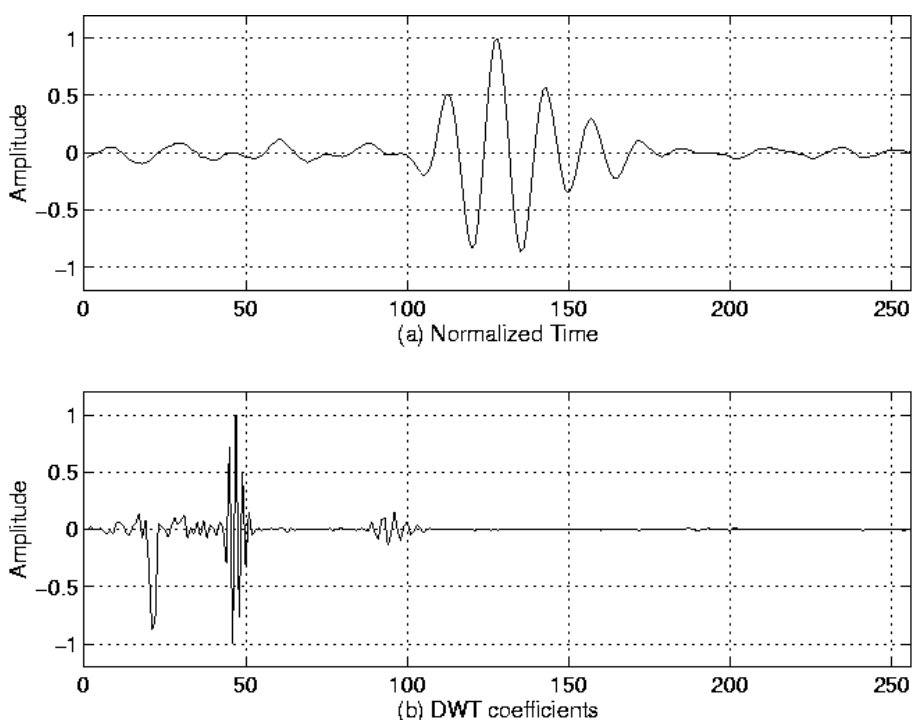


Рис.4-2. Пример ДВП.

Мы еще вернемся к этому примеру для лучшего понимания интерпретации ДВП. Пока что, однако, закончим математический анализ ДВП. Важным свойством ДВП является взаимосвязь между импульсными характеристиками ВЧ и НЧ фильтров. Эти фильтры ДВП связаны между собой отношением:

$$g[L-1-n] = (-1)^n \cdot h[n]$$

Уравнение 4-4.

где $g[n]$ есть ВЧ, $h[n]$ – НЧ фильтр, а L – длина фильтра (число точек). ВЧ фильтр получается из НЧ путем «переворота» вектора его коэффициентов и изменения знака у нечетных коэффициентов. Фильтры, удовлетворяющие этому условию, использовались в цифровой обработке сигналов задолго до вейвлетов и получили название квадратурно-зеркальных фильтров (КЗФ). Операции фильтрации и прореживания могут быть записаны как:

$$y_{high}[k] = \sum_n x[n] \cdot g[-n + 2k]$$

$$y_{low}[k] = \sum_n x[n] \cdot h[-n + 2k]$$

Уравнение 4-5.

Реконструкция в данном случае тривиальна, так как полуполосные фильтры образуют ортонормальный базис. Синтез выполняется в обратном порядке по сравнению с анализом. Сигналы на каждом уровне интерполируются в 2 раза, пропускаются через фильтры синтеза $g'[n]$ и $h'[n]$ (ВЧ и НЧ, соответственно) и затем складываются. Интересно отметить, что фильтры анализа и синтеза идентичны, за исключением обратного порядка следования коэффициентов. Поэтому формула реконструкции для каждого уровня ДВП может быть записана в виде:

$$x[n] = \sum_{k=-\infty}^{\infty} (y_{high}[k] \cdot g[-n + 2k]) + (y_{low}[k] \cdot h[-n + 2k])$$

Уравнение 4-6.

Однако если фильтры не являются идеальными полуполосными, то полное восстановление может быть и не достигнуто. Идеальные фильтры, конечно же, нельзя сконструировать на практике. Все же налагая дополнительные ограничения на импульсные характеристики фильтров, можно добиться полного восстановления при использовании «реальных» фильтров. Наиболее известными являются, наверное, фильтры Добеши, приводящие к вейвлетам Добеши.

Интерпретация коэффициентов ДВП несколько сложнее, чем коэффициентов Фурье. Рассмотрим пример:

Пусть анализируемый сигнал дискретизирован на частоте 10МГц и состоит из 256 отсчетов. Значит, верхняя частота сигнала 5 МГц. Тогда коэффициенты первого уровня разложения (128) занимают полосу частот [2.5 5] МГц. Эти коэффициенты отображаются в конце графика вейвлет-коэффициентов. Вейвлет-коэффициенты второго уровня (64) «отвечают» за полосу частот [1.25 2.5] МГц. Они отображаются перед вейвлет-коэффициентами первого уровня.

Процедура повторяется до тех пор, пока не останется 1 вейвлет-коэффициент и 1 отсчет аппроксимации на 9 уровне. Всего получается $(1+1+2+4+8+16+32+64+128) = 256$ коэффициентов. Т.е. число коэффициентов равно числу отсчетов в исходном сигнале. Если основная энергия сигнала была сосредоточена возле частоты 2 МГц, то вейвлет-коэффициенты второго уровня будут большими, а вейвлет-коэффициентами первого уровня можно пренебречь. Кажется странным, что вейвлет-коэффициенты, отвечающие за НЧ область могут также иметь большую величину. Это связано с тем, что они описывают огибающую сигнала.

На этом закончим наш мини-курс по вейвлетам. Я хотел бы напомнить вам, что этот учебник носит обзорный характер, не является полным и не охватывает всю область вейвлетов. Он предназначен для первоначального знакомства с темой. Возможно, здесь имеется много ошибок, и я буду рад, если вы сообщите мне о них. Спасибо за ваш интерес к учебнику.