

А.А.Коростелев

ПРОСТРАНСТВЕННО- ВРЕМЕННАЯ ТЕОРИЯ РАДИОСИСТЕМ

Учебное
пособие
для высших
учебных
заведений

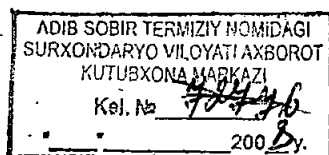
«Радио и связь»

32
А.А.Коростелев К68

ПРОСТРАНСТВЕННО- ВРЕМЕННАЯ ТЕОРИЯ РАДИОСИСТЕМ

Допущено Министерством высшего
и среднего специального образования СССР
в качестве учебного пособия
для студентов радиотехнических
специальностей вузов

398286



Москва
«Радио и СВЯЗЬ»
1987

ББК 32.95
К68
УДК 621.396.96

Рецензенты: Московский авиационный институт им. С. Орджоникидзе (д-р техн. наук, проф. П. А. Бакулев), д-р техн. наук, проф. В. Ф. Судаков

Редакция литературы по кибернетике и вычислительной технике

Коростелев А. А.

К 68 Пространственно-временная теория радиосистем:
Учеб. пособие для вузов. — М.: Радио и связь, 1987. —
320 с.: ил.

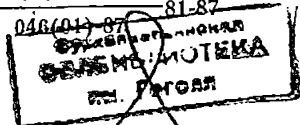
Излагается теория оптимального приема электромагнитной волны измерительной системой, свойства которой заданы только формой сигнала и апертурой антенны. Определяются потенциальные возможности радиосистемы по дальности действия, разрешающей способности и точности измерения всех координат наблюдаемого объекта и их производных при заданных характеристиках помех. Изложение ведется применительно к произвольному измеряемому параметру сигнала на основе обобщенной комплексной огибающей или ее спектра. Описывается структура оптимальной радиосистемы, включая антенну и выходные устройства.

Для студентов радиотехнических специальностей вузов. Может быть полезна специалистам в области радиолокации, радионавигации и радиоуправления.

2402020000-044

81-87

ББК 32.95



© Издательство «Радио и связь», 1987

ПРЕДИСЛОВИЕ

В настоящее время в учебниках и учебных пособиях, в которых излагается теория оптимальных методов радиоприема, сигнал рассматривается только как функция времени, что в определенной степени удовлетворяет специалистов в области радиосвязи, радиотелеметрии и передачи команд. Однако в системах радиолокации, радионавигации и радиоуправления многие измеряемые величины являются параметрами волны (поля), т. е. пространственно-временного колебания, а не просто электрического сигнала. Соответственно само оптимальное приемное устройство и систему в целом необходимо рассматривать как пространственно-временную структуру.

Развиваемый в книге пространственно-временной подход позволяет определить потенциальные возможности и оптимальную структуру радиосистемы, рассмотреть с единых позиций статистической теории оптимального приема все измеряемые параметры. Новый подход стал необходим в связи со все более широким использованием фазированных антенных решеток (ФАР), широкополосных и сверхширокополосных сигналов. При описании работы оптических процессоров, которые широко применяются в современных радиосистемах, пространственно-временной подход является также единственно возможным.

Расширенное представление о сигнале и самой системе потребовало изменения и методики изложения. При большом числе измеряемых параметров (а их обычно шесть: дальность, радиальная скорость, два угла и две угловые скорости) нерационально вести рассмотрение применительно к каждому конкретному параметру, как это обычно делается. Поэтому в книге сначала рассматривается произвольный параметр или совокупность любых параметров волны, для чего вводится понятие обобщенной комплексной огибающей и ее спектра. Затем простой подстановкой конкретных функций в общие формулы получаются результаты для каждого определенного параметра. Та же методика используется и при описании структуры оптимальных измерителей. Это дает возможность установить общность и отличительные

особенности процессов и устройств при измерении каждого конкретного параметра.

Развиваемая в книге теория позволяет найти оптимальные технические характеристики системы, временную (сигнал) и пространственную (апертура антенны), и определить при заданных технических характеристиках оптимальную структуру приемного тракта и потенциальные возможности системы (энергетические и метрические) в присутствии помех с заданными статистическими характеристиками. Оперативность извлечения информации и способы обзора пространства, которые в значительной степени зависят от назначения и конкретного вида системы, не могут быть охвачены общей теорией и поэтому не рассматриваются.

Изложение ведется на основе скалярной теории поля, учитывающей все характеристики волны, кроме поляризационных. Поляризация практически не влияет на процесс измерения указанных выше параметров, однако для общности изложения показаны пути построения теории оптимального приема с учетом поляризационных характеристик сигнала.

Практически впервые в учебной литературе излагается оптимальная фильтрация (линейная и нелинейная) случайных процессов, являющаяся теоретической базой для синтеза следящих измерителей. Кроме известных фильтров Винера и Калмана рассматривается фильтр, оптимальный по вероятностному критерию, как наиболее адекватный процессам в следящих измерителях. Приводится краткое описание общей структуры и элементов следящих и разомкнутых измерителей.

Изложение материала основано на оригинальных разработках автора и опыте чтения лекций в ряде вузов страны как самим автором, так и его последователями. При написании § 6.7, 7.7 и 7.8 были использованы материалы, любезно предоставленные Е. Э. Чернышевым, М. Г. Степановым и Б. Г. Мельниковым.

Излагаемый в книге материал затрагивает самые различные области радиотехники, оптики, теории автоматического регулирования, различающиеся методикой подхода, терминологией и системой обозначений. Это вызвало серьезные трудности при изложении материала, которые автор, по-видимому, не смог преодолеть до конца. Поэтому все пожелания и предложения по книге будут приняты с благодарностью.

Глава 1. ОБЩИЕ СВЕДЕНИЯ ОБ ОБЗОРНО-ИЗМЕРИТЕЛЬНЫХ СИСТЕМАХ

1.1. НАЗНАЧЕНИЕ ОБЗОРНО-ИЗМЕРИТЕЛЬНЫХ СИСТЕМ

Обзорно-измерительные системы служат для извлечений информации о положении и движении различных объектов в пространстве, осматриваемом с помощью электромагнитных волн различных диапазонов. Эти волны либо излучаются объектами самопроизвольно (пассивный метод) или специальными излучателями (беззапросный метод), либо образуются в результате отражения от объектов зондирующих колебаний, излучаемых самой станцией (активный метод) или со стороны (полуактивный метод), либо, наконец, переизлучаются установленным на объекте ответчиком с приходом запросных сигналов от станции (активный ответ). Принимаемая станцией волна содержит определенную информацию о наличии объектов в зоне обзора, их относительном положении и скорости, а также о других свойствах объектов. Эта информация, отображаемая в параметрах принимаемой волны, и извлекается системой. К обзорно-измерительным системам относят системы радиолокации, радионавигации, оптические (лазерные) локации, измерительные каналы систем радиоуправления и т. д. С помощью этих систем можно обнаруживать объекты, измерять их координаты и параметры движения на огромных расстояниях, благодаря чему стало возможным управление этими или другими объектами почти повсеместно, включая космос в пределах и за пределами Солнечной системы.

Не все среды прозрачны для электромагнитных волн. В жидких и твердых средах эти волны очень сильно затухают. Поэтому в гидролокации и устройствах дефектоскопии различных изделий или структур для измерительных целей используют ультразвуковые волны, законы распространения которых в пространстве те же, что и для электромагнитных волн. Основное отличие состоит в том, что объемные звуковые волны имеют продольную структуру сжатия-разрежения среды (скалярную), а электромагнитные волны имеют вид поперечных колебаний вектора поля, определенным образом ориентированного в пространстве (поляризованные волны). Поляризация волны несет дополнительную информацию об объекте, отсутствующую у неполяризованных

звуковых волн. Однако эта информация относится в основном к тонкой структуре и конфигурации объекта и очень слабо связана с его положением и движением в пространстве. Поэтому в обзорно-измерительных радиосистемах ограничиваются, за редким исключением, скалярным описанием процессов, которое является исчерпывающим для систем звуколокации.

Первоначально в обзорно-измерительных радиосистемах использовались электромагнитные волны только радиодиапазона, преимущественно метровые и дециметровые. Затем диапазон постепенно расширялся, главным образом в сторону более коротких волн. Имеется принципиальная возможность использовать почти весь диапазон электромагнитных волн, включая инфракрасные, видимые и ультрафиолетовые, а также гамма- и рентгеновское излучения. Переход в сторону все более коротких волн вызван тем, что при одинаковых габаритах антенн достигается более высокая точность измерений и информативность принимаемых сигналов. При этом, к сожалению, возрастает отрицательное влияние условий распространения, так как с частотой усиливается затухание волн в атмосфере.

1.2. КЛАССИФИКАЦИЯ РАДИОСИСТЕМ

Обзорно-измерительные радиосистемы предназначены в основном для определения относительных координат и относительных скоростей объекта и станции. Поэтому безразлично, где производят измерения: на борту самого объекта, положение которого определяется, или вне его. В первом случае имеем системы *самоопределения*, во втором — системы *иноопределения*. Наряду с этими признаками классификации радиосистем могут быть: диапазон волн (радиоволны всех диапазонов, инфракрасное, оптическое, гамма- и рентгеновское излучения); назначение системы (управление, наблюдение, исследование и т. п.); место установки (наземные, корабельные, самолетные и т. д.); для наземных систем дополнительно — мобильность (стационарные, подвижные) и размещение аппаратуры в пространстве (одно- и многопунктные, с совмещенными и разнесенными передатчиком и приемником). Наиболее существенными признаками классификации радиосистем, определяющими их структуру и возможности, являются источники сигналов и вид измеряемых координат.

При классификации по источникам сигналов необходимо учитывать, что радиосистемы измерения относительных координат и скоростей включают разнесенные в пространстве станцию и объект наблюдения. Как станция, так и объект наблюдения могут содержать или не содержать генераторы электромагнитных колебаний носителей информации, специально предназначенные для работы системы. Иными словами, станция и объект могут быть каждый в отдельности *активными* и *пассивными*. В соответствии с этим радиосистемы подразделяют на четыре основ-

ные группы: радиолокационные, системы с активным ответом, беззапросные и системы пассивной радиолокации.

В *радиолокационных системах* используется явление радиоэха — радиолокационный сигнал образуется в результате отражения объектом зондирующих колебаний, излучаемых передатчиком самой радиолокационной станции (активный метод) либо сторонним облучателем (полуактивный метод). Для того чтобы объект был обнаружен, отражающие свойства его и окружающей среды должны отличаться, т. е. объект должен обладать радиолокационным контрастом. Радиолокационные системы обладают наиболее широкими возможностями, они могут работать как по известным, так и по неизвестным объектам. Однако их применение связано с большими энергетическими затратами, так как отраженный сигнал составляет незначительную долю излучаемого. Кроме того, обмен информацией между станцией и объектом весьма ограничен, так как объект может лишь в небольших пределах преднамеренно изменять свои отражающие свойства.

Системы с активным ответом (запросные) принимают переизлученные (в ответ на запрос станции) сигналы специального ответчика, установленного на объекте. Системы этого типа используют тогда, когда наблюдение ведется только за определенными взаимодействующими объектами. Благодаря тому, что сигнал переизлучается, а не отражается, эти системы с точки зрения энергетических затрат значительно выгоднее радиолокационных, кроме того, позволяют по каналам запроса и ответа дополнительно передавать различного рода информацию.

В обеих системах передатчик, генерирующий зондирующие или запросные колебания, размещается либо рядом с приемником, либо на значительном удалении от него. В соответствии с этим указанные системы дополнительно подразделяют на *совмещенные* и *разнесенные*. У первых передатчик и приемник могут иметь общую антенну, у вторых один передатчик может обслуживать несколько разнесенных приемников.

Системы, у которых на подвижном объекте установлен передатчик, а сама станция является приемно-измерительным устройством без передатчика, называют *беззапросными*. Они потребляют мало энергии и позволяют передавать дополнительную информацию, но только по каналу объект — станция. К сожалению, в простейших беззапросных системах не все параметры сигнала можно эффективно измерить. Для точного измерения скорости и особенно дальности и станция, и объект должны иметь высокостабильные эталонные генераторы.

В *системах пассивной радиолокации* в качестве сигнала может использоваться самопроизвольное излучение объектов: тепловое в коротковолновой части радиодиапазона и длинноволновое. К ним же относят системы радионавигации, использующие радионизлучение небесных светил. Системы пассивной радиолокации содержат лишь приемно-измерительное устройство и характеризуются

минимальными энергетическими затратами и максимальной помехозащищенностью. Но для выделения чрезвычайно слабых тепловых сигналов на фоне более сильных шумов требуется значительное время приема. Поэтому системы теплорасположения очень инерционны, их использование при большой скорости взаимного перемещения станции и объекта затруднительно, а дальность действия невелика. Кроме того, они обычно не позволяют измерять дальность и скорость, а также передавать дополнительную информацию. Длинноволновые системы имеют большую дальность действия и позволяют, например, обнаруживать большие области ионизации из любой точки земного шара.

Положение объекта в пространстве можно характеризовать различными геометрическими величинами: расстоянием до объекта, направлением на него, разностью и суммой расстояний до двух заданных точек пространства. Поэтому важным признаком классификации обзорно-измерительных систем является вид измеряемых геометрических величин (координат) или, что то же самое, форма линий положения на плоскости или поверхностей положения в пространстве. *Линией* или *поверхностью положения* называют геометрическое место точек, в которых измеряемая геометрическая величина остается неизменной. Так, на плоскости постоянному углу соответствует радиус относительно станции, постоянной дальности — окружность, постоянной разности расстояний — гипербола, постоянной сумме — эллипс (рис. 1.1); в пространстве соответственно имеем коническую поверхность, сферу, гиперболоид и эллипсоид.

Каждой измеряемой величине соответствует свой измерительный канал станции. По виду измеряемой величины каналы подразделяют на угломерные, дальномерные, разностно-дальномерные и суммарно-дальномерные. Сами системы классифицируют по совокупности измерительных каналов, как однородных, например угломерные, дальномерные, так и разнородных, например угломерно-дальномерные системы, измеряющие дальность и два угла. При этом положение объекта O на плоскости определяется точкой пересечения двух линий положения (рис. 1.2), в пространстве — трех поверхностей положения.

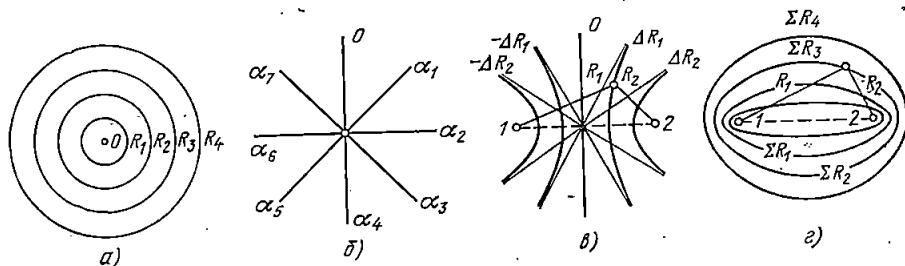


Рис. 1.1 Линии положения на плоскости:

а — круговые (дальномерных систем); б — радиальные (угломерных систем); в — гиперболические (разностно-дальномерных систем); г — эллиптические (суммарно-дальномерных систем)

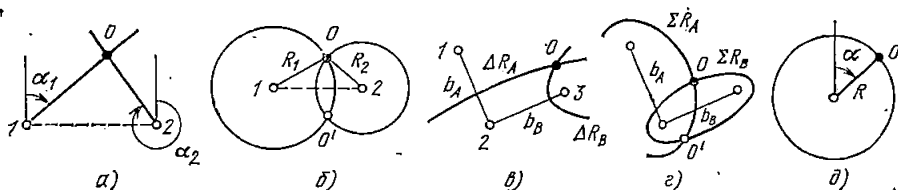


Рис. 1.2. Определение положения объекта на плоскости в угломерной (а), дальномерной (б), разностно-дальномерной (в), суммарно-дальномерной (г), смешанной угломерно-дальномерной (д) системах

Расстояние до объекта и направление на него (один угол на плоскости и два угла в пространстве) можно измерить из одной точки, разность и сумму расстояний — из двух точек пространства. Поэтому вид и число измеряемых геометрических величин определяют пространственную структуру системы: однопунктная и многопунктная. Так, однопунктными могут быть только смешанные системы — угломерно-дальномерные. Измерение всех трех одинаковых величин (трех дальностей, трех углов) возможно только в многопунктных системах.

Измерительные системы, классифицируемые по виду измеряемых величин (координат), существенно различаются по тактическим характеристикам, составу и объему аппаратуры и находятся в определенной связи с классификацией по наличию источников излучения. Так, радиолокационные системы и системы с активным ответом чаще всего являются угломерно-дальномерными или дальномерными, безапросные системы — угломерными или разностно-дальномерными, системы пассивной радиолокации — только угломерными.

Угломерно-дальномерные измерительные системы удобны тем, что положение объекта определяется из одной точки (рис. 1.2, д). Поэтому системы, работающие по отраженному и ответному сигналам, являются преимущественно угломерно-дальномерными. Они компактны, удобны при транспортировке и обслуживании, однако для их работы необходимы большие остронаправленные антенны, гарантирующие высокую точность угловых измерений. Для работы дальномерных и разностно-дальномерных станций не требуются остронаправленные антенны. Недостатком их является размещение аппаратуры в нескольких пунктах при определении всех координат объекта, а также неоднозначность (точки O и O' на рис. 1.2, б, г), для устранения которой необходимы специальные меры. Разностно-дальномерные системы удобны для определения собственного положения объекта относительно наземных маяков с помощью простого приемоиндикатора. Абсолютные расстояния до этих маяков измерить трудно, а разность расстояний — сравнительно просто. Суммарно-дальномерные измерения применяют в разнесенных системах, где сигнал проходит двойное расстояние передатчик — объект и объект — приемник.

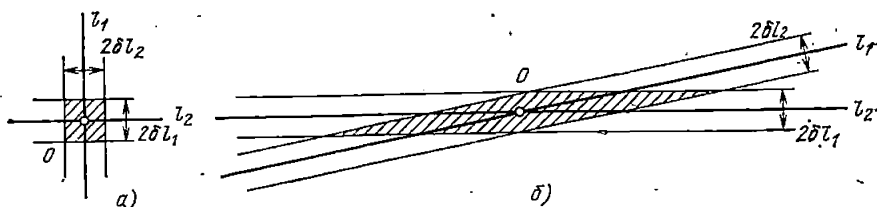


Рис. 1.3. Ошибки определения объекта на плоскости при пересечении линий положения под прямым (а) и острым (б) углами

Одним из наиболее важных показателей обзорно-измерительной системы является точность определения местоположения объекта. Эта точность определяется как линейными погрешностями δl измерения координаты l , так и взаимной ориентацией пересекающихся линий положения $l = \text{const}$ в районе расположения объекта O (рис. 1.3). Наивысшая точность определения местоположения объекта (заштрихованная на рисунке площадь) достигается, если линии положения пересекаются под прямым, а не под острым углом (ср. рис. 1.3, а и 1.3, б) при той же точности измерения каждой из координат. Системы с однородными измеряемыми координатами могут обеспечить высокую точность лишь в ограниченной области пространства, где углы пересечения близки к прямому. В дальномерно-угломерной системе все линии положения пересекаются под прямым углом, поэтому ее дополнительным преимуществом является высокая точность местоопределения, хотя эта точность снижается с дальностью ввиду возрастания линейной ошибки при заданной погрешности угловых измерений.

1.3. ПРИНЦИПЫ ИЗМЕРЕНИЯ ОТНОСИТЕЛЬНЫХ ПАРАМЕТРОВ — КООРДИНАТ И ИХ ПЕРВЫХ ПРОИЗВОДНЫХ

В основу измерения относительных координат объектов в радиосистемах положено постоянство вектора скорости электромагнитных волн, распространяющихся в однородной среде, т. е. постоянство скорости $c \approx 10^8$ м/с и прямолинейность распространения волн.

Измерение дальности R производится из одной точки пространства (рис. 1.4, а). В беззапросных системах время запаздывания сигнала прямого излучения

$$\tau = R/c. \quad (1.1)$$

Для его отсчета необходимо знать момент излучения сигнала объектом. Ввиду этого в пунктах излучения и приема нужно располагать высокостабильными эталонами времени — генераторами ко-

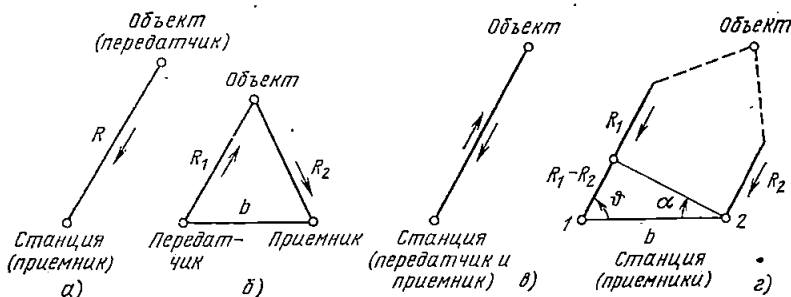


Рис. 1.4. Принцип измерения координат

лебаний, согласованными по частоте и фазе. Любое сколь угодно малое расхождение частот этих двух колебаний с течением времени приводит к расхождению их по фазе (т. е. по времени), в результате через достаточно большое время после сверки эталонов, правильный отсчет времени запаздывания сигнала становится невозможным. Таким образом, реализация беззапросного метода измерения дальности связана с большими трудностями.

В системах с разнесенными передатчиком и приемником (рис. 1.4, б), в которых сигнал проходит суммарный путь передатчик-объект и объект-приемник (суммарно-дальномерных), время запаздывания

$$\tau = (R_1 + R_2)/c. \quad (1.2)$$

Для его отсчета в общем случае также необходимы эталоны времени. Однако при известном расстоянии между передатчиком и приемником b для отсчета можно использовать сигнал, принимаемый от передатчика непосредственно, с поправкой на время его запаздывания b/c , и необходимость в эталонах времени отпадает.

Наиболее удобно измерять дальность R запросным методом, когда передатчик и приемник совмещены (рис. 1.4, в). Время запаздывания отраженного или переизлученного сигнала (без учета задержки в ответчике)

$$\tau = 2R/c. \quad (1.3)$$

При этом запросный сигнал является опорным для отсчета времени запаздывания.

Измерение разности расстояний или угловых координат производится сравнением времени прихода сигналом объекта в разнесенных точках приема (рис. 1.4, г)

$$\Delta\tau = (R_1 - R_2)/c. \quad (1.4)$$

Разнос b называют базой. При малых базах ($b \ll R_1, R_2$) лучи в точках приема можно считать параллельными и разностно-дальномерная система становится угломерной:

$$R_1 - R_2 \approx b \cos \varphi = b u_{x, y},$$

где $u_x, u_y = \cos \varphi_{x,y} = \sin \alpha_{x,y}$ — направляющие косинусы относительно баз, направленных вдоль осей x или y , а φ или α — углы, отсчитываемые относительно базы или нормали к базе соответственно. Разность времен приема сигналов в разнесенных точках угломерной системы

$$\Delta\tau \approx (b/c) u_{x,y}$$

пропорциональна направляющему косинусу. Направление на объект характеризуется двумя направляющими косинусами u_x и u_y , отсчитываемыми от взаимно перпендикулярных баз.

Таким образом, угломерные системы являются вырожденным случаем разностно-дальномерных систем: на больших удалениях от центра базы гиперболы совпадают с их асимптотами, проходящими через центр базы. Разностно-дальномерные системы используют большие базы, соизмеримые с расстоянием до объекта, а угломерные — очень малые. Поэтому первые относятся к многопунктным системам, вторые — к однопунктным.

При отсчете разности времен запаздывания двух принимаемых сигналов один из них используется в качестве опорного. Поэтому в угломерных и разностно-дальномерных системах не имеет значения, какой сигнал принимается — прямой, отраженный или переизлученный. Разностно-дальномерные системы удобны для навигации по сигналам наземных пунктов, поскольку на борту достаточно иметь приемно-измерительное устройство.

В основу измерения производных от координат по времени (составляющих вектора скорости объекта) положен эффект Доплера — дополнительный набег фазы при взаимном перемещении источника и приемника колебаний, равносильный смещению частоты принимаемых колебаний относительно излучаемых. Точные формулы для нахождения доплеровского сдвига частоты могут быть получены только в рамках теории относительности.

Рассмотрим предварительно две инерциальные системы отсчета (в плоскости): неподвижную (X, Y) и движущуюся вдоль оси X со скоростью $V(X', Y')$, внутри которой передается сигнал из точки O' в точку A' (рис. 1.5, а). В момент излучения сигнала начала координат O и O' обеих систем совмещены. Через время Δt в неподвижной системе отсчета (или $\Delta t'$ в движущейся) сигнал приходит в точку A (или A'), пройдя путь $c\Delta t$ (или $c\Delta t'$), так как скорость света не зависит от системы отсчета. При этом сама движущаяся система переместилась на расстояние $V\Delta t$ (или $V\Delta t'$).

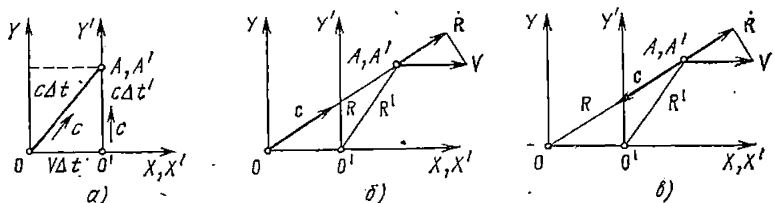


Рис. 1.5. Эффект Доплера

щаяся система переместится относительно неподвижной на расстояние $V\Delta t$ вдоль оси X . Расстояние, проходимое сигналом вдоль оси Y , в обеих системах одинаково и равно $c\Delta t'$, поскольку вдоль этой оси системы взаимно неподвижны. Из прямоугольного треугольника $ОАО'$ неподвижной системы отсчета находим, что

$$c\Delta t' = \sqrt{c^2\Delta t^2 - V^2\Delta t^2}$$

или

$$\Delta t' = \Delta t \sqrt{1 - V^2/c^2}. \quad (1.5)$$

Таким образом, в движущейся системе отсчета время течет медленнее, а цена единицы измерения времени (секунды), выше чем в неподвижной системе.

Рассмотрим теперь радиоканал, в котором передатчик совмещен с точкой O неподвижной системы, а приемник с точкой A' системы, движущейся относительно нее со скоростью V вдоль оси X (рис. 1.5, б). Фаза (состояние) сигнала в точке приема не зависит от системы отсчета: $2\pi f(t - R/c) = 2\pi f'(t' - R'/c)$. Поскольку в движущейся системе расстояние R' неизменно и $dR'/dt' = 0$, после дифференцирования последнего равенства получим

$$f(1 - \dot{R}/c) = f'(dt'/dt),$$

где $\dot{R} = dR/dt$ — радиальная скорость приемника относительно передатчика. Из формулы (1.5) следует, что $dt'/dt = \sqrt{1 - V^2/c^2}$. Поэтому частота принимаемого сигнала

$$f' = f \frac{1 - \dot{R}/c}{\sqrt{1 - V^2/c^2}} \quad (1.6)$$

отличается от частоты излучаемых колебаний на величину доплеровского сдвига $F = f' - f$, пропорционального радиальной скорости объекта.

Если же передатчик и приемник поменять местами (рис. 1.5, в — излучается f' , принимается f), формула (1.6) сохраняет свой вид, только перед слагаемыми $\dot{R}/c$ нужно поменять знак на противоположный, так как в этом случае векторы $\mathbf{R}$ и $\mathbf{c}$ направлены в противоположные стороны:

$$f = f' \frac{\sqrt{1 - V^2/c^2}}{1 + \dot{R}/c} \quad (1.7)$$

Заметим, что частота излучаемых и принимаемых колебаний при работе по сигналу прямого излучения неодинакова, даже в том случае, когда $\dot{R} = 0$ (объект движется со скоростью V в тангенциальном направлении). Частота как число периодов в единицу времени в этом случае неодинакова только потому, что сама единица времени зависит согласно (1.5) от системы отсчета. Возникающее при этом различие по частоте $F \approx fV^2/2c^2$ называют релятивистским поперечным доплеровским эффектом.

При работе по отраженному или переизлученному сигналу (запросные системы) частота принимаемых и переизлучаемых ответчиком сигналов определяется формулой (1.6), а частота принимаемых станций ответных сигналов — формулой (1.7). Обозначив частоту запросных сигналов в (1.6) как $f=f_0$ и подставив f' в формулу (1.7), получим

$$f = f_0 \frac{1 - \dot{R}/c}{1 + \dot{R}/c}. \quad (1.8)$$

В этом случае релятивистский доплеровский эффект отсутствует, так как частоты f_0 и f сравниваются в одной и той же (в данном случае неподвижной) системе отсчета.

У большинства реальных объектов $V \ll c$, поэтому вместо точных формул (1.6) и (1.7) для сигналов прямого излучения часто пользуются приближенной формулой $f = f_0(1 - \dot{R}/c)$, откуда находим доплеровский сдвиг частоты

$$F \approx f - f_0 = -f_0 \dot{R}/c = -\dot{R}/\lambda. \quad (1.9)$$

При работе по отраженным или ответным сигналам вместо (1.8) справедлива формула

$$F \approx -2f_0 \dot{R}/c = -2\dot{R}/\lambda. \quad (1.10)$$

При этом допускается погрешность порядка $f_0(\dot{R}/c)^2$. При работе по отраженному (ответному) сигналу доплеровский сдвиг частоты (1.10) вдвое больше, чем при сигналах прямого излучения (1.9), так как набег фазы сигнала, обусловленный движением объекта, происходит дважды — при запросе и ответе.

Таким образом, измерение радиальной скорости объекта сводится к определению доплеровского сдвига частоты принимаемых сигналов. В этом и состоит принцип измерения радиальной скорости. Знак доплеровского сдвига указывает направление движения объекта: при сближении ($\dot{R} < 0$) частота увеличивается, при удалении ($\dot{R} > 0$) уменьшается.

Разность доплеровских частот сигналов, принимаемых в разнесенных точках, пропорциональна разности радиальных скоростей объекта относительно этих точек:

$$\Delta F = F_1 - F_2 = -f_0(\dot{R}_1 - \dot{R}_2)/c \quad (1.11)$$

(аналогичное соотношение со знаком плюс справедливо для суммарной доплеровской частоты в разнесенных системах). При малой базе $b \ll R_1, R_2$, как отмечалось, $R_1 - R_2 \approx b \dot{u}_{x,y}$. Следовательно, согласно (1.11) разность доплеровских частот пропорциональна производной $\dot{u}_x$ или $\dot{u}_y$ направляющих косинусов относительно соответствующих осей (обычно взаимно перпендикулярных):

$$\Delta F \approx -f_0(b/c) \dot{u}_{x,y} = -(b/\lambda) \dot{u}_{x,y}. \quad (1.12)$$

Знак этой разности указывает сторону движения объекта: частота выше в той из двух точек приема, в сторону которой движется объект.

Для измерения радиальной скорости по сигналу прямого излучения в пункте приема необходим опорный сигнал, имеющий частоту излучаемых колебаний. Для этого на передающем и приемном концах используют высокостабильные эталоны частоты — кварцевые или молекулярные генераторы. Требования к стабильности частоты этих эталонов значительно ниже, чем для эталонов времени: небольшое расхождение частот двух эталонов сохраняется в заданных пределах, в то время как набег фазы эталонов времени, вызванный этим расхождением, непрерывно нарастает. Поэтому реализовать беззапросные системы измерения радиальной скорости значительно проще, чем беззапросные системы измерения дальности. При работе по отраженному (или ответному) сигналу запросный сигнал выполняет функции опорного, а требования к стабильности его частоты существенно снижаются, так как за малое время запаздывания она не успевает заметно измениться. При измерении угловой скорости по разности доплеровских частот один из принимаемых сигналов играет роль опорного, поэтому требования к стабильности частоты колебаний также невысокие.

Итак, принцип измерения относительных координат и их производных радиотехническими методами состоит в том, что аналогом дальности является время запаздывания, аналогом радиальной скорости — доплеровский сдвиг частоты принимаемого сигнала, аналогом направляющего косинуса — разность времен запаздывания и аналогом производной направляющих косинусов — разность доплеровских частот сигналов, принимаемых в разнесенных точках приема:

$$\tau \sim R, \Delta\tau \sim u_{x,y}, F \sim \dot{R}, \Delta F \sim \dot{u}_{x,y}. \quad (1.13)$$

Никакие другие принципы измерений относительных координат объектов и их первых производных в измерительных радиосистемах не используются. Например, измерение дальности по амплитуде принимаемого сигнала неприемлемо ввиду чрезвычайно низкой точности, так как на амплитуду влияет множество факторов, не поддающихся строгому учету. Так называемые косвенные методы не могут рассматриваться как самостоятельные измерения. Например, определение дальности интегрированием радиальной скорости, измеренной по доплеровскому сдвигу частоты, или определение радиальной скорости дифференцированием текущей дальности, измеренной по времени запаздывания, являются операциями вычисления, а не измерениями. Измеряется только радиальная скорость в первом случае и дальность во втором. Вычисления не несут дополнительной информации, не расширяют наших знаний о положении и движении объекта, а только видоизменяют их.

1.4. МЕТОДЫ ИЗМЕРЕНИЯ ОТНОСИТЕЛЬНЫХ КООРДИНАТ И СОСТАВЛЯЮЩИХ ВЕКТОРА СКОРОСТИ ОБЪЕКТА

При измерении производных координат непосредственно реализуется принцип их определения — фиксируется доплеровский сдвиг частоты или разность доплеровских частот, а отсчет производится по набегу фазы или разности фаз за некоторый интервал времени. Однако принципы измерения самих координат реализуются различными методами, так как измеряемые координаты могут быть закодированы в различных параметрах сигнала.

Известны три основных метода измерения дальности: временной (или импульсный), фазовый и частотный. В радиосистемах, использующих импульсные или другие более сложные сигналы, непосредственно реализуется принцип измерения дальности по времени запаздывания сигнала — временной метод: фиксируется интервал τ между принятым и излученным импульсами. Измерение дальности фазовым методом основано на том, что фаза некоторой гармонической составляющей f_m принимаемого сигнала относительно излученного пропорциональна времени запаздывания τ , и следовательно, дальности:

$$\varphi = 2\pi f_m \tau. \quad (1.14)$$

Ввиду существенной неоднозначности измерение дальности фазовым методом производят не на несущей, а на сравнительно низкой частоте модуляции или частоте биений двух высокочастотных колебаний f_m . Частотный метод измерения дальности основан на использовании непрерывных колебаний, частота которых определенным образом изменяется во времени. Так, при линейном изменении излучаемые колебания частоты $f = f_0 + at$ и принимаемые колебания частоты $f' = f_0 + a(t - \tau)$ дают биения с частотой

$$f_R = f - f' = a\tau, \quad (1.15)$$

которая пропорциональна времени запаздывания и, следовательно, дальности.

Для измерения разности и суммы расстояний используют те же методы, что и при измерении дальности, для измерения угловых координат — фазовый и амплитудный метод. Фазовый метод основан на пропорциональности между разностями фаз $\Delta\varphi$ и времен запаздывания $\Delta\tau$ сигналов (1.4), принимаемых в разнесенных точках:

$$\Delta\varphi = 2\pi f_0 \Delta\tau = 2\pi f_0 b u_{x,y}/c. \quad (1.16)$$

Разность фаз, пропорциональная направляющему косинусу u_x или u_y , обычно определяется непосредственно на несущей частоте f_0 , поскольку разность хода сигналов и времен их запаздывания в разнесенных точках антенны очень мала, так что неоднозначность либо отсутствует, либо легко устраняется. При амплитудном методе

измерения угловых координат зависящая от направления прихода волны разность фаз сигналов, принимаемых отдельными элементами антенны, преобразуется в амплитудную зависимость суммарного выходного сигнала антенны от направления приема. Эту зависимость — пеленгационную характеристику — можно сделать линейной функцией угла на некотором интервале. Чтобы исключить влияние неизвестных факторов, измеряют не абсолютное, а относительное значение амплитуды, сравнивая ее с амплитудой опорного сигнала, формируемого той же антенной.

Таким образом, определение относительных координат и их производных в радиосистемах сводится к измерению неэнергетических параметров сигнала: частотного сдвига, времени запаздывания, разности фаз и относительной амплитуды. Относительная амплитуда не является энергетическим параметром, так как ее зависимость от угла появляется как результат расхождения фаз принимаемых элементами антенны сигналов, а не изменения их интенсивности, влияние которой устраняется при сравнении измерительного и опорного сигналов (нормировке).

Следует отметить, что для независимого измерения координаты и ее производной нельзя использовать один и тот же параметр сигнала или антенны. Так, измерения дальности по фазе огибающей высокочастотного сигнала и радиальной скорости по частоте несущей — независимые. При этом частоты огибающей и несущей отличаются во много раз. Измерение же дальности и радиальной скорости соответственно по фазе огибающей и частоте той же огибающей не дает взаимно независимой информации, так как частота является производной фазы. Аналогично нельзя считать независимыми измерения угла по разности фаз и угловой скорости по разности частот одних и тех же сравниваемых сигналов при одной и той же базе антенны. Для того чтобы угловая скорость давала дополнительную информацию, ее нужно измерять по относительной базе b/λ , значительно превышающей угломерную, т. е. использовать дополнительную антенную систему с большей базой или дополнительный сигнал с большей частотой.

В заключение отметим, что ввиду нелинейной зависимости между измеряемым углом и разностью времен прихода сигналов в разнесенных точках приема, пропорциональной направляющему косинусу, для определения угловых скоростей и составляющих вектора скорости недостаточно измерить только производные направляющих косинусов. Так, измеренные относительно осей антенной системы, расположенных вдоль поверхности земли в направлениях север-юг (ось x) и запад-восток (ось y) (рис. 1.6), направляющие коси-

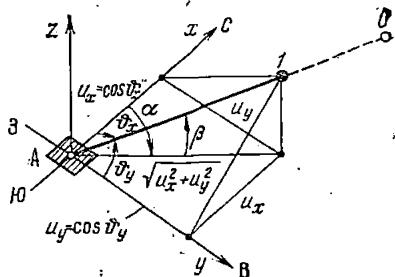


Рис. 1.6. Связь между угловыми координатами и направляющими косинусами объекта

Кол. № 70776
200 8

ОБЩЕИЗДАТЕЛЬСКАЯ
ИМ. ГОРЬКОГО 17

нуса $u_x = \cos \vartheta_x$ и $u_y = \cos \vartheta_y$ позволяют определить азимут и угол места объекта

$$\alpha = \arctg(u_y/u_x); \quad \beta = \arccos \sqrt{u_x^2 + u_y^2}. \quad (1.17)$$

Если, кроме того, измерить производные направляющих косинусов $\dot{u}_x$ и $\dot{u}_y$, можно определить азимутальную и угломестную составляющие угловой скорости объекта

$$\dot{\alpha} = \frac{\dot{u}_x \cos \alpha - \dot{u}_y \sin \alpha}{\cos \beta}; \quad \dot{\beta} = \frac{\dot{u}_x \sin \alpha - \dot{u}_y \cos \alpha}{\sin \beta}. \quad (1.18)$$

Для определения составляющих вектора скорости объекта — радиальной, тангенциальной по азимуту и тангенциальной по углу места

$$V_R = \dot{R}; \quad V_\alpha = \dot{\alpha} R; \quad V_\beta = \dot{\beta} R \quad (1.19)$$

(где R — дальность до объекта) с учетом (1.18) наряду с производными его координат необходимо измерить и сами координаты, т. е. нужно знать положение объекта.

1.5. ОБОБЩЕННАЯ СХЕМА И ОСНОВНЫЕ ХАРАКТЕРИСТИКИ РАДИОСИСТЕМ

При всем многообразии радиосистем у них существует и много общего. Поэтому можно ввести обобщенную схему радиосистемы, включающую аппаратуру станции и аппаратуру, установленную на объекте (рис. 1.7). В общем случае аппаратура как станции, так и объекта включает передатчик и приемник вместе с антеннами и эталоны частоты или времени; кроме того, станция содержит оконечное устройство (измеритель), предназначенное для обнаружения сигналов и измерения их параметров. Заметим, что в системах самоопределения станция размещается на самом объекте, а под аппаратурой объекта следует понимать аппаратуру вспомогательных опорных объектов (наземных ответчиков или маяков).

В частных, конкретных случаях тот или иной элемент обобщенной схемы может отсутствовать, а работа на передачу и прием обеспечиваться общей антенной через развязывающее устройство или антенный коммутатор. Так, в радиолокационных системах обычно

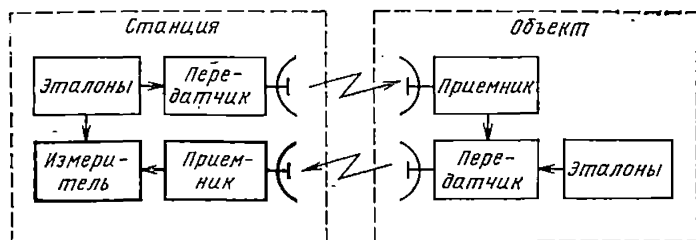


Рис. 1.7. Обобщенная структурная схема радиосистем

используется общая сканирующая антенна (антенная решетка), аппаратура на объекте полностью исключается, работа производится по отраженному от объекта сигналу. Для отсчета временного и частотного сдвигов (дальности, радиальной скорости) в качестве опорного можно использовать сам излучаемый сигнал, ввиду чего необходимость в эталонах времени и частоты практически исключается. Измеритель фиксирует факт наличия объекта в зоне обзора и его координаты; данные выдаются в аналоговой или цифровой форме.

Системы с активным ответом содержат ответчик, состоящий из передатчика, приемника и антенной системы (общей или раздельной), аппаратура самой станции практически не отличается от предыдущей. Необходимость в эталонах также отпадает; обычно ответный сигнал когерентно связан с запросным. Чтобы запросный и ответный сигналы не создавали взаимных помех, их обычно разносят по частоте путем деления (умножения) частоты запросного сигнала в дробное число раз. Это позволяет исключить связь между каналами запроса и ответа на гармониках сигнала и обеспечить малый разнос по частоте, а следовательно, использовать общий для обоих каналов антенно-фидерный тракт как в станции, так и в ответчике. В этих системах ответные сигналы могут отличаться от запросных не только по частоте, но и по форме. Это позволяет кодировать сигналы и использовать их для передачи дополнительной информации в обе стороны (например, в целях опознавания).

В беззапросных системах для измерения дальности и радиальной скорости обязательно наличие эталонов времени и частоты как на станции, так и на объекте; что существенно усложняет аппаратуру. Для измерения угловых координат любыми методами эталоны не нужны.

Системы пассивной радиолокации отличаются от беззапросных полным отсутствием специального оборудования на объекте и эталонов на станции. Сама станция содержит только приемно-измерительный тракт. Ввиду этого исключается возможность измерения дальности и радиальной скорости объекта.

Таким образом, неизменными элементами всякой измерительной радиосистемы являются антенна, приемник и окончное устройство (измеритель) (на рис. 1.7 выделены жирными линиями), все остальные элементы системы служат лишь для создания носителя информации — электромагнитного колебания и хранения данных о его параметрах. Электромагнитное колебание становится сигналом только после отражения, излучения или переизлучения его объектом. Параметры принятого сигнала содержат информацию о координатах объекта и их производных, для выделения которой в измерителе производится сравнение этих параметров с текущими или хранящимися параметрами излучаемых колебаний. Из сказанного следует, что наиболее существенную роль в радиосистемах играет приемно-измерительный тракт, которому преимущественно и посвящена настоящая книга.

Качество измерительной информации зависит от формы сигнала и распределения поля в раскрыве антенны, которые являются материальными носителями этой информации. При заданной форме сигнала одни параметры передаются хорошо, другие плохо. Так, короткий радиоимпульс хорошо передает информацию о дальности и плохо о скорости; колебание типа немодулированной несущей не позволяет однозначно измерить дальность, но хорошо передает информацию о радиальной скорости — по сдвигу частоты колебаний. Выбрав форму сигнала как носителя информации, можно обеспечить наилучшее качество ее в каждом конкретном случае. Очевидно, что системы пассивной радиолокации не дают возможности для такого выбора, поэтому они далеко не во всех случаях обеспечивают получение необходимых данных.

Качество получаемой в радиосистемах информации зависит от помех, которые, накладываясь на сигнал, искажают информацию. Основной задачей измерительных систем является выделение полезной информации на фоне помех. Эту задачу и выполняет приемно-измерительный тракт — основной элемент всякой обзорно-измерительной радиосистемы.

Свойства радиосистем описываются их основными характеристиками. Выходные (или внешние) характеристики радиосистемы определяют ее возможности при выполнении поставленной задачи. Входные (или внутренние) характеристики отвечают на вопрос, какими техническими средствами обеспечиваются указанные выше возможности радиосистемы.

Выходные характеристики подразделяют на энергетические, метрические и эксплуатационные.

Энергетические характеристики: размеры рабочей зоны и качество обнаружения, которые находятся в прямой связи с энергозатратами. Размеры рабочей зоны обычно характеризуются дальностью действия системы и секторами обзора по азимуту и углу места. Качество обнаружения сигналов чаще всего оценивают вероятностью правильного обнаружения при заданной достаточно низкой вероятности ложных обнаружений.

Метрические характеристики: число и вид измеряемых параметров, точность и однозначность их измерения, разрешающая способность по каждому из них. Точность измерений обычно характеризуется максимальной систематической погрешностью и среднеквадратическим значением случайной ошибки. Разрешающая способность обычно оценивается минимальным разномом сигналов близко расположенных объектов по данному измеряемому параметру, при котором еще возможно раздельное наблюдение без взаимных помех, если все остальные параметры сигналов одинаковы.

Эксплуатационные характеристики: надежность и помехозащищенность. Надежность оценивается вероятностью безотказной работы системы в течение установленного времени, или средним временем работы между смежными отказами, или частотой отказов. Помехозащищенность включает помехоустойчивость и скрытность

работы системы. Помехоустойчивостью называют способность системы сохранять работоспособность при воздействии помех. Под скрытностью работы понимают трудности определения противником параметров системы, знание которых необходимо для постановки наиболее эффективной помехи.

Временная характеристика — период обзора — среднее время, в течение которого производится однократный обзор зоны (интервал между засечками положения объекта). Своевременное обнаружение и правильное определение траектории объекта обеспечивается при достаточно малом периоде обзора.

Входные характеристики так же подразделяют на энергетические, метрические, структурные и массогабаритные.

Энергетические характеристики: энергия излучения за один обзор (или средняя мощность), эффективная площадь антенны и уровень шумов, пересчитанных на вход приемного устройства, который обычно оценивают спектральной плотностью шума.

Метрические характеристики: форма, длительность и несущая частота (длина волны) сигнала, форма функции раскрытия и относительные размеры антенны (по отношению к длине волны).

Структурные характеристики определяют аппаратную сложность системы, способ обзора рабочей зоны (последовательный одноканальный, параллельный многоканальный или смешанный) и степень автоматизации процессов обслуживания и контроля.

Массогабаритные характеристики: масса и габаритные размеры станции, потребляемая от внешних источников мощность.

Задание конкретных входных характеристик системы однозначно определяет выходные характеристики системы. Это так называемая прямая задача. Так, энергетические выходные характеристики (в частности, дальность действия при заданном качестве обнаружения сигнала) полностью определяются через энергетические входные характеристики. Метрические же выходные характеристики (точность, разрешающая способность) определяются и через метрические, и через энергетические входные характеристики. В случае одноканального последовательного обзора проще аппаратура, но велик период обзора; при многоканальном одновременном обзоре аппаратура существенно сложнее, зато период обзора минимален.

При синтезе (проектировании) радиосистем решается обратная задача: по заданным выходным характеристикам определяют необходимые для их обеспечения входные характеристики системы. Сложность этой задачи состоит в том, что она не имеет однозначного решения: одни и те же выходные характеристики системы обеспечиваются при различных сочетаниях входных характеристик. Например, большую дальность действия можно обеспечить, увеличив либо энергию излучения, либо эффективную площадь антенны; точность измерения угловых координат можно повысить, увеличив размеры антенны или уменьшив длину волны. Ис-

жусство проектировщика состоит в том, чтобы найти наиболее экономичное решение.

Развиваемая здесь теория позволяет произвести оптимальный синтез радиосистем лишь относительно метрических и энергетических характеристик, входных и выходных. Зависимость же между временными выходными и структурными входными характеристиками, как и ряд других, в решающей степени определяется конкретным назначением и видом системы и поэтому не может быть охвачена общей теорией. Все эти вопросы рассматриваются в специальных курсах, посвященных технической реализации радиосистем.

1.6. ПРИНЦИПЫ И МЕТОДЫ ИЗМЕРЕНИЯ АБСОЛЮТНЫХ ПАРАМЕТРОВ ДВИЖЕНИЯ

Координаты объекта и их первые производные являются относительными навигационными параметрами, учитывающими только кинематику относительного движения двух объектов, один из которых (безразлично какой) принимается в качестве опорного. Между этими двумя пунктами создается внешнее по отношению к ним электрическое поле, используемое в качестве носителя информации. Поэтому такие радиосистемы можно назвать системами внешних измерений.

Наряду с относительными существуют абсолютные параметры, связанные с динамикой движения объекта и действующими на него силами и полями, например угловые скорость и ускорение объекта, его линейное ускорение и гравитация. Для их измерения используют соответственно гиометры (гироскопы); акселерометры и гравиметры, основанные чаще всего на неинерциальных механических явлениях. Особенность этих измерителей состоит в том, что для отсчета абсолютных параметров не требуется второе опорное тело. Корпус измерителя жестко закрепляется на самом объекте. Чувствительным элементом механических измерителей является масса, не имеющая жесткой связи с корпусом измерителя, внутри которого она располагается.

Возможны радиотехнические аналоги механических измерителей. Некоторые уже созданы и успешно соперничают с механическими измерителями. В качестве чувствительного элемента радиотехнических измерителей используется электромагнитное колебание, распространяющееся от излучателя к приемнику, но не вне, как у радиосистем внешних измерений, а внутри объекта, параметры движения которого измеряются. Поэтому такие измерители можно назвать радиосистемами внутренних измерений. При наличии экранировки внешнее электромагнитное поле у них отсутствует и не может быть обнаружено никакими приборами.

В отличие от рассмотренных ранее радиосистем внешних измерений строгое описание процессов распространения электромагнитных волн в неинерциальных системах отсчета, каковыми являются радиосистемы внутренних изме-

рений, возможно только в рамках общей теории относительности, с использованием тензорного анализа. Однако в некоторых простейших ситуациях необходимые соотношения можно получить на основе классических представлений, если процессы, протекающие в неинерциальных системах, рассматривать в инерциальной (невращающейся) системе отсчета.

Принцип измерения угловой скорости изолированного объекта с помощью электромагнитных волн состоит в следующем. Пусть на объекте закреплен пассивный интерферометр Саньяка в виде круглого контура радиуса r ($L_K = 2\pi r$), вращающегося в плоскости контура с угловой скоростью Ω (рис. 1.8, а). С помощью полупрозрачного (ПЗ) и непрозрачного (НЗ) зеркал в контур вводятся и из контура выводятся электромагнитные колебания, разделяющиеся на две встречные волны. Ввод и вывод колебаний происходит в одной и той же точке вращающегося кольца. Однако если процессы рассматривать в инерциальной системе отсчета, то точки излучения (Изл) и приема (Прм) оказываются разнесенными на расстояние $\Delta L = \Omega r L_K / c$, которое равно пути, проходимому кольцом при вращении за время обхода контура волной L_K / c (линейная скорость любой точки кольца составляет Ωr).

Поэтому путь от точки излучения до точки приема и время его прохождения по контуру для волны, движущейся в сторону вращения контура (попутной), возрастают (по сравнению с неподвижным контуром):

$$\tau_{\text{п}} = \frac{L_K + \Delta L}{c} = \frac{L_K}{c} + \frac{\Omega r L_K}{c^2};$$

для встречной волны, наоборот, убывают:

$$\tau_{\text{в}} = \frac{L_K - \Delta L}{c} = \frac{L_K}{c} - \frac{\Omega r L_K}{c^2}.$$

Отсюда вытекает принцип измерения угловой скорости объекта, состоящий в том, что разность времен обхода контура встречными волнами

$$\Delta\tau = \tau_{\text{п}} - \tau_{\text{в}} = \frac{2rL_K}{c^2} \Omega = \frac{4S_K}{c^2} \Omega, \quad (1.20)$$

где $S_K = \pi r^2 = r L_K / 2$ — площадь контура, пропорциональная его угловой скорости Ω . Знак этой разности указывает направление вращения, поэтому ее отсчет должен производиться так, чтобы сохранить знак.

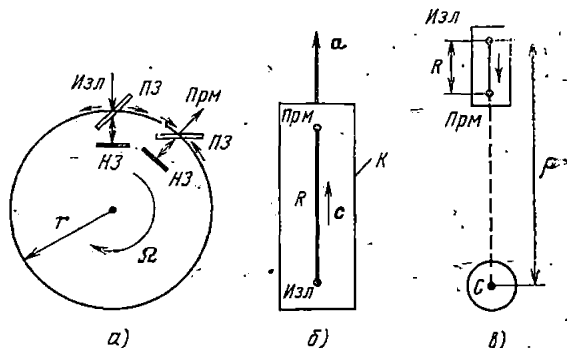


Рис. 1.8. Принципы измерения угловой скорости (а), линейного ускорения (б) и гравитации (в)

Принцип измерения линейного ускорения состоит в следующем. Пусть на объекте, подверженном линейному ускорению a , жестко закреплен корпус акселерометра (К) и связанные с ним излучатель и приемник колебаний, размещенные в направлении вектора a на расстоянии R друг от друга (рис. 1.8,б). Рассмотрим процессы в инерциальной системе отсчета, совпадающей с неинерциальной, ускоренной системой отсчета в момент излучения сигнала, когда скорость движения точки излучения в обеих системах одинакова. В неинерциальной системе, связанной с корпусом акселерометра и объектом, излучатель и приемник взаимно неподвижны. Однако в инерциальной системе отсчета точка приема движется относительно точки излучения со скоростью $\dot{R} = at = aR/c$, так как за время распространения волны между этими двумя точками $t = R/c$ скорость движения объекта изменилась под действием силы, вызывающей ускорение a . Появление относительной скорости между излучателем и приемником вызывает согласно формуле (1.9) доплеровский сдвиг частоты принимаемых колебаний относительно частоты излучения f_0 .

Отсюда вытекает принцип измерения линейного ускорения изолированного объекта — доплеровский сдвиг частоты в акселерометре пропорционален ускорению объекта:

$$F = -f_0 \frac{\dot{R}}{c} = -\frac{f_0 R}{c^2} a. \quad (1.21)$$

Ускорение положительно (сдвиг частоты отрицателен), если оно направлено от излучателя к приемнику.

Принцип измерения углового ускорения в кольцевом интерферометре Саньяка аналогичен: линейная скорость движения точки приема по кольцу относительно точки излучения, отсчитываемая в системе без углового ускорения (в ускоренной системе отсчета эти точки взаимно неподвижны), вызывает доплеровский сдвиг частоты между двумя встречными волнами, пропорциональный угловому ускорению объекта в плоскости кольца. Знак частотного сдвига противоположен знаку ускорения.

Отметим, что в заданной системе координат угловая скорость и ускорение имеют три составляющих (вектор), поэтому для их измерения нужно иметь три измерителя, размещаемые в трех взаимно перпендикулярных плоскостях.

Несколько подробнее остановимся на принципе измерения гравитации. В наземных условиях все тела имеют вес, который ощущается как реакция опоры на данный объект. Опорой является земля, наземные объекты, корабли, самолеты в горизонтальном полете и т. д. Реакцию опоры можно рассматривать как силу, вызывающую ускорение, равное ускорению земного притяжения g , но противоположное по знаку. Таким образом, для измерения гравитации в этом случае можно использовать рассмотренный выше принцип измерения линейных ускорений $a = -g$.

Согласно закону Всемирного тяготения Ньютона ускорение земного притяжения связано с геоцентрическим расстоянием ρ соотношением $g = \gamma M/\rho^2$, где γ — гравитационная постоянная; M — масса Земли. Следовательно, по измеренному ускорению g можно определить расстояние ρ от объекта до центра Земли.

Сложнее измерить g или ρ , когда опора и, следовательно, ее реакция (вес тела) отсутствуют, например у искусственных спутников Земли (ИСЗ). Здесь принцип акселерометра непригоден: результирующие силы и ускорения равны

нулю. Однако можно воспользоваться особенностью центрального поля тяготения Земли — неоднородностью по величине (зависит от ρ) и направлению. Эту особенность и можно в принципе использовать для измерений гравитации, разместив излучатель и приемник на ИСЗ на расстоянии R друг от друга. в направлении на центр Земли C (рис. 1.8, в).

Для определенности положим, что излучатель совмещен с центром тяжести ИСЗ и отстоит от центра Земли на расстоянии $\rho \gg R$. В центре тяжести ИСЗ центростремительное ускорение земного тяготения $g_{изл} = \gamma M / \rho^2$ уравновешено центробежным ускорением движения по орбите $a_{изл} = \Omega^2 \rho$, откуда находим, что $\Omega^2 = \gamma M / \rho^3$. В точке приема $g_{прм} = \gamma M / (\rho - R)^2$ уже не равно $a_{прм} = \Omega^2 (\rho - R)$. Найдем разности:

$$\Delta g = \gamma M / (\rho - R)^2 - \gamma M / \rho^2 \approx 2\gamma MR / \rho^3;$$

$$\Delta a = \Omega^2 \rho - \Omega^2 (\rho - R) = \Omega^2 R = \gamma MR / \rho^3.$$

Общее ускорение в точке приема $\omega_{прм} = \Delta g + \Delta a = 3\gamma MR / \rho^3$. Это ускорение, приобретаемое каждым излучаемым фотоном массы m при движении от точки излучения ($x=0$) к точке приема ($x=R$), линейно нарастает от 0 до $\omega_{прм}$, т. е. $\omega(x) = 3\gamma Mx / \rho^3$. Двигаясь к Земле, падая, излучаемый фотон приобретает дополнительную энергию

$$\Delta E = m \int_0^R \omega(x) dx = \frac{3\gamma MmR^2}{2\rho^3}.$$

Энергия фотона может возрасти только путем изменения частоты колебаний. Если f_0 — частота излучаемых колебаний, а f — принимаемых, то энергия фотона в соответствии с формулами Планка и Эйнштейна в точках излучения и приема

$$E_{изл} = hf_0 = mc^2,$$

$$E_{прм} = hf = mc^2 + \Delta E = mc^2 (1 + \Delta E / mc^2) = hf_0 (1 + \Delta E / mc^2),$$

где h — постоянная Планка. Отсюда находим $f = f_0 (1 + \Delta E / mc^2)$ и доплеровский сдвиг частоты

$$F = f - f_0 = f_0 \frac{\Delta E}{mc^2} = \frac{3\gamma MR^2}{2c^2 \rho^3} f_0. \quad (1.22)$$

Таким образом, принцип измерения гравитации с помощью электромагнитных волн состоит в том, что доплеровский сдвиг частоты F в разнесенных по высоте точках приема и излучения является функцией геоцентрического расстояния объекта ρ или местного значения ускорения земного притяжения $g = \gamma M / \rho^2$. Сдвиг частоты положителен, если точка излучения находится выше точки приема.

Из всех устройств внутренних измерений в настоящее время нашли применение только гиromетры (гироскопы). Акселерометры и особенно гравиметры трудно реализовать, потому что сдвиг частоты в устройствах приемлемых размеров чрезвычайно мал. Согласно формулам (1.21) и (1.22) он обратно пропорционален квадрату скорости света, а в гравиметрах, кроме того, кубу геоцентрического расстояния, а это, как известно, очень большие величины.

Однако околоземное и космическое пространство не является строго инерциальной системой, поэтому описанные здесь явления проявляются и в обычных

в беззапросных радиосистемах — в виде дополнительных ошибок (в запросных системах они отсутствуют). При большом различии высот передатчика и приемника $R = r_{\text{изл}} - r_{\text{прм}}$ частотный сдвиг (1.22) может оказаться определяющим, особенно в космических беззапросных радиосистемах высокой точности. Аналогичные ошибки возникают в беззапросных радиосистемах, движущихся с большим ускорением вдоль линии распространения радиоволн. Поэтому в результаты измерений необходимо вносить поправки.

Принцип измерения угловой скорости (1.20) реализуется двумя методами: частотным и фазовым.

Частотный метод используется в лазере с кольцевым резонатором, образованным в виде правильного многоугольника с помощью зеркал (лазерный гироскоп). Такой резонатор эквивалентен кольцевому с радиусом вписанной окружности $r_{\text{эк}} = 2S_{\text{к}}/L_{\text{к}}$, где $S_{\text{к}}$ — площадь, а $L_{\text{к}}$ — периметр многоугольника. При вращении гироскопа в плоскости резонатора его длина для встречного и попутного колебаний в инерциальной системе отсчета оказывается различной, в результате и частоты генерируемых колебаний отличаются на разностную частоту

$$F = -2f_0(r_{\text{эк}}/c)\Omega = (4S_{\text{к}}/\lambda_0 L_{\text{к}})\Omega, \quad (1.23)$$

где f_0 и λ_0 — частота и длина волны колебаний в неподвижном гироскопе. Таким образом, измеряемая частота биений двух колебаний пропорциональна угловой скорости.

Фазовый метод основан на измерении разности фаз

$$\Delta\varphi = 2\pi f_0 n \Delta t = (8\pi n S_{\text{к}} f_0 / c^2) \Omega \quad (1.24)$$

двух встречных колебаний, вводимых от внешнего лазера во вращающийся n -витковый светодоводный интерферометр, где время прохождения колебаний по одному витку Δt определяется соотношением (1.20).

Измеряемые разности частот (1.23) и фаз (1.24) приобретают заметную величину только на достаточно высоких световых частотах, поэтому рассматриваемые гироскопы были созданы только после появления лазеров и светодоводов.

Таким образом, измерение как относительных, так и абсолютных навигационных параметров в радиосистемах сводится к сравнению излучаемого и принимаемого сигналов по одному из параметров. Измерение параметров сигнала и обнаружение его производится в присутствии неизбежных помех (шумов), имеющих случайный характер. Поэтому при описании процессов в радиосистемах используется статистическая теория, ориентирующаяся на некоторый средний результат. Эта теория относится в основном к обычным радиосистемам, так как системы внутренних измерений используют достаточно сильные сигналы, когда влияние шумов пренебрежимо мало и определяющими являются аппаратные погрешности, полностью зависящие от конкретной технической реализации измерителя, что выходит за рамки рассматриваемой задачи.

Глава 2. ОПТИМАЛЬНЫЙ ПРИЕМ СИГНАЛОВ ИЗВЕСТНОЙ ФОРМЫ

2.1. ОЦЕНКА КАЧЕСТВА ПРИЕМА СИГНАЛОВ

В обзорно-измерительных радиосистемах используют сигналы заранее известной формы, являющиеся переносчиками информации, содержащейся в параметрах этих сигналов. Исключение составляют сигналы систем пассивной радиолокации, включая устройства инфракрасного диапазона, которые обычно рассматривают отдельно.

Прием сигналов происходит в присутствии неизбежных помех. Помехи, накладываясь на сигнал, искажают его, в том числе и параметры сигнала, несущие информацию о координатах объекта и его производных. Поэтому безошибочно определить эти параметры невозможно. По входному воздействию $x(t, \mathbf{r}) = s(t, \mathbf{r}) + n(t, \mathbf{r})$, представляющему собой смесь сигнала с помехой в функции от времени t и координат точек приема $\mathbf{r}$, можно получить лишь более или менее достоверное представление об измеряемом параметре — его статистическую оценку. Статистическая оценка формируется в результате определенных операций над входной смесью сигнала и помехи с учетом их статистических характеристик. Получение статистических оценок — *измерение* — является основной задачей приема сигналов в радиосистемах.

Однако принимаемый сигнал $s(t, \mathbf{r})$ может быть до такой степени искажен помехами, что не всегда удастся правильно установить сам факт его наличия. Поэтому измерению предшествует или сопровождает его *обнаружение* сигнала на фоне помех — дополнительная задача приема.

Таким образом, прием сигналов — обнаружение и измерение — неизбежно сопровождается ошибками (промахами). Для оптимального приема, обеспечивающего минимум промахов, необходимо ввести критерии качества.

При обнаружении сигналов известной формы возможны два исходных условия: либо сигнал есть (1), т. е. $x(t, \mathbf{r}) = s(t, \mathbf{r}) + n(t, \mathbf{r})$, либо его нет (0), т. е. $x(t, \mathbf{r}) = n(t, \mathbf{r})$, и два решения: «сигнал есть» (1*) или «сигнала нет» (0*). Оптимальным решением является такое, при котором обеспечивается одновременно максимум вероятности правильного обнаружения $P_D = P(1^*|1)$ и минимум вероятности ложной тревоги $P_F = P(1^*|0)$, т. е. ложного обнаружения. Оба эти требования можно объединить в единый весовой критерий оптимального обнаружения, при котором максимизируется взвешенная разность

$$P_D - \Lambda_0 P_F, \quad (2.1)$$

так как она растет с увеличением P_D и с уменьшением P_F . Здесь весовой коэффициент $\Lambda_0 \gg 1$ учитывает относительную цену приращения P_F по сравнению с приращением P_D . В реальных системах, например радиолокационных, допустимая вероятность ложной тревоги P_F очень мала ($10^{-6} \dots 10^{-8}$), иначе пропадет доверие к результатам обнаружения, а требуемая вероятность правильного обнаружения P_D обычно близка к единице. Изменение вероятности P_D на 1% (10^{-2}) практически не играет никакой роли, в то время как увеличение P_F на 1% недопустимо велико. Более того, часто P_F ограничивают строго определенным значением; тогда критерием оптимального обнаружения вместо (2.1) становится требование максимума P_D — широко распространенный критерий Неймана — Пирсона.

Сигнал известной формы (обычно в виде пакета импульсов, число которых постоянно) принимается при равномерном обзоре рабочей зоны. Существуют также более гибкие системы с неравномерной скоростью обзора, зависящей от случайно складывающейся ситуации. Эти системы здесь не рассматриваются, так как совокупный сигнал уже нельзя считать известным (число импульсов в пакете случайно).

Критерием эффективности измерения произвольного неэнергетического параметра сигнала является минимум дисперсии (среднего квадрата) погрешности измерений $\epsilon = \hat{v} - v$ этого параметра

$$\sigma_v^2 = \overline{(\hat{v} - v)^2} = \int \int_{\hat{v}} (\hat{v} - v)^2 W_1(v, \hat{v}) dv d\hat{v}, \quad (2.2)$$

где v и $\hat{v}$ — истинное и измеренное (оценка) значения параметра; $W_1(v, \hat{v})$ — плотность вероятностей их совместного распределения (погрешности измерений) при наличии сигнала. Погрешности измерений обычно подчиняются гауссовскому закону с нулевым средним (несмещенная оценка); точные измерения производятся при большом выходном отношении сигнал-шум, когда наложение гауссовских шумов на сигнал вызывает гауссовский разброс (погрешность) измеряемого параметра. Поэтому критерием эффективности измерений и является минимум дисперсии погрешности — единственного параметра гауссовского закона.

2.2. ОПТИМАЛЬНЫЙ ПРИЕМНИК

Установим вид операций над входной смесью сигнала и шума $x(t, r)$ в оптимальном приемнике обнаружения и измерения.

Заменим (по теореме Котельникова) входную реализацию $x(t, r)$ отсчетами $x_1, x_2, \dots, x_n$, которые можно считать проекциями вектора x на оси многомерной ортогональной системы координат. Иными словами, сложный образ $x(t, r)$ заменяем простым многомерным — точкой, соответствующей концу вектора x . Совокупность всех возможных входных реализаций занимает в многомер-

ном пространстве некоторую область A . Задача приемника обнаружения состоит в том, чтобы одни реализации x отнести к группе «сигнал есть», другие — к группе «сигнала нет», т. е. разбить область A на две: «сигнальную» $\hat{A}_1$ и «шумовую» $\hat{A}_0$.

Введем многомерные плотности вероятностей $W_1(x)$ для отсчетов функции $x(t, r) = s(t, r) + n(t, r)$, когда сигнал действительно есть, и $W_0(x)$ для отсчетов функции $x(t, r) = n(t, r)$, когда сигнала нет. Тогда по определению

$$P_D = \int_{\hat{A}_1} W_1(x) dx, \quad P_F = \int_{\hat{A}_1} W_0(x) dx, \quad (2.3)$$

где интегрирование ведется по области $\hat{A}_1$, при попадании в которую реализации x принимается решение «сигнал есть». Для обоснованного решения о том, какие же реализации следует отнести к «сигнальным», т. е. как выбрать область $\hat{A}_1$, воспользуемся оптимальным весовым критерием (2.1)

$$P_D - \Lambda_0 P_F = \int_{\hat{A}_1} [W_1(x) - \Lambda_0 W_0(x)] dx \rightarrow \max. \quad (2.4)$$

Максимум интеграла обеспечивается, если в область $\hat{A}_1$ включить такие x , для которых подынтегральная функция неотрицательна, т. е. $W_1(x) - \Lambda_0 W_0(x) \geq 0$. Отсюда, если ввести в рассмотрение отношение правдоподобия $\Lambda(x) = W_1(x)/W_0(x)$, условие оптимального обнаружения сигнала примет вид

$$\Lambda(x) = W_1(x)/W_0(x) \geq \Lambda_0. \quad (2.5)$$

Таким образом, операция обнаружения сигнала в оптимальном приемнике при поступлении на его вход реализации x сводится к формированию отношения правдоподобия и испытанию его на порог, выбранный по одному из оптимальных критериев. В случае превышения порога принимается решение «сигнал есть» при непревышении — «сигнала нет».

Для нахождения структуры оптимального приемника измерения примем во внимание, что оценка параметра $\hat{v}$ вырабатывается в результате определенных операций над принятой реализацией x , представляющей собой смесь сигнала и помехи, т. е. существует однозначная функциональная зависимость $\hat{v} = \hat{v}(x)$. От замены переменных вероятность не изменится:

$$W_1(v, \hat{v}) dv d\hat{v} = W_1(v, x) dv dx = W_1(x) W_1(v|x) dv dx,$$

где $W_1(x)$ — плотность распределения вероятностей входных воздействий, не зависящая от измеряемого параметра; $W_1(v|x)$ — апостериорная плотность распределения вероятностей измеряемого параметра.

Полученная после замены переменных дисперсия погрешности измерений (2.2)

$$\sigma^2_v = \int_x W_1(x) dx \int_v (\hat{v} - v)^2 W_1(v|x) dv \quad (2.6)$$

минимальна, если минимален зависящий от оценки $\hat{v}$ внутренний интеграл $J = \int_v (\hat{v} - v)^2 W_1(v|x) dv$, т. е.

$$\frac{dJ}{d\hat{v}} = 2 \int_v (\hat{v} - v) W_1(v|x) dv = 0.$$

Разбив интеграл на два, в первом из которых выносим за знак интеграла $\hat{v}$ и полагаем $\int_v W_1(v|x) dv = 1$, получим оптимальную оценку

$$\hat{v} = \int_v v W_1(v|x) dv. \quad (2.7)$$

Таким образом, оптимальная оценка параметра находится как центр тяжести (первый момент) его апостериорного распределения $W_1(v|x)$.

Согласно известной формуле Байеса

$$W_1(v, x) = W(v) W_1(x|v) = W_1(x) W_1(v|x),$$

откуда находим

$$W_1(v|x) = \frac{1}{W_1(x)} W(v) W_1(x|v) = \frac{W_0(x)}{W_1(x)} W(v) \Lambda(x|v),$$

где $W(v)$ — априорная плотность распределения вероятностей измеряемого параметра; $W_1(x|v)$ — условная плотность распределения вероятностей входных воздействий x при заданном v , которую как функцию v называют функцией правдоподобия; $\Lambda(x|v) = W_1(x|v)/W_0(x)$ — отношение правдоподобия. Тогда оптимальная оценка (2.7) примет вид

$$\hat{v} = k \int_v v W(v) \Lambda(x|v) dv, \quad (2.8)$$

где $k = W_0(x)/W_1(x)$ — коэффициент, не зависящий от v и определяемый из условия нормировки $W_1(v|x)$.

Функция $W(v)$ считается известной до приема сигнала. В процессе приема сигнала оптимальным приемником (измерения) формируется лишь отношение правдоподобия $\Lambda(x|v)$, как и при обнаружении сигнала.

Представляет интерес частный случай, когда $W(v) = \text{const}$, а отношение правдоподобия унимодально (имеет только один максимум) и симметрично относительно v_m . Тогда оценкой (2.8) становится абсцисса максимума отношения правдоподобия $\hat{v} = v_m$. Формирование оценки таким способом получило название метода максимума правдоподобия. При указанных выше допущениях оценка по методу максимума правдоподобия совпадает с оптимальной байесовской. Эти допущения на практике обычно выполняются при точных измерениях, когда пик отношения правдоподобия достаточ-

но узкий и симметричный (описывается нормальным распределением), а априорное распределение достаточно широкое, чтобы в пределах узкого пика его можно было считать равномерным (рис. 2.1).

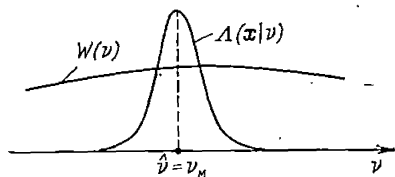


Рис. 2.1. Формирование оценки по методу максимума правдоподобия

Из проведенного анализа следует, что оптимальные приемники обнаружения и измерения выполняют одну и ту же операцию над входной смесью — формирование отношения правдоподобия, т. е. в радиосистемах существует лишь один оптимальный приемник. Различие заключается только в использовании выходного результата приемника: при обнаружении отношения правдоподобия испытывается на порог, а при измерении находится значение параметра, при котором отношение правдоподобия максимально.

2.3. ЛОГАРИФМ ОТНОШЕНИЯ ПРАВДОПОДОБИЯ И КОРРЕЛЯЦИОННЫЙ ИНТЕГРАЛ ВРЕМЕННЫХ СИГНАЛОВ

Формирование отношения правдоподобия является единственной операцией, производимой оптимальным приемником над входной смесью сигнала и шума. Испытание полученного отношения на порог при обнаружении или нахождения абсциссы его максимума при измерении параметров сигнала по методу максимума правдоподобия выполняются выходными устройствами приемника.

Вычислим отношение правдоподобия (2.6) сначала для случая чисто временных процессов $x(t)$, когда помехой является аддитивный белый гауссовский шум $n(t)$, а сигнал $s(t; v, \varphi)$ кроме интересующего нас измеряемого параметра v содержит случайный неизмеряемый параметр φ с заданной плотностью распределения $W(\varphi)$. Для этого определим сначала многомерную плотность распределения вероятностей $W(n)$ независимых отсчетов шума $n = \{n_1, n_2, \dots, n_N\}$ на интервале T , число которых $N = 2F_{\max}T$. По определению распределение i -го отсчета шума n_i нормальное с нулевым математическим ожиданием и дисперсией σ_n^2 :

$$W(n_i) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{n_i^2}{2\sigma_n^2}\right),$$

а многомерная плотность распределения вероятностей согласно теореме умножения вероятностей независимых событий

$$W(n) = \prod_{i=1}^N W(n_i) = (2\pi\sigma_n^2)^{-N/2} \exp\left(-\frac{1}{2\sigma_n^2} \sum_{i=1}^N n_i^2\right).$$

Учитывая, что $\sigma_n^2 = N_0 F_{\max}$, где N_0 — спектральная плотность шума; $F_{\max}$ — максимальная частота в его спектре, для перехода от

отсчетов n_i к непрерывной функции $n(t)$ положим $F_{\max} \rightarrow \infty$. Тогда интервал между отсчетами $\Delta t = 1/2F_{\max} \rightarrow dt$ и сумма обращается в интеграл:

$$\frac{1}{2\sigma_n^2} \sum_{i=1}^N n_i^2 = \frac{1}{N_0} \int_0^T n^2(t) dt,$$

откуда находим плотность распределения шума

$$W[n(t)] = (2\pi\sigma_n^2)^{-N/2} \exp \int_0^T n^2(t) dt.$$

При наличии сигнала во входной смеси $x(t)$ шум $n(t) = x(t) - s(t)$, при отсутствии $n(t) = x(t)$. Подставив значения $n(t)$ в последнюю формулу, получим плотность распределения входной смеси при наличии и отсутствии сигнала соответственно

$$W_1[x(t)] = (2\pi\sigma_n^2)^{-N/2} \exp \left\{ -\frac{1}{N_0} \int_0^T [x(t) - s(t)]^2 dt \right\}, \quad (2.9)$$

$$W_0[x(t)] = (2\pi\sigma_n^2)^{-N/2} \exp \left\{ -\frac{1}{N_0} \int_0^T x^2(t) dt \right\}. \quad (2.10)$$

Отсюда согласно формуле (2.5) находим отношение правдоподобия как функцию входной смеси, зависящую от измеряемого v и неизменяемого φ параметров:

$$\Lambda[x(t); v, \varphi] = \exp \left(-\frac{\partial}{N_0} \right) \exp \left\{ \frac{2}{N_0} \int_{-\infty}^{\infty} x(t) s(t; v, \varphi) dt \right\}, \quad (2.11)$$

где

$$\partial = \int_{-\infty}^{\infty} s^2(t) dt \quad (2.12)$$

— величина, пропорциональная энергии сигнала (на сопротивлении 1 Ом). Пределы в последних выражениях заменены бесконечными, что допустимо, так как вне интервала 0, T подынтегральная функция равна нулю.

Возьмем развернутую запись высокочастотного сигнала с несущей f_0 и неизвестной начальной фазой φ_0 (неизменяемый параметр), имеющий равномерное распределение $W(\varphi_0) = 1/2\pi$ на интервале $-\pi, +\pi$:

$$s(t) = A(t) \cos[2\pi f_0 t + \varphi(t) + \varphi_0],$$

где $A(t)$ и $\varphi(t)$ — функции, определяющие соответственно законы амплитудной и фазовой модуляции сигнала, медленно изменяющиеся по сравнению с $\cos 2\pi f_0 t$. Чтобы представить отношение правдоподобия явной функцией параметра φ_0 , подставим развернутое выражение для сигнала в корреляционный интеграл $z = \int x(t) \times$

$\times s(t)dt$ формулы (2.11) и произведем тригонометрические преобразования:

$$\begin{aligned} z &= \int_{-\infty}^{\infty} x(t) A(t) \cos[2\pi f_0 t + \varphi(t) + \varphi_0] dt = \\ &= \left\{ \int_{-\infty}^{\infty} x(t) A(t) \cos[2\pi f_0 t + \varphi(t)] dt \right\} \cos \varphi_0 - \\ &- \left\{ \int_{-\infty}^{\infty} x(t) A(t) \sin[2\pi f_0 t + \varphi(t)] dt \right\} \sin \varphi_0 = \\ &= Z_1 \cos \varphi_0 - Z_2 \sin \varphi_0 = Z \left(\frac{Z_1}{Z} \cos \varphi_0 - \frac{Z_2}{Z} \sin \varphi_0 \right) = Z \cos(\theta + \varphi_0), \end{aligned} \quad (2.13)$$

где

$$Z_1 = \int_{-\infty}^{\infty} x(t) A(t) \cos[2\pi f_0 t + \varphi(t)] dt,$$

$$Z_2 = \int_{-\infty}^{\infty} x(t) A(t) \sin[2\pi f_0 t + \varphi(t)] dt$$

— ортогональные составляющие корреляционного интеграла;

$$Z = \sqrt{Z_1^2 + Z_2^2} \quad (2.14)$$

— огибающая (модуль) корреляционного интеграла;

$$\theta = \arctg \frac{Z_2}{Z_1} \quad \left(\cos \theta = \frac{Z_1}{Z}, \quad \sin \theta = \frac{Z_2}{Z} \right)$$

— фаза (аргумент) корреляционного интеграла. Подставив $z = Z \cos(\theta + \varphi_0)$ в выражение (2.11) и усреднив его по всем φ_0 с учетом распределения $W(\varphi_0) = 1/2\pi$, получим отношение правдоподобия для сигнала с неизвестной начальной фазой

$$\begin{aligned} \Lambda[x(t); v] &= \exp \left\{ -\frac{\vartheta}{N_0} \right\} \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp \left\{ \frac{2Z}{N_0} \right\} \cos(\theta + \varphi_0) d\varphi_0 = \\ &= \exp \left\{ -\frac{\vartheta}{N_0} \right\} I_0 \left(\frac{2Z}{N_0} \right); \end{aligned} \quad (2.15)$$

где $I_0(2Z/N_0)$ — модифицированная функция Бесселя нулевого порядка.

Следует отметить, что в качестве выходного сигнала удобно брать не само отношение правдоподобия, а монотонную функцию от него — логарифм отношения правдоподобия

$$l[x(t)] = \ln \Lambda[x(t)] = -\frac{\vartheta}{N_0} + \ln I_0 \left(\frac{2Z}{N_0} \right) \quad (2.16)$$

или огибающую корреляционного интеграла Z . При больших отношениях сигнал-шум, обеспечивающих высокое качество обнаружения и измерения, огибающая корреляционного интеграла линейно

связана с логарифмом отношения правдоподобия $l[x(t)] \approx -\mathcal{Z}/N_0 + 2Z/N_0$, так как $\ln l_0(y) \approx y$ при $y \gg 1$. Удобство использования l или Z в качестве выходного сигнала приемника состоит в том, что в случае приема составного сигнала (например, пакета импульсов) суммарный выходной сигнал образуется как сумма частных выходных сигналов, а не как произведение частных отношений правдоподобия.

Замена отношения правдоподобия его логарифмом или огибающей корреляционного интеграла не изменяет результата измерений, поскольку их максимум наступает при одном и том же значении измеряемого параметра. Аналогично не изменятся и результаты обнаружения, если логарифм отношения правдоподобия испытывать на порог $l_0 = \ln \Lambda_0$, а огибающую корреляционного интеграла на порог $Z_0 = (\mathcal{Z} + l_0 N_0)/2$.

В дальнейшем в качестве выходного сигнала будем в основном пользоваться огибающей корреляционного интеграла Z , которую можно получить (исключив неизвестную начальную фазу φ_0), продетектировав корреляционный интеграл z — сигнал на выходе линейной части оптимального приемника.

При описании процессов в радиосистемах широко используется комплексная огибающая узкополосного радиосигнала. Так, радиосигнал можно представить как действительную часть произведения комплексной огибающей

$$S(t) = A(t)e^{j\varphi(t)} = A(t)\cos\varphi(t) + jA(t)\sin\varphi(t)$$

и осциллирующего высокочастотного множителя $\exp[j2\pi f_0 t]$, модуль которого равен единице:

$$\begin{aligned} s(t) &= A(t)\cos([2\pi f_0 t + \varphi(t)]) = \operatorname{Re}\{A(t)e^{j\varphi(t)}e^{j2\pi f_0 t}\} = \\ &= \operatorname{Re}\{S(t)e^{j2\pi f_0 t}\}. \end{aligned}$$

В комплексной огибающей заложена вся информация о структуре сигнала и его параметрах. Высокочастотное заполнение, используемое лишь как переносчик информации по радиоканалу, исключается при детектировании сигнала после приема. Описание процессов в радиосистемах с помощью комплексных огибающих, а не самих радиосигналов удобнее, поскольку оно не связано с тригонометрическими преобразованиями, а сигнал и его спектр при этом описываются в общем случае одинаково — комплексными функциями.

Через комплексную огибающую $X(t)$ выражается узкополосное входное воздействие приемника

$$x(t) = \operatorname{Re}\{X(t)e^{j2\pi f_0 t}\} = \frac{1}{2}[X(t)e^{j2\pi f_0 t} + X^*(t)e^{-j2\pi f_0 t}]$$

и корреляционный интеграл $Z = Z_1 + jZ_2$, где

$$Z_{1,2} = \int_{-\infty}^{\infty} x(t)A(t) \cos_{\sin}[2\pi f_0 t + \varphi(t)] dt$$

($X^*(t)$ — комплексно-сопряженная функция). При этом комплексная огибающая корреляционного интеграла

$$\begin{aligned} Z &= \int_{-\infty}^{\infty} x(t) A(t) \{ \cos [2\pi f_0 t + \varphi(t)] + j \sin [2\pi f_0 t + \varphi(t)] \} dt = \\ &= \int_{-\infty}^{\infty} x(t) A(t) \exp \{ j 2\pi f_0 t + \varphi(t) \} dt = \int_{-\infty}^{\infty} x(t) S(t) \exp \{ j 2\pi f_0 t \} dt = \\ &= \frac{1}{2} \int_{-\infty}^{\infty} X(t) S(t) \exp \{ j 2\pi \cdot 2f_0 t \} dt + \frac{1}{2} \int_{-\infty}^{\infty} X^*(t) S(t) dt \approx \\ &\approx \frac{1}{2} \int_{-\infty}^{\infty} X^*(t) S(t) dt, \end{aligned}$$

так как первый интеграл от быстро осциллирующей знакопеременной функции близок к нулю, поскольку модуль ее $|X(t)S(t)|$ практически не изменяется за период колебаний частоты $2f_0$.

В общем случае огибающую корреляционного интеграла можно представить двумя равноценными выражениями

$$Z = \frac{1}{2} \left| \int_{-\infty}^{\infty} X^*(t) S(t) dt \right| = \frac{1}{2} \left| \int_{-\infty}^{\infty} X(t) S^*(t) dt \right|. \quad (2.17)$$

Комплексная огибающая входного воздействия $X(t) = S(t; \nu) + N(t)$ представляет собой сумму комплексных огибающих сигнала и шума, причем сигнал зависит от измеряемого параметра ν . Второй сомножитель в интегралах (2.17) является комплексной огибающей опорной функции приемника $S(t; \nu^0)$, где ν^0 — значение измеряемого параметра, на которое настроен приемник. Следовательно, огибающая корреляционного интеграла зависит от расстройки $\nu - \nu^0$:

$$Z(\nu - \nu^0) = |Z(\nu - \nu^0)| = \frac{1}{2} \left| \int_{-\infty}^{\infty} X(t) S^*(t; \nu^0) dt \right|. \quad (2.18)$$

Выходной сигнал оптимального приемника, описываемый комплексной огибающей $Z = Z_s + Z_n$, содержит сигнальную и шумовую составляющие:

$$Z_s(\nu - \nu^0) = \frac{1}{2} \int_{-\infty}^{\infty} S(t; \nu) S^*(t; \nu^0) dt, \quad (2.19)$$

$$Z_n(\nu - \nu^0) = \frac{1}{2} \int_{-\infty}^{\infty} N(t) S^*(t; \nu^0) dt. \quad (2.20)$$

При точной настройке ($\nu^0 = \nu$) сигнальная составляющая выходного напряжения, пропорциональная полной энергии сигнала

$$\mathcal{S} = Z_s(0) = \frac{1}{2} \int_{-\infty}^{\infty} |S(t; \nu^0)|^2 dt, \quad (2.21)$$

вещественна и максимальна. По максимуму ее можно безошибочно определить истинное значение измеряемого параметра. Од-

нако неизбежное присутствие случайной шумовой составляющей Z_n смещает максимум огибающей (2.18) и оценка $\hat{v}$ выдается с погрешностью $\varepsilon = \hat{v} - v$.

2.4. ОПТИМАЛЬНЫЙ ПРИЕМНИК ВРЕМЕННЫХ СИГНАЛОВ

Существуют два метода формирования корреляционного интеграла, т. е. два способа реализации оптимального приемника: корреляционный и фильтрационный.

Структура оптимального корреляционного приемника (рис. 2.2, а) непосредственно вытекает из формул (2.13) и (2.14). Наряду с входной смесью $x(t)$ на два квадратурных канала подается опорный сигнал (реплика) $s(t)$, причем в каждом канале производится перемножение входного и опорного напряжений (на смесителях) и интегрирование полученных произведений (с помощью узкополосных фильтров нижних частот — ФНЧ), а затем видеосигналы обоих каналов объединяются в квадратуре. В выходном устройстве выполняются операции обнаружения (испытание выходного сигнала на порог) или измерения (испытание на экстремум).

Использование двух квадратурных каналов обусловлено незнанием начальной фазы принимаемого сигнала. Если неизвестна разность фаз между входным и опорным сигналами случайно окажется равной $\pi/2$ при перемножении входного и опорного сигналов частоты f_0 , видеосигнал на выходе узкополосного ФНЧ одного из квадратурных каналов может отсутствовать, зато во втором канале будет принят полный сигнал. В промежуточном случае в одном канале выдается видеосигнал Z_1 , пропорциональный косинусу разности фаз, в другом — сигнал Z_2 , пропорциональный синусу, а на выходе квадратурного сумматора получаем полный сигнал — огибающую корреляционного интеграла Z .

Можно отказаться от второго канала, если на смеситель подать опорный сигнал на частоте гетеродина f_r , отличающийся от частоты принимаемого сигнала f_0 на промежуточную частоту $f_{пр} = f_r - f_0$, а сигнал промежуточной частоты z подать на «линейный» амплитудный детектор, выделяющий огибающую Z корреляционного интеграла (рис. 2.2, б). В этом случае роль интегратора выполняет узкополосный резонансный фильтр промежуточной частоты (ФПЧ).

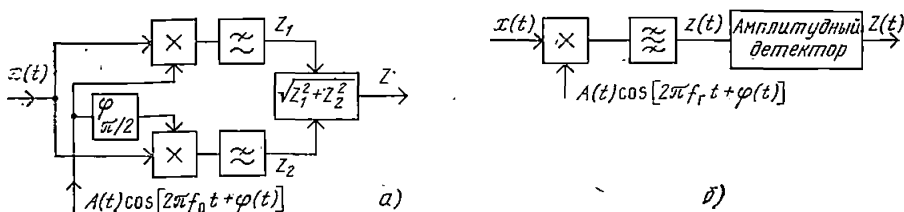


Рис. 2.2. Схемы оптимального корреляционного приемника

Особенность корреляционного приемника обоих видов — неинвариантность к времени прихода сигнала — является серьезным недостатком. Для того чтобы получить корреляционный интеграл, и принимаемый и опорный сигналы должны совпадать по времени. Так как время прихода принимаемого сигнала обычно неизвестно, приходится применять набор большого числа корреляторов, опорные сигналы которых перекрывают все возможные значения времени прихода сигнала. При этом по номеру коррелятора, в котором получен максимальный отклик, как раз и определяется время прихода сигнала (дальность до объекта). Ввиду чрезвычайной сложности многоканальные корреляционные приемники почти не применяют. Обычно используют одноканальные корреляционные приемники там, где отслеживается во времени единственный принимаемый сигнал.

В основу построения *фильтрационного* приемника положена известная зависимость между входным $x(t)$ и выходным $y(t)$ сигналами фильтра с заданной импульсной характеристикой $h(t)$:

$$y(t) = \int_{-\infty}^{\infty} x(t') h(t-t') dt', \quad (2.22)$$

где t' — время как переменная интегрирования.

Возьмем фильтр с импульсной характеристикой

$$h(t) = Ks(t_0 - t), \quad (2.23)$$

которая с точностью до постоянного коэффициента передачи K и постоянного временного сдвига t_0 является зеркальным отражением во времени опорного сигнала $s(t)$. Такой фильтр называют согласованным, а приемник, включающий согласованный полосовой фильтр и амплитудный детектор, оптимальным фильтрационным приемником (рис. 2.3).

Подставим выражение (2.23) для $h'' = t - t'$ в (2.22). Тогда $h(t-t') = h(t'') = Ks(t_0 - t'') = Ks(t_0 - t + t')$ и на выходе согласованного фильтра получаем функцию корреляции

$$y(t) = K \int_{-\infty}^{\infty} x(t') s(t_0 - t + t') dt', \quad (2.24)$$

которая в момент $t = t_0$ становится корреляционным интегралом

$$z = y(t_0) = K \int_{-\infty}^{\infty} x(t') s(t') dt'. \quad (2.25)$$

С выхода амплитудного детектора снимается огибающая функции корреляции $Z(t)$, максимум которой находится в окрестности точки $t = t_0$ (в отсутствие шумов в этой точке) и является искомой огибающей корреляционного интеграла Z .

Все сказанное относится к входному сигналу, началу отсчета которого $t = 0$. Если же вход-

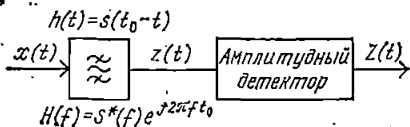


Рис. 2.3. Схема оптимального фильтрационного приемника

ной сигнал поступает с запаздыванием τ , то выходной сигнал имеет максимум вблизи точки $t_0 + \tau$ и по-прежнему является огибающей корреляционного интеграла Z . Отсюда следует, что фильтрационный приемник в отличие от корреляционного инвариантен ко времени прихода сигнала, т. е. оптимален для сигналов множества объектов, поступивших в любой момент времени, и, следовательно, всего один. Важно лишь, чтобы пики выходных сигналов разделялись по времени (разрешались). В этом огромное достоинство фильтрационного приемника перед корреляционным, что и предопределило его преимущественное использование.

В случае превышения пиком выходного сигнала, когда бы он ни появился, заданного порога выдается решение о наличии сигнала (обнаружение), а по моменту наступления максимума пика фиксируется время запаздывания (измерение дальности).

Необходимо отметить важное различие корреляционного и фильтрационного приемников. Свойства корреляционного приемника почти целиком определяются видом опорного сигнала, а структура его сохраняется одинаковой для сигналов любой формы. Наоборот, у оптимального фильтрационного приемника его свойства заложены именно в структуре фильтра и описываются импульсной характеристикой $h(t)$ или функцией передачи фильтра $H(f)$ (частотной характеристикой), которые связаны между собой преобразованием Фурье:

$$\begin{aligned} H(f) &= \int_{-\infty}^{\infty} h(t) e^{-j2\pi ft} dt = K \int_{-\infty}^{\infty} s(t_0 - t) e^{-j2\pi ft} dt = \\ &= K e^{j2\pi ft_0} \int_{-\infty}^{\infty} s(t') e^{j2\pi ft'} dt' = K S^*(f) e^{j2\pi ft_0}, \end{aligned} \quad (2.26)$$

где

$$S(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi ft} dt \text{ — спектр сигнала } s(t).$$

Таким образом, передаточная функция согласованного фильтра комплексно-сопряжена со спектром сигнала. Показатель степени экспоненциального множителя (2.26) учитывает фазовый сдвиг, вызываемый временным сдвигом t_0 на частоте f . Модуль функции передачи $|H(f)| = |S(f)|$. Это показывает, что в присутствии белого шума с равномерным спектром наиболее интенсивные спектральные составляющие сигнала, значительно превышающие уровень шума, передаются с большим весом, чем спектральные составляющие, слабые по сравнению с шумами.

Спектр выходного сигнала (без шумов)

$$Z_s(f) = S(f) H(f) = K e^{j2\pi ft_0} |S(f)|^2 \quad (2.27)$$

является вещественной величиной, т. е. в момент t_0 все спектральные составляющие синфазны и дают максимально возможный пик, а амплитуда выходного сигнала (функция корреляции (2.24)) спадает в обе стороны от t_0 . Скорость спадаания возрастает, а ширина

пика выходного сигнала (интервал корреляции) Δt убывает с возрастанием ширины спектра Δf , т. е. $\Delta t = 1/\Delta f$. Амплитуда выходного сигнала в максимуме, как следует из (2.22), пропорциональна полной энергии сигнала $\mathcal{E}$, т. е. в момент максимума вся энергия входного сигнала сосредоточивается в пике выходного сигнала (интегрируется). Энергии сигнала $\mathcal{E}$ противостоит шум, мощность которого пропорциональна σ_n^2 , а энергия на интервале корреляции $\Delta t = 1/\Delta f$ равна спектральной плотности $N_0 = \sigma_n^2/\Delta f = \sigma_n^2 \Delta t$. На выходе отношение сигнал-шум по напряжению равно энергетическому отношению сигнал-шум $\mathcal{E}/N_0$. Как следует из (2.16), при $\mathcal{E}/N_0 \gg 1$ среднее значение логарифма отношения правдоподобия в максимуме $l = l_s = \mathcal{E}/N_0$ как раз и отображает эту величину.

Дисперсия шумовой составляющей логарифма отношения правдоподобия в максимуме (т. е. в момент обнаружения и измерения)

$$\begin{aligned} \overline{l_n^2} - l_n^* l_n^* &= \frac{1}{N_0^2} \int N(t) S^*(t) dt \int N^*(t) S(t) dt = \\ &= \frac{1}{N_0^2} \iint S^*(t) S(t') \overline{N(t) N^*(t')} dt dt' = \\ &= \frac{1}{N_0^2} \iint S^*(t) S(t') \frac{N_0}{2} \delta(t-t') dt dt' = \frac{1}{2N_0} \int |S(t)|^2 dt = \frac{\mathcal{E}}{N_0}. \end{aligned} \quad (2.28)$$

Здесь произведение интегралов заменено двойным интегралом с разделяющимися переменными интегрирования и учтено фильтрующее свойство дельта-функции $\int S(t') \delta(t-t') dt' = S(t)$, которой описывается функция корреляции белого шума $\overline{N(t) N^*(t')} = 0,5 N_0 \delta(t-t')$.

Таким образом, отношение $\mathcal{E}/N_0$ характеризует и энергетическое отношение сигнал-шум, и отношение сигнал-шум в максимуме выходного сигнала оптимального приемника, и дисперсию его выходного шума.

Характерной особенностью оптимального фильтрационного приемника является то, что при прохождении через согласованный фильтр сигнал в общем случае не сохраняет свою форму: выходной сигнал (2.24) является корреляционной функцией входного, а спектр выходного сигнала (2.27) пропорционален квадрату амплитудного спектра входного сигнала. Сохраняют свою форму лишь такие сигналы ($\cos x$, $\sin Nx/N \sin x$, $\sin x/x$ и др.), двусторонний спектр которых (с учетом отрицательных частот) равномерный и поэтому остается равномерным при возведении в квадрат.

Искажение формы сигнала при прохождении через оптимальный приемник не влечет за собой никаких нежелательных последствий, так как форма сигнала заранее известна и сохранять ее нет никакой надобности. Задачей приемника сигнала известной формы, как отмечалось, является наивысшее качество обнаруже-

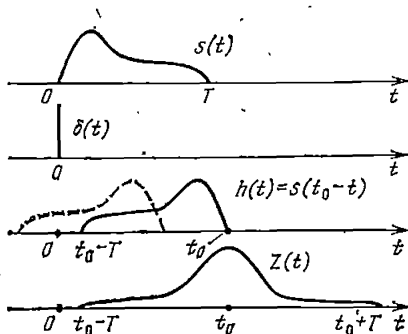


Рис. 2.4. Процессы в согласованном фильтре

ния сигнала и измерения его параметров, которое он и обеспечивает.

Отметим еще одну важную особенность фильтрационного оптимального приемника — условие физической реализуемости фильтра, которое для простоты проиллюстрируем на примере видеосигнала (рис. 2.4). Здесь показан входной сигнал $s(t)$ длительностью T и дельта-функция $\delta(t)$, поступающие в момент $t=0$, а также импульс-

ная характеристика согласованного фильтра $h(t) = s(t_0 - t)$ в виде зеркального отражения сигнала относительно $t=t_0$ (как реакция фильтра на указанную дельта-функцию) и выходной сигнал $Z(t)$ с пиком в момент $t=t_0$. Нетрудно видеть, что условием физической реализуемости согласованного фильтра является неравенство $t_0 \geq T$, т. е. время памяти фильтра должно быть не меньше длительности сигнала. В противном случае реакция фильтра на дельта-функцию появляется раньше входного воздействия, что нереально (на рис. 2.4 штриховой линией показана нереализуемая импульсная характеристика, у которой $t_0 < T$). Действительно, фильтр должен «помнить» все элементарные составляющие входного сигнала, от первой до последней, чтобы в момент максимума выходного сигнала образовать от них суммарный эффект. Условие физической реализуемости легко выполнить для коротких сигналов (импульсов) и сложно для сигналов большой длительности (пакета импульсов, длительность которого составляет несколько периодов повторения). Поэтому реализация элементов памяти фильтра для сигналов большой длительности представляет сложную техническую задачу.

В заключение заметим, что сигнал $s(t)$, имеющий спектр $S(f)$ и наблюдаемый на фоне небелого шума с известной спектральной плотностью $B(f)$, можно свести к эквивалентному сигналу $s_{\text{эк}}(t)$, наблюдаемому на фоне белого шума. Спектр этого сигнала можно найти из условия постоянства спектральной плотности шума во всем интервале частот: $S_{\text{эк}}(f) = S(f)/B(f)$.

Функцию передачи согласованного фильтра для такого сигнала находим согласно формуле (2.26):

$$H(f) = K S_{\text{эк}}^*(f) e^{i 2\pi f t_0} = K \frac{S^*(f)}{B(f)} e^{i 2\pi f t_0}. \quad (2.29)$$

Это и есть функция передачи оптимального фильтра при наложении шумов с неравномерной спектральной плотностью. В этом случае, чтобы найти импульсную характеристику оптимального

фильтра, умножим соотношение (2.29) на $B(f)$ и возьмем от произведения обратное преобразование Фурье:

$$\int_{-\infty}^{\infty} H(f) B(f) e^{j 2 \pi f t} df = \int_{-\infty}^{\infty} K S^*(f) e^{j 2 \pi f t_0} e^{j 2 \pi f t} df.$$

Правая часть в соответствии с формулой (2.26) дает исходную функцию $s(t_0 - t)$. Произведя замену в левой части

$$H(f) = \int_{-\infty}^{\infty} h(t') e^{-j 2 \pi f t'} dt',$$

получим

$$\int_{-\infty}^{\infty} h(t') \left[\int_{-\infty}^{\infty} B(f) e^{j 2 \pi f (t-t')} df \right] dt' = s(t_0 - t).$$

Внутренний интеграл как обратное преобразование Фурье от спектральной плотности дает функцию корреляции шума $B(t - t')$. Отсюда получаем интегральное уравнение для импульсной характеристики (в неявном виде) оптимального фильтра для смеси сигнала и шума с заданной функцией корреляции

$$\int_{-\infty}^{\infty} h(t') B(t - t') dt' = s(t_0 - t). \quad (2.30)$$

К сожалению, это уравнение имеет аналитическое решение только для белого шума, у которого $B(t - t') = 0,5 N_0 \delta(t - t')$ и, следовательно, $h(t) = K s(t_0 - t)$, так как интеграл от произведения какой-либо функции на $\delta(t - t')$ равен значению этой функции при $t' = t$. Для небелого шума также можно записать $h(t) = K s_{\Sigma K}(t_0 - t)$.

Таким образом, все соотношения, относящиеся к реальному сигналу в присутствии белого шума, в общем случае справедливы для эквивалентного сигнала $s_{\Sigma K}(t)$.

2.5. ПРИМЕРЫ СОГЛАСОВАННЫХ ФИЛЬТРОВ

Для иллюстрации процесса оптимальной фильтрации и общих принципов построения согласованных фильтров рассмотрим несколько характерных примеров. Общий принцип построения согласованного фильтра состоит в том, что сначала определяется его импульсная характеристика как зеркальное отражение входного сигнала без шумов, а затем по ней конструируется сам фильтр, реакцией которого на единичный импульс (дельта-функцию) является эта характеристика. На вход подается сам сигнал, с которым согласован фильтр, и находится выходной сигнал, имеющий вид функции корреляции входного сигнала.

Рассмотрим сначала сигнал в виде пакета n коротких видеоимпульсов одинаковой амплитуды, следующих с периодом повторения T_n (рис. 2.5, а). Время памяти фильтра $t_0 = T$, где $T = (n-1)T_n$ — общая длительность сигнала.

Реакцию фильтра вида $h(t) = s(t_0 - t)$, на единичный импульс $\delta(t)$ можно получить с помощью секционированной линии задержки, равноотстоящие отводы

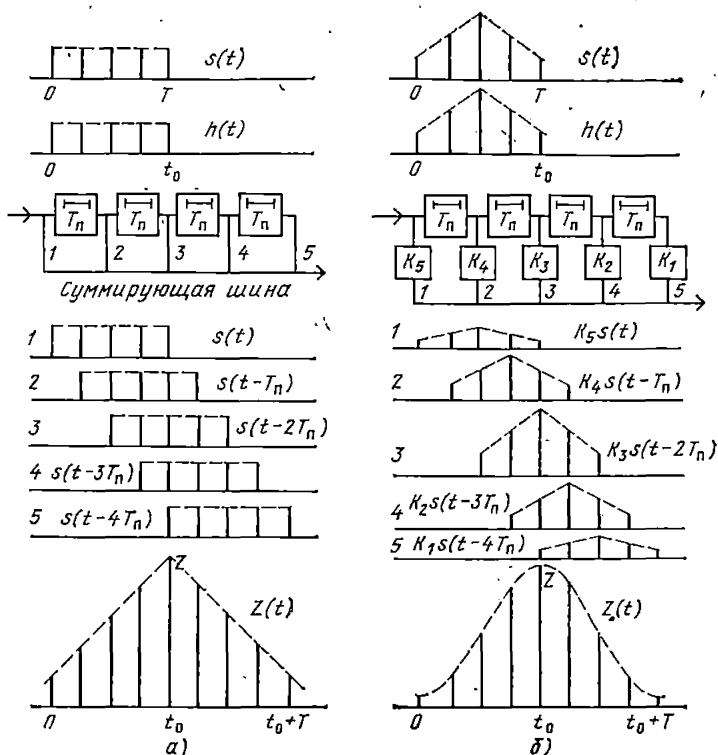


Рис. 2.5. Схемы согласованных фильтров для пакета коротких видеоимпульсов одинаковой (а) и неодинаковой (б) амплитуды

которой (по числу импульсов в пакете) подключены к выходной суммирующей шине. Если пренебречь затуханием в линии задержки или компенсировать его соответствующим усилением в каждом отводе, на выходе такого фильтра действительно получим импульсный пакет, повторяющий сигнал $s(t)$ в обратной последовательности (для пакетов с симметричной огибающей прямая и обратная последовательности совпадают). При поступлении на вход фильтра самого сигнала $s(t)$ с каждого отвода линии задержки на суммирующую шину поступают сигналы, отличающиеся только временем задержки, и на выходе фильтра образуется функция корреляции входного сигнала $Z(t)$, имеющая треугольную огибающую. В момент максимума t_0 она равна сумме всех входных импульсов — корреляционному интегралу Z .

Более общий случай представляет фильтр, согласованный с пакетом импульсов неодинаковой амплитуды, но, как правило, с симметричной огибающей (рис. 2.5, б). Поэтому в отличие от предыдущего случая суммирование сигналов на выходной шине производится с весами K_i , пропорциональными амплитуде соответствующего импульса A_i (нумерация весовых коэффициентов в соответствии с зеркальным характером импульсной характеристики фильтра ведется в обратном порядке). Таким образом, в момент максимума t_0 , когда сигнал полностью вошел в фильтр, амплитуда суммируемых импульсов пропорци-

ональна мощности (энергии) сигнала в импульсе $K_i A_i \equiv A_i^2$, а выходной сигнал (корреляционный интеграл Z) пропорционален общей энергии сигнала. Поэтому такие фильтры называют сумматорами (или интеграторами) импульсов, а также гребенчатыми фильтрами, так как их функция передачи $H(f)$ имеет вид бесконечной гребенки, «зубья» которой разнесены на $F_{\pi} = 1/T_{\pi}$, ширина их $\Delta f_{\pi} = 1/T$, обратная общей длительности сигнала $T = (n-1)T_{\pi}$, а форма определяется видом огибающей пакета импульсов (является преобразованием Фурье от огибающей).

Благодаря весовому суммированию при наличии неизбежных шумов импульсы с большим отношением сигнал-шум в момент максимума t_0 вносят больший вклад в выходной сигнал. Тем самым согласованный фильтр обеспечивает наивысшее отношение сигнал-шум на выходе. Заметим, что при весовом суммировании крайние импульсы пакета выходного сигнала имеют более низкий уровень, а огибающая пакета вблизи максимума изменяется более плавно, чем при равновесном суммировании.

Согласованный фильтр для пакета коротких радиоимпульсов отличается от видеофильтра тем, что он должен учитывать (компенсировать) фазовые соотношения между радиоимпульсами. Для этого закон изменения фазы от импульса к импульсу должен быть известен, т. е. гребенчатый фильтр можно построить лишь для пакета когерентных радиоимпульсов. Некогерентные радиоимпульсы, фазовые соотношения которых неизвестны, можно суммировать только после детектирования на видеочастоте, так как на радиочастоте вместо ожидаемого суммирования может произойти взаимное подавление сигналов из-за расхождения их по фазе.

Таким образом, в отводах согласованного радиофильтра кроме весовых усилителей необходимо ставить компенсирующие фазовращатели. В комплексной записи амплитуды импульсов $S_i = A_i \exp(j\varphi_i)$ и весовые коэффициенты $K_i \equiv S_i^* = A_i \exp(-j\varphi_i)$, поэтому в момент максимума t_0 слагаемые на выходной шине $K_i S_i \equiv |S_i|^2$ вещественны, т. е. импульсы суммируются синфазно и дают на выходе максимально возможный результат. Для пакета синфазных радиоимпульсов фазовращатели не нужны и весовые коэффициенты вещественны. При определенных фазовых соотношениях между импульсами входного сигнала амплитуды радиоимпульсов по обе стороны от максимума огибающей выходного пакета могут очень быстро спадать, так как компенсации фазовых сдвигов не происходит. Благодаря этому в согласованном фильтре достигается сжатие пакета импульсов.

Основной и наиболее технически сложной частью гребенчатого фильтра-интегратора является элемент памяти, который выше был представлен в виде секционированной линии задержки. Однако практически линии задержки используются крайне редко. Дело в том, что период повторения импульсов измеряется миллисекундами и столь большую задержку могут обеспечить только ультразвуковые линии задержки, которые представляют собой довольно сложное устройство, включающее кроме самой линии возбудителя, модуляторы, усилители и т. п. При этом таких линий в фильтре может быть несколько десятков и даже сотен. Поэтому в качестве элементов памяти гребенчатых фильтров чаще используют магнитные барабаны или ленты, иногда фотопленку, а также потенциалоскопы, экраны индикаторов и цифровую память ЭВМ.

Гребенчатый фильтр-интегратор на магнитном барабане или ленте подобен рассмотренным выше интеграторам на линиях задержки, только перемещение

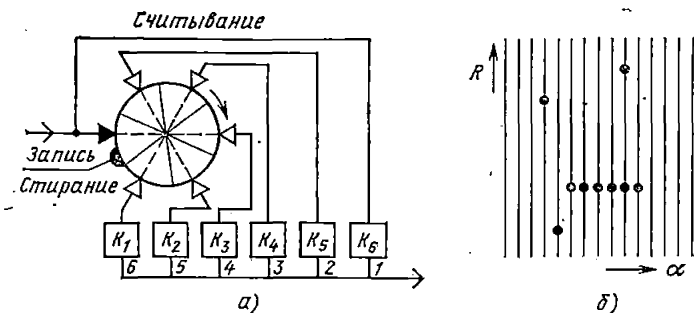


Рис. 2.6. Схема интеграторов импульсов

сигнала в фильтре обеспечивается движением дорожки относительно неподвижных головок, а не задержкой сигнала в линии (рис. 2.6,а). Благодаря тому, что записывающая и считывающие головки размещены на одинаковых расстояниях друг от друга, каждому моменту времени в очередном периоде посылок (углу поворота барабана между головками) соответствует запись сигнала с определенной дальности. В каждом периоде посылки производится запись вновь поступившего импульса и стирание последнего из n импульсов пакета. Так происходит плавное перемещение входного сигнала относительно весового окна, образуемого усилителями на входе суммирующей шины. В момент t_0 , когда пакет и весовое окно совмещаются, выходной сигнал максимален.

Несколько иначе суммируются сигналы на мишени потенциалоскопа или экране индикатора радиолокационной станции (рис. 2.6,б). Запись сигналов в каждом периоде посылки производится построчно, причем удаление от начала строки пропорционально текущей дальности R , а смещение строк — углу поворота антенны α . Импульсы от объекта создают заряды на мишени потенциалоскопа или засветы на экране индикатора. При плотной записи благодаря накоплению заряда мишени или энергии возбуждения люминофора экрана заряды или засветы перекрываются — импульсы суммируются. Однако суммирование импульсов происходит равномерно (т. е. неоптимально), притом интеграторам такого типа свойственно насыщение — эффективно интегрируется лишь ограниченное число импульсов.

Цифровые гребенчатые фильтры в общем случае выполняют над сигналом те же операции, что и аналоговые. В каждом периоде посылки входной сигнал путем квантования по времени и уровню преобразуется в цифры, которые заносятся в ячейку памяти, соответствующие одному и тому же времени (позиция дальности) в n смежных периодах посылок. Эти значения в арифметическом устройстве умножаются на весовые коэффициенты, хранящиеся в других ячейках памяти, и суммируются. В каждом периоде добавляются новый отсчет сигнала на данной позиции и отбрасывается последний из n предыдущих отсчетов, чем достигается продвижение цифровой записи относительно весового окна.

Однако часто в цифровых фильтрах эти операции заменяют простейшими. Например, сигнал квантуют лишь на два уровня 0 и 1, а весовое суммирование заменяют подсчетом числа единиц в n смежных периодах посылок (двоичное интегрирование) или фиксируют наличие единиц в k периодах подряд (метод совпадения). Благодаря этому упрощается устройство цифровой обра-

ботки при сравнительно небольших энергетических потерях. Хотя из-за потерь при квантовании и упрощении процедуры обработки эффективность цифровых фильтров всегда ниже потенциальных возможностей аналоговых устройств, реальный выигрыш обычно мал или отсутствует вообще, поскольку действительные характеристики аналоговых фильтров часто далеко не совершенны.

Индикаторы на электронно-лучевых трубках являются наиболее простыми гребенчатыми фильтрами-интеграторами, у которых съем данных (обнаружение цели, отсчет ее координат) производится человеком-оператором. Все остальные интеграторы используются в автоматизированных системах.

Значительно проще строятся согласованные фильтры для одиночных импульсов ввиду малой длительности сигнала и, следовательно, малой емкости памяти фильтра. Наименее сложны согласованные фильтры для простых импульсов, без внутримпульсной модуляции, база которых (произведение ширины спектра на длительность) $\Delta f \tau_n = 1$. Так, для гауссовского радиоимпульса требуется фильтр, функция передачи (частотная характеристика) которого близка к гауссовской. Именно такую характеристику имеет многокаскадный усилитель промежуточной частоты (УПЧ) на обычных резонансных контурах. Время памяти этого фильтра t_0 определяется крутизной его фазочастотной характеристики $\Phi(\omega) = t_0(\omega - \omega_0)$ вблизи резонансной частоты ω_0 .

Согласованные фильтры для импульсов сложной формы, модулированных по частоте или фазе, тем сложнее, чем больше их база $N = \Delta f T$. При очень большой базе эти фильтры по сложности и объему приближаются к гребенчатым. В согласованном фильтре импульсы сложной формы сжимаются до длительности $\tau_c = 1/\Delta f$, т. е. в $N = T/\tau_c$ раз. Максимальный коэффициент сжатия $N \approx 10^3 \dots 10^4$, максимальная память $t_0 \geq T$, достигаемая в искусственных длинных линиях задержки, порядка десятка микросекунд, а в линиях задержки на поверхностных акустических волнах — сотен микросекунд.

Согласованный фильтр для пакета импульсов конечной длительности представляется в виде последовательно соединенных фильтра для одиночного импульса (обычно — канал УПЧ) и гребенчатого фильтра (выходной интегратор). Часто сигналы (импульсы сложной формы) можно представить в виде простых; более коротких импульсов, вплотную примыкающих друг к другу и отличающихся по фазе или частоте. Ввиду этого структура согласованных фильтров для импульсов сложной формы примерно такая же, что и фильтров для пакета импульсов конечной длительности.

Рассмотрим два примера импульсов сложной формы. Амплитуда сигнала в пределах длительности импульса полагается постоянной, что исключает весовые множители в отводах фильтра. Фильтры имеют вид линии задержки с отводами, подключенными к суммирующей шине.

Пример 1. Импульс с линейной частотной модуляцией (рис. 2.7,а). Ввиду того, что периоды высокочастотного заполнения импульса при частотной модуляции неодинаковы, синфазное суммирование снимаемых с отводов линии задержки сигналов в момент t_0 обеспечивается, только если отводы расположены неравномерно. При этом наступает максимум выходного сигнала. Вправо и влево от точки t_0 фазы суммируемых сигналов расходятся, ввиду чего амплитуда суммарного (выходного) сигнала быстро падает. Образуются короткий пик сжатого выходного сигнала и боковые интерференционные лепестки по обе стороны от него (на рисунке для наглядности число периодов взято

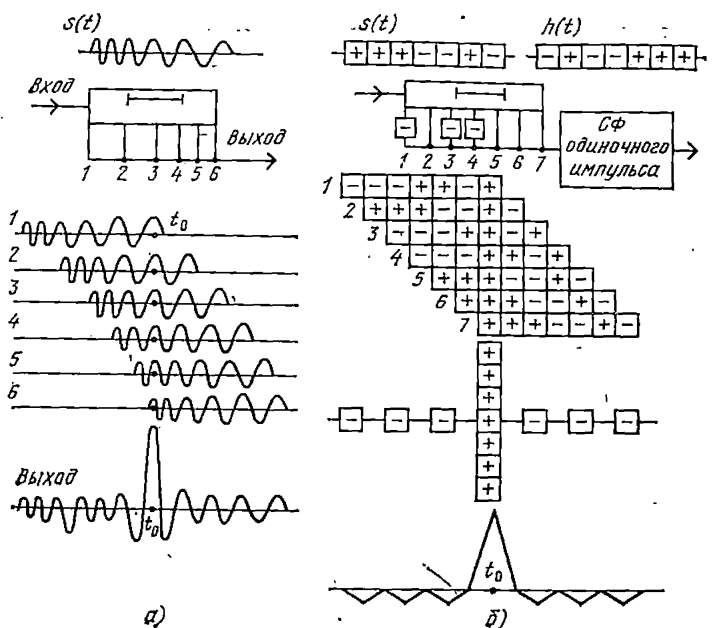


Рис. 2.7. Схемы согласованного фильтра для импульсов сложной формы и процессы в нем

очень малым; на самом деле в пределах узкого пика выходного сигнала укладывается несколько периодов высокочастотного заполнения).

Пример 2. Импульс с псевдослучайной фазовой манипуляцией (рис. 2.7,б). Сигнал такого типа представляется в виде набора высокочастотных тактовых импульсов длительностью τ_c , примыкающих один к другому, фаза которых скачком изменяется на 180° или не изменяется. Комплексную огибающую i -го тактового импульса единичной амплитуды можно представить как $S_i = \exp(j\pi\xi_i t)$, где ξ_i принимает значения 0 или 1 в псевдослучайной последовательности: 0 соответствует нулю фазы тактового импульса или значение $+1$ комплексной огибающей, а 1 — фаза π или -1 комплексной огибающей. Набор всех нулей и единиц в пределах импульса образует код. На рис. 2.7,б показано только чередование знака комплексной огибающей, что соответствует перевороту фазы сигнала на π в тактах 4, 5 и 7 или коду 0001101 (код Баркера, $N=7$).

На примере фазоманипулированного импульса особенно отчетливо прослеживается принцип определения структуры согласованного фильтра. Поскольку фазоманипулированный импульс можно рассматривать как пакет прямоугольных импульсов одинаковой длительности и периода τ_c , его можно построить в виде гребенчатого фильтра с эквидистантными отводами через интервал τ_c , на выходе (или входе) которого ставится согласованный фильтр (СФ) для одиночного тактового импульса длительностью τ_c . Учитывая, что импульсная характеристика фильтра $h(t)$ является зеркальным отражением самого сигнала, переворот фазы следует произвести в отводах 1, 3 и 4 линии задержки. В результате

на выходе гребенчатого фильтра выдается короткий импульс длительностью $\tau_c = T/N$ с боковыми лепестками минимально возможного уровня (что характерно для кода Баркера). На выходе второго фильтра прямоугольные импульсы превращаются в треугольные (на рисунке показана их комплексная огибающая).

2.6. КОГЕРЕНТНЫЙ И НЕКОГЕРЕНТНЫЙ ОБНАРУЖИТЕЛИ СИГНАЛОВ ИЗВЕСТНОЙ ФОРМЫ

При непрерывном обзоре пространства зондирующее колебание представляет собой, как правило, последовательность повторяющихся сигналов известной формы (обычно импульсов). За время облучения объекта, находящегося на данном направлении, от него принимается отраженный или переизлученный сигнал в виде пакета таких импульсов. Огибающая амплитуд импульсов в пакете определяется диаграммой направленности антенны в плоскости сканирования. Для расчетов его можно заменить эквивалентным по энергии пакетом с П-образной огибающей, состоящим из $n = F_n \tau_{обл}$ импульсов одинаковой амплитуды, равной максимальной амплитуде импульса реального пакета. Частота повторения импульсов F_n обычно выбирается из условия однозначного измерения дальности, а время облучения $\tau_{обл} = \theta_{эк} / \omega_A$ определяется угловой скоростью сканирования антенного луча ω_A при заданной ширине $\theta_{эк}$ эквивалентной П-образной диаграммы направленности.

Форма повторяющихся импульсов пакета считается заданной с точностью до неизвестного времени запаздывания и неизвестного доплеровского сдвига частоты несущей. Начальная фаза сигнала неизвестна, что и было учтено при нахождении алгоритма оптимального обнаружения сигнала. Однако закон изменения фазы сигнала от импульса к импульсу может быть как известным, так и неизвестным. В соответствии с этим сигналы подразделяются на когерентные и некогерентные (когерентная и некогерентная последовательности). Соответственно строятся и оптимальные приемники обнаружения — когерентный и некогерентный. В первом случае учитывается известный закон чередования фаз, во втором приходится мириться с незнанием этого закона, что непременно ведет к энергетическому проигрышу.

Когерентный приемник. Когерентными сигналами, закон изменения фазы которых от импульса к импульсу известен, являются сигналы, принимаемые от неподвижных относительно станции объектов, сигналы объектов, относительно которых сама станция перемещается с известной скоростью, а также сигналы любых объектов, если приемник имеет множество каналов приема, рассчитанных на все возможные значения доплеровского сдвига частоты. Необходимо отметить, что понятие «когерентный» относится к отраженным или ответным сигналам даже в том случае,

если сами зондирующие колебания некогерентны, т. е. имеют случайные фазовые соотношения между импульсами. Обычно такие колебания излучаются магнетронными передатчиками. Если станция имеет когерентный гетеродин, фазиремый в каждом периоде посылок импульсами магнетрона, а колебания этого гетеродина подводятся к смесителю приемника, то сигналы промежуточной частоты при указанных выше условиях сохраняют когерентность. Иными словами, закон изменения фазы от импульса к импульсу на промежуточной частоте сохраняется известным, так как при гетеродинировании исключается случайная фаза импульсов магнетрона.

Знание закона чередования фазы в пакете от импульса к импульсу позволяет производить интегрирование импульсов на высокой (промежуточной) частоте, до детектора. Для этого в отводах гребенчатого фильтра включаются фазовращатели, компенсирующие различие импульсов по фазе в момент t_0 , когда сигнал полностью вошел в фильтр. Таким образом, оптимальный приемник когерентной последовательности импульсов (когерентный приемник) состоит из последовательно соединенных согласованного фильтра (и усилителя) для одиночного импульса, гребенчатого фильтра-интегратора и детектора (рис. 2.8, а). Чтобы обеспечить высокое качество обнаружения (максимальное значение $P_D - \Lambda_0 P_F$), энергетическое отношение сигнал-шум $\mathcal{E}/N_0$ на выходе интегратора должно быть большим, поэтому детектор ведет себя как линейное устройство и энергетические потери практически отсутствуют.

Предположим, что пакет состоит из n прямоугольных импульсов одинаковой амплитуды A и длительности τ_n , а дисперсия шума равна σ_n^2 . Тогда на входе интегратора отношение сигнал-шум по напряжению $a = A/\sigma_n$, а по мощности $A^2/2\sigma_n^2 = a^2/2$ (мощность сигнала пропорциональна половине квадрата амплитуды). На выходе интегратора амплитуда максимального импульса равна nA (сумме средних значений), а дисперсия шума — сумме дисперсий $n\sigma_n^2$. Поэтому отношение сигнал-шум на выходе интегратора по напряжению $nA/\sqrt{n}\sigma_n = a\sqrt{n}$, а по мощности

$$\frac{na^2}{2} = \frac{nA^2}{2\sigma_n^2} = \frac{nA^2\tau_n}{2N_0\Delta f\tau_n} = \frac{n}{N_0} \left(\frac{A^2}{2} \tau_n \right) = \frac{n\mathcal{E}}{N_0} = \frac{\mathcal{E}}{N_0},$$

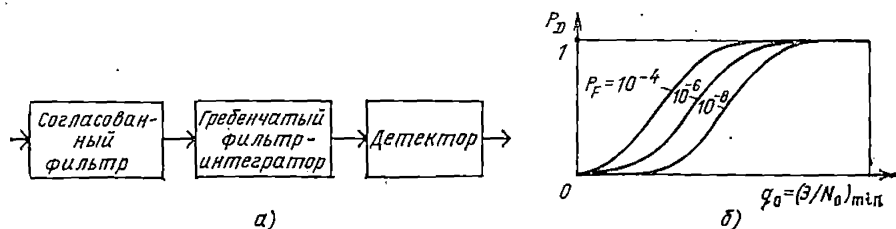


Рис. 2.8. Когерентный прием сигналов:

а — структурная схема приемника; б — характеристики обнаружения

где учтено, что $\sigma^2_n = N_0 \Delta f$, произведение полосы пропускания оптимального приемника на длительность импульса $\Delta f \tau_n = 1$, энергия импульса $\mathcal{E}_1 = A^2 \tau_n / 2$, а энергия сигнала $\mathcal{E} = n \mathcal{E}_1$ (на сопротивлении 1 Ом).

Заданные значения качества обнаружения P_D и P_F обеспечиваются при некотором пороговом отношении сигнал-шум $q_0 = (\mathcal{E}/N_0)_{\min}$ на выходе интегратора. Зависимость между тремя величинами P_D , P_F и q_0 называют характеристикой обнаружения.

Рассчитаем вероятности обнаружения и ложных тревог. Известно, что огибающая шума до детектора v (нормированная по σ_n) распределена по закону Рэлея, и вероятность превышения ею порога обнаружения v_0

$$P_F = \int_{v_0}^{\infty} v e^{-v^2/2} dv = e^{-v_0^2/2},$$

откуда находим сам порог $v_0 = \sqrt{2 \ln(1/P_F)}$. При большом отношении сигнал-шум (а оно должно быть большим, чтобы обеспечить заданное качество обнаружения!) нормированная огибающая максимального сигнала на входе детектора распределена нормально относительно среднего $a\sqrt{n}$, найденного выше. Поэтому вероятность превышения порога $v_0 = \sqrt{2 \ln(1/P_F)}$ огибающей сигнала

$$\begin{aligned} P_D &= \frac{1}{\sqrt{2\pi}} \int_{v_0}^{\infty} e^{-0,5(v - a\sqrt{n})^2} dv = \frac{1}{\sqrt{2\pi}} \int_{v_0 - a\sqrt{n}}^{\infty} e^{-t^2/2} dt = \\ &= \Phi(a\sqrt{n} - v_0) = \Phi\left(a\sqrt{n} - \sqrt{2 \ln\left(\frac{1}{P_F}\right)}\right), \end{aligned}$$

где $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-t^2/2} dt$ — табличный интеграл вероятностей.

Обозначим $y = \Phi(z)$, тогда $z = \Phi^{-1}(y)$ — обратная функция интеграла вероятностей (таблицу этого интеграла нужно читать наоборот — по функции y находить аргумент z). Тогда $\Phi^{-1}(P_D) = a\sqrt{n} - \sqrt{2 \ln(1/P_F)}$. Согласно показанному ранее $a\sqrt{n} = \sqrt{2(\mathcal{E}/N_0)_{\min}} = \sqrt{2q_0}$, поэтому из последнего равенства получаем характеристику обнаружения когерентного сигнала

$$q_0 = 0,5[\Phi^{-1}(P_D) + \sqrt{2 \ln(1/P_F)}]^2. \quad (2.31)$$

Обычно характеристика обнаружения (т. е. зависимость P_D , P_F и q_0) дается как функция $P_D(q_0)$ при различных $P_F = \text{const}$ (рис. 2.8,б).

Таким образом, качество оптимального обнаружения когерентного сигнала известной формы совершенно не зависит от его формы и определяется только энергетическим отношением сигнал-шум q_0 . С вероятностью не меньше заданной P_D при фиксированной вероятности ложных тревог P_F обнаруживаются все сигнала

лы, у которых $\mathcal{E}/N_0 \geq q_0$. В интеграторе в момент максимума достигается накопление всей энергии когерентного сигнала.

Часто при заданном качестве обнаружения оценивают чувствительность приемника через минимальную мощность сигнала на входе приемника $P_{\text{ДРМ min}} = m_p P_{\text{ш}}$, которая должна превышать в m_p раз среднюю мощность шума $P_{\text{ш}} = \sigma_n^2$, чтобы обеспечить заданные P_D и P_F . Величину m_p называют коэффициентом различимости. Согласно введенным выше обозначениям $a^2/2 = q_0/n$, а $m_p = P_{\text{ДРМ min}}/P_{\text{ш}} = a^2/2$, поэтому из формулы (2.31) находим

$$m_p = (1/2n) [\Phi^{-1}(P_D) + \sqrt{2 \ln(1/P_F)}]^2. \quad (2.32)$$

Таким образом, чувствительность когерентного приемника обнаружения повышается пропорционально числу импульсов в пакете.

Некогерентный приемник. В большинстве случаев радиальная скорость наблюдаемых объектов и, следовательно, несущая частота принимаемого сигнала неизвестна, а сам приемник имеет всего один частотный канал настройки. В этих условиях закон изменения фазы сигнала от импульса к импульсу оказывается неизвестным для наблюдателя, т. е. последний имеет дело с некогерентным сигналом (последовательностью импульсов). Так как неизвестные фазовые сдвиги между импульсами невозможно скомпенсировать, интегрирование на высокой или промежуточной частоте производить нельзя: вместо суммирования может произойти взаимное подавление сигналов. Поэтому в оптимальном некогерентном приемнике детектор включается перед гребенчатым фильтром-интегратором (рис. 2.9, а).

Включение детектора перед интегратором существенно изменяет эффективность приема сигнала. Если число интегрируемых импульсов n достаточно велико, то даже при большом отношении сигнал-шум на выходе интегратора, обеспечивающем высокое качество обнаружения, отношение сигнал-шум в каждом импульсе, т. е. на входе детектора, мало. В этих условиях детектор ведет себя как квадратичный: отношение сигнал-шум по напряжению на его выходе равно отношению сигнал-шум по мощности на входе приемника $a^2/2$. После интегратора (как было ранее показано) это отношение возрастает в $\sqrt{n}$ раз:

$$\frac{a^2}{2} \sqrt{n} = \frac{A^2 \sqrt{n}}{2\sigma_n^2} = \frac{1}{\sqrt{n}} \frac{A^2 n}{2\sigma_n^2} = \frac{\mathcal{E}}{N_0 \sqrt{n}} = \frac{q_0}{\sqrt{n}}.$$

Таким образом, энергетическое отношение сигнал-шум на выходе некогерентного приемника, пересчитанное на его вход, в $\sqrt{n}$ раз меньше, чем при когерентном приеме, что эквивалентно уменьшению энергии сигнала в $\sqrt{n}$ раз. Это уменьшение является платой за незнание закона чередования фазы импульсов в пакете (или отказ от многоканального приема по частоте).

Рассчитаем вероятности обнаружения и ложных тревог, полагая число импульсов в пакете большим. Шум на выходе детек-

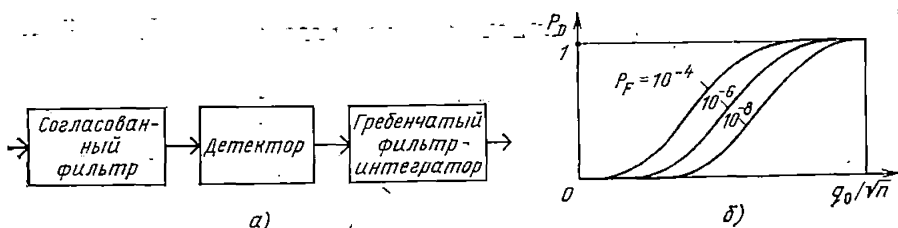


Рис. 2.9. Некогерентный прием сигналов:

а — структурная схема приемника; б — характеристики обнаружения

тора, пропорциональный мощности шума на входе, распределен экспоненциально; среднее значение шума равно его дисперсии или, в нормированном значении, 1. Шумовое напряжение на выходе интегратора как сумма большого числа n слагаемых распределено по нормальному закону. При этом вероятность превышения порога v_0 шумом

$$P_F = \frac{1}{\sqrt{2\pi}} \int_{v_0}^{\infty} e^{-\frac{1}{2}(v-1)^2} dv = \Phi(1-v_0),$$

откуда находим сам порог $v_0 = 1 - \Phi^{-1}(P_F)$. Распределение сигнала относительно вычисленного выше среднего значения $a_2 \sqrt{n}/2 = q_0/\sqrt{n}$ также нормальное, откуда вероятность превышения им порога v_0

$$\begin{aligned} P_D &= \frac{1}{\sqrt{2\pi}} \int_{v_0}^{\infty} e^{-0,5(v-0,5a^2\sqrt{n})^2} dv = \Phi(0,5a^2\sqrt{n}-v_0) = \\ &= \Phi\left[\frac{q_0}{\sqrt{n}} + \Phi^{-1}(P_F) - 1\right]. \end{aligned}$$

Решая эту зависимость относительно аргумента интеграла вероятностей, находим характеристику обнаружения некогерентного сигнала при большом n (рис. 2.9, б):

$$q_0 \approx \sqrt{n} [1 + \Phi^{-1}(P_D) - \Phi^{-1}(P_F)]. \quad (2.33)$$

При малых n характеристика обнаружения имеет промежуточный характер между когерентной и некогерентной, а при $n=1$ различие между когерентным и некогерентным приемом полностью исчезает. Таким образом, при некогерентном приеме отрицательно сказывается использование сигнала в виде пакета импульсов, а не одиночного импульса — дробление энергии сигнала между несколькими импульсами. Только в этом и проявляется влияние вида сигнала при некогерентном обнаружении. Однако в ряде случаев дробление сигнала принципиально необходимо (требование непрерывности обзора, определение угловой координаты объекта при обзоре по огибающей импульсного пакета и т. п.).

Из формулы (2.33) легко определить коэффициент различимости и чувствительность некогерентного приемника. В соответст-

вии с принятыми обозначениями $m_p \sqrt{n} = a^2 \sqrt{n}/2 = q_0 / \sqrt{n}$, откуда согласно (2.33) находим

$$m_p = [1 + \Phi^{-1}(P_D) - \Phi^{-1}(P_F)] / \sqrt{n}. \quad (2.34)$$

В отличие от когерентного чувствительность некогерентного приемника с увеличением числа импульсов растет пропорционально $\sqrt{n}$, а не n , т. е. значительно медленнее.

Необходимо отметить, что полученный результат для некогерентного сигнала является наилучшим, а сам приемник оптимальным. В случае же некогерентного приема когерентного сигнала имеем прямой проигрыш, так как такой приемник уже не является оптимальным.

Все сказанное выше относится к минимальному обнаруживаемому пороговому сигналу. При приеме сигналов большой интенсивности детектор ведет себя как линейное устройство, и поэтому эффективность когерентного и некогерентного приемников становится одинаковой.

2.7. ЭНЕРГЕТИЧЕСКИЙ ПОТЕНЦИАЛ И ДАЛЬНОСТЬ ДЕЙСТВИЯ РАДИОСИСТЕМ

Обнаружение сигнала либо предшествует измерению его параметров, либо производится одновременно с измерением. Поэтому энергетический потенциал радиосистем обычно определяется из решения задачи обнаружения.

Энергетический потенциал радиосистемы $Q = \mathcal{E}_{изл} A_{прм}$ характеризуется двумя энергетическими параметрами: временным — излучаемой за цикл обзора энергией сигнала $\mathcal{E}_{изл}$ и пространственным — эффективной площадью приемной антенны $A_{прм}$. Минимально необходимый энергетический потенциал существенно различен в системах, работающих по прямому и отраженному сигналам.

В радиосистемах, использующих прямой или переизлученный сигнал, излучаемая за цикл обзора энергия $\mathcal{E}_{изл}$ (при равномерном рассеянии ее в пределах заданного телесного угла рабочей зоны системы Ω_0) создает в точке приема, удаленной на расстояние R_m , соответствующее заданной дальности действия, плотность потока энергии $\mathcal{E}_{изл} / (\Omega_0 R_m^2)$. Тогда энергия сигнала, воспринимаемая приемной антенной с эффективной площадью $A_{прм}$ в течение цикла обзора, равна $\mathcal{E}_{изл} A_{прм} / \Omega_0 R_m^2$, а энергетическое отношение сигнал-шум в оптимальном приемнике обнаружения должно удовлетворять условию

$$\mathcal{E}_{изл} A_{прм} / (\Omega_0 R_m^2 N_0) \geq q_0,$$

где N_0 — спектральная плотность шумов приемника, пересчитанная на его вход, а $q_0 = (\mathcal{E}/N_0)_{\min}$ — минимальное отношение сигнал-шум, при котором обеспечивается заданное качество обнаружения сигнала в соответствии с формулами (2.31) и (2.33).

Из этого условия находим минимально необходимый энергетический потенциал радиосистем, работающих по прямому или переизлученному сигналам,

$$Q = \mathcal{E}_{\text{изл}} A_{\text{прм}} = \Omega_0 R^2 M N_0 q_0. \quad (2.35)$$

Заметим, что при работе по переизлученному сигналу энергетический потенциал каналов запроса и ответа должен быть одинаков (условие энергетического баланса системы), иначе допускаются неоправданные энергетические затраты в том канале, где он больше.

В радиосистемах, использующих отраженный сигнал, плотность потока излучаемой за цикл обзора энергии (мощности) в точке расположения объекта, удаленного на расстояние R_M , составляет $\Pi_1 = \mathcal{E}_{\text{изл}} / (\Omega_0 R^2 M)$, если она равномерно рассеивается в пределах телесного угла Ω_0 рабочей зоны. Плотность потока отраженной энергии (мощности) Π_2 в точке приема, совмещенной с точкой излучения, находится из определения эффективной площади рассеяния (ЭПР) объекта (цели)

$$\sigma_{\text{ц}} = 4\pi R^2 M \Pi_2 / \Pi_1,$$

откуда определяем энергию сигнала, перехватываемую приемной антенной за цикл обзора,

$$\Pi_2 A_{\text{прм}} = \mathcal{E}_{\text{изл}} A_{\text{прм}} \sigma_{\text{ц}} / (4\pi \Omega_0 R^4 M).$$

Отношение сигнал-шум в оптимальном приемнике обнаружения должно удовлетворять условию

$$\mathcal{E}_{\text{изл}} A_{\text{прм}} \sigma_{\text{ц}} / (4\pi \Omega_0 R^4 M N_0) \geq q_0,$$

где q_0 по-прежнему определяется формулами (2.31) или (2.33) в зависимости от того, когерентный или некогерентный сигнал принимается. Из последнего условия находим минимально необходимый энергетический потенциал радиосистем, работающих по отраженному сигналу,

$$Q = \mathcal{E}_{\text{изл}} A_{\text{прм}} = 4\pi \Omega_0 R^4 M N_0 q_0 / \sigma_{\text{ц}}. \quad (2.36)$$

Таким образом, в общем случае минимально необходимый энергетический потенциал радиосистемы независимо от способа обзора определяется заданными размерами рабочей зоны ($\Omega_0 R^2 M$), спектральной плотностью собственных шумов приемника и минимальным отношением сигнал-шум, обеспечивающим заданное качество обнаружения. При работе по отраженному сигналу дополнительно учитывается ЭПР объекта и площадь сферы $4\pi R^2 M$, в пределах которой рассеяние объекта полагается изотропным. Фактический энергетический потенциал системы должен компенсировать потери в приемно-передающем тракте, затухание в атмосфере и другие энергетические потери.

Как следует из формул (2.35) и (2.36), энергетический потенциал при работе по прямому сигналу пропорционален квадрату дальности действия системы, а при работе по отраженному сиг-

налу — четвертой степени. В последнем случае дальность действия с увеличением энергетического потенциала растет очень медленно: для удвоения дальности действия требуется 16-кратное его увеличение. При работе по отраженному сигналу энергетический потенциал соответствует заданной вероятности обнаружения нефлуктуирующего объекта, учитываемой формулой (2.31) или (2.33). Вероятность обнаружения флуктуирующего объекта учитывается более сложными формулами, в которые входят распределения вероятностей флуктуирующего сигнала и шума.

В заключение заметим, что при расчете энергетического потенциала $Q = \mathcal{E}_{\text{изл}} A_{\text{прм}}$ только из условия оптимального обнаружения соотношение между составляющими $\mathcal{E}_{\text{изл}}$ и $A_{\text{прм}}$ может быть любым. Неоднозначность раскрывается после выбора частоты несущей и обеспечения требуемой угловой разрешающей способности (точности), о которой речь пойдет ниже.

Глава 3. ПРОСТРАНСТВЕННО-ВРЕМЕННОЕ ОПИСАНИЕ ПРОЦЕССОВ ПРИ ОПТИМАЛЬНОМ ПРИЕМЕ

3.1. ЭЛЕКТРОМАГНИТНАЯ ВОЛНА КАК ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ СИГНАЛ

В обзорно-измерительных радиосистемах сигналом является не электрическое колебание в цепях приемника, а электромагнитная волна, перехватываемая антенной. Именно к ней относятся такие важные измеряемые параметры, как направление прихода, характеризующее угловые координаты наблюдаемого объекта, или относительная угловая скорость перемещения объекта. Приходящая от объекта электромагнитная волна $s(t, r)$ является в общем случае функцией не только времени t , но и координат точек приема r , т. е. является пространственно-временным континуумом. Это более емкое понятие, чем электрический сигнал. Поэтому и описание процесса приема волны значительно усложняется.

Если высокочастотный электрический сигнал имеет четыре информативных параметра — амплитуду, частоту, фазу и время прихода, то электромагнитная волна — восемь: добавляются два параметра, характеризующие направление прихода волны, и два параметра, характеризующие ее поляризационную структуру. Эти параметры являются носителями информации о положении и скорости объекта, его конфигурации, ориентации и материале, из которого он состоит.

В обзорно-измерительных радиосистемах всех четырех рассмотренных ранее видов основной интерес представляют координаты объекта и их производные. Геометрические и отражающие свойства объектов выявляются лишь в собственно радиолокационных системах, специально предназначенных для идентификации (распознавания) объектов путем анализа тонкой структуры отраженной волны, включая поляризационную. Однако такие системы применяются пока ограниченно. Поскольку поляризационная структура волны практически не влияет на процесс измерения координат и их производных, в обзорно-измерительных радиосистемах оказалось возможным отказаться от ее учета. Отказ от учета поляризации в собственно радиолокационных системах проявляется в основном в том, что теряется часть энергии, заключенная в ортогональной поляризационной составляющей сигнала, возникающей при отражении от объекта. Обычно этой составляющей пренебрегают, поскольку она невелика (за редким исключением), а для ее приема и использования приходится существенно усложнять структуру системы, включая антенну.

Отказ от учета поляризационной структуры волны существенно упрощает описание как входного сигнала, так и процедуры его оптимального приема. Наиболее важной причиной этого отказа является то, что к поляризованным отраженным сигналам неприменимо понятие сигнала известной формы, справедливое без учета поляризации для всех объектов, размеры которых малы по сравнению с размерами объема пространства, разрешаемого системой. Поляризационная же структура отраженной волны у объектов различной конфигурации различна, даже если они малы по сравнению с разрешаемым объемом. Поэтому нельзя говорить о едином для всех объектов поляризованном отраженном сигнале точечной цели (сигнале известной формы) и строить единый для всех объектов оптимальный приемник таких сигналов.

Поэтому воспользуемся в основном скалярным описанием волны, игнорирующим ее поляризационную структуру, полагая, что от объекта принимается волна той же поляризации, что у излучаемой. (Краткое описание особенностей поляризованных волн дается в § 5.6.) В системах, использующих сигнал прямого излучения или переизлученный, поляризационная структура принимаемой волны почти всегда известна (если пренебречь ее искажением при распространении). В таких системах без потери общности можно также ограничиться скалярными представлениями. В звуколокации, где используются объемные акустические волны, скалярное описание волны вообще является исчерпывающим.

Для скалярного описания волны, воспринимаемой раскрывом антенны, необходимо знать две скалярные входные характеристики системы: временную — сигнал и пространственную — апертуру антенны. Для определения структуры оптимального приемника как пространственно-временного фильтра или коррелятора и описания выходного сигнала этого приемника дополнительно не-

обходимо знать также статистические характеристики шума в рамках корреляционной теории.

Применительно к сигналу как функции только времени структура оптимального приемника и его выходной сигнал широко известны и кратко описаны в предыдущей главе. Так как приходящая волна в каждой точке раскрыва антенны описывается чисто временной зависимостью, изложенные ранее положения теории оптимального приема распространяются и на пространственно-временной входной сигнал как совокупность обычных, временных сигналов по всем точкам раскрыва антенны. Ниже приводятся краткие сведения о пространственной и временной входных характеристиках радиосистем и пространственно-временное описание входного и выходного сигналов оптимального приемника. На этой основе в дальнейшем определяется пространственно-временная структура оптимального приемника и основные выходные характеристики радиосистем по максимально достижимой точности и разрешающей способности.

3.2. СИГНАЛ КАК ВРЕМЕННАЯ ВХОДНАЯ ХАРАКТЕРИСТИКА РАДИОСИСТЕМЫ

Сигнал известной формы, используемый в радиосистеме, может быть представлен двумя равноценными функциями: комплексной огибающей $S(t)$ или комплексным спектром $S(f)$. Эти функции взаимно однозначно связаны между собой прямым и обратным преобразованием Фурье

$$S(f) = \int_{-\infty}^{\infty} S(t) e^{-j2\pi ft} dt; \quad S(t) = \int_{-\infty}^{\infty} S(f) e^{j2\pi ft} df. \quad (3.1)$$

Здесь и далее, кроме специально оговоренных случаев, имеется в виду двусторонний спектр, содержащий и положительные, и отрицательные частоты (физический смысл знака частоты будет раскрыт в гл. 5). В таком представлении прямое и обратное преобразования Фурье различаются только знаком в ядре преобразования.

Обратное преобразование Фурье описывает операцию синтеза комплексной огибающей сигнала $S(t)$ любой формы из гармонических колебаний $\exp(j2\pi ft)$ частоты f , амплитуда и фаза которых характеризуется комплексным множителем $S(f) df = A(f) \times \exp[j\varphi(f)] df$. Прямое преобразование Фурье отображает операцию спектрального анализа сигнала $S(t)$ — вычисление корреляции между комплексными огибающими этого сигнала и гармонического колебания $\exp(j2\pi ft)$ с частотой f и единичной амплитудой (знак минус у экспоненты объясняется тем, что при образовании корреляционных функций от комплексных сигналов один из них берется комплексно-сопряженным). Так как все спектральные составляющие сигнала, кроме одной, не коррелированы с гармоническим колебанием данной частоты f , полученный коэф-

коэффициент корреляции пропорционален комплексной амплитуде $S(f)df$ гармоники этой частоты.

Возможны и другие разложения сигнала по ортогональным функциям, кроме гармонических. Однако для гармонических функций характерно, что при прохождении через любую линейную систему их форма не меняется, изменяется только амплитуда и фаза данной гармоники. Это и определило преимущественное использование спектрального представления по Фурье. Следует отметить, что преобразования Фурье применимы только к сигналам с конечной энергией. Этому условию удовлетворяют все реальные сигналы. Однако при математическом описании процессов иногда используют идеализированные функции, вообще говоря, не удовлетворяющие этому условию, но тем не менее к ним также применяется преобразование Фурье: это гармоническое колебание, бесконечное по времени, и дельта-функция, бесконечная по спектру. Однако, если сначала взять интервалы конечными, а потом произвести предельный переход, все становится на свое место.

Ниже рассматриваются некоторые свойства функций и их фурье-преобразований на примере временных сигналов. Эти свойства в дальнейшем используются при описании более общих зависимостей.

Комплексная огибающая сигнала $S(t) = A(t) \exp[j\varphi(t)]$ описывает амплитудную и фазовую (частотную) модуляцию сигнала. При чисто амплитудной модуляции модулирующей функцией $v(t)$ пропорциональна амплитуда сигнала и, следовательно, сама огибающая $S(t) = A(t) = kv(t)$. Поэтому спектр комплексной огибающей сохраняет форму спектра модулирующей функции, только переносится в область высоких частот. В этом смысле амплитудную модуляцию можно считать линейной операцией, как и операции преобразования частоты и «линейного» детектирования. При фазовой модуляции $\varphi(t) = kv(t)$ комплексная огибающая $S(t) = A \exp[jkv(t)]$ нелинейно связана с модулирующей функцией $v(t)$. Поэтому спектр комплексной огибающей отличается от спектра модулирующей функции — появляются высшие гармоники модулирующих частот. Еще сложнее обстоит дело при частотной модуляции, когда модулирующей функцией пропорциональна частота $f(t) = kv(t)$ и комплексная огибающая имеет вид

$$S(t) = A \exp \left\{ j \left[\varphi_0 + 2\pi \int_0^t f(t') dt' \right] \right\} = \\ = A \exp \left\{ j \left[\varphi_0 + 2\pi k \int_0^t v(t') dt' \right] \right\},$$

так как фаза есть интеграл от частоты плюс начальная фаза φ_0 . Следовательно, фазовую и частотную модуляцию относят к нелинейным преобразованиям, при которых спектр сигнала отличается от спектра модулирующей функции. Однако при малой глу-

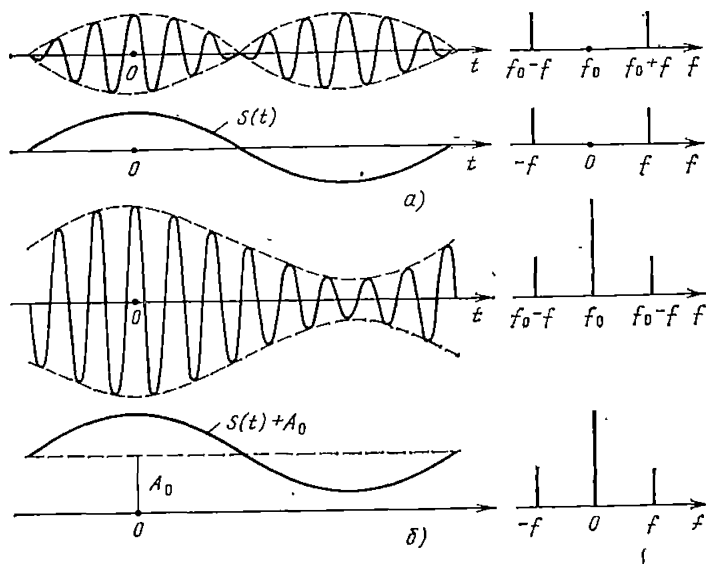


Рис. 3.1. Высокочастотные сигналы, их гармоническая огибающая и спектры сигналов и огибающих

бине фазовой модуляции $\exp[j\varphi(t)] \approx 1 + j\varphi(t)$ спектральный состав сигнала и модулирующей функции примерно одинаков, а при частотной модуляции он лишь убывает обратно пропорционально частоте модуляции (интегрирование функции соответствует делению спектра на $j\omega$).

При амплитудной модуляции простейшей является гармоническая огибающая (вещественная функция). Такую огибающую имеет как сигнал частоты f_0 , модулированный по амплитуде гармоникой частоты f (рис. 3.1,б), так и биения двух гармонических сигналов одинаковой амплитуды, разнесенные по

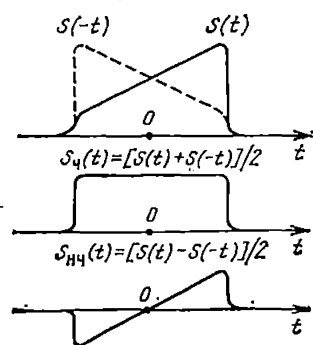


Рис. 3.2 Представление асимметричного сигнала в виде суммы четной и нечетной функций

частоте на интервал $2f$ (рис. 3.1,а). Биения можно рассматривать как амплитудно-модулированное колебание, у которого удалена несущая частота f_0 (для удобства сравнения на рис. 3.1 приведен односторонний спектр высокочастотных колебаний и двусторонний спектр огибающей). Соответственно комплексную огибающую биений можно рассматривать как огибающую амплитудно-модулированного сигнала, у которой удалена постоянная составляющая. Ввиду этого комплексная огибающая биений $S(t)$ в отличие от огибающей амплитуды $A(t)$ через полупериод меняет знак, что соответствует перевороту фазы на π у биений высокочастотного заполнения. Таким образом, комплексная огибающая даже при чисто амплитудной модуляции учитывает не только

амплитудные, но и фазовые соотношения в сигнале тем, что меняет знак, когда фаза высокочастотного заполнения изменяется на π .

В общем случае комплексная огибающая амплитудно-модулированного сигнала является вещественной функцией, характер спектра которой зависит от симметрии огибающей. У четной-симметричной (четной) вещественной огибающей $S(-t) = S(t)$ спектр также вещественный и четный:

$$\begin{aligned} S(f) &= \int_{-\infty}^{\infty} S(t) e^{-j 2\pi f t} dt = \int_0^{\infty} S(t) [e^{-j 2\pi f t} + e^{j 2\pi f t}] dt = \\ &= 2 \int_0^{\infty} S(t) \cos 2\pi f t dt = S_{\text{ч}}(f), \end{aligned} \quad (3.2)$$

так как сумма только четных вещественных функций (косинусов) с вещественными весами $S(t)$ дает четную функцию (здесь интеграл был разбит на два — отдельно для положительных и отрицательных t). У нечетно-симметричной функции $S(-t) = -S(t)$ спектр также нечетно-симметричный, но чисто мнимый:

$$\begin{aligned} S(f) &= \int_{-\infty}^{\infty} S(t) e^{-j 2\pi f t} dt = \int_0^{\infty} S(t) [e^{-j 2\pi f t} - e^{j 2\pi f t}] dt = \\ &= -j 2 \int_0^{\infty} S(t) \sin 2\pi f t dt = -j S_{\text{нч}}(f), \end{aligned} \quad (3.3)$$

так как с вещественными весами $S(t)$ суммируются только нечетные вещественные функции (синусы). Множитель $-j = \exp(-j\pi/2)$ у спектра огибающей означает, что высокочастотное заполнение сигнала сдвинуто по фазе на $-\pi/2$ относительно заполнения четной огибающей ($-\sin 2\pi f_0 t$ вместо $\cos 2\pi f_0 t$).

Асимметричная вещественная огибающая всегда может быть представлена в виде суммы (рис. 3.2)

$$S(t) = S_{\text{ч}}(t) + S_{\text{нч}}(t), \quad (3.4)$$

где $S_{\text{ч}}(t) = 0,5[S(t) + S(-t)]$, $S_{\text{нч}}(t) = 0,5[S(t) - S(-t)]$ — четная и нечетная вещественные функции, определяемые соответственно как полусумма и полуразность огибающей сигнала $S(t)$ и ее зеркального отражения $S(-t)$. Согласно доказанному выше спектр такой огибающей комплексный:

$$S(f) = S_{\text{ч}}(f) - j S_{\text{нч}}(f), \quad (3.5)$$

где $S_{\text{ч}}(f)$ и $S_{\text{нч}}(f)$ — четная и нечетная вещественные функции (3.2) и (3.3).

От комплексного спектра $S(f)$ легко перейти к амплитудному $A(f)$ и фазовому $\varphi(f)$ спектрам:

$$S(f) = A(f) e^{-j \varphi(f)}, \quad (3.6)$$

где $A(f) = |S(f)| = \sqrt{S_{\text{ч}}^2(f) + S_{\text{нч}}^2(f)}$;

$$\varphi(f) = \arg S(f) = \arctg[S_{\text{нч}}(f)/S_{\text{ч}}(f)].$$

В частности, у четных вещественных функций $A(f) = S(f)$, $\varphi = 0$; у нечетных $A(f) = |S_{\text{нч}}(f)|$, $\varphi(f) = 0$, $\varphi(-f) = \pi$.

Таким образом, амплитудный спектр любой вещественной функции всегда четный, так как даже нечетная функция $S_{\text{нч}}(-f) = -S_{\text{нч}}(f)$ дает под корнем выражения (3.6) квадрат $S_{\text{нч}}^2(-f) = S_{\text{нч}}^2(f)$, являющийся четной функцией.

Прямое и обратное преобразование Фурье симметричны: все что справедливо при переходе от функции к спектру, справедливо и при переходе от спектра к функции. Поэтому четному вещественному спектру соответствует четная вещественная функция, нечетному вещественному спектру — нечетная чисто мнимая функция, симметричный вещественный спектр имеет любая комплексная функция, действительная часть которой четная, а мнимая нечетная. Указанные комплексные функции имеют также четную огибающую (модуль комплексной огибающей). Иными словами, если спектр функции

$$S(f) = S_{\text{ч}}(f) + S_{\text{нч}}(f), \quad (3.7)$$

где обе составляющие вещественные, то сама функция

$$S(t) = S_{\text{ч}}(t) - j S_{\text{нч}}(t), \quad (3.8)$$

комплексная, ее модуль

$$A(t) = |S(t)| = \sqrt{S_{\text{ч}}^2(t) + S_{\text{нч}}^2(t)} = A(-t) \quad (3.9)$$

имеет четную симметрию, а аргумент

$$\varphi(t) = \arg S(t) = \arctg [S_{\text{нч}}(t)/S_{\text{ч}}(t)] = -\varphi(-t) \quad (3.10)$$

нечетную.

Из сказанного, в частности, следует, что корреляционная функция любого действительного сигнала как обратное преобразование Фурье от его энергетического спектра $0,5A^2(f) = 0,5|S(f)|^2$ всегда вещественна и четна, так как амплитудный и энергетический спектры всегда вещественны, а в данном случае и четны. Комплексные сигналы, обладающие асимметричными амплитуд-

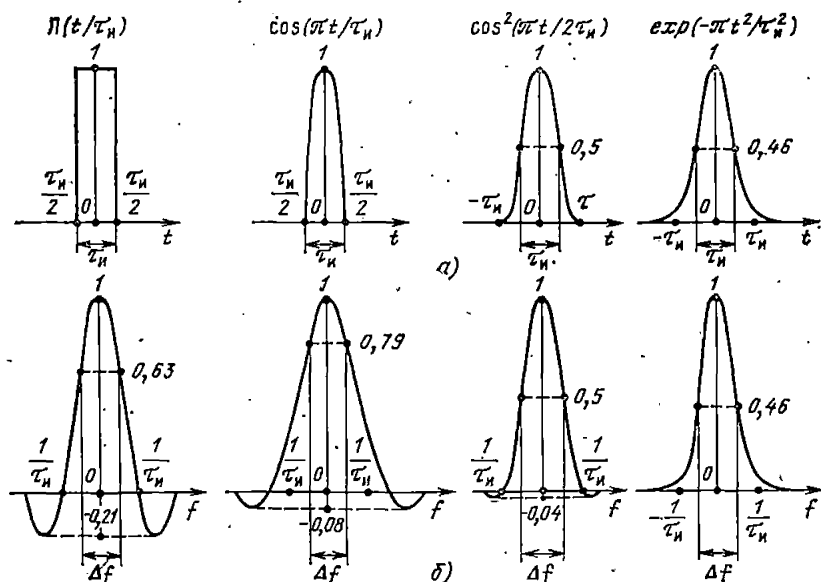


Рис. 3.3. Вещественные функции (а) и их спектры (б)

ным и энергетическим спектрами, имеют комплексную корреляционную функцию, но всегда с четной огибающей.

Заметим также, что преобразованиями Фурье связаны не только сигнал и его спектр, но и импульсная характеристика (временная характеристика) и функция передачи (частотная характеристика) фильтра:

$$H(t) = \int_{-\infty}^{\infty} H(f) e^{j 2\pi f t} dt, \quad H(f) = \int_{-\infty}^{\infty} H(t) e^{-j 2\pi f t} dt \quad (3.11)$$

(применительно к комплексным огибающим, а не к реальным сигналам, обе приведенные функции в общем случае комплексны).

Для отработки навыков нахождения спектров конкретных функций и выявления общих закономерностей рассмотрим несколько частных примеров. При этом ограничимся только четными вещественными функциями (рис. 3.3), что упрощает решение поставленной задачи.

Пример 1. Прямоугольный импульс $P(t/\tau_n)$ амплитудой A и длительностью τ_n

$$S(t) = A, \quad -\tau_n/2 \leq t \leq \tau_n/2.$$

Его спектр

$$S(f) = A \int_{-\tau_n/2}^{\tau_n/2} e^{-j 2\pi f t} dt = \frac{A}{-j 2\pi f} e^{-j 2\pi f t} \Big|_{-\tau_n/2}^{\tau_n/2} = A \tau_n \frac{\sin(\pi f / \Delta f)}{\pi f / \Delta f}$$

имеет ширину $\Delta f = 1/\tau_n$ на уровне 0,63.

Пример 2. Импульс вида полупериода косинусоиды амплитудной A и длительностью τ_n на нулевом уровне

$$S(t) = A \cos\left(\frac{\pi t}{\tau_n}\right) = \frac{A}{2} (e^{j \pi t / \tau_n} + e^{-j \pi t / \tau_n}), \quad -\frac{\tau_n}{2} \leq t \leq \frac{\tau_n}{2}.$$

Его спектр

$$\begin{aligned} S(f) &= \frac{A}{2} \int_{-\tau_n/2}^{\tau_n/2} [e^{-j 2\pi (f - 1/2\tau_n) t} + e^{-j 2\pi (f + 1/2\tau_n) t}] dt = \\ &= \frac{A \tau_n}{2} \left[\frac{\sin \pi (f / \Delta f - 1/2)}{\pi (f / \Delta f - 1/2)} + \frac{\sin \pi (f / \Delta f + 1/2)}{\pi (f / \Delta f + 1/2)} \right] \end{aligned}$$

имеет ширину $\Delta f = 1/\tau_n$ на уровне 0,79.

Пример 3. Импульс вида квадрата полупериода косинусоиды амплитудой A и длительностью τ_n на уровне 0,5

$$\begin{aligned} S(t) &= A \cos^2\left(\frac{\pi t}{2\tau_n}\right) = \frac{A}{2} \left(1 + \cos \pi \frac{t}{\tau_n}\right) = \\ &= \frac{A}{4} (2 + e^{j \pi t / \tau_n} + e^{-j \pi t / \tau_n}), \quad -\tau_n \leq t \leq \tau_n. \end{aligned}$$

Его спектр

$$S(f) = \frac{A}{4} \int_{-\tau_n}^{\tau_n} [2e^{-j 2\pi f t} + e^{-j 2\pi (f - 1/2\tau_n) t} + e^{-j 2\pi (f + 1/2\tau_n) t}] dt = \\ = \frac{A \tau_n}{2} \left[2 \frac{\sin 2\pi (f/\Delta f)}{2\pi (f/\Delta f)} + \frac{\sin 2\pi (f/\Delta f - 1/2)}{2\pi (f/\Delta f - 1/2)} + \frac{\sin 2\pi (f/\Delta f + 1/2)}{2\pi (f/\Delta f + 1/2)} \right]$$

имеет ширину $\Delta f = 1/\tau_n$ также на уровне 0,5.

Пример 4. Колоколообразный импульс амплитудой A и длительностью τ_n на уровне $\exp(-\pi/4) \approx 0,46$

$$S(t) = A \exp(-\pi t^2/\tau_n^2), \quad -\infty < t < \infty.$$

Его спектр

$$S(f) = A \int_{-\infty}^{\infty} \exp\left(\frac{-\pi t^2}{\tau_n^2} - j 2\pi f t\right) dt = A \tau_n \exp\left(\frac{-\pi f^2}{\Delta f^2}\right)$$

имеет ширину $\Delta f = 1/\tau_n$ также на уровне 0,46 (здесь показатель экспоненты был дополнен до полного квадрата $(-\pi t^2/\tau_n^2 - j 2\pi f t + \pi f^2 \tau_n^2) - \pi f^2 \tau_n^2 = -\pi(t/\tau_n + j f \tau_n)^2 - \pi f^2 \tau_n^2$, а интеграл от экспоненты с первым слагаемым дает τ_n).

Общим для всех сигналов является то, что: 1) спектральная плотность пропорциональна площади импульса (множитель $A \tau_n$); 2) уровень боковых лепестков тем меньше, чем более плавно спадает функция (у спектров приведенных функций максимальный уровень боковых лепестков составляет соответственно 0,21, 0,08, 0,04 и 0); 3) между длительностью вещественной функции τ_n и шириной ее спектра Δf существует обратная зависимость

$$\Delta f = 1/\tau_n, \quad (3.12)$$

так что при уменьшении длительности в несколько раз спектр расширяется во столько же раз. Однако соотношение (3.12) справедливо, если только протяженность функции и ширину спектра отсчитывать на указанных выше уровнях, различных для разных функций. Если же взять общепринятый уровень отсчета 0,71, то коэффициент в формуле (3.12) будет несколько отличаться от единицы — также по-разному у разных функций. Однако существует интегральная, эффективная протяженность функции и ее спектра, не зависящая от уровня отсчета (она будет введена далее), для которой соотношение (3.12) справедливо всегда независимо от вида функции.

Вещественный характер спектра приведенных функций указывает на то, что отдельные спектральные составляющие, суммируясь в некоторый момент времени $t=0$ синфазно, дают в этот момент максимально возможное значение функции. Так как частота отдельных спектральных составляющих различна, в стороне от максимума фазы гармонических слагаемых расходятся и функция спадает тем быстрее, чем шире спектр. Таким образом, вещественный спектр обеспечивает максимально возможную амплитуду

литуду и минимальную длительность сигнала в соответствии с формулой (3.12).

При комплексном спектре с нелинейной фазочастотной характеристикой пик функции убывает по высоте и расширяется, так как гармонические составляющие ни в один момент времени не дают синфазного суммирования, а дополнительное расхождение фаз гармоник в стороне от максимума проявляется слабее.

Для сигналов с комплексным спектром (частотно- и фазомодулированных) справедливо следующее соотношение между их длительностью и шириной спектра:

$$\Delta f = N/\tau_n, \quad (3.13)$$

где $N = \Delta f \tau_n$ — база сигнала. Как уже отмечалось, сигналы с базой $N = 1$ называются простыми, а при $N \gg 1$ — сложными. Последние в согласованном фильтре, корректирующем фазу гармоник, можно сжать в N раз — до длительности $\tau_c = 1/\Delta f$, удовлетворяющей условию (3.12).

Рассмотрим в качестве примера сигнала сложной формы колоколообразный импульс амплитуды A и длительностью τ_n на уровне 0,46 с линейной частотной модуляцией (ЛЧМ) $f(t) = -at$, где a — крутизна частотной модуляции, что соответствует квадратичному фазовому сдвигу $\varphi(t) = -\pi a t^2$:

$$S(t) = A \exp(-\pi t^2/\tau_n^2 - j \pi a t^2).$$

Его спектр

$$S(f) = A \int_{-\infty}^{\infty} \exp\left(-\frac{\pi t^2}{\tau_n^2} - j \pi a t^2 - j 2 \pi f t\right) dt.$$

Дополнив показатель экспоненты до полного квадрата, получим

$$\begin{aligned} & -\pi \left[(1 + j a \tau_n^2) \frac{t^2}{\tau_n^2} + 2 f t - \frac{f^2 \tau_n^2}{1 + j a \tau_n^2} \right] - \pi \frac{f^2 \tau_n^2}{1 + j a \tau_n^2} = \\ & = -\pi \left(\sqrt{1 + j a \tau_n^2} \frac{t}{\tau_n} + j \frac{f \tau_n}{\sqrt{1 + j a \tau_n^2}} \right)^2 - \pi \frac{f^2 \tau_n^2 - j a f a f^2 \tau_n^4}{1 + a^2 \tau_n^4}. \end{aligned}$$

Экспонента со вторым слагаемым в показателе определяет форму спектра (она выходит из-под знака интеграла), а сам интеграл дает некоторый множитель $K \tau_n$. Таким образом, такой сигнал имеет комплексный спектр вида

$$\begin{aligned} S(f) &= K A \tau_n \exp\left(-\frac{\pi f^2 \tau_n^2}{1 + a^2 \tau_n^4} - j \frac{\pi a f^2 \tau_n^4}{1 + a^2 \tau_n^4}\right) = \\ &= K A \tau_n \exp\left(-\frac{\pi f^2}{\Delta f^2} - j \frac{\pi a \tau_n^2 f^2}{\Delta f^2}\right), \end{aligned} \quad (3.14)$$

где

$$\Delta f = \sqrt{1 + a^2 \tau_n^4} / \tau_n = N/\tau_n \quad (3.15)$$

— ширина спектра на уровне 0,46, а $N = \sqrt{1 + a^2 \tau_n^2}$ — база ЛЧМ сигнала (при отсутствии частотной модуляции $N=1$, а при глубокой линейной частотной модуляции $N \approx a \tau_n^2$ и ширина спектра $\Delta f \approx a \tau_n$ определяется только частотной модуляцией). Такой импульс можно сжать в N раз до длительности

$$\tau_c = \frac{1}{\Delta f} = \frac{1}{\sqrt{1/\tau_n^2 + a^2 \tau_n^2}} \approx \frac{1}{a \tau_n}, \quad (3.16)$$

обратной длительности исходного сигнала.

Комплексный спектр ЛЧМ сигнала (3.14) можно представить в виде амплитудного и фазового спектра

$$A(f) = K A \tau_n \exp(-\pi f^2 / \Delta f^2); \quad \varphi(f) = -\pi a \tau_n^2 f^2 / \Delta f^2. \quad (3.17)$$

При глубокой частотной модуляции ($\Delta f \approx a \tau_n$) амплитудный спектр $A(f) \approx K A \tau_n \exp(-\pi f^2 / a^2 \tau_n^2)$ подобен огибающей ЛЧМ сигнала $A(t) = A \exp(-\pi t^2 / \tau_n^2)$ и в соответствии с законом частотной модуляции $f = at$ получается из огибающей простой заменой $t = f/a$. Это свойство ЛЧМ сигналов появляется при любой форме огибающей, а не только колоколообразной. Его можно использовать для воспроизведения спектра высокочастотных сигналов преобразованием частоты с помощью гетеродинного ЛЧМ напряжения и пропускания преобразованного сигнала через фазокорректирующий фильтр. После преобразования частоты сигнал становится линейно-модулированным по частоте, и его спектр по аналогии с (3.17)

$$S_{\text{ЛЧМ}}(f) = A(f/a) \exp j \varphi(f/a)$$

по модулю повторяет огибающую исходного сигнала $A(t)$. После фазокорректирующего фильтра с фазочастотной характеристикой $-\varphi(f/a)$ спектр становится вещественным. Поскольку этот спектр повторяет огибающую исходного сигнала, то сам выходной сигнал повторяет спектр огибающей этого сигнала, его можно наблюдать на осциллооскопе.

Устранение фазовых сдвигов между спектральными составляющими в фазокорректирующем фильтре создает эффект сжатия сигнала. Поэтому длительность выходного сигнала фильтра τ_c , т. е. ширина спектра исходного сигнала, обратна длительности этого сигнала τ_n , что раскрывает физический смысл соотношения (3.16).

3.3. АПЕРТУРА АНТЕННЫ КАК ПРОСТРАНСТВЕННАЯ ВХОДНАЯ ХАРАКТЕРИСТИКА РАДИОСИСТЕМЫ

Рассмотрим вначале линейную антенну с комплексным (в общем случае) распределением поля по раскрыву $G(x)$, причем $|G(x)|_{\max} = G(0) = 1$. На раскрыв шириной d под углом α к нормали падает плоская гармоническая волна частоты f_0 от удаленного источника (рис. 3.4). Разность хода волны между центром раскрыва 0 и произвольной точкой x на нем $\Delta R = x \sin \alpha = x u_x$, где

$u_x = \sin \alpha = \cos \vartheta_x$ — направляющий косинус угла относительно оси x . Разность хода вызывает относительный фазовый сдвиг сигнала $2\pi \Delta R / \lambda = 2\pi x u_x / \lambda$, где $\lambda = c/f_0$ (фазу сигнала в центре раскрыва принимаем равной нулю).

Таким образом, комплексная амплитуда сигнала, воспринимаемого элементом dx в точке x раскрыва антенны, пропорциональна $G(x) \exp(j 2\pi x u_x / \lambda) dx$, а полный сигнал, принимаемый раскрывом антенны с направления u_x , с точностью до постоянного коэффициента k описывается диаграммой направленности

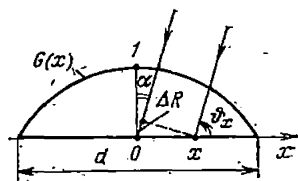


Рис. 3.4. Прием сигнала из антенну с линейным раскрывом

$$G(u_x) = k \int_{-d/2}^{d/2} G(x) e^{j 2\pi x u_x / \lambda} dx, \quad (3.18)$$

в общем случае комплексной.

Множитель x/λ при u_x в показателе экспоненты играет такую же роль, как частота f при t в обратном преобразовании Фурье (3.1). По аналогии введем относительную (электрическую) координату раскрыва $v_x = x/\lambda$ и назовем ее *угловой пространственной частотой*. Кроме того, положим $k = \lambda$, а пределы интегрирования продлим до бесконечности, что не изменит значения интеграла, так как вне раскрыва функция $G(x) = 0$. При таких условиях полученная формула ничем, кроме обозначений, не отличается от обратного преобразования Фурье (3.1) и функцию $G(v_x) = G(x/\lambda)$ можно назвать *угловым пространственным спектром*.

Следовательно, как сигнал описывается комплексной огибающей и ее комплексным спектром, связанными между собой преобразованиями Фурье, так и раскрыв (апертуру) антенны можно описать комплексной диаграммой направленности $G(u_x)$ и ее угловым спектром $G(v_x)$ — функцией раскрыва, которые также связаны преобразованиями Фурье. То же самое справедливо относительно $G(u_y)$ и $G(v_y)$, где $u_y = \cos \vartheta_y$, а $v_y = y/\lambda$ — угловая частота линейного раскрыва вдоль оси y . Функции $G(u_{x,y})$ или $G(x,y)$ описывают пространственную характеристику измерительной системы — апертуру линейной антенны. Обобщая полученные выводы, запишем прямое и обратное преобразования Фурье для апертуры линейной антенны:

$$G(v_{x,y}) = \int_{-\infty}^{\infty} G(u_{x,y}) e^{-j 2\pi v_{x,y} u_{x,y}} du_{x,y}; \quad (3.19)$$

$$G(u_{x,y}) = \int_{-\infty}^{\infty} G(v_{x,y}) e^{j 2\pi v_{x,y} u_{x,y}} dv_{x,y}.$$

Соотношения (3.19) для апертуры антенны подобны соотношениям (3.1) для сигнала. Однако сигнал обычно ограничен по длительности, и поэтому его спектр теоретически бесконечен, а антенна всегда имеет ограниченный раскрыв — конечную ширину

углового спектра, и поэтому сама функция — диаграмма направленности — простирается в теоретически бесконечных пределах.

При синфазном раскрыве — вещественном угловом спектре $G(v_{x,y})$ — обеспечивается наибольшая амплитуда сигнала в направлении максимума и наиболее узкая диаграмма направленности, $G(u_{x,y})$. Так как связанные по Фурье функции имеют взаимно обратную протяженность, угловая ширина диаграммы направленности $\Delta u_{x,y}$ определяется шириной углового спектра $\Delta v_{x,y} = d_{x,y}/\lambda$:

$$\Delta u_{x,y} = 1/\Delta v_{x,y} = \lambda/d_{x,y}, \quad (3.20)$$

если отсчет ширины производить на соответствующем уровне. Здесь полная аналогия с формулой (3.12) для длительности и ширины спектра сигнала. При этом диаграмма направленности вещественная и четная, если функция раскрыва также вещественная и четно-симметричная. Нечетная вещественная функция раскрыва дает нечетную, но чисто мнимую диаграмму направленности, что соответствует фазовому сдвигу сигнала на $-\pi/2$. Поэтому антенны с асимметричной вещественной функцией раскрыва имеют комплексную диаграмму направленности. Амплитудная диаграмма направленности (модуль комплексной диаграммы) антенны с любой вещественной функцией раскрыва, всегда четная.

Комплексная функция раскрыва $G(v_{x,y})$ (т. е. несинфазного пространственного спектра поля по раскрыву антенны) кроме линейного набега фазы, соответствующего простому сдвигу (повороту) неискаженной диаграммы направленности $G(u_{x,y})$, вызывает ее расширение в N раз:

$$\Delta u_{x,y} = N/\Delta v_{x,y}, \quad (3.21)$$

что полностью соответствует формуле (3.13) для сигнала и его спектра. Коррекцией фазы сигналов по раскрыву — пропусканием совокупного сигнала антенны через фазокорректирующий пространственный фильтр — диаграмма направленности сжимается до минимально возможной ширины, определяемой формулой (3.20).

Таким образом, между характеристиками сигнала $S(f)$ и $S(t)$ и раскрыва антенны $G(u_{x,y})$ и $G(v_{x,y})$ существует глубокая аналогия. Это позволяет все свойства сигналов переносить на апертуру антенны. Однако у апертуры антенны имеются существенные особенности.

1. Амплитуда сигнала антенны дополнительно зависит от направления прихода волны $\vartheta_{x,y}$, так как на раскрыв воздействует только составляющая поля $E_{x,y} = E \sin \vartheta_{x,y} = EV \sqrt{1-u_{x,y}^2}$, параллельная линейному раскрыву (рис. 3.5). При этом предполагается, что вектор E лежит

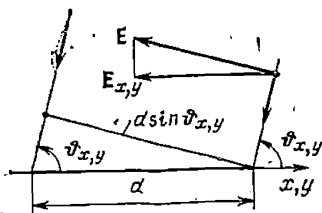


Рис. 3.5. Зависимость амплитуды сигнала от направления прихода волны

в плоскости раскрыва; если же он перпендикулярен этой плоскости, то сигнал вообще не будет принят — в этом проявляется влияние поляризации волны. Таким образом, зависимость амплитуды сигнала антенны от угла определяется произведением $\sqrt{1-u_{x,y}^2} G(u_{x,y})$.

При фазировании антенны на направление u_{x_0, y_0} амплитуда выходного сигнала

$$U(u_{x,y}) = K \sqrt{1-u_{x,y}^2} G(u_{x,y} - u_{x_0, y_0}) \quad (3.22)$$

уменьшается, но сама диаграмма направленности антенны только смещается, сохраняя форму и ширину по направляющим косинусам $u_{x,y}$. Если же отсчет вести в реальных углах $\vartheta_{x,y}$, то ширина диаграммы направленности остронаправленных антенн

$$\Delta\vartheta_{x,y} \approx \frac{\Delta u_{x,y}}{\sin \vartheta_{x_0, y_0}} = \frac{1}{\Delta\vartheta_{x,y} \sin \vartheta_{x_0, y_0}} = \frac{\lambda}{d \sqrt{1-u_{x_0, y_0}^2}} \quad (3.23)$$

возрастает в такой же степени, в какой снижается амплитуда сигнала (3.22) в направлении максимума. При этом основной лепесток у остронаправленных антенн сохраняет свою форму, а у слабонаправленных из симметричного становится асимметричным, расширяясь в сторону малых направляющих углов $\vartheta_{x,y}$. Снижение амплитуды сигнала (3.22) и расширение диаграммы направленности при ее смещении (3.23) можно трактовать как уменьшение раскрыва антенны $d \sin \vartheta_{x,y} = d \sqrt{1-u_{x,y}^2}$, видимого с данного направления (см. рис. 3.5).

2. В отличие от сигнала и его спектра, которые могут иметь неограниченную протяженность, диаграмма направленности антенны $G(u_{x,y})$ как функция переменной $u_{x,y}$ — направляющего косинуса угла ($-1 \leq u_{x,y} \leq 1$) — имеет конечную протяженность

$$[u_{x,y \max} - u_{x,y \min}] = 2, \quad (3.24)$$

а угловой пространственный спектр $G(v_{x,y})$ всегда ограничен раскрывом антенны ($\Delta v_{x,y} = d_{x,y}/\lambda$). Отсюда вытекают важные следствия.

Первое следствие состоит в том, что диаграмма направленности $G(u_{x,y})$ (по крайней мере теоретически) бесконечна, так как при $|u_{x,y}| > d/2\lambda$ функция раскрыва $G(v_{x,y}) = 0$. Она не равна нулю и при $|u_{x,y}| > 1$, что соответствует мнимым углам (косинус мнимого угла — гиперболический косинус — может превышать единицу). Той части диаграммы направленности, которая принадлежит мнимым углам, соответствует реактивная колебательная мощность. Эта мощность у остронаправленных антенн практически отсутствует, так как в область мнимых углов попадают только дальние боковые лепестки, уровень которых очень мал. У вибраторов очень малых размеров ($d/\lambda \ll 1$) большая часть диаграммы направленности приходится на мнимые углы; их входное сопротивление в основном реактивное, что и указывает на наличие пре-

зимушественно колебательной, а не излучаемой или воспринимаемой ими мощности.

У элементарного диполя, относительный размер которого $\Delta v_{x,y} = d/\lambda \ll 1$ очень мал, комплексная диаграмма направленности $G(u_{x,y})$ как преобразование Фурье от функции раскрыва $G(v_{x,y})$ имеет очень большую протяженность в направляющих косинусах $\Delta u_{x,y} = 1/\Delta v_{x,y} \gg 1$ и далеко выходит за пределы реальных углов $-1, 1$ (рис. 3.6, а). В пределах же реальных углов $G(u_{x,y}) \approx 1$, так как сигналы в раскрыве, приходящие с любого реального направления, практически синфазны ввиду пренебрежимо малой разности хода. Поэтому реальная зависимость амплитуды сигнала антенны от угла (3.22) описывается только множителем $\sin \vartheta_{x,y} = \sqrt{1 - u_{x,y}^2}$. Элементарный диполь крайне неэффективно принимает и излучает колебания: мощность и входное сопротивление диполя почти чисто реактивные — емкостные у электрического и индуктивные у магнитного диполя. Полуволновой вибратор ($d = \lambda/2$ или $\Delta v_{x,y} = d/\lambda = 1/2$) имеет функцию раскрыва $G(v_{x,y}) = \cos 2\pi v_{x,y}$, преобразование Фурье которой дает комплексную диаграмму направленности $G(u_{x,y})$ шириной $\Delta u_{x,y} = 1/\Delta v_{x,y} = 2$ на уровне 0,79 в пределах реальных углов $-1, 1$ (рис. 3.6, б). Так как в этих пределах сама функция мало отличается от единицы, реальные зависимости амплитуды сигнала от направления у полуволнового вибратора и элементарного диполя (штриховые линии) очень близки. Таким образом, формула (3.23) неприменима для определения ширины реальных диаграмм направленности слабонаправленных антенн. Значительная часть диаграммы $G(u_{x,y})$ полуволнового вибратора также приходится на мнимые углы, однако в отличие от элементарного диполя колебания мощности имеют резонансный характер: емкостная и индуктивная составляющие реактивной мощности одинаковы. Поэтому полуволновой вибратор имеет чисто активное входное сопротивление, эффективно излучает и принимает колебания. Это и определило его

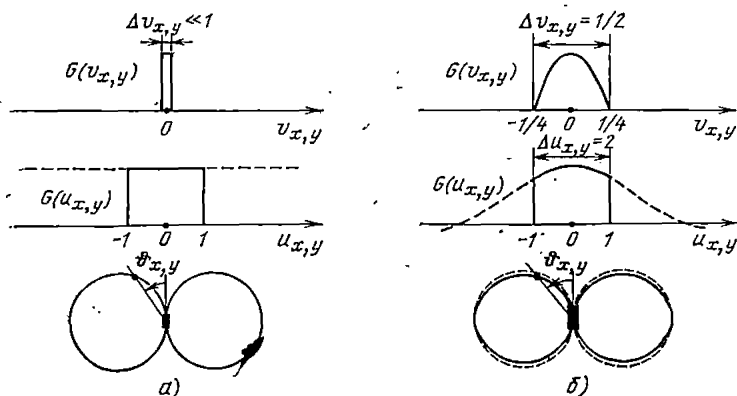


Рис. 3.6. Распределение поля, его преобразование Фурье и реальная диаграмма направленности элементарного диполя (а) и полуволнового вибратора (б)

широкое использование как самостоятельно, так и в составе антенных решеток.

Необходимо отметить, что формула (3.22), описывающая реальную диаграмму направленности слабонаправленных антенн в пределах реальных углов, $-1 \leq u_{x,y} \leq 1$, у остронаправленных антенн имеет вид $G(u_{x,y})$, а функция $\sqrt{1-u_{x,y}^2} = \sin \vartheta_{x,y}$ в пределах основного лепестка диаграммы ($u_{x,y} \approx u_{x_0, y_0}$) принимает вид амплитудного множителя $\sqrt{1-u_{x_0, y_0}^2}$, соответствующего направлению максимума диаграмм направленности u_{x_0, y_0} . Поскольку в обзорно-измерительных системах используют в основном остронаправленные антенны, далее речь будет идти только о диаграммах направленности вида $G(u_{x,y})$.

Здесь были рассмотрены диаграммы направленности только в плоскости, проходящей через ось вибратора. Фактически диаграмма направленности линейной антенны является объемной, полученной как результат вращения рассмотренных плоских диаграмм вокруг оси вибратора, т. е. диаграмма направленности в плоскости, перпендикулярной оси вибратора, имеет вид окружности.

Другим важным следствием, вытекающим из второй особенности апертуры антенны, является то, что непрерывный раскрыв можно заменить дискретным, взятым через интервалы не более полу волны, не опасаясь неоднозначности отсчета угла. У временных функций временной интервал неограничен, и дискретный спектр частоты обязательно приводит к неоднозначности: даже при очень малом периоде дискретизации δf функция повторяется через достаточно большие интервалы времени $T = 1/\delta f$, обратные периоду дискретизации частоты. У углового спектра $G(u_{x,y})$ с периодом дискретизации $\delta u_{x,y} = b/\lambda$, где b — расстояние между эквидистантными излучающими элементами антенной решетки, интервал возможных углов ограничен: $[u_{x,y \max} - u_{x,y \min}] = 2$. Поэтому повторения диаграммы направленности не будет, если пространственный период $1/\delta u_{x,y}$ не меньше полного интервала углов, т. е.

$$\delta u_{x,y} = \frac{b}{\lambda} \leq \frac{1}{[u_{x,y \max} - u_{x,y \min}]} = \frac{1}{2}. \quad (3.25)$$

Это и означает, что для однозначного воспроизведения угла расстояние между вибраторами должно удовлетворять условию $b \leq \lambda/2$. При этом, если вибраторы решетки имеют чисто активное входное сопротивление (например, полуволновые), повторяющиеся пики множителя решетки в области мнимых углов не играют никакой роли.

Использование фазированных антенных решеток (ФАР) позволяет управлять амплитудой и фазой сигналов в каждой точке раскрыва, что невозможно в антеннах со сплошным раскрывом.

3. Большинство реальных антенн имеет плоский, двумерный раскрыв. Комплексная диаграмма направленности плоской антенны $G(u_x, u_y)$ и ее двумерный спектр — функция раскрыва

$G(v_x, v_y)$ связаны между собой двумерными преобразованиями Фурье, прямым и обратным:

$$G(v_x, v_y) = \iint_{-\infty}^{\infty} G(u_x, u_y) e^{-j 2\pi (v_x u_x + v_y u_y)} du_x du_y; \quad (3.26)$$

$$G(u_x, u_y) = \iint_{-\infty}^{\infty} G(v_x, v_y) e^{j 2\pi (v_x u_x + v_y u_y)} dv_x dv_y,$$

причем $|G(v_x, v_y)|_{\max} = G(0, 0) = 1$. Ширина двумерной диаграммы направленности синфазной плоской антенны $\Delta(u_x, u_y)$ определяется шириной двумерного углового спектра $\Delta(v_x, v_y) = S/\lambda^2$ (S — площадь раскрыва):

$$\Delta(u_x, u_y) = 1/\Delta(v_x, v_y) = \lambda^2/S. \quad (3.27)$$

Здесь, как и у линейной антенны, отсчет функций $\Delta(u_x, u_y)$ и $\Delta(v_x, v_y)$ ведется на уровне, зависящем от конкретного вида функций $G(u_x, u_y)$ и $G(v_x, v_y)$. При отсчете обеих функций на уровне 0,71 в формуле (3.27) единица заменяется близким к ней коэффициентом, также зависящим от вида функций.

При комплексном двумерном угловом спектре $G(v_x, v_y)$, т. е. несинфазном плоском раскрыве*, справедливо соотношение

$$\Delta(u_x, u_y) = N_x N_y / \Delta(v_x, v_y) \quad (3.28)$$

и диаграмма направленности расширяется в N раз по каждой из двух координат. Фазированием сигналов ее можно сжать до минимальной ширины (3.27).

Если функции раскрыва по осям x и y независимы, т. е. $G(v_x, v_y) = G(v_x) G(v_y)$ и, следовательно, $G(u_x, u_y) = G(u_x) G(u_y)$, то $S = d_x d_y$, $\Delta(v_x, v_y) = \Delta v_x \Delta v_y$ и $\Delta(u_x, u_y) = \Delta u_x \Delta u_y$, где Δv_x и Δu_x , Δv_y и Δu_y связаны между собой соотношением (3.20). Условие независимости выполняется только для прямоугольного раскрыва, стороны которого параллельны осям x и y , а у антенны с круглым или эллиптическим раскрывом такое разделение невозможно.

Телесный угол, занимаемый диаграммой направленности плоской синфазной антенны,

$$\Omega_A = \frac{\Delta(u_x, u_y)}{\cos \vartheta_{z_0}} = \frac{\Delta(u_x, u_y)}{\sqrt{1 - u_{x_0}^2 - u_{y_0}^2}} = \frac{\lambda^2}{S \sqrt{1 - u_{x_0}^2 - u_{y_0}^2}} \quad (3.29)$$

зависит от направления фокусирования остронаправленной антенны относительно нормали к раскрыву — направляющего косинуса $u_{z_0} = \sqrt{1 - u_{x_0}^2 - u_{y_0}^2}$ и определяется видимой с этого направления площадью раскрыва $S \sqrt{1 - u_{x_0}^2 - u_{y_0}^2}$. Расширение смещенной диаграммы направленности по каждой из двух угловых координат отдельно согласно (3.28) определяется множителя-

* Имеется в виду квазислучайный закон распределения фазы $\varphi(v_x, v_y)$; линейное изменение фазы вызывает лишь смещение диаграммы направленности $G(u_x, u_y)$.

ми $\sqrt{1-u_{x_0}^2}$ и $\sqrt{1-u_{y_0}^2}$. Таким образом, плоская неподвижная антенна обеспечивает постоянное разрешение $\Delta(u_x, u_y)$ по направляющим косинусам, а не по реальным углам: телесный угол луча возрастает при отклонении от нормали к раскрыву.

В радиосистемах очень широкое применение получили круглые плоские апертуры. В этом случае как функция раскрыва, так и диаграмма направленности антенны обладают осевой симметрией. Поэтому для вычисления двумерных преобразований Фурье (3.26) целесообразно перейти от декартовых координат v_x, v_y и u_x, u_y к полярным v, θ и u, δ . При этом

$$v_x = v \cos \theta \text{ при } -\infty < v_x < \infty, \quad v_y = v \sin \theta \text{ при } -\infty < v_y < \infty, \quad v = \sqrt{v_x^2 + v_y^2} \\ \text{при } 0 \leq v < \infty;$$

$$u_x = u \cos \delta \text{ при } -\infty < u_x < \infty, \quad u_y = u \sin \delta \text{ при } -\infty < u_y < \infty, \quad u = \sqrt{u_x^2 + u_y^2} \\ \text{при } 0 \leq u < \infty;$$

$$\theta = \arctg(v_y/v_x) \text{ при } -\pi < \theta \leq \pi; \quad \delta = \arctg(u_y/u_x) \text{ при } -\pi < \delta \leq \pi;$$

$$dv_x dv_y = v dv d\theta; \quad du_x du_y = u du d\delta;$$

$$v_x u_x + v_y u_y = vu \cos \theta \cos \delta + vu \sin \theta \sin \delta = vu \cos(\theta - \delta).$$

Учитывая также, что

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j 2\pi vu \cos(\theta - \delta)} d\delta = J_0(2\pi vu)$$

— функция Бесселя нулевого порядка, после замены переменных в формулах (3.26) получим

$$G(v, \theta) = \int_0^{\infty} \int_{-\pi}^{\pi} u G(u, \delta) e^{-j 2\pi vu \cos(\theta - \delta)} du d\delta =$$

$$= 2\pi \int_0^{\infty} u G(u, \delta) J_0(2\pi vu) du;$$

$$G(u, \delta) = 2\pi \int_0^{\infty} v G(v, \theta) J_0(2\pi vu) dv.$$

Очевидно, что функция $G(v, \theta)$ и $G(u, \delta)$ равномерно распределены в пределах 2π по всем направлениям θ и δ соответственно. Поэтому положим $G(v) = G(v, \theta)/2\pi$; $G(u) = G(u, \delta)/2\pi$ и получим прямое и обратное преобразования Ганкеля для круглой плоской апертуры антенны

$$G(v) = \int_0^{\infty} u G(u) J_0(2\pi vu) du, \quad u > 0; \\ G(u) = \int_0^{\infty} v G(v) J_0(2\pi vu) dv, \quad v > 0. \quad (3.30)$$

Здесь $v = r/\lambda$ — линейное отклонение от центра апертуры (в электрических единицах), а u — угловое отклонение от направления максимума диаграммы направленности. В отличие от преобразований Фурье прямое и обратное преобра-

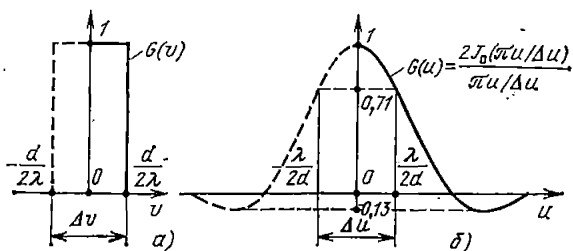


Рис. 3.7. Распределение поля в антенне с круглым плоским раскрывом (а) и ее диаграмма направленности (б)

зования Ганкеля полностью симметричны, а аргументы v и u только положительны. Двумерные функции раскрыва и диаграммы направленности круглой плоской антенны образуются вращением функции $G(v)$ и $G(u)$ вокруг оси ординат.

Возьмем в качестве примера круглую антенну диаметром d с равномерным распределением поля по раскрыву шириной $\Delta v = d/\lambda$. Подставив $G(v) = 1$ при $0 \leq v \leq \Delta v/2$ и $G(v) = 0$ при $v > \Delta v/2$ в формулу (3.30), получим

$$G(u) = \int_0^{\Delta v/2} v J_0(2\pi vu) dv = \frac{J_1(2\pi vu)}{2\pi u} \Big|_0^{\Delta v/2} = \frac{\Delta v}{4} \frac{2J_1(\pi \Delta v u)}{\pi \Delta v u},$$

где $J_1(\cdot)$ — функция Бесселя первого порядка (рис. 3.7). Ширина этой диаграммы на уровне 0,71 от максимума $\Delta u = 1/\Delta v = \lambda/d$, а наибольший из боковых лепестков имеет уровень 0,13 (на рисунке функции $G(v)$ и $G(u)$ симметрично дополнены штриховой линией, чтобы показать осевое сечение двумерных функций). Обычно функция раскрыва плавно спадает к его краям, что уменьшает уровень боковых лепестков и несколько расширяет основной лепесток. Но в целом уровень боковых лепестков у антенны с осевой симметрией раскрыва всегда ниже, чем у антенны с прямоугольным раскрывом при одинаковом законе облучения. Ширина диаграммы направленности и уровень боковых лепестков круглых антенн одинаковы в любом сечении в отличие от антенн любого другого вида.

В заключение отметим, что плоская антенная решетка, все элементы которой размещены в одной плоскости, образует не один, а два луча, симметричных относительно плоскости раскрыва. Для получения единственного луча (исключение неоднозначности) антенна должна иметь конечный продольный размер (см. гл. 6).

3.4. ПРИНИМАЕМЫЙ СИГНАЛ КАК ФУНКЦИЯ ВРЕМЕНИ И ПРОСТРАНСТВА

Рассмотрим взаимодействие пространственной и временной характеристик системы: комплексную огибающую $S_E(t) \equiv S(t)$ волны, падающей на плоскую антенну ($z=0$) с функцией раскрыва $G(x, y)$ или $G(v_x, v_y)$. За начало отсчета сигнала примем его середину, а за начало отсчета координат раскрыва — его центр. Пусть в начальный момент $t=0$, когда удаленный точечный объект излучает (переизлучает, отражает) сигнал в начале его отсчета, расстояние от объекта до центра раскрыва равно R , а нап-

равление на объект относительно плоскости раскрыва характеризуется направляющими косинуса $u_x = \cos \vartheta_x$ и $u_y = \cos \vartheta_y$ (рис. 3.8). Тогда расстояние от объекта до произвольной точки раскрыва x, y можно представить как $R - \Delta R$, где разность хода лучей как проекция вектора $\mathbf{r} = ix + jy$ на направление $\mathbf{u} = iu_x + ju_y$ определяется векторным произведением $\Delta R = \mathbf{r} \cdot \mathbf{u} = xu_x + yu_y$.

Если объект движется и производные $\dot{R}, \dot{u}_x, \dot{u}_y$ на небольших интервалах времени можно считать постоянными, то к моменту $t \geq 0$ его координатами станут $R + \dot{R}t$, $u_x + \dot{u}_x t$, $u_y + \dot{u}_y t$, а расстояние от объекта до произвольной точки раскрыва x, y

$$R(x, y, t) = R(t) - \Delta R(t) = R + \dot{R}t - [x(u_x + \dot{u}_x t) + y(u_y + \dot{u}_y t)] \quad (3.31)$$

становится функцией времени и пространства, а также значений координат объекта R, u_x, u_y и их производных $\dot{R}, \dot{u}_x, \dot{u}_y$ в момент излучения $t=0$:

Таким образом, напряженность поля падающей на раскрыв антенны волны в произвольной точке x, y в произвольный момент времени t описывается функцией

$$S_E[t - R(x, y, t)/c] \exp[-j2\pi f_0 R(x, y, t)/c],$$

где аргумент первого сомножителя учитывает запаздывание огибающей, а показатель экспоненты — фазовый сдвиг несущей f_0 за время распространения сигнала от объекта до данной точки раскрыва.

Направление прихода волны относительно нормали z к раскрыву и вектора поля $\mathbf{E}$ относительно плоскости раскрыва характеризуется направляющим косинусом $u_z = \sqrt{1 - u_x^2 - u_y^2}$. Раскрывом антенны воспринимается составляющая напряженности поля, пропорциональная этому направляющему косинусу. При наклоне фронта волны только в плоскости $\mathbf{E}$ это очевидно; в общем случае нужно учесть наклон и в плоскости $\mathbf{H}$, что приводит к тому же результату, поскольку $S_H(x, y, t) = S_E(x, y, t)/\rho$ (ρ — волновое сопротивление свободного пространства). Необходимо учесть также степень взаимодействия между полем и раскрывом в каждой его точке, определяемую функцией раскрыва $G(x, y)$, в общем случае комплексной. Обычно эта функция максимальна в центре раскрыва $G(0, 0) = 1$ и спадает к его краям. Учитывая влияние наклона волны и функции раскрыва, находим комплексную оги-

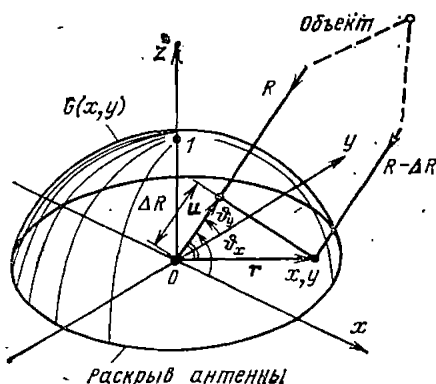


Рис. 3.8. Разность хода сигналов при падении волны на раскрыв антенны

бающую напряженности поля входного сигнала приемного устройства радиосистемы, воспринимаемого произвольной точкой раскрыва в произвольный момент времени:

$$S_E(x, y, t) = \sqrt{1 - u_x^2 - u_y^2} G(x, y) S_E[t - R(x, y, t)/c] \times \exp[-j2\pi f_0 R(x, y, t)/c], \quad (3.32)$$

где $R(x, y, t)$ определяется соотношением (3.31).

На входной сигнал (3.32) накладывается шум $N_E(x, y, t)$, как внешний, так и внутренний, пересчитанный на вход приемного устройства. Внешний шум $N''_E(x, y, t)$ представляет собой случайное электромагнитное поле, воспринимаемое антенной со всех направлений. Свойства этого шума описываются пространственно-временной функцией корреляции или пространственно-временной спектральной плотностью, которые однозначно связаны между собой многомерными преобразованиями Фурье. Будем считать внешнее шумовое поле центрированным процессом, дельта-коррелированным во времени и пространстве, т. е. имеющим равномерную спектральную плотность во всем диапазоне временных и пространственных частот. Такое предположение, вообще говоря, не совсем верно, так как радиоизлучение отдельных звезд и галактик дает пространственно зависимый вклад во внешнее шумовое поле. Однако эти дискретные составляющие при необходимости можно учесть отдельно. С учетом степени восприятия поля раскрывом $G(x, y)$ внешние шумы на входе системы принимают вид $N''_E(x, y, t) G(x, y)$, а средняя мощность внешнего шума и, следовательно, его спектральная плотность пропорциональны квадрату функции раскрыва: $N''_0 |G(x, y)|^2$. Внутренний шум линейной части приемного устройства, отнесенный ко входу системы, можно представить как временной шум $N'(t)$, равномерно распределенный по всей площади раскрыва антенны $S = \int dx dy$, т. е. $N'_E = N'(t)/S$, а его спектральная плотность N'_0 не зависит от координат раскрыва.

Таким образом, входной шум системы как функция времени и координат раскрыва принимает вид

$$N_E(x, y, t) = N'_E(t)/S + N''_E(x, y, t) G(x, y), \quad (3.33)$$

а спектральная плотность выходного шума антенны

$$N_0(x, y) = N'_0 + N''_0 |G(x, y)|^2 \quad (3.34)$$

оказывается неравномерной по ее раскрыву. Следовательно, входной шум радиосистем в общем случае коррелирован в пространстве. Однако в реальных системах, особенно при использовании ФАР, которые обладают высоким уровнем собственных шумов, внешними шумами можно пренебречь, положив $N''_0 \ll N'_0$ и $N'_0 \approx N_0$. Тогда входной шум системы $N_E(x, y, t) \approx N'(t)/S$, имеющий спектральную плотность

$$N_{E0}(x, y, f) \approx N_0/S, \quad (3.35)$$

оказывается некоррелированным как во времени, так и в пространстве.

Указанное предположение позволяет распространить на пространственно-временные входные сигналы развитую только для временных сигналов теорию оптимальной, согласованной фильтрации. В соответствии с известной формой входного сигнала (3.32) определяется пространственно-временная структура оптимального приемного устройства и выходной сигнал в виде многомерной пространственно-временной функции корреляции входного сигнала. Эта функция позволяет определить потенциальные возможности системы в отношении точности и разрешающей способности по всей совокупности измеряемых параметров — координат объекта и их производных. Для этого кроме энергетических параметров достаточно знать только две входные метрические характеристики системы, временную и пространственную.

3.5. ВЫХОДНОЙ СИГНАЛ ОПТИМАЛЬНОГО ПРИЕМНИКА КАК ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ КОРРЕЛЯЦИОННЫЙ ИНТЕГРАЛ

В соответствии с выражениями (3.31) и (3.32) входное воздействие приемного устройства

$$U_E(x, y, t) = S_E(x, y, t; \mathbf{v}) + N_E(x, y, t) \quad (3.36)$$

является функцией времени, координат пространства и шести неизвестных, измеряемых параметров волны $\mathbf{v} = \{\tau, F; u_x, u_y; \dot{u}_x, \dot{u}_y\}$, где $\tau = R/c$ и $F = -\dot{R}f_0/c$ — время запаздывания и доплеровский сдвиг частоты сигнала в системах прямого излучения (при работе по отраженному или переизлученному сигналу обе величины удваиваются). Входной шум $N_E(x, y, t)$ согласно (3.35) полагаем белым (т. е. дельта-коррелированным во времени и пространстве), спектральная плотность которого N_0 равномерно распределена по раскрытию антенны.

По аналогии с обычным, временным оптимальным приемником выходной сигнал описывается комплексной огибающей пространственно-временного корреляционного интеграла

$$Z(\mathbf{v}, \mathbf{v}^0) = \frac{k}{2} \int_{-\infty}^{\infty} \int \int U_E(x, y, t) S_E^*(x, y, t; \mathbf{v}^0) dx dy dt, \quad (3.37)$$

взято по всем значениям времени и координат раскрытия антенны. Опорный сигнал $kS_E^*(x, y, t; \mathbf{v}^0)$ является комплексно-сопряженной функцией входного сигнала (3.32), с точностью до постоянного множителя и постоянного сдвига, и зависит от шести опорных значений параметров $\mathbf{v}^0 = \{\tau^0, F^0, u_x^0, u_y^0; \dot{u}_x^0, \dot{u}_y^0\}$, на которые настроен приемник. В общем случае $\mathbf{v}^0 \neq \mathbf{v}$, т. е. выходной сигнал является функцией расстройки оптимального приемника по измеряемым параметрам.

Поскольку входное воздействие (3.36) есть сумма сигнала и шума, выходной сигнал оптимального пространственно-временного приемника также состоит из сигнальной и шумовой составляющих:

$$Z(\mathbf{v}, \mathbf{v}^0) = Z_s(\mathbf{v}, \mathbf{v}^0) + Z_n. \quad (3.38)$$

Сигнальная составляющая комплексной огибающей выходного сигнала (3.37) в соответствии с формулами (3.32), (3.36)

$$\begin{aligned} Z_s(\mathbf{v}, \mathbf{v}^0) = & 0,5 k \sqrt{1 - u_x^2 - u_y^2} \int_{-\infty}^{\infty} \int \int |G(x, y)|^2 S_E[t - R(x, y, t)/c] \times \\ & \times S_E^*[t - R^0(x, y, t)] \exp\{-j 2\pi f_0 [R(x, y, t) - \\ & - R^0(x, y, t)]/c\} dx dy dt. \end{aligned} \quad (3.39)$$

Здесь $R^0(x, y, t)$ определяется формулой (3.31), где измеряемые параметры заменены их опорными значениями.

Поскольку речь идет об узкополосных сигналах ($\Delta f/f_0 \ll 1$), не все слагаемые $R(x, y, t)$ и $R^0(x, y, t)$ в выражении (3.39) играют одинаковую роль. Так, основной сдвиг огибающей $S_E(t)$ по времени дает член $\tau = R/c$, а остальными можно пренебречь. Этот член в показателе экспоненты характеризует неизвестную начальную фазу и дает несущественный множитель, выходящий из-под знака интеграла, а остальные слагаемые существенно влияют на результат интегрирования. Необходимо также учесть, что в показателе экспоненты (3.39) фигурирует линейная разность действительных и опорных значений измеряемых параметров, а для огибающих $S_E(t - \tau)$ и $S_E^*(t - \tau^0)$ имеет значение только сдвиг по времени $\tau - \tau^0$. Поэтому интеграл (3.39) является функцией разности действительной и опорной совокупности значений $\mathbf{v}' = \mathbf{v} - \mathbf{v}^0 = \{\tau', F', u'_x, u'_y; \dot{u}'_x, \dot{u}'_y\}$, где $\tau' = \tau - \tau^0$, $F' = F - F^0$ и т. д.

Учитывая вышеизложенное, после замены переменных $v_x = x/\lambda$, $v_y = y/\lambda$ и отбрасывания несущественных множителей получаем развернутое выражение для сигнальной составляющей пространственно-временного корреляционного интеграла

$$\begin{aligned} Z_s(\mathbf{v}') = & \frac{\lambda^2}{2} \sqrt{1 - u_x^2 - u_y^2} \int_{-\infty}^{\infty} \int \int |G(v_x, v_y)|^2 S_E(t) \times \\ & \times S_E^*(t - \tau') \exp\{j 2\pi [F' t + v_x(u'_x + \dot{u}'_x t) + \\ & + v_y(u'_y + \dot{u}'_y t)]\} dv_x dv_y dt. \end{aligned} \quad (3.40)$$

При нулевом рассогласовании $\mathbf{v}' = 0$, как и в чисто временной задаче, интеграл (3.40) дает полную энергию принимаемого сигнала

$$\begin{aligned} \mathcal{E} = & \lambda^2 \sqrt{1 - u_x^2 - u_y^2} \int_{-\infty}^{\infty} \int \int |G(v_x, v_y)|^2 dv_x dv_y 0,5 \int_{-\infty}^{\infty} |S_E(t)|^2 dt = \\ = & A \sqrt{1 - u_x^2 - u_y^2} \int_{-\infty}^{\infty} 0,5 |S_E(t)|^2 dt = \mathcal{E}_0 \sqrt{1 - u_x^2 - u_y^2}. \end{aligned} \quad (3.41)$$

Здесь

$$A = \iint_{-\infty}^{\infty} |G(x, y)|^2 dx dy = \lambda^2 \iint_{-\infty}^{\infty} |G(u_x, u_y)|^2 du_x du_y \quad (3.42)$$

— эффективная площадь антенны; $0,5|S_E(t)|^2$ — плотность потока мощности (на волновом сопротивлении 1 Ом); $0,5|S_E(t)|^2 \sqrt{1-u_x^2-u_y^2}$ — составляющая плотности потока мощности в направлении нормали к раскрытию;

$0,5|S(t)|^2 = 0,5|S_E(t)|^2 A \sqrt{1-u_x^2-u_y^2}$ — мощность принимаемого высокочастотного сигнала (на сопротивлении 1 Ом);

$$S(t) = S_E(t) \sqrt{A} \sqrt{1-u_x^2-u_y^2} \quad (3.43)$$

— комплексная амплитуда суммарного высокочастотного сигнала фазированной антенны;

$$\mathcal{E}_0 = \frac{1}{2} A \int_{-\infty}^{\infty} |S_E(t)|^2 dt \quad (3.44)$$

— максимальная энергия сигнала, воспринимаемая антенной, когда прием производится с направления нормали к раскрытию.

Таким образом, при оптимальном приеме отношение сигнал-шум $\mathcal{E}/N_0 = (\mathcal{E}_0/N_0) \sqrt{1-u_x^2-u_y^2}$ зависит от направления прихода волны относительно нормали к плоскости раскрытия антенны. В антеннах с механической настройкой на направление приема путем поворота антенны всегда справедливо $\mathcal{E} = \mathcal{E}_0$. Энергию принимаемого сигнала $\mathcal{E}$ находим из уравнений дальности (см. § 3.4) и при заданных размерах зоны обзора определяется двумя энергетическими входными параметрами системы, пространственным и временным.

Временным энергетическим параметром является средняя мощность излучения или энергия излучаемого за один период обзора сигнала $\mathcal{E}_{изл}$. Излучение энергии в пределах зоны обзора может производиться как последовательно, так и одновременно в зависимости от принятого метода обзора. В обоих случаях в зоне расположения приемной антенны создается плотность потока энергии $0,5 \int |S_E(t)|^2 dt$, которая входит в качестве множителя в соотношение (3.41).

Пространственным энергетическим параметром системы является эффективная площадь антенны A . Эффективная площадь антенны, видимая с направления прихода волны, $A \sqrt{1-u_x^2-u_y^2}$ входит в качестве второго множителя в (3.41). Эффективная площадь связана с физической площадью раскрытия антенны S известным соотношением

$$A = K_A S, \quad (3.45)$$

где

$$K_A = \frac{1}{S} \iint_{-\infty}^{\infty} |G(x, y)|^2 dx dy = \frac{\lambda^2}{S} \iint_{-\infty}^{\infty} |G(u_x, u_y)|^2 du_x du_y \quad (3.46)$$

— коэффициент использования раскрыва антенны (без учета потерь в антенно-фидерном тракте). С учетом энергетических потерь $K_A = 0,4 \dots 0,8$.

Огибающая корреляционного интеграла (3.37) является функцией расстройки по всем шести измеряемым параметрам и в момент максимума $Z = |Z(v \simeq v^0)|$ определяется формулой (2.14). Ее получают детектированием сигнала на выходе линейного тракта приемника. В выходном устройстве производится обнаружение объектов путем испытания огибающей корреляционного интеграла на порог или измерения параметров их движения по методу максимума правдоподобия, как и в приемнике временных сигналов.

При определении вида операций над входным воздействием в оптимальном пространственно-временном приемнике в первую очередь необходимо учитывать наличие пространственной и временной когерентности сигналов. Отсутствие хотя бы одной из них не позволяет производить полностью когерентную обработку сигналов по высокой частоте, она возможна лишь в пределах интервала когерентности. Затем сигналы детектируются и дальнейшая их обработка производится некогерентно суммированием огибающих частных сигналов. О временной когерентности уже была речь в § 3.4, поэтому остановимся на пространственной когерентности.

Оптимальная когерентная обработка сигналов при очень больших апертурах, когда отдельные элементы антенн разнесены на значительные расстояния, возможна лишь в том случае, если обеспечивается когерентная связь (известны фазовые соотношения) между сигналами разнесенных элементов. В противном случае их фазирование и когерентное суммирование неосуществимо. Так, при условии когерентности сигналов нескольких станций, далеко отстоящих друг от друга, их антенны в совокупности можно считать раскрывом одной большой пространственной системы, сигналы которой обрабатываются (фазировются) по высокой частоте. Если же когерентной связи нет, система представляет собой совокупность отдельных станций, сигналы которых обрабатываются порознь когерентно, а результаты обработки затем объединяются некогерентно. Аналогично сигналы большой длительности можно разбить на участки, в пределах которых когерентность сохраняется и возможна когерентная обработка, но результаты обработки частных сигналов можно суммировать лишь некогерентно.

Таким образом, наличие пространственной и временной когерентности позволяет наиболее полно реализовать возможности в системах со сколь угодно большой протяженностью во времени и пространстве. Если же на всем интервале когерентность не обеспечивается, он делится на участки когерентности. Дробление энергии сигнала между участками когерентности ведет (как уже отмечалось) к энергетическим потерям в $\sqrt{n}$ раз (n — число участков), отрицательно сказывается не только на качестве обнаружения сигналов, но и на точности измерений параметров движения объектов.

Соотношения, описывающие выходные сигналы пространственно-временного и чисто временного оптимальных приемников, идентичны, если входные шумы имеют равномерную спектральную плотность как по временным, так и по пространственным частотам.

там. Применительно к пространственным частотам это условие выполняется (как отмечалось) только в том случае, когда внешними шумами можно пренебречь. Однако если учитывать и внешнюю, и внутреннюю составляющие шума, то спектральная плотность входного шума согласно формуле (3.34) при замене координат x и y пространственными частотами v_x и v_y оказывается неравномерной по раскрыву:

$$N_0(v_x, v_y) = N'_0 + N''_0 |G(v_x, v_y)|^2 = N_0 \left(\frac{N'_0}{N_0} + \frac{|G(v_x, v_y)|^2 N''_0}{N_0} \right), \quad (3.47)$$

где $N_0 = N'_0 + N''_0$. Неравномерный шум можно свести к равномерному со спектральной плотностью N_0 , если вместо действительной ввести эквивалентную функцию раскрыва

$$G_{\text{эк}}(v_x, v_y) = N_0 G(v_x, v_y) / [N'_0 + N''_0 |G(v_x, v_y)|^2]. \quad (3.48)$$

Аналогично можно учесть и действительную неравномерность спектральной плотности внешнего шумового поля, вызванную радиоизлучением звезд и галактик, рассматривая N''_0 в формуле (3.48) как функцию пространственных частот.

Таким образом, если учитывать не только внутренние шумы системы, но и внешний шум, необходимо в качестве пространственной характеристики системы использовать эквивалентную функцию раскрыва, так же как в качестве временной характеристики описанный ранее эквивалентный спектр огибающей сигнала. Именно эквивалентные характеристики определяют в общем случае свойства системы.

Необходимо отметить, что по мере увеличения доли внешних шумов (отношения N''_0/N'_0) эквивалентная функция раскрыва приближается к равномерной (рис. 3.9). Это равноценно некоторому сжатию основного лепестка диаграммы направленности и росту уровня боковых лепестков выходного сигнала по угловым координатам. В пределе, когда внутренними шумами можно пренебречь, антенна с любой функцией раскрыва эквивалентна антенне с равномерным раскрывом, т. е. выбор желаемой пространственной характеристики системы — функции раскрыва антенны — становится невозможным.

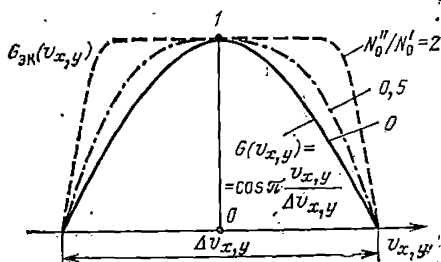


Рис. 3.9. Связь эквивалентной и реальной функций раскрыва при различных соотношениях между внешними и внутренними шумами

Глава 4. ОПТИМАЛЬНЫЙ ПРИЕМНИК КАК СОГЛАСОВАННЫЙ ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ ФИЛЬТР ИЛИ КОРРЕЛЯТОР

4.1. ОПТИМАЛЬНАЯ СТРУКТУРА ПРОСТРАНСТВЕННО-ВРЕМЕННОГО ПРИЕМНОГО УСТРОЙСТВА

Радиолокационный сигнал (3.36) как пространственно-временной континуум, воспринимаемый раскрывом антенны при воздействии аддитивных гауссовских шумов, равномерно распределенных по временным и пространственным частотам, позволяет определить структуру оптимального приемника в соответствии с теорией оптимальной фильтрации, разработанной для временных сигналов. Согласно этой теории оптимальный приемник представляет собой согласованный с чистым входным сигналом (без шумов) фильтр или коррелятор. Функция передачи такого фильтра комплексно-сопряжена со спектром сигнала, или импульсная характеристика его является зеркальным отражением собственно сигнала, что равносильно, поскольку обе функции однозначно связаны между собой преобразованиями Фурье.

Ввиду полного подобия пространственных и временных характеристик, используемых при описании входного сигнала как функции пространства и времени, оптимальный приемник такого сигнала представляется в виде пространственно-временного согласованного фильтра или коррелятора. Поскольку в данном случае входной сигнал и его спектр описываются комплексными функциями, импульсная характеристика и функция передачи оптимального пространственно-временного приемника представляют собой функции, комплексно-сопряженные с характеристиками сигнала.

В общем случае входной сигнал и структура согласованного с ним пространственно-временного приемника описываются довольно сложно. Однако задача существенно упрощается, если учесть предположения, сделанные ранее при выводе формулы (3.40), которые справедливы для сравнительно узкополосных сигналов и не слишком больших антенн, используемых в большинстве реальных систем, т. е. если можно пренебречь переходными процессами при падении волны на раскрыв антенны. Сущность переходных процессов состоит в том, что при наклонном падении на раскрыв антенны короткого волнового пакета (радиоимпульса) в некоторые моменты времени (особенно в начале и в конце этого пакета) облучается не весь раскрыв, как это было бы при облучении его непрерывной несущей, а только часть раскрыва. Поэтому диаграмма направленности, соответствующая полностью облученному

раскрыву, устанавливается не сразу или вообще не успевает сформироваться.

Переходные процессы на раскрыве можно не учитывать, если время прохождения волной раскрыва антенны d/c много меньше интервала корреляции сигнала $\tau_k = 1/\Delta f$, т. е.

$$\Delta f d/c \ll 1. \quad (4.1)$$

Это условие выполняется для сигналов с шириной спектра Δf до нескольких мегагерц и антенн с раскрывом d до нескольких десятков метров.

При условии (4.1) в соответствии с соотношением (3.32) после замены реальных координат раскрыва x, y угловыми пространственными частотами u_x, u_y сигнал на входе пространственно-временного приемника с точностью до постоянного множителя описывается комплексной функцией

$$S_E(u_x, u_y, t) \approx G(u_x, u_y) S_E(t - \tau) \exp \{ j 2\pi [F t + u_x(u_x + \dot{u}_x t) + u_y(u_y + \dot{u}_y t)] \}. \quad (4.2)$$

Соответственно характеристика согласованного приемника является комплексно-сопряженной с сигналом (4.2) функцией. Однако вместо истинных параметров сигнала $\tau, F, u_x, u_y, \dot{u}_x, \dot{u}_y$ при этом нужно вписать опорные значения $\tau^0, F^0, u_x^0, u_y^0, \dot{u}_x^0, \dot{u}_y^0$, на которые настроен приемник, в общем случае не равные истинным. Отсюда следует, что для достижения максимального выходного эффекта приемник должен быть не только согласованным, но и настроенным на параметры входного сигнала. Поскольку эти параметры обычно неизвестны и подлежат измерению, оптимальный приемник должен быть многоканальным или последовательно перестраиваемым по всем значениям измеряемых параметров (если перестройка не связана с энергетическими потерями). При этом необходимо иметь в виду, что оптимальный приемник в виде фильтра (но не коррелятора!) не требует настройки по дальности, перестройка по угловым координатам в принципе не связана с энергетическими потерями, а перестройка по радиальной скорости всегда ведет к энергетическим потерям, когда спектр сигнала находится вне полосы пропускания. Таким образом, оптимальный приемник обязательно должен быть многоканальным по радиальной скорости.

Необходимо особо остановиться на настройке по угловой скорости удаленных объектов, угловая скорость которых обычно невелика (не более 10^{-2} рад/с), а время облучения объекта (общая длительность сигнала) в обзорных станциях также ограничено (сотые доли секунды). В итоге угловое перемещение объекта за время облучения не превышает долей миллирадиана, т. е. направление на объект можно считать постоянным и, следовательно, влиянием угловой скорости (расстройки приемника по производным направляющих косинусов) при относительно небольших раскрывах можно пренебречь, положив в выражении (4.1) $\dot{u}_x = \dot{u}_y = 0$. В этих условиях входной сигнал (4.2) представляется в

виде произведения двух независимых функций, временной и пространственной:

$$S_E(v_x, v_y; t) \simeq S_E(t - \tau) \exp(j2\pi Ft) \times \\ \times G(v_x, v_y) \exp[j2\pi(v_x u_x + v_y u_y)]. \quad (4.3)$$

Следовательно, и приемник пространственно-временных сигналов можно представить в виде последовательно соединенных независимых трактов: пространственного и временного. При этом порядок следования трактов в линейной части приемника не имеет значения. Обычно сначала производится пространственная обработка сигналов по раскрытию антенны, затем временная. Однако в некоторых системах, наоборот, сначала сигналы каждого элемента раскрытия подвергаются временной фильтрации, а затем производится их пространственная обработка.

Структура временного тракта, т. е. оптимального приемника чисто временных сигналов, уже была рассмотрена в гл. 2. Поэтому далее основное внимание будет уделено построению пространственного тракта оптимального приемного устройства.

Пространственно зависимый член входного сигнала (4.3) с учетом ранее введенного множителя $\lambda^2 \sqrt{1 - u_{x0}^2 - u_{y0}^2}$, учитывающего зависимость сигнала антенны от направления прихода волны относительно раскрытия антенны, представляет собой угловой пространственный спектр входного сигнала

$$S_E(v_x, v_y) = \lambda^2 \sqrt{1 - u_{x0}^2 - u_{y0}^2} G(v_x, v_y) \times \\ \times \exp[j2\pi(v_x u_x + v_y u_y)]. \quad (4.4)$$

Функция передачи пространственного тракта согласованного приемника в присутствии шума с равномерным пространственным спектром является комплексно-сопряженной со спектром (4.4) характеристикой (с точностью до постоянного множителя)

$$H_0(v_x, v_y) = G^*(v_x, v_y) \exp[-j2\pi(v_x u_{x0} + v_y u_{y0})]. \quad (4.5)$$

Произведение функций (4.4) и (4.5) дает пространственный спектр выходного сигнала пространственного тракта приемника (без учета шумов) в функции расстройек $u'_{x,y} = u_{x,y} - u_{x0,y0}$:

$$Z_s(v_x, v_y) = \lambda^2 \sqrt{1 - u_{x0}^2 - u_{y0}^2} |G(v_x, v_y)|^2 \times \\ \times \exp[j2\pi(v_x u'_x + v_y u'_y)]. \quad (4.6)$$

Сам выходной сигнал пространственного тракта как результат суммирования функции (4.6) по всему раскрытию

$$Z_s(u'_x, u'_y) = \lambda^2 \sqrt{1 - u_{x0}^2 - u_{y0}^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(v_x, v_y)|^2 \times \\ \times \exp[j2\pi(v_x u'_x + v_y u'_y)] dv_x dv_y \quad (4.7)$$

является обратным преобразованием Фурье от функции $|G(v_x, v_y)|^2$, т. е. описывает некую результирующую выходную диаграмму направленности. Ее размеры в максимуме ($u'_x = u'_y = 0$) харак-

теризуют согласно формуле (3.42) видимую эффективную площадь антенны

$$A \sqrt{1 - u_{x0}^2 - u_{y0}^2} = \sqrt{1 - u_{x0}^2 - u_{y0}^2} \lambda^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(v_x, v_y)|^2 dv_x dv_y, \quad (4.8)$$

которая является пространственным энергетическим параметром системы, а нормированная функция (4.7) характеризует форму выходного сигнала оптимального приемника в зависимости от угловых координат.

Соотношение (4.6) описывает процесс пространственной согласованной фильтрации, подобной временной согласованной фильтрации в обычных приемниках. Однако обратное преобразование Фурье (4.7), происходящее при интегрировании пространственно отфильтрованных сигналов по всем координатам раскрыва антенны, является дополнительной операцией, отличающей пространственный тракт приема от временного, где такая операция отсутствует. Заметим, что в обычных антеннах также выполняется обратное преобразование Фурье — формирование диаграммы направленности антенны в результате когерентного суммирования поля в раскрыве. Однако пространственная согласованная фильтрация в обычных антеннах не производится, что при наличии внутренних шумов приемника ведет к энергетическим потерям, а также (как будет ясно из дальнейшего изложения) к снижению точности измерений и разрешающей способности по угловым координатам.

Чтобы осуществить оптимальную пространственную фильтрацию, нужно разделять управлять амплитудой и фазой сигнала в каждой точке раскрыва, что в антеннах со сплошным раскрывом невыполнимо. Однако отличительной особенностью пространственных характеристик является то, что непрерывную функцию раскрыва можно заменить практически эквивалентной ей дискретной функцией, если линейный интервал дискретизации не превышает $\lambda/2$. Отсюда следует, что пространственный согласованный фильтр или коррелятор можно построить только на основе ФАР. Для описания же общих свойств устройства можно по-прежнему пользоваться более удобными для анализа непрерывными функциями.

Определим вид оптимальной функции передачи пространственного тракта системы (4.5) для эквидистантной ФАР с линейным шагом b между ее элементами, число которых N по любой из двух координат раскрыва может быть четным или нечетным. При этом положение i -го элемента относительно центра раскрыва

$$v_{x, y} = \begin{cases} \frac{ib}{\lambda}, & i = 0, \pm 1, \pm 2, \dots, \pm \frac{N-1}{2}, \quad N - \text{нечетное,} \\ i(1 - i/2|i|) \frac{b}{\lambda}, & i = \pm 1, \pm 2, \dots, \pm \frac{N}{2}, \quad N - \text{четное,} \end{cases}$$

где $|i|$ — абсолютное значение числа i . Настройка системы на заданное направление может быть как раздельной по каждой из двух угловых координат, так и совместной по обеим координатам.

При раздельной настройке функцию раскрыва по координате x или y (в общем случае комплексную) и линейные по этой координате фазовые сдвиги принимаемого с направления u_x или u_y сигнала в канале i -го элемента ФАР заменяем дискретными отсчетами:

$$G_i = G(v_{x,y} i); \quad \varphi_i = 2\pi v_{x,y} i u_{x,y}.$$

Тогда функция передачи пространственного согласованного фильтра или коррелятора, настроенного на направление $u^0_{x,y}$, согласно формуле (4.5) применительно к одной координате определяется совокупностью комплексных чисел, общий член которой для i -го элемента ФАР имеет вид

$$H_{0i} = B_i \exp(-j\varphi^0_i), \quad (4.9)$$

где

$$B_i = k G^*_{i_i} \quad (4.10)$$

— весовой множитель (k — вещественный коэффициент);

$$\varphi^0_i = \begin{cases} i \psi_{x,y}, & N \text{ — нечетное,} \\ i(1 - i/2 |i|) \psi_{x,y}, & N \text{ — четное,} \end{cases} \quad (4.11)$$

— вводимый фазовый сдвиг;

$$\psi_{x,y} = -2\pi(b/\lambda) u^0_{x,y} \quad (4.12)$$

— фазовый сдвиг между сигналами соседних элементов при настройке на направление u^0_x или u^0_y , который при $b = \lambda/2$ принимает значение $\psi_{x,y} = -\pi u^0_{x,y}$.

Аналогично при совместной настройке по обеим координатам для ij -го элемента ФАР

$$G_{ij} = G(v_{x,i}, v_{y,j}); \quad \varphi_{ij} = 2\pi(v_{x,i} u_x + v_{y,j} u_y).$$

При этом функция передачи согласованного пространственного фильтра или коррелятора, настроенного на направление u^0_x, u^0_y , в соответствии с формулой (4.5) принимает вид матрицы, общий член которой для ij -го элемента ФАР

$$H_{0ij} = B_{ij} \exp(-j\varphi^0_{ij}), \quad (4.13)$$

где

$$B_{ij} = k G^*_{ij}, \quad (4.14)$$

$$\varphi^0_{ij} = \begin{cases} i \psi_x + j \psi_y, & N \text{ — нечетное,} \\ i \left(1 - \frac{i}{2|i|}\right) \psi_x + j \left(1 - \frac{j}{2|j|}\right) \psi_y, & N \text{ — четное,} \end{cases} \quad (4.15)$$

а величины ψ_x и ψ_y описываются формулой (4.12).

В соответствии с полученными формулами определяется структура пространственного тракта в виде согласованного фильтра

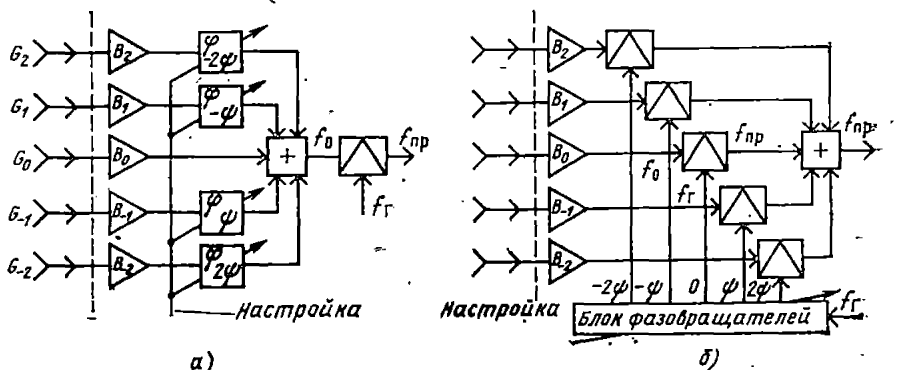


Рис. 4.1. Схема пространственного тракта согласованного приемного устройства на линейной ФАР

или коррелятора. Пространственный согласованный фильтр (рис. 4.1,а) содержит в канале каждого элемента ФАР усилитель с комплексным (в общем случае) коэффициентом передачи и фазовращатель для настройки на заданное направление приема (на рисунке приведен одномерный случай). Пространственный коррелятор (рис. 4.1,б) отличается от фильтра тем, что требуемые для настройки фазовые сдвиги придаются в блоке фазовращателей гетеродинным напряжениям частоты f_r , подаваемым на смесители в каналах элементов решетки, в результате настройка коррелятора на заданное направление приема достигается не на высокой f_0 , а на промежуточной f_{np} частоте.

Обычно коэффициенты G_{ij} и B_{ij} симметричны (по модулю) по обеим координатам и, за исключением специальных случаев, вещественны, что соответствует синфазной антенне. Расфазировку и последующее оптимальное сжатие диаграммы направленности (комплексные коэффициенты) используют иногда в многоканальных РЛС с активными ФАР для облучения широкой зоны обзора при сохранении высокой разрешающей способности по направлению.

Система типа пространственно-временного фильтра или коррелятора на ФАР обладает рядом преимуществ перед устройствами на антеннах со сплошным раскрывом. Эти преимущества объясняются как достоинствами самих ФАР, так и результатом совместной пространственно-временной обработки, недостижимым при раздельной обработке. К достоинствам ФАР относится возможность одновременной работы по нескольким объектам в самых различных режимах (поиск, захват, слежение). Благодаря гибкому электронному управлению лучом или несколькими лучами одновременно значительно возрастает темп обзора пространства (возможен практически мгновенный обзор всей заданной зоны), а также качество обнаружения и точность измерений. Они обладают высокой механической прочностью, не требуют антенного привода и громоздких дорогих антенных обтекателей, не

нуждаются в прецизионной обработке поверхности по лекалам и благодаря большому числу однотипных элементов удобны для массового поточного производства. Снижаются энергетические затраты, так как в РЛС со сплошными подвижными антеннами иногда до 50% мощности расходуется на громоздкий антенный привод, а наличие люфтов, трения и влияние ветровой нагрузки приводит к дополнительным угловым ошибкам. Надежность РЛС на ФАР исключительно высока: выход из строя до 20...30% элементов в случайной последовательности практически не снижает угловой точности и разрешающей способности. К сожалению, ФАР пока еще очень дороги, тяжелы, имеют повышенную шумовую температуру и большие потери мощности на высокочастотных элементах. Однако эти недостатки с течением времени преобладают.

Совмещение пространственной и временной обработки при наличии ФАР позволяет реализовать такие дополнительные преимущества, как сверхбыстрый обзор в течение одного импульса, гибкий последовательный поиск, учитывающий возможность принятия решения на каждом очередном шаге и поэтому дающий дополнительное снижение затрат энергии и времени, а также другие результаты, которые нельзя получить при отдельной обработке.

4.2. ПРОСТРАНСТВЕННЫЙ ТРАКТ СОГЛАСОВАННОГО ПРИЕМНИКА РАДИОСИСТЕМ С АНТЕННЫМИ РЕШЕТКАМИ

Многие характеристики и даже сама возможность реализации оптимального пространственно-временного приемника зависят от вида ФАР, применяемых в системе. По виду связи с приемно-передающим трактом ФАР подразделяются на облучаемые и кондуктивные.

*Облучаемые ФАР** (рис. 4.2) бывают линзовые (рефракторные) и зеркальные (рефлекторные). Они выполняют соответственно роль линзы или зеркала и по ряду показателей сходны со сплошными антеннами соответствующего вида. В частности, излучение и прием в каждом периоде посылок возможны только в одном направлении. Однако благодаря безынерционному электрическому сканированию они допускают быстрое (от импульса к импульсу) переключение с одного направления на другое, что в обычном масштабе времени равносильно параллельной работе в нескольких направлениях.

По сложности зеркальные и линзовые ФАР примерно одинаковы, только линзовые содержат не одну, а две решетки — входную и выходную. Управляющие фазовращатели используют для

* Облучаемые антенны часто называют антеннами с оптическим облучением.

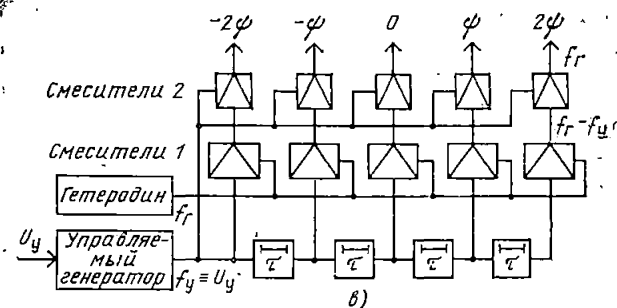
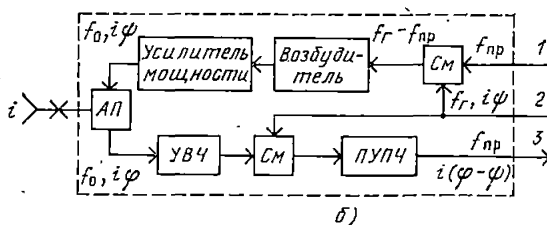
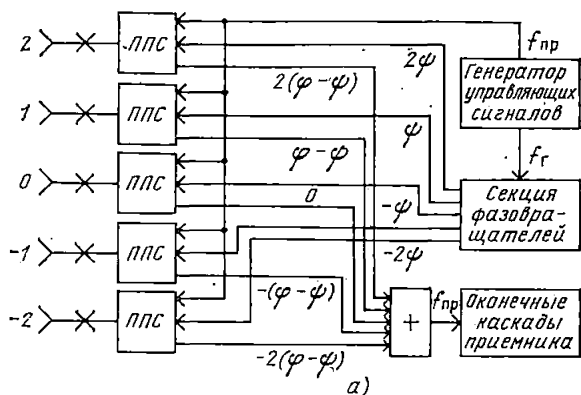


Рис. 4.4. Схемы сканирующей системы с пространственным трактом корреляционного типа на приемно-передающих секциях в канале каждого элемента ФАР (а), ФАПЧ (б) и секция фазовращателей (в)

на секционированной линии задержки путем изменения частоты управляемого генератора f_y пропорционально приложенному напряжению U_y . С отводов линии задержки снимаются напряжения с нужными фазовыми соотношениями, но переменной частотой f_y . Чтобы перенести эти фазовые соотношения на гетеродинное напряжение постоянной частоты f_g , используют два набора смесителей. В первом полученная разность фаз переносится на колебания разностной частоты $f_g - f_y$, которая также непостоянна, а во втором наборе — на колебания частоты f_g стабилизированного гетеродина, суммарной по отношению к частотам входных напряжений вторых смесителей.

Замена фазовращателей пространственного фильтра смесителями пространственного коррелятора полностью исключает потери энергии излучаемого и принимаемого сигналов в фазовращателях. Секция фазовращателей работает всегда в благоприятном режиме в отличие от управляющих фазовращателей пространственных фильтров, где проходящий через них сигнал либо слишком сильный (излучение), либо слишком слабый (прием). Следует при этом отметить, что фазовращатели, рассчитанные на большую мощность излучения, обладают низкими шумовыми свойствами, что снижает чувствительность приемника.

Достоинства приемных устройств типа пространственного коррелятора наиболее отчетливо проявляются в системах с плоскими двумерными ФАР, принцип управления настройкой которых на заданное направление схематически иллюстрируется рис. 4.5 для решетки из 3×3 элементов.

В системах с приемниками типа пространственного фильтра возможно как непосредственное управление настройкой системы на заданное направление, так и с помощью пересчетной схемы. Непосредственное управление настройкой разделяется по каждой из двух угловых координат достигается с помощью пары фазовращателей, последовательно включенных в каналах каждого элемента ФАР (рис. 4.5, а). Все азимутальные и все угломестные фазовращатели совершенно идентичны и управляются как одно целое по заданным значениям азимута и угла места соответственно согласно формулам (4.11), (4.12). Таким образом, общее число фазовращателей в устройстве этого типа равно удвоенному числу элементов ФАР. В устройстве управления с пересчетной схемой (рис. 4.5, б) фазовращателей вдвое меньше, поскольку в каналах элементов ФАР включается только по одному фазовращателю. Пересчет углов настройки u^σ_x и u^σ_y в фазовый сдвиг φ_{ij} канала ij -го элемента производится в соответствии с формулами (4.12), (4.15), где i и j — номера рядов и этажей решетки, отсчитывае-

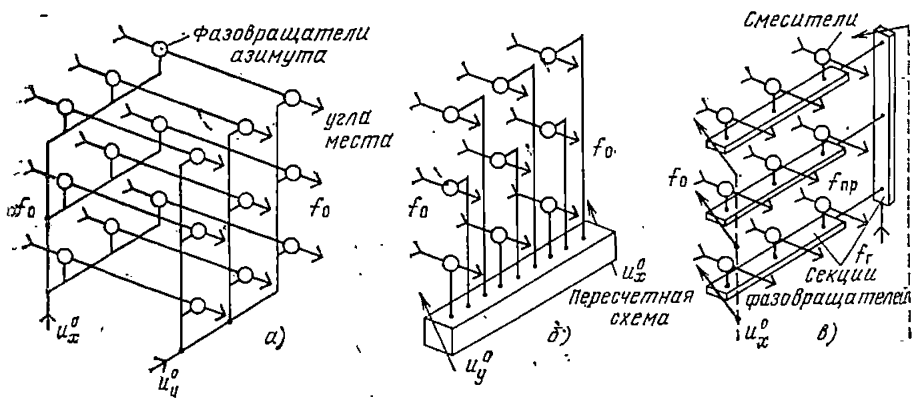


Рис. 4.5. Управление двумерной антенной решеткой в приемниках типа пространственного фильтра с независимой (а) и совместной (б) настройкой и в приемнике типа пространственного коррелятора (в)

мые от середины. Фазовые сдвиги в фазовращателях всех каналов отличаются друг от друга, поэтому каждый фазовращатель управляется индивидуально. Обычно вносимые фазовращателями фазовые сдвиги квантуются, а пересчетная схема выполняется в виде специализированной ЭВМ.

В системах с приемниками типа пространственного коррелятора вместо фазовращателей используют смесители, на которые через секции фазовращателей подаются гетеродинные напряжения частоты f с нужными фазовыми сдвигами (рис. 4.5, в). Один из вариантов секции фазовращателей приведен на рис. 4.4, в. Поскольку схема управления настройкой системы по углу имеет только один входной сигнал гетеродина, число секций фазовращателей существенно меньше числа фазовращателей в пространственном фильтре, где входной сигнал у каждого фазовращателя свой. В данном случае включается одна секция фазовращателей: настройки по углу места и набор секций настройки по азимуту — по числу этажей решетки, причем все секции идентичны и управляются как одно целое. Тем самым в пространственном корреляторе обеспечивается независимая настройка по азимуту (ψ°_α) и углу места (ψ°_γ) согласно формулам (4.11), (4.12) непосредственно, без пересчетной схемы. К сожалению, при гетеродинировании каналов всех элементов ФАР существенно удлиняется высокочастотный фидерный тракт.

Системы на кондуктивных ФАР с приемниками типа пространственного фильтра или коррелятора допускают мгновенный многоканальный обзор всей зоны при условии, что в режиме излучения она целиком облучается одним из указанных выше способов. Затраты времени на обзор минимальны и определяются только временем запаздывания сигнала от самого удаленного объекта. Отличительной особенностью таких систем является многоканальный выход (рис. 4.6). В этих системах для пространственной об-

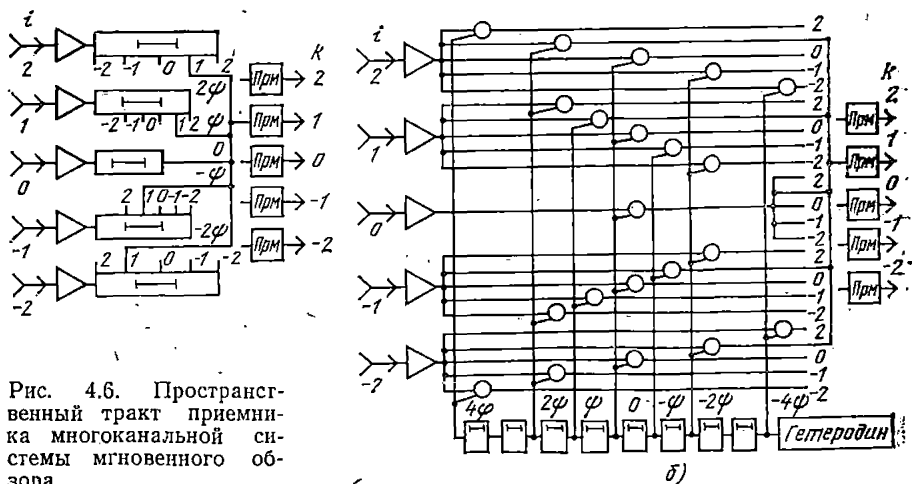


Рис. 4.6. Пространственный тракт приемника многоканальной системы мгновенного обзора

работки сигналов ФАР можно использовать также оптические устройства, которые по своей природе эквивалентны многоканальным электронным устройствам.

Структура многоканальных систем мгновенного обзора непосредственно вытекает из структуры сканирующих систем, если единственный последовательно перестраиваемый канал обработки сигналов заменить набором таких каналов, настроенных на несколько фиксированных направлений. Для плотного перекрытия зоны обзора по любой из двух угловых координат направления настройки разделяются интервалом, примерно равным ширине основного лепестка диаграммы направленности ФАР:

$$\Delta u_{x,y} = 1/\Delta v_{x,y} = \lambda/d_{x,y}, \quad (4.16)$$

где $d_{x,y}$ — линейный размер решетки, а $\Delta v_{x,y} = d_{x,y}/\lambda$ — ширина углового спектра (относительный размер). Для исключения связи между лучами угол разноса выбирают таким, чтобы направление максимума одного луча совпадало с направлением нуля соседнего луча. Для фазирования i -го элемента на k -е направление $u^0_{x,y,k} = k\Delta u_{x,y}$ вводится фазовый сдвиг

$$\varphi^0_i(k) = 2\pi i (b/\lambda) k \Delta u^0_{x,y} = ik\psi, \quad (4.17)$$

где

$$\psi = 2\pi (b/\lambda) \Delta u^0_{x,y} \quad (4.18)$$

— вносимый фазовый сдвиг между сигналами соседних лепестков ФАР; b — шаг решетки.

Фазирование (настройка) системы на заданные направления производится в пространственном фильтре по высокой частоте с помощью фазовращателей (рис. 4.6,а), а в пространственном корреляторе по промежуточной частоте с помощью смесителей, на которые подаются гетеродинные напряжения через фиксированные фазовращатели (рис. 4.6,б). Фазовращатели бывают узкополосные и широкополосные. Широкополосные фазовращатели, выполненные на линиях задержки (показаны на рис. 4.6,а), пригодны для любых сигналов, включая сигналы с большой относительной шириной спектра $\Delta f/f_0$. Разность времен прихода на существенно различных частотах спектра широкополосных сигналов дает различный фазовый сдвиг, ограничиться компенсацией фазы лишь на частоте несущей нельзя, поэтому перед суммированием сигналов решетки компенсируют непосредственно разность времен их запаздывания. У узкополосных сигналов ввиду малого различия частот в их спектре достаточно сфазировать сигналы решетки на частоте несущей.

Для снижения уровня шума в канале каждого элемента ФАР обычно включают усилитель с согласованной полосой пропускания и достаточно большим коэффициентом усиления или же полностью приемно-передающие секции рассмотренного выше типа. В последнем случае благодаря наличию генераторов в каждом канале решетка излучает очень большую мощность, необходимую для одновременного облучения всей зоны обзора.

Однако большие возможности многоканальных систем мгновенного обзора пространства на кондуктивных ФАР достигаются дорогой ценой. Особенной сложностью отличается фидерный тракт, общая протяженность которого в некоторых системах составляет несколько десятков километров. Ввиду этого для таких систем характерны значительные энергетические потери и фазовые искажения в фидерах, а наличие в них большого числа гибридных сочленений и других высокочастотных элементов создает серьезные трудности при изготовлении, настройке и эксплуатации системы.

Как компромисс между сложностью многоканальных систем мгновенного обзора и недопустимо большим периодом обзора заданной зоны в одноканальных системах со сканированием по обеим угловым координатам используют системы смешанного обзора. Они сканируют только по одной угловой координате (обычно по азимуту), а по второй производится многоканальный мгновенный обзор. При этом период обзора сокращается во столько раз, сколько угловых каналов укладывается по второй координате. Следует также указать на одно важное для использования смешанных систем обстоятельство. Во многих случаях желателен круговой обзор по азимуту и обзор в сравнительно узком секторе по углу места. Поскольку в плоских неподвижных антенных решетках сектор мгновенного или последовательного обзора ограничен ($90 \dots 100^\circ$) из-за недопустимого расширения луча при больших углах отклонения его максимума от нормали к раскрыву и соответствующего снижения коэффициента направленного действия (КНД), для кругового обзора по азимуту практически требуется не одна, а четыре ФАР системы мгновенного обзора с громоздким пространственным трактом. Если же использовать круговое вращение антенны по азимуту, можно ограничиться одной системой, имеющей ограниченное число каналов лишь по углу места. Таким образом, замена многоканальных систем смешанными дает огромный выигрыш в объеме и сложности аппаратуры при достаточно малом периоде обзора, обеспечивающем нужную частоту обновления информации об объектах, находящихся в зоне обзора. При всем этом даже системы смешанного обзора по двум угловым координатам оказываются в 4...10 раз дороже систем, измеряющих только одну угловую координату (и дальность).

4.3. ПОДОПТИМАЛЬНЫЕ СТРУКТУРЫ ПРОСТРАНСТВЕННОГО ТРАКТА РАДИОСИСТЕМ

Ввиду сложности пространственного тракта оптимального (согласованного) приемника на практике используют устройства, в которых за счет некоторого ухудшения характеристик системы или ограничений ее использования существенно упрощается сама структура системы. Наибольшие возможности для этого имеют системы с кондуктивными антенными решетками. Среди большого числа вариантов реализации таких подоптимальных систем рассмотрим три

характерных примера: частотное сканирование, скрытое сверхбыстрое сканирование и применение диаграммообразующих матриц.

Частотно-зависимые системы обзора. При описании оптимальных систем предполагалось, что средняя частота спектра сигнала в процессе работы системы остается постоянной, благодаря чему характеристики системы сохраняются неизменными, а принимаемые сигналы когерентными. Это, однако, не исключает перестройки системы на другую достаточно близкую частоту в случае появления помехи на данной частоте. Близость перестраиваемых частот позволяет использовать одни и те же элементы антенно-фидерного тракта и сохранить практически неизменными основные показатели системы (точность, разрешающую способность).

Если отказаться от условия постоянства средней частоты излучаемого сигнала и пренебречь связанными с этим потерями, можно использовать частотное сканирование антенного луча, при котором переменный фазовый сдвиг между сигналами соседних элементов решетки создается при прохождении сигнала с изменяющейся частотой несущей через элементы задержки.

Принцип частотного сканирования в ФАР с параллельным возбуждением состоит в следующем. Диаграмма направленности имеет максимум в том направлении $u^0_{x,y}$, где разность фаз сигналов, принимаемых соседними элементами решетки, $2\pi(b/\lambda)u^0_{x,y}$ совпадает с разностью фаз $2\pi l/\Lambda$, вызванной разностью их хода l до общей точки фидера, длина волны в котором равна Λ , или отличается на целое число 2π , т. е. $2\pi(b/\lambda)u^0_{x,y} = 2\pi l/\Lambda - 2\pi n$. Целое число n выбирают таким, чтобы этот фазовый сдвиг не превышал заданного. Из последнего выражения находим направление максимума луча

$$u^0_{x,y} = (l/b)\lambda(1/\Lambda - n/l). \quad (4.19)$$

Таким образом, зависимость углового положения луча от длины волны (частоты) сигнала тем сильнее, чем больше длина фидера l между элементами решетки по сравнению с их шагом b . Для повышения l/b — чувствительности углового отклонения к изменению частоты — питающий решетку фидер выполняют в виде пространственной спирали или плоской змейки (рис. 4.7, а). На конце замедляющих систем включается поглощающая нагрузка для создания режима бегущей волны в фидере. Конструкция в виде змейки более удобна. Однако подключение к ней излучателей в местах изгиба (как показано на рисун-

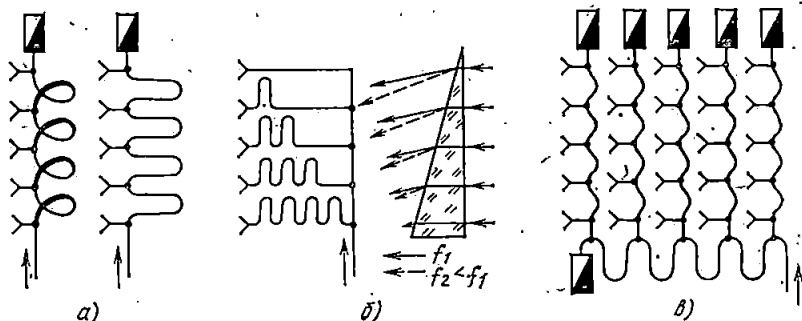


Рис. 4.7. Частотное сканирование по одной угловой координате (а, б) и по обеим угловым координатам (в)

ке) трудно реализовать, так как на малом шаге решетки (полволны) необходимо разместить два волновода по узкой стороне, и, кроме того, условия возбуждения излучателей решетки в местах изгиба волновода неблагоприятны. Поэтому излучатели обычно делают в виде щелей по узкой стороне волновода; при этом шаг при той же длине петли сокращается вдвое (два волновода в петле), а разность фаз π в двух встречных волноводах петли компенсируется тем, что щели выполняют с противоположным наклоном.

Как известно, длина волны в волноводе (для основного типа волны) и коаксиальном кабеле соответственно определяется известными соотношениями $\Lambda = \lambda / \sqrt{1 - (\lambda/2a)^2}$ и $\Lambda = \lambda / \sqrt{\epsilon_0}$, где a — размер волновода по широкой стороне; ϵ_0 — диэлектрическая проницаемость изоляции коаксиального кабеля. Тогда закон сканирования (4.19) в функции частоты $f = c/\lambda$ для волновода

$$u_{x,v}^0 = (l/bf) \left(\sqrt{f^2/c^2 - 1/4a^2} - nc/l \right), \quad (4.20)$$

а для коаксиального кабеля

$$u_{x,v}^0 = (l/bf) \left(f \sqrt{\epsilon_0} - nc/l \right). \quad (4.21)$$

Среднее значение частоты сигнала при сканировании обычно выбирают из условия $u_{x,v}^0 = 0$. Девияция частоты и длина отрезка l определяют пределы сканирования по углу.

Частотное сканирование возможно также в решетке с параллельным возбуждением и эквивалентной ей антенне в виде диэлектрической линзы, облучаемой параллельным пучком лучей (рис. 4.7.б). Различная степень замедления сигналов по раскрыву антенны приводит к тому, что направление излучения на ее выходе отклоняется в сторону большего замедления. Отклонение тем больше, чем длиннее волна, т. е. меньше частота, что и позволяет производить сканирование, изменяя частоту.

Решетки рассмотренного вида, возбуждаемые широкополосными сигналами (например, ЛЧМ импульсами), можно использовать для мгновенного одновременного обзора заданного сектора пространства. Попадающий в зону обзора объект отражает сигнал только той частоты из всего спектра излучаемого сигнала, которая излучается в данном направлении. Измеряя частоту этого сигнала, можно определить угловое положение объекта. Именно это происходит при разложении широкополосного белого света призмой на все цвета радуги, где колебания наиболее низкочастотного, красного цвета отклоняются на больший угол, чем все остальные. Поскольку каждому цвету на выходе призмы соответствует свое направление, можно определить угловое положение неокрашенного предмета, помещенного в освещенную зону, по цвету (частоте) отраженных от него колебаний, наблюдаемых визуально.

В частотно-зависимых системах мгновенного обзора, где каждому направлению соответствует своя частота принимаемого сигнала, спектр и форма сигнала отличаются от излучаемого. Ввиду этого в таких системах не реализуются потенциальные возможности. Во-первых, снижается точность измерения дальности, поскольку из всего спектра сигнала вырезается лишь некоторая его часть. Во-вторых, возникает неопределенность угол — радиальная скорость, так как средняя частота спектра принимаемого сигнала зависит и от углового положения, и от радиальной скорости объекта. Эта неопределенность проявляется и в сканирующих частотно-зависимых системах.

Частотное сканирование возможно одновременно по двум угловым координатам, если чувствительность к изменению частоты по одной из координат сделать в несколько раз выше, чем по другой, включив питающий фидер с большим замедлением (рис. 4.7, в). При этом закон сканирования подобен телевизионной развертке, быстрой в горизонтальной плоскости — по строке (высокая чувствительность к изменению частоты) и медленной в вертикальной — по кадру (низкая чувствительность). При изменении частоты сигнала фаза питающих отводы напряжений столько раз переходит через 2π за время полного отклонения луча по вертикали, сколько строк в кадре. Сектор сканирования в горизонтальной плоскости лежит при этом в пределах $\pm\pi/2$ по углу (± 1 по направляющему косинусу), т. е. больше, чем обычно требуется. Это приводит к напрасным затратам энергии и времени на обзор пространства вне заданного сектора. Потери энергии можно исключить, если предусмотреть выключение передатчика при нахождении луча вне заданного сектора, но потери времени неизбежны. Если же сектор обзора по направляющим косинусам лежит в пределах $\pm 0,5$ ($\pm\pi/6$ по углу), можно, изменяя фазу сигнала на π в момент достижения лучом этих границ, повторно за цикл изменения фазы произвести сканирование по строке в тех же пределах. При этом потери времени и энергии полностью исключаются, а требования к девиации частоты снижаются вдвое.

Для частотного сканирования необходимо обеспечить высокую точность выполнения и настройки тех элементов системы, которые определяют закон сканирования. При сканировании по обеим угловым координатам эти требования и отступления системы от оптимальности значительно возрастают.

Системы со скрытым сверхбыстрым сканированием. Важным достоинством приемников типа пространственного коррелятора в системах на кондуктивных ФАР является возможность сверхбыстрого (за длительность импульса) скрытого сканирования заданного сектора в режиме приема. Если при передаче облучается весь сектор, то такой метод работы равносителен мгновенному обзору пространства с помощью приемных устройств, имеющих одноканальный перестраиваемый выход с временным или частотным разделением сигналов по угловой координате (рис. 4.8).

Принцип действия одной из систем с временным разделением сигналов (рис. 4.8, а) состоит в том, что гетеродинное напряжение частоты f_r смешивается со сложным периодическим напряжением частоты f_c , богатым гармониками (например, меандром), поступающим от модулятора. Гармоники биений $f_r, f_r \pm f_c, f_r \pm 2f_c$ и т. д. выделяются фильтрами и подаются на соответствующие смесители в каналах каждого элемента ФАР. Поскольку разность фаз сигналов, которые принимаются с направления $u_{x,y}$ в каналах двух соседних элементов решетки, разнесенных на интервал b , равна $2\pi(bu_{x,y}/\lambda - fct)$, максимум выходного сигнала наступает, когда сигналы всех элементов решетки в канале УПЧ синфазны, т. е. выражение в скобках равно нулю. Отсюда находим, что направление максимума приема

$$u_{x,y} = (\lambda/b) fct, \quad -T_{ск}/2 \leq t \leq T_{ск}/2, \quad (4.22)$$

является кусочно-линейной функцией времени: диаграмма направленности ФАР в режиме приема сканирует по углу с периодом $T_{ск} = 1/f_c$ в пределах $-1 \leq u_{x,y} \leq 1$.

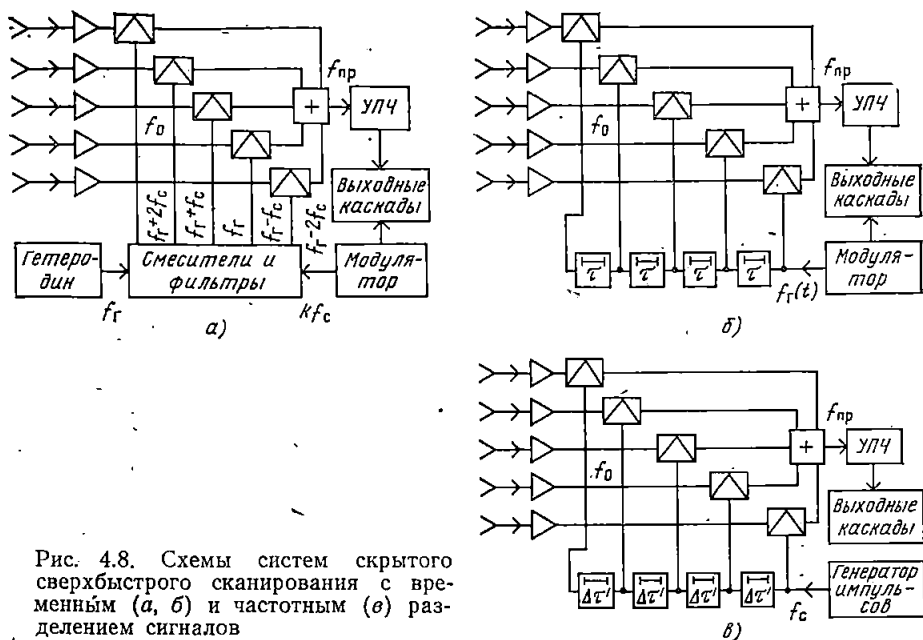


Рис. 4.8. Схемы систем скрытого сверхбыстрого сканирования с временным (а, б) и частотным (в) разделением сигналов

Для нормальной работы системы необходимо, чтобы спектры сигналов соседних каналов, имеющие ширину Δf и разнесенные на f_c , не перекрывались, т. е. $f_c \geq \Delta f$. При $f_c = \Delta f$ период сканирования $T_{ск} = 1/f_c$ равен интервалу корреляции сигнала $\Delta\tau = 1/\Delta f$, т. е. длительности импульса на выходе согласованного фильтра (для простых импульсных сигналов, без внутримпульсной модуляции, длительности излучаемого импульса). Таким образом, обзор всего сектора происходит за время приема одного импульса. Если число элементов ФАР равно N , то общая ширина спектра сигнала в канале УПЧ составляет $N\Delta f$ (частотное уплотнение сигналов). Фазовые соотношения между сигналами элементов ФАР по промежуточной частоте на интервале $T_{ск} = \Delta\tau$ для объекта, находящегося на некотором направлении $u^0_{x,y} = (\lambda/b)f_c t$, благоприятны только в течение небольшого отрезка времени $\Delta\tau' = \Delta\tau/N = 1/N\Delta f$. В результате из импульса длительностью $\Delta\tau = 1/\Delta f$ вырезается в N раз более короткий импульс $\Delta\tau'$, временное положение которого на интервале $T_{ск}$ указывает направление объекта $u^0_{x,y}$. Происходит временное разделение сигналов, принимаемых с различных направлений.

Съем угловой координаты можно произвести цифровым и аналоговым способами. При цифровом способе промежуток времени между началом периода $T_{ск}$ и коротким импульсом на выходе заполняется счетными импульсами. Показания счетчика импульсов пропорциональны направляющему косинусу. При аналоговом способе используется быстрая линейная развертка, начало которой совпадает по времени с началом периода $T_{ск}$. Угловая координата определяется по положению отметки, создаваемой на экране индикатора коротким импульсом, одновременно с измерением дальности по положению этой отметки относительно зондирующего импульса на второй координате экрана с более медленной разверткой.

Сокращение длительности импульса и расширение спектра по промежуточной частоте в N раз не ведет к энергетическим потерям, если во входной цепи каждого элемента ФАР стоит усилитель с согласованной полосой пропускания и большим коэффициентом усиления, достаточным для того, чтобы на сигнал не влияли шумы последующих цепей. Амплитуда же короткого импульса равна амплитуде исходного импульса, из которого он вырезается при сканировании, а форма повторяет диаграмму направленности ФАР.

В другой системе с временным разделением (частотным уплотнением) сигналов (рис. 4.8,б) используется гетеродинное ЛЧМ напряжение с периодом $T_{\text{ск}}$, подаваемое через секционированную линию задержки. Здесь частота гетеродинного напряжения периодически (с периодом $T_{\text{ск}}$) изменяется по линейному закону $f_{\text{г}}(t) = f_{\text{г}0} + at$, где $f_{\text{г}0}$ — среднее значение частоты гетеродина; a — скорость ее изменения. При этом разность фаз гетеродинных напряжений, снимаемых с двух соседних отводов линии, при условии $2\pi f_{\text{г}0} \tau = 2\pi n \equiv 0$ (где n — целое число) оказывается кусочно-линейной функцией времени $\psi = 2\pi(f_{\text{г}0} + at)\tau = 2\pi at\tau$. Эта разность фаз переносится на сигналы соседних элементов ФАР, снимаемые с выхода смесителей. Результирующая разность фаз этих сигналов в канале УПЧ с учетом разности времени их прихода с направления $u_{x,y}$ равна $2\pi(bu_{x,y}/\lambda - at\tau)$, а выходной сигнал сумматора максимален, когда она достигает нуля. Отсюда следует, что направление максимума приема

$$u_{0x,y} = (\lambda/b)at\tau, \quad -T_{\text{ск}}/2 \leq t \leq T_{\text{ск}}/2, \quad (4.23)$$

линейно изменяется в течение интервала $T_{\text{ск}}$.

Если заданы симметричный сектор сканирования $\pm |u_{x,y}|_{\text{м}}$ и период сканирования $T_{\text{ск}} = \Delta\tau = 1/\Delta f$, то скорость изменения частоты определяется соотношением $a = 2\Delta f b |u_{x,y}|_{\text{м}}/\lambda \tau$ и зависимость (4.23) принимает вид

$$u_{0x,y}/2 |u_{x,y}|_{\text{м}} = \Delta f t, \quad (4.24)$$

т. е. относительное изменение угла определяется только шириной спектра импульса. В данном случае, как и в предыдущем, в канале УПЧ происходит расширение спектра и сжатие импульса в N раз. Однако второй способ имеет серьезные преимущества перед первым. Во-первых, исключаются потери времени и излишне высокая скорость сканирования, так как обычно требуемый сектор сканирования $2|u_{x,y}|_{\text{м}} < \pi$, а при первом способе он сверх требуемого простирается до π . Во-вторых, облегчается формирование гетеродинных напряжений, подводимых к смесителям в каналах элементов ФАР (исключаются резонансные фильтры, предназначенные для выделения гармоник биений). Недостатком второго способа является резкое изменение частоты гетеродина при переходе от одного периода к другому, что приводит к появлению нестационарных процессов. Для исключения последних пилообразный закон изменения частоты заменяют более плавным, что также ведет к непроизводительным затратам времени, хотя и меньшим.

В системе с частотным разделением (рис. 4.8,в) в качестве гетеродинных используются импульсные напряжения с несущей $f_{\text{г}}$, частотой следования $f_{\text{с}} = \Delta f$ и длительностью $\Delta\tau' = 1/Nf_{\text{с}}$, в N раз меньшей длительности сигнала импульса $\Delta\tau = 1/\Delta f$. Гетеродинные напряжения подаются на смесители каналов элементов ФАР с эквидистантных отводов линий задержки через интервалы $\Delta\tau'$, равные длительности гетеродинных импульсов. Таким образом, каждый

импульс сигнала перекрывается N гетеродинными импульсами (временное уплотнение). Разность фаз сигналов двух соседних элементов ФАР на промежуточной частоте $f_{\text{пр}}$ составляет $2\pi(bu_{x,y}/\lambda - f_{\text{пр}}\Delta\tau') = 2\pi bu_{x,y}/\lambda$, так как $2\pi f_{\text{пр}}\tau' = 2\pi n \equiv 0$, а на частоте $f_{\text{пр}} + if_c$ спектра последовательности гетеродинных импульсов $2\pi(bu_{x,y}/\lambda - if_c\Delta\tau') = 2\pi(bu_{x,y}/\lambda - i/N)$, так как $f_c\Delta\tau' = 1/N$. Поэтому благоприятные фазовые соотношения (нулевая разность фаз) для направления прихода сигнала

$$u^0_{x,y} = (\lambda/b) (i/N) \quad (4.25)$$

существуют только на биениях с определенной гармоникой частоты следования if_c гетеродинных импульсов (частотное разделение). В этой схеме для отсчета направления на объект выходной каскад приемника в принципе должен включать спектроанализатор, однако частотное разделение можно свести к временному, несколько усложнив эту схему.

Рассмотренные системы со скрытым сверхбыстрым сканированием обладают рядом достоинств. Они реализуют действительно оптимальный прием сигнала, однако возникает неопределенность вида угол — дальность в пределах длительности импульса, что приводит к снижению точности измерения дальности. На сканирование по углу не требуется дополнительного времени (в отличие от частотного сканирования), благодаря чему они оказываются эквивалентными значительно более сложным многоканальным системам мгновенного обзора. Ввиду значительного расширения спектра сигнала в канале УПЧ и связанных с этим технических трудностей сверхбыстрое сканирование обычно осуществляется только по одной угловой координате. Однако и этого достаточно для существенного сокращения (в несколько раз!) времени обзора заданной зоны по сравнению с обычными сканирующими системами. Необходимое для работы по отраженному или переизлученному сигналу облучение сектора скрытого сканирования обеспечивается либо отдельным слабонаправленным излучателем, либо расфазировкой решетки (комплексная функция раскрыва).

Многоканальные и коммутируемые системы с диаграммообразующими матрицами. В многоканальных и коммутируемых по угловым координатам радиосистемах с ФАР иногда используют диаграммообразующие матрицы; выполняемые в виде волноводных сочленений. В этой матрице совокупность сигналов отдельных каналов приемно-передающего тракта преобразуется в совокупность сигналов, излучаемых элементами ФАР при передаче, а при приеме происходит обратное преобразование этих совокупностей. Наиболее просто реализовать матрицу, когда распределение поля по раскрытию решетки равномерное, а число элементов решетки и антенных лучей равно $N=2^n$ (матрица $N \times N$). Такая матрица содержит n рядов делителей мощности и некоторое количество фиксированных фазовращателей. Делители мощности представляют собой волноводные мосты с двумя входами и двумя выходами. Мощность подводимых к любому входу колебаний делится между двумя выходами поровну, поэтому в n рядах делителей входная мощность уменьшается в $N=2^n$ раз. Волноводные мосты, число которых в каждом ряду равно $N/2$, бывают трех типов: щелевые, кольцевые и двойные T -образные.

Структура матрицы определяется типом используемых в ней волноводных мостов и формулами, которые связывают вводимые фазовые сдвиги с по-

ложением дискретных лучей. Направление максимума k -го луча, отсчитываемое от нормали к раскрыву, при четных N

$$u_{x,y}^0 = k(1 - k/2|k|) \Delta u_{x,y}, \quad k = \pm 1, \pm 2, \dots, \pm N/2, \quad (4.26)$$

где $\Delta u_{x,y}$ — разнос лучей, рассчитываемый из условия развязки сигналов в смежных каналах, при котором максимум любого луча совпадает с первым нулем соседнего:

$$\Delta u_{x,y} = \lambda/d_{x,y} = \lambda/Nb, \quad (4.27)$$

где $d_{x,y} = Nb$ — размер решетки. При этом условии направление максимума k -го луча

$$u_{x,y}^0 = k(1 - k/2|k|) \lambda/Nb. \quad (4.28)$$

Фазовый сдвиг, вводимый между сигналами соседних элементов решетки для формирования k -го луча, в соответствии с формулами (4.11) и (4.26), — (4.28) определяется соотношением

$$\psi_k = -2\pi \left(\frac{c}{\lambda} \right) k \left(1 - \frac{1}{2k} |k| \right) \Delta u_{x,y} = - \frac{2\pi k (1 - k/2|k|)}{N}. \quad (4.29)$$

Отсюда согласно (4.12) находим фазовый сдвиг, вводимый в канал i -го элемента решетки для образования k -го луча,

$$\varphi_i^0(k) = \psi_k i(1 - i/2|i|) + \varphi_0, \quad (4.30)$$

где φ_0 — постоянный набег фазы в волноводных мостах матрицы, обеспечиваемый при минимальном числе фиксированных фазовращателей.

В качестве примера на рис. 4.9 приведена диаграммообразующая матрица Батлера на щелевых мостах для решетки из $N=8$ элементов, входы которой пронумерованы в соответствии с номером луча k . Колебания по любому из входов щелевого моста проходят на прямой выход без сдвига по фазе, а на смещенный выход — со сдвигом на $\pi/2$. В результате образуется линейный набег фазы по раскрыву решетки, причем скорость его нарастания или спадаания определяется номером луча. Так, для луча $k=1$ (первый справа от нормали к раскрыву) вводимый фазовый сдвиг от одного элемента к другому спадает согласно формуле (4.29) на $\psi_1 = \pi/8$ и при $\varphi_0 = 17\pi/16$ в соответствии с формулой (4.30) у каждого элемента антенной решетки последовательно принимает значения $\varphi_{-4}^0 \dots \varphi_4^0 = 12\pi/8, 11\pi/8, 10\pi/8, 9\pi/8, 8\pi/8, 7\pi/8, 6\pi/8, 5\pi/8$. Скорость спадаания фазы по раскрыву для луча 2 вдвое больше, для луча 3 — вчетверо и т. д. Для лучей слева (отрицательные k) фаза нарастает.

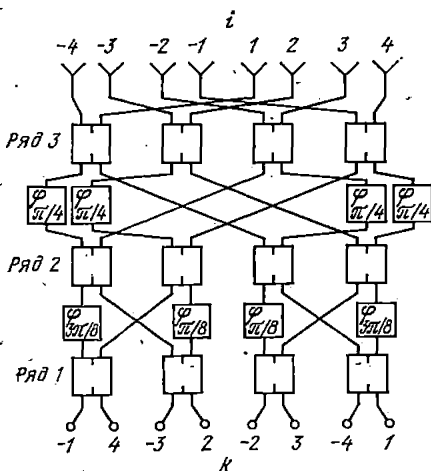


Рис. 4.9. Диаграммообразующая матрица на щелевых волноводных мостах

Если к каждому входу матрицы подключить свой приемопередатчик,

возбуждаемый от общего генератора управляющих сигналов, получим многоканальную систему мгновенного обзора пространства, как бы состоящую из отдельных одноканальных систем, каждая из которых действует по своему лучу. Если же использовать только один приемопередатчик, подключаемый ко входам матрицы через коммутатор каналов, получим одноканальную коммутируемую систему последовательного обзора. В коммутируемой системе в отличие от сканирующих пакет импульсов принимаемого сигнала имеет прямоугольную огибающую, а не повторяет диаграмму направленности решетки вида $\sin Nx/\sin x$.

Недостатком систем, использующих диаграммообразующие матрицы, является жесткое ограничение на вид функции раскрытия — только равномерная, что определяет очень высокий уровень боковых лепестков (21% по напряженности поля). Несмотря на относительную простоту матриц высокочастотный тракт получается очень громоздким, значительно сложнее, чем в системах с частотным или скрытым сканированием. Особенно громоздкой получается матрица у систем с плоской двумерной решеткой.

Описанные выше пространственно-временные приемные устройства радиосистем формируют выходной сигнал (в оптимальном случае логарифм отношения правдоподобия), который используется обычно для обнаружения объектов с одновременной сравнительно грубой оценкой параметров их положения и движения. Для точного измерения этих параметров по методу максимального правдоподобия строятся дополнительные каналы, рассматриваемые в последующих главах.

Глава 5. ВЗАИМОСВЯЗЬ И ПРЕОБРАЗОВАНИЕ ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ХАРАКТЕРИСТИК В РАДИОСИСТЕМАХ

5.1. ФИЗИЧЕСКИЙ СМЫСЛ ПРОСТРАНСТВЕННЫХ ЧАСТОТ

Временная частота f (или ω) определяется однозначно: характеризует скорость изменения текущей фазы гармонических колебаний во времени. Аналогично определяется пространственная частота: характеризует скорость изменения фазы гармонической радиоволны по пространственной координате. Однако пространство имеет не одну, а три независимые координаты, поэтому приходится говорить о трех пространственных частотах применительно к каждой из них. Кроме того, координаты подразделяют на линейные и угловые (направляющие косинусы), поэтому в зависимости от того, в функции какой из этих координат описываются процессы в радиосистемах, используют линейные и угловые пространственные частоты, а также соответствующие спектры процессов.

Угловые пространственные частоты

$$v_x = x/\lambda = fx/c; \quad v_y = y/\lambda = fy/c \quad (5.1)$$

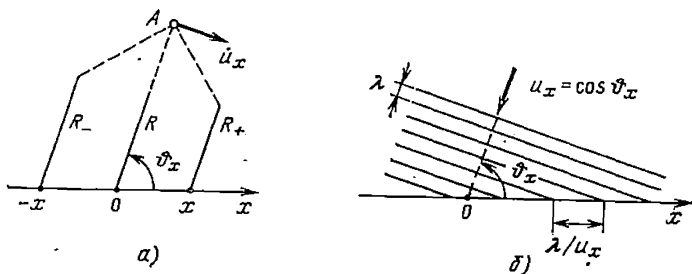


Рис. 5.1. К пояснению физического смысла угловых (а) и линейных (б) пространственных частот

характеризуют скорость изменения фазы плоской гармонической волны частоты f вдоль направляющих косинусов u_x и u_y , отсчитываемых относительно осей x и y раскрыва антенны соответственно:

$$\begin{aligned}\varphi(u_x) &= -2\pi v_x u_x = -2\pi x u_x / \lambda; \\ \varphi(u_y) &= -2\pi v_y u_y = -2\pi y u_y / \lambda.\end{aligned}\quad (5.2)$$

Для пояснения физического смысла угловой пространственной частоты v_x (или v_y) рассмотрим точечный источник, размещенный в точке x , и удаленный анализатор поля A , который мысленно перемещаем в направлении возрастания угловой координаты u_x с постоянной скоростью $\dot{u}_x$, сохраняя постоянным удаление R относительно центра раскрыва O (рис. 5.1, а). Вследствие уменьшения расстояния R_+ между анализатором и источником колебаний происходит набег фазы воспринимаемых колебаний, причем тем быстрее, чем больше расстояние до точки x и меньше λ , т. е. скорость набег фазы определяется пространственной частотой $v_x = x/\lambda$. Если излучатель перенести в центр раскрыва, расстояние от которого до анализатора R не изменяется, набег фазы исчезнет, что соответствует нулевой пространственной частоте. Если же излучение производить из точки $-x$, симметричной x , то вследствие возрастания расстояния R_- возникает запаздывание по фазе, скорость нарастания которого определяется отрицательной пространственной частотой $v_x = -x/\lambda$, т. е. угловые пространственные частоты имеют знак. Аналогичный рассмотренному результат получим, если точки излучения и анализа поля поменять местами.

В оптике широко используют линейные пространственные частоты

$$f_x = u_x / \lambda = f u_x / c; \quad f_y = u_y / \lambda = f u_y / c, \quad (5.3)$$

характеризующие скорость изменения фазы плоской волны вдоль линейных координат x и y , перпендикулярных оптической оси системы z :

$$\begin{aligned}\varphi(x) &= -2\pi f_x x = -2\pi x u_x / \lambda; \\ \varphi(y) &= -2\pi f_y y = -2\pi y u_y / \lambda.\end{aligned}\quad (5.4)$$

Правые части выражений (5.2) и (5.4) полностью совпадают, следовательно, введение угловых или линейных пространственных частот совершенно равноправно; все определяется удобством их использования в конкретных обстоятельствах. Если в качестве аргумента какой-либо функции берется угол (направляющий косинус), то пространственной частотой следует считать линейную координату, отнесенную к длине волны, если же линейная координата, то пространственной частотой становится направляющий косинус, отнесенный к длине волны. Общей для обоих видов пространственных частот является обратная пропорциональность длине волны, т. е. прямая пропорциональность частоте колебаний. С повышением частоты расширяется спектр пространственных частот и, следовательно, сужается функция самой координаты (например, диаграмма направленности).

Физический смысл линейной пространственной частоты легко раскрывается через серию эквифазных фронтов (отстоящих друг от друга на λ) волны удаленного источника, приходящей под углом ϑ_x (направляющий косинус $u_x = \cos \vartheta_x$) к раскрыву (рис. 5.1, б). При этом интервал между точками равных фаз вдоль линейной координаты x — пространственный период λ/u_x — тем меньше, а пространственная частота $f_x = u_x/\lambda$ тем выше, чем больше наклон фронта (u_x) и меньше λ . При замене угла прихода волны симметричным (отрицательные u_x) фаза вдоль линейной координаты x не нарастает, а спадает с той же скоростью. Поэтому линейная пространственная частота также имеет знак и равна нулю при падении волны по нормали к раскрыву (пространственный период $\lambda/0$ бесконечен).

Вдоль направления распространения волны R (или оптической оси системы z) пространственным периодом становится длина волны λ , а пространственной частотой —

$$f_R = 1/\lambda = f/c. \quad (5.5)$$

Закон изменения фазы волны в радиальном направлении

$$\varphi(R) = -2\pi f_R R = -2\pi R/\lambda = -2\pi f R/c \quad (5.6)$$

повторяет закон изменения ее во времени (положим $t = -R/c$) на любой частоте f . Следовательно, любой узкополосный радиосигнал как сумма гармоник с заданными амплитудами и фазами повторяет вдоль пути распространения в свободном пространстве свою временную структуру (с масштабным множителем c).

Отметим, что продольную пространственную частоту (5.5) можно получить из формул (5.3), подставив $u_x = u_y = 1$, что соответствует нулевому углу относительно нормали к раскрыву. Широкоизвестный волновой коэффициент

$$\kappa = 2\pi/\lambda = 2\pi f_R = \omega_R \quad (5.7)$$

представляет собой круговую продольную пространственную частоту (в векторной форме он отображает и пространственную частоту, и направление распространения волны).

Выше для пояснения физического смысла пространственных частот мы мысленно перемещали в пространстве приемник (анализатор поля) или источник колебаний. Если такое перемещение происходит в действительности, т. е. координаты точек приема или излучения становятся функциями времени $R=\dot{R}t$, $u_x=\dot{u}_x t$, $u_y=\dot{u}_y t$ (где $\dot{R}$, $\dot{u}_x$, $\dot{u}_y$ — производные координат R , u_x , u_y), то возникает дополнительный линейный набег фазы принимаемого сигнала относительно излученного:

$$\begin{aligned}\varphi_R(t) &= -2\pi \dot{f}_R \dot{R}t = -2\pi \dot{R}t/\lambda = 2\pi \dot{R}t/c = 2\pi F t, \\ \varphi_x(t) &= -2\pi v_x \dot{u}_x t = -2\pi x \dot{u}_x t/\lambda = 2\pi x \dot{u}_x t/c = 2\pi F_x t, \\ \varphi_y(t) &= -2\pi v_y \dot{u}_y t = -2\pi y \dot{u}_y t/\lambda = -2\pi y \dot{u}_y t/c = 2\pi F_y t.\end{aligned}\quad (5.8)$$

Линейному набегу фазы соответствует доплеровский сдвиг частоты — радиальный и два тангенциальных:

$$\begin{aligned}F &= -\dot{f}_R \dot{R} = -\dot{R}/\lambda = -\dot{R}f/c, \\ F_x &= -v_x \dot{u}_x = -\dot{u}_x x/\lambda = -\dot{u}_x f_x/c, \\ F_y &= -v_y \dot{u}_y = -\dot{u}_y y/\lambda = -\dot{u}_y f_y/c.\end{aligned}\quad (5.9)$$

Следовательно, доплеровский сдвиг есть результат перехода пространственных частот во временные, когда сами координаты источника или приемника колебаний становятся функциями времени. Это и положено в основу измерения производных координат — радиальной и угловых скоростей — радиотехническими методами. Согласно соотношениям (5.9) пространственные частоты играют роль коэффициентов пропорциональности между производными координат и доплеровским сдвигом частоты колебаний.

Системы, в которых происходит относительное перемещение приемника и источника колебаний в свободном пространстве, можно назвать пространственно-временными линейными структурами. Отличительной особенностью таких структур, вытекающей из соотношений (5.9), является возможность появления на выходе гармонических составляющих сигнала, которых не было на входе. Рассмотрим пространственно-временную линейную структуру в виде излучателя (вход), движущегося в радиальном направлении относительно приемной антенны (выход) (рис. 5.2,а). Согласно первой формуле (5.9) излучаемое колебание частоты f воспринимается приемником как колебание другой частоты $f+F$ (где $F=-\dot{R}/\lambda$ — доплеровский сдвиг). Это еще более ярко проявляется при тангенциальном движении источника колебаний частоты f относительно центра раскрыва приемной антенны, которую для определенности представим в виде пятиэлементной эквидистантной антенной решетки с шагом b (рис. 5.2,б). Согласно второй формуле (5.9) центральный элемент решетки ($x=0$) принимает колебание частоты f , первый справа ($x=b$) и первый слева ($x=-b$) — колебания частот $f+F_{x1}$ и $f-F_{x1}$ соответственно, где $F_{x1}=-\dot{u}_x f b/c$, а второй справа ($x=2b$) и второй слева ($x=-2b$) — колебания частот $f+F_{x2}$ и $f-F_{x2}$, где $F_{x2}=-2\dot{u}_x f b/c$. Таким образом, вместо единственной излучаемой гармоники при-

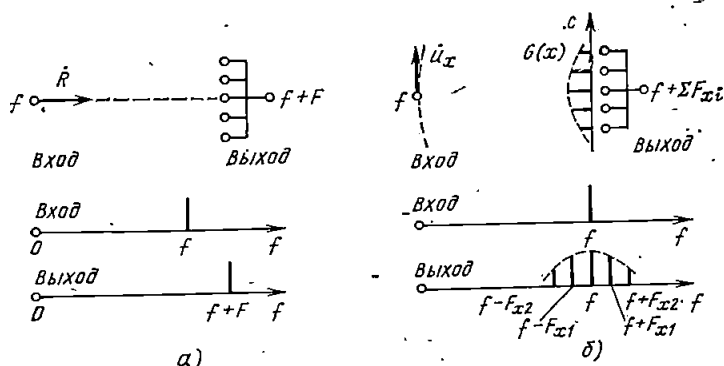


Рис. 5.2. Спектр сигнала на входе и на выходе пространственно-временной структуры при радиальном (а) и тангенциальном (б) движении источника относительно центра раскрыва антенны

нимается спектр гармонических колебаний, амплитуды которых пропорциональны функции раскрыва антенны $G(x)$ в соответствующих точках. Этот спектр тем шире, чем больше угловая скорость $\dot{\alpha}_x$. Суммируясь на выходе решетки, эти гармоники образуют сигнал, огибающая которого повторяет диаграмму направленности антенны, т. е. вместо излучаемого непрерывного гармонического колебания принимается модулированный по амплитуде сигнал. Тем самым реализуется один из известных способов снятия диаграммы направленности антенны — путем ее облета.

В линейных пространственно-временных структурах в отличие от линейных электрических цепей (чисто временные структуры) возможно появление на выходе гармонических составляющих, отсутствующих на входе. Поэтому появление новых гармоник не может служить признаком нелинейности пространственно-временной структуры, а только лишь неприменимость принципа суперпозиции.

Применение аппарата пространственных частот позволило построить единую теорию радиосистем, охватывающую все элементы, включая антенну. Особенно успешно этот аппарат используется в устройствах обработки принимаемых сигналов, где временные функции записываются на пространственном носителе (магнитной и фотопленке, экране индикатора и т. п.) и, наоборот, пространственные функции с помощью развертки переводятся во временные. При этом пространственный и временный (двусторонний) спектры отличаются только масштабом шкалы частот.

5.2. ПРОСТРАНСТВЕННО-ВРЕМЕННАЯ ЭКВИВАЛЕНТНОСТЬ

В определении угловых пространственных частот $\nu_x = f x / c$ и $\nu_y = f y / c$ в качестве равноценных сомножителей участвуют частота колебаний f и координаты раскрыва x или y . Это является осно-

вой пространственно-частотной эквивалентности при формировании функции раскрыва антенных решеток, являющейся пространственной характеристикой радиосистемы. Наоборот, пространственная структура в определенных случаях приобретает свойства частотно-избирательной временной системы. Ниже рассмотрим оба случая.

Дискретной функции раскрыва в виде дельта-функций координаты u_x

$$G(u_x) = \delta\left(u_x - \frac{b}{2\lambda}\right) + \delta\left(u_x - \frac{3b}{2\lambda}\right) + \delta\left(u_x + \frac{b}{2\lambda}\right) + \delta\left(u_x + \frac{3b}{2\lambda}\right) \quad (5.10)$$

соответствует диаграмма направленности

$$\begin{aligned} G(u_x) &= 2 \left(\frac{\cos \pi b u_x}{\lambda} + \frac{\cos 3\pi b u_x}{\lambda} \right) = \\ &= 2 \left(\frac{\cos \pi b f_0 u_x}{c} + \frac{\cos 3\pi b f_0 u_x}{c} \right). \end{aligned} \quad (5.11)$$

Коэффициент 3 во втором слагаемом можно в равной степени отнести и к шагу решетки b , и к частоте f_0 . Поэтому функцию раскрыва (5.10) можно реализовать либо с помощью четырехэлементной эквидистантной решетки на единственной гармонике f_0 , либо с помощью двухэлементной решетки на двух гармониках частоты f_0 и $3f_0$ (рис. 5.3, а). Таким образом, выбором спектра сигнала можно на простейшей антенной решетке получить сложную функ-

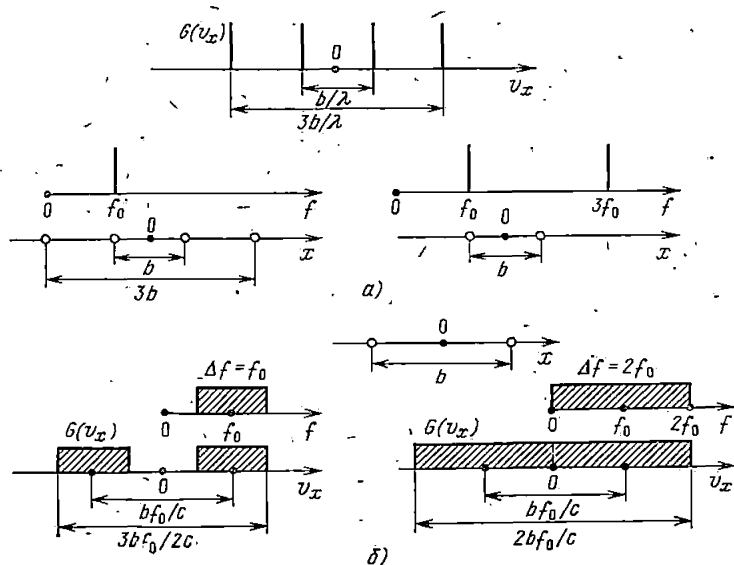


Рис. 5.3. Пространственно-временная эквивалентность в антенных решетках при использовании сигналов с дискретным (а) и сплошным (б) спектрами

цию раскрыва. Действительно, возьмем простейшую решетку из двух изотропных излучателей с шагом b между ними и будем использовать широкополосный сигнал с несущей f_0 и равномерным спектром в полосе частот $\Delta f = f_0$ и $\Delta f = 2f_0$ (рис. 5.3,б). В первом случае такая структура эквивалентна двум антеннам, разнесенным на интервал b , со сплошным раскрывом $d = b/2$ у каждой, причем функция раскрыва каждой антенны повторяет форму спектра сигнала. Случай $\Delta f = 2f_0$ является предельно возможным, когда спектр сигнала перекрывает все частоты от 0 до $2f_0$. При этом рассматриваемая двухэлементная решетка эквивалентна антенне со сплошным раскрывом $d = 2b$, равном удвоенному шагу решетки.

Однако эквивалентность пространственных и временных характеристик в приведенных примерах относительна, за исключением случая чисто синусоидального сигнала, когда изменение частоты сигнала и линейного разноса вибраторов совершенно одинаково влияет на функцию раскрыва и, следовательно, диаграмму направленности. При любом другом сигнале функция раскрыва зависит от времени, т. е. в пространственной структуре поля антенны происходят нестационарные переходные процессы. Так, две синусоиды кратных частот (например, f_0 и $2f_0$) образуют биения, в такт с изменением амплитуды которых изменяется во времени и диаграмма направленности двухэлементной решетки, которая эквивалентна диаграмме направленности четырехэлементной решетки на частоте f_0 только в те моменты времени, когда колебания обеих частот оказываются синфазными. При сигнале со сплошным спектром шириной $\Delta f = 2$ эквивалентность наступает в единственный момент времени, когда все составляющие спектра синфазны.

Таким образом, при очень широкополосных сигналах пространственные и временные характеристики системы жестко связаны, их уже нельзя рассматривать независимо. Можно говорить лишь о формировании диаграммы направленности, относящейся к заданному моменту времени (например, к моменту приема радиоимпульса). При использовании же узкополосных сигналов можно считать, что все гармоники в раскрыве по частоте практически не отличаются от несущей, и рассчитывать диаграмму направленности только на этой частоте.

Рассмотрим теперь антенну с продольной фокусировкой (антенна бегущей волны), протяженность которой d вдоль оси распространения волны z значительна, а сама антенна настроена на частоту f_0 . При настройке принимаемый (и излучаемый) вдоль оси сигнал максимален, так как принимаемые (или излучаемые) отдельными участками раскрыва элементарные сигналы суммируются синфазно. При расстройке $f - f_0$ набег фазы элементарных сигналов $2\pi(f - f_0)z/c$ линейно зависит от продольной координаты z элементарного участка dz , отсчитываемой от середины раскрыва. Полагая функцию раскрыва на интервале d постоянной, а напряженность поля принимаемой волны единичной, находим амплитуду

литуду сигнала, воспринимаемого всей антенной с направления вдоль оси z , как функцию расстройки:

$$U(f-f_0) = \int_{-d/2}^{d/2} \exp \left[\frac{j 2\pi (f-f_0) z}{c} \right] dz = d \frac{\sin [\pi (f-f_0) d/c]}{\pi (f-f_0) d/c}. \quad (5.12)$$

Если раскрыв линейный, то снижение амплитуды при изменении частоты относительно f_0 объясняется тем, что диаграмма направленности отклоняется от продольной оси z и эффективно принимает сигналы с других направлений (частотное сканирование). Если же антенна объемная, т. е. имеет значительный поперечный раскрыв, прием (и излучение) с других направлений не происходит (кроме боковых лепестков). Таким образом, объемная антенна с продольной фокусировкой эквивалентна временному фильтру, настроенному на частоту f_0 , с амплитудно-частотной характеристикой вида $\sin x/x$ и полосой пропускания $\Delta f = c/d$, отсчитываемой на уровне 0,64, близком к уровню половинной мощности. Относительная полоса пропускания такой пространственной структуры $\Delta f/f_0 = \lambda/d$ определяется теми же параметрами системы, что и ширина диаграммы направленности.

Приведенные примеры являются наглядным доказательством того, что временные функции (сигналы) при определенных условиях могут определять пространственные характеристики системы (функцию раскрыва), а чисто пространственные структуры выполнять роль временных фильтров (в антенной технике такие структуры называют узкодиапазонными).

Ранее было показано, что продольная пространственная структура волны повторяет временную структуру колебания и (у гармонических и узкополосных сигналов) в процессе распространения волны не изменяется. Совершенно иначе обстоит дело с поперечной структурой волны. Чтобы показать это, рассмотрим сначала фазовую структуру поля точечного излучателя (ИЗЛ) в поперечной плоскости (рис. 5.4). Расстояние от облучателя до произвольной точки x этой плоскости ($x \ll R$) $r = \sqrt{R^2 + x^2} \simeq R + x^2/2R$, поэтому закон изменения фазы в зависимости от поперечной координаты x с точностью до постоянной φ_0 квадратичный и зависит от расстояния R :

$$\varphi(x) = \varphi_0 - 2\pi r/\lambda \approx -\pi x^2/\lambda R. \quad (5.13)$$

Это означает, что поперечная фазовая структура поля (и в меньшей степени амплитудная) в процессе распространения волны претерпевает непрерывные изменения: по мере удаления от источника фронт волны становится все более плоским. Поэтому можно говорить о поперечной функции передачи свободного пространства, фазовая характеристика которой описывается соотношением (5.13). Знак минус в (5.13) означает, что волна расходящаяся, т. е. с увеличением $|x|$ фаза испытывает отставание. У сходящейся волны, которую можно получить с помощью собирающей линзы или параболического зеркала, фаза с увеличением $|x|$ испытывает опережение, и знак меняется на противоположный.

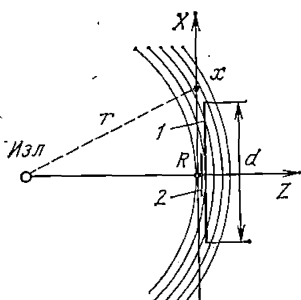


Рис. 5.4. Поперечная фазовая структура поля точечного излучателя

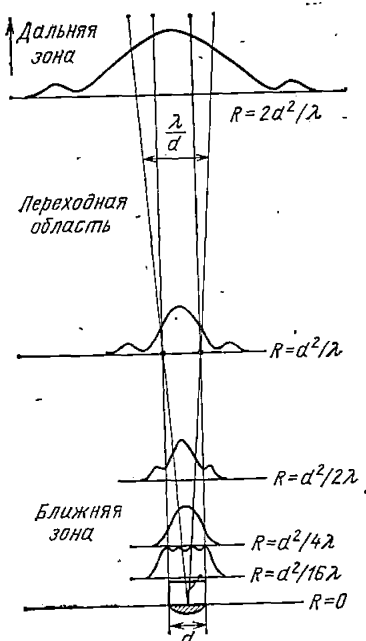


Рис. 5.5. Изменение поперечной структуры поля (диаграммы направленности) антенны в зависимости от расстояния

Сигналы, принимаемые антенной 1 (с большим раскрытием d), для которой излучатель находится в ближней зоне ($R < d^2/\lambda$), несинфазные. Для получения максимального выходного сигнала необходимо скорректировать фазу сигналов по раскрытию перед суммированием. Эта операция называется фокусировкой (в фотоаппаратах она достигается настройкой на резкость). Сигналы, принимаемые антенной 2 (с малым раскрытием), для которой излучатель находится в дальней зоне ($R > d^2/\lambda$), не требуют фокусировки, так как волну в пределах ее раскрытия можно считать плоской, а сигналы синфазными (фотоаппарат устанавливают на бесконечность независимо от дальности).

Проследим теперь качественное изменение поперечной структуры поля, создаваемого протяженным излучателем — раскрытием антенны (рис. 5.5). Она определяется как результат суммирования волн, возбуждаемых каждой точкой раскрытия, при котором фазовые соотношения между слагаемыми переходят в амплитудную структуру суммарной волны. Амплитудная структура плавно изменяется с расстоянием, переходя от равномерного синфазного поля в раскрытии антенны к диаграмме направленности в дальней зоне. При этом в пределах ближней зоны (до границы $R = d^2/\lambda$) резко изменяется вид функции, но практически не изменяется ее ширина d , а после этой границы вид функции изменяется очень мало, но ее линейная протяженность растет пропорционально расстоянию R при постоянном угле раствора $\theta = \lambda/d$. Начиная с расстояния $R = 2d^2/\lambda$ (в дальней зоне) диаграмма направ-

ленности в функции угловой координаты практически не изменяется, но ее линейное сечение непрерывно возрастает. Итак, при чисто гармонических или узкополосных сигналах ($\Delta f \ll f_0$) поперечная структура поля по мере удаления от антенны непрерывно изменяется, а продольная сохраняется неизменной.

У сверхширокополосных сигналов, ширина спектра которых соизмерима с несущей ($\Delta f \approx f_0$), структура поля существенно отличается от рассмотренной, хотя для каждой гармоники все закономерности полностью сохраняются. Такие сигналы (при отсутствии внутриимпульсной модуляции) содержат всего от одного до нескольких периодов несущей и занимают широкий спектр радиочастот — от самых низких до очень высоких, перекрывая весь радиодиапазон. Некоторые особенности использования таких сигналов уже были указаны при описании пространственно-временной эквивалентности, а именно сильная зависимость между пространственными и временными характеристиками системы (зависимость диаграммы направленности от времени). Теперь рассмотрим дополнительные особенности поля таких сигналов.

1. При использовании столь коротких сигналов даже небольшой сдвиг во времени между сигналами, принимаемыми или излучаемыми в данном направлении различными точками раскрыва, вызывает существенное изменение формы результирующего сигнала антенны (условие (4.1) не выполняется). Таким образом, каждому направлению в пределах диаграммы направленности передающей антенны соответствует не только определенная амплитуда, но и определенная форма сигнала. Это означает, что и продольная и поперечная структуры поля изменяются в зависимости от координат и времени, хотя спектральный состав сигнала остается неизменным. Такие процессы (в электродинамике их называют переходными) оказываются существенно нестационарными.

2. В процессе распространения волны от передатчика к приемнику изменяется и сам спектральный состав сигнала. Действительно, энергетика в радиолинии определяется коэффициентом направленного действия (КНД) передающей антенны $D = 4\pi A/\lambda^2$, существенно зависящим от длины волны, и эффективной площадью приемной антенны A , которая в первом приближении от длины волны не зависит. Таким образом, мощность принимаемого сигнала в радиолинии определяется произведением

$$D_{\text{прд}} A_{\text{прм}} = 4\pi A_{\text{прд}} A_{\text{прм}}/\lambda^2 = A_{\text{прд}} A_{\text{прм}} \omega^2/\pi c^2, \quad (5.14)$$

пропорциональным квадрату частоты. Следовательно, амплитуда принимаемого сигнала пропорциональна $|\omega|$. Но $|\omega|$ есть модуль функции передачи дифференцирующей цепочки. Поэтому происходит дифференцирование передаваемого сверхширокополосного сигнала. В частности, сигнал нулевой частоты вообще не может быть передан. В силу этих и других обстоятельств предъявляются особые требования к форме сверхширокополосного сигнала, который уже нельзя назвать ни высокочастотным, ни видеосигналом; его нельзя протектировать, понятие огибающей к нему

применимо лишь с особой оговоркой. Сверхширокополосные сигналы являются удобным инструментом для анализа тонкой структуры отражающих объектов и выполнения других сходных задач. Однако вопросы их применения, генерирования и приема выходят за пределы тематики настоящей книги. Заметим, что формула (5.14) в равной степени применима и к узкополосным сигналам, но их спектральные составляющие различаются по частоте столь незначительно, что эффект дифференцирования практически не проявляется.

5.3. ПРОСТРАНСТВЕННО-ВРЕМЕННЫЕ ПРЕОБРАЗОВАНИЯ И ОПТИЧЕСКАЯ ОБРАБОТКА СИГНАЛОВ

В процессе приема и обработки сигналов непрерывно происходит преобразование пространственных параметров во временные и наоборот как в силу естественных причин, так и преднамеренно. В частности, вследствие конечной скорости распространения радиоволн координата дальности преобразуется во время запаздывания принимаемого сигнала, а относительная радиальная скорость перемещения источника сигнала — в доплеровский сдвиг частоты. При сканировании антенного луча угловая координата объекта также преобразуется во временное положение принимаемого от него сигнала. При записи полученных временных последовательностей на пространственном носителе происходит обратное преобразование временных параметров сигнала в пространственные. Так, с помощью радиальной развертки записывающего луча $r = vt$ (где v — линейная скорость записи) на экране индикатора воспроизводится изображение O' точечного объекта O , время запаздывания сигнала которого $\tau = 2R/c$ пропорционально дальности. Если при этом направление радиальной развертки с помощью следящей системы, поворачивающей отклоняющую катушку вокруг электронно-лучевой трубки индикатора, повторяет угловое положение α оси сканирующей антенны, на экране индикатора создается изображение наблюдаемой пространственной картины в масштабе $M = r/R = 2v/c$ (рис. 5.6).

В автоматизированных устройствах обработки и хранения радиотехнической информации в качестве пространственных носителей сигнала широко используют мишень потенциалоскопа, магнитные ленты и барабаны, а также линии задержки.

Пространственно-временные преобразования сопровождаются переходом пространственных частот во временные и наоборот, ввиду чего временную фильтрацию сигналов можно заменить пространственной. На этом основано использование пространственных оптических процессоров — фильтров и корреляторов — для обработки сигналов радиотехнических систем.

В когерентных оптических процессорах в качестве переносчика информации используется широкий однородный по всей площа-

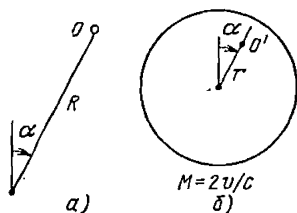


Рис. 5.6. Восстановление пространственной картины (а) на экране индикатора с радиально-круговой разверткой (б)

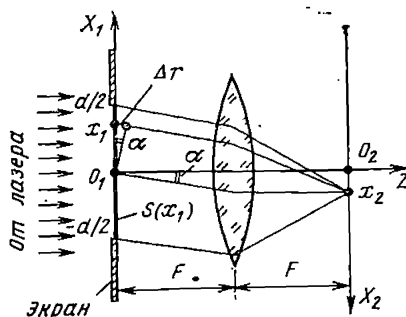


Рис. 5.7. Связь между сигналами в передней и задней фокальных плоскостях оптической линзы

ди входного окна коллимированный (параллельный) пучок когерентного (мономатического) света лазера (рис. 5.7). Во входном окне устанавливается транспарант с записью сигнала $S(x)$, благодаря чему проходящий через него поток модулируется вдоль координаты x по амплитуде и фазе записанным входным сигналом, в общем случае комплексным (способы и виды записи сигналов радиотехнических систем на транспарант кратко рассмотрены ниже). Фильтрация в оптическом процессоре заключается в изменении амплитудно-фазовой картины на выходе устройства в соответствии с заданным законом.

Оптические процессоры обладают рядом достоинств по сравнению с обычными, электронными:

1) вместо единственной независимой переменной электрического сигнала — времени — электромагнитное световое поле имеет дополнительно две независимые переменные — пространственные координаты в плоскости, перпендикулярной направлению распространения волны;

2) световая волна характеризуется большим числом параметров — носителей информации, чем электрический сигнал, что позволяет производить многоканальную обработку большого числа одномерных сигналов или многомерного сигнала (поля) сравнительно простыми средствами;

3) предельно просто выполняются операции умножения, интегрирования, преобразования Фурье и другие для любых функций, включая комплексные.

Действительно, однородный световой поток при прохождении через два транспаранта с записью сигналов $S_1(x)$ и $S_2(x)$ имеет на выходе последнего амплитудно-фазовое распределение, пропорциональное произведению $S_1(x) \cdot S_2(x)$. Так выполняется операция умножения функций. Ввиду того, что справа от транспаранта с записью сигнала $S(x_1)$ параллельный световой поток с помощью линзы с фокусным расстоянием F собирается в точке O_2 задней фокальной плоскости (рис. 5.7), напряженность поля в этой точ-

ке равна интегралу от функции $S(x_1)$. Так выполняется операция интегрирования функции.

Поскольку каждая точка на выходе транспаранта является источником колебаний, распространяющихся во всех направлениях, рассмотрим суммарный световой поток от всего транспаранта $S(x_1)$ в произвольном направлении $\alpha \ll \pi/2$. При этом воспользуемся правилом, что луч, исходящий из главного фокуса (точка O_1), после линзы идет параллельно ее оси, а все параллельные лучи на входе линзы сходятся в одной точке (x_2) задней фокальной плоскости. Если фазу сигнала, приходящего в точку x_2 из точки O_1 , принять равной нулю, то из-за разности хода $\Delta r = x_1 u_x = x_1 \sin \alpha \approx x_1 \alpha$ фаза сигнала, приходящего из произвольной точки x_1 транспаранта в точку x_2 , равна $-2\pi f_0 \Delta r / c \approx -2\pi f_0 x_1 \alpha / c$. Поэтому напряженность поля, наводимого в точке x_2 элементом dx_1 транспаранта, можно характеризовать комплексной огибающей $S(x_1) \exp(-j2\pi f_0 x_1 \alpha / c) dx_1$, а напряженность поля, создаваемого в этой точке всем транспарантом, — интегралом по всему раскры-
ву окна d

$$S(\alpha) = \int_{-d/2}^{d/2} S(x_1) e^{-j2\pi f_0 x_1 \alpha / c} dx_1. \quad (5.15)$$

Введем пространственную частоту

$$f_x = u_x / \lambda \approx \alpha / \lambda_{св} \approx x_2 / \lambda_{св} F, \quad (5.16)$$

где $\lambda_{св}$ — длина световой волны, положим $x = x_1$, а пределы интегрирования в формуле (5.15) продолжим до бесконечности, так как при $|x_1| > d/2$ $S(x_1) \equiv 0$. Тогда интеграл представляется как преобразование Фурье

$$S(f_x) = \int_{-\infty}^{\infty} S(x) e^{-j2\pi f_x x} dx, \quad (5.17)$$

которое является линейным пространственным спектром функции $S(x)$, записанной на транспаранте. Согласно соотношению (5.16) координата x_2 задней фокальной плоскости линзы с точностью до постоянного множителя представляет собой линейную пространственную частоту f_x .

Поскольку, как отмечалось, световой поток от транспаранта распространяется во всех направлениях α , соотношения (5.15) и (5.17) справедливы для всех точек задней фокальной плоскости. Следовательно, если в передней фокальной плоскости линзы разместить транспарант с записью какого-то сигнала и осветить его равномерным коллимированным пучком когерентного света, то в задней фокальной плоскости воспроизводится спектр этого сигнала — преобразование Фурье. Поэтому в данном случае переднюю фокальную плоскость линзы называют сигнальной, а заднюю — спектральной.

Необходимо отметить, что указанные выше преобразования являются приближенными. Они справедливы только для области

малых углов (параксиальной) относительно оси процессора ($\sin \alpha \approx \alpha \approx \operatorname{tg} \alpha$). При больших углах возникают искажения.

Для передачи амплитудно-фазовых соотношений сигналов радиотехнических систем, поступающих в оптический процессор, в транспарантах (масках) используют три вида записи, соответствующие трем видам модуляции проходящего через них светового потока: амплитудной, фазовой и комплексной. Амплитудная запись характеризуется степенью прозрачности транспаранта в данной точке (например, пропусканием фотопленки), фазовая — набегом фазы в слое диэлектрика, из которого выполнен транспарант (толщиной слоя), комплексная охватывает одновременно и то и другое.

Равномерной прозрачности транспаранта в пределах окна шириной d соответствует нулевая пространственная гармоника — яркое дифракционное пятно нулевого порядка в главном фокусе задней фокальной плоскости линзы (рис. 5.8, а). Благодаря прохождению через окно синфазного светового потока спектр имеет вид $\sin x/x$ шириной $F\lambda_{\text{св}}/d$, что соответствует диаграмме направленности антенны оптического диапазона (окна).

Гипотетический транспарант с амплитудной записью сигнала в виде отрезка синусоиды с пространственным периодом λ , соответствующим длине волны несущей радиоимпульса (рис. 5.8, б), дает дифракционные пятна в точках f_{x0} и $-f_{x0}$. В этих точках разность хода лучей от двух смежных полупериодов синусоиды одного знака равна длине световой волны $\lambda_{\text{св}}$, а разных знаков — ее половине, в результате именно в этих двух точках происходит суммирование колебаний, создаваемых входным сигналом (дифракция порядка $+1$ и -1). Частоты f_{x0} и $-f_{x0}$ называют пространственными несущими. В главном фокусе (область нулевых пространственных частот) засвет отсутствует, так как сигналы положительных и отрицательных полупериодов синусоиды гасят друг друга. Ширина спектральных линий обратна длительности высокочастотного импульса, а вид функции повторяет форму спектра его огибающей.

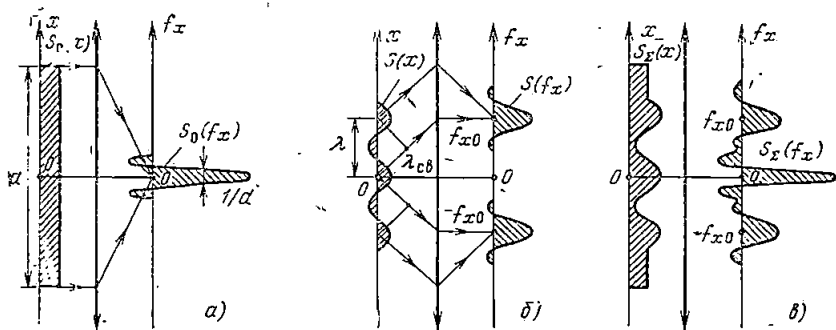


Рис. 5.8. Пространственный вещественный сигнал и его спектр

На самом деле синусоиду на транспаранте непосредственно записать нельзя, так как прозрачность не имеет знака. Чтобы обойти эту трудность, вводят средний уровень прозрачности, равный половине диапазона прозрачности — от полного пропускания до полного подавления. Тогда отрицательным полупериодам синусоиды соответствует уровень прозрачности ниже среднего, а положительным — выше среднего. Таким образом, реальный амплитудный транспарант содержит ненужную постоянную составляющую (рис. 5.8,а), которую легко устранить в выходном сигнале оптического процессора, поставив непрозрачный экран (режекторный фильтр) на пути светового потока в области нулевых частот спектральной плоскости оптической системы. Следует отметить, что столь же просто можно подавить (или, наоборот, выделить) спектральные составляющие сигнала в любой области спектра. Отсюда вытекает, что в общем случае для фильтрации сигнала по любому наперед заданному закону достаточно поставить в спектральной плоскости оптического процессора транспарант с записью соответствующей функции передачи $H(f_x)$, в общем случае комплексной.

Таким образом, пространственный спектр входного сигнала содержит и положительные и отрицательные частоты, причем обе половины спектра совершенно одинаковы, и для фильтрации можно использовать только одну, подавив вторую (и нулевую гармонику) с помощью непрозрачного экрана. Для пояснения физического смысла знака линейной пространственной частоты f_x рассмотрим некоторый сигнал $S(x)$, размещенный в центре оптического окна и имеющий спектр $S(f_x)$. Тогда согласно теореме смещения спектр смещенного на x_0 сигнала $S(x-x_0)$ имеет вид $S(f_x) \exp(-j2\pi f_x x_0)$, где показатель экспоненты определяет набег фазы в различных точках частотной плоскости оптической системы в функции смещения. По одну сторону оптической оси системы смещение x_0 вызывает опережение, по другую — отставание по фазе, что и определяется знаком пространственной частоты. Если же смещение сделать линейной функцией времени $x_0 = vt$, т. е. транспарант с записью сигнала протягивать с постоянной скоростью v , то показатель экспоненты будет функцией времени $-2\pi f_x vt = -2\pi F_x t$, где $F_x = -f_x v$ — доплеровский сдвиг реальной, временной частоты сигнала в данной точке f_x спектральной плоскости. Сигнал $S(x-vt)$ по отношению к одним точкам этой плоскости приближается, по отношению к другим удаляется, что и определяет знак доплеровской и пространственной частот сигнала. Значит, при движении транспаранта частота световых колебаний в частотной плоскости оптического процессора отличается от частоты облучения на доплеровский сдвиг $F_x = -vf_x$, пропорциональный скорости протяжки транспаранта и значению пространственной частоты в каждой точке спектральной плоскости. Таким образом, вместо монохроматического луча на входе процессора получаем многочастотное световое колебание.

При фазовой модуляции светового потока используется про-

зрачный транспарант с фазовой записью сигнала $S(x)$. Так как индекс фазовой модуляции ψ_m обычно очень мал, комплексный коэффициент пропускания (передачи) такого транспаранта

$$T(x) = \exp[j\psi_m S(x)] \approx 1 + j\psi_m S(x) \quad (5.18)$$

содержит постоянную составляющую, соответствующую полной прозрачности транспаранта, и полезную составляющую, сдвинутую относительно постоянной по фазе на $\pi/2$ ($j = \exp(j\pi/2)$). Запись (5.18) в спектральной плоскости процессора дает ее пространственный спектр

$$T(f_x) = \delta(f_x) + j\psi_m S(f_x), \quad (5.19)$$

где $\delta(f_x)$ — дельта-функция нулевой частоты, соответствующая постоянной составляющей светового потока (фактически в области нулевой частоты спектр имеет конечную ширину ввиду ограниченной ширины окна оптического процессора). В отличие от спектра амплитудного транспаранта сигнальная составляющая спектра (5.18) сдвинута по фазе относительно нулевой гармоники на $\pi/2$. Если в области нулевой частоты спектральной плоскости оптического процессора поставить четвертьволновую пластину, коэффициент пропускания которой равен $j = \exp(j\pi/2)$, то справа от спектральной плоскости световой поток будет иметь пространственный спектр

$$T_v(f_x) \simeq j[\delta(f_x) + \psi_m S(f_x)], \quad (5.20)$$

который отличается от спектра сигнала с амплитудной записью только несущественным множителем j , означающим постоянный сдвиг начальной фазы светового потока $\pi/2$ по всей спектральной плоскости. Тем самым фазовая модуляция становится равноценной амплитудной и невидимая на фазовом транспаранте запись на выходе процессора становится видимой. Если к тому же размещенная в области нулевых частот четвертьволновая пластина вносит затухание в K раз, во столько же раз возрастает контрастность изображения на выходе. При большем индексе фазовой модуляции в отличие от амплитудной имеются спектральные составляющие на гармониках пространственной несущей $2f_{x0}$, $3f_{x0}$ и т. д., которые легко подавляются с помощью непрозрачного экрана. Таким образом, не очень глубокую фазовую модуляцию светового потока можно свести к амплитудной, которая не изменяет спектрального состава модулирующей функции, т. е. записанного на транспаранте сигнала.

При комплексной записи на транспаранте вместо самого высокочастотного сигнала фиксируется его комплексная огибающая. Ввиду отсутствия пространственной несущей спектр комплексной огибающей сигнала группировается в центре спектральной плоскости, около нулевой частоты, и имеет такой же вид, как и спектр сигнала в области каждой из двух пространственных несущих при амплитудной записи. Отсюда вытекают преимущества комплексной записи по сравнению с амплитудной или фазовой записью самого высокочастотного сигнала.

1. В частотной плоскости отсутствует паразитный сигнал, соответствующий постоянной составляющей функции передачи транспаранта, который по интенсивности (яркости) во много раз превышает полезный и мешает наблюдению (выделению) последнего.

2. Эффективнее (примерно втрое) используется частотная плоскость: при записи высокочастотного сигнала на ней воспроизводятся два одинаковых спектра записанного сигнала и дополнительно спектр нулевых частот (а необходим только один спектр сигнала), при записи комплексной огибающей — только единственный спектр сигнала, причем не в стороне, а вдоль оси оптического процессора, где оптические искажения минимальны.

3. Снижаются требования к разрешающей способности или размерам сигнального транспаранта: для записи высокочастотного сигнала необходимо отдельно воспроизвести каждый полупериод его высокочастотного заполнения, при комплексной записи достаточно воспроизвести лишь огибающую сигнала.

К сожалению, комплексную запись очень сложно реализовать. Для этого нужно, например, чтобы прозрачность фотоэмульсии передавала амплитудные соотношения, а толщина подложки фотопленки — фазовые. Поэтому комплексные транспаранты обычно используют не для записи сигналов, а только в качестве оптических фильтров-масок, обеспечивающих заданную функцию передачи оптических фильтров.

В транспарантах-модуляторах светового потока используют оптические носители, статические и динамические.

Наиболее распространенным типом статического носителя является фотопленка. В режиме записи фотопленку протягивают устройством 1 с некоторой постоянной скоростью v относительно электронно-оптического преобразователя 2 (электронно-лучевые трубки, неоновые лампочки, лазеры и т. п.), к которому подводится записываемый сигнал (рис. 5.9, а). Луч преобразователя с помощью оптической системы 3 фокусируется на фотопленку. Записанный сигнал становится функцией координаты $x=vt$, являю-

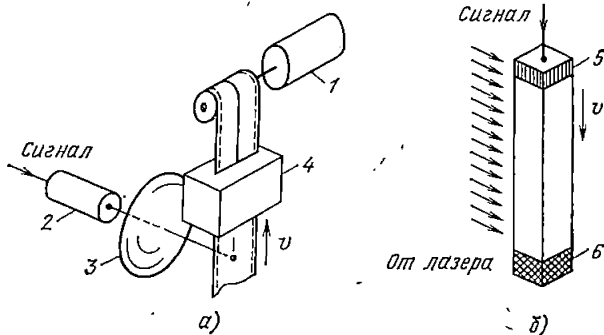


Рис. 5.9. Схемы модуляторов света электрическим сигналом на статических (а) и динамических (б) носителях

щейся аналогом времени. После проявления и закрепления пленки в устройстве 4 можно производить обработку записанного сигнала, протягивая пленку мимо окна, освещенного широким коллимированным пучком света лазера. Если скорость протяжки при обработке сигнала равна скорости протяжки пленки при записи, то выходной сигнал оптического фильтра выдается в том же темпе, что и входной, но с задержкой во времени.

К достоинствам фотопленки следует отнести простоту записи сигнала на ней, возможность использования как комплексной записи, так и записи на высокой (промежуточной) частоте, линейной и нелинейной записи, выбора любой скорости записи и воспроизведения и, что особенно важно, неограниченные емкость памяти и время хранения информации. Однако очень серьезным недостатком фотопленки является разрыв во времени между приемом сигнала и его обработкой, что во многих случаях совершенно недопустимо. Недостатком является также большой расход пленки, часто двойной — для записи как входного, так и обработанного сигналов, а также непригодность ее для повторного использования (необратимость). Амплитудная запись связана с большими потерями света при пропускании его через пленку. Однако этот недостаток можно устранить, отбелив пленку фотохимическим способом — в соответствии с прозрачностью эмульсии изменяют толщину подложки, а эмульсию обесцвечивают. Линейная, амплитудная запись становится нелинейной, фазовой. К сожалению, отбеливание усложняет и затягивает процесс фотохимической обработки пленки.

Кроме фотопленки используют также термопластические и фотохромные материалы, в которых большинство приведенных недостатков отсутствует: их можно многократно использовать, производить обработку в реальном масштабе времени (без задержки) и без потерь света. У нагретой и помещенной в вакуум термопластической пленки запиывающий электронный луч создает потенциальный рельеф, сжимающий ее пропорционально местному заряду (фазовая модуляция). Под воздействием сигналов фотохромные материалы изменяют свои отражающие свойства. Возможно также использование светомодуляторов на полупроводниковых кристаллах. Однако необходимость работы в вакууме затрудняет применение термопластиков, а другие материалы еще находятся в стадии освоения.

Динамическим модулятором света является ультразвуковая линия задержки в виде прозрачного бруска диэлектрика, просвечиваемого сбоку коллимированным пучком когерентного света (рис. 5.9,б). На входе линии ставят пьезопреобразователь 5, преобразующий входной сигнал в ультразвуковую волну, распространяющуюся вдоль линии. Изменения плотности среды — сжатие и разрежение — изменяют коэффициент преломления среды. Это приводит к фазовой модуляции проходящего сквозь линию пучка света входным сигналом в реальном масштабе времени, без задержки. Для создания режима бегущей волны в конце линии ставят

поглощающую нагрузку 6. Ультразвуковая линия задержки в принципе является идеальным светомодулятором и в отличие от статических пленочных носителей не нуждается в механическом, лентопротяжном механизме. К сожалению, скорость распространения ультразвука и затухание в линии слишком велики, поэтому в настоящее время в видимом свете воспроизводятся лишь сигналы длительностью не больше нескольких десятков микросекунд.

5.4. КОГЕРЕНТНЫЕ ОПТИЧЕСКИЕ ПРОЦЕССОРЫ

Отмеченные выше свойства когерентной оптики позволяют строить оптические процессоры — фильтры и корреляторы. Фильтрацию можно производить как в сигнальной, так и в спектральной плоскости. Рассмотрим сначала одномерный случай.

В оптическом процессоре с фильтрацией в сигнальной плоскости во входном окне передней фокальной плоскости линзы размещаются вплотную транспарант с записью сигнала $S(x)$ и маска с записью импульсной характеристики фильтра $H(x)$ (рис. 5.10, а). Входное окно облучается однородным коллимированным пучком света лазера (1) с расширяющей (2) и коллимирующей (3) линзами. Фильтр-маска $H(x)$ закреплен неподвижно, а транспарант с записью сигнала $S(x)$ протягивается вдоль оси x со скоростью v , воспроизводя в момент t функцию $S(x-vt)$. Световой поток модулируется произведением этих двух функций, в результате на фотоприемник (4), расположенный в задней фокальной плоскости линзы ($f_x=0$), воздействует поток напряженностью

$$S_{\text{вых}}(t) = \int_{-\infty}^{\infty} S(x-vt) H(x) dx. \quad (5.21)$$

Этот интеграл является сверткой входного сигнала и импульсной характеристики фильтра, и поэтому представляет собой выходной сигнал фильтра с заданной характеристикой. В частности, характеристику можно сделать согласованной с сигналом, в результате на выходе получим сигнал, прошедший через согласованный

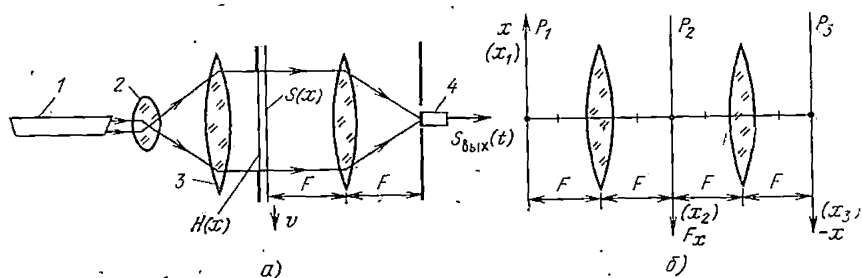


Рис. 5.10. Схемы одномерных оптических процессоров с фильтрацией в сигнальной (а) и спектральной (б) областях

фильтр, т. е. функцию корреляции. Выходной сигнал выдается последовательно, по мере поступления входного.

Для описания процесса фильтрации в спектральной области рассмотрим систему из двух линз, фокальные плоскости которых совмещены в плоскости P_2 (рис. 5.10,б). Если в плоскости P_1 разместить сигнальный транспарант $S(x)$, освещенный пучком когерентного света, то в плоскости P_2 согласно ранее доказанному получим спектр сигнала $S(f_x)$. Поскольку во второй линзе происходят такие же процессы, как и в первой, то сигнал в плоскости P_3

$$S(x_3) = \int_{-\infty}^{\infty} S(f_x) e^{-j2\pi f_x x_3} df_x.$$

Если бы не знак минус в показателе экспоненты, мы получили бы вновь исходный сигнал как обратное преобразование Фурье. Знак минус можно исключить, если повернуть координату x_3 в противоположном направлении. Следовательно, в плоскости P_3 мы вновь получим входной сигнал, но перевернутый на 180° .

Поставим теперь в плоскости P_2 фильтр-маску с функцией передачи $H(f_x)$. Тогда справа от нее световой поток описывается произведением спектра выходного сигнала фильтра и функции передачи: $S_{\text{вых}}(f_x) = S(f_x)H(f_x)$. При этом в плоскости P_3 воспроизводится (при $x = -x_3$) сам выходной сигнал этого фильтра как обратное преобразование Фурье:

$$S_{\text{вых}}(t) = \int_{-\infty}^{\infty} S_{\text{вых}}(f_x) e^{j2\pi f_x x} df_x. \quad (5.22)$$

Если взять $H(f_x) = S^*(f_x)$, то фильтр становится согласованным (оптимальным) и на его выходе получим функцию корреляции сигнала. Эта функция в отличие от предыдущего процессора воспроизводится в плоскости P_3 целиком, а не последовательно. Однако само устройство менее компактно, а для автоматического съема полученных результатов нужно специальное устройство, которое существенно сложнее фотоприемника в предыдущем процессоре.

Рассмотрим теперь вторую координату фокальных плоскостей — координату y . Ее можно использовать для создания процессоров в виде набора одномерных фильтров, параллельных оси x , или одного двумерного фильтра (рис. 5.11).

В наборе одномерных фильтров используют цилиндрическую линзу (1), входные сигналы (2) и характеристики фильтров (3) записывают параллельными строками, разнесенными по оси y (рис. 5.11,а). Каждый сигнал можно обрабатывать по своему закону. В зависимости от разрешающей способности и размеров транспаранта число одновременно обрабатываемых сигналов может быть очень большим. Например, если в качестве транспаранта используем обычную 36-мм фотопленку с разрешающей способностью 50 линий/мм, то на ней можно записать и обработать по самым различным законам до 1800 сигналов одновременно!

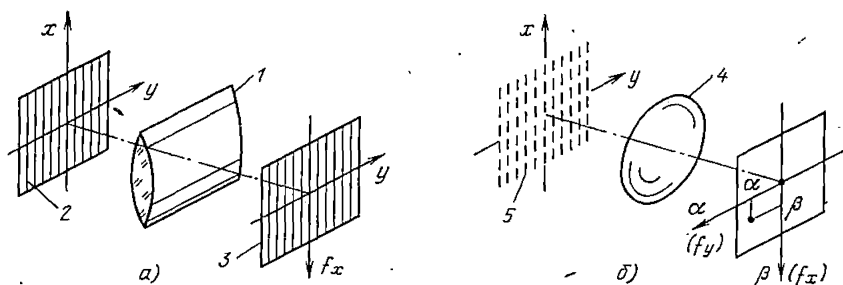


Рис. 5.11. Схемы двухкоординатных оптических процессоров в виде набора одномерных фильтров (а) и в виде двухмерного фильтра для сигналов антенной решетки (б)

Реально сигналы записывают с меньшей плотностью, но все же их число остается большим.

В двумерных процессорах используют сферические линзы (такие же, как в фотоаппаратах), которые выполняют двумерное преобразование Фурье

$$S(f_x, f_y) = \iint_{-\infty}^{\infty} S(x, y) e^{-j 2\pi (f_x x + f_y y)} dx dy. \quad (5.23)$$

Двумерные процессоры можно использовать для многоканальной обработки сигналов антенной решетки, для распознавания образов и т. п.

Пусть в формуле (5.23) функция $S(x, y)$ представляет собой поле, наводимое отражающим или излучающим объектом в раскрыве антенной решетки радиосистемы. С помощью специальных светомодуляторов сигналы, принимаемые отдельными элементами решетки, преобразуются в световые колебания с сохранением амплитудных и фазовых соотношений. Таким образом, в передней фокальной плоскости линзы (4 на рис. 5.11, б) воспроизводится уменьшенная модель раскрыва решетки (5). Так как $f_x \approx \alpha/\lambda$, $f_y \approx \beta/\lambda$ (где α и β — азимут и угол места объекта), то в выходной фокальной плоскости линзы воспроизводится пятно, положение которого указывает угловое положение объекта в пространстве, а форма и размеры пятна повторяют диаграмму направленности плоской антенной решетки. Если в пространстве наблюдается несколько объектов, то все они одновременно воспроизводятся в виде отметок, указывающих их угловое положение. Таким образом, измерительную систему вместе с антенной решеткой и рассмотренным оптическим процессором можно уподобить линзе фотоаппарата, которая полностью воспроизводит обстановку в поле зрения объектива. К сожалению, пригодные для воспроизведения сигналов антенные решетки в реальном масштабе времени — динамические светомодуляторы на прозрачных ультразвуковых линиях задержки — имеют большую продольную протяженность, что

затрудняет их использование в оптических аналогах двумерных решеток.

Двумерные процессоры можно использовать также для распознавания образов с определением их местоположения на изучаемой картине. Для этого применяют двумерный процессор с фильтрацией в спектральной плоскости (плоскость P_2 на рис. 5.10,б), где ставится фильтр-маска с записью комплексно-сопряженной функции спектра распознаваемого образа. Согласно доказанному ранее в выходной плоскости P_3 воспроизводится функция корреляции образа, записанного на входном транспаранте (в плоскости P_1). Положение пика функции корреляции (дифракционного пятна) соответствует положению распознаваемого образа на входном транспаранте. Другие образы с данным фильтром-маской не дадут пика корреляции. Таким образом, если на входном транспаранте среди множества образов имеется интересующий нас (например, буква на книжной странице), то он будет обнаружен, причем положение светящегося пятна указывает его местоположение.

Подобные системы можно использовать для вторичной (логической) обработки радиолокационных данных. К сожалению, задача распознавания даже плоских образов затруднена из-за критичности процессора к масштабу изображения и ориентации образа.

Оптические процессоры позволяют производить некоторые операции чисто оптическими средствами (призмами, зеркалами, линзами) без применения специальных фильтров-масок. И наоборот, включение транспарантов в оптическую схему может быть эквивалентно использованию оптических линз и призм. Это положение в равной мере относится и к временным функциям (сигналам), записанным на транспаранте. Так, поворот светового потока с помощью призмы на некоторый угол α_0 эквивалентен переносу изображения на пространственную несущую $f_{x0} = \sin \alpha_0 / \lambda_{св} \approx \alpha_0 / \lambda_{св}$. Эту же операцию можно выполнить с помощью фильтра-маски с линейным изменением набега фазы $2\pi f_{x0}x$ вдоль координаты x , что соответствует переносу на несущую $f_0 = v f_{x0}$ временного видеосигнала, запечатленного с масштабом $t = x/v$ на световом потоке, проходящем через эту маску.

Наиболее показательным примером общности свойств оптических структур, транспарантов и электрических цепей является эквивалентность фокусировки светового потока линзой и сжатия ЛЧМ сигнала в фильтре (рис. 5.12).

Фокусирующие свойства линз основаны на запаздывании фазы колебаний, проходящих через оптически плотную среду (коэффициент преломления $n > 1$). Рассмотрим плоско-выпуклую собирающую линзу с радиусом кривизны r (рис. 5.12,а). В зависимости от координаты x толщина такой линзы изменяется как $\Delta z = \Delta - \sqrt{r^2 - x^2} \approx \Delta - r + x^2/2r$, так как в рабочей, параксиальной области $x/r \ll 1$. Отсюда дополнительный набег фазы световых ко-

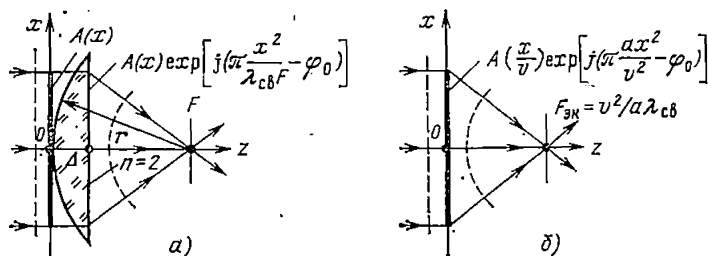


Рис. 5.12. Эквивалентность фокусирующих свойств линзы (а) и транспаранта с записью ЛЧМ сигнала (б) при облучении когерентным коллимированным пучком света

лебаний, проходящих через линзу в произвольном сечении x , составляет

$$\varphi(x) = 2\pi\Delta z(n-1)/\lambda_{св} = \pi x^2(n-1)/\lambda_{св}r - \varphi_0, \quad (5.24)$$

где $\varphi_0 = 2\pi(r-\Delta)(n-1)/\lambda_{св} = \text{const}$, причем $r \gg \Delta$. Если перед линзой дополнительно поставить маску с амплитудным коэффициентом прозрачности $A(x)$, то проходящий через них однородный коллимированный поток когерентного света приобретает амплитудно-фазовую структуру

$$\begin{aligned} S(x) &= A(x) \exp[j\pi x^2(n-1)/(\lambda_{св}r) - \varphi_0] = \\ &= A(x) \exp[j\pi x^2/(\lambda_{св}F) - \varphi_0], \end{aligned} \quad (5.25)$$

где $F = r/(n-1)$ — фокусное расстояние линзы (при $n=2$, как показано на рисунке, равно радиусу ее кривизны). Двояковыпуклая линза с различными радиусами кривизны эквивалентна системе из двух линз, для которой справедливо известное соотношение

$$\frac{1}{F} = \frac{1}{F_1} + \frac{1}{F_2} = (n-1) \left(\frac{1}{r_1} + \frac{1}{r_2} \right). \quad (5.26)$$

Соотношение (5.25) описывает волну, сходящуюся в дифракционное пятно в главном фокусе линзы F (знак плюс перед квадратичным показателем экспоненты). Как было показано в § 5.3, размер выходного пятна

$$d_{\text{вых}} = F\lambda_{св}/d \quad (5.27)$$

тем меньше, чем больше ширина окна прозрачности на входе, описываемого функцией $A(x)$. Происходит сжатие светового потока в задней фокальной плоскости линзы (хотя за ее пределами поток вновь расходится).

Возьмем теперь некоторый сигнал с глубокой линейной частотной модуляцией (ЛЧМ сигнал)

$$\begin{aligned} S(t) &= 2A(t) \cos(\pi a t^2 - \varphi_0) = A(t) \exp[j(\pi a t^2 - \varphi_0)] + \\ &+ A(t) \exp[-j(\pi a t^2 - \varphi_0)], \end{aligned} \quad (5.28)$$

где a — скорость изменения частоты сигнала $f(t) = at$, а $2A(t)$ — закон изменения его амплитуды. Если сигнал имеет длительность

$\tau_{\text{н}}$, то при $a \gg 1$ ширина спектра $\Delta f \approx a\tau_{\text{н}}$. Пусть этот сигнал записан на транспаранте вдоль координаты $x = vt$ (где v — скорость записи). Тогда функция прозрачности транспаранта примет вид

$$T(x) = 2A \frac{x}{v} \cos\left(\frac{\pi ax^2}{v^2} - \varphi_0\right) = A \frac{x}{v} \exp\left[j\left(\frac{\pi ax^2}{v^2} - \varphi_0\right)\right] + A \frac{x}{v} \exp\left[-j\left(\frac{\pi ax^2}{v^2} - \varphi_0\right)\right]. \quad (5.29)$$

При облучении транспаранта однородным коллимированным пучком света первое слагаемое в правой части соотношения (5.29) дает сходящуюся волну, а второе — расходящуюся. Следовательно, операция, описываемая экспоненциальным множителем сходящейся волны

$$A(x/v) \exp[j(\pi ax^2/v^2 - \varphi_0)], \quad (5.30)$$

эквивалентна действию собирающей линзы с фокусным расстоянием

$$F_{\text{эк}} = v^2/(a\lambda_{\text{св}}), \quad (5.31)$$

которое находится приравниванием коэффициентов при x^2 в формулах (5.25) и (5.29). Таким образом, световой поток (5.30) также фокусируется в световое пятно в некоторой точке $F_{\text{эк}}$, хотя собирающая линза отсутствует (рис. 5.12,б).

Размеры выходного пятна $d_{\text{вых}} = v\tau_{\text{вых}}$ определяются формулой (5.27) при заданных размерах окна прозрачности на входе $d = d_{\text{н}}$, если в эту формулу вместо F подставить $F_{\text{эк}}$ согласно соотношению (5.31). После указанных подстановок находим, что длительность выходного импульса процессора обратна длительности ЛЧМ сигнала $\tau_{\text{н}}$ на входе, т. е. ширине спектра сигнала $\Delta f \approx a\tau_{\text{н}}$:

$$\tau_{\text{вых}} = 1/a\tau_{\text{н}} \approx 1/\Delta f.$$

Таким образом, операции фокусировки и сжатия ЛЧМ сигнала тождественны.

Однако транспарант с функцией (5.29) создать невозможно, поскольку косинус принимает отрицательные значения, а прозрачность не может быть отрицательной. Если для передачи отрицательных всплесков ввести средний уровень прозрачности, то на выходе процессора будут одновременно три волны: сходящаяся и расходящаяся сигнальная и сходящаяся среднего уровня. Эти волны создадут взаимные помехи, причем последняя забьет остальные. Комплексную запись (5.30), как отмечалось, технически трудно реализовать.

Поэтому для записи используется высокочастотный сигнал

$$s(t) = 2A(t) \cos[2\pi f_{\text{н}}t + \pi at^2 - \varphi_0] = A(t) \exp[j(2\pi f_{\text{н}}t + \pi at^2 - \varphi_0)] + A(t) \exp[-j(2\pi f_{\text{н}}t + \pi at^2 - \varphi_0)] \quad (5.32)$$

с несущей частотой $f_{\text{н}}$ и текущей частотой $f(t) = f_{\text{н}} + at$. Эта запись

с учетом среднего уровня A_0 описывается функцией пропускания транспаранта

$$T(x) = A_0 + A(x) \exp[j(2\pi f_{x0}x + \pi x^2/\lambda_{св}F_{вк} - \varphi_0)] + \\ + A(x) \exp[-j(2\pi f_{x0}x + \pi x^2/\lambda_{св}F_{вк} - \varphi_0)], \quad (5.33)$$

где $f_{x0} = f_{\pi}/v$ — пространственная несущая; $x = vt$, а $F_{вк}$ определяется формулой (5.31). Тем самым все три волны разделяются по пространственным несущим: постоянная составляющая дает яркий засвет на оптической оси (нулевая несущая) в главном фокусе, сходящаяся волна фокусируется в фокальной же плоскости выше оптической оси (несущая $+f_{x0}$), а расходящаяся идет под углом к оптической оси вниз (несущая $-f_{x0}$). Это обстоятельство позволяет с помощью непрозрачного экрана, имеющего щель в области пятна от сходящейся сигнальной волны, подавить мешающие волны (пространственная фильтрация).

Наиболее ярко достоинства оптических процессоров рассмотренного типа проявляются при использовании их для оптимальной фильтрации ЛЧМ сигналов большой длительности в когерентных РЛС бокового обзора. Эти станции устанавливают на самолетах и предназначают для получения радиолокационного изображения местности по маршруту полета. Луч, узкий в горизонтальной плоскости и широкий (веерный) в вертикальной, формируется с помощью протяженной антенны, размещенной параллельно продольной оси самолета (линии пути) под фюзеляжем или в самом фюзеляже под прямым углом к этой оси.

Используют также два таких луча, благодаря которым одновременно осматривают две полосы местности — справа и слева от линии пути (полоса непосредственно под самолетом не просматривается ввиду крайне низкой разрешающей способности по горизонтальной дальности). Обзор вдоль пути достигается за счет последовательного облучения местности при движении самолета, а в поперечном направлении — по наклонной дальности — за счет естественного запаздывания импульсных сигналов (рис. 5.13).

В когерентных РЛС бокового обзора все принимаемые от любого объекта сигналы суммируются с учетом их фазовых соотношений. Результат когерентного суммирования сигналов, последовательно принимаемых за время облучения объекта, эквивалентен одновременному приему этих сигналов на антенну с большим раскрытием, равным поперечному сечению луча в районе расположения объекта. Такую антенну называют синтезированной, а сами радиолокационные станции получили название радиолокаторов с синтезированием апертуры (РСА). Очень высокая разрешающая способность в продольном направлении обеспечивается синтезированной апертурой больших размеров (порядка километра), а в поперечном

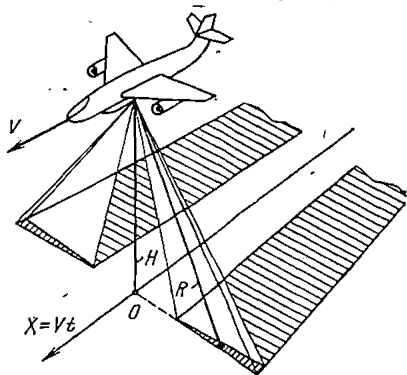


Рис. 5.13. Обзор местности с помощью РЛС бокового обзора

направлении — благодаря использованию коротких импульсов или сложных сигналов с широким спектром. В результате получают подробное изображение местности, сходное с аэрофотоснимком, но при этом отсутствуют перспективные искажения и экранировка облаками.

Чтобы описать структуру сигнала от точечного объекта, удаленного от линии пути на расстояние R_0 (траверзное расстояние), рассмотрим, как изменяется его текущая дальность $R(t) = \sqrt{R_0^2 + V^2 t^2}$ и текущий угол $\alpha(t) = \arctg(Vt/R_0)$, где V — продольная скорость самолета; Vt — путь, проходимый самолетом, если за начало отсчета принять момент пролета траверза объекта. Ширина диаграммы направленности $G(\alpha)$ в плоскости визирования объекта очень мала, поэтому можно принять

$$R(t) \approx R_0 + V^2 t^2 / 2R_0, \quad \alpha(t) \approx Vt/R_0. \quad (5.34)$$

Амплитуда принимаемого сигнала пропорциональна квадрату диаграммы направленности, так как сигнал дважды — при излучении и приеме — умножается на нее, а фаза сигнала

$$\varphi(t) = -2\pi \cdot 2R(t)/\lambda = \varphi_0 - 2\pi V^2 t^2 / \lambda R_0, \quad (5.35)$$

где введена начальная фаза $\varphi_0 = -4\pi R_0/\lambda$ и учтено двойное прохождение сигналом расстояния $R(t)$. Таким образом, с учетом текущей фазы $2\pi f_0 t$ на частоте несущей и амплитудного множителя U_0 , определяемого из уравнения дальности действия, временная структура принимаемого сигнала имеет вид

$$s(t) = 2U_0 G^2(Vt/R_0) \cos(2\pi f_0 t - 2\pi V^2 t^2 / \lambda R_0 + \varphi_0). \quad (5.36)$$

На самом деле принимаемый сигнал представляет собой короткие вырезки из этого колебания — импульсы, разделенные интервалом T_n , но на процессе синтезирования апертуры это почти не отражается.

Полученное выражение показывает, что изменение амплитуды и фазы сигнала зависит от дальности R_0 , т. е. поперечной координаты объекта. Следовательно, для оптимального приема сигналов со всего диапазона дальности необходимо иметь набор из множества (сотни или тысячи) согласованных фильтров, причем каждый такой фильтр рассчитывается на суммирование сотен или тысяч импульсов. Поэтому и реализовать устройства оптимальной обработки сигналов в когерентных станциях бокового обзора с помощью традиционных электронных схем оказалось практически невозможно. Единственно пригодными остаются пока оптические процессоры.

Из сравнения формул (5.32) и (5.36) следует, что сигнал (5.36) представляет собой широкий импульс с глубокой линейной частотной модуляцией, причем скорость изменения частоты $a = -2V^2/\lambda R_0$ (частота убывает), а огибающая сигнала $A(t) = U_0 G^2(Vt/R_0)$. Однако для записи на оптический носитель нужно вместо высокой частоты f_0 перенести колебания на очень низкую частоту подставки f_n , удовлетворяющую условию $\Delta f \gg f_n > \Delta f_m$ или $1/\tau_n \gg f_n > a\tau_{обл}$, где $\Delta f = 1/\tau_n$ — ширина спектра импульсного сигнала, а $\Delta f_m = a\tau_{обл}$ — ширина спектра доплеровских частот, возникающих в течение всего времени облучения объекта $\tau_{обл}$. Перенос сигнала (5.36) с f_0 на f_n производится с помощью фазового детектора, частота опорного колебания которого отличается от частоты входного сигнала на малую величину f_n .

Выходной сигнал фазового детектора записывается на фотопленку с помощью индикатора, перед экраном (I) которого перпендикулярно развертке

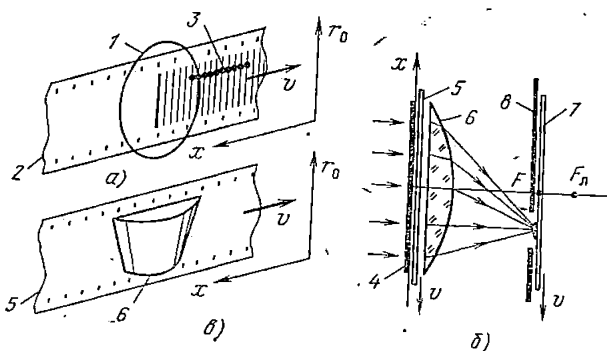


Рис. 5.14. Процессы записи (а), обработка сигнала (б) и протяжки пленки (в) в PCA

дальности протягивается фотопленка (рис. 5.14,а). Скорость протяжки пленки v пропорциональна скорости V самолета-носителя. При этом вдоль пленки в масштабе $M=v/V$, откладывается продольная координата $x=MVt=vt$, а поперек пленки — поперечная $r_0=MR_0$. Отметка (3) состоит из импульсов засвета, вплотную примыкающих один к другому, а не разделенных интервалом, соответствующим периоду повторения T_n , как раз поэтому продольная структура отметки адекватно описывается непрерывным сигналом (5.36).

После проявления пленки этот сигнал передается в указанном выше масштабе — по аналогии с формулой (5.33) — функцией прозрачности

$$T(x) = A_0 + U_0 G^2 \left(\frac{x}{r_0} \right) \exp \left[j \left(2\pi f_{x0} x - \frac{2\pi x^2}{\lambda_{св} r_0} + \varphi_0 \right) \right] + \\ + U_0 G^2 \left(\frac{x}{r_0} \right) \exp \left[-j \left(\pi f_{x0} x - \frac{2\pi x^2}{\lambda_{св} r_0} + \varphi_0 \right) \right]. \quad (5.37)$$

Сходящейся волне соответствует последнее слагаемое суммы (5.37), у которого квадратичный фазовый член положителен (свойство самофокусировки), а пространственная несущая f_{x0} отрицательна. Фокусное расстояние эквивалентной линзы, как следует из сравнения формул (5.33) и (5.37), $F_{вк}=r_0/2=MR_0/2$. Из условия равенства продольного и поперечного масштабов $M=v/V$ находим эквивалентное фокусное расстояние

$$F_{вк}=R_0 v/2V. \quad (5.38)$$

При таком фокусном расстоянии оказываются недопустимо большими габаритные размеры устройства и потери света. Поэтому рядом с пленкой ставят дополнительно фокусирующую линзу с фокусным расстоянием F_n . Фокусное же расстояние всей оптической системы F , определяемое известной формулой

$$1/F=1/F_n+1/F_{вк}, \quad (5.39)$$

меньше фокусного расстояния линзы.

Таким образом, оптический процессор для оптимальной фильтрации сигнала PCA (рис. 5.14,б) должен включать амплитудный транспарант (4) с функцией пропускания $A(x)=G^2(x/r_0)$, неподвижно закрепленный во входном окне, пленку с записью входного сигнала (5), протягиваемую со скоростью v относительно окна, линзу (6) и пленку для записи сфокусированного выходного

сигнала (7), протягиваемую синхронно со входной пленкой в фокальной плоскости оптической системы. Чтобы исключить засвет выходной пленки двумя лишними составляющими световой волны, перед ней ставится непрозрачный экран (8) с окном в области фокусировки пятна $-f_{\text{ж}}$.

Фокусное расстояние системы (5.39) через эквивалентное фокусное расстояние (5.38) зависит от дальности R_0 . Чтобы обеспечить $F = \text{const}$ для всего диапазона дальностей, фокусное расстояние $F_{\text{л}}$ линзы делают переменным в поперечном направлении, т. е. используют коническую линзу (рис. 5.14,е). Коническая линза и амплитудный транспарант образуют набор из большого числа одномерных сжимающих фильтров, имеющих различные амплитудно-фазовые характеристики. Необходимо отметить, что частичная самофокусировка сигнала обеспечивается только по продольной координате, а в поперечном направлении пятно оказывается размытым. Для исключения этого явления, называемого анаморфотностью, в оптическую систему дополнительно вводят корректирующую линзу (на рисунке не показана).

Синтезированный раскрыв антенной решетки $d_{\text{синт}} = \theta R_0$ равен линейному сечению луча бортовой антенны с раскрывом d , где $\theta = \lambda/d$ — ширина луча бортовой антенны. С возрастанием дальности R_0 до объекта синтезированный раскрыв увеличивается, поэтому ширина диаграммы направленности синтезированной решетки

$$\theta_{\text{синт}} = \lambda/d_{\text{синт}} = \lambda/\theta R_0 \quad (5.40)$$

сужается, а продольное разрешение в потенциале не зависит от дальности и равно раскрыву бортовой антенны:

$$\Delta x = \theta_{\text{синт}} R_0 = \lambda/\theta = d. \quad (5.41)$$

Это большое преимущество РСА перед обычными антеннами, у которых линейная разрешающая способность с дальностью ухудшается:

В заключение отметим, что когерентное суммирование сигналов в РСА на большом интервале времени можно обеспечить только при высокой стабильности частоты излучаемых сигналов, минимальных фазовых искажениях принимаемых сигналов в трактах приема и обработки, а также если строго выдерживается прямолинейность полета и постоянство путевой скорости на всем интервале синтеза. Поэтому реальные возможности РСА ниже потенциальных.

5.5. ГОЛОГРАФИЧЕСКИЕ МЕТОДЫ

Прямое излучение источников и отраженные от объектов колебания распространяются со скоростью света, образуя в открытом пространстве электромагнитное поле (волну). Эту волну на фотопластинку или другой светочувствительный носитель непосредственно записать нельзя, так как при этом пластинка равномерно засвечивается. Для записи необходимо «остановить» волну, локализовать ее энергию в пространстве. Тогда засвет в одних участках фотопластинки будет сильнее, в других слабее. Полученная запись может иметь или не иметь видимое сходство с наблюдаемым объектом (сюжетом), но всегда однозначно ему соответствует.

Известны два способа записи изображений: фотографический и голографический.

Фотографический способ основан на локализации энергии в пространстве путем фокусировки волны с помощью линз. Сфокусированная волна создает на сетчатке глаза, фотопластинке или фотокатоде телевизионной трубки изображение, идентичное наблюдаемому объекту, но плоское (бинокулярное зрение — двумя глазами — создает иллюзию объемности). Кроме того, оно сфокусировано только на определенную дальность, поэтому в зависимости от удаления одни объекты дают четкое изображение, а другие размытое (быстрая аккомодация хрусталика глаза делает этот недостаток зрения не очень заметным).

Голографический способ записи основан на локализации энергии в пространстве путем образования стоячей волны, возникающей при биениях двух когерентно связанных волн: опорной и отраженной от объекта (объектной). В результате биений в пространстве образуются неподвижные узлы и пучности энергии, которые на фотопластинке, помещенной в это поле, создают узор из чередующихся темных и светлых полос, подобный отпечатку пальцев человека. Этот узор, называемый голограммой, описывает амплитудно-фазовую структуру отраженной от объекта волны, не имеет видимого сходства с объектом. Чтобы голограмма отражала только свойства объекта, объектную и опорную волны берут от общего источника. При биениях из объектной волны исключаются все фазовые сдвиги, общие для облучающей и опорной волн (амплитуды их постоянны), в результате сохраняется однозначное соответствие между объектом и его записанным образом — голограммой. Перед попаданием на фотопластинку опорная и объектная волны проходят от общего первичного источника света разные пути, т. е. сдвинуты во времени. Поэтому излучаемый источником свет не может быть модулированным (изменяемым во времени); в голографии используется когерентный свет, имеющий вид гармонического колебания. Расстояние между полосами на голограмме, характеризующими конфигурацию фазового фронта стоячей волны, очень мало — порядка длины световой волны. Чтобы исключить смазывание изображения, необходима высокая когерентность излучаемых первичным источником колебаний, а также высокая разрешающая способность фотопластин, используемых для записи. Поэтому голографический способ записи практически стал возможен лишь с появлением лазеров.

Голограмма является «замороженным» изображением стоячей волны в том ее сечении, где при записи была установлена фотопластинка, т. е. она определяет граничные условия объектной волны в плоскости фотопластинки. По граничным условиям можно восстановить и всю волну в пространстве. Восстановление, «оживление» волны достигается пропусканием через голограмму или отражением от нее восстанавливающей волны. Если восстанавливающая волна подобна опорной, то в результате модуляции ее го-

логограммой получаем волну, не отличающуюся от объектной волны при записи. Визуально наблюдается объемное изображение объекта, которое, как и реальный объект, можно рассматривать с разных направлений. Полученное изображение не связано с фокусировкой, поэтому для голографической записи не требуется линз, получается одинаково четкое изображение объектов, находящихся на любом удалении.

Размеры голограммы влияют на характер получаемого изображения. Голограмму можно трактовать как оконное стекло, через которое наблюдатель рассматривает объекты. Поскольку волны от всех объектов, находящихся за окном, проходят через это стекло, а голограмма является застывшей записью волны, то при «оживлении» волны на этапе воспроизведения наблюдатель как бы непосредственно видит эти объекты. Количество наблюдаемых объектов при заданных размерах голограммы (окна) возрастает по мере приближения к ней (к нему). Более того, разбив голограмму на малые куски, по любому из них можно наблюдать те же самые объекты, что и по полной голограмме, если вплотную приблизиться к данному куску голограммы (даже через замочную скважину можно увидеть, что делается за дверью!). Однако возможности рассматривания объектов с различных направлений по мере уменьшения размеров голограммы сокращаются. Следовательно, в отличие от фотоснимка голограмма в любом ее участке содержит информацию о всех облученных объектах, видимых и не видимых из данной точки пространства. Отсюда и название голографии — $\eta\omega\lambda\omicron\sigma$ (холос) по-гречески означает «весь, полный». Голограмму можно рассматривать также как плоскую антенну, воспринимающую световую волну. Отсюда разрешающая способность изображения определяется размерами голограммы: при линейном размере голограммы d и заданной длине световой волны $\lambda_{\text{св}}$ разрешаются такие элементы объекта, угловые размеры которых на данном расстоянии не меньше $\lambda_{\text{св}}/d$. Таким образом, с уменьшением размеров голограммы сохраняются все наблюдаемые объекты, однако сокращается подробность изображения и возможность рассматривать наблюдаемую картину с различных направлений.

Процессы записи и воспроизведения в голографии протекают следующим образом (рис. 5.15). При записи падающая от когерентного источника (лазера) плоская волна (1), отражаясь от зеркала (2), образует опорную волну (3), а облучая объект (4), — расходящуюся объектную волну (5). Результат интерференции этих двух волн фиксируется на фотопластинке (6), проявив которую, получают голограмму. Если при воспроизведении плоскую восстанавливающую волну (7) подать с того же направления, что и опорную при записи, то справа от голограммы (8) возникнет расходящаяся волна (9), подобная объектной при записи. Эта волна дает мнимое изображение (10), т. е. наблюдателю кажется, что он видит сам объект, хотя по левую сторону голограммы кроме однородной восстанавливающей волны ничего нет.

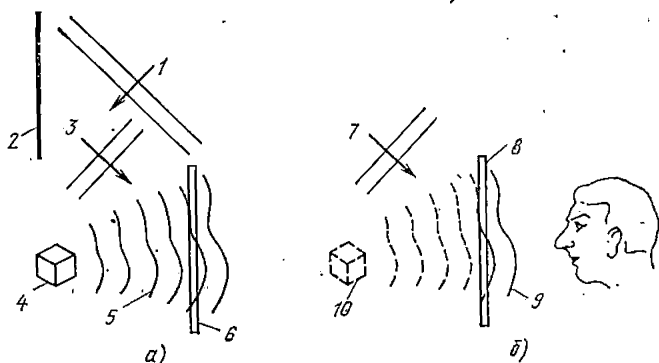


Рис. 5.15. Запись голограммы (а) и воспроизведение изображения (б)

Итак, голография — это двухэтапный процесс записи голограммы и воспроизведения изображения.

Запись. Запись в голографии является линейной операцией над полем, хотя в принципе представляет собой нелинейный процесс, так как все фотоматериалы реагируют на интенсивность (плотность потока мощности), а не на само поле (напряженность). Амплитудно-фазовая картина поля передается комплексной огибающей в функции координат. В зависимости от толщины эмульсионного слоя голограммы подразделяют на плоские (тонкий слой) и объемные (толстый слой), свойства которых существенно различаются. Ограничимся сначала тонкими голограммами.

Пусть задано амплитудно-фазовое распределение объектной S и опорной A волн в плоскости голограммы:

$$S(x, y) = |S(x, y)| \exp[-j\varphi_S(x, y)], \quad (5.42)$$

$$A(x, y) = |A(x, y)| \exp[j\varphi_A(x, y)],$$

где знак минус означает расходящуюся (от объекта) волну. Тогда интенсивность суммарного сигнала в плоскости фотопластины

$$\begin{aligned} J(x, y) &= |A + S|^2 = (A + S)(A^* + S^*) = |A|^2 + |S|^2 + A^*S + AS^* = \\ &= |A|^2 + |S|^2 + 2|A(x, y)||S(x, y)|\cos[\varphi_A(x, y) - \varphi_S(x, y)]. \end{aligned} \quad (5.43)$$

Биекции опорной и объектной волн напоминают операции в фазовом детекторе, когда роль сигнала выполняет объектная волна, а роль напряжения гетеродина — опорная. Как и у напряжения гетеродина, интенсивность опорной волны $|A|^2$ значительно больше, чем объектной, и постоянна по всей плоскости фотопластины. Поэтому составляющая $|A|^2$ интенсивности $J(x, y)$ дает сплошной равномерный фон. Используя соотношение (5.43), можно представить прозрачность полученной голограммы как

$$T(x, y) = \beta J(x, y) = T_0 + \beta(|S|^2 + A^*S + AS^*), \quad (5.44)$$

где $T_0 = \beta|A|^2$ — средняя прозрачность (фон); β — контрастная чувствительность фотопластины.

Воспроизведение. Пусть восстанавливающая волна дает в плоскости голограммы амплитудно-фазовое распределение $B(x, y)$. Тогда по другую сторону голограммы получим три составляющие:

$$C(x, y) = BT = C_1 + C_2 + C_3 = B(T_0 + \beta|S|^2) + \beta BA^*S + \beta BAS^*. \quad (5.45)$$

Составляющая $C_1 = \beta B(|A|^2 + |S|^2)$ дает почти равномерный мешающий фон, так как $|A|^2 \gg |S|^2$, а $B(x, y)$, как и $A(x, y)$, имеет в плоскости голограммы постоянную интенсивность. Составляющая C_2 — расходящаяся волна, подобная волне $S(x, y)$, исходящей от самого объекта, — дает мнимое изображение, видимое глазом. Составляющая C_3 — комплексно-сопряженная сходящаяся волна $S^*(x, y)$ — дает действительное изображение объекта, которое можно наблюдать на экране или в облаке дыма, а также непосредственно зафиксировать на фотопленке без объектива. Изображение получается неискаженным или искаженным в зависимости от структуры восстанавливающей волны.

Чтобы одно из двух изображений, мнимое или действительное, наблюдать без помех, вызванных двумя другими волнами, необходимо эти изображения на выходе плоской голограммы разделить. Разделение изображений производится по линейным пространственным частотам, т. е. по угловым координатам. Для этого при записи однородную плоскую опорную волну A направляю под углом α к оси y плоскости фотопластины (1), а с осью x она составляет прямой угол (рис. 5.16, а). Тогда распределение фазы опорной волны в плоскости фотопластины примет вид

$$\varphi_A(x, y) = 2\pi y \sin \alpha / \lambda = 2\pi f_y y, \quad (5.46)$$

где $f_y = \sin \alpha / \lambda$ — пространственная частота (по y). Так как интенсивность опорной волны на фотопластине принимаем постоянной,

$$A(x, y) = A_0 \exp(j 2\pi f_y y). \quad (5.47)$$

По-прежнему полагая

$$S(x, y) = |S(x, y)| \exp[-j\varphi_S(x, y)], \quad (5.48)$$

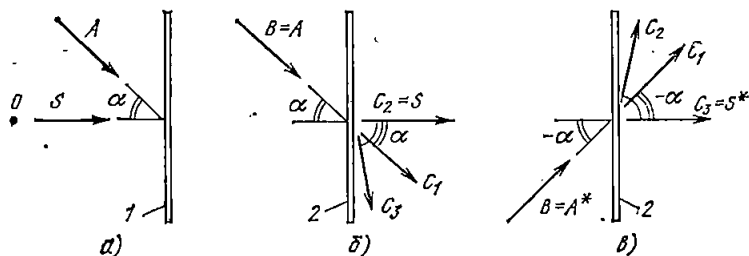


Рис. 5.16. Запись плоской голограммы объекта (а) и пространственное разделение изображений при воздействии восстанавливающей волны с направления опорной волны (б) и с симметричного ему направления (в)

находим интенсивность биений в плоскости фотопластины

$$I(x, y) = A_0 + |S(x, y)|^2 + 2A_0|S(x, y)| \cos[2\pi f_y y - \varphi_S(x, y)]. \quad (5.49)$$

Сравнение полученного результата с электрическим сигналом

$$u(t) = U_0 + U(t) + U_m(t) \cos[2\pi f t - \varphi(t)], \quad (5.50)$$

где U_0 — постоянная составляющая; $U(t)$ — видеосигнал; $U_m(t)$ и $\varphi(t)$ — амплитуда и фаза модулированного гармонического сигнала частоты f , показывает их полное сходство. Следовательно, низкочастотную составляющую $A_0 + |S(x, y)|^2$ можно отделить от высокочастотного изображения по пространственной частоте, как разделяют видео- и радиосигнал на обычных частотах. Кроме того, у пространственных сигналов несущая имеет знак, что дополнительно позволяет разделить действительное и мнимое изображения.

При воспроизведении изображения с помощью восстанавливающей волны, совпадающей с опорной ($B=A$), образуются три составляющих (рис. 5.16, б):

фон

$$C_1(x, y) = A(x, y)[T_0 + \beta|S(x, y)|^2] - A_0[T_0 + \beta|S(x, y)|^2] \times \\ \times \exp(j2\pi f_y y) = A_0[T_0 + \beta|S(x, y)|^2] \exp(j2\pi y \sin \alpha / \lambda); \quad (5.51)$$

копия расходящейся объектной волны

$$C_2(x, y) = \beta A_0 \exp(j2\pi f_y y) A_0 \exp(-j2\pi f_y y) S(x, y) = \beta A_0^2 S(x, y); \quad (5.52)$$

сходящаяся объектная волна

$$C_3(x, y) = \beta A_0 \exp(j2\pi f_y y) A_0 \exp(j2\pi f_y y) S^*(x, y) = \\ = \beta A_0^2 S^*(x, y) \exp[j2\pi(2f_y y)] = \beta A_0^2 S^*(x, y) \exp[j2\pi y (2 \sin \alpha) / \lambda]. \quad (5.53)$$

Таким образом, составляющая C_2 воспроизводит неискаженное мнимое изображение объекта в том месте относительно голограммы, где находился сам объект при записи. Составляющая C_1 создает почти равномерный фон на пространственной несущей f_y , т. е. волну, идущую под тем же углом α , что и восстанавливающая волна. Составляющая C_3 представляет собой волну на пространственной несущей $2f_y$, т. е. идущую вниз под углом $\arcsin(2 \sin \alpha)$, которая создает искаженное (сжатое по углу) действительное изображение объекта.

Если при воспроизведении восстанавливающую волну направить под углом $-\alpha$, симметричным направлению опорной волны при записи ($B=A^*$), то также получим три составляющие (рис. 5.16, в):

фон

$$C_1(x, y) = A^*(x, y)[T_0 + \beta|S(x, y)|^2] = \\ = A_0[T_0 + \beta|S(x, y)|^2] \exp(-j2\pi f_y y); \quad (5.54)$$

расходящуюся объектную волну

$$C_2(x, y) = \beta A_0^2 S(x, y) \exp[-j2\pi(2f_y)y]; \quad (5.55)$$

копию сходящейся объектной волны

$$C_3(x, y) = \beta A_0^2 S^*(x, y). \quad (5.56)$$

Составляющая C_3 дает неискаженное действительное изображение, симметричное самому объекту при записи. Фон C_1 представляет собой волну на пространственной несущей $-f_y$, т. е. идущую под углом $-\alpha$, как и восстанавливающая волна. Составляющая C_2 идет на пространственной несущей $-2f_y$, т. е. под углом $\arcsin(-2\sin \alpha)$, и дает искаженное мнимое изображение.

Таким образом, все три составляющие разделяются по пространственным частотам, подобно частотной селекции электрических сигналов. На нулевой несущей (в районе $\alpha=0$) воспроизводится нужное нам неискаженное изображение, что соответствует электрическому видеосигналу. Остальные составляющие оттесняются вниз или вверх, сжимаясь и обеспечивая наблюдение интересующего нас изображения без взаимных помех.

Как следует из приведенной теории, плоские голограммы дают при воспроизведении не одно, а два изображения объекта. Это обусловлено тем, что на тонкой эмульсии в виде линий записываются лишь следы поверхностей, образованных пучностями биений, которые в сечении голограммы имеют вид точек (рис. 5.17, а). Эти следы в равной степени могут принадлежать как объекту O , так и симметричному ему объекту O' . При воспроизведении и образуются две симметричные волны, описывающие поле обоих объектов.

Плоские голограммы были впервые получены в 1948 г. Д. Габором, который за открытие голографии в 1971 г. был удостоен Нобелевской премии. Однако им не было получено отчетливое изображение объекта, так как ему не удалось отделить одну волну от другой. Разделение волн путем записи голограммы на пространственной несущей, описанное выше, было предложено в 1962 г. Е. Лейтом и Дж. Упатниексом, которые впервые получили хорошее изображение на плоских голограммах. В том же году советский ученый Ю. Н. Денисюк предложил и получил принципиально новый вид голограмм — объемные, которые в толще эмульсии несут не следы поверхностей биений, а сами поверхности (рис. 5.17, б). В результате голограмма приобрела ряд ценных свойств.

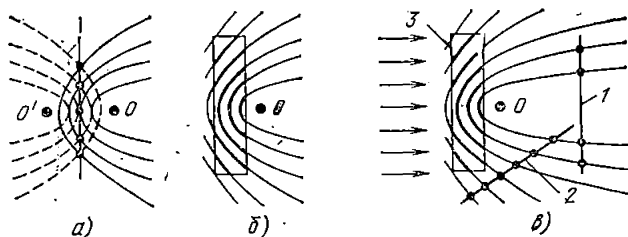


Рис. 5.17. Сечение эмульсии плоской (а) и объемной (б) голограмм и положение пластины относительно облучаемого объекта O при записи голограмм Габора (1), Лейта — Упатниекса (2) и Денисюка (3)

Во-первых, при воспроизведении образуется только одна волна, расходящаяся или сходящаяся, в зависимости от стороны облучения.

Во-вторых, объемная голограмма обладает частотной избирательностью, она подобна объемной антенне больших продольных размеров с записью падающей на нее волны. При воспроизведении ее можно облучать некогерентным, но параллельным пучком света и получить изображение только на той длине волны, на которой, была произведена запись. Можно также получить цветное объемное изображение объекта по методу цветной фотографии Липпмана: на фотопластину последовательно (до проявления) записываются голограммы в трех основных цветах — красном, зеленом и синем, а каждая из голограмм при облучении белым светом выбирает свой цвет.

В-третьих, объемная голограмма несет значительно больше информации, чем плоская. Для сравнительной оценки информационной ценности голограмм Габора, Лейта — Упатниекса и Денисюка (без учета цветности последней) рассмотрим их положение относительно картины биений, создаваемой облучаемым объектом при записи (рис. 5.17, в). Поле биений имеет наибольшую плотность линий впереди облучаемого объекта, где производится запись голограммы Денисюка, и наименьшую плотность позади объекта (голограмма Габора). Поэтому голограмма Денисюка несет наибольшее количество информации о поле, а следовательно, и о самом объекте. Однако увеличение плотности линий повышает требования к разрешающей способности фотоэмульсии: для тонких голограмм — 600...800, для толстых — 3000...5000 линий на миллиметр (в обычной фотографии пленка имеет не более 50 линий на миллиметр).

Голографические методы находят очень широкое применение в различных областях науки, техники, медицины, искусства. Среди многих направлений использования голографии в радиотехнике рассмотрим только наиболее близкий к теме книги голографический способ получения фильтров-масок, согласованных с сигналом заданной формы. В отличие от описанного ранее, при голографировании используется линза, так как в качестве объекта берется сигнальный транспарант, а записать нужно его спектральную функцию.

Устройство для получения согласованного фильтра-маски состоит из линзы, в передней фокальной плоскости которой располагается транспарант с записью сигнала $S(x)$ и схема формирования опорного луча, а в задней плоскости устанавливается фотопластина, на которой записывается голограмма (рис. 5.18). Все устройство облучается когерентным коллимированным пучком света.

Схема формирования коллимированного опорного пучка состоит из линзы малого диаметра, задний фокус которой совпадает с отверстием в экране, размещенном, как и сигнальный транспарант, в передней фокальной плоскости основной линзы. Последняя превращает расходящийся пучок лучей в параллельный (коллимированный) опорный луч. Отверстие малого диаметра в экране линзы смещено относительно центра сигнального транспаранта на некоторую величину b , благодаря чему обеспечивается разнос сигнального и опорного лучей по пространственной частоте $f_x \approx b/\lambda F$ {т. е. по углу $\alpha \approx b/F$ }.

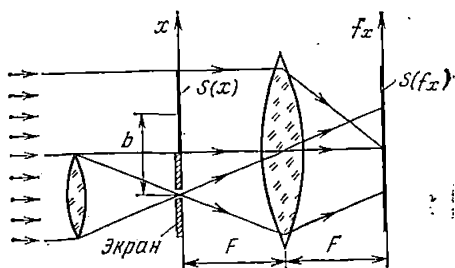


Рис. 5.18. Устройство для получения согласованного голографического фильтра-маски

Верхняя часть схемы выполняет преобразование Фурье над сигналом, записанным на транспаранте. Поэтому волна в плоскости фотопластины представляет собой спектр этого сигнала, которая дает в результате биений с опорной волной голограмму спектра.

При воспроизведении уже в самом оптическом процессоре фильтр-маска облучается под углом α , поэтому от голограммы в направлении оптической оси процессора распространяется только волна, отображающая комплексно-сопряженную со спектром сигнала функцию, что и требуется. Голографические фильтры часто создают с помощью ЭВМ.

В заключение заметим, что формирование синтезированной антенной решетки, описанное в предыдущем разделе, является примером использования голографии не в оптическом, а в радиодиапазоне. Действительно, в РСА сигнал точечного объекта, принимаемый за время облучения, представляет собой объектную волну, а подаваемое на фазовый детектор опорное колебание — опорную волну. Записанная вдоль пленки отметка объекта по яркости пропорциональна квадрату выходного напряжения фазового детектора и является, таким образом, одномерной голограммой объекта, записанной в уменьшенном масштабе по продольной координате. Поперек пленки записывается множество таких голограмм, соответствующих различным значениям наклонной дальности. Самофокусирующееся пятно, получаемое при облучении входной пленки процессора светом лазера (восстанавливающей волной), представляет собой действительное изображение объекта, воспроизводимое на выходной пленке.

5.6. ПОЛЯРИЗОВАННЫЕ СИГНАЛЫ, ИХ СПЕКТРЫ И ЗАКОНЫ ПРОХОЖДЕНИЯ ЧЕРЕЗ РАДИОТРАКТ

Изложенное выше описание пространственно-временных сигналов и теории их оптимального приема не является исчерпывающим — игнорируется поляризационная структура электромагнитных колебаний. Поскольку в радиосистемах полезная информация вырабатывается именно в процессе распространения и отражения (переизлучения) электромагнитных колебаний объектами, размещенными в пространстве, это неизбежно ведет к потере части полезной информации. Обычно поляризованный сигнал описывается двумерным вектором, свойства среды распространения и размещенных в ней объектов — матрицей размера 2×2 , а законы прохождения этого сигнала — векторно-матричным произведением. Законы прохождения поляризованного сигнала и радиосигнала через линейные цепи приемника существенно отличаются, что не позволяет создать единую теорию оптимального приема с учетом поляризационной структуры сигнала. Описание прохождения поляризованных сигналов векторно-матричным произведением относится к единственному элементу тракта, через который он проходит (излучающая антенна, среда распространения сигнала туда и обратно, объект и приемная антенна), а в совокупности представляет собой цепочку таких произведений. Кроме того, все это спра

ведливо лишь для единственной гармоник сигнала (приближенно для очень узкополосных сигналов). Ввиду этого векторно-матричное описание прохождения реальных сигналов через радиотракт чрезвычайно промоздко и необозримо, особенно применительно к широкополосным и сверхширокополосным сигналам.

Ниже приводится спектральное описание поляризованных сигналов представленных комплексными функциями времени в плоскости фронта волны (дальняя зона). Благодаря этому известные для неполяризованных сигналов законы прохождения (которые значительно проще векторно-матричных) и основные положения теории оптимального приема можно распространить и на поляризованные сигналы, узкополосные и широкополосные. Основное назначение данного параграфа — расширить понятие о сигналах и их спектрах, раскрыть физический смысл некоторых представлений, не просматривающийся у обычных, чисто временных сигналов, через свойства пространственно-временных сигналов.

Рассмотрим вначале поведение во времени гармонического сигнала с круговой поляризацией в комплексной плоскости фронта волны, нормальной направлению ее распространения в дальней зоне. Если смотреть на эту плоскость в направлении движения волны, горизонтальную ось считать действительной, а перпендикулярную ей ось — мнимой и представлять фазу колебаний (как это принято) углом φ влево от горизонтальной оси, то при левом вращении вектора напряженности поля фаза возрастает и угловая скорость — круговая частота $\omega = d\varphi/dt > 0$, а при правом вращении фаза убывает и $\omega = d\varphi/dt < 0$. Следовательно, знак частоты (во времени) или угловой скорости (в пространстве) указывает направление вращения вектора поля гармонического сигнала с круговой поляризацией: при левом вращении частота положительна, а при правом — отрицательна. Модуль вектора поля a и абсолютное значение частоты $|\omega|$ у гармонического сигнала постоянны.

Таким образом, гармонический сигнал частоты $\pm\omega$ с круговой поляризацией в плоскости фронта волны можно представить комплексной функцией времени

$$s(t) = a_{\pm} e^{\pm j(\omega t \pm \varphi_{\pm})} = A [\cos(\omega t \pm \varphi_{\pm}) \pm j \sin(\omega t \pm \varphi_{\pm})] / \sqrt{2}, \quad (5.57)$$

где знаки плюс и минус относятся к сигналам, лево- и правосторонне поляризованным. При этом за каждый период колебаний конец вектора поля описывает окружность в направлении, указанном стрелкой (рис. 5.19, а). Радиус окружности равен модулю напряженности поля $a_{\pm}$, а начальная фаза $\varphi_{\pm}$ определяется начальным угловым положением вектора или точкой на окружности. В направлении распространения волны конец вектора поля описывает пространственную спираль с временным $T_0 = 2\pi/\omega$ или пространственным периодом — длиной волны $\lambda = cT_0 = 2\pi c/\omega$.

Комплексный двусторонний спектр сигнала (5.57) находим как прямое преобразование Фурье при фиксированной частоте $\omega = \omega_1$:

$$s(\omega) = \int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt = a_{\pm} e^{j\varphi_{\pm}} \int_{-\infty}^{\infty} e^{-j(\omega \mp \omega_1)t} dt = \\ = a_{\pm} e^{j\varphi_{\pm}} \delta(\omega \mp \omega_1), \quad (5.58)$$

где $\delta(\omega \mp \omega_1)$ — дельта-функция, равная единице при $\omega = \omega_1$ или $\omega = -\omega_1$ и нулю при всех остальных ω . Следовательно, спектр сигнала (5.57) содержит спектральные линии на данной частоте $\omega = \omega_1$ только одного знака (рис. 5.19, б). Сигналы с таким двусторонним спектром называются аналитическими, поэтому гармоническое колебание с круговой поляризацией является физическим образом аналитического гармонического сигнала.

Суммирование двух колебаний с круговой поляризацией встречного вращения, имеющих одинаковую частоту $|\omega|$, но различающихся по модулю a и начальной фазе φ , дает эллиптически-поляризованный сигнал, описываемый комплексной функцией

$$s(t) = a_+ e^{+j(\omega t + \varphi_+)} + a_- e^{-j(\omega t - \varphi_-)} = \\ = A e^{j\Phi} [\cos \varepsilon \cos(\omega t + \varphi_0) + j \sin \varepsilon \sin(\omega t + \varphi_0)], \quad (5.59)$$

где

$$A = \sqrt{A_I^2 + A_{II}^2}, \quad A \geq 0, \quad (5.60)$$

— амплитуда эллиптически-поляризованного сигнала, характеризующая размер эллипса; $A_I = A \cos \varepsilon = a_+ + a_-$ и $A_{II} = A \sin \varepsilon = a_+ - a_-$ — большая и малая полуоси эллипса (амплитуды двух ортогональных линейно-поляризованных составляющих сигнала), при-

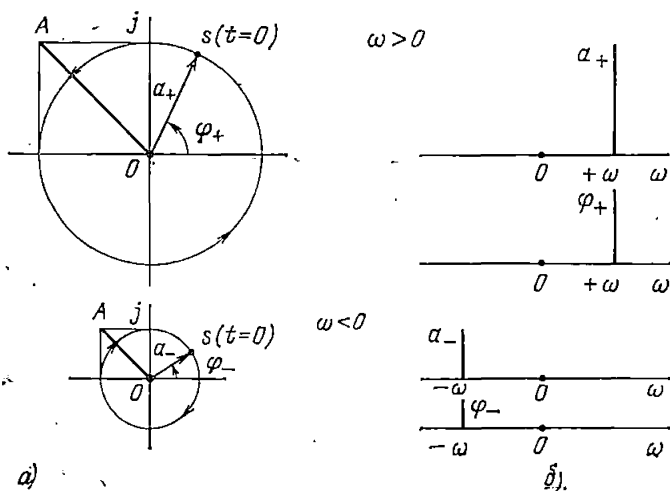


Рис. 5.19. Поляризационная структура и двусторонние спектры гармонических сигналов, левосторонние (а) и правосторонние (б) поляризованных

чем $\cos \varepsilon$ и $\sin \varepsilon$ представляют большую и малую полуоси единичного эллипса ($A=1$);

$$\varphi_0 = (\varphi_+ - \varphi_-)/2, \quad -\pi < \varphi_0 \leq \pi, \quad (5.61)$$

— начальная фаза суммарной напряженности поля;

$$\Phi = (\varphi_+ + \varphi_-)/2, \quad -\pi/2 < \Phi \leq \pi/2, \quad (5.62)$$

— угол ориентации эллипса (наклона его большой полуоси относительно горизонтальной оси);

$$\varepsilon = \operatorname{arctg} \frac{A_{II}}{A_I} = \operatorname{arctg} \frac{a_+ - a_-}{a_+ + a_-}, \quad -\frac{\pi}{4} \leq \varepsilon \leq \frac{\pi}{4}, \quad (5.63)$$

— угол округлости эллипса, под которым видна малая полуось из вершины большой, характеризующий форму эллипса, как и коэффициент округлости эллипса

$$r = \operatorname{tg} \varepsilon = \frac{A_{II}}{A_I} = \frac{a_+ - a_-}{a_+ + a_-}, \quad 1 \leq r \leq 1. \quad (5.63a)$$

Пределы в формулах (5.61) — (5.63) обеспечивают однозначность отсчета параметров эллиптически-поляризованного сигнала (рис. 5.20, а). Если полусумма (5.62) выходит за указанные пределы, значения Φ и φ_0 следует изменить на π .

Направление вращения (знак ω) и начальная фаза (положение) в пространстве суммарного вектора $s(t=0)$ определяются

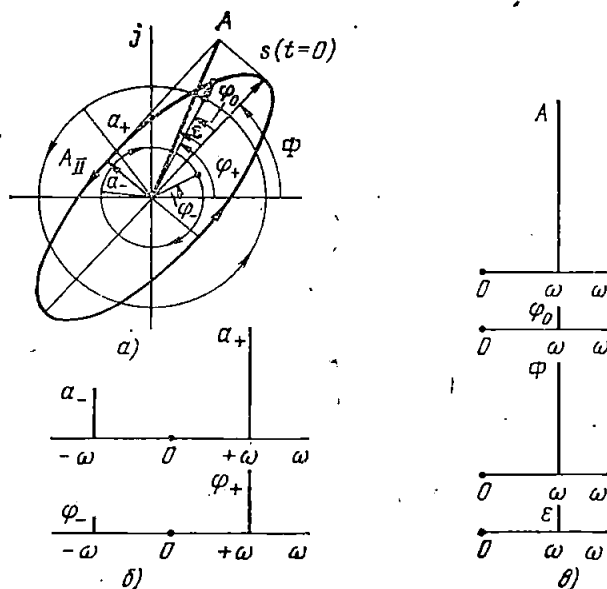


Рис. 5.20. Поляризационная структура (а), двусторонний (б) и односторонний (в) спектры эллиптически-поляризованного гармонического сигнала как суммы двух сигналов противоположной поляризации

наибольшей из двух составляющих, в данном случае левосторонне поляризованной ($\epsilon > 0$). Однако начальная фаза суммарной напряженности поля φ_0 уже не равна начальной фазе (φ_+) наибольшей составляющей и отсчитывается от большой полуоси эллипса, а не от горизонтальной оси. При обратном соотношении значений составляющих ($A_{II} = a_+ - a_- < 0$) вращение изменяется на противоположное ($\epsilon < 0$), и начальная фаза φ_0 отсчитывается от большой полуоси эллипса в направлении φ_- .

Первое выражение в формуле (5.59) описывает эллиптически-поляризованный гармонический сигнал в параметрах двустороннего спектра a_+ , φ_+ и a_- , φ_- , определяемого по формуле (5.58) как сумма спектров составляющих сигналов с круговой поляризацией встречного вращения (рис. 5.20, б). Второе выражение (5.59) описывает этот сигнал в параметрах одностороннего спектра A , φ_0 и Φ , ϵ , непосредственно характеризующих поляризационную структуру сигнала (рис. 5.20, в). Параметры $a_{\pm}$ и A являются энергетическими, остальные неэнергетическими.

Формула (5.59) дает самое общее описание поляризованного гармонического сигнала. При $a_- = 0$ или $a_+ = 0$ эллиптически-поляризованный сигнал обращается в один из сигналов (5.57), а при $a_+ = a_- = a$ — в линейно-поляризованный гармонический сигнал

$$s(t) = a [e^{+j(\omega t + \varphi_+)} + e^{-j(\omega t - \varphi_-)}] = A e^{j\Phi} \cos(\omega t + \varphi_0). \quad (5.64)$$

При этом эллипс обращается в отрезок прямой длиной $2A$, угол наклона Φ которого относительно горизонтальной оси характеризует положение плоскости поляризации, $\epsilon = 0$ ($A_I = 2A$, $A_{II} = 0$), а двусторонний спектр модуля напряженности поля приобретает четную симметрию относительно нулевой частоты (рис. 5.21, а). Если условие четной симметрии спектра модуля дополнить условием нечетной симметрии фазы $\varphi_- = -\varphi_+ = -\varphi_0$, получим горизон-

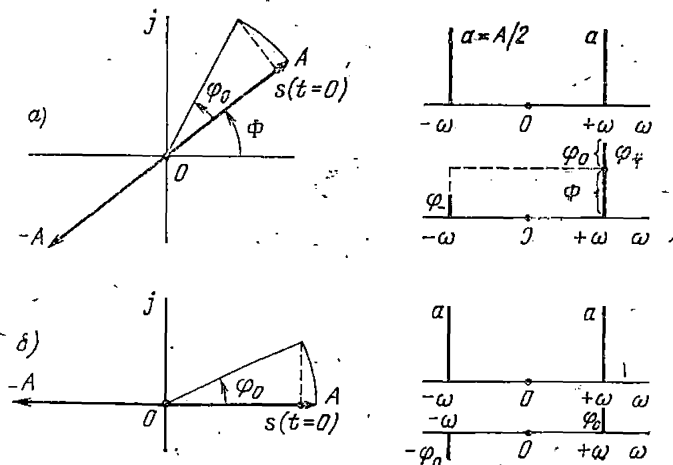


Рис. 5.21. Поляризационная структура и двусторонние спектры линейно-поляризованных сигналов: а — общий случай; б — горизонтальная поляризация

тально линейно-поляризованный или обычный, неполяризованный гармонический сигнал

$$s(t) = a[e^{+j(\omega t + \varphi_0)} + e^{-j(\omega t + \varphi_0)}] = A \cos(\omega t + \varphi_0), \quad (5.65)$$

являющийся действительной функцией времени как сумма двух комплексно-сопряженных сигналов (рис. 5.21, б).

Практически действительную функцию (5.65) целесообразно использовать для описания чисто временных, неполяризованных сигналов, а горизонтально поляризованный сигнал условно описывать функцией $s(t) = Ae^{j0} \cos(\omega t + \varphi_0)$, что не меняет сути дела, но позволяет отличить поляризованный сигнал от неполяризованного.

Поляризованный сигнал изменяет как свое значение, так и ориентацию в плоскости фронта волны; у него лишь в частном случае $\varepsilon=0$ и $\Phi=0$. Обычный, временной сигнал, изменяющий только свое значение, следует рассматривать как вырожденный случай, когда $\varepsilon=0$ и $\Phi=0$, потому, что к нему они вообще не применимы. Временной гармонический сигнал характеризуется лишь двумя параметрами A и φ_0 , а поляризованный, пространственно-временной сигнал содержит четыре параметра A , φ_0 и Φ , ε , из которых первые два временные (отображают комплексное значение напряженности поля), а вторые два пространственные (отображают поляризационную структуру поля).

Энергетические параметры $a_{\pm}$ или A определяют текущую мощность (плотность потока мощности) сигнала на волновом сопротивлении 1 Ом. Текущая мощность сигнала с круговой поляризацией (5.57) $P = |s(t)|^2 = a_{\pm}^2 = A^2/2$ есть величина постоянная. У эллиптически- и линейно-поляризованных сигналов текущая мощность в течение каждого периода колебаний переменная, а средняя за период мощность определяется как $P = A^2/2 = (A_1^2 + A_2^2)/2 = = [(a_+ + a_-)^2 + (a_+ - a_-)^2]/2 = a_+^2 + a_-^2$ и $P = A^2/2 = 2a^2$ соответственно, т. е. равна сумме постоянных мощностей составляющих левого и правого вращения. Здесь традиционное истолкование амплитуды полностью применимо лишь к линейно-поляризованным и неполяризованным сигналам; у эллиптически-поляризованных сигналов и сигналов с круговой поляризацией амплитуда вводится через мощность — путем суммирования в квадратуре амплитуд двух ортогональных линейно-поляризованных составляющих (см. рис. 5.19 и 5.20).

По виду двустороннего спектра легко оценить поляризационную структуру сигнала. Форма эллипса, определяемая коэффициентом r или углом округлости ε , является функцией асимметрии спектра модуля поля. При полной асимметрии ($a_- = 0$ или $a_+ = 0$) эллипс вырождается в окружность ($|r| = 1$ или $|\varepsilon| = \pi/4$ максимумы), при отсутствии асимметрии ($a_+ = a_- = a$) — в отрезок прямой ($r = \varepsilon = 0$). Угол ориентации эллипса Φ (и начальная фаза колебаний φ_0) зависит от соотношения фаз в положительной и отрицательной областях спектра: $|\Phi| = 0$ при нечетной симметрии

фазового спектра ($\varphi_- = -\varphi_+$) и возрастает с уменьшением степени нечетной симметрии.

Поляризационная структура сигнала, характеризующаяся параметрами Φ и ε , графически представляется единичным вектором, а совокупность концов всех этих векторов, исходящих из общего центра, образует геометрическое место точек (сферу Пуанкаре). Определенной поляризации соответствует точка на сфере, образуемая пересечением меридиана $2\Phi = \text{const}$ и параллели $2\varepsilon = \text{const}$. Левому вращению соответствует верхняя полусфера ($\varepsilon > 0$), правому — нижняя ($\varepsilon < 0$); на экваторе $2\varepsilon = 0$, что указывает на линейную поляризацию. Окружность эллипса от нулевой на экваторе возрастает до полной на полюсах ($2\varepsilon = \pm \pi/2$). Нулевому меридиану ($2\Phi = 0$) соответствует горизонтальная ориентация эллипса (линии), а противоположному ($2\Phi = \pi$) — вертикальная. Тем самым вся сфера представляет любой вид поляризации.

Временная структура поляризованного гармонического сигнала частоты ω описывается комплексной амплитудой $S = A e^{j\varphi_0}$, которую графически можно представить вектором в комплексной плоскости, касательной к сфере Пуанкаре в точке, характеризующей поляризацию сигнала. Комбинация сферы и плоскости определяет полную совокупность поляризованных гармонических сигналов различной амплитуды и фазы. Неполаризованным сигналам соответствует только комплексная амплитуда и комплексная плоскость.

Конкретный вид поляризации имеет, строго говоря, только гармонический сигнал. С некоторым допущением это можно отнести и к узкополосным негармоническим сигналам, у которых составляющие спектра почти не отличаются по частоте от несущей ω_0 . Однако к широкополосным и сверхширокополосным сигналам такое представление о поляризации совершенно неприемлемо. Поляризационные (и не поляризационные) параметры таких сигналов изменяются во времени медленно или быстро в зависимости от ширины спектра сигнала — гармонический поляризованный сигнал общего вида (5.59) оказывается модулированным по любому из параметров или по всем одновременно:

$$s(t) = a_+(t) e^{+j[\omega_0 t + \varphi_+(t)]} + a_-(t) e^{-j[\omega_0 t + \varphi_-(t)]} = \\ = A(t) e^{j\Phi(t)} \{ \cos \varepsilon(t) \cos [\omega_0 t + \varphi_0(t)] + j \sin \varepsilon(t) \sin [\omega_0 t + \varphi_0(t)] \},$$

где $A(t)$ — огибающая сигнала. Модуляция вида $A(t)$ называется амплитудной, $\varphi_0(t)$ — угловой, $\Phi(t)$ и $\varepsilon(t)$ — поляризационной модуляцией ориентации и формы эллипса с учетом смены стороны вращения (знака).

Угловая модуляция, как отмечалось, подразделяется на фазовую, когда модулирующей функции пропорциональна сама фаза, и частотную, когда модулирующую функцию повторяет частота — производная от фазы $\varphi_0(t)$.

Спектр негармонических сигналов, вообще говоря, непрерывный. Поэтому энергетические параметры $a_{\pm}$ и A заменяют на спектральные плотности модуля $|s(\omega)|$ и амплитуды $|S(\omega)|$:

$$da_{\pm} = |s(\omega)| d\omega/2\pi \text{ и } dA = |S(\omega)| d\omega/2\pi. \quad (5.66)$$

Тогда вместо параметров гармонических сигналов

$$a_{\pm}, \varphi_{\pm}; a, \varphi_0 \text{ и } A, \varphi_0; \Phi, \epsilon \quad (5.67)$$

вводятся соответственно двусторонний и односторонний спектры широкополосных сигналов:

$$|s(\omega)| = |s(\omega_+)| + |s(\omega_-)|, \quad \varphi(\omega) = \varphi(\omega_+) + \varphi(\omega_-) \quad (5.68)$$

и

$$|S(\omega)|, \varphi_0(\omega); \Phi(\omega), \epsilon(\omega),$$

где индексы плюс и минус относятся к положительной и отрицательной областям частот двустороннего спектра соответственно. Переход от двустороннего спектра к одностороннему производится по формулам (5.60) — (5.63) с соответствующей заменой обозначений (5.67) на (5.68) и учетом соотношений (5.66).

Таким образом, двусторонний спектр широкополосных сигналов в общем случае несимметричный по энергетической составляющей, содержит две составляющих, а односторонний — четыре (рис. 5.22, а и б соответственно). Односторонний спектр можно заменить пространственной фигурой, в каждом сечении которой воспроизводятся все параметры частного гармонического сигнала (рис. 5.22, в). Суммарный широкополосный поляризованный сиг-

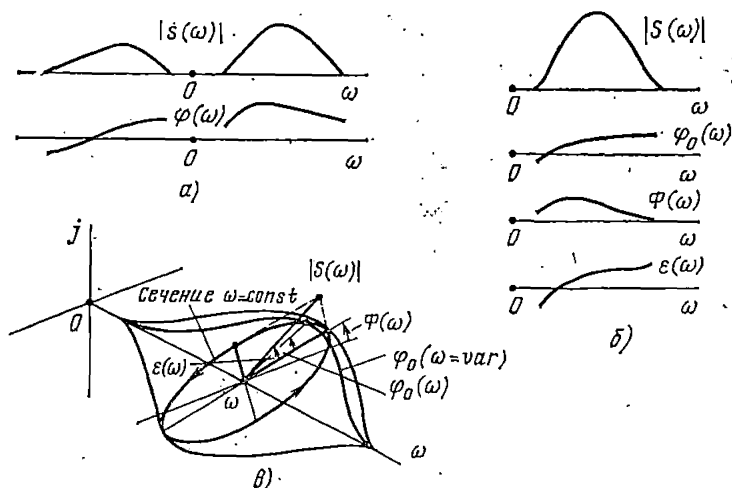


Рис. 5.22. Двусторонний (а), односторонний (б) спектры и объемный образ одностороннего спектра (в) широкополосного поляризованного сигнала

нал находится как обратное преобразование Фурье от двустороннего или одностороннего спектра:

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} s(\omega) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_0^{\infty} S(\omega) e^{j\omega t} d\omega. \quad (5.69)$$

В первом случае он является суперпозицией по всей оси частот гармонических сигналов с круговой поляризацией, во втором частные гармонические сигналы имеют произвольную поляризацию и простираются только в области частот, не имеющих знака.

Спектральное фурье-разложение поляризованных сигналов позволяет почти так же легко описать их прохождение через линейный тракт, как и неполяризованных, если пользоваться двусторонним спектром. Поскольку в линейной системе гармонические сигналы сохраняют свою форму, выходной комплексный двусторонний спектр находится простым умножением входного спектра на комплексную двустороннюю функцию передачи радиотракта, т. е. среды распространения вместе с размещенными в ней объектами. В радиолокационном случае радиотракт включает пять основных элементов: излучающая антенна, среда распространения от излучателя к объекту, отражающий объект, среда распространения от объекта до приемной антенны и сама приемная антенна.

Однако прохождение поляризованных волн имеет важную особенность. Наряду с прямым прохождением спектральных составляющих сигнала через радиотракт возможно перекрестное прохождение через один или несколько его элементов. Последнее состоит в том, что любая гармоническая составляющая поляризованного сигнала может изменить направление вращения, т. е. знак частоты, например, при облучении идеально отражающего плоского объекта гармоническим сигналом с круговой поляризацией по нормали к его плоскости (сигнал левого вращения — частота положительна — приобретает правое вращение — частота отрицательна, и наоборот).

Для описания прямого и перекрестного прохождения сигналов введем вектор-столбец двустороннего спектра входного сигнала $S_0(\omega) = [s_0(\omega), s_0(-\omega)]^T$, где $s_0(\omega)$ — входной спектр, $s_0(-\omega)$ — его зеркальное отражение относительно нулевой частоты, а « T » означает транспонирование. Одновременно введем вектор-столбец двусторонней функции передачи $H(\omega) = [h(\omega), \tilde{h}(\omega)]^T$, где $h(\omega)$ и $\tilde{h}(\omega)$ — прямая и перекрестная функции передачи. Тогда двусторонний спектр выходного сигнала находится как произведение векторов:

$$s(\omega) = H^T(\omega) S_0(\omega) = h(\omega) s_0(\omega) + \tilde{h}(\omega) s_0(-\omega). \quad (5.70)$$

Если в каком-либо элементе радиотракта перекрестная передача отсутствует (например, в среде без ионосферы), то выходной спектр этого элемента $s(\omega) = h(\omega) s_0(\omega)$, как и в обычных радиоцепях. Закон изменения спектра при отражении от объекта позволяет определить поляризационные свойства объекта, а по ним подробные

характеристики самого объекта (конфигурацию, ориентацию, материал).

В заключение отметим, что структура излучаемого и принимаемого (входного) поляризованного сигналов различна (в отличие от неполяризованных и узкополосных сигналов). В воспринимающем элементе приемной антенны возможны любые преобразования поляризации входного сигнала, в том числе перевод произвольной поляризации в линейную. Поскольку структура такого сигнала и неполяризованного одинакова (оба сигнала описываются комплексной огибающей $S(t)$ или ее спектром $S(\omega)$), теория оптимального приема неполяризованных сигналов, включая обнаружение и измерение параметров, естественным образом распространяется и на поляризованные сигналы.

Глава 6. ПОТЕНЦИАЛЬНАЯ ТОЧНОСТЬ И РАЗРЕШАЮЩАЯ СПОСОБНОСТЬ РАДИОСИСТЕМ

6.1. МНОГОМЕРНАЯ ПРОСТРАНСТВЕННО-ВРЕМЕННАЯ ФУНКЦИЯ КОРРЕЛЯЦИИ И ЧАСТНЫЕ ФУНКЦИИ КОРРЕЛЯЦИИ

Выходной сигнал оптимального приемника — корреляционный интеграл — содержит сигнальную и шумовую составляющие. Нормирование сигнальной составляющей корреляционного интеграла (3.40) по энергии (3.41) дает многомерную пространственно-временную функцию корреляции

$$\Psi_{\text{нв}}(\mathbf{v}') = \frac{Z_s(\mathbf{v}')}{\mathcal{E}} = \frac{1}{2\mathcal{E}} \int \int \int_{-\infty}^{\infty} |G(v_x, v_y)|^2 S(t) S^*(t - \tau') \exp \{j 2\pi [F' t + v_x(u'_x + \dot{u}'_x t) + v_y(u'_y + \dot{u}'_y t)]\} dv_x dv_y dt. \quad (6.1)$$

Наряду с отношением сигнал-шум $\mathcal{E}/N_0$ эта функция определяет потенциальные возможности системы в отношении точности и однозначности измерений параметров движения наблюдаемых объектов, а также разрешающей способности системы применительно и к совокупности этих параметров, и к каждому параметру отдельно. В свою очередь, многомерная функция корреляции (6.1) описывается только двумя входными характеристиками системы: временной (сигналом) $S(t)$ или $S(f)$ и пространственной (апертурой) $G(u_x, u_y)$ или $G(v_x, v_y)$. Отсюда следует, что свойства радиосистемы определяются выбором ее временной и пространственной характеристик (при заданном отношении сигнал-шум). В много-

мерной функции корреляции (6.1) связь временных и пространственных характеристик проявляется через угловые скорости, которые, как уже отмечалось, приводят к искажению спектра сигнала и, следовательно, расстройке приемного устройства. Однако при работе по удаленным объектам угловые скорости малы, и поэтому во многих практических случаях (сравнительно узкополосания по угловым скоростям $\dot{u}_x$ и $\dot{u}_y$ на выходной сигнал измерительной системы можно пренебречь. Тогда пространственно-временную функцию корреляции (6.1) можно представить произведением двух независимых функций:

$$\Psi_{\text{пв}}(\mathbf{v}') = \Psi(\tau', F') \Psi(u'_x, u'_y), \quad (6.2)$$

где

$$\begin{aligned} \Psi(\tau', F') &= \frac{1}{2\vartheta} \int_{-\infty}^{\infty} S(t) S^*(t - \tau') e^{j 2\pi F' t} dt = \\ &= \frac{1}{2\vartheta} \int_{-\infty}^{\infty} S^*(f) S(f - F') e^{j 2\pi f \tau'} df \end{aligned} \quad (6.3)$$

— временная функция корреляции;

$$\begin{aligned} \Psi(u'_x, u'_y) &= \frac{\lambda^2}{A} \iint_{-\infty}^{\infty} G(u_x, u_y) G^*(u_x - u'_x, u_y - u'_y) du_x du_y = \\ &= \frac{\lambda^2}{A} \iint_{-\infty}^{\infty} |G(v_x, v_y)|^2 e^{j 2\pi (v_x u'_x + v_y u'_y)} dv_x dv_y \end{aligned} \quad (6.4)$$

— пространственная функция корреляции (A — эффективная площадь антенны). При этом временная функция корреляции определяется только временной характеристикой системы (сигналом), а пространственная — пространственной характеристикой (апертурой).

Здесь даны две формы представления функции корреляции: первая через сами функции временных, частотных или угловых сдвигов, вторая через энергетические спектры, так как функция корреляции сигнала и его энергетический спектр связаны преобразованием Фурье, которое и выражает правая часть формул (6.3) и (6.4). В частности, при нулевых расстройках интегралы дают два равенства Парсеваля:

$$\vartheta = \frac{1}{2} \int_{-\infty}^{\infty} |S(t)|^2 dt = \frac{1}{2} \int_{-\infty}^{\infty} |S(f)|^2 df; \quad (6.5)$$

$$A = \lambda^2 \iint_{-\infty}^{\infty} |G(u_x, u_y)|^2 du_x du_y = \lambda^2 \iint_{-\infty}^{\infty} |G(v_x, v_y)|^2 dv_x dv_y, \quad (6.6)$$

представляющие энергию сигнала и эффективную площадь антенны соответственно.

Временная функция корреляции $\Psi(\tau', F')$ (квадрат модуля которой $|\Psi(\tau', F')|^2$ называют функцией неопределенности Вудворда) определяет взаимную зависимость точности и разрешающей способности при измерении дальности и скорости с помощью одного и того же сигнала. Пространственная функция корреляции

$\Psi(u'_x, u'_y)$, как и пространственная функция неопределенности $|\Psi(u'_x, u'_y)|^2$, определяет те же зависимости при измерении двух независимых угловых координат с помощью общей антенны при заданной частоте несущей.

Положив в формулах (6.3) и (6.4) расстройку по одному из параметров нулевой, получим функции корреляции только по дальности, только по скорости или только по одной из двух угловых координат. Чтобы найти расстройку по составляющим угловой скорости, необходимо воспользоваться самой пространственно-временной функцией корреляции (6.1), положив в ней нулевыми все расстройки, кроме расстройки по данной составляющей угловой скорости. В итоге, опуская бесконечные пределы интегрирования, получаем частные функции корреляции по дальности, радиальной скорости, угловым координатам и их производным соответственно:

$$\Psi(\tau) = \frac{\int S(t) S^*(t - \tau) dt}{\int |S(t)|^2 dt} = \frac{\int |S(f)|^2 \exp(j 2\pi f \tau) df}{\int |S(f)|^2 df}, \quad (6.7)$$

$$\Psi(F') = \frac{\int S^*(f) S(f - F') df}{\int |S(f)|^2 df} = \frac{\int |S(t)|^2 \exp(j 2\pi F' t) dt}{\int |S(t)|^2 dt}, \quad (6.8)$$

$$\begin{aligned} \Psi(u'_{x,y}) &= \frac{\int G(u_{x,y}) G^*(u_{x,y} - u'_{x,y}) du_{x,y}}{\int |G(u_{x,y})|^2 du_{x,y}} = \\ &= \frac{\int |G(v_{x,y})|^2 \exp(j 2\pi v_{x,y} u'_{x,y}) dv_{x,y}}{\int |G(v_{x,y})|^2 dv_{x,y}}, \end{aligned} \quad (6.9)$$

$$\Psi(\ddot{u}'_{x,y}) = \frac{\iint |G(v_{x,y})|^2 |S(t)|^2 \exp(j 2\pi v_{x,y} \dot{u}'_{x,y} t) dv_{x,y} dt}{\int |G(v_{x,y})|^2 dv_{x,y} \int |S(t)|^2 dt}. \quad (6.10)$$

Последние две формулы справедливы только при независимых функциях раскрыва по каждой из двух осей, когда $G(v_x, v_y) = G(v_x) G(v_y)$ и, следовательно, $G(u_x, u_y) = G(u_x) G(u_y)$. В общем случае эти функции по осям x и y зависимы и вместо формул (6.9), (6.10) лучше пользоваться наиболее общими функциями

$$\begin{aligned} \Psi(u'_x) &= \frac{\iint G(u_x, u_y) G^*(u_x - u'_x, u_y) du_x du_y}{\iint |G(u_x, u_y)|^2 du_x du_y} = \\ &= \frac{\iint |G(v_x, v_y)|^2 \exp(j 2\pi v_x u'_x) dv_x dv_y}{\iint |G(v_x, v_y)|^2 dv_x dv_y}, \\ \Psi(u'_y) &= \frac{\iint G(u_x, u_y) G^*(u_x, u_y - u'_y) du_x du_y}{\iint |G(u_x, u_y)|^2 du_x du_y} = \\ &= \frac{\iint |G(v_x, v_y)|^2 \exp(j 2\pi v_y u'_y) dv_x dv_y}{\iint |G(v_x, v_y)|^2 dv_x dv_y}; \end{aligned} \quad (6.11)$$

$$\left. \begin{aligned} \Psi(u'_x) &= \frac{\iint \iint |G(v_x, v_y)|^2 |S(t)|^2 \exp(j 2\pi v_x u'_x t) dv_x dv_y dt}{\iint |G(v_x, v_y)|^2 dv_x dv_y \int |S(t)|^2 dt} \\ \Psi(u'_y) &= \frac{\iint \iint |G(v_x, v_y)|^2 |S(t)|^2 \exp(j 2\pi v_y u'_y t) dv_x dv_y dt}{\iint |G(v_x, v_y)|^2 dv_x dv_y \int |S(t)|^2 dt} \end{aligned} \right\} \quad (6.12)$$

Можно пользоваться функциями вида (6.9), (6.10) и в общем случае, если вместо $G(v_x)$ и $G(v_y)$ подставить усредненные по второй координате эквивалентные функции раскрыва

$$G_{\text{эк}}(v_x) = \int G(v_x, v_y) dv_y, \quad G_{\text{эк}}(v_y) = \int G(v_x, v_y) dv_x. \quad (6.13)$$

В частности, при равномерном облучении $G(v_x, v_y) = 1$ круглого раскрыва диаметром d получим $G_{\text{эк}}(v_{x,y}) = 2 \sqrt{(d/2\lambda)^2 - v_{x,y}^2}$, $|v_{x,y}| \leq d/2\lambda$.

Частные функции корреляции описывают поведение выходного сигнала оптимального приемника измерительной системы и потенциальные возможности системы в том случае, когда все параметры принимаемого сигнала, кроме данного, известны или расстройки по ним отсутствуют.

6.2. ОБОБЩЕННАЯ ФУНКЦИЯ КОРРЕЛЯЦИИ

Ввиду взаимного сходства частных функций корреляции (6.7) — (6.10) их можно получить простой заменой обозначений, только функция корреляции по угловой скорости имеет некоторые особенности, которые будут оговорены ниже. Основываясь на этом сходстве, можно ввести обобщенную функцию корреляции для произвольного измеряемого неэнергетического параметра v , что позволяет определить потенциальные возможности измерителя по точности и разрешающей способности, не рассматривая каждый конкретный параметр отдельно.

Для этого введем понятия обобщенной огибающей $S(v)$, обобщенной частоты μ , определяющей скорости изменения фазы принимаемого сигнала (волны) $\phi(v) = 2\pi\mu v$ по параметру v , и спектр обобщенной огибающей $S(\mu)$, являющийся преобразованием Фурье от огибающей $S(v)$. При указанных обозначениях введенных понятий обобщенная функция корреляции принимает вид

$$\Psi(v') = \frac{\int_{-\infty}^{\infty} S(v) S^*(v - v') dv}{\int_{-\infty}^{\infty} |S(v)|^2 dv} = \frac{\int_{-\infty}^{\infty} |S(\mu)|^2 \exp(j 2\pi \mu v') d\mu}{\int_{-\infty}^{\infty} |S(\mu)|^2 d\mu}, \quad (6.14)$$

причем

$$\int_{-\infty}^{\infty} |S(v)|^2 dv = \int_{-\infty}^{\infty} |S(\mu)|^2 d\mu. \quad (6.15)$$

Таблица 6.1

Измеряемый параметр v	Обобщенная огнибающая $S(v)$	Текущая фаза $2\pi v$	Обобщенная частота μ	Спектр обобщенной огнибающей $S(\mu)$
τ F u_x, v $\dot{u}_x, v$	$S(t)$ $S(f)$ $G(u_x, v)$ —	$2\pi f\tau$ $2\pi tF$ $2\pi v_{x,y} u_{x,y}$ $2\pi v_{x,y} t \dot{u}_{x,y}$	f t $v_{x,y}$ $v_{x,y} t$	$S(f)$ $S(t)$ $G(v_{x,y})$ $G(v_{x,y}) S(t)$

Обобщенная функция корреляции (6.14) принимает конкретный вид (6.7) — (6.10), если обобщенные обозначения заменить истинными согласно табл. 6.1.

Данные таблицы показывают, что только у функции корреляции по времени запаздывания (дальности) обобщенные понятия полностью соответствуют истинным. У функции корреляции по частотному сдвигу (радиальной скорости) обобщенные и истинные понятия огнибающей и ее спектра меняются местами. У функции корреляции по угловым координатам роль обобщенной огнибающей играет комплексная диаграмма направленности, а роль ее спектра — функция раскрыва, т. е. угловой пространственный спектр. Ввиду особого положения угловой скорости как пространственного и временного понятия одновременно обобщенная частота и обобщенный спектр представляются в виде двух сомножителей, пространственного и временного, а одномерная корреляционная функция (6.10) представляется интегралом по двум переменным: координате (угловой пространственной частоте) и времени. В этом случае сама обобщенная огнибающая не выражается через известные функции, хотя формально ее можно ввести как двумерное преобразование Фурье от произведения $G(v_{x,y})S(t)$.

Необходимо отметить, что в частных функциях корреляции следует, вообще говоря, различать текущий параметр сигнала и измеряемый параметр, имеющий одинаковую с текущим физическую природу: время и запаздывание по времени (дальность объекта), частоту и доплеровский сдвиг частоты (радиальная скорость объекта), направляющие косинусы и направление прихода волны (угловые координаты объекта). Однако в данном случае для сокращения числа вводимых обозначений в обобщенной функции корреляции (6.14) (а также в частных функциях корреляции по угловым координатам и их производным) текущий и измеряемый параметры обозначены одной буквой v , как и сдвиги v' по ним. Допустимость такого обозначения наглядно видна на примере частной функции корреляции (6.7): в ее подынтегральном выражении $S(t)S^*(t-\tau')$ смещение по измеряемому параметру — времени запаздывания τ' — можно заменить смещением по текущему времени t' .

При наблюдении точечного объекта, представляемого дельта-функцией $\delta(v-v_0)$ текущего параметра v (где v_0 — истинное значение этого параметра), входная и выходная обобщенные харак-

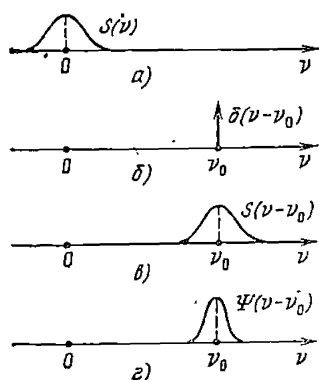


Рис. 6.1. Обобщенная огибающая сигнала (а), импульсная характеристика точечного объекта (б), входной (в) и выходной (г) сигналы согласованного фильтра.

характеристики системы полностью описывают входной и выходной сигналы оптимального приемника в отсутствие шумов (рис. 6.1). Действительно, входная характеристика системы — обобщенная огибающая $S(\nu)$ или ее спектр $S(\mu)$ — описывает входной сигнал приемника $S(\nu - \nu_0)$ или $S(\mu) \exp(-j2\pi\mu\nu_0)$, где спектральное представление входного сигнала получено из $S(\mu)$ по теореме смещения. Соответственно сигнал точечного объекта на выходе оптимального приемника описывается через смещенную на ν_0 выходную характеристику приемника — обобщенную функцию корреляции или энергетический спектр: $\Psi(\nu - \nu_0)$ или $|S(\mu)|^2 \exp(-j2\pi\mu\nu_0)$. Таким образом, входной и выходной сигналы оптимального приемника являются входной и выходной реакцией системы на точечный объект.

Огибающую выходного сигнала точечного объекта $\Psi(\nu - \nu_0)$ можно непосредственно наблюдать на экране индикатора. При этом положение пика огибающей вдоль развертки по ν характеризует истинное значение параметра ν_0 (в отсутствие шумов), а ширина пика, обратная ширине обобщенного спектра $\Delta\mu$, определяет точность и разрешающую способность по данному параметру (при наличии шумов), как будет показано ниже.

Заметим также, что при измерении одновременно двух неизвестных параметров ν_1 и ν_2 выходной сигнал можно наблюдать на экране индикатора в виде пятна, яркость которого пропорциональна мощности выходного сигнала $|\Psi(\nu_1 - \nu_{10}, \nu_2 - \nu_{20})|^2$, причем положение центра пятна указывает истинные значения параметров ν_{10} и ν_{20} , а размеры его, определяемые протяженностью двумерного спектра $\Delta(\mu_1, \mu_2)$, характеризуют совместную точность измерений и разрешающую способность по обоим параметрам. Выходной сигнал как функцию трех и более измеряемых параметров удовлетворительно воспроизвести невозможно.

6.3. ПОТЕНЦИАЛЬНАЯ ТОЧНОСТЬ ИЗМЕРЕНИЙ ФАЗЫ, АМПЛИТУДЫ И ПРОИЗВОЛЬНОГО НЕЭНЕРГЕТИЧЕСКОГО ПАРАМЕТРА СИГНАЛА

Рассмотрим гармоническую обобщенную огибающую, амплитуда которой A много больше среднеквадратического значения шума σ_n (рис. 6.2, а). Здесь и далее предполагается, что измерения проводятся при достаточно большом отношении сигнал-шум,

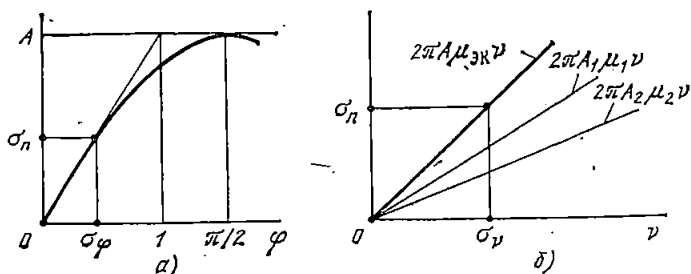


Рис. 6.2. Погрешности измерений фазы гармонической огибающей (а) и произвольного параметра сигнала ν с огибающей в виде суммы двух синусов с частотами μ_1 и μ_2 и эквивалентной им гармонической огибающей с частотой $\mu_{\text{эк}}$ (б)

что справедливо для точных измерений. При этом среднеквадратическая погрешность измерения фазы $\sigma_\varphi = \sigma_n/A$, а дисперсия

$$\sigma_\varphi^2 = \frac{\sigma_n^2}{A^2} = \frac{N_0 \Delta f}{A^2} = \frac{N_0 \Delta f \tau_{\text{н}}}{A^2 \tau_{\text{н}}}, \quad (6.16)$$

где N_0 — спектральная плотность шума; Δf — полоса пропускания приемника; $\tau_{\text{н}}$ — длительность сигнала (импульса).

Учитывая, что $\mathcal{E} = A^2 \tau_{\text{н}}/2$ есть энергия сигнала (на сопротивлении 1 Ом), а в оптимальном приемнике $\Delta f \tau_{\text{н}} = 1$, находим дисперсию погрешности измерения фазы сигнала

$$\sigma_\varphi^2 = \frac{1}{2\mathcal{E}/N_0}. \quad (6.17)$$

Таким образом, потенциальная точность измерения фазы (как и качество обнаружения) определяется только энергетическим отношением сигнал-шум.

Через дисперсию фазы можно определить точность измерения любого параметра сигнала, энергетического (амплитуды) и неэнергетического. Дисперсия колебаний амплитуды есть дисперсия шумов, накладываются на сигнал: $\sigma_A^2 = \sigma_n^2$. Поэтому, воспользовавшись формулами (6.16) и (6.17), находим

$$\sigma_A^2 = A^2 \sigma_\varphi^2 = \frac{A^2}{2\mathcal{E}/N_0}, \quad (6.18)$$

т. е. погрешность измерения амплитуды при заданном отношении сигнал-шум пропорциональна амплитуде сигнала.

Переходя от фазы к произвольному неэнергетическому параметру сигнала (волны) ν , рассмотрим вначале гармоническую обобщенную огибающую $S(\nu)$ частоты μ . Так как текущая фаза огибающей $\varphi = 2\pi\mu\nu$, то среднеквадратическое отклонение фазы $\sigma_\varphi = 2\pi\mu\sigma_\nu$ и дисперсия погрешности измерения неэнергетического параметра сигнала с гармонической огибающей

$$\sigma_\nu^2 = \frac{\sigma_\varphi^2}{(2\pi\mu)^2} = \frac{1}{(2\mathcal{E}/N_0)(2\pi\mu)^2}. \quad (6.19)$$

Таким образом, потенциальная точность измерения неэнергетического параметра кроме отношения сигнал-шум определяется еще частотой гармонической огибающей: чем выше частота, тем выше точность.

Покажем теперь, что огибающая сигнала произвольного вида $S(\nu)$ при данном отношении сигнал-шум эквивалентна гармонической огибающей с частотой $\mu_{\text{эк}}$. Для этого возьмем сначала огибающую в виде суммы двух гармонических составляющих частоты μ_1 и μ_2 и амплитуды A_1 и A_2 соответственно. Полная энергия сигнала $\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 = (A_1^2 + A_2^2) \tau_{\text{ш}} / 2 = A^2 \tau_{\text{ш}} / 2$, где длительность сигнала $\tau_{\text{ш}}$ полагаем достаточно большой. Поэтому с энергетической точки зрения сложную огибающую можно заменить гармонической, амплитуда которой $A = \sqrt{A_1^2 + A_2^2}$. Более того, результат наложения начальных линейных участков двух синусоид частоты μ_1 и μ_2 и амплитуды A_1 и A_2 при совпадении их нулевых фаз можно рассматривать как начальный участок некоторой синусоиды частоты $\mu_{\text{эк}}$ и амплитуды $A = \sqrt{A_1^2 + A_2^2}$ (рис. 6.2,б). Действительно, линейные участки синусоид в равной степени определяются их частотой и амплитудой:

$$\begin{aligned} A_1 \sin 2\pi\mu_1\nu &\approx 2\pi\mu_1 A_1\nu, & A_2 \sin 2\pi\mu_2\nu &\approx 2\pi\mu_2 A_2\nu, \\ A \sin 2\pi\mu_{\text{эк}}\nu &\approx 2\pi\mu_{\text{эк}} A\nu. \end{aligned} \quad (6.20)$$

Суммируя текущие мощности этих участков, получаем

$$(2\pi\mu_{\text{эк}} A\nu)^2 = (2\pi\mu_1 A_1\nu)^2 + (2\pi\mu_2 A_2\nu)^2$$

или после сокращений

$$\mu_{\text{эк}}^2 (A_1^2 + A_2^2) = \mu_1^2 A_1^2 + \mu_2^2 A_2^2,$$

так как $A^2 = A_1^2 + A_2^2$. Отсюда находим частоту эквивалентной гармонической огибающей такого сигнала

$$\mu_{\text{эк}}^2 = (\mu_1^2 A_1^2 + \mu_2^2 A_2^2) / (A_1^2 + A_2^2). \quad (6.21)$$

Распространяя этот результат на любое число дискретных гармоник огибающей, получаем

$$\mu_{\text{эк}}^2 = \sum \mu_i^2 A_i^2 / \sum A_i^2. \quad (6.22)$$

Если обобщенная огибающая произвольной формы имеет сплошной энергетический спектр $|S(\mu)|^2$, симметричный относительно нулевой частоты, то, разбив его на элементарные гармоники, квадрат амплитуды которых $A_i^2 = |S(\mu)|^2 d\mu$, и заменив суммы интегралами, вместо формулы (6.22) получим

$$\mu_{\text{эк}}^2 = \frac{\int_{-\infty}^{\infty} \mu^2 |S(\mu)|^2 d\mu}{\int_{-\infty}^{\infty} |S(\mu)|^2 d\mu}. \quad (6.23)$$

Таким образом, огибающую произвольной формы с симметричным энергетическим спектром при заданном отношении сигнал-шум можно заменить согласно формуле (6.23) гармонической огибающей частоты $\mu = \mu_{\text{эк}}$, обеспечивающей такую же потенциальную

точность измерения неэнергетического параметра ν , что и реальная огибающая. Поэтому для оценки потенциальной точности измерений при комплексной огибающей с произвольным, но симметричным спектром можно воспользоваться формулой (6.19), подставив в нее $\mu_{\text{ЭК}}$ вместо μ :

$$\sigma_{\nu}^2 = \frac{1}{(2\partial/N_0)(2\pi\mu_{\text{ЭК}})^2} \quad (6.24)$$

Как следует из формул (6.23) и (6.24), частота $\mu_{\text{ЭК}}$ и потенциальная точность измерений определяются только энергетическим спектром огибающей, а не ее формой, зависящей от фазовой структуры спектра. Однако, поскольку обобщенная функция корреляции (6.14) взаимно однозначно связана с энергетическим спектром сигнала, существует связь между потенциальной точностью измерений и обобщенной функцией корреляции сигнала $\Psi(\nu')$. Действительно, продифференцируем дважды функцию корреляции (6.14) по параметру ν' в точке максимума:

$$\begin{aligned} a &= -\frac{d^2}{d\nu'^2} \Psi(\nu')_{\nu'=0} = \frac{\int |dS(\nu)/d\nu'|^2 d\nu}{\int |S(\nu)|^2 d\nu} = \\ &= (2\pi)^2 \frac{\int \mu^2 |S(\mu)|^2 d\mu}{\int |S(\mu)|^2 d\mu} = (2\pi\mu_{\text{ЭК}})^2. \end{aligned} \quad (6.25)$$

Здесь при двойном дифференцировании левой части выражения (6.14) производная под интегралом числителя бралась сначала по ν' от первого сомножителя, затем по $-\nu'$ от второго сомножителя, что допустимо, так как сам интеграл зависит только от взаимного сдвига ν' этих сомножителей. Сопоставив формулы (6.24) и (6.25), получим искомую связь между дисперсией погрешности измерений и ее обобщенной функцией корреляции:

$$\sigma_{\nu}^2 = \frac{1}{2\partial/N_0} \frac{1}{a} = -\frac{1}{(2\partial/N_0) [d^2 \Psi(\nu')/d\nu'^2]_{\nu'=0}} \quad (6.26)$$

Эта формула, справедливая только для огибающих с симметричным энергетическим спектром, позволяет перейти к общему случаю, когда спектр не обязательно симметричен.

Рассмотрим свойства модуля обобщенной функции корреляции $|\Psi(\nu')|$ и квадрата ее модуля $|\Psi(\nu')|^2$. Поскольку энергетический спектр $|S(\mu)|^2$ — вещественная функция, связанная с функцией корреляции $\Psi(\nu')$ преобразованием Фурье, то квадрату модуля $|\Psi(\nu')|^2$ как произведению комплексно-сопряженных функций соответствует свертка одинаковых энергетических спектров. Но (как было показано в § 3.2) свертка (функция корреляции) одинаковых вещественных функций всегда вещественная и симметричная. Поэтому и связанная со сверткой спектров преобразованием Фурье функция $|\Psi(\nu')|^2$ также вещественная и симметричная, как и модуль функции корреляции $|\Psi(\nu')|$. Следовательно, функции $|\Psi(\nu')|$ и $|\Psi(\nu')|^2$ обладают симметричными спек-

рами и к ним применимы все выводы, полученные для функции $\Psi(v')$ с симметричным энергетическим спектром. В частности, подставив $|\Psi(v')|$ в формулу (6.25) вместо $\Psi(v')$, получим:

$$a = -\frac{d^2}{dv'^2} |\Psi(v')|_{v'=0}$$

или

$$\begin{aligned} \frac{d^2}{dv'^2} |\Psi(v')|^2_{v'=0} &= 2 \left[|\Psi(v')| \frac{d^2}{dv'^2} |\Psi(v')| \right]_{v'=0} + \\ &+ 2 \left[\frac{d}{dv'} |\Psi(v')| \right]_{v'=0}^2 = 2 \frac{d^2}{dv'^2} |\Psi(v')|_{v'=0} = -2a, \end{aligned} \quad (6.27)$$

так как в максимуме $|\Psi(v')|_{v'=0} = 1$ и $\left[\frac{d}{dv'} |\Psi(v')| \right]_{v'=0} = 0$. Учитывая, что $|\Psi(v')|^2 = \Psi(v') \Psi^*(v')$, и переходя к сокращенной записи функции и ее производных, получим

$$\begin{aligned} \frac{d}{dv'} |\Psi(v')|^2 &= \Psi' \Psi^* + \Psi \Psi^{*'}, \\ \frac{d^2}{dv'^2} |\Psi(v')|^2 &= \Psi'' \Psi^* + \Psi' \Psi^{*'} + \Psi' \Psi^{*'} + \Psi \Psi^{*''}. \end{aligned} \quad (6.28)$$

При однократном дифференцировании функция $\Psi(v')$ согласно правой части формулы (6.14) выдает множитель $j2\pi\mu$, а функция $\Psi^*(v')$ множитель $-j2\pi\mu$, поэтому $\Psi^{*'}(0) = -\Psi'(0)$ и $\Psi^{*''}(0) = \Psi''(0)$, причем $\Psi(0) = \Psi^*(0) = 1$. Возвращаясь к полной записи, из соотношений (6.27) и (6.28) с учетом (6.25) находим

$$\begin{aligned} a &= -\frac{d^2}{dv'^2} |\Psi(v')|_{v'=0} = -\frac{1}{2} \frac{d^2}{dv'^2} |\Psi(v')|^2_{v'=0} = \\ &= -\left\{ \frac{d^2}{dv'^2} \Psi(v') - \left[\frac{d}{dv'} \Psi(v') \right]^2 \right\}_{v'=0} = (2\pi\mu_{\text{эк}})^2. \end{aligned} \quad (6.29)$$

Выполнив дифференцирование обобщенной функции корреляции (6.14) в соответствии с формулой (6.29), получим соотношение для частоты эквивалентной гармонической огибающей в самом общем случае

$$\mu_{\text{эк}}^2 = \frac{\int_{-\infty}^{\infty} |dS(v)/dv'|^2 dv}{(2\pi)^2 \int_{-\infty}^{\infty} |S(v)|^2 dv} = \mu_0^2 = \frac{\int_{-\infty}^{\infty} \mu^2 |S(\mu)|^2 d\mu}{\int_{-\infty}^{\infty} |S(\mu)|^2 d\mu} \quad (6.30)$$

где

$$\mu_0 = \frac{\int_{-\infty}^{\infty} [dS(v)/dv'] S^*(v) dv}{2\pi \int_{-\infty}^{\infty} |S(v)|^2 dv} = \frac{\int_{-\infty}^{\infty} \mu |S(\mu)|^2 d\mu}{\int_{-\infty}^{\infty} |S(\mu)|^2 d\mu} \quad (6.31)$$

— первый начальный момент (смещение центра тяжести) энергетического спектра огибающей $|S(\mu)|^2$, а $\mu_{\text{ЭК}}^2$ является его вторым центральным моментом (моментом инерции). Таким образом, значение $\mu_{\text{ЭК}}^2$, пропорциональное абсолютному значению второй производной модуля функции корреляции огибающей выходного сигнала оптимального приемника, характеризует ширину энергетического спектра и, следовательно, остроту пика функции корреляции, т. е. скорость спадания выходного сигнала приемника при расстройке по измеряемому параметру. Простое (без изменения формы) смещение спектра μ_0 вдоль оси частот не влияет на его ширину и, следовательно, на значение $\mu_{\text{ЭК}}$ и точность измерений.

По аналогии с формулой (6.26) можно получить равноценные формулы для оценки потенциальной точности измерений произвольного параметра сигнала при любой форме спектра

$$\begin{aligned} \sigma_v^2 &= \frac{1}{\frac{2\mathcal{E}}{N_0}} \cdot \frac{1}{a} = \frac{1}{\frac{2\mathcal{E}}{N_0} (2\pi\mu_{\text{ЭК}})^2} = \frac{1}{\frac{\mathcal{E}}{N_0} \frac{d^2}{dv'^2} |\Psi(v')|_{v'=0}^2} = \\ &= \frac{1}{\frac{2\mathcal{E}}{N_0} \frac{d^2}{dv'^2} |\Psi(v')|_{v'=0}} = \frac{1}{\frac{2}{N_0} \frac{d^2}{dv'^2} Z_s(v')_{v'=0}} = \\ &= \frac{1}{\frac{d^2}{dv'^2} l_s(v')_{v'=0}}, \end{aligned} \quad (6.32)$$

где $Z_s(v') = \mathcal{E} |\Psi(v')|$ и $l_s(v') = (2\mathcal{E}/N_0) |\Psi(v')|$ — сигнальные составляющие корреляционного интеграла и логарифма отношения правдоподобия соответственно.

Если энергетический спектр огибающей симметричен относительно нулевой частоты (как это обычно бывает), $\mu_0 = 0$ и формулы (6.23) и (6.30) полностью совпадают, так же как и формулы (6.26) и (6.32).

Заменяя в формуле (6.32) произвольный неэнергетический параметр v конкретными параметрами τ , F , $u_{x,y}$, $\dot{u}_{x,y}$, находим дисперсии оценки дальности, радиальной скорости, направляющих косинусов и их первых производных соответственно:

$$\begin{aligned} \sigma_\tau^2 &= 1 / \left(\frac{2\mathcal{E}}{N_0} \right) (2\pi f_{\text{ЭК}})^2; & \sigma_F^2 &= 1 / \left(\frac{2\mathcal{E}}{N_0} \right) (2\pi t_{\text{ЭК}})^2; \\ \sigma_{u_{x,y}}^2 &= 1 / \left(\frac{2\mathcal{E}}{N_0} \right) (2\pi u_{x,y \text{ ЭК}})^2; & \sigma_{\dot{u}_{x,y}}^2 &= 1 / \left(\frac{2\mathcal{E}}{N_0} \right) (2\pi u_{x,y \text{ ЭК}} t_{\text{ЭК}})^2. \end{aligned} \quad (6.33)$$

Здесь конкретизированы понятия эквивалентной частоты обобщенной огибающей $\mu_{\text{ЭК}}$: эквивалентная частота огибающей сигнала $f_{\text{ЭК}}$, эквивалентное время отсчета $t_{\text{ЭК}}$ и эквивалентные про-

странственные частоты $v_{x \text{ эк}}$ и $v_{y \text{ эк}}$. В соответствии с общей формулой (6.28) и табл. 6.1 при $\mu_0 = 0$ имеем

$$\begin{aligned} f_{\text{эк}}^2 &= \frac{\int f^2 |S(f)|^2 df}{\int |S(f)|^2 df} = \frac{\int |dS(t)/dt|^2 dt}{(2\pi)^2 \int |S(t)|^2 dt}, \\ t_{\text{эк}}^2 &= \frac{\int t^2 |S(t)|^2 dt}{\int |S(t)|^2 dt} = \frac{\int |dS(f)/df|^2 df}{(2\pi)^2 \int |S(f)|^2 df}, \\ v_{x,y \text{ эк}}^2 &= \frac{\int v_{x,y}^2 |G(v_{x,y})|^2 dv_{x,y}}{\int |G(v_{x,y})|^2 dv_{x,y}} = \frac{\int |dG(u_{x,y})/du_{x,y}|^2 du_{x,y}}{(2\pi)^2 \int |G(u_{x,y})|^2 du_{x,y}}, \end{aligned} \quad (6.34)$$

где опущены бесконечные пределы.

Как следует из формул (6.30), дисперсия оценки производных направляющих косинусов (угловых скоростей) определяется произведением параметров $v_{x,y \text{ эк}}$ и $t_{\text{эк}}$, которые порознь характеризуют точность измерения самих углов и радиальной скорости соответственно. Поэтому при той же антенне и том же сигнале нет необходимости наряду с угловыми координатами измерять их производные: производные не содержат новой информации и могут быть получены простым дифференцированием измеренных текущих угловых координат. Однако измерение угловых скоростей, если оно производится при значительно больших величинах $v_{x,y \text{ эк}} = f_0 d_{\text{эк}}/c$, чем измерение углов, т. е. либо при другой антенне (увеличенных размеров $d_{\text{эк}}$), либо при другом сигнале (с более высокой частотой несущей f_0) дает новую, независимую от измеренных угловых координат информацию. Ввиду этого измерительные системы с заданным видом сигнала и апертуры рассчитаны на измерение только четырех параметров: дальности, радиальной скорости и двух угловых координат. Для измерения двух составляющих угловой скорости необходима дополнительная система.

В заключение в качестве примера вычислим частоту эквивалентной гармонической огибающей по формуле (6.30) для трех видов симметричного спектра:

равномерного

$$\begin{aligned} S(\mu) &= \begin{cases} 1, & |\mu| \leq \Delta\mu/2, \\ 0, & |\mu| > \Delta\mu/2, \end{cases} \\ \mu_{\text{эк}}^2 &= \frac{1}{[\Delta\mu]} \int_{-\Delta\mu/2}^{\Delta\mu/2} \mu^2 d\mu = \frac{\Delta\mu^2}{12}, \quad \mu_{\text{эк}} = \frac{\Delta\mu}{2\sqrt{3}}, \end{aligned}$$

параболического

$$S(\mu) = \begin{cases} 1 - (2\mu/\Delta\mu)^2, & |\mu| \leq \Delta\mu/2, \\ 0, & |\mu| > \Delta\mu/2; \end{cases}$$

$$\mu_{\text{ЭК}}^2 = \frac{\int_{-\Delta\mu/2}^{\Delta\mu/2} \mu^2 \left[1 - \left(\frac{2\mu}{\Delta\mu} \right)^2 \right]^2 d\mu}{\int_{-\Delta\mu/2}^{\Delta\mu/2} \left[1 - \left(\frac{2\mu}{\Delta\mu} \right)^2 \right]^2 d\mu} = \frac{\Delta\mu^2}{40}, \quad \mu_{\text{ЭК}} = \frac{\Delta\mu}{2\sqrt{10}};$$

гауссовского

$$S(\mu) = \exp(-\pi\mu^2/\Delta\mu^2);$$

$$\mu_{\text{ЭК}}^2 = \frac{\int_{-\infty}^{\infty} \mu^2 \exp\left(-\frac{2\pi\mu^2}{\Delta\mu^2}\right) d\mu}{\int_{-\infty}^{\infty} \exp\left(-\frac{2\pi\mu^2}{\Delta\mu^2}\right) d\mu} = \frac{\Delta\mu^2}{4\pi}, \quad \mu_{\text{ЭК}} = \frac{\Delta\mu}{2\sqrt{\pi}}.$$

В двух первых случаях ширина спектра $\Delta\mu$ отсчитывается на нулевом уровне, в третьем — на уровне $\exp(-\pi/4) \approx 0,46$. Последний результат можно получить, не прибегая к вычислению табличных интегралов, из известного для среднего квадрата (дисперсии) $\mu_{\text{ЭК}}^2$ представления гауссовской кривой $\exp(-2\pi\mu^2/\Delta\mu^2) = \exp(-\mu^2/2\mu_{\text{ЭК}}^2)$, подставив в это равенство $\Delta\mu/2$ вместо μ . Чтобы найти конкретные параметры $f_{\text{ЭК}}, t_{\text{ЭК}}, v_{x,y \text{ЭК}}$, в полученные результаты соответственно подставим $\Delta f, \Delta t, \Delta v_{x,y} = d/\lambda$.

6.4. ОПТИМАЛЬНАЯ ПО ТОЧНОСТИ. ФОРМА ОГИБАЮЩЕЙ

Как следует из полученных формул, при заданном отношении сигнал-шум S/N_0 потенциальная точность измерений произвольного параметра ν определяется частотой эквивалентной гармонической огибающей $\mu_{\text{ЭК}}$, которая зависит лишь от структуры обобщенного энергетического спектра огибающей $|S(\mu)|^2$. Поэтому необходимо оценить влияние параметров энергетического спектра на точность измерений и определить оптимальную по точности форму спектра, а следовательно, и самой огибающей.

В первую очередь оценим влияние смещения симметричного спектра относительно нулевой частоты. Для этого сместим симметричный относительно $\mu=0$ энергетический спектр $|S(\mu)|^2$ по оси частот на μ_0 и найдем второй начальный момент смещенного спектра

$$\mu_{00}^2 = \int \mu^2 |S(\mu - \mu_0)|^2 d\mu / \int |S(\mu - \mu_0)|^2 d\mu$$

(здесь и далее пределы интегрирования полагаем бесконечными). Произведем замену переменной: $\mu' = \mu - \mu_0$. Тогда

$$\begin{aligned} \mu_{00}^2 &= \frac{\int (\mu' + \mu_0)^2 |S(\mu')|^2 d\mu'}{\int |S(\mu')|^2 d\mu'} = \frac{\int \mu'^2 |S(\mu')|^2 d\mu'}{\int |S(\mu')|^2 d\mu'} + \\ &+ 2\mu_0 \frac{\int \mu' |S(\mu')|^2 d\mu'}{\int |S(\mu')|^2 d\mu'} + \mu_0^2. \end{aligned}$$

Здесь интеграл во втором слагаемом равен нулю, как первый момент функции $|S(\mu')|^2$, симметричной относительно $\mu' = 0$. Отсюда (отбросив штрих у переменной) находим

$$\mu_{\text{эк}}^2 = \mu_{00}^2 - \mu_0^2 = \int \mu^2 |S(\mu)|^2 d\mu / \int |S(\mu)|^2 d\mu. \quad (6.35)$$

Таким образом, смещение симметричного энергетического спектра огибающей относительно нулевой частоты не влияет на точность измерений: формула (6.35) совпадает с формулой (6.23) для несмещенного спектра.

Чтобы оценить влияние асимметрии энергетического спектра на точность измерений, разложим его на четную и нечетную составляющие:

$$|S(\mu)|^2 = |S(\mu)|_{\text{ч}}^2 + |S(\mu)|_{\text{нч}}^2, \quad (6.36)$$

где четная и нечетная составляющие находятся как полусумма и полуразность исходного спектра, обращенного относительно $\mu = 0$:

$$\begin{aligned} |S(\mu)|_{\text{ч}}^2 &= 0,5[|S(\mu)|^2 + |S(-\mu)|^2], \\ |S(\mu)|_{\text{нч}}^2 &= 0,5[|S(\mu)|^2 - |S(-\mu)|^2] \end{aligned} \quad (6.37)$$

(рис. 6.3). При этом

$$\int_{-\infty}^{\infty} |S(\mu)|_{\text{ч}}^2 d\mu = \int_{-\infty}^{\infty} |S(\mu)|^2 d\mu, \quad \int_{-\infty}^{\infty} |S(\mu)|_{\text{нч}}^2 d\mu = 0, \quad (6.38)$$

т. е. четная составляющая содержит всю энергию огибающей, а нечетная (введенная чисто формально) не несет энергии (энергия не может быть отрицательной). Таким образом, асимметрию энергетического спектра можно трактовать как появление дополнительной нечетной составляющей, не изменяющей полной энергии сигнала. Асимметрия энергетического спектра проявляется в

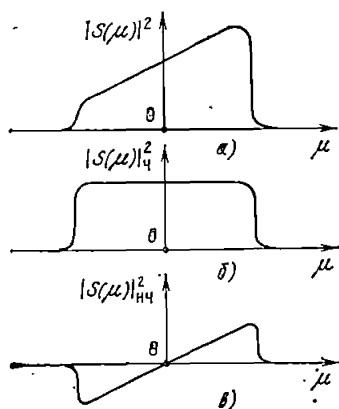
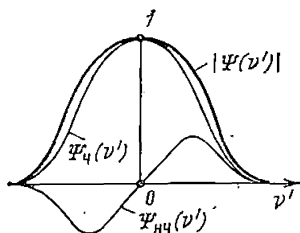


Рис. 6.3. Энергетический спектр сигнала (а), его четная (б) и нечетная (в) составляющие

Рис. 6.4. Расширение огибающей обобщенной функции корреляции за счет составляющей $\Psi_{\text{нч}}(\nu')$ при асимметрии энергетического спектра



том, что обобщенная функция корреляции принимает комплексный характер

$$\begin{aligned}\Psi(v') &= \int_{-\infty}^{\infty} |S(\mu)|^2 e^{j 2\pi\mu v'} d\mu = \\ &= \int_{-\infty}^{\infty} |S(\mu)|_c^2 e^{j 2\pi\mu v'} d\mu + \int_{-\infty}^{\infty} |S(\mu)|_{nc}^2 e^{j 2\pi\mu v'} d\mu = \\ &= \Psi_c(v') + j \Psi_{nc}(v'),\end{aligned}\quad (6.39)$$

где

$$\Psi_c(v') = 2 \int_0^{\infty} |S(\mu)|_c^2 \cos 2\pi\mu v' d\mu, \quad (6.40)$$

$$\Psi_{nc}(v') = 2 \int_0^{\infty} |S(\mu)|_{nc}^2 \sin 2\pi\mu v' d\mu$$

— четная и нечетная составляющие комплексной функции корреляции соответственно. Огибающая функции корреляции

$$|\Psi(v')| = \sqrt{\Psi_c^2(v') + \Psi_{nc}^2(v')} \quad (6.41)$$

при асимметрии расширяется за счет появления нечетной составляющей $\Psi_{nc}(v')$ по сравнению с огибающей с симметричным спектром, у которой $|\Psi(v')| = \Psi(v') = \Psi_c(v')$ (рис. 6.4). В результате точность измерений снижается. Заметим, что огибающая функции корреляции $|\Psi(v')|$ всегда симметрична, так как в формуле (6.41) функция $\Psi_{nc}^2(v')$ в отличие от $\Psi_{nc}(v')$ четная (это справедливо для любых функций, имеющих вещественный спектр).

Чтобы определить степень снижения точности вследствие асимметрии энергетического спектра, найдем его второй начальный момент

$$\mu_{00}^2 = \frac{\int \mu^2 |S(\mu)|^2 d\mu}{\int |S(\mu)|^2 d\mu} = \frac{\int \mu^2 |S(\mu)|_c^2 d\mu}{\int |S(\mu)|_c^2 d\mu}. \quad (6.42)$$

Этот момент определяется только четной составляющей спектра, так как нечетная составляющая при интегрировании по области положительных и отрицательных частот дает нуль. Таким образом, асимметрия спектра не изменяет его второго начального момента, однако дает ненулевой первый момент

$$\mu_0 = \frac{\int \mu |S(\mu)|^2 d\mu}{\int |S(\mu)|^2 d\mu} = \frac{\int \mu |S(\mu)|_{nc}^2 d\mu}{\int |S(\mu)|_c^2 d\mu}, \quad (6.43)$$

который обусловлен только нечетной составляющей спектра, так как произведение $\mu |S(\mu)|_c^2$ нечетно и при интегрировании по всей области частот дает нуль.

Острота пика огибающей функции корреляции $|\Psi(v')|$, определяющая точность измерений, зависит от ширины спектра этой функции, т. е. его второго центрального момента

$$\mu_{2\kappa}^2 = \mu_{00}^2 - \mu_0^2, \quad (6.44)$$

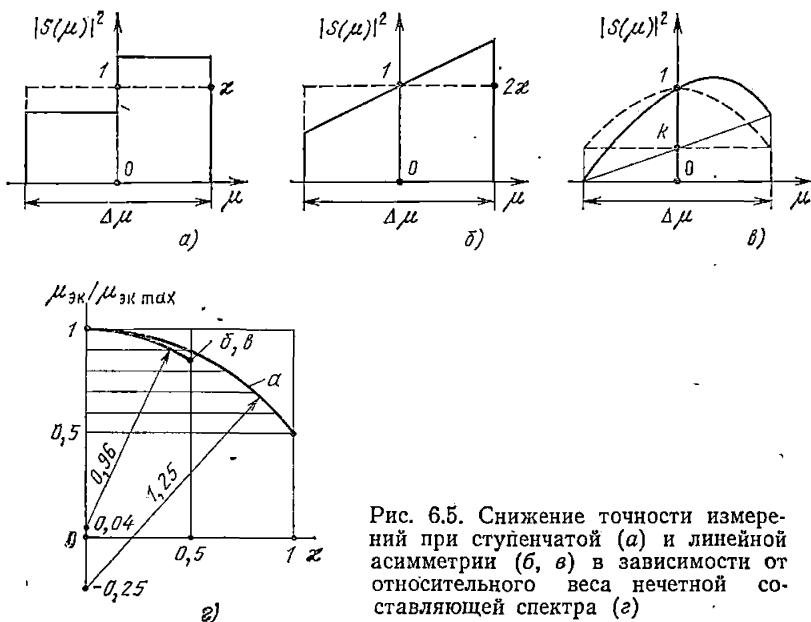


Рис. 6.5. Снижение точности измерений при ступенчатой (а) и линейной асимметрии (б, в) в зависимости от относительного веса нечетной составляющей спектра (г)

что совпадает с формулой (6.30) с учетом (6.31). При этом первое слагаемое формулы (6.44) обусловлено только четной составляющей энергетического спектра, второе только нечетной. Поскольку первое слагаемое не зависит от асимметрии спектра; снижение точности измерений при асимметрии определяется только вторым слагаемым.

Как следует из формулы (6.43), снижение точности при асимметрии за счет появления первого момента спектра μ_0 зависит от удельного веса площади $|S(\mu)|^2_{нч}$ по сравнению с полной площадью, определяемой $|S(\mu)|^2$. Но так как у $|S\mu|^2_{нч}$ одна половина положительна, а другая отрицательна, введем в качестве весового коэффициента отношение полуплощадей

$$\kappa = \frac{\int_0^{\infty} |S(\mu)|^2_{нч} d\mu}{\int_0^{\infty} |S(\mu)|^2 d\mu} = \frac{\int_0^{\infty} |S(\mu)|^2_{нч} d\mu}{\int_0^{\infty} |S(\mu)|^2_{ч} d\mu}. \quad (6.45)$$

К сожалению, количественная оценка ухудшения точности в общем виде затруднена тем, что она зависит не только от коэффициента κ , но и от формы спектра. Поэтому рассмотрим три частных примера (рис. 6.5):

$$а) |S(\mu)|^2 = \begin{cases} 1 + \kappa, & 0 < \mu \leq \Delta\mu/2, \\ 1 - \kappa, & -\Delta\mu/2 \leq \mu < 0, \end{cases} \quad 0 \leq \kappa \leq 1;$$

$$б) |S(\mu)|^2 = 1 + 4\kappa\mu/\Delta\mu, \quad |\mu| \leq \Delta\mu/2, \quad 0 \leq \kappa \leq 1/2;$$

$$в) |S(\mu)|^2 = 1 - 4(1 - \kappa)\mu^2/\Delta\mu^2 + 2\kappa\mu/\Delta\mu, \\ |\mu| \leq \Delta\mu/2, \quad 0 \leq \kappa \leq 1, \quad 0 \leq \kappa \leq 3/(4\kappa + 2),$$

где $\Delta\mu$ — общая ширина спектра. В первом примере нечетная составляющая ступенчатая, во втором и третьем — плавная, линейная и нелинейная. В последнем примере введена постоянная составляющая k , чтобы не нарушить условие $|S(\mu)|^2 \geq 0$.

В результате вычислений получаем (рис. 6.5,2)

$$a) \mu_{00}^2 = \frac{\Delta\mu^2}{12}, \quad \mu_0 = \frac{\kappa\Delta\mu}{4}, \quad \mu_{\text{эк}} = \frac{\Delta\mu}{2\sqrt{3}} \sqrt{1 - \frac{3\kappa^2}{4}},$$

$$\mu_{\text{эк max}} = \frac{\Delta\mu}{2\sqrt{3}} \quad \text{при } \kappa = 0, \quad \mu_{\text{эк min}} = \frac{\Delta\mu}{4\sqrt{3}} \quad \text{при } \kappa = 1;$$

$$б) \mu_{00}^2 = \frac{\Delta\mu^2}{12}, \quad \mu_0 = \frac{\kappa\Delta\mu}{3}, \quad \mu_{\text{эк}} = \frac{\Delta\mu}{2\sqrt{3}} \sqrt{1 - \frac{4\kappa^2}{3}},$$

$$\mu_{\text{эк max}} = \frac{\Delta\mu}{2\sqrt{3}} \quad \text{при } \kappa = 0, \quad \mu_{\text{эк min}} = \frac{\Delta\mu}{3\sqrt{2}} \quad \text{при } \kappa = 0,5;$$

$$в) \mu_{00}^2 = \frac{(2+3k)\Delta\mu^2}{20(2+k)} = \frac{(3-8\kappa)\Delta\mu^2}{60}, \quad \mu_0 = \frac{k\Delta\mu}{4+2k} = \frac{\kappa\Delta\mu}{3},$$

$$\mu_{\text{эк}} = \frac{\Delta\mu}{2\sqrt{5}} \sqrt{1 + \frac{4\kappa}{3} - \frac{20\kappa^2}{9}},$$

$$\mu_{\text{эк max}} = \Delta\mu/2\sqrt{5} \quad \text{при } \kappa = 0; \quad \mu_{\text{эк min}} = \Delta\mu/3 \quad \text{при } \kappa = 0,5.$$

Полученные результаты показывают, что точность измерений наиболее существенно может снижаться при ступенчатой асимметрии (случай *а*). В пределе при $\kappa=1$, когда полностью подавляется одна половина спектра, точность снижается вдвое. Более близкой к реальной является линейная асимметрия, которая оказывает практически одинаковое воздействие как на равномерный, так и на плавно спадающий на концах спектр (случаи *б* и *в*). Хотя линейная асимметрия при заданном κ проявляется несколько сильнее, максимальное снижение точности не превышает 18%, что вызвано ограничением удельного веса нечетной составляющей спектра величиной $\kappa=0,5$, вытекающим из условия $|S(\mu)|^2 \geq 0$. Рассмотренные примеры показывают также, что снижение точности при асимметрии спектра определяется в основном относительным весом κ нечетной составляющей и слабо зависит от формы спектра.

Таким образом, асимметрия энергетического спектра приводит к снижению точности измерений. Однако практически это снижение невелико, так как нечетная составляющая ограничена условием неотрицательности энергетического спектра и дает на выходе измерительной системы нечетную составляющую функции корреляции, которая, складываясь в квадратуре с большой четной составляющей согласно формуле (6.41), вызывает лишь незначительное расширение огибающей выходного сигнала и соответствующее снижение точности.

Исходя из приведенных выше соображений оптимальную по точности обобщенную огибающую при заданном отношении сигнал-шум будем искать в классе функций с симметричным энергетическим спектром (рис. 6.6). Искомая огибающая должна обеспечить максимальное значение $\mu_{\text{эк}}$ при условии, что ширина спектра $\Delta\mu = 2\mu_{\text{max}}$ задана, так как в противном случае с увели-

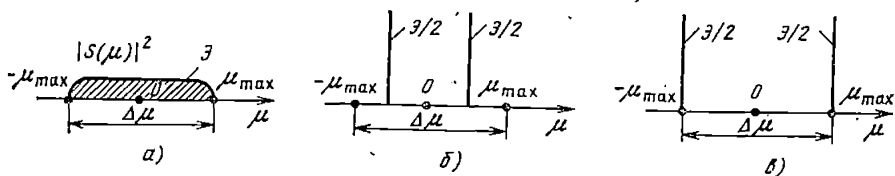


Рис. 6.6. Энергетический спектр обобщенной огибающей сигнала произвольной формы (а), его эквивалентной гармонической огибающей (б) и огибающей оптимального по точности сигнала (в)

чением $\Delta\mu$ параметр $\mu_{\text{эк}}$ будет неограниченно возрастать. Нетрудно видеть, что при заданном $\Delta\mu$ максимально возможное $\mu_{\text{эк max}} = \mu_{\text{max}} = \Delta\mu/2$. Таким образом, оптимальной огибающей, обеспечивающей максимум $\mu_{\text{эк}}$, является гармоническая огибающая частоты $\mu = \mu_{\text{max}} = \Delta\mu/2$. Так, при измерении дальности наилучшую точность обеспечивает сигнал в виде двух высокочастотных синусоид, разнесенных по частоте на Δf ($f_{\text{эк}} = \Delta f/2$), при измерении радиальной скорости — сигнал в виде двух коротких высокочастотных импульсов, разделенных интервалом Δt ($t_{\text{эк}} = \Delta t/2$).

Заметим, однако, что полученный вывод относительно оптимальной формы обобщенной огибающей справедлив, только если от вида ее энергетического спектра не зависит энергия сигнала, которая входит в качестве второго множителя в формулу потенциальной точности (6.24). Это условие не выполняется для угломерных систем, у которых от вида пространственного спектра $G(v_{x,y})$ зависит эффективная площадь антенны, являющаяся энергетическим параметром. Поэтому вопрос о выборе оптимального по точности вида функции $G(v_{x,y})$ требует дополнительного уточнения. Возможны два случая.

1. Заданы общий размер антенны L (или ширина пространственного спектра $\Delta v_{x,y} = L/\lambda$) и суммарная протяженность ее сплошного раскрыва d , причем $L \gg d$. Так как раскрывом d определяется энергия принимаемого сигнала при заданной плотности потока мощности в точке приема, условие $d = \text{const}$ означает $\mathcal{E} = \text{const}$, что соответствует общему условию при нахождении оптимальных характеристик. При этом максимальная эквивалентная пространственная частота $v_{x,y \text{ эк}} \approx \Delta v_{x,y}/2 = L/2\lambda$, а оптимальная антенна имеет вид двух полотен размером $d/2$, разнесенных на максимально возможное расстояние $b \approx L$ (рис. 6.7, а).

2. Задан только размер антенны L (и плотность потока мощности в точке приема). При этом оптимальна антенна, у которой максимально произведение $\mathcal{E} v_{x,y \text{ эк}} = d v_{x,y \text{ эк}}$, что достигается при $d = L$ (рис. 6.7, б). В этом случае эквивалентная пространственная частота $v_{x,y \text{ эк}} = \Delta v_{x,y}/2\sqrt{3}$, т. е. в $\sqrt{3}$ раз меньше, зато энергия принимаемого сигнала в L/d раз больше, чем в первом случае, когда $L \gg d$.

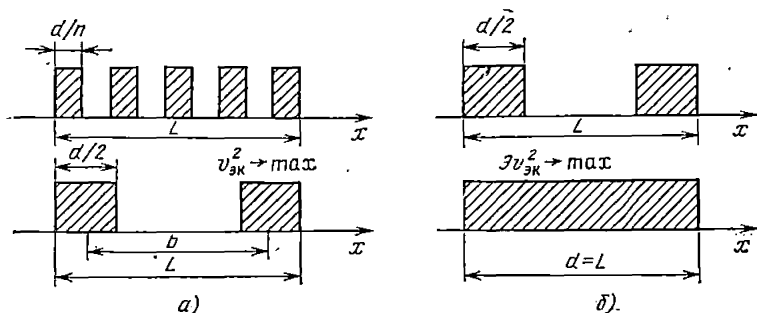


Рис. 6.7. Выбор оптимальной по точности функции раскрыва при заданной энергии принимаемого сигнала (а) и при заданной плотности потока мощности (б)

Однако качество измерений характеризуется не только точностью, определяемой σ_v или $\mu_{\text{эк}}$, но и однозначностью измерений. В частности, оптимальная по точности гармоническая огибающая частоты $\mu = \mu_{\text{max}}$ не обеспечивает однозначности измерений в пределах заданного диапазона $[v_{\text{max}} - v_{\text{min}}]$. Действительно, для обеспечения заданной точности измерений частота гармонической огибающей (обозначим ее μ_t) должна быть достаточно большой. При этом измерение параметра v производится по сдвигу фазы огибающей $\phi = 2\pi\mu_tv$, который при большой частоте μ_t очень быстро набегает с увеличением v , достигая значения 2π и многократно переходя его. Поскольку $\phi > 2\pi$ измерить невозможно, появляется неоднозначность. Положив $\phi = 2\pi$, найдем диапазон однозначности точных измерений

$$v_{\text{одн}} = 1/\mu_t. \quad (6.46)$$

Так как частота μ_t велика, он мал.

Для однозначного измерения во всем диапазоне возможных значений измеряемого параметра $[v_{\text{max}} - v_{\text{min}}]$ частоту огибающей (обозначим ее $\mu_{\text{гр}}$) нужно взять достаточно малой. Из условия $2\pi\mu_{\text{гр}}[v_{\text{max}} - v_{\text{min}}] = 2\pi$ находим частоту

$$\mu_{\text{гр}} = 1/[v_{\text{max}} - v_{\text{min}}], \quad (6.47)$$

обеспечивающую однозначность измерений. Так как частота мала, измерения получаются грубыми.

Отсюда следует, что требованиям точности и однозначности невозможно удовлетворить на одной частоте огибающей. Поэтому приходится использовать две гармонические огибающие на частотах μ_t и $\mu_{\text{гр}}$, образующих точную неоднозначную и грубую, но однозначную шкалы соответственно. Однако частоты μ_t и $\mu_{\text{гр}}$ обычно настолько резко различаются, что по огибающей, являющейся их суммой, не удастся однозначно выбрать максимальный пик точности шкалы, по которому производится отсчет (рис. 6.8). За правильный можно принять один из соседних пиков высокой частоты μ_t , так как огибающая пиков убывает очень

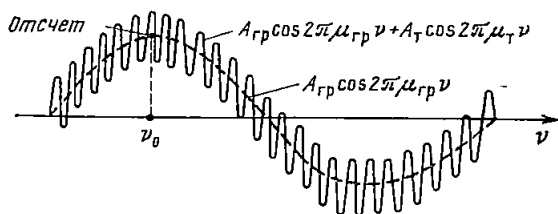


Рис. 6.8. Огибающая функции корреляции (выходного сигнала согласованного приемника) в виде суммы двух гармоник, частоты которых отличаются во много раз

медленно. Для того чтобы убывание происходило быстрее и позволяло произвести безошибочный отсчет, дополнительно вводят среднюю шкалу на частоте $\mu_{ср}$, определяемую из условия $\mu_{т}/\mu_{ср} = \mu_{ср}/\mu_{гр}$ (иногда несколько средних шкал).

Многошкальная система позволяет получить уверенный точный и однозначный отсчет, однако такая огибающая имеет спектр (рис. 6.9, а), существенно отличающийся от оптимального (рис. 6.9, б). В формуле (6.24) для оценки потенциальной точности должна участвовать не вся энергия принимаемого сигнала $\mathcal{E}$, как это требуется в оптимальном случае, а только часть энергии $\mathcal{E}_т$, заключенной в гармонике частоты $\mu_{т}$; энергия же составляющих частоты $\mu_{гр}$ и $\mu_{ср}$ служит лишь для исключения неоднозначности.

Обычно многошкальная система используется в нелинейных фазовых измерителях, где отсчет производят не по положению пика огибающей, а по сдвигу фазы на каждой ее гармонике отдельно, затем результаты этих измерений объединяют. В отличие от оптимальных измерений по пику огибающей, в фазовых системах при каждом измерении используется только n -я часть общей энергии, где n — число шкал. Однако фазовые многошкальные измерители широко используют там, где ограничения по энергии отсутствуют.

6.5. РАЗРЕШАЮЩАЯ СПОСОБНОСТЬ ПО ОДНОМУ ПАРАМЕТРУ И ВЫБОР ФОРМЫ ОБОБЩЕННОЙ ОГИБАЮЩЕЙ

При создании радиотехнических измерителей необходимо предусмотреть возможность одновременного приема нескольких сигналов, которые при взаимном наложении искажаются, в результате появляются дополнительные погрешности измерений или даже сами измерения становятся невозможными. Поэтому для сохранения высокой точности измерений необходимо, чтобы сигналы, принимаемые от нескольких объектов, разделялись, т. е. чтобы система обладала разрешающей способностью по измеряемому параметру.

Хорошей мерой разрешающей способности является функция корреляции $\Psi(\nu')$, которая характеризует степень взаимного перекрытия (по площади) двух одинаковых сигналов $S(\nu - \nu_1)$ и $S(\nu - \nu_2)$, разнесенных на $\nu' = \nu_2 - \nu_1$. При $\nu' = 0$ сигналы полно-

стью перекрываются ($\Psi(0)=1$) и разрешить их невозможно; при достаточной больших расстройках обычно $\Psi(v' \rightarrow \infty)=0$ и сигналы полностью разрешаются. Поскольку функция $\Psi(v')$ описывает также поведение выходного сигнала оптимального приемника при расстройке по измеряемому параметру, можно считать, что последний реализует потенциальные возможности по разрешающей способности: если сигналы разрешаются вообще, то они разрешаются на выходе оптимального приемника и наоборот.

Результат взаимодействия двух выходных сигналов оптимального приемника зависит от случайных фазовых соотношений между ними. Поэтому для оценки разрешающей способности в среднем удобно пользоваться не самой функцией корреляции $\Psi(v')$, а квадратом ее модуля $|\Psi(v')|^2$, пропорциональным мощности выходного сигнала оптимального приемника в функции расстройки, поскольку мощность не зависит от фазовых соотношений.

Функция $|\Psi(v')|^2$ наиболее полно характеризует разрешаемую способность: по ней просто можно определить ширину и форму основного пика (область высокой корреляции) и уровень боковых лепестков (область низкой корреляции). Однако часто требуется охарактеризовать разрешающую способность одним числом. В качестве такого числа удобно принять протяженность области высокой корреляции — разрешаемый интервал

$$\Delta v = \int_{-\infty}^{\infty} |\Psi(v')|^2 dv', \quad (6.48)$$

который численно равен площади функции $|\Psi(v')|^2$ и площади его прямоугольного эквивалента, высота которого равна 1 (рис. 6.10).

Для вычисления интеграла (6.48) воспользуемся правой частью формулы (6.14) и соотношением

$$\begin{aligned} |\Psi(v')|^2 &= \Psi(v') \Psi^*(v') = \\ &= \frac{\int |S(\mu)|^2 e^{j 2\pi \mu v'} d\mu \int |S(\mu)|^2 e^{-j 2\pi \mu v'} d\mu}{\left[\int |S(\mu)|^2 d\mu \right]^2} = \\ &= \frac{\iint |S(\mu)|^2 |S(\mu')|^2 e^{j 2\pi (\mu - \mu') v'} d\mu d\mu'}{\left[\int |S(\mu)|^2 d\mu \right]^2}, \end{aligned}$$

где произведение двух интегралов заменено двойным интегралом с разделяющимися переменными μ и μ' . Подставим полученное выражение в (6.48):

$$\Delta v = \frac{\iiint |S(\mu)|^2 |S(\mu')|^2 e^{j 2\pi (\mu - \mu') v'} d\mu d\mu' dv'}{\left[\int |S(\mu)|^2 d\mu \right]^2}.$$

От переменной v' зависит только экспонента. Поэтому в числителе выделим внутренний интеграл $\int_{-\infty}^{\infty} e^{j 2\pi (\mu - \mu') v'} dv' = \delta(\mu - \mu')$, ко-

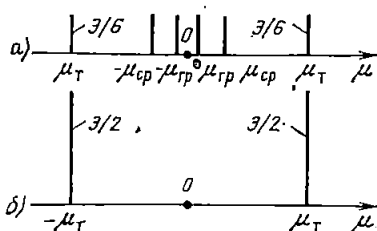


Рис. 6.9. Распределение энергии в двустороннем энергетическом спектре при многошпальном отсчете (а) и при огибающей оптимальной по точности формы (б)

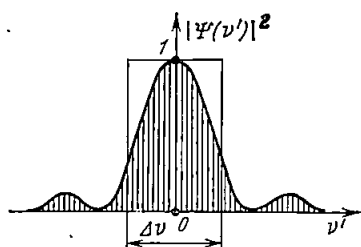


Рис. 6.10. Квадрат модуля функции корреляции и ее прямоугольный эквивалент

торый по определению является дельта-функцией, равной единице при $\mu' = \mu$ и нулю при $\mu' \neq \mu$. Ввиду этого дальнейшее интегрирование числителя по μ' дает функцию $\int_{-\infty}^{\infty} |S(\mu')|^2 \delta(\mu - \mu') d\mu' = |S(\mu)|^2$, остающуюся под знаком единственного интеграла по μ . Отсюда находим разрешаемый интервал

$$\Delta v = 1/\Delta \mu_{\text{эф}}, \quad (6.49)$$

где

$$\Delta \mu_{\text{эф}} = \left[\int_{-\infty}^{\infty} |S(\mu)|^2 d\mu \right]^2 / \int_{-\infty}^{\infty} |S(\mu)|^4 d\mu \quad (6.50)$$

— эффективная протяженность обобщенного спектра.

Подставив в (6.49) конкретные величины, получим разрешаемые интервалы по дальности, радиальной скорости и направляющим косинусам:

$$\Delta \tau = \frac{1}{\Delta f_{\text{эф}}}; \quad \Delta F = \frac{1}{\Delta \nu_{\text{эф}}}; \quad \Delta u_{x,y} = \frac{1}{\Delta v_{x,y \text{ эф}}}, \quad (6.51)$$

где эффективные протяженности спектра $\Delta f_{\text{эф}}$, сигнала $\Delta t_{\text{эф}}$ и пространственного спектра $\Delta v_{x,y \text{ эф}}$ определяются выражением (6.50) при подстановке вместо $S(\mu)$ соответственно $S(f)$, $S(t)$ и $G(v_{x,y})$.

Сплошной спектр прямоугольной формы и ширины $\Delta \mu$ имеет максимальную эффективную протяженность $\Delta \mu_{\text{эф max}} = \Delta \mu$. Все другие формы спектра при заданном $\Delta \mu$ дают $\Delta \mu_{\text{эф}} \leq \Delta \mu$, где $\Delta \mu$ — общая ширина спектра.

Эквивалентная протяженность спектра (в отличие от ширины спектра) не включает участки, где спектральная плотность равна нулю. У дискретного спектра с нулевой шириной спектральных линий (гармоник) $\Delta \mu_{\text{эф}} = 0$ и разрешающая способность отсутствует, какой бы большой ни была общая ширина $\Delta \mu$. Действительно, два гармонических сигнала одинаковой частоты да-

ют общий сигнал той же частоты, по которому невозможно определить измеряемый параметр ни одного из этих сигналов. Поэтому многошкальные системы, использующие набор гармоник, применяются только там, где исключено появление более одного сигнала или сигналы разрешаются не по измеряемому, а по другому параметру.

Таким образом, оптимальной по разрешающей способности системой является такая, у которой обобщенный спектр равномерный и сплошной на всей заданной ширине $\Delta\mu$. Определим, насколько совершенна эта система по точности, положив $|S(\mu)|^2 = 1$ на интервале $\Delta\mu = 2\mu_{\max}$ в формуле (6.14):

$$\mu_{\text{эк}}^2 = \frac{\int_{-\mu_{\max}}^{\mu_{\max}} \mu^2 d\mu}{\int_{-\mu_{\max}}^{\mu_{\max}} d\mu} = \mu_{\max}^2/3, \quad (6.52)$$

откуда находим $\mu_{\text{эк}} = \mu_{\max}/\sqrt{3} = \Delta\mu/2\sqrt{3}$, т. е. точность у этой системы всего в $\sqrt{3}$ раз меньше, чем у системы с оптимальной по точности характеристикой ($\mu_{\text{эк max}} = \Delta\mu/2$). Кроме того, при сплошном обобщенном спектре всегда обеспечивается однозначность измерений: такую систему можно считать многошкальной с бесконечным набором частот, вплотную прилегающих друг к другу.

Следовательно, оптимальная по разрешающей способности система оптимальна почти по всем показателям, только точность ее несколько ниже максимально возможной (менее чем в два раза). Заметим, что система с высокой разрешающей способностью всегда обеспечивает и высокую точность измерений, но не всегда справедливо обратное утверждение. Многие системы предназначаются не для измерений, а для получения изображения (например, карты местности), качество которого (подробность) целиком зависит от разрешающей способности. Разрешающая способность определяет также возможность отстройки от помех. Поэтому эту характеристику относят к наиболее общим и важным для всякой измерительной системы.

Однако оценивание разрешающей способности одним числом — разрешаемым интервалом $\Delta\nu$ — часто дает лишь грубое представление о действительных свойствах системы. Поэтому в ряде случаев дополнительно учитывают уровень боковых лепестков и наличие периодичностей. Оценка разрешающей способности интервалом $\Delta\nu$ применима лишь к сигналам приблизительно одинаковой интенсивности. У сигналов, интенсивность которых отличается на порядок или несколько порядков, боковые лепестки сильного сигнала могут замаскировать основной пик более слабого сигнала. Равномерный обобщенный спектр, оптимальный по минимуму разрешаемого интервала $\Delta\nu$, наихудший, поскольку при крутых срезах спектра велика интенсивность боковых лепестков у обобщенной огибающей. Так, равномерный спектр имеет огибающую вида $\sin x/x$, у которой уровень первого бокового лепестка превышает 0,2 от основного. Кроме того, трудно создать устройства для формирования и фильтрации сигналов, огибающие

которых или их спектры имеют крутые срезы. Сложно также реализовать равномерное облучение раскрыва (пространственный спектр с крутыми срезами) ввиду затекания энергии за края раскрыва, ведущего к энергетическим потерям и повышению уровня шума. Поэтому обычно используют обобщенный спектр с плавными срезами, при котором несколько расширяется основной пик огибающей (уменьшается $\Delta\mu_{\text{эф}}$), зато резко снижается уровень боковых лепестков. Разработаны специальные методы, обеспечивающие либо наибо́льшее снижение уровня боковых лепестков, либо их равномерно минимальный уровень во всем диапазоне или части диапазона измеряемого параметра, либо минимальное отношение энергии боковых лепестков к энергии основного лепестка.

Наконец, в радиосистемах часто используют периодическую последовательность импульсов, обеспечивающую непрерывность наблюдения. Периодическая обобщенная огибающая пакета повторяющихся сигналов (без шумов) на выходе оптимального приемника, пропорциональная по мощности функции $|\Psi(v')|^2$, содержит в своем спектре $|S(\mu)|^2$ полосы шириной $\Delta\mu_n = 1/Nv_n$ каждая, где v_n — период повторения, а N — число периодов в пакете, причем полосы чередуются с интервалами $\mu_n = 1/v_n$. Длительность каждого пика огибающей Δv_n и общая ширина спектра $\Delta\mu$ находятся в обратной зависимости: $\Delta v_n = 1/\Delta\mu$ (рис. 6.11).

Эффективная протяженность прерывистого спектра, приближенно равная сумме протяженностей всех спектральных полос:

$$\Delta\mu_{\text{эф}} = N\Delta\mu_n = \Delta\mu\Delta v_n/v_n = \Delta\mu/q \quad (6.53)$$

в q раз меньше общей ширины спектра $\Delta\mu$, где $q = v_n/\Delta v_n$ — скважность. При этом $\Delta\mu_{\text{эф}}$ определяет разрешающую способность в целом, учитывая как протяженность отдельного пика огибающей, так и число пиков, т. е. степень неоднозначности. Однако, когда интервал измерений не превышает периода повторения, т. е.

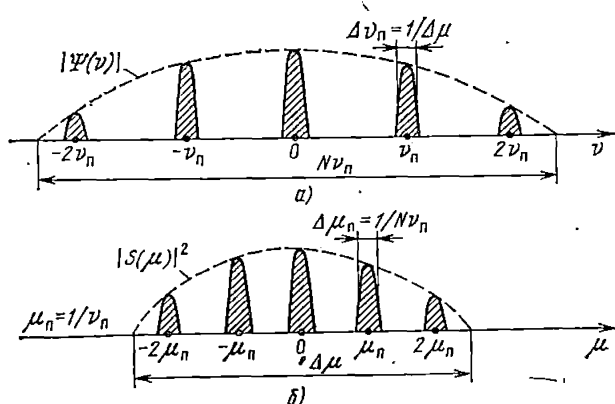


Рис. 6.11. Квадрат модуля функции корреляции (а) и энергетический спектр (б) пакета повторяющихся сигналов

$[v_{\max} - v_{\min}] \leq v_n$, чего иногда добиваются для исключения неоднозначности, разрешающая способность практически определяется только протяженностью пика огибающей $\Delta v_n = 1/\Delta \mu$, т. е. обратно пропорциональна полной ширине спектра. В частности, при измерении дальности R по отраженному или ответному сигналу интервал измерений времени запаздывания $\tau = 2R/c$ лежит в пределах $0 < \tau \leq \tau_{\max}$. Тогда однозначность измерений обеспечивается выбором периода повторения $T_n \geq [\tau_{\max} - \tau_{\min}] = \tau_{\max} = 2R_{\max}/c$, где $R_{\max}$ определяется из уравнения дальности действия. При измерении радиальной скорости $\dot{R}$ по этому же сигналу доплеровский сдвиг частоты $F = -2\dot{R}/\lambda$ лежит в пределах $-|F|_{\max} \leq F \leq |F|_{\max}$ и однозначность измерений обеспечивается выбором частоты повторения $F_n = 1/T_n$ из условия $F_n \geq 2|F|_{\max} = 4|\dot{R}|_{\max}/\lambda$. При выполнении указанных условий, которые часто оказываются взаимно противоречивыми, разрешающая способность по дальности полностью оценивается только шириной спектра импульса ($\Delta \tau = 1/\Delta f$), а по скорости только длительностью импульса ($\Delta F = 1/\Delta t$).

Интервал измерений дальности и радиальной скорости хотя и конечен, но может быть сколь угодно большим. Однако направляющие косинусы изменяются в ограниченных пределах $-1 \leq u_{x,y} \leq 1$, т. е. полный интервал их изменения $[u_{x,y \max} - u_{x,y \min}] = 2$ и период однозначных измерений удовлетворяет условию $u_{x,y n} = 2$. Поэтому, не опасаясь неоднозначности, сплошной угловой спектр $G(v_{x,y})$ можно заменить дискретным с интервалом $v_{x,y n} = 1/u_{x,y n} \leq 1/2$. Это означает, что антенная решетка с шагом $b \leq \lambda/2$, т. е. $v_{x,y n} = b/\lambda = 1/2$, практически эквивалентна сплошной антенне.

В заключение отметим, что точность измерений определяется (через $\mu_{\text{эк}}$) общей шириной энергетического спектра, в то время как разрешающая способность (с учетом неоднозначности) — суммарной протяженностью $\Delta \mu_{\text{эф}}$ только «заполненных» участков спектра, т. е. его общей шириной за вычетом «пустых» участков.

6.6. СЛОЖНАЯ ОГИБАЮЩАЯ И ОПТИМАЛЬНОЕ СЖАТИЕ СИГНАЛА

Одной из важных характеристик обобщенной огибающей является ее база — произведение ширины спектра огибающей $\Delta \mu$ на ее исходную (до фильтрации) протяженность $v_{\text{исх}}$: $N = \Delta \mu v_{\text{исх}}$. По этой характеристике огибающие подразделяются на два вида: простые, у которых $N \approx 1$, и сложные, у которых $N \gg 1$. Простые имеют вещественный (синфазный) спектр, а сложные — комплексный (спектр расфазирован). Независимо от этого протяженность Δv основного лепестка выходного сигнала оптимального приемника согласно формуле (6.49) определяется только эффективной протяженностью спектра $\Delta \mu_{\text{эф}} \approx \Delta \mu$, т. е. $\Delta v \approx 1/\Delta \mu$. Следовательно, после прохождения через согласованный фильтр простые огибающие практически сохраняют свою протяженность ($\Delta v \approx 1/\Delta \mu \approx$

$\approx v_{исх}$), а сложные сжимаются в N раз ($\Delta v \approx 1/\Delta \mu \approx v_{исх}/N$). Это объясняется тем, что спектр выходного сигнала оптимального приемника $|S(\mu)|^2$ всегда вещественный, поэтому протяженность пика огибающей выходного сигнала $\Psi(v)$ минимальна независимо от протяженности огибающей входного сигнала.

При измерении дальности обобщенной огибающей является реальная огибающая сигнала. Сложная огибающая, например, колоколообразного импульса с глубокой линейной частотной модуляцией $S(t) = \exp[-\pi t^2/\tau_n^2 - j\pi a t^2]$ (как было показано в § 3.2) имеет комплексный спектр, ширина которого равна девиации частоты $\Delta f \approx a\tau_n$, изменяемой со скоростью a в течение длительности импульса τ_n , а база $N = \Delta f \tau_n \approx a\tau_n^2$. Длительность выходного сигнала $\Delta \tau = 1/\Delta f \approx 1/a\tau_n = \tau_n/N$ сжимается в N раз. Во столько же раз возрастает мощность сигнала, так как энергия в импульсе сохраняется неизменной. Таким образом, сложный сигнал большой длительности и малой мощности эквивалентен в оптимальной системе простому сигналу малой длительности и большой мощности при одинаковой ширине спектра. Выходной сигнал будет иметь одинаковую форму и энергию.

При измерении радиальной скорости роль обобщенной комплексной огибающей играет спектр огибающей реального сигнала, а обобщенным спектром является сама огибающая. Поэтому сжатая обобщенная огибающая получается в результате свертки спектральных характеристик сигнала и фильтра или (что то же самое) умножения комплексной огибающей сигнала на сопряженную с ней импульсную характеристику фильтра. Таким образом, для оптимального сжатия спектра сигнала нужно перемножить, например на смесителе, два сложных сигнала, у которых изменение частоты и фазы происходит синхронно. В результате получаем сигнал постоянной разностной частоты (или видеосигнал), имеющий узкий спектр, так как происходит демодуляция входного сигнала. В приведенном выше примере гауссовского импульса с глубокой линейной частотной модуляцией при измерении радиальной скорости (частотного сдвига) роль обобщенной огибающей играет его спектр $S(f) \approx \exp[-\pi f^2/\Delta f^2 - j\pi f^2/a]$, закон изменения фазы которого с частотой (мнимый член в показателе экспоненты) также квадратичный. После перемножения с опорным сигналом происходит сжатие спектра $\Delta f \approx a\tau_n$ в $N = a\tau_n^2$ раз. В результате ширина спектра на выходе $\Delta F = \Delta f/N = a\tau_n/a\tau_n^2 = 1/\tau_n$, как и у немодулированного, простого сигнала. Благодаря узкому пику спектра обеспечивается высокая разрешающая способность по частоте, т. е. по радиальной скорости, и высокая точность ее измерения.

Нужно, однако, отметить, что свертку спектров можно выполнить только с помощью коррелятора, а поэтому необходимо обеспечить одновременность поступления входного и опорного сигналов.

При измерении угловой координаты роль обобщенной огибающей играет диаграмма направленности антенны, т. е. функция уг-

ла, а роль обобщенного спектра — функция раскрыва, т. е. угловой пространственный спектр. По отношению к ним справедливы все положения, которые были высказаны выше применительно к сигналу как функции времени и его спектру. В частности, комплексная функция раскрыва (несинфазный угловой спектр) дает широкую диаграмму направленности, которую можно сжать до минимально возможной, пропустив через согласованный пространственный фильтр (фазированием).

Применение обобщенной огибающей сложной формы позволяет обеспечить, совместить энергетические и метрические требования к измерительной системе, которые часто противоречивы. Энергетическим требованием является обеспечение заданной дальности действия в заданном секторе, а метрическими — заданная точность измерений и разрешающая способность по измеряемому параметру. При огибающих простой формы эти противоречивые требования часто неразрешимы. Покажем это на примере временной и пространственной характеристик системы.

Пусть временная характеристика — огибающая излучаемого сигнала — имеет вид пакета из n простых импульсов длительностью $\tau_n = 1/\Delta f$ и мощностью P_n каждый. Так как дальность действия системы определяется полной энергией излучаемого сигнала $\mathcal{E} = P_n \tau_n n$, ее можно повысить, увеличив все три параметра или хотя бы один из них. Однако число импульсов, принимаемых за время облучения объекта, ограничено допустимым временем обзора всей зоны. Излучаемая мощность ограничена опасностью пробоя в антенно-фидерной системе. Остается единственная возможность — увеличение длительности импульса. Однако с увеличением длительности τ_n простых импульсов ($\Delta f \tau_n = 1$) уменьшается ширина спектра $\Delta f = 1/\tau_n$ и, следовательно, ухудшается точность измерений и разрешающая способность по дальности ($\Delta \tau \approx 1/\Delta f = \tau_n$). Это противоречие разрешается при использовании импульсов сложной формы, база которых $N = \Delta f \tau_n \gg 1$ за счет внутриимпульсной модуляции. При достаточно большой базе заданная дальность действия обеспечивается за счет большой длительности импульса τ_n , а заданная точность и разрешающая способность ($\Delta \tau \approx 1/\Delta f = \tau_n/N$) — за счет выбора ширины спектра Δf .

Обратимся теперь к пространственной характеристике системы — функции раскрыва ФАР. Предположим, что система предназначена для одновременного мгновенного обзора всей зоны и обладает высокими метрическими характеристиками, т. е. высокой точностью и разрешающей способностью по угловой координате. Это обеспечивается большой шириной пространственного спектра $\Delta u_{x,y} = d/\lambda$, т. е. большим относительным размером раскрыва антенны. Если при этом использовать синфазный раскрыв, т. е. вещественный угловой спектр с базой $N = \Delta u_{x,y} u_{x,y} \approx 1$, то ФАР может излучать (и принимать) сигналы только в пределах узкого сектора $u_{x,y} \approx 1/\Delta u_{x,y}$.

Таким образом, простая апертура ($N=1$) не обеспечивает облучения в широком секторе, которое необходимо для мгновенно-

го обзора в заданной зоне. Если же ввести сдвиг по фазе между сигналами, излучаемыми элементами ФАР (комплексный пространственный спектр), ширина диаграммы направленности $u_{x,y}$ и возрастет во много раз и перекроет всю зону обзора. База полученной сложной апертуры $N = \Delta u_{x,y} u_{x,y} \gg 1$. При оптимальном приеме сдвиг по фазе между пространственными спектральными составляющими устраняется и диаграмма направленности сжимается до предельно малой ширины $\Delta u_{x,y} \approx 1/\Delta u_{x,y} = u_{x,y}/N$, что обеспечивает высокую угловую точность и разрешающую способность. Снижение дальности действия, вызванное введением фазовых сдвигов по раскрыву, при излучении компенсируется установкой передатчиков в канале каждого элемента ФАР. Мощность излучения каждого передатчика остается много ниже предела, определяемого опасностью пробоя, хотя при большом числе элементов мощность излучения всей ФАР очень высока.

Облучение заданной зоны можно обеспечить и отдельной излучающей антенной малых размеров, создающей широкую диаграмму направленности без расфазирования сигналов по раскрыву ($N = \Delta u_{x,y} u_{x,y} \approx 1$). Однако при большой дальности действия системы требуемая мощность излучения может значительно перекрыть допустимую.

Схема системы, в которой используется принцип оптимального сжатия диаграммы направленности (рис. 6.12, а), содержит в канале каждого элемента ФАР усилители мощности (УМ), возбуждаемые от общего задающего генератора (ЗГ), антенные переключатели (АП) и фиксированные фазовращатели φ_i , обеспечивающие требуемое распределение фаз по раскрыву. В приемном тракте каждого элемента включают компенсирующие фазовращатели $-\varphi_i$, которые вместе с многоканальным приемником (многоканальный вход и выход) образуют оптимальный пространственный фильтр, обеспечивающий одновременный прием сигналов со всех направлений с высокой точностью пеленгации и

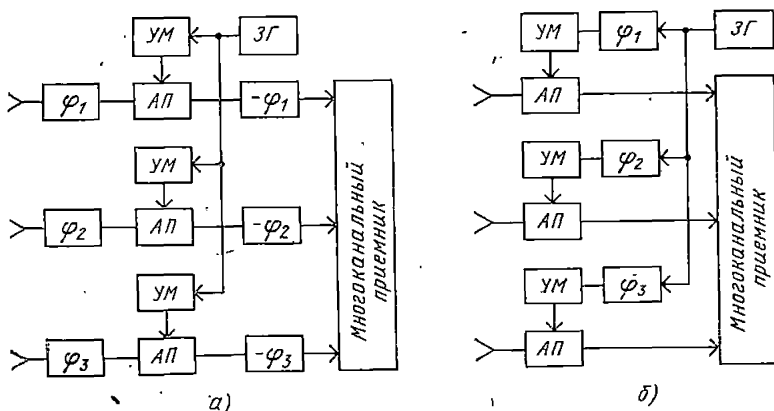


Рис. 6.12. Схемы системы мгновенного образа с оптимальным сжатием диаграммы направленности

разрешающей способностью по направлению благодаря сжатию диаграммы направленности. Поскольку расфазировка сигналов по раскрыву необходима только при излучении, возможна более простая система, которая содержит фазовращатели только в передающем тракте (рис. 6.12,6).

6.7. НЕОПТИМАЛЬНОЕ СЖАТИЕ ОГИБАЮЩЕЙ И УТОЧНЕНИЕ ПОНЯТИЯ РАЗРЕШАЮЩЕЙ СПОСОБНОСТИ

Разрешаемый интервал $\Delta\nu$, характеризующий ширину обобщенной функции корреляции $\Psi(\nu')$ (т. е. выходного сигнала оптимального приемника вдоль измеряемого параметра), является объективной оценкой разрешающей способности системы при двух условиях: объект считается точечным, т. е. аппроксимируется дельта-функцией измеряемого параметра $\delta(\nu)$, и сама система обеспечивает максимальное отношение сигнал-шум на выходе, благодаря чему достигается наивысшее качество обнаружения объекта. Если отказаться от этих условий и рассматривать задачу приема сигнала несколько шире, то понятие разрешающей способности потребует уточнения. Эти вопросы, связанные с задачей распознавания формы объектов, кратко рассматриваются ниже.

Для учета влияния конечных размеров, конфигурации и отражающих свойств объекта на входной и выходной сигналы оптимальной системы введем понятие импульсной реакции объекта $Q(\nu)$. Реакция $Q(\nu)$ отображает поведение комплексной огибающей (амплитуды и фазы) отраженных колебаний вдоль измеряемой координаты при облучении объекта радиоволной, огибающая которой вдоль этой координаты описывается дельта-функцией $\delta(\nu)$. Так, по дальности роль такого колебания играет радиоимпульс, длительность которого много меньше времени прохождения волной этого объекта или его элементов, а по угловым координатам — узкая диаграмма направленности, ширина которой много меньше видимых угловых размеров объекта или его элементов.

Импульсная реакция объекта является комплексной функцией, зависящей от частоты облучающих колебаний и учитывающей фазовые соотношения между сигналами отдельных элементов объекта, которые определяются разностью радиальных расстояний до них, выраженной в длинах волн. Квадрат модуля импульсной реакции объекта $|Q(\nu)|^2 = \sigma_0(\nu)$, пропорциональный мощности, отображает поведение эффективной площади рассеяния элементов этого объекта вдоль данной координаты $\sigma_0(\nu)$. Преобразование Фурье от функции $Q(\nu)$ является спектром импульсной реакции объекта $Q(\mu)$. Если же объект облучается волной, огибающая которой описывается не δ -функцией, а временной характеристикой системы $S(\nu)$ конечной протяженности, то огибающая отраженного сигнала имеет вид свертки

$$S_0(\nu) = \int_{-\infty}^{\infty} Q(\nu') S(\nu - \nu') d\nu', \quad (6.54)$$

а ее спектр является произведением

$$S_0(\mu) = Q(\mu) S(\mu). \quad (6.55)$$

В частном случае, когда объект точечный ($Q(v) = \delta(v)$), отраженный сигнал и его спектр описываются только характеристиками самой системы: $S_0(v) = S(v)$ и $S_0(\mu) = S(\mu)$. Именно этот случай рассматривается в данной книге. В другом частном случае, когда $S(v) = \delta(v)$, имеем $S_0(v) = Q(v)$ и $S_0(\mu) = Q(\mu)$, т. е. реакция объекта воспроизводится без искажений, а сама система называется идеальной. Однако в приведенных определениях не учитывается поляризация (т. е. векторный характер облучающей и отраженной волн), поэтому реакция $Q(v)$ не является исчерпывающей характеристикой объекта. Следовательно, термин «идеальная система» нужно понимать как условный. Кроме того, необходимо учесть, что импульсная характеристика любого объекта (кроме шара) зависит от его ориентации относительно направления наблюдения и должна быть представлена для каждой из трех координат при всех возможных углах ориентации (поворота относительно трех его осей). Все это делает задачу описания совокупной импульсной реакции и отраженного сигнала протяженного объекта исключительно сложной. Мы же задачу существенно упростим, исключив из рассмотрения поляризацию и полагая неизменной (независимой от ориентации) импульсную реакцию объекта по единственной координате — измеряемому параметру v .

Нормированный выходной сигнал оптимальной системы находим как обобщенную функцию взаимной корреляции сигнала протяженного объекта $S_0(v)$ и реакции системы $S^*(v)$ с учетом соотношения (6.54)

$$\Psi_0(v') = \frac{1}{2\theta} \int_{-\infty}^{\infty} S_0(v) S^*(v-v') dv = \int_{-\infty}^{\infty} Q(v) \Psi(v-v') dv, \quad (6.56)$$

где

$$\Psi(v') = \frac{1}{2\theta} \int_{-\infty}^{\infty} S(v) S^*(v-v') dv \quad (6.57)$$

— обобщенная функция корреляции сигнала точечного объекта. Соответственно из формул (6.55) и (6.56) находим обобщенный спектр выходного сигнала

$$\Psi_0(\mu) = Q(\mu) |S(\mu)|^2. \quad (6.58)$$

Как следует из (6.56), сигнал протяженного объекта на выходе оптимальной системы является сверткой импульсной реакции объекта $Q(v)$ и выходной реакции системы на точечный объект. Функция $\Psi(v)$ маскирует, сглаживает истинный вид реакции $Q(v)$, характеризующей индивидуальные свойства протяженного объекта. У идеальной системы $\Psi(v) = \delta(v)$, и форма ее выходного сигнала $\Psi_0(v) = Q(v)$ воспроизводит реакцию объекта без искажений.

В принципе и по выходному сигналу реальной системы $\Psi_0(v)$ можно восстановить истинный вид функции $Q(v)$, поскольку ре-

акция системы на точечный объект $\Psi(v)$ известна априори. Математическое восстановление этой функции сводится к решению интегрального уравнения (6.56) относительно функции $Q(v)$ при известных зависимостях $\Psi_0(v)$ и $\Psi(v)$. Указанная операция называется редукцией (сведением) реальной системы к идеальной. В результате редукции система как бы приобретает неограниченную разрешающую способность, позволяющую различать объекты или детали протяженного объекта в пределах разрешаемого интервала $\Delta v = 1/\Delta \mu_{эф}$ оптимальной системы. Практически и после редукции разрешающая способность системы всегда ограничена, поскольку на ее вход сигнал $S_0(v)$ поступает не в чистом виде, а в смеси с шумом $N(v)$. Во входном воздействии $X(v) = S_0(v) + N(v)$ шум маскирует истинный вид функции $S_0(v)$, а следовательно, и $Q(v)$. Таким образом, достижимая разрешающая способность, как и точность измерений, зависит от отношения сигнал-шум S/N_0 .

Технически редукция, т. е. восстановление истинного вида функции $Q(v)$ по входному $S_0(v)$ или выходному $\Psi_0(v)$ сигналу, выполняется с помощью восстанавливающего фильтра. Его задачей является придание характеристике системы $S(v)$ или $\Psi(v)$ свойств дельта-функции или, что равносильно, придание спектрам $S(\mu)$ или $|S(\mu)|^2$ равномерного вида в широком диапазоне частот. Отсюда вытекает условие для функции передачи восстанавливающего фильтра

$$H_{вф}(\mu)S(\mu) = 1, \quad (6.59)$$

из которого находим вид самой функции

$$H_{вф}(\mu) = 1/S(\mu). \quad (6.60)$$

Таким образом, функция передачи восстанавливающего фильтра прямо противоположна функции передачи приемника оптимальной системы $H(\mu) = S^*(\mu)$. В оптимальном согласованном фильтре с наибольшим весом передаются те участки спектра, где сигнал велик по сравнению с шумом, в результате максимизируется общее отношение сигнал-шум на выходе. В восстанавливающем фильтре, наоборот, усиливаются в большей степени те участки спектра, где сигнал относительно слаб, и результирующее отношение сигнал-шум резко снижается. Это и ограничивает возможность редукции. (Редукция, сопровождающаяся поднятием спектральных составляющих в области высоких частот, используется в телевидении для повышения четкости изображения, где получила название частотной коррекции.)

К сожалению, идеальный восстанавливающий фильтр нереализуем (рис. 6.13, а). Во-первых, потребовался бы недопустимо большой коэффициент передачи фильтра на некоторых участках спектра, а во-вторых (и это главное), интенсивность шумов на выходе фильтра возрастала бы безгранично. Поэтому вводят некоторую

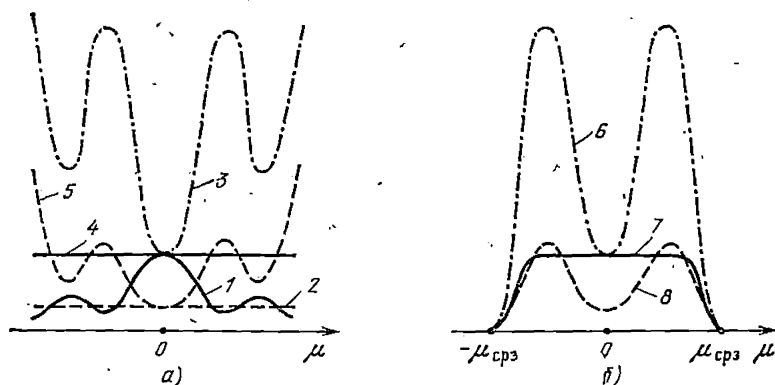


Рис. 6.13. Частотные зависимости при идеальной (а) и реальной (б) редукции: 1, 2 — спектры огибающих входного сигнала и шума; 3, 6 — функции передачи идеального и реального восстанавливающих фильтров; 4, 5 и 7, 8 — спектры огибающих сигнала и шума на выходе идеального и реального фильтров соответственно

частоту среза $\mu_{срз}$, т. е. характеристику (6.60) умножают на функцию

$$K_{ид}(\mu) = \begin{cases} 1, & |\mu| \leq \mu_{срз}, \\ 0, & |\mu| > \mu_{срз}. \end{cases} \quad (6.61)$$

Однако такая идеальная функция нереализуема. Реальная функция $K(\mu)$ должна плавно спадать на краях полосы прозрачности (рис. 6.13, б). Эта функция выбирается достаточно широкой, чтобы в минимальной степени искажалась требуемая характеристика восстанавливающего фильтра (6.60) и, следовательно, определяемая функция $Q(v)$. В то же время она должна быть достаточно узкой, чтобы и шумы не искажали функцию $Q(v)$. Отыскание функции $K(\mu)$, обеспечивающей минимум искажений импульсной реакции объекта $Q(v)$, называется регуляризацией восстанавливающего фильтра, а сама функция $K(\mu)$ регуляризующей.

Рассмотрим самые общие закономерности редукции. Отсекаемая регуляризующей функцией $K(\mu)$ часть спектра полезной составляющей входного воздействия искажает искомую функцию $Q(v)$. Искажения можно существенно снизить, если учесть априорную информацию об объекте и общих свойствах функции $Q(v)$ реальных объектов. Интересен частный случай, когда протяженный объект состоит из нескольких точечных объектов, попадающих в интервал $\Delta v = 1/\Delta\mu_{эф}$ и не разрешаемых оптимальной системой. В этом случае можно говорить о разрешающей способности внутри интервала Δv , оцениваемой шириной пиков выходного сигнала восстанавливающего фильтра. Ширина пиков, в свою очередь, определяется шириной полосы пропускания этого фильтра $\Delta\mu_{вф} = 2\mu_{срз}$ или частотой среза. Поскольку последняя выбирается из условия равенства искажений функции $Q(v)$, вносимых шумами и отсечением части спектра полезного сигнала, с повышением уровня шума оптимальная ширина полосы пропускания,

удовлетворяющая этому условию, сужается и разрешающая способность редуцированной системы снижается. В этом проявляется зависимость предельно возможной разрешающей способности от отношения сигнал-шум. Точность редукции с уменьшением отношения сигнал-шум снижается очень быстро.

У оптимальной системы, без редукции, энергетические соотношения наиболее благоприятны и дальность действия максимальна, но разрешающая способность ограничена шириной пика выходного сигнала $\Delta\nu$. На максимальной дальности редукция невозможна, так как пороговое отношение сигнал-шум мало. Она применима и дает выигрыш в разрешающей способности на меньших дальностях, когда сигнал достаточно велик по сравнению с шумами.

Таким образом, редукция является операцией обмена между энергетическими и метрическими параметрами системы. В частности, она позволяет увеличить метрический параметр $\mu'_{\text{ЭК}}$, характеризующий точность измерений, снизив энергетическое отношение сигнал-шум $\mathcal{E}/N_0$. Однако в целом потенциальная точность, определяемая согласно формуле (6.24) произведением $(\mathcal{E}/N_0)\mu'^2_{\text{ЭК}}$, не может быть превышена. Нельзя также повысить и разрешающую способность выше величины, определяемой произведением $(\mathcal{E}/N_0)\Delta\mu_{\text{эф}}$, независимо от того, производится редукция или нет. Если сигнал теряется в шумах, уверенное разрешение или измерение вообще становится невозможным. Это относится ко всем системам, с редукцией или без редукции.

Можно показать, что при измерениях редукция дает полезный эффект, когда результирующая точность измерений ограничена не шумом, а другими факторами (например, инструментальной погрешностью). Для этого сравним среднеквадратические погрешности измерений, вызванные шумами (σ_v) и нешумовыми помехами ($\sigma_{\text{ш}}$) в оптимальной и редуцированной системах (рис. 6.14). Как уже отмечалось, в результате редукции возрастает метрический параметр системы ($\mu'_{\text{ЭК}} > \mu_{\text{ЭК}}$), но ухудшается энергетическое отношение сигнал-шум из-за возрастания уровня шума ($\sigma'_{\text{ш}} > \sigma_{\text{ш}}$). В итоге потенциальная точность системы после редукции остается практически неизменной ($\sigma'_v \approx \sigma_v$), но реальная, ограниченная нешумовыми помехами, возрастает ($\sigma'_{\text{ш}} < \sigma_{\text{ш}}$). Повышение результирующей точности при редукции возможно до тех пор, пока возрастающий уровень шума не станет главным фактором, огра-

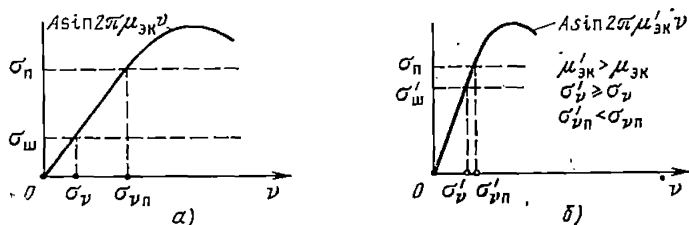


Рис. 6.14. Погрешности измерений в оптимальной (а) и редуцированной (б) системах, обусловленные шумами и нешумовыми помехами

ничивающим точность, т. е. пока сохраняется условие $\sigma'_ш < \sigma_n$. Это справедливо и для разрешающей способности; если ее ограничивают не шумы, сжатие выходного сигнала при редукции даст полезный эффект.

Необходимо отметить, что низкая эффективность редукции при наличии заметных шумов объясняется тем, что одновременно с повышением разрешающей способности из-за поднятия высокочастотных составляющих спектра, передающих мелкие детали объекта, повышаются и шумы в указанной части спектра, которые маскируют эти детали. В отсутствие шумов любая система редуцируется к идеальной, но при заданном уровне шумов более высокая разрешающая способность после редукции сохраняется у той системы, у которой она была таковой и до редукции.

В заключение отметим, что редукция является линейной операцией над сигналом, а на выходе системы регистрируется огибающая сигнала, получаемая в результате нелинейной операции. Поэтому редукция должна предшествовать получению огибающей. Обычно перед редукцией высокочастотный сигнал с помощью фазовых детекторов разделяется на две ортогональные составляющие, после раздельной редукции которых происходит их объединение в квадратуре. Для выполнения редукции используются обычно оптические и цифровые процессоры (в частности, голография намечалась вначале как средство повышения четкости оптических изображений, полученных с помощью систем с низкой разрешающей способностью).

6.8. ПОТЕНЦИАЛЬНАЯ ТОЧНОСТЬ ИЗМЕРЕНИЙ СОВОКУПНОСТИ НЕИЗВЕСТНЫХ ПАРАМЕТРОВ СИГНАЛА

Обычно у принимаемого сигнала неизвестно сразу несколько параметров, образующих совокупность $\mathbf{v} = \{v_1, v_2, \dots, v_n\}$. Расстройка по каждому из них влияет на выходной сигнал оптимального приемника, что учитывается многомерной функцией корреляции $\Psi(\mathbf{v}')$. Распределение вероятностей совокупности измеряемых параметров при достаточно большом отношении сигнал-шум, обеспечивающем заданную точность измерений, подчинено многомерному нормальному закону. В отличие от нормального распределения единственного параметра сигнала, описываемого дисперсией оценки через вторую производную одномерной функции корреляции $a = -0,5 [d^2 |\Psi(\mathbf{v}')|^2 / d\mathbf{v}'^2]_{\mathbf{v}'=0}$, для определения характеристик многомерного нормального распределения необходимо вычислить целый набор вторых смешанных частных производных квадрата многомерной функции корреляции по совокупности всех измеряемых параметров

$$a_{ij} = -\frac{1}{2} \frac{\partial^2}{\partial v'_i \partial v'_j} |\Psi(\mathbf{v}')|^2_{\mathbf{v}'=0}. \quad (6.62)$$

Упорядоченное расположение этих производных образует квадратную матрицу

$$A = \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \dots \dots \\ a_{n1} & a_{n2} \dots a_{nn} \end{bmatrix},$$

которая симметрична ($a_{ij}=a_{ji}$), так как порядок дифференцирования роли не играет.

Как и при измерении единственного параметра, для определения потенциальной точности оценки k -го параметра при совместном измерении нескольких неизвестных параметров можно воспользоваться формулой (6.32), только вместо $1/a$ необходимо подставить a^{-1}_{kk} :

$$\sigma^2_{v_k} = \frac{1}{2\mathcal{N}/N_0} a^{-1}_{kk}, \quad (6.63)$$

где a^{-1}_{kk} — диагональный элемент матрицы A^{-1} , обратной матрице A . Элементы обратной матрицы определяются известным соотношением

$$a^{-1}_{ij} = A_{ij}/|A|, \quad (6.64)$$

где $|A|$ — определитель матрицы A ; $A_{ij} = (-1)^{i+j}|A_{ij}|$ — алгебраическое дополнение элемента a_{ij} матрицы A ; $|A_{ij}|$ — определитель матрицы, полученной из матрицы A вычеркиванием i -й строки и j -го столбца. Матрица A^{-1} также симметрична ($a^{-1}_{ij}=a^{-1}_{ji}$).

Для вычисления элемента a_{ij} непосредственно через многомерную функцию корреляции $\Psi(\mathbf{v}')$, а не через квадрат ее модуля $|\Psi(\mathbf{v}')|^2 = \Psi(\mathbf{v}')\Psi^*(\mathbf{v}')$, введем обозначения:

$$\Psi = \Psi(\mathbf{v}'), \quad \Psi'_i = \frac{\partial}{\partial v'_i} \Psi(\mathbf{v}'), \quad \Psi''_{ij} = \frac{\partial^2}{\partial v'_i \partial v'_j} \Psi(\mathbf{v}') \text{ и т. д.}$$

При этом получим

$$\frac{\partial}{\partial v'_i} |\Psi(\mathbf{v}')|^2 = \Psi'_i \Psi^* + \Psi \Psi'^*_{i*},$$

$$\frac{\partial^2}{\partial v'_i \partial v'_j} |\Psi(\mathbf{v}')|^2 = \Psi''_{ij} \Psi^* + \Psi'_i \Psi'^*_{j*} + \Psi'_j \Psi'^*_{i*} + \Psi \Psi''^*_{ij*}.$$

Так как произведения $\Psi''_{ij}\Psi^*$ и $\Psi\Psi''^*_{ij*}$, $\Psi'_i\Psi'^*_{j*}$ и $\Psi'_j\Psi'^*_{i*}$ попарно являются комплексно-сопряженными, при суммировании их вещественные части удваиваются, а мнимые взаимно исключаются:

$$\Psi''_{ij}\Psi^* + \Psi\Psi''^*_{ij*} = 2 \operatorname{Re}(\Psi''_{ij}\Psi^*),$$

$$\Psi'_i\Psi'^*_{j*} + \Psi'^*_{i*}\Psi'_j = 2 \operatorname{Re}(\Psi'_i\Psi'^*_{j*}).$$

Возвращаясь к полной записи и учитывая, что $\Psi(0) = \Psi^*(0) = 1$, вместо формулы (6.62) получаем

$$a_{ij} = -\operatorname{Re} \left[\frac{\partial^2 \Psi(\mathbf{v}')}{\partial v'_i \partial v'_j} + \frac{\partial \Psi(\mathbf{v}')}{\partial v'_i} \frac{\partial \Psi^*(\mathbf{v}')}{\partial v'_j} \right]_{\mathbf{v}'=0}. \quad (6.65)$$

Формула (6.63) совместно с формулой (6.65) и матричным преобразованием (6.64) определяет характеристики многомерного нормального закона и, следовательно, потенциальную точность измерений k -го параметра сигнала при наличии нескольких неизвестных параметров.

Вычисление диагонального элемента a_{kk} ничем не отличается от вычисления коэффициента a для единственного параметра, так как расстройки по всем остальным параметрам можно положить нулевыми еще до дифференцирования, а $\Psi'^*_{\dot{k}} = -\Psi'_{\dot{k}}$, поэтому получаем вещественную производную

$$a_{kk} = - \left\{ \frac{\partial^2 \Psi(\nu'_k)}{\partial \nu'^2_k} - \left[\frac{\partial \Psi(\nu'_k)}{\partial \nu'_k} \right]^2 \right\}_{\nu'_k=0} \quad (6.66)$$

При этом эквивалентную частоту для k -го параметра находим из соотношения

$$\mu^2_{k \text{ эк}} = a_{kk} / (2\pi)^2. \quad (6.67)$$

Так как многомерная обобщенная функция корреляции обычно распадается на две двумерные, временную и пространственную, вида

$$\begin{aligned} \Psi(\nu'_1, \nu'_2) &= \frac{\iint S(\nu_1, \nu_2) S^*(\nu_1 - \nu'_1, \nu_2 - \nu'_2) d\nu_1 d\nu_2}{\iint |S(\nu_1, \nu_2)|^2 d\nu_1 d\nu_2} = \checkmark \\ &= \frac{\iint |S(\mu_1, \mu_2)|^2 \exp[j2\pi(\mu_1 \nu'_1 + \mu_2 \nu'_2)] d\mu_1 d\mu_2}{\iint |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2}, \end{aligned} \quad (6.68)$$

каждой из них соответствует двумерная матрица $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, у которой $a_{21} = a_{12}$, $|A| = a_{11}a_{22} - a_{12}^2$, $A_{11} = a_{22}$, $A_{22} = a_{11}$. В результате дифференцирования функции (6.68) получим

$$\begin{aligned} a_{11} &= -[\partial^2 \Psi(\nu'_1, \nu'_2) / \partial \nu'^2_1]_{\nu'_1, \nu'_2=0} = (2\pi\mu_{1 \text{ эк}})^2, \\ a_{22} &= -[\partial^2 \Psi(\nu'_1, \nu'_2) / \partial \nu'^2_2]_{\nu'_1, \nu'_2=0} = (2\pi\mu_{2 \text{ эк}})^2, \\ a_{12} = a_{21} &= -\text{Re} [\partial^2 \Psi(\nu'_1, \nu'_2) / \partial \nu'_1 \partial \nu'_2]_{\nu'_1, \nu'_2=0} = (2\pi)^2 R_{12}, \end{aligned} \quad (6.69)$$

где

$$\mu_{1 \text{ эк}} = \sqrt{\mu_{100}^2 - \mu_{10}^2}, \quad \mu_{2 \text{ эк}} = \sqrt{\mu_{200}^2 - \mu_{20}^2} \quad (6.70)$$

— частоты эквивалентных гармонических огибающих для каждого из двух измеряемых параметров, определяемые через вторые начальные моменты энергетического спектра

$$\begin{aligned} \mu_{100}^2 &= \frac{\iint \mu_1^2 |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2}{\iint |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2}, \\ \mu_{200}^2 &= \frac{\iint \mu_2^2 |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2}{\iint |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2} \end{aligned} \quad (6.71)$$

и первые начальные моменты (смещение) энергетических спектров по каждой из двух обобщенных частот

$$\mu_{10} = \frac{\iint \mu_1 |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2}{\iint |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2}, \quad (6.72)$$

$$\mu_{20} = \frac{\iint \mu_2 |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2}{\iint |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2};$$

$$R_{12} = \frac{\iint \mu_1 \mu_2 |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2}{\iint |S(\mu_1, \mu_2)|^2 d\mu_1 d\mu_2} \quad (6.73)$$

— ненормированный коэффициент корреляции между параметрами ν_1 и ν_2 . Введя нормированный коэффициент корреляции

$$\rho_{12} = R_{12}/\mu_{1\text{ЭК}} \mu_{2\text{ЭК}} = a_{12}/\sqrt{a_{11} a_{22}}, \quad (6.74)$$

получим

$$a_{11}^{-1} = \frac{a_{22}}{a_{11} a_{22} - a_{12}^2} = \frac{1}{a_{11} (1 - a_{12}^2/a_{11} a_{22})} = \frac{1}{a_{11} (1 - \rho_{12}^2)},$$

$$a_{22}^{-1} = \frac{a_{11}}{a_{11} a_{22} - a_{12}^2} = \frac{1}{a_{22} (1 - a_{12}^2/a_{11} a_{22})} = \frac{1}{a_{22} (1 - \rho_{12}^2)}.$$

Отсюда в соответствии с общей формулой (6.63) окончательно находим дисперсию, определяющую потенциальную точность оценки одного из двух неэнергетических параметров волны при совместном измерении,

$$\sigma_{\nu_1}^2 = \frac{1}{(2\mathcal{J}/N_0) (2\pi\mu_{1\text{ЭК}})^2 (1 - \rho_{12}^2)}, \quad (6.75)$$

$$\sigma_{\nu_2}^2 = \frac{1}{(2\mathcal{J}/N_0) (2\pi\mu_{2\text{ЭК}})^2 (1 - \rho_{12}^2)}.$$

Сравнение полученной формулы с формулой (6.24) для одного параметра показывает, что они отличаются только множителем $1 - \rho_{12}^2$, учитывающим взаимосвязь измеряемых параметров. Так как $\rho_{12}^2 \geq 0$, корреляция между параметрами ($\rho_{12} \neq 0$) ведет к снижению точности измерений. При $a_{12} = R_{12} = \rho_{12} = 0$ матрица диагональная, измерения независимые, точность измерения максимальна и равна точности измерения одиночного параметра ($a^{-1}_{hh} = 1/a_{hh}$).

При совместном измерении двух и более параметров сигнала проявляются факторы, которые дополняют и видоизменяют закономерности, справедливые для одиночного параметра. В частности, по-иному сказывается влияние асимметрии и смещения спектра по оси частот. Рассмотрим вначале частный случай, когда взаимная связь между измеряемыми параметрами ν_1 и ν_2 отсутст-

вует, т. е. двумерный энергетический спектр является произведением одномерных спектров:

$$|S(\mu_1, \mu_2)|^2 = |S(\mu_1)|^2 |S(\mu_2)|^2. \quad (6.76)$$

В этом случае влияние асимметрии, т. е. нечетной составляющей энергетического спектра, на частоту соответствующей эквивалентной огибающей $\mu_{1\text{ЭК}}$ и $\mu_{2\text{ЭК}}$ проявляется, как и в одномерном случае, так как формулы (6.71) переходят в формулу (6.43). Однако несмотря на взаимную независимость параметров появляется коэффициент взаимной корреляции, который согласно формуле (6.72) при условии (6.76) равен произведению первых моментов μ_{10} и μ_{20} (с учетом перехода формулы (6.71) в (6.43)):

$$R_{12} = \frac{\int \mu_1 |S(\mu_1)|^2 d\mu_1}{\int |S(\mu_1)|^2 d\mu_1} \frac{\int \mu_2 |S(\mu_2)|^2 d\mu_2}{\int |S(\mu_2)|^2 d\mu_2}. \quad (7.76)$$

Таким образом, асимметрия обоих одномерных спектров, образующих двумерный спектр согласно формуле (6.76), приводит к снижению точности измерений не только из-за уменьшения $\mu_{1\text{ЭК}}$ и $\mu_{2\text{ЭК}}$ в соответствии с формулами (6.70), но и из-за множителя $1 - \rho_{12}^2$. Если же хотя бы один из двух спектров симметричен (μ_{10} (или μ_{20}) = 0), $\rho_{12} = 0$, и снижение точности вследствие второй причины не происходит. Однако, в отличие от одномерного случая, даже простое смещение симметричных спектров относительно их нулевых частот одновременно на μ_{10} и μ_{20} приводит к снижению точности из-за появления коэффициента корреляции (6.77).

В общем случае, когда двумерный энергетический спектр нельзя представить в виде произведения одномерных спектров, влияние его асимметрии на частоты эквивалентных гармонических огибающих $\mu_{1\text{ЭК}}$ и $\mu_{2\text{ЭК}}$ и коэффициент корреляции ρ_{12} , которые определяют потенциальную точность совместных измерений, существенно отличается от случая независимых параметров ν_1 и ν_2 . В этом случае влияние асимметрии спектра на частоту эквивалентной гармонической огибающей (6.70), как при измерении единственного параметра или независимых параметров, проявляется только через первые начальные моменты μ_{10} и μ_{20} , так как на вторые начальные моменты μ_{100}^2 и μ_{200}^2 асимметрия не влияет. Однако при этом смещения μ_{10} и μ_{20} значительно меньше, так как условием неотрицательности энергетического спектра $|S(\mu_1, \mu_2)|^2 \geq 0$ ограничены уже не одна, а две его нечетные составляющие в сумме. Таким образом, влияние асимметрии спектра на снижение точности совместных измерений из-за уменьшения $\mu_{1\text{ЭК}}$ и $\mu_{2\text{ЭК}}$ еще более ограничено.

Проиллюстрируем этот вывод на простейшем примере ступенчатой асимметрии двумерного спектра, одинаковой для каждой из двух обобщенных частот μ_1 и μ_2 , когда скачки интенсивности спектра располагаются на границе квадрантов (рис. 6.15, а). Относительное значение скачка равно 2κ , где κ — одинаковая в данном случае для каждой из двух частот степень асимметрии спектра, оцениваемая объемом нечетной составляющей по сравнению с общим

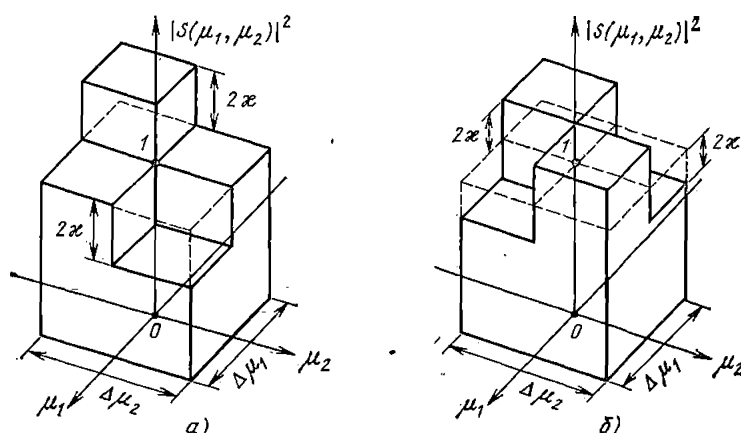


Рис. 6.15. Асимметрия двумерного обобщенного энергетического спектра: внутридиагональная (а), междиагональная (б)

объемом спектра (объемом симметричного спектра единичной интенсивности). Из условия $|S(\mu_1, \mu_2)|^2 \geq 0$ максимальное абсолютное значение отрицательного скачка не превышает $2\kappa_{\max} = 1$. Отсюда находим, что максимально возможная степень асимметрии двумерного спектра для каждой из двух частот, взятых вместе, $\kappa_{\max} = 0,5$ в отличие от одномерного спектра, где $\kappa_{\max} = 1$ (см. рис. 6.5,а и 6.5,б). Ввиду этого снижения точности измерений одновременно по двум измеряемым параметрам вследствие асимметрии κ , характеризующее, как и в одномерном случае, коэффициентом $\sqrt{1-3\kappa^2/4}$, в предельном случае ($\kappa_{\max} = 0,5$) составляет лишь 10, а не 50%. Однако снижение точности только ввиду асимметрии по одному из параметров сохраняется таким же, как и у единственного измеряемого параметра.

Необходимо отметить, что одновременное уменьшение $\mu_{1\text{эк}}$ и $\mu_{2\text{эк}}$ определяется только асимметрией спектра по двум диагональным квадрантам и не зависит от средней интенсивности спектра в каждой паре диагональных квадрантов, так как снижающие точность величины μ_{10} и μ_{20} определяются только отклонением от среднего уровня вверх и вниз в квадрантах каждой пары.

Наряду с эквивалентными частотами $\mu_{1\text{эк}}$ и $\mu_{2\text{эк}}$ точность совместных измерений согласно формуле (6.75) определяется также коэффициентом корреляции (6.73), (6.74). При вычислении коэффициента корреляции R_{12} в подынтегральном выражении числителя (6.73) спектральные составляющие суммируются со знаком «плюс» в тех диагональных квадрантах, где частоты μ_1 и μ_2 имеют одинаковый знак ($\mu_1\mu_2 > 0$), и со знаком «минус» в тех квадрантах, где у частот разные знаки ($\mu_1\mu_2 < 0$). Поэтому коэффициент корреляции (6.73) зависит только от разности средних уровней каждой пары диагональных квадрантов и не зависит от нечетно-симметричных отклонений от среднего внутри каждой диагональной пары (рис. 6.16,б).

Таким образом, влияние асимметрии спектра на эквивалентные частоты и коэффициент корреляции прямо противоположно: $R_{12} > 0$, когда средний уровень спектра выше у четной пары диагональных квадрантов, где $\mu_1\mu_2 > 0$, и $R_{12} < 0$, когда этот уровень выше у нечетной пары, где $\mu_1\mu_2 < 0$. Для приведен-

ного примера согласно формуле (6.73) $|R_{12}| = \kappa \Delta\mu_1 \Delta\mu_2 / 8$, где κ — междиagonalная асимметрия спектра. В пределе $\kappa_{\max} = 0,5$ и $|R_{12}|_{\max} = \Delta\mu_1 \Delta\mu_2 / 16$. При этом величины $\mu_{1,2 \text{ эк}} = \Delta\mu_{1,2} / 2 \sqrt{3}$ максимальны (внутри любой пары нечетная симметрия отсутствует и $\mu_{10} = \mu_{20} = 0$), а $|\rho_{12}|_{\max} = |R_{12}|_{\max} / \mu_{1 \text{ эк}} \mu_{2 \text{ эк}} = 3/4$. В результате дисперсия погрешности измерений возрастает из-за корреляции в $(1 - \rho_{12 \text{ max}}^2) = 7/16$ раз и составляет $\sigma_{\mu_{1,2}}^2 = [64(2\vartheta/N_0) \Delta\mu_{1,2}^2/3]^{-1}$. Заметим, что в другом предельном случае, когда максимальна внутридиагональная асимметрия ($\kappa_{\max} = 0,5$, $R_{12} = 0$), возрастание дисперсии погрешности измерений из-за уменьшения $\mu_{1,2 \text{ эк}}$ оценивается коэффициентом $1 - 3\kappa_{\max}^2/4 = 3/16$.

Приведенные примеры показывают, что междиagonalная асимметрия двумерного энергетического спектра проявляется (через коэффициент корреляции ρ_{12}) значительно сильнее, чем внутридиагональная (через $\mu_{1,2 \text{ эк}}$ за счет смещения центра тяжести μ_{10} и μ_{20}). Для предельных случаев асимметрии относительное влияние этих факторов на точность оценивается коэффициентом $\sqrt{13/7} = 1,35$. Одновременное появление максимальной асимметрии обоих видов взаимно исключено ввиду существования упомянутого условия $|S(\mu_1, \mu_2)|^2 \geq 0$. В реальных ситуациях асимметрия обычно незначительна. Поэтому в целом влияние асимметрии двумерного спектра на точность измерений невелико, слабее, чем для одномерного спектра. Снижение точности, вызванное случайными неконтролируемыми отклонениями реального спектра от заданной симметричной формы, пренебрежимо мало.

6.9. ПОТЕНЦИАЛЬНАЯ ТОЧНОСТЬ СОВМЕСТНОГО ИЗМЕРЕНИЯ УГЛОВЫХ КООРДИНАТ

В соответствии с общими формулами (6.70) — (6.75) потенциальная точность совместного измерения направляющих косинусов u_x и u_y определяется формулами

$$\sigma_{u_x}^2 = \frac{1}{(2\vartheta/N_0) (2\pi v_{x \text{ эк}})^2 (1 - \rho_{xy}^2)}, \quad (6.78)$$

$$\sigma_{u_y}^2 = \frac{1}{(2\vartheta/N_0) (2\pi v_{y \text{ эк}})^2 (1 - \rho_{xy}^2)},$$

где эквивалентные угловые пространственные частоты

$$v_{x \text{ эк}} = \sqrt{v_{x00}^2 - v_{x0}^2}, \quad v_{y \text{ эк}} = \sqrt{v_{y00}^2 - v_{y0}^2} \quad (6.79)$$

определяются через вторые начальные моменты

$$v_{x00}^2 = \frac{\iint v_x^2 |G(v_x, v_y)|^2 dv_x dv_y}{\iint |G(v_x, v_y)|^2 dv_x dv_y},$$

$$v_{y00}^2 = \frac{\iint v_y^2 |G(v_x, v_y)|^2 dv_x dv_y}{\iint |G(v_x, v_y)|^2 dv_x dv_y} \quad (6.80)$$

и первые начальные моменты энергетического углового спектра по осям x и y раскрыва антенны

$$\begin{aligned} v_{x0} &= \frac{\int \int v_x |G(v_x, v_y)|^2 dv_x dv_y}{\int \int |G(v_x, v_y)|^2 dv_x dv_y}, \\ v_{y0} &= \frac{\int \int v_y |G(v_x, v_y)|^2 dv_x dv_y}{\int \int |G(v_x, v_y)|^2 dv_x dv_y}, \end{aligned} \quad (6.81)$$

а также коэффициенты пространственной корреляции

$$R_{xy} = \frac{\int \int v_x v_y |G(v_x, v_y)|^2 dv_x dv_y}{\int \int |G(v_x, v_y)|^2 dv_x dv_y}, \quad \rho_{xy} = \frac{R_{xy}}{v_{x\text{ЭК}} v_{y\text{ЭК}}}. \quad (6.82)$$

Если функции раскрыва по осям x и y независимы: $G(v_x, v_y) = G(v_x)G(v_y)$, то формулы (6.78)–(6.82) переходят в формулы (6.33), (6.34) для каждой угловой координаты. Все сказанное выше относительно совместного измерения двух произвольных параметров справедливо и для угловых координат, однако необходимо ввести некоторые уточнения и ограничения.

1. Независимые функции раскрыва по осям x и y могут быть только у антенны с раскрывом в виде прямоугольника, стороны которого параллельны этим осям. Большинство же реальных антенн имеет раскрыв в виде круга или эллипса, усеченного или неусеченного. Поэтому такие метрические характеристики антенны, как $v_{x\text{ЭК}}$ и $v_{y\text{ЭК}}$, зависят от распределения поля не только по данной, но и по второй координате раскрыва. В частности, для круглого раскрыва с равномерным облучением ($G(v_x, v_y) = 1$)

$$v_{x\text{ЭК}}^2 = \int \int v_x^2 dv_x dv_y / \int \int dv_x dv_y,$$

где интегрирование ведется по всей площади круга. Переходя от декартовых координат v_x, v_y к полярным $v = \sqrt{v_x^2 + v_y^2}$ и $\varphi = \arctg(v_y/v_x)$, после замены $v_x = v \cos \varphi$ и $v_y = v \sin \varphi$ получим

$$v_{x\text{ЭК}}^2 = \frac{\int_0^{2\pi} \cos^2 \varphi d\varphi \int_0^{d/2\lambda} v^3 dv}{\pi d^2/4\lambda^2} = \frac{d^2}{16\lambda^2}, \quad (6.83)$$

т. е. $v_{x\text{ЭК}} = d/4\lambda$ (здесь у $v_{x\text{ЭК}}$ опущен индекс x , так как круглая антенна обеспечивает одинаковую точность независимо от ориентации осей координат). У антенны с равномерным прямоугольным раскрывом эквивалентная частота $v_{x\text{ЭК}} = d/2\sqrt{3}\lambda$, т. е. в $2/\sqrt{3} \approx 1,15$ раза больше, чем у круглой. В этом сказалось (весьма слабо) уменьшение размеров круглого раскрыва вдоль одной координаты при смещении от его центра по второй координате.

2. Смещение симметричной функции раскрыва относительно нулевой пространственной частоты физически нереально. Поэтому появление корреляционной связи по осям раскрыва и вызываемого ею снижения точности возможно только при междиагональной

асимметрии пространственного энергетического спектра $|G(v_x, v_y)|^2$. Однако указанная корреляционная связь устраняется простым поворотом осей раскрыва на некоторый угол. Поэтому сохраняется только влияние внутридиагональной асимметрии на точность угловых измерений.

3. Так как максимальное значение пространственного энергетического спектра $|G(v_x, v_y)|^2$, характеризующего степень отбора мощности падающей волны в каждой точке раскрыва, не может быть больше единицы, любая асимметрия функции $|G(v_x, v_y)|^2$ возможна только при понижении мощности (энергии) воспринимаемого антенной сигнала, поэтому снижение потенциальной точности, определяемой формулой (6.78), вызывается уменьшением не только эквивалентной частоты $v_{\text{эк}}$ или $v_{y \text{ эк}}$, но и сомножителя $2\mathcal{E}/N_0$. Таким образом, асимметрия пространственного энергетического спектра невыгодна с энергетической точки зрения. Энергетически выгоднее для создания диаграмм направленности специальной несимметричной формы (например, косекансных) подбор фазовых соотношений по раскрыву, а не асимметрия функции $|G(v_x, v_y)|^2$.

4. При заданных размерах антенны и длине волны почти полностью исключено повышение угловой точности сверх потенциальной, так как при этом оказывается заданной ширина углового пространственного спектра. Так называемая сверхнаправленность является операцией обмена энергетического и метрического параметров в пределах потенциальной точности, определяемой соотношением (6.78), и разрешающей способности. Необходимые для создания сверхнаправленности резкие перепады фазы поля в соседних точках раскрыва приводят к появлению уравнивающих токов по раскрыву и, следовательно, к энергетическим потерям. В отличие от этого для достижения желаемой точности временных функций ширину спектра можно сделать сколь угодно большой с помощью внутриимпульсной или частотной модуляции.

5. Энергетические и метрические характеристики антенн определяются одной и той же функцией раскрыва. Различие их состоит в том, что энергетический параметр характеризуется абсолютными размерами раскрыва, а метрический — относительными (в длинах волн).

6. При отклонении направления настройки от нормали к раскрыву диаграмма направленности антенны, отсчитываемая в реальных углах, расширяется. При этом точность угловых измерений снижается в результате как уменьшения крутизны пеленгационной характеристики, так и снижения отношения $\mathcal{E}/N_0$ вследствие уменьшения видимого раскрыва антенны.

Результатом этих особенностей и ограничений, накладываемых на пространственный энергетический спектр, является то, что реально все факторы, приводящие к снижению точности угловых измерений, практически исключаются. Единственным ограничением на форму углового пространственного спектра является требование малого уровня боковых лепестков у пространственной функ-

ции корреляции $\Psi(u'_x, u'_y)$, т. е. выходного сигнала как функции угловых координат. Это требование удовлетворяется при спадании функции раскрыва на краях, благодаря чему исключается затекание энергии за края раскрыва, которое ведет к возрастанию шумовой температуры антенны, но расширяется основной лепесток диаграммы направленности.

6.10. ПОТЕНЦИАЛЬНАЯ ТОЧНОСТЬ СОВМЕСТНОГО ИЗМЕРЕНИЯ ДАЛЬНОСТИ И РАДИАЛЬНОЙ СКОРОСТИ

Отличительной особенностью совместного измерения дальности и радиальной скорости (т. е. временного и частотного сдвигов сигнала) является то, что комплексная огибающая для частоты является спектром комплексной огибающей для времени и наоборот. Поэтому комплексные огибающие частоты и времени оказываются взаимно однозначно связанными между собой преобразованиями Фурье: поведение одной из них предопределяет поведение другой, что выражено так называемым принципом неопределенности (будет рассмотрен ниже). В частности, следствием этого принципа является жесткая взаимная зависимость точности совместного измерения дальности и радиальной скорости с помощью одного и того же сигнала. Ввиду этого временная двумерная функция корреляции (6.3) выражается через одномерные функции — сам сигнал или его спектр, в отличие от пространственной двумерной функции корреляции, где и огибающая и ее спектр двумерные.

Через временную функцию корреляции (6.3) в соответствии с общими формулами (6.70) — (6.75) определяется точность совместного измерения дальности и радиальной скорости:

$$\sigma_\tau^2 = \frac{1}{(2\theta/N_0)(2\pi f_{\text{эк}})^2 (1 - \rho_{\tau F}^2)}, \quad \sigma_F^2 = \frac{1}{(2\theta/N_0)(2\pi t_{\text{эк}})^2 (1 - \rho_{\tau F}^2)},$$

$$\rho_{\tau F} = R_{\tau F} / f_{\text{эк}} t_{\text{эк}}, \quad (6.84)$$

где $R_{\tau F}$ — ненормированный коэффициент времячастотной корреляции, зависящий от формы сигнала $S(t)$ или его спектра $S(f)$; $f_{\text{эк}}$ и $t_{\text{эк}}$ — эквивалентные частота и время, определяемые формулами (6.34).

Рассмотрим отдельно влияние амплитудной $|S(t)|$ и частотной $f(t)$ модуляции у комплексной огибающей сигнала

$$S(t) = |S(t)| \exp \left\{ j 2\pi \int_0^t f(t') dt' \right\}. \quad (6.85)$$

Функция корреляции этого сигнала

$$\Psi(\tau', F') = \frac{1}{2\theta} \int_{-\infty}^{\infty} |S(t)| |S(t - \tau')| \exp \left\{ j 2\pi \left[F' t + \int_{t-\tau'}^t f(t') dt' \right] \right\} dt, \quad (6.86)$$

так как произведение $S(t)S^*(t-\tau') \exp(j2\pi F't)$ дает в показателе экспоненты разность двух одинаковых интегралов с пределами 0, t и 0, $t-\tau'$, которая равна интегралу с пределами $t-\tau'$, t .

Продифференцируем эту функцию сначала по временному сдвигу:

$$\begin{aligned} \frac{\partial}{\partial \tau'} \Psi(\tau', F') = & -\frac{1}{2\vartheta} \int_{-\infty}^{\infty} |S(t)| \frac{\partial}{\partial \tau'} |S(t-\tau')| \times \\ & \times \exp \left\{ j2\pi \left[F' t + \int_{t-\tau'}^t f(t') dt' \right] \right\} dt + \\ & + \frac{j2\pi}{2\vartheta} \int_{-\infty}^{\infty} f(t) |S(t)| |S(t-\tau')| \exp \left\{ j2\pi \left[F' t + \int_{t-\tau'}^t f(t') dt' \right] \right\} dt, \end{aligned} \quad (6.87)$$

так как дифференцирование интеграла по нижнему пределу дает подынтегральную функцию с обратным знаком, да и сам параметр дифференцирования в нижнем пределе отрицателен. В результате повторного дифференцирования по F' находим, что вторая смешанная частная производная функции $\Psi(\tau', F')$ по времени и частоте

$$\begin{aligned} \frac{\partial^2 \Psi(\tau', F')}{\partial \tau' \partial F'} = & -j \frac{2\pi}{2\vartheta} \int_{-\infty}^{\infty} t |S(t)| \frac{\partial}{\partial \tau'} |S(t-\tau')| \times \\ & \times \exp \left\{ j2\pi \left[F' t + \int_{t-\tau'}^t f(t') dt' \right] \right\} dt - \\ & - \frac{(2\pi)^2}{2\vartheta} \int_{-\infty}^{\infty} t f(t) |S(t)| |S(t-\tau')| \times \\ & \times \exp \left\{ j2\pi \left[F' t + \int_{t-\tau'}^t f(t') dt' \right] \right\} dt \end{aligned}$$

при $\tau'=F'=0$ дает действительную составляющую $a_{12}=a_{21}=$
 $= (2\pi)^2 R_{\tau F}$, где

$$R_{\tau F} = \frac{\int_{-\infty}^{\infty} t f(t) |S(t)|^2 dt}{\int_{-\infty}^{\infty} |S(t)|^2 dt} \quad (6.88)$$

— искомая зависимость ненормированного коэффициента корреляции от формы сигнала.

Из полученной формулы вытекает, что чисто амплитудная модуляция ($f(t) \equiv 0$) и частотная модуляция в виде четной функции времени при условии симметричности $|S(f)|^2$ дают $R_{\tau F}=0$ (влияние асимметрии спектра $|S(f)|^2$ уже было оценено ранее). Следовательно, коэффициенты $R_{\tau F}$ и $\rho_{\tau F}$ не равны нулю только при

линейном, кубическом и другом нечетном законе частотной модуляции.

Таким образом, с учетом ранее полученных соотношений можно сделать вывод о том, что асимметрия законов как амплитудной, так и частотной модуляции снижает точность совместных измерений дальности и радиальной скорости.

Взяв вторую частную производную от функции (6.87) по временному сдвигу, получим

$$a_{11} = -\frac{\partial^2}{\partial \tau'^2} \Psi(\tau', F')_{\tau', F'=0} = \frac{1}{2\vartheta} \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial \tau'} |S(t)| \right]^2 dt + \\ + \frac{(2\pi)^2}{2\vartheta} \int_{-\infty}^{\infty} f^2(t) |S(t)|^2 dt = (2\pi f_{\text{ЭК}})^2.$$

Отсюда находим, что эквивалентная частота сигнала определяется суммой

$$f_{\text{ЭК}}^2 = f_{\text{ЭК ам}}^2 + f_{\text{ЭК чм}}^2, \quad (6.89)$$

где

$$f_{\text{ЭК ам}}^2 = \frac{\int_{-\infty}^{\infty} (\partial |S(t)| / \partial \tau')^2 dt}{(2\pi)^2 \int_{-\infty}^{\infty} |S(t)|^2 dt}; \quad f_{\text{ЭК чм}}^2 = \frac{\int_{-\infty}^{\infty} f^2(t) |S(t)|^2 dt}{\int_{-\infty}^{\infty} |S(t)|^2 dt} \quad (6.90)$$

— составляющие, обусловленные только амплитудной и только частотной модуляцией соответственно. В частности, для ЛЧМ сигнала с колоколообразной огибающей вида $\exp[-\pi(t^2/\tau_n^2 + jat^2)]$ (см. (3.15) и пример в § 6.3)

$$f_{\text{ЭК}}^2 = \Delta f^2 / 4\pi = (1/\tau_n^2 + a^2 \tau_n^2) / 4\pi;$$

$$f_{\text{ЭК ам}}^2 = 1/4\pi \tau_n^2, \quad f_{\text{ЭК чм}}^2 = a^2 \tau_n^2 / 4\pi.$$

Определение второй частной производной по частотному сдвигу при частотной модуляции особенностей не имеет и дает $a_{22} = -[\partial^2 \Psi(\tau', F') / \partial F'^2]_{\tau', F'=0} = (2\pi t_{\text{ЭК}})^2$, где

$$t_{\text{ЭК}}^2 = \int_{-\infty}^{\infty} t^2 |S(t)|^2 dt / \int_{-\infty}^{\infty} |S(t)|^2 dt. \quad (6.91)$$

Полученные соотношения (6.84) и (6.88)—(6.91) позволяют оценить степень времячастотной корреляции и, следовательно, точность совместного измерения дальности и скорости при любом законе частотной (фазовой) модуляции.

Рассмотрим в качестве примера комплексную огибающую ЛЧМ сигнала с $f(t) = at$. Согласно (6.88), (6.90) и (6.84) с учетом (6.89) соответственно получаем

$$R_{\tau F} = \frac{a \int_{-\infty}^{\infty} t^2 |S(t)|^2 dt}{\int_{-\infty}^{\infty} |S(t)|^2 dt} = a t_{\text{ЭК}}^2, \quad (6.92)$$

$$f_{\text{ЭК}}^2_{\text{ЧМ}} = \frac{a^2 \int_{-\infty}^{\infty} t^2 |S(t)|^2 dt}{\int_{-\infty}^{\infty} |S(t)|^2 dt} = a^2 t_{\text{ЭК}}^2,$$

$$\rho_{\tau F}^2 = \frac{a^2 t_{\text{ЭК}}^4}{f_{\text{ЭК}}^2 t_{\text{ЭК}}^2} = \frac{a^2 t_{\text{ЭК}}^2}{a^2 t_{\text{ЭК}}^2 + f_{\text{ЭК ам}}^2} = \frac{1}{1 + f_{\text{ЭК ам}}^2 / a^2 t_{\text{ЭК}}^2}. \quad (6.93)$$

Раскроем отношение $f_{\text{ЭК ам}}^2 / t_{\text{ЭК}}^2$, входящее в последнее выражение. Известно, что $f_{\text{ЭК ам}} = 1/K_1 \tau_{\text{н}}$, а $t_{\text{ЭК}} = \tau_{\text{н}}/K_2$, где коэффициенты $K_1, K_2 > 1$, а их отношение $K_1/K_2 \approx 1$ (у гауссовской огибающей $K_1 = K_2 = \sqrt{4\pi}$). Поэтому в общем случае можно положить $f_{\text{ЭК ам}}^2 / t_{\text{ЭК}}^2 \approx 1/\tau_{\text{н}}^2$, и

$$\rho_{\tau F}^2 \approx \frac{1}{1 + 1/a^2 \tau_{\text{н}}^4} \approx \frac{1}{1 + 1/N^2},$$

где $N = a\tau_{\text{н}}^2$ — база ЛЧМ сигнала (коэффициент сжатия).

Таким образом, с ростом глубины частотной модуляции a или базы сигнала N коэффициент времячастотной корреляции $\rho_{\tau F}$ очень быстро стремится к единице, что резко снижает точность совместных измерений дальности и скорости. Однако одновременно с этим возрастает $f_{\text{ЭК}}^2 = f_{\text{ЭК ам}}^2 + a^2 t_{\text{ЭК}}^2$, что ведет к возрастанию точности измерений дальности. Для оценки воздействия обоих этих факторов подставим значение коэффициента корреляции при линейной частотной модуляции (6.93) в формулы (6.84):

$$\sigma_{\tau}^2 = \frac{1}{\frac{2\theta}{N_0} (2\pi f_{\text{ЭК ам}})^2} = \frac{1}{\frac{2\theta}{N_0} (2\pi f_{\text{ЭК}})^2 \left(\frac{f_{\text{ЭК ам}}}{f_{\text{ЭК}}} \right)^2}, \quad (6.94)$$

$$\sigma_F^2 = \frac{1}{\frac{2\theta}{N_0} (2\pi t_{\text{ЭК}})^2 \left(\frac{f_{\text{ЭК ам}}}{f_{\text{ЭК}}} \right)^2}.$$

Таким образом, точность совместного измерения (оба параметра неизвестны) по сравнению с отдельным (когда один из двух параметров известен) убывает во столько раз, во сколько возрастает частота эквивалентной огибающей при частотной модуляции $f_{\text{ЭК}}$ по сравнению с эквивалентной частотой немодулированного по частоте сигнала $f_{\text{ЭК ам}}$. Если сравнить эти результаты с сигналом, не имеющим частотной модуляции ($f_{\text{ЭК}} = f_{\text{ЭК ам}}$), когда неизвестны оба параметра, то можно видеть, что точность измерения дальности при линейной частотной модуляции не возрастает

ет, а точность измерения радиальной скорости снижается в $f_{\text{эк}}/f_{\text{эк ам}}$ раз, что равно коэффициенту сжатия N .

Поэтому для одновременного повышения точности по дальности и скорости используют либо четную модуляцию частоты (например, симметрично нарастающую и спадающую линейную модуляцию), либо шумоподобные частотно- или фазомодулированные сигналы, у которых в силу их максимальной нерегулярности коэффициент времячастотной корреляции близок к нулю.

6.11. СОВМЕСТНОЕ РАЗРЕШЕНИЕ ПО ДВУМ ПАРАМЕТРАМ. РАЗРЕШАЮЩАЯ СПОСОБНОСТЬ ПО НАПРАВЛЕНИЮ

Совместная разрешающая способность по двум измеряемым параметрам, пространственным или временным, полностью описывается квадратом модуля двумерной функции корреляции $|\Psi(v'_1, v'_2)|^2$, который над плоскостью v'_1, v'_2 образует тело неопределенности с единичным пиком в центре (рис. 6.16).

Для оценки совместной разрешающей способности одним числом тело неопределенности заменяют равновеликим по объему цилиндром с единичной высотой, основанием которого является сечение тела на некотором уровне. Тогда площадь сечения

$$\Delta(v_1, v_2) = \iint_{-\infty}^{\infty} |\Psi(v'_1, v'_2)|^2 dv'_1 dv'_2, \quad (6.95)$$

численно равная объему тела и равновеликого цилиндра, является числовой характеристикой — постоянной совместного разрешения. Сечения очень удобны для графического описания разрешающей способности. В частности, положение сечений тел неопре-

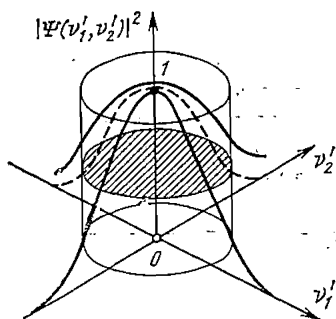


Рис. 6.16. Двумерное тело неопределенности и равновеликий цилиндр, площадь основания которого равна площади сечения (заштрихована) и объему тела

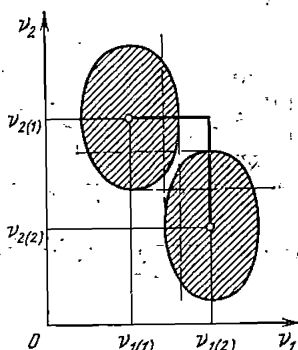


Рис. 6.17. Сечения тел неопределенности двух разрешаемых сигналов, которые не разрешаются по каждому из двух параметров в отдельности

ленности двух взаимно смещенных сигналов (рис. 6.17) показывает, что совместное разрешение возможно, даже если сигналы ни по одному из параметров в отдельности не разрешаются.

Таким образом, в отличие от точности, которая при совместном измерении нескольких неизвестных параметров по сравнению с точностью измерения единственного параметра может только ухудшаться, совместная разрешающая способность по сравнению с разрешением по каждому параметру отдельно только повышается.

Более детальное изучение совместного разрешения по двум параметрам в общем виде затруднительно. Поэтому вначале рассмотрим конкретно совместное разрешение по пространственным, а затем по временным параметрам.

Для описания совместной разрешающей способности по направлению, характеризуемому двумя независимыми угловыми координатами, используется пространственная функция корреляции (6.4). В соответствии с общей формулой (6.95) определим постоянную совместного разрешения по обоим угловым координатам

$$\Delta(u_x, u_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\Psi(u'_x, u'_y)|^2 du'_x du'_y. \quad (6.96)$$

Вычисление проведем в том же порядке, что и при определении разрешаемого интервала по одному параметру (6.48). Сначала в подынтегральной функции произведение комплексно-сопряженных двойных интегралов (6.4) по v_x и v_y заменяем четырехкратным интегралом с разделяющимися переменными v_x, v_y и v'_x, v'_y .

$$\begin{aligned} |\Psi(u'_x, u'_y)|^2 &= \Psi(u'_x, u'_y) \Psi^*(u'_x, u'_y) = \\ &= \frac{\lambda^4}{A^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(v_x, v_y)|^2 |G(v'_x, v'_y)|^2 \times \\ &\times \exp \{j 2\pi [(v_x - v'_x) u'_x + (v_y - v'_y) u'_y]\} dv_x dv_y dv'_x dv'_y. \end{aligned}$$

Интегрируя экспоненту последнего выражения согласно формуле (6.96) по переменным u'_x и u'_y , получаем дельта-функцию двух переменных $\delta(v_x - v'_x, v_y - v'_y)$, которая двойной интеграл по v'_x и v'_y от $|G(v'_x, v'_y)|^2$ обращает в функцию $|G(v_x, v_y)|^2$. Наконец, остающийся двойной интеграл дает постоянную совместного разрешения по обоим угловым координатам

$$\Delta(u_x, u_y) = \frac{\lambda^4}{A^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(v_x, v_y)|^4 dv_x dv_y = \frac{\lambda^2}{A_{\text{эф}}} = \frac{1}{\Delta(v_x, v_y)_{\text{эф}}}, \quad (6.97)$$

где

$$\Delta(v_x, v_y)_{\text{эф}} = \frac{\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(v_x, v_y)|^2 dv_x dv_y \right]^2}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(v_x, v_y)|^4 dv_x dv_y} \quad (6.98)$$

— эффективная протяженность двумерного углового пространственного спектра (относительная эффективная площадь антенны);

$$A_{\text{эф}} = \left[\iint_{-\infty}^{\infty} |G(x, y)|^2 dx dy \right]^2 / \iint_{-\infty}^{\infty} |G(x, y)|^4 dx dy \quad (6.99)$$

— эффективная по разрешению площадь антенны, несколько отличающаяся от эффективной по энергии площади антенны

$$A = \iint_{-\infty}^{\infty} |G(x, y)|^2 dx dy. \quad (6.100)$$

Метрические параметры $\Delta(v_x, v_y)$ и $A_{\text{эф}}$ характеризуют угловую разрешающую способность (направленность) антенны, эффективная площадь A является энергетическим параметром.

При равномерной функции раскрыва $|G(v_x, v_y)|^2 = 1$ антенны с физической площадью S эффективная площадь антенны $A_{\text{эф}} = A = S$ и эффективная ширина углового спектра $\Delta(v_x, v_y) = A_{\text{эф}}/\lambda^2 = A/\lambda^2 = S/\lambda^2$ максимальны и определяются его физической площадью. Однако при этом уровень боковых лепестков также максимален и происходит затекание энергии за края раскрыва. Поэтому практически используют спадающие до нуля по краям раскрыва функции, у которых $A_{\text{эф}} < S$ и $A < S$, но форма диаграммы направленности улучшается и затекание энергии отсутствует.

Если функции раскрыва по обоим осям антенны независимы, т. е. $G(v_x, v_y) = G(v_x) G(v_y)$, тогда

$$\Delta(u_x, u_y) = \Delta u_x \Delta u_y, \quad \Delta u_x = 1/\Delta v_{x \text{ эф}}, \quad \Delta u_y = 1/\Delta v_{y \text{ эф}},$$

где $\Delta(v_x, v_y)_{\text{эф}} = \Delta v_{x \text{ эф}} \Delta v_{y \text{ эф}} = (d_{x \text{ эф}}/\lambda) (d_{y \text{ эф}}/\lambda)$; $A_{\text{эф}} = d_{x \text{ эф}} d_{y \text{ эф}}$.

Нужно отметить, что при заданных значениях $\Delta(v_x, v_y)_{\text{эф}}$ или $A_{\text{эф}}$ разрешающая способность по направлению (как и точность угловых измерений) сохраняется постоянной при отсчете ее в направляющих косинусах, а не в реальных углах. В реальных углах постоянная разрешения определяется телесным углом Ω , соответствующим постоянной разрешения $\Delta(u_x, u_y)$, которая определяется из формулы (6.97). Однако телесный угол Ω при заданном $\Delta(u_x, u_y)$ зависит от многих факторов: вида апертуры (линейная, плоская, объемная), способа сканирования (механическое, электрическое) и направления настройки антенны (по нормали, под углом или вдоль раскрыва). Грубо говоря, телесный угол Ω , соответствующий постоянной разрешения $\Delta(u_x, u_y)$, есть телесный угол, занимаемый основным лепестком диаграммы направленности антенны, настроенной (сфазированной) на данное направление. Как уже отмечалось, телесный угол измеряется площадью сферы единичного радиуса, на которую он опирается. Так, полный телесный угол 4π равен всей площади сферы $S = 4\pi r^2$ при $r = 1$. Реально основной лепесток любой антенны (линейной, плоской или объемной) занимает лишь некоторый угол $\Omega < 4\pi$, в результате ослабления поля в области боковых лепестков в основном лепестке происходит усиление напряженности поля, которое

оценивается отношением $D=4\pi/\Omega$, названным коэффициентом направленного действия (КНД).

Определение телесного угла Ω , характеризующего угловую разрешающую способность при заданной постоянной разрешения $\Delta(u_x, u_y) = 1/\Delta(v_x, v_y)_{\text{эф}}$, начнем с наиболее простой, линейной антенны (рис. 6.18). Положим для определенности, что линейный раскрыв размером d расположен вдоль оси z . Эффективная протяженность углового спектра $\Delta v_{z \text{ эф}} = d_{\text{эф}}/\lambda$, где $d_{\text{эф}} \leq d$, определяется видом функции раскрыва $G(v_z)$ в соответствии с формулой (6.98). Через нее вычисляется постоянная разрешения $\Delta u_z = 1/\Delta v_{z \text{ эф}}$ по направляющему косинусу $u_z = \cos \vartheta_z$. Отложим эту постоянную на оси u_z (совпадающей с осью z , ибо u_z есть проекция единичного вектора направления на эту ось). При этом диаграмма направленности антенны в зависимости от направления настройки $u_z^0 = 0$ или $u_z^0 \neq 0$ имеет вид тарелки или воронки (рис. 6.18, а), сечения которых заштрихованы.

Телесный угол Ω , занимаемый диаграммой направленности, численно равен, как отмечалось, площади той части поверхности сферы единичного радиуса, на которую он опирается. Это — боковая поверхность шарового пояса, площадь которой $S_6 = 2\pi r h$ в нашем случае при $r=1$ и $h=\Delta u_z$ есть разрешаемый телесный угол

$$\Omega = 2\pi \Delta u_z = 2\pi / \Delta v_{z \text{ эф}} = 2\pi \lambda / d_{\text{эф}}. \quad (6.101)$$

Таким образом, разрешаемый линейной антенной телесный угол Ω , как и сам интервал Δu_z , есть величина постоянная, не зависящая от направления настройки антенны. Однако плоский угол $\Delta \alpha$, занимаемый диаграммой направленности в большом сечении сферы и численно равный длине стягивающей его дуги $l = r \Delta \alpha$ при $r=1$ (на рис. 6.18 выделено жирной линией), зависит от направления настройки $u_z^0 = \cos \vartheta_z^0 = \sin \alpha^0$: он наименьший при настройке антенны в направлении нормали к раскрыву $u_z^0 = \alpha^0 = 0$ и возрастает по мере увеличения $|u_z^0| = |\sin \alpha^0|$. В зависимости от реальных плоских углов одновременно с расширением искажается и форма диаграммы направленности, становясь асимметрич-

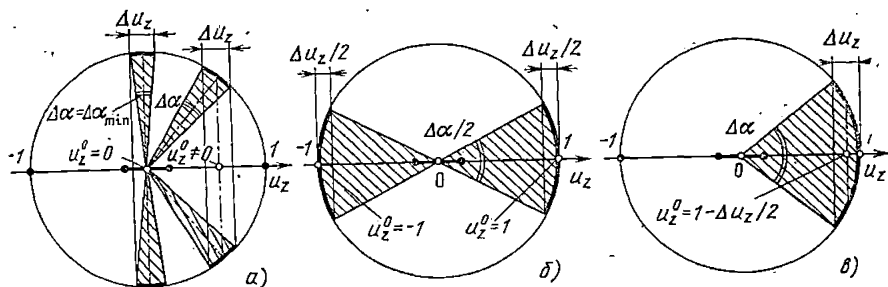


Рис. 6.18. Сечение диаграммы направленности линейной антенны, настроенной по нормали и под углом к раскрыву (а), вдоль раскрыва в обе стороны (б) и в одну сторону (в)

ной. Однако у антенн с большим относительным раскрытием $\Delta u_{\text{эф}} = d_{\text{эф}}/\lambda \gg 1$, у которых $\Delta u_z \ll 1$, искажением формы основного лепестка диаграммы направленности можно пренебречь, а учитывать только расширение углового разрешаемого интервала

$$\Delta \alpha \approx \Delta u_z / u_z^\circ = \Delta u_z / \cos \alpha^\circ = \lambda / d_{\text{эф}} \cos \alpha^\circ, \quad (6.102)$$

обратно пропорциональное косинусу угла отклонения диаграммы от нормали к раскрытию. Отклонение достигается электрическим способом — изменением угла наклона линейного распределения фазы поля по раскрытию.

Для управления фазой поля антенна выполняется в виде решетки с шагом между элементами, равными $\lambda/2$, обеспечивающим однозначность отсчета угловой координаты u_z . При возбуждении такой решетки токами с взаимным фазовым сдвигом $\psi = \pi$ получаем линейную антенну с продольной фокусировкой (рис. 6.18,б). Так как сдвиг фазы на $+\pi$ и $-\pi$ дает одинаковый эффект, антенна образует два продольных луча, настроенных на $u_z^\circ = 1$ и $u_z^\circ = -1$, а шаровой слой разбивается на два, высотой $\Delta u_z/2$ каждый. Однако общий телесный угол Ω при данном Δu_z по-прежнему определяется формулой (6.101).

Для продольного излучения (и приема) только в одном направлении необходимо антенну настроить на угол $u_z^\circ = 1 - \Delta u_z/2$ (рис. 6.18,в) или $u_z^\circ = -(1 - \Delta u_z/2)$. Это можно сделать двумя путями: либо уменьшить разность фаз питающих токов ($|\psi| < \pi$), сохранив шаг решетки равным $\lambda/2$, либо уменьшить расстояние между элементами по сравнению с $\lambda/2$, сохранив $|\psi| = \pi$. В последнем случае все элементы решетки с продольной фокусировкой, кроме одного, можно сделать пассивными, но уменьшить шаг решетки вдвое (меньше $\lambda/4$), так как пассивные элементы работают на отражение, при котором проходимый волной путь удваивается. При этом направление продольного излучения (приема) зависит от знака фазового сдвига наводимого пассивными элементами поля. Если пассивный элемент сделать короче $\lambda/4$, то наводимое поле приобретает емкостный характер и продольное излучение будет происходить в сторону пассивного элемента — директора. Пассивный элемент размерами больше $\lambda/4$ — рефлектор, поле которого имеет индуктивный характер, обеспечивает продольное излучение в сторону, противоположную его положению относительно активного элемента решетки. Рефлектор можно заменить отражающим экраном.

Линейные антенны с продольной фокусировкой обычно имеют несколько директоров и один рефлектор. Они нашли широкое применение, например, в телевизионных приемных устройствах. Однако в измерительных радиосистемах они используются сравнительно редко ввиду слабой направленности. Хотя разрешаемый телесный угол Ω , определяемый соотношением (6.101), сохраняет свое значение и в линейных антеннах с продольной фокусировкой, разрешаемый плоский угол $\Delta \alpha$ максимален. Действительно, как следует из рис. 6.18,в $\cos(\Delta \alpha/2) = 1 - \Delta u_z$. Для малых $\Delta \alpha \ll \pi/2$

справедливо разложение $\cos(\Delta\alpha/2) \approx 1 - (\Delta\alpha/2)^2/2$, поэтому последнее равенство принимает вид $1 - (\Delta\alpha/2)^2/2 \approx 1 - \Delta u_z$, откуда находим разрешаемый линейный угол антенн с продольной фокусировкой

$$\Delta\alpha \approx 2\sqrt{2\Delta u_z} = 2\sqrt{2\lambda/d_{\text{эф}}}. \quad (6.103)$$

Этот угол значительно больше, чем $\Delta\alpha \approx \lambda/d_{\text{эф}}$ у антенн с поперечной настройкой по нормали к раскрытию.

Приведенный анализ позволяет перейти к оценке разрешающей способности по направлению в реальных углах у антенн с раскрытием любого вида. Для этого рассмотрим последовательно пространственную картину, иллюстрирующую на сфере единичного радиуса угловую разрешающую способность антенн с линейным, крестообразным и плоским раскрытием (рис. 6.19).

Воронкообразное тело углового разрешения линейной антенны, в данном случае горизонтальной (рис. 6.19,а), позволяет, комбинируя его с точно таким же телом вертикальной антенны, получить картину углового разрешения для крестообразной антенны (рис. 6.19,б). Если горизонтальная антенна при этом настроена на направление u°_x (косинус угла между осью x и образующей косинуса воронки), а вертикальная на u°_y , то область перекрытия воронок образуется в направлении u°_x, u°_y . Таких областей две: одна в передней, другая в задней полусфере относительно плоскости раскрытия крестообразной антенны, причем они симметричны относительно этой плоскости. С увеличением u°_x и u°_y ширина шарового пояса и, следовательно, плоский угол $\Delta\alpha$ согласно формуле (6.102) возрастают, так как при отмеченном выше постоянстве площади поверхности шарового пояса его периметр убывает. Ввиду этого растет площадь перекрытия двух тел и численно равный ей телесный угол перекрытия.

Если теперь взять антенную решетку с плоским прямоугольным раскрытием, стороны которого соответственно равны горизонтальному и вертикальному размерам крестообразной антенны, причем настройка по каждой из двух координат, u°_x и u°_y , сохра-

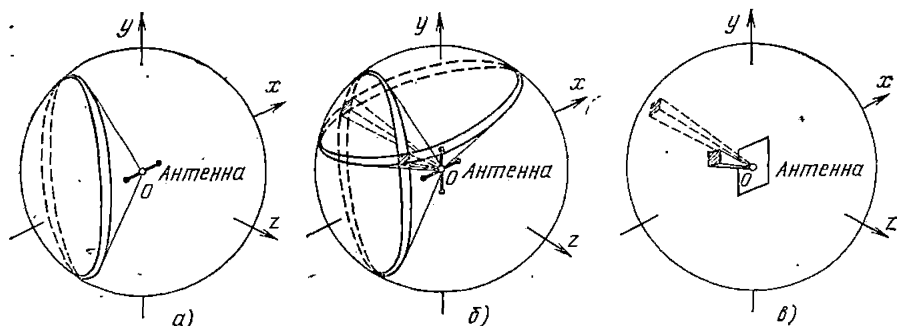


Рис. 6.19. Диаграммы направленности антенн с линейным (а), крестообразным (б) и плоским (в) раскрытием.

няется прежней, то она образует два узких луча, по размерам и положению полностью совпадающие с зонами перекрытия диаграмм направленности крестообразной антенны (рис. 6.19, в), передний и задний. Эти лучи также симметричны относительно плоскости раскрыва решетки.

Таким образом, действительно плоская антенная решетка не обеспечивает однозначного определения углового положения объектов. Однозначность достигается только в том случае, если антенна имеет конечные размеры и по третьей координате — продольной z . Как уже отмечалось, однозначность измерений любой решетки, в том числе и продольной, обеспечивается, если расстояние между ее активными элементами составляет $\lambda/2$, а при наличии пассивных элементов — примерно $\lambda/4$. Следовательно, любые антенны для обеспечения однозначности (формирования единственного луча) должны занимать некоторый объем, причем минимальный продольный размер при наличии только активных элементов вдоль него порядка $\lambda/2$, а при наличии пассивных элементов $\lambda/4$ (решетка с рефлектором). У большинства антенн (рупорных, зеркальных) продольные размеры больше минимальных. Поэтому на практике под антеннами с плоским раскрывом понимают антенны, у которых продольные размеры исключают образование заднего луча.

Телесный угол, характеризующий потенциальную разрешающую способность остронаправленной плоской антенны по направлению.

$$\Omega = \frac{\Delta(u_x, u_y)}{u_{z0}} = \frac{\Delta(u_x, u_y)}{\sqrt{1-u_{x0}^2-u_{y0}^2}} = \frac{1}{\Delta(v_x, v_y)_{\text{эф}} \sqrt{1-u_{x0}^2-u_{y0}^2}} = \frac{\lambda^2}{A_{\text{эф}} \sqrt{1-u_{x0}^2-u_{y0}^2}} \quad (6.104)$$

зависит от косинуса угла между направлением максимума луча и нормалью к раскрыву антенны $u_{z0} = \sqrt{1-u_{x0}^2-u_{y0}^2}$ и определяется видимой с этого направления эффективной по разрешению площадью антенны $A_{\text{эф}} \sqrt{1-u_{x0}^2-u_{y0}^2}$. Эффективный КНД плоской антенной решетки

$$D_{\text{эф}} = 4\pi/\Omega = 4\pi A_{\text{эф}} \sqrt{1-u_{x0}^2-u_{y0}^2}/\lambda^2$$

также зависит от смещения луча относительно нормали к раскрыву.

Наряду с антеннами, имеющими условно плоский раскрыв, продольный размер которых лишь исключает образование заднего луча, существуют поверхностные антенны сложной конфигурации и объемные. К поверхностным антеннам сложной конфигурации относятся, например, антенные решетки, элементы которых огибают поверхность объекта, где установлена станция (корабль, самолет), благодаря чему исключается искажение формы этих объектов антенной. Кроме того, сферическая антен-

ная решетка в отличие от плоской позволяет обеспечить одинаковую угловую точность и разрешающую способность во всех направлениях. Объемными считают такие антенны, у которых направленные свойства вследствие продольной фокусировки соизмеримы с направленными свойствами площади их поперечного раскрытия $A_{\perp}$, т. е. $\Delta\varphi_{z\text{эф}} \lesssim \Delta(v_x, v_y)_{\text{эф}}$ или $d_{z\text{эф}} \lesssim A_{\perp\text{эф}}/\lambda$. Постоянная углового разрешения таких антенн $\Delta(u_x, u_y) = 1/\Delta(v_x, v_y, v_z)_{\text{эф}}$ определяется эффективной протяженностью их трехмерного пространственного спектра $\Delta(v_x, v_y, v_z)_{\text{эф}}$, которая вычисляется также через двумерную угловую пространственную функцию корреляции.

При использовании поверхностных и объемных антенн двумерная угловая функция корреляции

$$\begin{aligned} \Psi(u'_x, u'_y | u_{x0}, u_{y0}) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(v_x, v_y, v_z)|^2 \times \\ &\times \exp \{ j 2\pi [v_x u'_x + v_y u'_y + v_z (\sqrt{1 - u_{x0}^2 - u_{y0}^2} - \\ &- \sqrt{1 - (u_{x0} + u'_x)^2 - (u_{y0} + u'_y)^2})] \} dv_x dv_y dv_z \times \\ &\times \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(v_x, v_y, v_z)|^2 dv_x dv_y dv_z \right]^{-1} \end{aligned} \quad (6.105)$$

зависит не только от рассогласованной u'_x, u'_y , но и от абсолютных значений u_{x0}, u_{y0} направляющих косинусов, так как набег фазы по продольной координате $2\pi v_z u_z = 2\pi v_z \sqrt{1 - u_x^2 - u_y^2}$ нелинейно связан с направляющими косинусами u_x, u_y . Ввиду этого вычисление постоянной совместного углового разрешения поверхностных и объемных антенн, а также эффективной протяженности трехмерного пространственного спектра в соответствии с формулой (6.96) в общем виде невозможно.

Объемные и сложные поверхностные антенны, как и плоские, имеющие конечную продольную протяженность, не создают зеркального луча в задней полусфере, так же как объемные голограммы не дают второго, зеркального изображения.

Приведенные выше характеристики наиболее употребительных линейных и плоских антенн отражают лишь общие свойства радиотехнических систем в отношении их угловой точности и разрешающей способности. Эти характеристики являются только первым (хотя и достаточно показательным) приближением к истинным характеристикам систем, так как основаны на ряде допущений.

Во-первых, предполагалось, что функция раскрытия антенны $G(v_x, v_y)$ и диаграмма направленности в функциях направляющих косинусов $G(u_x, u_y)$ не изменяют своей формы при изменении направления настройки. На самом деле между элементами антенной решетки существует взаимная связь, которая, к сожалению, изменяется при настройке антенны на различные направле-

ния. Чтобы функции $G(v_x, v_y)$ и $G(u_x, u_y)$ не зависели от направления настройки, необходимо либо свести к нулю эту связь, либо сделать ее независимой от направления настройки. Ни то, ни другое в полной мере реализовать не удастся, поэтому при перестройке происходят некоторые изменения диаграммы направленности, особенно в области боковых лепестков.

Во-вторых, игнорировалось то обстоятельство, что при конечных размерах раскрыва антенны (ширины углового спектра) диаграмма направленности по направляющим косинусам как преобразование Фурье от углового спектра имеет теоретически бесконечную протяженность. Однако сами направляющие косинусы реальных углов простираются в пределах $-1 \leq u_x, u_y \leq 1$. Выход диаграммы за эти пределы в область мнимых углов $|u_x, u_y| > 1$ означает появление колебательной мощности, расходуемой на тепловые потери. Обычно в эту область попадают только крайние боковые лепестки, а смещение диаграммы направленности при настройке на новое направление вызывает изменение состава боковых лепестков, попадающих в область реальных углов, т. е. существующих реально.

В-третьих, конечный продольный размер условно плоских антенн не просто устраняет задний луч, а изменяет форму диаграммы направленности во всем диапазоне углов. При сканировании степень искажения формы изменяется. Таким образом, принятое нами допущение о формировании диаграммы направленности только плоским раскрывом, без учета продольной фокусировки, и сохранении ее формы при сканировании на практике выполняется нестрого.

Однако, как следует из приведенных рассуждений, нестрогость принятых допущений относится в основном к области боковых лепестков. Отмеченные факторы практически не влияют на основной лепесток диаграммы направленности. Поэтому описанные выше общие закономерности, определяемые в основном поведением пространственных характеристик системы в области основного лепестка, можно считать достаточно точными.

6.12. СОВМЕСТНОЕ РАЗРЕШЕНИЕ ПО ДАЛЬНОСТИ И РАДИАЛЬНОЙ СКОРОСТИ. ПРИНЦИП НЕОПРЕДЕЛЕННОСТИ

Совместная потенциальная разрешающая способность по дальности и радиальной скорости, так же как и точность их совместного измерения, определяется временной функцией корреляции $R(\tau', F')$, отражающей поведение выходного сигнала оптимального приемника при расстройке по временному (τ') и частотно-сдвигам (F') одновременно. Независимо от фазы высокочастотного заполнения двух сигналов совместную разрешающую способность определяют через квадрат модуля этой функции

$|\Psi(\tau', F')|^2$ (функцию неопределенности Вудворда), который описывает поведение мощности выходного сигнала оптимального приемника при расстройках τ' и F' . Особенность совместного разрешения (и измерения) параметров τ и F состоит в том, что они являются аргументами функций $S(t-\tau)$ и $S(f-F)$, связанных между собой преобразованиями Фурье.

Совместное разрешение по дальности (времени) и радиальной скорости (частоте) характеризуется постоянной

$$\Delta(\tau, F) = \iint_{-\infty}^{\infty} |\Psi(\tau', F')|^2 d\tau' dF', \quad (6.106)$$

определяющей площадь сечения тела неопределенности, численно равную его объему и объему равновеликого цилиндра, основанием которого является эта площадь.

Аналогично определению постоянной (6.96) запишем подынтегральное выражение в формуле (6.106) как

$$|\Psi(\tau', F')|^2 = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(t) S^*(t-\tau') \times \\ \times S^*(t') S(t'-\tau') \exp[j2\pi F'(t-t')] dt dt'.$$

Подставив его в формулу (6.106) и интегрируя экспоненту по F' , получим дельта-функцию $\delta(t-t')$. Последняя при интегрировании произведения $S^*(t') S(t'-\tau')$ по t' дает $S^*(t) S(t-\tau')$. Полученный двойной интеграл с разделяющимися переменными интегрирования t и τ' заменим произведением двух интегралов, каждый из которых равен удвоенной энергии сигнала, откуда находим

$$\Delta(\tau, F) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |S(t)|^2 |S(t-\tau')|^2 dt d\tau' = \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} |S(t)|^2 dt \right]^2 = 1.$$

Итак, постоянная разрешения, т. е. объем тела неопределенности и площадь его сечения,

$$\Delta(\tau, F) = \iint_{-\infty}^{\infty} |\Psi(\tau', F')|^2 d\tau' dF' = 1. \quad (6.107)$$

Формула (6.107) представляет собой соотношение неопределенности в радиолокации: любое изменение формы сигнала $S(t)$ деформирует тело неопределенности, не изменяя его объема и высоты $|\Psi(0, 0)|^2 = 1$ в начале координат. Тело неопределенности нельзя независимо сжимать и по оси времени, и по оси частот, поскольку (как уже отмечалось) функции времени и частоты сигнала $S(t)$ и $S(f)$ связаны между собой преобразованием Фурье. Сжатие по времени неизбежно ведет к расширению тела по частоте и наоборот. Отсюда следует, что при измерении дальности и радиальной скорости с помощью общего сигнала невозможно удовлетворить сколь угодно высокие требования к точности и разрешающей способности по обоим измеряемым параметрам одновре-

менно, какой бы формы сигнал мы не применяли. В этом физическая сущность принципа неопределенности.

Принцип неопределенности, однако, отнюдь не означает, что вообще невозможно выбрать удовлетворительную форму сигнала для совместного измерения дальности и скорости. В каждой конкретной ситуации существует определенная рабочая область обзорно-измерительной системы, что позволяет подобрать сигнал такой формы, при которой большая часть объема тела неопределенности оказывается за ее пределами либо образует равномерный фон, практически не препятствующий наблюдению других сигналов.

Изучению свойств сигналов путем анализа их тела неопределенности посвящено много публикаций. Поэтому ограничимся беглым обзором свойств нескольких характерных видов сигналов по форме сечений тела неопределенности (рис. 6.20, где плотной штриховкой показана область высокой корреляции (сечение основного лепестка) тела неопределенности, редкой штриховкой — область низкой корреляции (боковые лепестки)). Для большинства сигналов (кроме шумоподобных) общая площадь сечения области высокой корреляции тела неопределенности, как и его объем, равна единице. Форму тела неопределенности шумоподобных

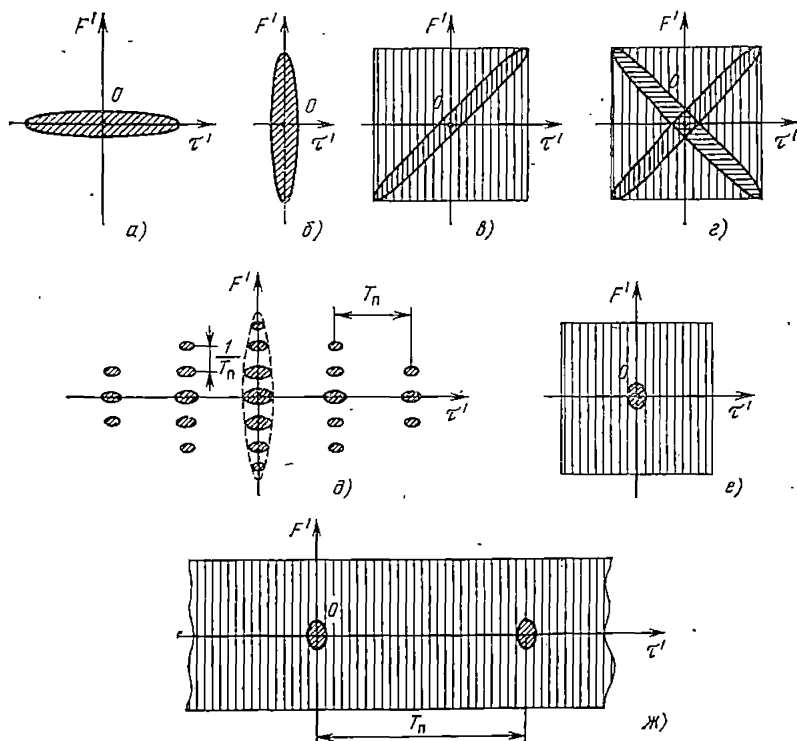


Рис. 6.20. Сечения тел неопределенности

сигналов таким сечением передать нельзя, поскольку большая часть его объема сосредоточена в области низкой корреляции.

Простые сигналы ($\Delta f_{\tau} = 1$) большой длительности (рис. 6.20,а) обеспечивают хорошее разрешение и точность по частоте (радиальной скорости), короткие импульсы — хорошее разрешение и точность по дальности, а не по скорости, так как огибающая тела по оси частот спадает медленно (рис. 6.20,б). ЛЧМ сигналы ($\Delta f_{\tau} \gg 1$) обеспечивают хорошее разрешение только по дальности или только по скорости, как и точность их измерения (рис. 6.20,в); однако при некотором сочетании сдвигов τ' и F' , когда смещение происходит вдоль «хребта» тела неопределенности, совместные разрешающая способность и точность резко снижаются: неизвестное смещение по одной координате вызывает сдвиг по другой ввиду жесткой корреляции время-частота. Симметричная (четная) частотная модуляция (как отмечалось) устраняет корреляцию и обеспечивает высокую точность совместного измерения дальности и скорости. Однако симметричная, например линейно нарастающая и спадающая, частотная модуляция дает плохую с точки зрения разрешающей способности форму тела неопределенности (рис. 6.20,г, где для наглядности показаны сечения на трех уровнях). Заметим, что во всех рассмотренных случаях совместная неопределенность (6.107) остается неизменной.

Обычно для непрерывного обзора пространства используют не одиночные импульсы, а их периодические последовательности, и от каждого объекта принимается пачка импульсов. Тело неопределенности пачки импульсов дробится (рис. 6.20,д), сохраняя неизменными объем и общую площадь сечения (штриховой линией обведена площадь сечения тела неопределенности одиночного импульса). Дробление приводит к неоднозначности измерений, когда запаздывание сигнала превышает период повторения импульсов $T_{\text{п}}$, а доплеровский сдвиг частоты выходит за пределы $\pm F_{\text{п}}/2$, и отсутствию разрешения, когда смещение двух сигналов по времени кратно периоду повторения $T_{\text{п}}$, а смещение по частоте — частоте повторения $F_{\text{п}}$. Однако период повторения обычно выбирают из условия однозначности измерения дальности $T_{\text{п}} \leq 2R_{\text{max}}/c$, при выполнении которого все пики тела неопределенности вправо и влево от оси частот не играют практической роли. Если при этом дополнительно наблюдаются малоподвижные объекты, сигналы которых имеют доплеровский сдвиг $|F| \leq F_{\text{п}}/2$, отпадает влияние и всех пиков по оси частот, кроме центрального. При указанных ограничениях совместная разрешающая способность определяется только центральным пиком, объем которого, характеризующий реальную неопределенность, много меньше единицы.

Таким образом, одним из путей преодоления принципа неопределенности является ограничение рабочей области время-частота для конкретных условий использования измерительных систем. К сожалению, при наблюдении скоростных удаленных объ-

ектов условие однозначного определения и дальности, и скорости одновременно не выполняется.

Другой способ — применение шумоподобных сигналов. Все пики тела неопределенности, кроме центрального, как например, у только что рассмотренной пачки импульсов, являются следствием упорядоченного чередования импульсов, и поэтому для «размывания» этих пиков по широкой области время-частота в виде тонкого слоя нужно придать сигналам максимальную неупорядоченность — использовать шумоподобные сигналы. Почти весь объем тела неопределенности таких сигналов (рис. 6.20,е) сосредоточен в области низкой корреляции, влияние которой на точность и разрешающую способность очень мало. Площадь области низкой корреляции (произведение ширины спектра на длительность сигнала) равна базе шумоподобного сигнала $N = \Delta f T$. Поэтому при почти единичном объеме области низкой корреляции уровень боковых лепестков по мощности (высота слоя) равен $1/N$, а по напряжению в $1/\sqrt{N}$ раз ниже пика. Объем тела центрального пика и площадь сечения области высокой корреляции также равны $1/N$. Таким образом, с увеличением базы сигнала уменьшается влияние боковых лепестков, а совместная точность и разрешающая способность пропорционально повышаются. Результирующая неопределенность зависит только от центрального пика, объем которого много меньше единицы.

Необходимо отметить существенное различие минимально достижимого уровня боковых лепестков функции неопределенности только по дальности и одновременно по дальности и/скорости. В первом случае он составляет $1/N^2$ ($1/N$ по напряжению), во втором он равен $1/N$ ($1/\sqrt{N}$ по напряжению), т. е. в N раз больше. Однако это относится к равномерно низкому уровню боковых лепестков во всем диапазоне. Если же ограничиться достаточно малой областью, то в ее пределах можно получить практически нулевой уровень боковых лепестков в обоих случаях за счет их возрастания в остальной области.

Для создания шумоподобных сигналов наиболее целесообразна частотная или фазовая модуляция при постоянной амплитуде. При этом длительность сигнала может быть сколь угодно большой — вплоть до периода повторения T_p . Тогда образуется непрерывная шумоподобная последовательность, пики корреляции которой повторяются через интервалы T_p (рис. 6.20,ж). Таким образом, при использовании шумоподобных сигналов теряется существенное различие между импульсным и непрерывным излучением.

Поскольку закон модуляции непрерывной последовательности заранее известен, при приеме ее можно демодулировать, превратив в синусоидальное колебание (свертка по частоте). По измеренной частоте этого колебания однозначно и точно определяется радиальная скорость; острые пики корреляции, интервал T_p между которыми можно выбрать сколь угодно большим, позво-

ляют точно и однозначно измерить дальность. Поэтому непрерывные шумоподобные последовательности являются единственным видом сигнала, где заданным требованиям по дальности и скорости можно удовлетворить полностью. Однако реализовать измерительные системы с такими сигналами очень трудно. Так, при работе по отраженному сигналу не удастся эффективно развязать тракты передачи и приема, поэтому принимают полумеры — разнос передатчика и приемника в пространстве, использование режима малой скважности и т. п. При работе же по сигналу прямого излучения или переизлученному (с разнесом каналов запроса и ответа по частоте путем дробно-частотного преобразования запросного сигнала) использование таких сигналов не встречает ограничений. Однако для образования свертки по частоте (демодуляции) необходимо использовать корреляционный приемник, многоканальный или отслеживающий единственный сигнал по времени.

Применение шумоподобных сигналов в радиолокации является средством борьбы с распределенными пассивными помехами — облаком отражателей, естественными (гидрометеоры, фон земли) или искусственными (дипольные помехи). Каждый отражатель облака дает на выходе приемника точно такой же сигнал, как и интересующий нас точечный объект, но смещенный по времени и частоте. Геометрическим образом мощности выходного сигнала оптимального приемника в частотно-временной области является его тело неопределенности. Тела неопределенности облака отражателей, перекрываясь, образуют суммарный сигнал фона, мешающего наблюдению интересующего нас объекта. Очевидно, что средняя мощность фона тем меньше, чем меньше число перекрывающихся сигналов, т. е. чем выше разрешающая способность по времени и частоте.

Если бы облако представляло собой совокупность отражателей, движущихся в разные стороны с большими скоростями, когда доплеровские сдвиги частоты их сигналов перекрывали бы всю ширину полосы спектра сигнала, то области низкой корреляции шумоподобных мешающих сигналов на выходе приемника образовали бы в результате взаимного перекрытия точно такой же фон, как и любой другой сигнал, простой или сложный. Таким образом, в общем случае распыление тела неопределенности шумоподобных сигналов по широкой области низкой корреляции не повышает разрешающей способности и, следовательно, помехозащищенности. Принцип неопределенности сохраняет свою силу. Однако облаков с указанными свойствами в природе не существует. Реальные облака движутся со сравнительно небольшой скоростью (скоростью ветра) и преимущественно в одном направлении. Разнос спектров таких сигналов занимает не всю область низкой корреляции, а лишь ее небольшую часть. Поэтому перекрытие выходных сигналов реальных пассивных помех происходит практически лишь по временной оси, а не по всей площади

области время-частота. Именно благодаря этому и снижается воздействие пассивных помех при использовании шумоподобных сигналов.

В заключение рассмотрим способы формирования шумоподобных сигналов, импульсных и непрерывных.

Известны два способа формирования импульсных шумоподобных сигналов: активный и пассивный. При активном способе высокочастотное несущее колебание модулируется по частоте или фазе по заданному закону. При пассивном способе для формирования излучаемого сигнала используют тот же самый фильтр, что и для оптимального сжатия принятых сигналов. Отличие состоит только в том, что вход и выход сжимающего и формирующего фильтра меняются местами, а процессы при формировании происходят в обратном порядке: на вход подается простой короткий импульс, а на выходе получают широкий модулированный импульс. Так, для получения импульса, модулированного кодом Баркера (см. рис. 2.7), к выходу секционированной линии задержки подводят прямоугольный высокочастотный импульс, длительность которого равна длительности такта, а на входе линии задержки образуется импульс заданной формы, который после усиления мощности поступает в антенну и излучается.

В отличие от импульсов для генерирования и оптимального сжатия псевдослучайных непрерывных последовательностей ввиду их большой длины фильтры не используют, а применяют только модуляцию несущей (активный способ). Наиболее часто применяется π -манипуляция фазы от одного тактового интервала к другому в псевдослучайной последовательности. Комплексную огибающую сигнала, модулированного такой последовательностью, в k -м тактовом интервале можно представить как

$$S_k(t) = \exp(j\varphi_k) \Pi(t - k\tau_c) = \exp(j\pi\xi_k) \Pi(t - k\tau_c), \quad (6.108)$$

где $\varphi_k = \pi\xi_k$ — фаза высокочастотного заполнения в k -м такте, которая принимает значение 0 или π в зависимости от двоичного числа ξ_k , принимающего значение 0 или 1 соответственно, а $\Pi(t)$ — прямоугольный импульс, амплитуду которого принимаем равной единице. Тогда комплексная амплитуда сигнала (6.108) в k -м такте S_k принимает значение 1 или -1 в соответствии с двоичным числом 0 или 1. Псевдослучайный двоичный код длиной N $\xi_N = \{\xi_1, \xi_2, \dots, \xi_N\}$ образует непрерывный шумоподобный сигнал с комплексной огибающей

$$S(t) = \sum_{k=1}^N S_k(t) = \sum_{k=1}^N \exp(j\pi\xi_k) \Pi(t - k\tau_c) = \sum_{k=1}^N (-1)^{\xi_k} \Pi(t - k\tau_c), \quad (6.109)$$

имеющей период повторения $T_N = N\tau_c$. Спектр такого сигнала (за один период повторения)

$$S(f) = \tau_c \sum_{k=1}^N (-1)^{\xi_k} e^{-j2\pi f k \tau_c} \frac{\sin \pi f \tau_c}{\pi f \tau_c}. \quad (6.110)$$

Перемножению комплексных амплитуд такого сигнала в разных тактовых интервалах соответствует суммирование двоичных чисел по модулю два:

$$S_k S_l = \exp[j(\varphi_k + \varphi_l)] = \exp[j\pi(\xi_k \oplus \xi_l)], \quad (6.111)$$

так как двойное изменение фазы на π дает нулевую фазу.

Наиболее широкое применение нашли аддитивно-циклические, или рекуррентные последовательности, обладающие свойством

$$\xi(t) \oplus \xi_N(t - k\tau_c) = \xi_N(t - k\tau_c), \quad (6.112)$$

т. е. сложение по модулю два двух взаимно смещенных одинаковых последовательностей дает такую же последовательность, но смещенную на другое число тактов. Это свойство и положено в основу формирования таких последовательностей с помощью n -разрядного регистра сдвига, в цепь обратной связи которого включен сумматор по модулю два. На один вход сумматора подается сигнал с выхода последнего (n -го разряда), а на другой — с выхода одного или нескольких предшествующих разрядов. Поскольку каждый разряд регистра может находиться в состоянии 0 или 1, число возможных состояний регистра составляет 2^n . Однако, состояние «все нули» должно быть исключено, так как регистр будет непрерывно находиться в этом состоянии и никакой последовательности не образуется. Поскольку число возможных состояний регистра сдвига не превышает $2^n - 1$, он формирует на выходе периодическую последовательность двоичных чисел. Последовательность длиной $N = 2^n - 1$ называется двоичной последовательностью максимальной длины или M -последовательностью.

M -последовательность может быть сформирована только при условии, что на второй вход сумматора по модулю два подается сигнал с выхода определенного i -го разряда регистра сдвига (или нескольких разрядов одновременно). Только при этом условии в первый разряд регистра сдвига вводится двоичное число, принадлежащее M -последовательности. Номера i указанных разрядов вычислены для всех n или N (табл. 6.2).

В качестве примера рассмотрим работу генератора M -последовательности для $n=3$ ($N=7$) совместно с модулятором, выполненным на схемах совпадений, управляемых последним разрядом регистра сдвига (рис. 6.21). В этом случае к сумматору по модулю два кроме последнего разряда регистра подключен согласно данным табл. 6.2 выход первого разряда регистра. При формировании M -последовательности начальное состояние регистра сдвига может быть любым, кроме 000. Возьмем для определенности состояние 101 (табл. 6.3).

С приходом первого тактового импульса на выходе третьего и первого разрядов регистра появляются импульсы переброса 1, которые на выходе сумматора дают 0, вводимый в первый разряд. К приходу второго тактового импульса в регистре установится состояние 010 и т. д. В течение $N=7$ тактов выходной разряд регистра выдает последовательность 1010011. В восьмом такте состояние регистра 101 вновь повторяется, что является началом нового периода M -последовательности.

Таблица 6.2

n	2	3	4	5	6	7	8
N	3	7	15	32	63	127	255
i	1	1	1	1-2-3	1-3-4	3	2-3-4

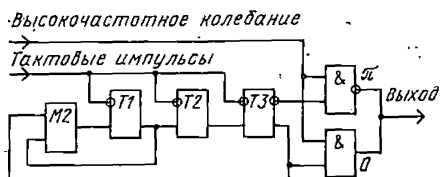


Рис. 6.21. Формирование сигнала типа M -последовательности

Таблица 6.3

Номер такта		1	2	3	4	5	6	7	8
Состояние триггера	T1	1	0	0	1	1	1	0	1
	T2	0	1	0	0	1	1	1	0
	T3	1	0	1	0	0	1	1	1
Фаза выхода		π	0	π	0	0	π	π	π

Модуляция входного высокочастотного колебания непрерывной M -последовательности происходит следующим образом. При состоянии 1 выходного триггера регистра высокочастотное колебание проходит через верхний вентиль с поворотом фазы на π , при состоянии 0 — через нижний вентиль без поворота фазы. Тем самым закон фазовой манипуляции $\pi 0 \pi 0 \pi 1$ определяется двоичной последовательностью 1010011.

Точно так же работает устройство демодуляции (свертки по частоте) принимаемой M -последовательности, совпадающей по времени с формируемой в регистре сдвига. Принимаемая последовательность подается на вентили совпадений. Перевернутый по фазе сигнал проходит через верхний вентиль с повторным перевертыванием фазы, неперевернутый проходит через нижний вентиль без поворота фазы; в результате восстанавливается несущее немодулированное колебание, частоту которого можно измерить. В самой M -последовательности несущая почти полностью подавлена, и частоту ее измерить нельзя.

С учетом свойства (6.112) аддитивно-циклических последовательностей, которое состоит в том, что перемножение взаимно сдвинутых сигналов (суммирование последовательностей по модулю два) дает точно такой же сигнал (последовательность), но с другим сдвигом, корреляционную функцию M -последовательности для $|\tau'| > \tau_c$ можно представить как

$$\begin{aligned} \Psi(\tau') &= \frac{1}{T_n} \int_0^{T_n} S(t) S(t - \tau') dt = \frac{1}{T_n} \int_0^{T_n} S(t - \tau'') dt = \\ &= \frac{1}{N} \sum_{k=1}^N (-1)^{\xi_k} = -\frac{1}{N}. \end{aligned} \quad (6.113)$$

Действительно, число нечетных k в сумме из N слагаемых (где N — всегда нечетное) на единицу больше, чем четных. Каждая пара смежных слагаемых, четного и нечетного, в сумме дает нуль. Поэтому вся сумма в формуле (6.113) равна -1 , т. е. одному нечетному слагаемому, оставшемуся вне пары. При $\tau' = 0$ сумма равна N , так как суммируются квадраты $S^2_k = 1$ независимо от знака S_k , и нормированная корреляционная функция равна 1. При $0 \leq |\tau'| \leq \tau_c$ она линейно падает от 1 до $-1/N$.

Таким образом, выходной сигнал оптимальной системы, использующей M -последовательность, при отсутствии расстройки по частоте имеет постоянный уровень $-1/N$, кроме пиков, повторяющихся через интервал $T_n = N\tau_c$. Расстройка по частоте F' приводит к появлению дополнительных, линейно нара-

стоящих фазовых сдвигов между слагаемыми. В результате пики функции $\Psi(\tau', F')$ начинают резко падать, а боковые лепестки возрастать.

Рассмотренные устройства применяют и для формирования шумоподобных импульсов в виде одного цикла M -последовательности, так как дающие минимальный уровень боковых лепестков коды Баркера известны только для некоторых $N \leq 13$. При использовании импульсов в виде цикла M -последовательности уровень боковых лепестков зависит от начального состояния регистра сдвига. Поэтому выбирают такое начальное состояние, при котором минимизируется максимальный уровень боковых лепестков (в рассмотренном выше примере именно таким является состояние 101). Этот уровень не превышает $1/\sqrt{N}$ по напряжению или $1/N$ по мощности.

6.13. ОБОБЩЕННОЕ СООТНОШЕНИЕ НЕОПРЕДЕЛЕННОСТИ. ОСОБЕННОСТИ ПРОЯВЛЕНИЯ УГЛОВОЙ СКОРОСТИ

Для описания совместной разрешающей способности более чем по двум параметрам плоскость реального пространства заменяется гиперплоскостью многомерного пространства, имеющей более двух независимых координат. Для определения постоянной совместного разрешения по четырем параметрам — дальности, радиальной скорости и двум углам — при наличии рассогласования по угловым скоростям вычислим интеграл от квадрата модуля пространственно-временной функции корреляции (6.1)

$$\Delta(\tau, F, u_x, u_y | \dot{u}'_x, \dot{u}'_y) = \\ = \int \int \int \int |\Psi(\tau', F', u'_x, u'_y; \dot{u}'_x, \dot{u}'_y)|^2 d\tau' dF' du'_x du'_y. \quad (6.114)$$

Эта постоянная равна объему и площади сечения равновеликого цилиндра, построенного на четырехмерной гиперплоскости.

Учитывая, что квадрат модуля многомерной функции корреляции (6.1) как произведение двух комплексно-сопряженных трехкратных интегралов можно заменить шестикратным интегралом с попарно разделяющимися переменными интегрирования t и t' , v_x и v'_x , v_y и v'_y , получаем

$$|\Psi(\tau', F', u'_x, u'_y; \dot{u}'_x, \dot{u}'_y)|^2 = \\ = \frac{1}{4\pi^2} \int \int \int \int |G(v_x, v_y)|^2 |G(v'_x, v'_y)|^2 S_E(t) S_E^*(t - \tau') \times \\ \times S^*(t') S(t' - \tau') \exp \{j 2\pi [F'(t - t') + (v_x - v'_x) u'_x + \\ + (v_y - v'_y) u'_y + (v_x - v'_x) \dot{u}'_x (t - t') + \\ + (v_y - v'_y) \dot{u}'_y (t - t')]\} dv_x dv'_x dv_y dv'_y dt dt'. \quad (6.115)$$

Подстановка функции (6.115) в (6.114) дает десятикратный интеграл. Выделим из него сначала экспоненциальный множитель, содержащий $t-t'$, и произведем интегрирование по F' , в результате получим дельта-функцию

$$\delta(t-t') = \int_{-\infty}^{\infty} \exp \{j 2\pi [F' + (v_x - v'_x) u'_x + (v_y - v'_y) u'_y] (t-t')\} dF',$$

так как дополнительные слагаемые у F' дают лишь сдвиг по оси частот, который при бесконечных пределах не изменяет величины интеграла. Затем выделим все сомножители, содержащие t' , интегрирование по которому дает

$$\int_{-\infty}^{\infty} S_E^*(t') S_E(t'-\tau') \delta(t-t') dt' = S_F^*(t) S_E(t-\tau'),$$

а из двукратного интеграла по t и τ' получим произведение двух одинаковых интегралов

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |S_E(t)|^2 |S_E(t-\tau')|^2 dt d\tau' = \\ & = \int_{-\infty}^{\infty} |S_E(t)|^2 dt \int_{-\infty}^{\infty} |S_E(t-\tau')|^2 d\tau' = \left[\int_{-\infty}^{\infty} |S_E(t)|^2 dt \right]^2. \end{aligned} \quad (6.116)$$

Интегрирование оставшегося экспоненциального множителя по u'_x и u'_y дает двумерную дельта-функцию

$$\begin{aligned} & \delta(v_x - v'_x, v_y - v'_y) = \\ & = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \{j 2\pi [(v_x - v'_y) u'_x + (v_y - v'_y) u'_y]\} du'_x du'_y, \end{aligned}$$

при интегрировании сомножителей, зависящих от v'_x и v'_y , получим

$$|G(v_x, v_y)|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(v'_x, v'_y)|^2 \delta(v_x - v'_x, v_y - v'_y) dv'_x dv'_y,$$

в результате остается в формуле (6.114) второй двумерный интеграл

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G(v_x, v_y)|^4 dv_x dv_y. \quad (6.117)$$

Таким образом, вычисление десятикратного интеграла дает произведение функций (6.116) и (6.117), деленных на $(2\vartheta)^2$. Так как

$$2\vartheta = \int_{-\infty}^{\infty} |G(v_x, v_y)|^2 dv_x dv_y \int_{-\infty}^{\infty} |S_E(t)|^2 dt,$$

получаем постоянную совместного разрешения по дальности, радиальной скорости и углам при рассогласовании по угловым скоростям,

$$\Delta(\tau, F, u_x, u_y | \dot{u}'_x, \dot{u}'_y) = \frac{\iint |G(v_x, v_y)|^2 dv_x dv_y [\int |S_E(t)|^2 dt]^2}{[\iint |G(v_x, v_y)|^2 dv_x dv_y]^2 [\int |S_E(t)|^2 dt]^2} = \frac{1}{\Delta(v_x, v_y)_{\text{эф}}}$$

или с учетом (6.97)

$$\Delta(\tau, F, u_x, u_y | \dot{u}'_x, \dot{u}'_y) = \Delta(\tau, F, u_x, u_y) = \Delta(u_x, u_y) = 1/\Delta(v_x, v_y)_{\text{эф}}, \quad (6.118)$$

где $\Delta(v_x, v_y)_{\text{эф}}$ определяется формулой (6.98).

Таким образом, постоянная совместного разрешения по дальности, радиальной скорости и углам при наличии рассогласования по угловым скоростям не зависит от этого рассогласования и равна постоянной разрешения только по угловым координатам. В этом сущность обобщенного соотношения неопределенности. В частном случае, при отсутствии рассогласования по угловой скорости, постоянная совместного разрешения (6.118) распадается на произведение двух постоянных, временной и пространственной:

$$\Delta(\tau, F, u_x, u_y) = \Delta(\tau, F) \Delta(u_x, u_y) = \Delta(u_x, u_y) = 1/\Delta(v_x, v_y)_{\text{эф}}, \quad (6.119)$$

так как $\Delta(\tau, F) = 1$. Хотя конечный результат в формулах (6.118) и (6.119) одинаков, поскольку временная постоянная разрешения $\Delta(\tau, F) = 1$, между ними есть существенное различие: постоянную (6.118) нельзя представить в виде произведения временной и пространственной постоянных разрешения, поскольку угловая скорость одновременно является и временным, и пространственным параметром.

Первое следствие обобщенного соотношения неопределенности (6.118) состоит в том, что если в пространстве имеется большое число равномерно распределенных по объему и скоростям отражателей, создающих мешающий фон, то единственным путем для снижения мешающего влияния этого фона является уменьшение телесного угла, занимаемого диаграммой направленности антенны, т. е. повышение угловой разрешающей способности. Действительно, общий отраженный сигнал создают те отражатели, которые попадают в разрешаемый объем пространства, ограниченный разрешаемым телесным углом $\Delta(u_x, u_y)$ и разрешаемым интервалом по дальности $\Delta R \equiv \Delta \tau$. Из всех сигналов этих отражателей только те попадают в полосу пропускания приемника, которые имеют достаточно малый разброс по радиальной скорости $\Delta R \equiv \Delta F$, т. е. по доплеровской частоте. С повышением разрешающей способности по дальности, которое достигается за счет расширения спектра сигнала, разрешаемый интервал дальности и общее число отражателей в разрешаемом объеме уменьшается,

однако вследствие расширения полосы пропускания оптимального приемника при расширении спектра сигнала-возрастает доля тех отражателей, сигналы которых попадают в полосу пропускания. Общее число отражателей, образующих сигнал фона в оптимальном приемнике, в среднем сохраняется неизменным. Однако, если отражатели неподвижны или малоподвижны, повышение разрешающей способности по дальности приводит к снижению мешающего влияния фона.

Второе следствие обобщенного соотношения неопределенности состоит в отсутствии энергетической селекции сигнала по угловым скоростям. Приведенная выше постоянная совместного разрешения по четырем параметрам (6.118) не охватывает производные направляющие косинусов, а лишь учитывает расстройку по ним. Дополнительное интегрирование квадрата модуля пространственно-временной функции корреляции по угловым скоростям показывает, что постоянная разрешения по всем шести измеряемым параметрам пропорциональна полному диапазону угловых скоростей. Это как раз и указывает на отсутствие разрешающей способности по угловым скоростям в энергетическом смысле, а не вообще, так как интеграл от квадрата модуля функции корреляции характеризует именно энергию выходного сигнала.

Действительно, падение волны на настроенное и ненастроенное по угловой скорости приемное устройство приводит к изменению фазовых соотношений в спектре и, следовательно, формы выходного сигнала, но не заключенной в нем энергии. В настроенном приемном устройстве все сигналы по раскрыву антенны суммируются синфазно, что дает выходной сигнал в виде узкого пика с максимальной амплитудой. При расстройке по угловой скорости происходит несинфазное суммирование, пик выходного сигнала расплывается, амплитуда его уменьшается, т. е. возможна амплитудная селекция по угловой скорости. Энергетическая же селекция отсутствует, так как в обоих случаях антенна перехватывает одну и ту же энергию падающей волны, на что и указывает отсутствие зависимости постоянной разрешения (6.118) от расстройки по угловой скорости.

Практическим результатом второго следствия является невозможность селекции сигналов фона от облака отражателей по угловым скоростям, поскольку ввиду случайных фазовых соотношений между сигналами множества отражателей имеет смысл говорить только об энергетической селекции.

Поясним сущность энергетической селекции по радиальной скорости и невозможность такой селекции по угловой скорости. При наличии радиальной скорости $\dot{R}$, как отмечалось ранее, частота каждой гармонической составляющей сигнала изменяется в $k=1-\dot{R}/c$ раз, в том числе частота несущей и ширина спектра принимаемого сигнала. Расширением спектра обычно пренебрегают, так как $\Delta f \ll f_0$. Следовательно, полная энергетическая селек-

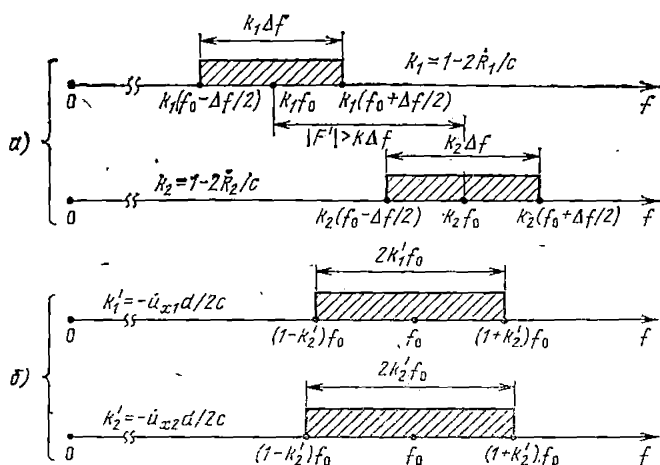


Рис. 6.22. Спектры принимаемых сигналов при различной радиальной (а) и угловой (б) скорости объектов

ция двух сигналов по радиальной скорости достигается, когда их несущие сдвинуты по доплеровской частоте на $F' > \Delta f$ (рис. 6.22,а). Когда непрерывная синусоидальная волна частоты f_0 поступает на раскрыв антенны $d = 2x_m$ от неподвижной цели, частота принимаемого сигнала в каждой точке x раскрыва одинакова и равна f_0 . При движении цели относительно центра раскрыва $x = 0$ с постоянной угловой скоростью $\dot{\omega}_x$ частота принимаемого сигнала в произвольной точке раскрыва x изменяется на величину $-\dot{\omega}_x f_0 x / c$, оставаясь равной f_0 в центре раскрыва. Бесконечно узкий спектр сигнала приобретает конечную ширину $\Delta f = |\dot{\omega}_x| d / c$ и повторяет функцию раскрыва $G(v_x = x/\lambda)$. Следовательно, угловая скорость только расширяет, но не смещает спектр (расширение очень мало и практически неощутимо у сигналов с конечной шириной спектра). Поэтому отстроиться только по угловой скорости невозможно (рис. 6.22,б).

Для оценки амплитудной селекции по угловой скорости воспользуемся пространственно-временной функцией неопределенности (6.1), положив в ней нулевыми все расстройки, кроме расстройки по угловой скорости. Так как движение объекта с постоянной угловой скоростью происходит всегда в плоскости, проходящей через его траекторию и точку расположения станции, расстройку можно брать только по производной направляющего косинуса, отсчитываемого лишь в этой плоскости. Тогда зависимость комплексной амплитуды выходного сигнала от расстройки только по производной $\dot{\omega}_x$ принимает вид

$$\Psi(\dot{\omega}_x) = \frac{1}{2\theta} \iint_{-\infty}^{\infty} |G(v_x)|^2 |S_E(t)|^2 e^{j 2\pi v_x \dot{\omega}_x t} dv_x dt. \quad (6.120)$$

Отсюда находим постоянную разрешения по угловой скорости

$$\Delta \dot{u}_x = \int_{-\infty}^{\infty} |\Psi(\dot{u}'_x)|^2 d\dot{u}'_x. \quad (6.121)$$

К сожалению, интегралы (6.120) и (6.121) в общем виде вычислить не удастся. Поэтому ограничимся частным примером, когда сигнал и распределение поля по раскрыву имеют прямоугольные огибающие:

$$S_E(t) = \begin{cases} 1, & -\tau_n/2 \leq t \leq \tau_n/2, \\ 0, & |t| > \tau_n/2; \end{cases}$$

$$G(v_x) = \begin{cases} 1, & -d/2\lambda \leq v_x \leq d/2\lambda, \\ 0, & |v_x| > d/2\lambda. \end{cases}$$

При этом функция (6.120) принимает вид (рис. 6.23)

$$\begin{aligned} \Psi(\dot{u}'_x) &= \frac{\lambda}{d\tau_n} \int_{-d/2\lambda}^{d/2\lambda} dv_x \int_{-\tau_n/2}^{\tau_n/2} \exp(j2\pi v_x \dot{u}'_x t) dt = \\ &= \frac{\lambda}{d} \int_{-d/2\lambda}^{d/2\lambda} \frac{\sin \pi v_x \dot{u}'_x \tau_n}{\pi v_x \dot{u}'_x \tau_n} dv_x = \frac{\text{Si}(\pi \dot{u}'_x \tau_n/2\lambda)}{\pi \dot{u}'_x \tau_n/2\lambda} = \frac{\text{Si}(\pi \kappa/2)}{\pi \kappa/2}, \end{aligned}$$

где $\text{Si}(z)$ — интегральный синус, а $\kappa = d\tau_n \dot{u}'_x / \lambda$ — относительное значение расстройки по угловой скорости, равное единице, когда угловая скорость такова, что за время τ_n перемещение объекта по углу равно ширине диаграммы направленности λ/d .

Численным интегрированием находим постоянную разрешения по угловой скорости (6.121)

$$\Delta \dot{u}_x = \int_{-\infty}^{\infty} \left[\frac{\text{Si}(\pi \dot{u}'_x \tau_n/2\lambda)}{\pi \dot{u}'_x \tau_n/2\lambda} \right]^2 d\dot{u}'_x = \frac{3,6}{\tau_n d / \lambda} = \frac{3,6}{\Delta t_{\text{эф}} \Delta v_{x \text{эф}}},$$

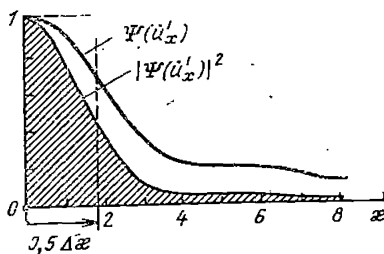


Рис. 6.23. Зависимость амплитуды и мощности выходного сигнала оптимальной системы от расстройки по угловой скорости (сигнал и функция раскрыва антенны имеют прямоугольную огибающую)

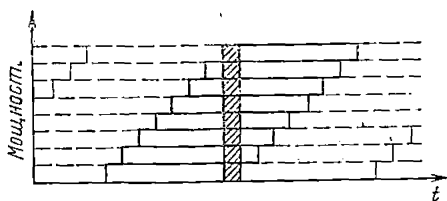


Рис. 6.24. Отсутствие селекции согласованного с фильтром сигнала (заштриховано) на фоне совокупности несфокусированных сигналов распределенных отражателей (энергия сигналов, отображаемая схематично площадью прямоугольников, одинакова)

где $\Delta t_{\text{эф}} = \tau_n$ и $\Delta v_{x \text{ эф}} = d/\lambda$ — эффективные протяженности сигнала и углового пространственного спектра соответственно.

Из приведенного примера можно видеть, что заметное снижение амплитуды выходного сигнала в согласованном пространственно-временном фильтре или корреляторе, не учитывающем расстройку по угловой скорости, достигается только при очень большой угловой скорости.

В общем случае постоянную разрешения по угловой скорости можно представить как

$$\Delta u_x = K \Delta u_x \Delta F = K / \Delta v_{x \text{ эф}} \Delta t_{\text{эф}}, \quad (6.122)$$

где $\Delta u_x = 1/\Delta v_{x \text{ эф}}$ и $\Delta F = 1/\Delta t_{\text{эф}}$ — постоянные разрешения по угловой координате и частоте сигнала соответственно; K — коэффициент, зависящий от формы сигнала и вида функции раскрыва антенны, который имеет порядок половины десятка.

Таким образом, если в пространстве наблюдается ограниченное число объектов, то возможна амплитудная селекция по угловой скорости. Если же объект наблюдается на фоне большого числа движущихся отражателей, сигналы которых суммируются (в среднем) по мощности (энергии), то селекция сигнала этого объекта по угловой скорости (подавление фона) невозможна (рис. 6.24).

Наглядным примером отсутствия энергетической селекции по угловой скорости и, следовательно, подавления сигналов фона является невозможность применения непрерывного немодулированного излучения в когерентных РЛС бокового обзора, формирующих синтезированную апертуру. В них наземные объекты в зависимости от удаления R_0 от линии пути РЛС имеют различную угловую скорость $\omega \approx V/R_0$, где V — скорость движения носителя РЛС. Ввиду этого комплексная амплитуда таких сигналов

$$S(t) = G^2(Vt/R_0) \exp(-j 2\pi V^2 t^2 / \lambda R_0),$$

т. е. закон изменения их амплитуды (фазы), зависит от отношения V/R_0 , являющегося угловой скоростью. При пропускании сигнала с данной дальности через согласованный для него фильтр образуется пик выходного сигнала, имеющий максимально возможную амплитуду, так как фазовые сдвиги окажутся скомпенсированными. Сигналы объектов, находящихся на другом удалении, дают меньший пик или вообще остаются размытыми. Следовательно, существует амплитудная селекция сигналов по дальности R_0 , т. е. угловой скорости. Однако размытые сигналы множества отражателей (фона земли) дают такую же суммарную мощность фона, какую дали бы и неразмытые сигналы объектов, не различающихся по дальности (угловой скорости). В результате сигналу объекта, находящегося на данной дальности, на выходе согласованного фильтра противостоит сигнал фона со всего диапазона дальностей, маскирующий его. Действительно, при непрерывном излучении спектр принимаемого сигнала равен спектру доплеровских частот, ширина которого, пропорциональная углу раствора диаграммы направленности бортовой антенны, не зависит от дальности. Следовательно, через согласованный фильтр в равной степени проходят спектральные составляющие сигналов с любой дальности, без подавления, т. е. энергетическая селекция отсутствует.

Глава 7. ОПТИМАЛЬНАЯ ЛИНЕЙНАЯ ФИЛЬТРАЦИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

7.1. ФИЛЬТРАЦИЯ И ЭКСТРАПОЛЯЦИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

В радиотехнике, автоматике и других областях полезным сигналом может быть любой случайный для наблюдателя процесс, который доставляет полезную информацию. Другой случайный процесс (помеха), накладываясь на этот сигнал, мешает его наблюдению. В радиотехнике роль сигнала играет высокочастотное колебание, несущее неизвестное получателю сообщение, а помехой является излучение посторонних источников. В системах автоматического регулирования сигналом является задающее воздействие, которое должно быть отработано системой, а помехой — всякое постороннее возмущающее воздействие. Вообще в любом физическом эксперименте на основной процесс накладывается постороннее возмущение, искажающее его, и задача экспериментатора — выделить (отфильтровать) полезную информацию из наблюдаемой им смеси, поскольку полезный сигнал в чистом виде недоступен для наблюдения. В такой постановке задача фильтрации затрагивает самые различные области применения.

В отличие от рассмотренной ранее задачи фильтрации сигналов известной формы с неизвестными параметрами при обработке случайных сигналов (процессов) мы имеем дело не с самим случайным процессом, а лишь с одной его конкретной реализацией, являющейся суммой реализаций двух случайных процессов — сигнала $s(t)$ и шума (помехи) $n(t)$:

$$x(t) = s(t) + n(t). \quad (7.1)$$

Поскольку форма сигнала неизвестна, в этом случае уже нельзя говорить о фильтре, оптимальном для данной реализации сигнала. Можно говорить лишь о фильтре, оптимальном в среднем по всем реализациям, т. е. оптимальном для случайного процесса в целом.

Второе отличие от фильтрации сигнала известной формы заключается в том, что необходимо произвести оценку именно сигнала или какой-либо функции от него. При фильтрации сигнала известной формы мы старались не сохранить его форму (она и так известна!), а наилучшим образом определить параметры сигнала (измерение) или установить сам факт его присутствия (обнаружение). Чтобы получить оценку $\hat{s}(t)$ — произвести фильтрацию сигнала неизвестной формы, достаточно близкую к истинному виду данной реализации сигнала $s(t)$, необходимо, чтобы между двумя случайными процессами — сигналом и шумом (в целом, а не для данных их реализаций!) — существовало какое-то

различие и это различие было известно. Иными словами, должны различаться статистические (усредненные) характеристики этих двух процессов. Наиболее полными статистическими характеристиками случайных процессов являются их многомерные функции распределения вероятностей. Однако в такой постановке задача фильтрации оказывается очень сложной. Во многих практических приложениях достаточно описание характеристик случайных процессов в рамках корреляционной теории (для гауссовских процессов такое описание является полным). Поэтому в дальнейшем будем полагать, что известны средние значения сигнала и шума, их корреляционные функции $B_s(\tau)$ и $B_n(\tau)$ и взаимокорреляционная функция $B_{sn}(\tau)$.

Задание корреляционной функции равносильно заданию энергетического спектра процесса, который указывает лишь, как распределена энергия по частотному диапазону, без учета фазовой структуры спектра, от которой зависит форма сигнала. Таким образом, задавая корреляционную функцию или энергетический спектр сигнала (и помехи), мы задаем множество реализаций, имеющих указанный общий признак процесса. Фильтрация дает полезный эффект лишь тогда, когда корреляционные функции сигнала и шума различны, т. е. их энергетические спектры не перекрываются или перекрываются лишь частично; в первом случае шум можно подавить полностью, во втором часть шума сохраняется.

Фильтрация бывает линейной и нелинейной. При линейной фильтрации оценивается сам сигнал или линейная функция от него (например, интеграл или дифференциал), при нелинейной фильтрации — процесс, являющийся нелинейной функцией входного сигнала. Именно нелинейные связи между сигналами и оцениваемыми процессами, закодированными в них, характерны для радиосистем. Однако решить задачу нелинейной фильтрации в общем виде чрезвычайно трудно. На практике в большинстве случаев нелинейную задачу фильтрации можно свести к линейной, если помеха представляет собой гауссовский процесс и отношение сигнал-шум достаточно велико. Поэтому в данной главе рассматриваются общие принципы линейной оптимальной фильтрации случайных процессов, а конкретные применения в измерительных радиосистемах, включая линеаризацию задачи нелинейной фильтрации, будут описаны в следующей главе. Первоначально задача линейной фильтрации будет рассмотрена применительно к центрированным ($\bar{s} = \bar{n} = 0$) случайным процессам, а затем постепенно усложняться.

Задача линейной оптимальной фильтрации сигнала неизвестной формы ставится следующим образом: заданы корреляционные функции сигнала и помехи; на основании их и результатов наблюдения некоторой входной реализации, представляющей смесь сигнала и помехи, требуется сформировать оценку сигнала

$$\hat{s}(t) = L[x(t)], \quad (7.2)$$

т. е. подвергнуть входную смесь (7.1) некоторому линейному функциональному преобразованию $L[x(t)]$. Эта задача решается с помощью линейного фильтра, на вход которого поступает смесь $x(t)$ (7.1), а на выходе выдается оценка сигнала $\hat{s}(t)$.

Для получения оптимальной оценки и определения оптимального линейного оператора необходимо ввести критерий оптимизации. Очевидно, что любая оценка отличается в общем случае от истинного сигнала и возникает ошибка фильтрации

$$\varepsilon(t) = \hat{s}(t) - s(t) = L[x(t)] - s(t). \quad (7.3)$$

Эта ошибка изменяется от реализации к реализации, т. е. является случайной функцией. Поэтому критерий оптимальности оператора $L[x(t)]$ должен быть некоторой статистической (усредненной) характеристикой $\varepsilon(t)$. В качестве критерия можно было бы использовать сведения к минимуму одной из следующих величин: среднего значения $M[\varepsilon(t)]$, модуля среднего значения $|M[\varepsilon(t)]|$, дисперсии $\sigma_\varepsilon^2(t) = M[\varepsilon^2(t)] - \{M[\varepsilon(t)]\}^2$ и т. п., где операция $M(\cdot)$ обозначает математическое ожидание (усреднение). Если сигнал и шум являются стационарными процессами, то указанные величины не зависят от времени. Если, кроме того, они центрированы, т. е. их средние значения равны нулю, $\sigma_\varepsilon^2 = M(\varepsilon^2)$.

Однако при использовании в качестве критерия оптимальности минимума $M(\varepsilon)$ возможны большие положительные и отрицательные ошибки, которые взаимно компенсируют друг друга. Минимум $|M(\varepsilon)|$ не имеет этого недостатка, но приводит к большим математическим трудностям. Поэтому в качестве критерия оптимальности, полагая оптимальную оценку несмещенной, т. е. $M[\varepsilon] = 0$ и $\sigma_\varepsilon^2 = M[\varepsilon^2]$, выбирают минимум дисперсии ошибки фильтрации

$$\sigma_\varepsilon^2 = M\{L[x(t)] - s(t)\}^2. \quad (7.4)$$

Если сигнал и шум — центрированные процессы, условие $M(\varepsilon) = 0$ выполняется всегда, так как при отсутствии постоянной составляющей на входе постоянная составляющая на выходе линейного фильтра появиться не может. Если же $M(s) \neq 0$ ($M(n) = 0$ практически всегда), оценка может быть и смещенной, но этот случай требует отдельного рассмотрения.

Следовательно, задачу оптимальной линейной фильтрации случайного сигнала можно сформулировать следующим образом: требуется отыскать такой вид функционального преобразования (7.2) для входного процесса, являющегося суммой двух центрированных случайных процессов с известными корреляционными функциями, при котором для данной реализации $x(t)$ достигается минимум дисперсии ошибки фильтрации (7.4).

Может возникнуть сомнение в практической ценности поставленной задачи, поскольку вряд ли корреляционные функции сигнала и помехи всегда априори известны. Но это не так, корреляционные свойства помехи можно выявить по результатам приема $x(t)$, когда сигнала заведомо нет. Что касается корреляционных

функций или энергетических спектров самих сигналов, то и они часто известны, хотя бы приближенно. При этом нужно иметь в виду, что небольшие отклонения предполагаемого спектра сигнала от истинного слабо влияют на вид функционального преобразования.

К сформулированной здесь задаче фильтрации оказывается очень близкой как по постановке, так и по методам решения задача оптимальной экстраполяции (предсказания). Формулируется она следующим образом: требуется отыскать такой вид функционального преобразования входной смеси $x(t)$, полученной за прошедшее время, включая настоящий момент, чтобы получить оценку

$$\hat{s}(t+t_s) = L[x(t)], \quad (7.5)$$

максимально близкую к будущему значению сигнала $s(t+t_s)$ в соответствии с критерием

$$\sigma_s^2 = M\{L[x(t)] - s(t+t_s)\}^2 \rightarrow \min. \quad (7.6)$$

Задача оптимальной экстраполяции решается сначала при дополнительном условии, что процесс стационарный, корреляционные связи, установленные на основании прошлого поведения сигнала, справедливы и для будущего. Таким образом, в фильтре используются связи между прошлым, настоящим и будущим значениями сигнала.

Если в формулах (7.5) и (7.6) положить время экстраполяции $t_s = 0$, приходим к задаче оптимальной фильтрации как частному случаю экстраполяции. Следовательно, фильтрация и экстраполяция случайных сигналов представляют собой единый процесс, различающийся только параметром $t_s > 0$. Так как при этом учитываются только корреляционные свойства случайных сигналов, задача оптимальной фильтрации и экстраполяции сигналов неизвестной формы в указанной постановке относится к классу линейных, поскольку корреляционные функции учитывают лишь линейную связь между прошлым и будущим. Соответственно сама задача решается с помощью линейного фильтра.

Задача оптимальной фильтрации и экстраполяции имеет очень большое значение для систем автоматического управления самого различного назначения. В частности, управление космическими объектами без предсказания их будущего положения принципиально невыполнимо, так как на прохождение измерительных радиосигналов от объекта в центр управления и командных радиосигналов из центра управления к объекту, вырабатываемых по данным измерений, затрачивается значительное время. За это время положение самого объекта может существенно измениться, и команды, вырабатываемые без учета его упрежденного положения, безнадежно устаревают.

Задача оптимальной фильтрации и экстраполяции стационарных случайных процессов была впервые решена академиком А. Н. Колмогоровым еще до второй мировой войны. Во время

войны ее решение было повторено Н. Винером уже не только в научном, но и в прикладном плане — для управления зенитной стрельбой (предсказание точки и момента встречи снаряда с самолетом после выстрела). Впоследствии решение было распространено на нестационарные, векторные процессы и т. п., включая нелинейные преобразования.

7.2. ФИЛЬТРАЦИЯ И ЭКСТРАПОЛЯЦИЯ НЕПРЕРЫВНЫХ ПРОЦЕССОВ. ИНТЕГРАЛЬНОЕ УРАВНЕНИЕ ВИНЕРА — ХОПФА

Определим импульсную характеристику $h(t)$ оптимального фильтра и экстраполятора для стационарного центрированного случайного процесса с реализацией $x(t) = s(t) + n(t)$, если заданы корреляционные функции сигнала и шума $B_s(\tau)$ и $B_n(\tau)$ и взаимокорреляционная функция сигнала и шума $B_{sn}(\tau)$, причем $M(s) = M(n) = M(x) = 0$. Пусть к настоящему моменту времени t поступила реализация входного процесса $x(t)$ (рис. 7.1). Выделим из нее элементарный сигнал, поступивший τ времени тому назад. Если δ -импульс единичной площади дает на выходе реакцию $h(t)$, то элементарный сигнал площадью $x(t-\tau)d\tau$ в момент t дает реакцию $h(\tau)x(t-\tau)d\tau$, а в момент $t+t_3$ — реакцию $h(\tau+t_3)x(t-\tau)d\tau$. Таким образом, суммарный выходной эффект от всей поступившей к моменту t реализации входного процесса в упрежденный момент времени $t+t_3$ определяется интегралом свертки

$$\hat{s}(t+t_3) = \int_0^t h(\tau+t_3)x(t-\tau)d\tau = L[x(t)]. \quad (7.7)$$

Это и есть оценка будущего значения входного сигнала. Математическое ожидание случайной ошибки фильтрации-экстраполяции $\varepsilon = \hat{s}(t+t_3) - s(t+t_3)$

$$M(\varepsilon) = \int_0^t h(\tau+t_3)M[x(t-\tau)]d\tau - M[s(t+t_3)] = 0, \quad (7.8)$$

т. е. оценка $\hat{s}(t+t_3)$ несмещенная.

Дисперсия ошибки

$$\begin{aligned} \sigma_\varepsilon^2 &= M(\varepsilon^2) = M\{L[x(t)] - s(t+t_3)\}^2 = \\ &= M\{L[x(t)]\}^2 - 2M\{L[x(t)]s(t+t_3)\} + M[s(t+t_3)]^2. \end{aligned}$$

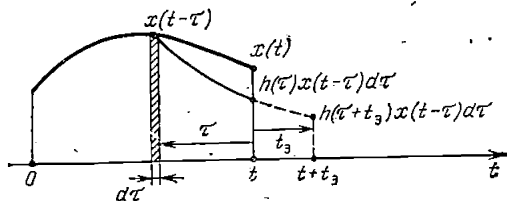


Рис. 7.1. Процессы в линейном фильтре-экстраполяторе при воздействии непрерывного входного сигнала

Вычислим слагаемые дисперсии в соответствии с формулой (7.7), заменив в первом слагаемом произведение одинаковых интегралов двойным интегралом с разделяющимися переменными:

$$M\{L[x(t)]\}^2 = \int_0^t \int_0^t h(\tau + t_0) h(\tau' + t_0) M[x(t-\tau)x(t-\tau')] d\tau d\tau' =$$

$$= \int_0^t \int_0^t h(\tau + t_0) h(\tau' + t_0) B_x(\tau - \tau') d\tau d\tau';$$

$$M\{L[x(t)]s(t+t_0)\} = \int_0^t h(\tau + t_0) M[s(t+t_0)x(t-\tau)] d\tau =$$

$$= \int_0^t h(\tau + t_0) B_{sx}(\tau + t_0) d\tau;$$

$$M[s(t+t_0)]^2 = B_s(0) = \sigma_s^2.$$

Здесь $B_x(\tau - \tau') = M[x(t-\tau)x(t-\tau')]$ — корреляционная функция входной смеси, а $B_{sx}(\tau + t_0) = M[s(t+t_0)x(t-\tau)]$ — взаимокорреляционная функция чистого сигнала и смеси сигнала с помехой, где $s(t+t_0)$ было внесено под знак интеграла, так как не зависит от переменной интегрирования τ . Ввиду стационарности процессов корреляционные функции зависят только от разности аргументов, а дисперсия (мощность) сигнала σ_s^2 не зависит от времени.

Таким образом, дисперсия ошибки фильтрации-экстраполяции

$$\sigma_e^2 = \int_0^t \int_0^t h(\tau + t_0) h(\tau' + t_0) B_x(\tau - \tau') d\tau d\tau' -$$

$$- 2 \int_0^t h(\tau + t_0) B_{sx}(\tau + t_0) d\tau + \sigma_s^2 \quad (7.9)$$

является функцией характеристики фильтра $h(t)$, т. е. $\sigma_e^2 = \sigma_s^2 [h(t)]$.

Чтобы найти оптимальную импульсную характеристику фильтра, при которой дисперсия ошибки фильтрации-экстраполяции минимальна, решим вариационную задачу. Предположим, что существует такая характеристика $h_0(t)$, при которой дисперсия σ_e^2 минимальна. Тогда характеристику произвольного фильтра можно представить в виде суммы

$$h(t) = h_0(t) + \mu \kappa(t),$$

где $\mu \kappa(t)$ — вариация (разность между произвольной и оптимальной характеристиками фильтра); μ — коэффициент; $\kappa(t)$ — некоторая функция того же класса, что и $h(t)$. При $\mu = 0$ $h(t) =$

$= h_0(t)$ и дисперсия ошибки σ_e^2 минимальна, т. е. $\{d\sigma_e^2[h(t)]/d\mu\} = 0$ или в соответствии с формулой (7.9)

$$\frac{d}{d\mu} \left\{ \int_0^t \int_0^t [h_0(\tau + t_0) + \mu \kappa(\tau + t_0)] [h_0(\tau' + t_0) + \mu \kappa(\tau' + t_0)] B_x(\tau - \tau') d\tau d\tau' - \right. \\ \left. - 2 \int_0^t [h_0(\tau + t_0) + \mu \kappa(\tau + t_0)] B_{sx}(\tau + t_0) d\tau + \sigma_s^2 \right\}_{\mu=0} = 0.$$

Взяв производную (в первом слагаемом сначала по одному, затем по другому выражению в квадратных скобках) и положив $\mu = 0$, получим

$$\int_0^t \int_0^t \kappa(\tau + t_0) h_0(\tau' + t_0) B_x(\tau - \tau') d\tau d\tau' + \\ + \int_0^t \int_0^t \kappa(\tau' + t_0) h_0(\tau + t_0) B_x(\tau - \tau') d\tau d\tau' - \\ - 2 \int_0^t \kappa(\tau + t_0) B_{sx}(\tau + t_0) d\tau = 0.$$

Первые два интеграла совершенно симметричны относительно τ и τ' , поэтому они одинаковы. Отсюда, отбросив индекс 0 у $h_0(t)$ и объединив интегралы по временной переменной τ , находим

$$\frac{d}{d\mu} \sigma_e^2[h(t)]_{\mu=0} = \\ = 2 \int_0^t \left[\int_0^t h(\tau' + t_0) B_x(\tau - \tau') d\tau' + B_{sx}(\tau + t_0) \right] \kappa(\tau + t_0) d\tau = 0. \quad (7.10)$$

Интеграл по τ с переменным пределом равен нулю тогда и только тогда, когда равно нулю само подынтегральное выражение. Приравняв последнее к нулю и сократив на $\kappa(\tau + t_0)$, в общем случае не равное нулю, приходим к интегральному уравнению Винера — Хопфа

$$\int_0^t h(\tau' + t_0) B_x(\tau - \tau') d\tau' = B_{sx}(\tau + t_0). \quad (7.11)$$

Это уравнение в неявном виде определяет оптимальную характеристику одномерного фильтра Винера $h(t)$, обеспечивающего минимум дисперсии σ_e^2 ошибки фильтрации-экстраполяции стационарного центрированного процесса.

Известно, что $B_x(\tau) = B_s(\tau) + B_n(\tau) + 2B_{sn}(\tau)$ и $B_{sx}(\tau) = B_s(\tau) + B_{sn}(\tau)$. Обычно сигнал и шум не коррелированы, т. е. $B_{sn}(\tau) = 0$. Учитывая также, что $B(\tau) = \sigma^2 \rho(\tau)$, где $\rho(\tau)$ — норми-

рованный коэффициент корреляции, получим различные модификации уравнения Винера — Хопфа:

$$\begin{aligned} \int_0^t h(\tau' + t_0) [B_s(\tau - \tau') + B_n(\tau - \tau')] d\tau' &= B_s(\tau + t_0), \\ \int_0^t h(\tau' + t_0) [\sigma_s^2 \rho_s(\tau - \tau') + \sigma_n^2 \rho_n(\tau - \tau')] d\tau' &= \sigma_s^2 \rho_s(\tau + t_0), \\ \int_0^t h(\tau' + t_0) \left[\rho_s(\tau - \tau') + \frac{\sigma_n^2}{\sigma_s^2} \rho_n(\tau - \tau') \right] d\tau' &= \rho_s(\tau + t_0), \end{aligned} \quad (7.12)$$

которые справедливы при отсутствии корреляции между сигналом и шумом.

Для общности запишем, что в случае векторного нестационарного процесса интегральное уравнение Винера — Хопфа для оптимальной характеристики системы с n входами и m выходами имеет матричный вид

$$\int_0^t \mathbf{H}(t, \tau' + t_0) B_x(t - \tau', \tau - \tau') d\tau' = B_{sv}(\tau + t_0), \quad (7.13)$$

где $\mathbf{H}$ — матрица размера $n \times m$; $\mathbf{x} = \mathbf{s} + \mathbf{n}$ — n -мерный входной вектор; $\mathbf{v}$ — m -мерный выходной вектор ($m \leq n$), причем критерием оптимальности является минимум нормы вектора ошибок $\|\epsilon^2\|$, т. е. суммы дисперсий всех его составляющих.

7.3. ФИЛЬТРАЦИЯ И ЭКСТРАПОЛЯЦИЯ ДИСКРЕТНЫХ ОТСЧЕТОВ ПРОЦЕССА

Полученные выше соотношения и выводы проявляются более наглядно, если непрерывно наблюдаемый процесс $x(t - \tau)$ заменить дискретными отсчетами $x(t - \tau_i)$ (рис. 7.2). На основании наблюдений дискретных отсчетов входного процесса требуется определить оптимальную по минимуму дисперсии оценку будущего значения сигнала $\hat{s}(t + t_0)$. Эту оценку получаем с помощью линейного оператора $L[x(t)]$:

$$\begin{aligned} \hat{s}(t + t_0) &= \sum_{i=0}^n h_i x(t - \tau_i) = \\ &= h_0 x(t) + h_1 x(t - \tau_1) + h_2 x(t - \tau_2) + \dots + h_n x(t - \tau_n). \end{aligned} \quad (7.14)$$

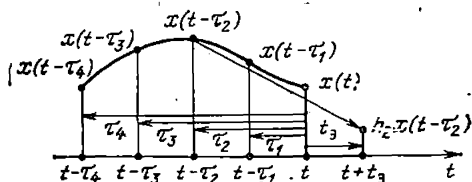


Рис. 7.2. Процессы в линейном фильтре-экстраполяторе при дискретных отсчетах входного сигнала

Таким образом, оценка является линейной комбинацией отсчетов $x(t-\tau_i)$ с весами h_i , что и говорит о линейности преобразования (при непрерывном наблюдении сумма переходит в интеграл).

Дисперсия ошибки фильтрации-экстраполяции

$$\sigma_e^2 = M \{ \hat{s}(t+t_0) - s(t+t_0) \}^2 = M \left\{ \sum_{i=0}^n h_i x(t-\tau_i) - s(t+t_0) \right\}^2. \quad (7.15)$$

Чтобы найти оптимальные весовые коэффициенты h_j , при которых дисперсия минимальна, необходимо взять $n+1$ частную производную от дисперсии ошибки по этим коэффициентам и приравнять производные к нулю:

$$\frac{\partial \sigma_e^2}{\partial h_j} = 2M \left\{ \left[\sum_{i=0}^n h_i x(t-\tau_i) - s(t+t_0) \right] \frac{\partial}{\partial h_j} \sum_{i=0}^n h_i x(t-\tau_i) \right\},$$

$$j=0, \dots, n.$$

Так как все слагаемые, кроме j -го, при дифференцировании дают нуль, производная

$$\frac{\partial}{\partial h_j} \sum_{i=0}^n h_i x(t-\tau_i) = x(t-\tau_j).$$

Подставляя полученный результат в предыдущую формулу, после переноса второго слагаемого в правую часть равенства получаем

$$\sum_{i=0}^n h_i M [x(t-\tau_i) x(t-\tau_j)] = M [s(t+t_0) x(t-\tau_j)].$$

Так как процессы стационарны,

$$M [x(t-\tau_i) x(t-\tau_j)] = B_x(\tau_i - \tau_j); \quad M [s(t+t_0) x(t-\tau_j)] = B_{sx}(\tau_j + t_0)$$

есть функции корреляции, зависящие только от разности аргументов. Отсюда получаем дискретный аналог интегрального уравнения Винера — Хопфа

$$\sum_{i=0}^n h_i B_x(\tau_i - \tau_j) = B_{sx}(\tau_j + t_0), \quad j=0, \dots, n, \quad (7.16)$$

— систему из $n+1$ уравнения для определения оптимальных значений весовых коэффициентов.

Чтобы найти минимальное значение дисперсии ошибки фильтрации-экстраполяции, преобразуем формулу (7.15):

$$\begin{aligned} \sigma_e^2 &= M \left[\sum_{i=0}^n h_i x(t-\tau_i) \right]^2 - \\ &- 2 \sum_{i=0}^n h_i M [s(t+t_0) x(t-\tau_i)] + M [s^2(t+t_0)] = \\ &= \sum_{i=0}^n \sum_{j=0}^n h_i h_j M [x(t-\tau_i) x(t-\tau_j)] - 2 \sum_{i=0}^n h_i B_{sx}(\tau_i + t_0) + B_s(0) = \\ &= \sum_{i=0}^n \sum_{j=0}^n h_i h_j B_x(\tau_i - \tau_j) - 2 \sum_{i=0}^n h_i B_{sx}(\tau_i + t_0) + B_s(0) \end{aligned}$$

(здесь произведение сумм заменено двойной суммой с разделяющимися индексами суммирования и принято обозначение $\sigma_s^2 = B_s(0)$). Подставим выражение $B_{sx}(\tau_i + t_0)$ из уравнения (7.16), поменяв индексы i и j местами:

$$\sigma_{e \min}^2 = \sum_{i=0}^n \sum_{j=0}^n h_i h_j B_x(\tau_i - \tau_j) - \\ - 2 \sum_{i=0}^n \sum_{j=0}^n h_i h_j B_x(\tau_i - \tau_j) + B_s(0).$$

Отсюда находим минимальное значение дисперсии

$$\sigma_{e \min}^2 = B_s(0) - \sum_{i=0}^n \sum_{j=0}^n h_i h_j B_x(\tau_i - \tau_j), \quad (7.17)$$

где оптимальные весовые коэффициенты определяются системой уравнений (7.16).

Проиллюстрируем эффективность оптимальных линейных фильтров и сделаем практические выводы на простейших частных примерах.

Пример 1. Наблюдение ограничивается только одним отсчетом $x(t)$. Полагаем сигнал и шум независимыми: $B_x(\tau) = B_s(\tau) + B_n(\tau)$, $B_{sx}(\tau) = B_s(\tau)$, причем $B_s(\tau) = \sigma_s^2 \rho_s(\tau)$, $B_n(\tau) = \sigma_n^2 \rho_n(\tau)$. Согласно формуле (7.14) оптимальная оценка будущего значения случайного сигнала $\hat{s}(t + t_0) = h_0 x(t)$, а уравнение (7.16) принимает вид $h_0 B_x(0) = B_s(t_0)$, откуда находим оптимальный весовой коэффициент $h_0 = B_s(t_0)/B_x(0)$ и оптимальную оценку

$$\hat{s}(t + t_0) = \frac{B_s(t_0)}{B_x(0)} x(t) = \frac{\sigma_s^2 \rho_s(t_0)}{\sigma_s^2 + \sigma_n^2} x(t). \quad (7.18)$$

При этом дисперсия ошибки минимальна и согласно формуле (7.17) равна

$$\sigma_{e \min}^2 = B_s(0) - h_0^2 B_x(0) = B_s(0) - \frac{B_s^2(t_0)}{B_x(0)} = \sigma_s^2 - \frac{\sigma_s^4 \rho_s^2(t_0)}{\sigma_s^2 + \sigma_n^2}. \quad (7.19)$$

Рассмотрим, в частности, случай «чистой» экстраполяции, когда шумами можно пренебречь ($\sigma_n^2 = 0$). Тогда из формул (7.18) и (7.19) получим

$$\hat{s}(t + t_0) = \rho_s(t_0) s(t); \quad \sigma_{e \min}^2 = \sigma_s^2 [1 - \rho_s^2(t_0)].$$

Ошибка экстраполяции возрастает с увеличением дисперсии сигнала и уменьшается с возрастанием корреляции между прошедшим и будущим значениями. Иными словами, предсказание тем точнее, чем медленнее протекает процесс.

Покажем, что это предсказание оптимально. На первый взгляд кажется, что поскольку, имея всего один отсчет, мы не знаем, как будет изменяться процесс — в сторону возрастания или убывания, то в качестве будущего следовало бы принять текущее значение, т. е. положить $\hat{s}(t + t_0) = s(t)$. При этом мы допускаем ошибку, дисперсия которой

$$\sigma_e^2 = M[\hat{s}(t + t_0) - s(t + t_0)]^2 = M[s(t) - s(t + t_0)]^2 = \\ = 2\sigma_s^2 - 2\sigma_s^2 \rho_s(t_0) = 2\sigma_s^2 [1 - \rho_s(t_0)].$$

Отношение

$$\frac{\sigma_e^2}{\sigma_{e \min}^2} = \frac{2}{1 + \rho_s(t_0)} \geq 1,$$

т. е. произвольное, решение $\hat{s}(t+t_0)=s(t)$ неоптимально. Оптимальное решение $\hat{s}(t+t_0)=\rho_s(t_0)\hat{s}(t)$ всегда занижено, так как малые значения сигнала при нормальном распределении с нулевым математическим ожиданием более вероятны.

При фильтрации ($t_0=0$) согласно формуле (7.18) оптимальная оценка

$$\hat{s}(t) = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} x(t),$$

т. е. чем больше отношение сигнал-шум σ_s^2/σ_n^2 , тем выше оценка сигнала. Дисперсия ошибки оптимальной фильтрации согласно формуле (7.19)

$$\sigma_{e \min}^2 = \sigma_s^2 - \frac{\sigma_s^4}{\sigma_s^2 + \sigma_n^2} = \sigma_n^2 \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2}.$$

Если же принять произвольное решение $\hat{s}(t)=x(t)$, то

$$\sigma_e^2 = M[x(t) - s(t)]^2 = M[n^2(t)] = \sigma_n^2.$$

Отношение дисперсий

$$\frac{\sigma_e^2}{\sigma_{e \min}^2} = \frac{\sigma_s^2 + \sigma_n^2}{\sigma_s^2} = 1 + \frac{\sigma_n^2}{\sigma_s^2} \geq 1,$$

т. е. выигрыш при оптимальном решении тем больше, чем сильнее шум.

Приведенный пример наглядно показывает преимущества оптимальной фильтрации и экстраполяции, хотя количественно выигрыш невелик; поскольку для прогнозирования использовался лишь единственный текущий отсчет. Результаты прогнозирования, естественно, улучшаются, если учитывать и прошлые отсчеты входного процесса.

Пример 2. Фильтрация и экстраполяция производится по двум отсчетам: текущему $x(t)$ и прошлому $x(t-\tau)$. Сигнал и шум по-прежнему полагаем некоррелированными. При этом оценка (7.14) принимает вид

$$\hat{s}(t+t_0) = h_0 x(t) + h_1 x(t-\tau), \quad (7.20)$$

где оптимальные коэффициенты определяются из системы (7.16):

$$h_0 B_x(0) + h_1 B_x(\tau) = B_s(t_0),$$

$$h_0 B_x(\tau) + h_1 B_x(0) = B_s(\tau+t_0).$$

Решение дает

$$\begin{aligned} h_0 &= \frac{B_s(t_0) B_x(0) - B_s(\tau+t_0) B_x(\tau)}{B_x^2(0) - B_x^2(\tau)}, \\ h_1 &= \frac{B_s(\tau+t_0) B_x(0) - B_s(t_0) B_x(\tau)}{B_x^2(0) - B_x^2(\tau)}. \end{aligned} \quad (7.21)$$

Можно видеть, что поправка, вносимая прошлым отсчетом $x(t-\tau)$, существенна только в том случае, если это прошлое не очень отдаленное, когда

корреляционные связи еще сохраняются. Если же τ велико, то всеми величинами, кроме $B_x(0)$ и $B_x(t_0)$, в последних формулах можно пренебречь. Тогда $h_0 \approx B_x(t_0)/B_x(0)$, $h \approx 0$, т. е. получаем такой же результат, как и при прогнозировании по единственному текущему отсчету $x(t)$.

Аналогично можно получить указанные зависимости при любом числе отсчетов. Однако уже приведенных здесь зависимостей достаточно для качественной оценки вклада, вносимого отдельными отсчетами.

В качестве примера практического использования линейного сглаживания и предсказания рассмотрим экономичную систему связи, поскольку в измерительных радиосистемах обычно используют нелинейные фильтры, которые рассмотрены в гл. 8. В обычных системах связи, где закон изменения амплитуды сигнала передается сообщением $s(t)$, средняя мощность передаваемого сигнала $P_s = \sigma_s^2$. Если же применять систему с оптимальной экстраполяцией данных на передающем и приемном концах канала связи, то в качестве сигнала достаточно передавать только ошибку экстраполяции; все остальное можно воспроизвести, не используя линии связи. Средняя мощность передаваемого сигнала в оптимальной системе с экстраполяцией определяется минимальной дисперсией ошибки экстраполяции (7.17)

$$P_e = \sigma_e^2 \min = \sigma_s^2 - \sigma_s^2 \sum_{i=0}^n \sum_{j=0}^n h_i h_j \rho_s(\tau_i - \tau_j). \quad (7.22)$$

Здесь второе слагаемое как раз и показывает выигрыш в мощности передаваемого сигнала по сравнению с линией связи без предсказания.

В канале связи с экстраполяцией (рис. 7.3) сообщение $s(t)$, подлежащее передаче в момент t , сравнивается с выходным сигналом экстраполятора $\hat{s}(t)$, выработанным на основе знания ранее переданных значений сигнала $s(t - \tau_i)$. Снимаемая с выхода схемы вычитания ошибка экстраполяции $e(t) = \hat{s}(t) - s(t)$ передается по каналу связи. На приемном конце принятый сигнал $e(t)$ добавляется к выходному сигналу экстраполятора, в результате на выходе приемного устройства восстанавливается действительное значение сигнала $s(t) = \hat{s}(t) + e(t)$. При постоянном интервале между выборками фильтр-экстраполятор имеет вид классического линейного фильтра на устройствах памяти с равноотстоящими отводами и оптимальными весовыми коэффициентами. С выхода суммирующей шины фильтр снимается предсказанное значение сигнала.

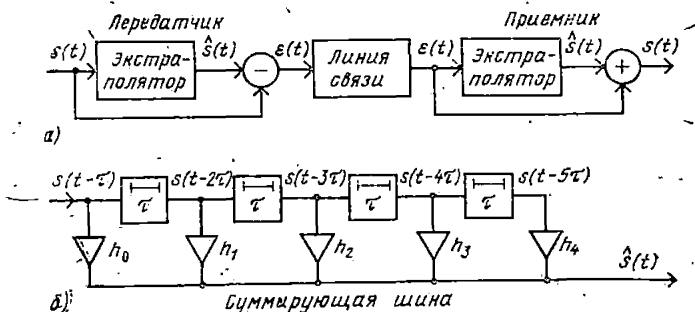


Рис. 7.3 Система связи с предсказанием (а) и структура фильтра-экстраполятора (б)

Рассмотрим в качестве примера систему связи с оптимальным чистым предсказанием по одному отсчету. Как было показано в примере 1, дисперсия ошибки экстраполяции, а следовательно, и мощность передаваемого в оптимальном канале связи сигнала

$$P_e = \sigma_s^2 [1 - \rho_s^2(\tau)] = P_s [1 - \rho_s^2(\tau)],$$

где P_s — мощность канала связи без предсказания. Если же использовать не оптимальное предсказание $\hat{s}(t) = \rho_s(\tau)s(t-\tau)$, а просто предыдущий отсчет в качестве оценки, т. е. $\hat{s}(t) = s(t-\tau)$, то получим широкоизвестную систему с дельта-модуляцией, когда в очередном такте по каналу связи передается разность текущего и предыдущего отсчетов $\delta(t) = s(t) - s(t-\tau)$. Это есть случай неоптимального чистого предсказания, рассмотренный в примере 1. В этом случае, как было показано, дисперсия ошибки экстраполяции и, следовательно, мощность передаваемого по линии связи сигнала

$$P_\delta = 2\sigma_s^2 [1 - \rho_s(\tau)] = 2P_s [1 - \rho_s(\tau)]$$

больше, чем при использовании оптимального предсказания (здесь τ — интервал между соседними отсчетами, равный времени упреждения $t_0 = \tau$).

7.4. РАСЧЕТ ПАРАМЕТРОВ ФИЛЬТРА ВИНЕРА

Для определения оптимальных параметров фильтра — весовых коэффициентов h_j — необходимо, как отмечалось, решить систему уравнений (7.16). Коэффициенты левой части этой системы образуют матрицу размера $(n+1) \times (n+1)$, а правая часть — вектор-столбец:

$$\begin{bmatrix} B_x(0) & B_x(\tau_1) & \dots & B_x(\tau_j) & \dots & B_x(\tau_n) \\ B_x(-\tau_1) & B_x(0) & \dots & B_x(\tau_j - \tau_1) & \dots & B_x(\tau_n - \tau_1) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ B_x(-\tau_i) & B_x(\tau_1 - \tau_i) & \dots & B_x(\tau_j - \tau_i) & \dots & B_x(\tau_n - \tau_i) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ B_x(-\tau_n) & B_x(\tau_1 - \tau_n) & \dots & B_x(\tau_j - \tau_n) & \dots & B_x(0) \end{bmatrix} \begin{bmatrix} B_{sx}(t_0) \\ B_{sx}(\tau_1 + t_0) \\ \dots \\ B_{sx}(\tau_i + t_0) \\ \dots \\ B_{sx}(\tau_n + t_0) \end{bmatrix}$$

Для этого находим определитель D матрицы и определители D_j матриц, полученных из данной матрицы заменой j -го столбца вектор-столбцом правой части. После этого вычислим сами коэффициенты

$$h_j = D_j / D, \quad j = 0, \dots, n. \quad (7.23)$$

Таким образом, нахождение параметров дискретного фильтра Винера — процедура несложная, но при больших n довольно громоздкая, причем возможны большие отклонения.

Сложнее определить характеристику оптимального фильтра для непрерывного сигнала. Воспользуемся интегральным уравнением Винера — Хопфа (7.11) и возьмем от него преобразование Фурье

$$\int_0^t h(\tau' + t_0) \left[\int_{-\infty}^{\infty} B_x(\tau - \tau') e^{-j 2\pi f \tau} d\tau \right] d\tau' = \\ = \int_{-\infty}^{\infty} B_{sx}(\tau + t_0) e^{-j 2\pi f \tau} d\tau.$$

Если $B(\tau) \stackrel{\mathcal{F}}{\Leftrightarrow} B(f)$, то согласно теореме смещения $B_x(\tau - \tau') \stackrel{\mathcal{F}}{\Leftrightarrow} B_x(f) \exp(-j 2\pi f \tau')$ и $B_{sx}(\tau + t_0) \stackrel{\mathcal{F}}{\Leftrightarrow} B_{sx}(f) \exp(j 2\pi f t_0)$, где $B(f)$ — энергетический спектр процесса, связанный с функцией корреляции $B(\tau)$ преобразованиями Фурье. С учетом этого последняя формула принимает вид

$$B_x(f) \int_0^t h(\tau' + t_0) e^{-j 2\pi f \tau'} d\tau' = B_{sx}(f) e^{j 2\pi f t_0}.$$

Если бы интеграл имел бесконечный верхний предел, он представлял бы собой функцию передачи

$$H_\Phi(f) = [B_{sx}(f)/B_x(f)] e^{j 2\pi f t_0}. \quad (7.24)$$

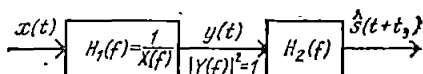
Это — функция передачи некоего фиктивного фильтра, учитывающего и будущие значения входного процесса, которые поступают после текущего момента t , что нереально. От функции (7.24) через преобразование Фурье можно получить импульсную характеристику фиктивного фильтра $h_\Phi(\tau)$. Необходимо отметить, что если бы характеристика $h_\Phi(\tau)$ для прошедших и будущих значений входного процесса выбиралась независимо, то можно было бы определить характеристику реального фильтра Винера как

$$h(\tau) = \begin{cases} h_\Phi(\tau), & \tau \geq 0, \\ 0, & \tau < 0 \end{cases} \quad (7.25)$$

(смещение $\tau < 0$ соответствует будущим неизвестным значениям входного процесса $x(t)$). Чтобы будущие неизвестные значения каким-то образом учесть, лучше всего положить их равными $M(x) = 0$, что и учтено формулой (7.25).

Однако у реального фильтруемого процесса прошедшие и будущие значения коррелированы (некоррелированный сигнал $s(t)$ отфильтровать от шумов невозможно). Поэтому отбрасывание слагаемых, соответствующих будущим значениям, т. е. принятие для них веса $h=0$, повлияет и на выбор весовых коэффициентов для прошедших значений. В этом можно убедиться, рассмотрев, например, уравнение (7.16), согласно которому значения h_j зависят от всех значений h_i . Возникает противоречие: метод, описываемый формулой (7.25), пригоден лишь для некоррелированных процессов, в то время как все фильтруемые процессы коррелированы и к ним он неприменим. Выход из этого положения был найден Боде и Шенноном, которые предложили разбить фильтр на две части (рис. 7.4). В первом, обесцараживающем фильтре входной процесс $x(t)$ декоррелируется, принимая вид белого шума, что по-

Рис. 7.4. Представление фильтра Винера в виде последовательности двух фильтров: обеляющего и формирующего



звоялет применить ко второму фильтру, где производится формирование оценки сигнала, изложенные выше идеи.

Представим функцию передачи фиктивного фильтра (7.24) в виде произведения функций передачи двух названных выше фильтров:

$$H_{\Phi}(f) = H_1(f) H_{2\Phi}(f), \quad (7.26)$$

где первый фильтр реализуемый. Представим энергетический спектр входного процесса как $B_x(f) = \tilde{X}(f) X^*(f)$ и выберем функции передачи первого фильтра

$$H_1(f) = 1/X(f). \quad (7.27)$$

Энергетический спектр выходного сигнала первого фильтра

$$B_y(f) = B_x(f) |H_1(f)|^2 = B_x(f)/B_x(f) = 1,$$

т. е. выходной сигнал его представляет собой белый шум, что и требовалось.

Физически нереализуемая характеристика второго фильтра согласно (7.26) и (7.24) имеет вид

$$H_{2\Phi}(f) = \frac{H(f)}{H_1(f)} = \frac{B_s(f)}{B_x(f)} X(f) e^{j 2\pi f t_0} = \frac{B_s(f)}{X^*(f)} e^{j 2\pi f t_0}. \quad (7.28)$$

Для входного процесса этого фильтра — белого шума — вместо будущих значений $y(t)$ можно принять $M(y) = 0$. При этом импульсная характеристика реального фильтра

$$h_2(\tau) = \begin{cases} h_{2\Phi}(\tau), & \tau \geq 0, \\ 0, & \tau < 0, \end{cases} \quad (7.29)$$

где $h_{2\Phi}(\tau) \Leftrightarrow H_{2\Phi}(f)$. После этого находим $\tilde{H}_2(f) \Leftrightarrow h_2(\tau)$ и характеристику всего фильтра

$$H(f) = H_1(f) H_2(f). \quad (7.30)$$

К сожалению, интегральные уравнения Винера — Хопфа в большинстве случаев не имеют аналитического решения. Решают их с помощью ЭВМ.

В заключение для усвоения изложенной методики рассмотрим один из немногих примеров аналитического расчета характеристики фильтра Винера. Расчет проводим в следующем порядке.

1. Задаем вид функции корреляции $B_x(\tau) = B_s(\tau) + B_n(\tau)$ или энергетического спектра $B_x(f) = B_s(f) + B_n(f)$, где $B_s(f) = 2\alpha\sigma_s^2/[\alpha^2 + (2\pi f)^2]$, $B_n(f) = N_0/2$;

$$B_x(f) = \frac{N_0}{2} + \frac{2\alpha\sigma_s^2}{\alpha^2 + (2\pi f)^2} = \frac{\alpha^2 + (2\pi f b)^2}{\alpha^2 + (2\pi f)^2},$$

$$\alpha^2 = 2\alpha\sigma_s^2 + \alpha^2 N_0/2; \quad b^2 = N_0/2.$$

2. Полагаем

$$B_x(f) = X(f) X^*(f),$$

$$\text{где } X(f) = \frac{a + j 2\pi f b}{a + j 2\pi f}; \quad X^*(f) = \frac{a - j 2\pi f b}{a - j 2\pi f}.$$

3. Принимаем в соответствии с формулой (7.27) функцию передачи обеляющего фильтра

$$H_1(f) = \frac{1}{X(f)} = \frac{a + j 2\pi f}{a + j 2\pi f b}.$$

4. Находим по формуле (7.28) функцию передачи нереализуемого формирующего фильтра

$$H_{2\Phi}(f) = \frac{B_s(f)}{X^*(f)} e^{j 2\pi f t_0} = \frac{\alpha \sigma_s^2 e^{j 2\pi f t_0}}{(a + j 2\pi f)(a - j 2\pi f b)}$$

(учитывая, что $\alpha^2 + (2\pi f)^2 = (a + j 2\pi f)(a - j 2\pi f)$).

5. Взяв обратное преобразование Фурье от $H_{2\Phi}(f)$, по теореме вычетов находим импульсную характеристику нереализуемого формирующего фильтра

$$h_{2\Phi}(\tau) = 2\alpha \sigma_s^2 \int_{-\infty}^{\infty} \frac{e^{j 2\pi f (\tau + t_0)} df}{(a + j 2\pi f)(a - j 2\pi f b)} = \frac{2\alpha \sigma_s^2 e^{-\alpha(\tau + t_0)}}{a + \alpha b}.$$

6. Взяв прямое преобразование Фурье от $h_2(\tau)$, определяемое через $h_{2\Phi}(\tau)$ согласно условию (7.29), получим функцию передачи реализуемого формирующего фильтра

$$H_2(f) = \frac{2\alpha \sigma_s^2 e^{-\alpha t_0}}{a + \alpha b} \int_0^{\infty} e^{-(\alpha + j 2\pi f) \tau} d\tau = \frac{2\alpha \sigma_s^2 e^{-\alpha t_0}}{(a + \alpha b)(\alpha + j 2\pi f)}$$

(здесь интегрирование ведется только по положительным значениям τ , так как согласно условию (7.29) при отрицательных τ $h_2(-\tau) = 0$).

7. Окончательно находим функцию передачи фильтра Винера

$$H(f) = H_1(f) H_2(f) = \frac{2\alpha \sigma_s^2 e^{-\alpha t_0}}{(a + \alpha b)(a + j 2\pi f b)} =$$

$$= \frac{2\alpha \sigma_s^2 e^{-\alpha t_0}}{a(a + \alpha b)} \frac{1}{1 + j 2\pi f b/a} = \frac{K_0}{1 + j 2\pi f T}.$$

Таким образом, в данном примере искомый оптимальный фильтр-экстраполятор представляет собой безынерционный усилитель с коэффициентом усиления $K_0 = 1/e^{\alpha t_0} [1 + (1 + \sqrt{1 + 4\sigma_s^2/\alpha N_0})/(4\sigma_s^2/\alpha N_0)]$ и RC-цепь с постоянной времени $T = 1/\alpha \sqrt{1 + 4\sigma_s^2/\alpha N_0}$. При заданном интервале t_0 характеристики фильтра определяются шириной спектра сигнала α и отношением сигнал-шум $\sigma_s^2/(\alpha N_0)$.

Приведенный пример показывает, что даже в простейшем случае аналитический расчет параметров фильтра Винера для непрерывного входного процесса вызывает серьезные трудности. Более сложные примеры вообще не поддаются аналитическому расчету.

7.5. ДИФФЕРЕНЦИАЛЬНОЕ УРАВНЕНИЕ ОДНОМЕРНОГО ФИЛЬТРА КАЛМАНА — БЬЮСИ

Фильтрация случайных процессов методом Винера сопряжена с рядом трудностей. Во-первых, как уже отмечалось, отсутствует аналитическое решение интегрального уравнения Винера — Хопфа, за исключением нескольких простейших случаев. Применение ЭВМ в значительной мере снимает этот недостаток при построении фильтра, но не его анализе. Во-вторых, при отыскании структуры фильтра возникают осложнения, связанные с его физической реализуемостью. В-третьих, интегральное описание фильтрации случайного процесса, взятого в целом, усложняется, когда процесс нестационарен, т. е. его параметры изменяются в течение времени наблюдения.

Все это послужило причиной того, что в последнее время предпочитают производить фильтрацию случайных сигналов на основе решения не интегральных, а дифференциальных уравнений с заданными начальными условиями. Этот метод получил название фильтрации Калмана — Бьюси. Предложенная Калманом и Бьюси рекуррентная форма алгоритма фильтрации оказалась очень удобной в прикладном отношении: не требуется большой емкости памяти, естественным образом учитывается изменение характеристик входного процесса по мере его поступления и фильтрации и т. д. Ввиду этого описание фильтра Калмана — Бьюси начнем непосредственно с нестационарных случайных процессов. Для удобства функцию корреляции нестационарного процесса $x(t)$ представим как $M[x(t)x(t')]$, где время t' может изменяться от нуля до текущего момента t . Описание процесса фильтрации начнем с наиболее простого случая — одномерного фильтра для централизованного процесса.

Предположим, что интересующий нас случайный сигнал $s(t)$, содержащийся во входной смеси $x(t) = s(t) + n(t)$, образуется в результате пропускания порождающего его белого шума $v(t)$ через формирующий фильтр, дифференциальное уравнение которого имеет вид

$$ds(t)/dt = a(t)s(t) + v(t), \quad s(0) = s_0. \quad (7.31)$$

Здесь $a(t)$ — коэффициент, медленно изменяющийся во времени в соответствии с изменением дисперсии $\sigma_s^2(t)$ входного сигнала $s(t)$. Полагаем также, что процессы $s(t)$, $n(t)$, $v(t)$ взаимно независимы и центрированы:

$$\begin{aligned} M[s(t)n(t')] &= M[s(t)v(t')] = M[n(t)v(t')] = 0, \\ M[s(t)] &= M[n(t)] = M[v(t)] = 0, \end{aligned} \quad (7.32)$$

а корреляционные функции входного и порождающего шумов имеют вид

$$M[n(t)n(t')] = q(t)\delta(t-t'); \quad M[v(t)v(t')] = g(t)\delta(t-t'). \quad (7.33)$$

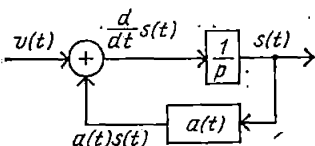


Рис. 7.5. Схема формирующего фильтра для одномерного центрированного процесса

Отсюда следует также, что

$$M[n(t)x(t')] = M[v(t)x(t')] = 0. \quad (7.34)$$

По виду уравнения (7.31) легко установить структуру формирующего фильтра (рис. 7.5). Входной шум $v(t)$ суммируется с выходным сигналом фильтра $s(t)$, прошедшим через безынерционный усилитель с переменным коэффициентом $a(t)$, в результате на выходе сумматора образуется производная выходного сигнала $ds(t)/dt$; поэтому между сумматором и выходом фильтра нужно поставить интегратор $1/p$.

Теперь определим структуру оптимального фильтра Калмана — Бьюси. Пусть искомый фильтр Калмана — Бьюси имеет импульсную характеристику $h(t, t')$, которая медленно изменяется во времени, так как входной процесс нестационарный (это обстоятельство как раз и отображается двойной зависимостью характеристики от времени: одна отражает вид характеристики, другая закон ее изменения). Формируемая на выходе фильтра оценка процесса связана с входным воздействием $x(t)$ интегралом свертки:

$$\hat{s}(t) = \int_0^t h(t, t') x(t') dt'. \quad (7.35)$$

Дисперсия ошибки фильтрации $\sigma_e^2 = M[\hat{s}(t) - s(t)]^2$ минимальна, если характеристика фильтра удовлетворяет интегральному уравнению Винера — Хопфа

$$\int_0^t h(t, t'') M[x(t'') x(t')] dt'' = M[s(t) x(t')]. \quad (7.36)$$

Продифференцировав выражения (7.35) и (7.36) по текущему времени в соответствии с правилом Лейбница

$$\frac{d}{dt} \int_0^t f(t, t') dt' = \int_0^t \frac{d}{dt} f(t, t') dt' + f(t, t), \quad (7.37)$$

получим

$$\frac{d}{dt} \hat{s}(t) = \int_0^t \frac{d}{dt} h(t, t') x(t') dt' + h(t, t) x(t); \quad (7.38)$$

$$\begin{aligned} & \int_0^t \frac{d}{dt} h(t, t'') M[x(t'') x(t')] dt'' + h(t, t) M[x(t) x(t')] = \\ & = M \left[\frac{d}{dt} s(t) x(t') \right] \end{aligned}$$

(в правой части последнего выражения порядок дифференцирования и усреднения изменен). Заменяя производную $ds(t)/dt$ пра-

вой частью уравнения формирующего фильтра (7.31), из последнего уравнения получим

$$\int_0^t \frac{d}{dt} h(t, t'') M[x(t'') x(t')] dt'' + h(t, t) M[x(t) x(t')] = \\ = a(t) M[s(t) x(t')] + M[v(t) x(t')].$$

Здесь $M[x(t) x(t')] = M[s(t) x(t')] + M[n(t) x(t')]$. Учитывая, что согласно условию (7.34) $M[n(t) x(t')] = M[v(t) x(t')] = 0$, после переноса всех слагаемых в левую часть последнее уравнение представим в виде

$$\int_0^t \frac{d}{dt} h(t, t'') M[x(t'') x(t')] dt'' + h(t, t) M[s(t) x(t')] - \\ - a(t) M[s(t) x(t')] = 0.$$

Заменив в нем $M[s(t) x(t')]$ левой частью уравнения (7.36) и обозначив

$$b(t) = h(t, t), \quad (7.39)$$

от суммы одинаковых интегралов перейдем к одному интегралу

$$\int_0^t \left\{ \frac{d}{dt} h(t, t'') + b(t) h(t, t'') - \right. \\ \left. - a(t) h(t, t'') \right\} M[x(t'') x(t')] dt'' = 0. \quad (7.40)$$

Интеграл (7.40) с переменными пределами равен нулю только в том случае, если равно нулю подынтегральное выражение. Приравняв последнее нулю и сократив его на $M[x(t'') x(t')]$, в общем случае не равное нулю, получим

$$\frac{d}{dt} h(t, t') = [a(t) - b(t)] h(t, t'). \quad (7.41)$$

Подставим полученный результат с учетом (7.39) в формулу (7.38):

$$\frac{d}{dt} \hat{s}(t) = [a(t) - b(t)] \int_0^t h(t, t') x(t') dt' + b(t) x(t).$$

Здесь согласно соотношению (7.35) интеграл является оценкой процесса $\hat{s}(t)$. Заменив его на $\hat{s}(t)$ и переставив слагаемые, найдем дифференциальное уравнение одномерного фильтра Калмана — Бьюси для нестационарного центрированного процесса

$$\frac{d}{dt} \hat{s}(t) = a(t) \hat{s}(t) + b(t) [x(t) - \hat{s}(t)], \quad s(0) = s_0. \quad (7.42)$$

По уравнению (7.41) легко установить структуру одномерного фильтра Калмана — Бьюси (рис. 7.6). На входе его из входного воздействия $x(t)$ вычитается оценка процесса $\hat{s}(t)$; результат вычитания после умножения на переменный весовой коэффициент

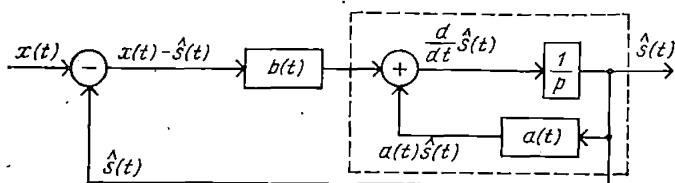


Рис. 7.6. Схема одномерного фильтра Калмана—Бьюси для центрированного процесса

$b(t)$ и суммирования с оценкой $\hat{s}(t)$, помноженной на переменный коэффициент $a(t)$, дает производную оценки $d\hat{s}(t)/dt$, после интегрирования которой на выходе вырабатывается искомая оценка процесса $\hat{s}(t)$.

Таким образом, фильтр Калмана — Бьюси включает в качестве составной части формирующий фильтр (обведен штриховой линией), кроме которого имеется еще входной компаратор (схема вычитания) и входной каскад.

Необходимо отметить, что полученная оценка процесса $\hat{s}(t)$ является несмещенной, т. е. $M[\varepsilon(t)] = M[\hat{s}(t) - s(t)] = 0$ и, следовательно, $M[\hat{s}(t)] = s(t)$. Иными словами, статистические (усредненные) характеристики входного и выходного процессов совпадают. Ввиду этого разностный процесс на выходе компаратора представляет собой входной белый шум, из которого и формируется сигнал на выходе фильтра Калмана. При этом сам фильтр можно представить в виде обеляющей и формирующей частей, где функцию обеления выполняет цепь обратной связи через компаратор. Дифференциальное уравнение фильтра Калмана (7.42) однозначно определяет его структуру, в то время как уравнение Винера — Хопфа описывает лишь свойства фильтра, не предпреляя его структуры. Параметры фильтра Калмана вычисляются путем решения дифференциальных уравнений, в отличие от решения на ЭВМ плохо обусловленной системы уравнений Винера — Хопфа, которое сопровождается большими искажениями при обращении матриц ввиду ограничения разрядной сетки.

7.6. РАСЧЕТ ПАРАМЕТРОВ ОДНОМЕРНОГО ФИЛЬТРА КАЛМАНА — БЬЮСИ

Расчет параметров одномерного фильтра Калмана — Бьюси сводится к нахождению коэффициентов $a(t)$ и $b(t)$ двух звеньев, входящих в него. Коэффициент $a(t)$ можно определить через функцию корреляции входного процесса $s(t)$ из дифференциального уравнения формирующего фильтра (7.31). Для этого преобразуем указанное уравнение по Лапласу:

$$ps(p) = \int_{-\infty}^{\infty} a(p') s(p-p') dp' + v(p) \approx a(p) s(p) + v(p). \quad (7.43)$$

Здесь учтено, что коэффициент $a(t)$ изменяется очень медленно по сравнению с $s(t)$, и поэтому свертка функций $a(p)$ и $s(p)$ приближенно заменена их произведением. Уравнение (7.43) можно представить в виде $[p - a(p)]s(p) = v(p)$, откуда находим изображение по Лапласу функции передачи формирующего фильтра

$$H_{\Phi\Phi}(p) = s(p)/v(p) = 1/[p - a(p)]$$

и квадрат модуля этой функции

$$\begin{aligned} |H_{\Phi\Phi}(p)|^2 &= H_{\Phi\Phi}(p) H_{\Phi\Phi}(-p) = \frac{1}{p - a(p)} \frac{1}{-p - a(-p)} = \\ &= \frac{1}{|a(p)|^2 - p^2}. \end{aligned} \quad (7.44)$$

Известно, что энергетический спектр процесса на выходе линейной системы $B_s(p)$ равен спектру процесса на входе $B_v(p) = g$, помноженному на квадрат модуля функции передачи (7.44):

$$B_s(p) = B_v(p) |H_{\Phi\Phi}(p)|^2 = \frac{B_v(p)}{|a(p)|^2 - p^2}. \quad (7.45)$$

Поскольку функция корреляции $B_s(\tau)$ однозначно связана с энергетическим спектром процесса, задание коэффициента передачи усилителя формирующего фильтра $a(t)$ равносильно заданию функции корреляции $B_s(\tau)$. Таким образом, фильтры Винера и Калмана — Бьюси идентичны по своим свойствам: одинаково задаются характеристики фильтруемого процесса и используется один и тот же критерий оптимизации — минимум дисперсии ошибки фильтрации.

Используя приведенные выше зависимости, можно перейти от корреляционной функции процесса к коэффициенту передачи формирующего фильтра. Переход производится в следующем порядке: взяв преобразование Фурье от функции корреляции $B_s(\tau)$, получим энергетический спектр процесса $B_s(f)$, по которому найдем изображение спектра по Лапласу $B_s(p)$; используя формулу (7.45), определим $a(p)$ и по изображению найдем искомый коэффициент $a(t)$.

Пример 1. Задана функция корреляции (рис. 7.7) стационарного случайного процесса $B_s(\tau) = \sigma_s^2 \exp(-\alpha|\tau|)$; требуется определить коэффициент a . Для этого в соответствии с описанной выше методикой находим сначала энергетический спектр процесса как преобразование Фурье от его функции корреляции:

$$\begin{aligned} B_s(f) &= \sigma_s^2 \int_{-\infty}^{\infty} e^{-\alpha|\tau| - j2\pi f \tau} d\tau = \\ &= \sigma_s^2 \left[\int_{-\infty}^0 e^{\alpha\tau - j2\pi f \tau} d\tau + \int_0^{\infty} e^{-\alpha\tau - j2\pi f \tau} d\tau \right] = \\ &= \sigma_s^2 \left[\frac{1}{\alpha - j2\pi f} + \frac{1}{\alpha + j2\pi f} \right] = \frac{2\sigma_s^2 \alpha}{\alpha^2 + (2\pi f)^2}. \end{aligned}$$

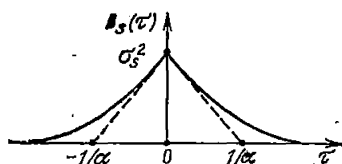


Рис. 7.7. Экспоненциальная функция корреляции входного процесса

Полагая $p = j2\pi f$, получаем изображение спектра по Лапласу

$$B_s(p) = 2\sigma_s^2 \alpha / (\alpha^2 - p^2).$$

Подставим полученный результат в формулу (7.45):

$$2\sigma_s^2 \alpha / (\alpha^2 - p^2) = g / (\alpha^2 - p^2).$$

Отсюда находим: $a = \alpha$; $g = 2\sigma_s^2 \alpha = 2\sigma_s^2 a$.

Так как a имеет смысл ширины полосы спектра фильтруемого сигнала, коэффициент передачи $a = \alpha$ управляет шириной полосы пропускания формирующего фильтра. Параметр g порождающего шума, из которого по предположению формируется случайный сигнал $s(t)$, определяется дисперсией этого сигнала σ_s^2 (и коэффициентом a), т. е. исключается. В нестационарном случае параметры сигнала α и σ_s^2 , а следовательно, и параметр a формирующего фильтра являются медленно меняющимися функциями времени.

Интенсивность белого шума на входе формирующего фильтра, входящего в фильтр Калмана — Бьюси, определяется коэффициентом передачи $b(t)$ входного каскада, который найти значительно сложнее.

Для вычисления коэффициента $b(t)$ помножим ошибку фильтрации

$$\varepsilon(t) = \hat{s}(t) - s(t) \quad (7.46)$$

на $x(t')$ и, усреднив, получим

$$M[\varepsilon(t)x(t')] = M[\hat{s}(t)x(t')] - M[s(t)x(t')]. \quad (7.47)$$

Аналогично из формулы для оценки сигнала

$$\hat{s}(t) = \int_0^t h(t, t'') x(t'') dt'' \quad (7.48)$$

находим

$$M[\hat{s}(t)x(t')] = \int_0^t h(t, t'') M[x(t'')x(t')] dt'' = M[s(t)x(t')], \quad (7.49)$$

где правая часть следует из уравнения Винера — Хопфа (7.36). Из соотношений (7.47) и (7.49) вытекает, что

$$M[\varepsilon(t)x(t')] = 0. \quad (7.50)$$

Из (7.48) находим также следующие зависимости:

$$M[\hat{s}(t)\varepsilon(t')] = \int_0^t h(t, t'') M[x(t'')\varepsilon(t')] dt'' = 0, \quad (7.51)$$

так как согласно (7.50) $M[x(t'')\varepsilon(t')] = 0$;

$$\begin{aligned}
M[\dot{s}(t)s(t')] &= \int_0^t h(t, t'') M[x(t'')s(t')] dt'' = \\
&= \int_0^t h(t, t'') M[s(t'')s(t')] dt'' + \int_0^t h(t, t'') M[n(t'')s(t')] dt'' = \\
&= \int_0^t h(t, t'') M[s(t'')s(t')] dt'', \quad (7.52)
\end{aligned}$$

так как согласно (7.32) $M[n(t'')s(t')] = 0$;

$$\begin{aligned}
M[\dot{s}(t)x(t')] &= \int_0^t h(t, t'') M[x(t'')x(t')] dt'' = \\
&= \int_0^t h(t, t'') M[s(t'')s(t')] dt'' + \int_0^t h(t, t'') M[n(t'')n(t')] dt'' = \\
&= M[\dot{s}(t)s(t')] + q(t) \int_0^t h(t, t'') \delta(t'' - t') dt'' = \\
&= M[\dot{s}(t)s(t')] + q(t)h(t, t'), \quad (7.53)
\end{aligned}$$

где учтено, что $M[x(t'')x(t')] = M[s(t'')s(t')] + M[n(t'')n(t')]$ — из условия независимости сигнала и шума (7.32), а также равенство (7.52), формула (7.33) и фильтрующее свойство дельта-функции.

При $t' = t$ $h(t, t) = b(t)$ и равенство (7.53) принимает вид

$$M[\dot{s}(t)x(t)] = M[\dot{s}(t)s(t)] + q(t)b(t). \quad (7.54)$$

Однако согласно (7.49)

$$M[\dot{s}(t)x(t)] = M[s(t)x(t)] = M[s(t)s(t)] + M[s(t)n(t)] = M[s(t)s(t)],$$

так как $M[s(t)n(t)] = 0$ из условия независимости $s(t)$ и $n(t)$. Учитывая это, а также формулы (7.46) и (7.51), из равенства (7.54) получаем

$$\begin{aligned}
q(t)b(t) &= M[s(t)s(t)] - M[\dot{s}(t)s(t)] = \\
&= M[(s(t) - \dot{s}(t))s(t)] = -M[\varepsilon(t)s(t)] = -M[\varepsilon(t)(\dot{s}(t) - \varepsilon(t))] = \\
&= -M[\varepsilon(t)\dot{s}(t)] + M[\varepsilon^2(t)] = M[\varepsilon^2(t)] = \sigma_\varepsilon^2(t).
\end{aligned}$$

Отсюда находим коэффициент передачи входного усилителя

$$b(t) = \sigma_\varepsilon^2(t)/q(t), \quad (7.55)$$

где остается пока неизвестной дисперсия ошибки фильтрации σ_ε^2 .

Для вычисления σ_ε^2 примем сокращенные обозначения. Вычитая из входной смеси $x = s + n$ оценку $\hat{s} = s + \varepsilon$, получаем

$$x - \hat{s} = n - \varepsilon. \quad (7.56)$$

Произведя соответствующую замену в уравнении фильтра Калмана — Бьюси

$$\hat{\dot{s}} = a\hat{s} + b(x - \hat{s}) = a\hat{s} + b(n - \varepsilon) \quad (7.57)$$

и вычитая из него уравнение формирующего фильтра $\dot{s} = as + v$, находим дифференциальное уравнение ошибки фильтрации

$$\dot{\varepsilon} = a\varepsilon + b(n - \varepsilon) - v = (a - b)\varepsilon + bn - v. \quad (7.58)$$

Из этого уравнения можно получить дифференциальное уравнение дисперсии ошибки фильтрации $\sigma_e^2 = M[\varepsilon^2]$:

$$\begin{aligned}\dot{\sigma}_e^2 &= 2M[\varepsilon\dot{\varepsilon}] = 2\{(a-b)M[\varepsilon^2] + bM[n\varepsilon] - M[v\varepsilon]\} = \\ &= 2\{(a-b)\sigma_e^2 + bM[n\varepsilon] - M[v\varepsilon]\}.\end{aligned}\quad (7.59)$$

Остается вычислить $M[n\varepsilon]$ и $M[v\varepsilon]$. Для этого рассмотрим динамическую систему, на вход которой поступает процесс $v(t)$, а на выходе имеем $\xi(t)$. Дифференциальное уравнение такой системы $\dot{\xi}(t) = \beta(t)\xi(t) + v(t)$, $\xi(0) = \xi_0$ имеет решение

$$\xi(t) = \int_0^t \eta(t, t') v(t') dt' + \eta(t, t') \xi_0.$$

Характеристика $\eta(t, t')$ этой системы дает интеграл свертки с дельта-функцией

$$\int_0^t \eta(t, t') \delta(t - t') dt' = \frac{1}{2} \eta(t, t') = \frac{1}{2},$$

где коэффициент $1/2$ появился потому, что в точке t функция $\eta(t, t')$ испытывает разрыв 1-го рода — площадь дельта-функции делится пополам (рис. 7.8).

В нашем случае уравнение (7.58) является уравнением такой динамической системы, если положить $v = bn - v$, а $\xi = \varepsilon$. Поэтому его решение имеет вид

$$\varepsilon(t) = \int_0^t \eta(t, t') b(t') n(t') dt' - \int_0^t \eta(t, t') v(t') dt' + \eta(t, t') \varepsilon_0.$$

Домножим поочередно это уравнение на $v(t)$ и $n(t)$, а затем усредним. При этом все члены справа, кроме одного, будут равны нулю. Учитывая вид корреляционных функций порождающего и входного шумов (7.33), получаем

$$\begin{aligned}M[\varepsilon(t)v(t)] &= - \int_0^t \eta(t, t') M[v(t')v(t)] dt' = \\ &= -g(t) \int_0^t \eta(t, t') \delta(t, t') dt' = -\frac{1}{2} g(t); \\ M[\varepsilon(t)n(t)] &= \int_0^t \eta(t, t') b(t') M[n(t')n(t)] dt' = \\ &= q(t) \int_0^t \eta(t, t') b(t') \delta(t-t') dt' = \frac{1}{2} b(t) q(t).\end{aligned}$$

Подставив эти результаты в уравнение (7.59) с учетом соотношения (7.55), получим дифференциальное уравнение

$$\begin{aligned}\dot{\sigma}_e^2 &= 2a\sigma_e^2 - 2b\sigma_e^2 + b^2q + g = 2a\sigma_e^2 - 2\sigma_e^2/q + \\ &+ \sigma_e^2/q + g = 2a\sigma_e^2 - \sigma_e^2/q + g.\end{aligned}$$

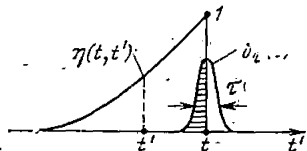


Рис. 7.8. К вычислению интеграла свертки импульсной реакции системы с дельта-функцией $\delta(t) = \lim_{\tau \rightarrow 0} \delta_\tau(t)$

В полной записи дифференциальное уравнение для дисперсии ошибки калмановской фильтрации имеет вид

$$\frac{d}{dt} \sigma_e^2(t) = 2a(t) \sigma_e^2(t) - \sigma_e^4(t)/q(t) + g(t), \quad \sigma_e^2(0) = \sigma_e^2. \quad (7.60)$$

Это существенно нелинейное уравнение Риккати. В общем случае оно не имеет явного аналитического решения и обычно интегрируется численно. Эти расчеты составляют большую часть вычислений, связанных с определением параметров фильтра Калмана — Бьюси, однако их часто можно выполнить заранее. В ряде случаев, например в установившемся режиме фильтрации, когда $d\sigma_e^2(t)/dt=0$, возможно аналитическое решение.

Пример 2. Для стационарного процесса ($d\sigma_e^2(t)/dt=0$) дифференциальное уравнение (7.60) превращается в обычное квадратное уравнение: $\sigma_e^4 - 2aq\sigma_e^2 - qg = 0$, откуда находим

$$\sigma_e^2 = aq + \sqrt{a^2q^4 + qg} = aq[1 + \sqrt{1 + g/a^2q}], \quad b = \sigma_e^2/q = a[1 + \sqrt{1 + g/a^2q}].$$

При экспоненциальной функции корреляции входного процесса (см. пример 1) $g=2a\sigma_s^2$ и $b=a(1 + \sqrt{1 + 2\sigma_s^2/aq})$.

Таким образом, коэффициент передачи входного усилителя $b(t) > 2a(t)$, причем его значение возрастает с увеличением отношения сигнал-шум σ_s^2/aq . Параметры фиктивного, порождающего шума в окончательном выражении для коэффициента $b(t)$, как и дисперсии $\sigma_e^2(t)$, исключаются.

7.7. ВЕКТОРНО-МАТРИЧНОЕ ДИФФЕРЕНЦИАЛЬНОЕ УРАВНЕНИЕ МНОГОМЕРНОГО ФИЛЬТРА КАЛМАНА — БЬЮСИ

Любой действительный случайный сигнал $s(t)$ (или совокупность сигналов $s_1(t), s_2(t), \dots, s_n(t)$), описываемый линейным стохастическим дифференциальным уравнением n -го порядка, однородным (с нулевой правой частью) или неоднородным (правая часть не равна нулю), можно представить в нормальной форме Коши, т. е. в виде n дифференциальных уравнений 1-го порядка. Свойства такого процесса в векторно-матричной форме описываются дифференциальным уравнением многомерного формирующего фильтра

$$\frac{d}{dt} s(t) = A(t)s(t) + v(t), \quad s(0) = s_0, \quad (7.61)$$

которое по виду не отличается от скалярного уравнения одномерного формирующего фильтра (7.31). Здесь $A(t)$ — $(n \times n)$ -матрица состояния формирующего фильтра; $s(t)$ — n -мерный вектор со-

стояния, соответствующий процессу $s(t)$; $v(t)$ — n -мерный порождающий гауссовский шум с характеристиками

$$M[v(t)] = 0, \quad M[v(t)v^T(t')] = G(t)\delta(t-t'), \quad (7.62)$$

где $G(t)$ — симметричная положительно определенная $(n \times n)$ -матрица интенсивности вектора $v(t)$, а индекс t означает транспонирование вектора или матрицы (строки исходной матрицы становятся столбцами транспонированной). Вектор состояния

$$s(t) = \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_n(t) \end{bmatrix} = \begin{bmatrix} s(t) \\ dt(s)/dt \\ \vdots \\ d^n s(t)/dt^n \end{bmatrix}, \quad (7.63)$$

т. е. первая его составляющая равна исходному процессу ($s_1(t) = s(t)$), вторая $s_2(t) = ds(t)/dt$ и т. д. Обычно полагают

$$M[s_0] = 0, \quad M[s_0, s_0^T] = \Theta_{s_0}, \quad (7.64)$$

т. е. начальные условия системы уравнений (7.61) представляют собой центрированную случайную векторную величину с корреляционной матрицей Θ_{s_0} .

Пример 1. Рассмотрим случайный сигнал $s(t)$, формируемый двумя последовательно включенными интеграторами при поступлении на вход скалярного белого гауссовского шума $v(t)$ с плотностью g , т. е. $M[v(t)v(t')] = g\delta(t-t')$ — шум стационарный. Процесс $s(t)$ описывается линейным неоднородным дифференциальным уравнением 2-го порядка $d^2s(t)/dt^2 = v(t)$, которому соответствует система двух дифференциальных уравнений 1-го порядка

$$\begin{aligned} ds_1(t)/dt &= s_2(t), \\ ds_2(t)/dt &= v(t), \quad s_1(t) = s(t). \end{aligned}$$

Полученную систему уравнений представим в матричном виде

$$\begin{bmatrix} ds_1(t)/dt \\ ds_2(t)/dt \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ v(t) \end{bmatrix}$$

или сокращенно

$$ds(t)/dt = As(t) + v(t), \quad s(0) = s_0.$$

где

$$s(t) = \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix}; \quad A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; \quad v(t) = \begin{bmatrix} 0 \\ v(t) \end{bmatrix}; \quad s_0 = \begin{bmatrix} s_0 \\ \dot{s}_0 \end{bmatrix};$$

$$M[v(t)v^T(t')] = G\delta(t-t') = \begin{bmatrix} 0 & 0 \\ 0 & g \end{bmatrix} \delta(t-t').$$

Заметим, что при умножении матрицы A на вектор-столбец $s(t)$ необходимо первый столбец матрицы умножить на первую строку вектор-столбца, а второй столбец матрицы на вторую строку:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} = \begin{bmatrix} 0 \cdot s_1(t) & 1 \cdot s_2(t) \\ 0 \cdot s_1(t) & 0 \cdot s_2(t) \end{bmatrix} = \begin{bmatrix} s_2(t) \\ 0 \end{bmatrix}.$$

При этом числе столбцов первого сомножителя всегда должно быть равно числу строк второго (сомножители менять местами нельзя). Тогда приведенная выше матричная форма системы уравнений

$$\begin{bmatrix} ds_1(t)/dt \\ ds_2(t)/dt \end{bmatrix} = \begin{bmatrix} s_2(t) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ v(t) \end{bmatrix}$$

переходит в исходную систему уравнений (построчно). Из приведенных рассуждений определяется вид матрицы A при переходе от системы уравнений к матричной записи.

Пример 2. Опишем в пространстве состояний случайный сигнал в виде степенного полинома

$$s(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2/2! + \alpha_3 t^3/3!,$$

где $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ — центрированные случайные величины с известной корреляционной матрицей $M[\alpha_i \alpha_j] = \sigma^2_{ij}$, $i, j = 0, \dots, 3$. Этот сигнал можно представить как выходной результат формирующего фильтра, состоящего из трех последовательно включенных интеграторов, начальные условия которого определяются приведенными коэффициентами, а порождающий шум на входе отсутствует (рис. 7.9).

Ввиду отсутствия порождающего шума формирование сигнала приведенной схемой описывается линейным однородным дифференциальным уравнением 4-го порядка $d^4 x(t)/dt^4 = 0$, которое можно заменить системой четырех уравнений 1-го порядка

$$\begin{aligned} ds_1(t)/dt &= s_2(t), \\ ds_2(t)/dt &= s_3(t), \\ ds_3(t)/dt &= s_4(t), \\ ds_4(t)/dt &= 0, \end{aligned} \quad s_1(t) = s(t).$$

В матричной форме эта система уравнений имеет вид

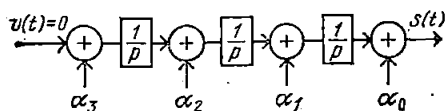
$$\begin{bmatrix} ds_1(t)/dt \\ ds_2(t)/dt \\ ds_3(t)/dt \\ ds_4(t)/dt \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \\ s_4(t) \end{bmatrix}$$

или сокращенно

$$ds(t)/dt = As(t), \quad s(0) = s_0, \quad \text{где}$$

$$s(t) = \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \\ s_4(t) \end{bmatrix}; \quad A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \quad s_0 = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}.$$

Рис. 7.9. Схема формирования сигнала в виде степенного полинома со случайными коэффициентами



На практике измеряются в общем случае не все составляющие вектора состояния $s(t)$, а лишь отдельные. Иначе говоря, на интервале $0, t$ измеряется вектор $x(t)$ размерности $m \leq n$, связанный с вектором $s(t)$ линейным матричным уравнением.

$$x(t) = C(t)s(t) + n(t), \quad (7.65)$$

где $C(t)$ — заданная модуляционная $(m \times n)$ -матрица (матрица наблюдения); $n(t)$ — m -мерный векторный гауссовский белый шум с характеристиками

$$M[n(t)] = 0, \quad M[n(t)n^T(t')] = Q(t)\delta(t-t'), \quad (7.66)$$

$Q(t)$ — симметричная $(m \times m)$ -матрица интенсивности шума измерений $n(t)$.

Пример 3. Предположим, что наблюдаемый процесс $s(t)$ описывается, как в примере 1, неоднородным дифференциальным уравнением 2-го порядка и $s_1(t) = s(t)$ передает координату, некоторого объекта, $s_2(t) = ds_1(t)/dt = ds(t)/dt$ — скорость объекта вдоль этой координаты, а $ds_2(t)/dt = d^2s(t)/dt^2 = v(t)$ — ускорение объекта, представляющее собой белый шум (подобная модель часто используется при описании параметров движения летательных аппаратов на участке маневра).

Если определяется только координата $s(t)$, то

$$x(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} + n_{s_1}(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s(t) \\ \dot{s}(t) \end{bmatrix} + n_s(t)$$

или $x(t) = s(t) + n_s(t)$,

т. е. $C(t) = \begin{bmatrix} 1 & 0 \end{bmatrix}$.

Если определяется только производная координаты $\dot{s}(t)$, то

$$x(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} + n_{s_2}(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} s(t) \\ \dot{s}(t) \end{bmatrix} + n_{\dot{s}}(t)$$

или $x(t) = \dot{s}(t) + n_{\dot{s}}(t)$,

т. е. $C(t) = \begin{bmatrix} 0 & 1 \end{bmatrix}$.

Если же определяется и координата, и ее производная, то

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s(t) \\ \dot{s}(t) \end{bmatrix} = \begin{bmatrix} n_s(t) \\ n_{\dot{s}}(t) \end{bmatrix}$$

или $x(t) = C(t)s(t) + n(t)$,

где

$$C(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad s(t) = \begin{bmatrix} s(t) \\ \dot{s}(t) \end{bmatrix}; \quad n(t) = \begin{bmatrix} n_s(t) \\ n_{\dot{s}}(t) \end{bmatrix};$$

$$Q(t) = \begin{bmatrix} M[n_s^2(t)] & M[n_s(t)n_{\dot{s}}(t)] \\ M[n_s(t)n_{\dot{s}}(t)] & M[n_{\dot{s}}^2(t)] \end{bmatrix}$$

— (2×2) -матрица интенсивности шумов измерений $n(t)$ (обычно $M[n_s(t)n_{\dot{s}}(t)] = 0$).

Векторно-матричное выражение оценки векторного процесса внешне сходно с выражением (7.35) для оценки скалярного процесса:

$$\hat{s}(t) = \int_0^t \mathbf{H}(t, t') \mathbf{x}(t') dt', \quad (7.67)$$

где $\mathbf{H}(t, t')$ — $(n \times m)$ -матрица весовых функций, непрерывно дифференцируемая по обоим аргументам. Оптимальная оценка удовлетворяет условию минимума нормы $\|\varepsilon^2\|$ n -мерного вектора ошибок фильтрации

$$\varepsilon(t) = \hat{\mathbf{s}}(t) - \mathbf{s}(t) = \begin{bmatrix} \hat{s}_1(t) \\ \hat{s}_2(t) \\ \vdots \\ \hat{s}_n(t) \end{bmatrix} - \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_n(t) \end{bmatrix} = \begin{bmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \\ \vdots \\ \varepsilon_n(t) \end{bmatrix}. \quad (7.68)$$

Норма вектора ошибок представляет собой сумму средних квадратов (дисперсий) всех составляющих этого вектора:

$$\begin{aligned} \|\varepsilon^2\| &= M[\varepsilon^T(t) \varepsilon(t)] = M \left\{ \begin{bmatrix} \varepsilon_1(t) & \varepsilon_2(t) & \dots & \varepsilon_n(t) \end{bmatrix} \begin{bmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \\ \vdots \\ \varepsilon_n(t) \end{bmatrix} \right\} = \\ &= M[\varepsilon_1^2(t) + \varepsilon_2^2(t) + \dots + \varepsilon_n^2(t)]. \end{aligned} \quad (7.69)$$

Оптимальная по этому критерию матрица весовых функций $\mathbf{H}(t, t')$ является решением векторно-матричного интегрального уравнения Винера — Хопфа

$$\int_0^t \mathbf{H}(t, t'') M[\mathbf{x}(t'') \mathbf{x}^T(t')] dt'' = M[\mathbf{s}(t) \mathbf{x}^T(t')], \quad 0 \leq t' \leq t. \quad (7.70)$$

Аналогично одномерному случаю получаем векторно-матричное дифференциальное уравнение многомерного фильтра Калмана — Бьюси для централизованного процесса

$$\frac{d}{dt} \hat{\mathbf{s}}(t) = \mathbf{A}(t) \hat{\mathbf{s}}(t) + \mathbf{B}(t) [\mathbf{x}(t) - \mathbf{C}(t) \hat{\mathbf{s}}(t)], \quad \hat{\mathbf{s}}(0) = 0. \quad (7.71)$$

Этому уравнению соответствует матричная структура многомерного фильтра Калмана — Бьюси (рис. 7.10, где штриховой линией выделен формирующий фильтр).

Уравнение (7.71) и матричное представление многомерного фильтра Калмана — Бьюси внешне отличаются от уравнения (7.42) и схемы одномерного фильтра только включением модуляционной матрицы $\mathbf{C}(t)$, когда измеряются не все составляющие векторного процесса (при измерении всех составляющих она становится единичной матрицей).

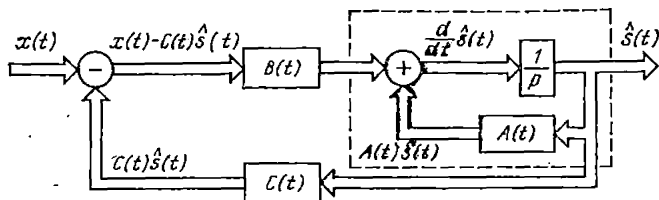


Рис. 7.10. Матричная схема многомерного фильтра Калмана-Бьюси для централизованного процесса

Аналогично нахождению коэффициента $b(t)$ согласно формуле (7.55) определим матричный коэффициент $B(t)$ передачи многомерного фильтра:

$$B(t) = R(t) C^T(t) Q^{-1}(t), \quad (7.72)$$

где

$$R(t) = M[\varepsilon(t) \varepsilon^T(t)] \quad (7.73)$$

— $(n \times n)$ -матрица корреляции вектора ошибки фильтрации $\varepsilon(t)$, которая вычисляется в результате решения нелинейного матричного дифференциального уравнения Риккати

$$\frac{d}{dt} R(t) = A(t) R(t) + R(t) A^T(t) -$$

$$- R(t) C^T(t) Q^{-1}(t) C^T(t) R(t) + G(t), \quad R(0) = \Theta_{s.} \quad (7.74)$$

Здесь $Q^{-1}(t)$ — матрица, обратная матрице интенсивностей шумов измерений (7.66), матрицы $A(t)$ и $C(t)$ заданы, а матрицы $G(t)$ и $R(t)$ определяются соотношениями (7.62) и (7.73) соответственно.

Матричное уравнение (7.74) мало походит на одномерное уравнение (7.60) ввиду особенностей перемножения матриц, но представляет собой систему таких одномерных уравнений. Решение нелинейного уравнения Риккати в многомерном случае еще более усложняется по сравнению с одномерным.

Матрицу корреляции вектора ошибок фильтрации (7.74) необходимо знать не только для определения матрицы коэффициентов передачи $B(t)$, но и для оценки качества фильтрации наблюдаемого процесса $s(t)$.

Заметим, что хотя матричное описание многомерного фильтра Калмана — Бьюси внешне очень сходно с описанием одномерного фильтра, его действительная структура имеет существенные отличия. Покажем это на частном примере фильтрации процесса, рассмотренного в примере 1.

Пример 4. Определим структуру и параметры фильтра Калмана — Бьюси для процесса $s(t)$, описываемого неоднородным дифференциальным уравне-

нием 2-го порядка, в предположении, что наблюдается первая составляющая вектора состояния $s(t)$. При этом

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; \quad S(t) = \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} = \begin{bmatrix} s(t) \\ \frac{ds(t)}{dt} \end{bmatrix}; \quad G = \begin{bmatrix} 0 & 0 \\ 0 & g \end{bmatrix};$$

$$C = [1 \quad 0], \quad Q = q,$$

а также

$$x(t) = [1 \quad 0] \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} + n(t)$$

или $x(t) = s_1(t) + n(t) = s(t) + n(t)$.

Так как вектор $s(t)$ содержит две составляющие, корреляционные матрицы ошибок имеет размер 2×2 :

$$R(t) = \begin{bmatrix} \sigma_1^2(t) & \rho_{12}(t) \\ \rho_{21}(t) & \sigma_2^2(t) \end{bmatrix},$$

где σ_1^2 , σ_2^2 — дисперсия ошибок оценки составляющих $s_1(t)$, $s_2(t)$, а $\rho_{12} = \rho_{21}$ — момент их взаимной корреляции.

В соответствии с формулой (7.72) находим матричный коэффициент передачи

$$B(t) = \begin{bmatrix} \sigma_1^2(t) & \rho_{12}(t) \\ \rho_{12}(t) & \sigma_2^2(t) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{q} = \begin{bmatrix} \sigma_1^2(t) \\ \rho_{12}(t) \end{bmatrix} \frac{1}{q}.$$

Составляем согласно формуле (7.71) векторно-матричное управление Калмана — Бьюси

$$\begin{bmatrix} \frac{d\hat{s}_1(t)}{dt} \\ \frac{d\hat{s}_2(t)}{dt} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{s}_1(t) \\ \hat{s}_2(t) \end{bmatrix} + \begin{bmatrix} \frac{\sigma_1^2(t)}{q} \\ \frac{\rho_{12}(t)}{q} \end{bmatrix} \left\{ x(t) - [1 \quad 0] \begin{bmatrix} \hat{s}_1(t) \\ \hat{s}_2(t) \end{bmatrix} \right\}$$

или после выполнения операций умножения

$$\begin{bmatrix} \frac{d\hat{s}_1(t)}{dt} \\ \frac{d\hat{s}_2(t)}{dt} \end{bmatrix} = \begin{bmatrix} \hat{s}_2(t) \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{\sigma_1^2(t)}{q} \\ \frac{\rho_{12}(t)}{q} \end{bmatrix} [x(t) - \hat{s}_1(t)],$$

которое сводится к системе двух уравнений 1-го порядка

$$\frac{d\hat{s}_1(t)}{dt} = \hat{s}_2(t) + \frac{\sigma_1^2(t)}{q} [x(t) - \hat{s}_1(t)],$$

$$\frac{d\hat{s}_2(t)}{dt} = \frac{\rho_{12}(t)}{q} [x(t) - \hat{s}_1(t)].$$

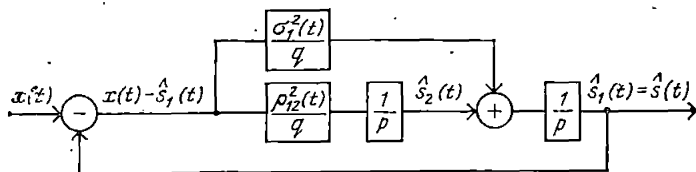


Рис. 7.11. Схема многомерного фильтра Калмана-Бьюси для центрированного процесса 2-го порядка

Этой системе уравнений соответствует структурная схема многомерного фильтра Калмана—Бьюси (рис. 7.11). Нетрудно заметить, что она существенно отличается от матричной структуры одномерного фильтра (рис. 7.6).

Чтобы найти коэффициенты $\sigma_1^2(t)$ и $\rho_{12}^2(t)$, решим дифференциальное уравнение Риккати (7.74), которое в данном конкретном случае принимает вид

$$\begin{bmatrix} \frac{d\sigma_1^2(t)}{dt} & \frac{d\rho_{12}^2(t)}{dt} \\ \frac{d\rho_{12}^2(t)}{dt} & \frac{d\sigma_2^2(t)}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_1^2(t) & \rho_{12}^2(t) \\ \rho_{12}^2(t) & \sigma_2^2(t) \end{bmatrix} + \\ + \begin{bmatrix} \sigma_1^2(t) & \rho_{12}^2(t) \\ \rho_{12}^2(t) & \sigma_2^2(t) \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} - \\ - \begin{bmatrix} \sigma_1^2(t) & \rho_{12}^2(t) \\ \rho_{12}^2(t) & \sigma_2^2(t) \end{bmatrix} \left[\begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{q} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \sigma_1^2(t) & \rho_{12}^2(t) \\ \rho_{12}^2(t) & \sigma_2^2(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & q \end{bmatrix} \right].$$

Перемножив матрицы, получим

$$\begin{bmatrix} \frac{d\sigma_1^2(t)}{dt} & \frac{d\rho_{12}^2(t)}{dt} \\ \frac{d\rho_{12}^2(t)}{dt} & \frac{d\sigma_2^2(t)}{dt} \end{bmatrix} = \begin{bmatrix} \rho_{12}^2(t) & \sigma_2^2(t) \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \rho_{12}^2(t) & 0 \\ \sigma_2^2(t) & 0 \end{bmatrix} - \\ - \frac{1}{q} \begin{bmatrix} \sigma_1^4(t) & \sigma_1^2(t) \rho_{12}^2(t) \\ \rho_{12}^2(t) \sigma_1^2(t) & \rho_{12}^4(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & g \end{bmatrix}.$$

Этому матричному уравнению соответствует система трех скалярных уравнений (поскольку матрица $R(t)$ симметричная)

$$\begin{aligned} d\sigma_1^2(t)/dt &= 2\rho_{12}^2(t) - \sigma_1^4(t)/q, \\ d\rho_{12}^2(t)/dt &= \sigma_2^2(t) - \sigma_1^2(t)\rho_{12}^2(t)/q, \\ d\sigma_2^2(t)/dt &= g - \rho_{12}^2(t)/q. \end{aligned}$$

В установившемся режиме производные равны нулю и дифференциальные уравнения переходят в алгебраические:

$$\begin{aligned} 2\rho_{12}^2 - (\sigma_1^2)^2/q &= 0, \\ \sigma_2^2 - \sigma_1^2\rho_{12}^2/q &= 0, \\ g - \rho_{12}^2/q &= 0. \end{aligned}$$

В результате решения этой системы уравнений последовательно находим искомые коэффициенты

$$\rho_{12} = \sqrt{qg};$$

$$\sigma_1^2 = \sqrt{2\rho_{12}q} = \sqrt[4]{4q^3g};$$

$$\sigma_2^2 = \sigma_1^2 \rho_{12}/q = \sqrt[4]{4qg^3}.$$

Коэффициент передачи разомкнутого фильтра Калмана — Бьюси (рис. 7.11) в установившемся режиме определяется как произведение коэффициентов передачи двух последовательно соединенных звеньев, до сумматора включительно и после сумматора:

$$H(q) = \left(\frac{\rho_{12}}{q} \frac{1}{p} + \frac{\sigma_1^2}{q} \right) \frac{1}{p} = \frac{\rho_{12}}{q} \frac{1}{p^2} \left(1 + p \frac{\sigma_1^2}{\rho_{12}} \right) = \frac{K_a(1+pT)}{p^2},$$

где $K_a = \rho_{12}/g = \sqrt{g/q}$, $T = \sigma_1^2 \rho_{12} = \sqrt{4q/g}$, а параметр g фиктивного порождающего процесса определяется через характеристики самого фильтруемого процесса, как это было показано в примере 2 § 7.6.

7.8. ВЕРОЯТНОСТНЫЙ ФИЛЬТР

Рассмотренные выше фильтры используются для оптимальной фильтрации только центрированных процессов по критерию минимума дисперсии ошибки фильтрации. Так как при фильтрации центрированных процессов в шумах $x(t) = s(t) + n(t)$ математическое ожидание $M[s(t)] = M[n(t)] = M[x(t)] = 0$, оценка $\hat{s}(t)$ получается несмещенной, т. е. $M[\varepsilon(t)] = M[\hat{s}(t) - s(t)] = 0$ (если нет постоянной составляющей на входе линейного фильтра, то ее не может быть и на выходе его). Однако многие реальные процессы $s(t) = s_p(t) + s_{cl}(t)$ наряду со случайной составляющей $s_{cl}(t)$ имеют большую регулярную составляющую $s_p(t)$. Априори составляющая $s_p(t)$ точно неизвестна, поэтому свести процесс к центрированному простым вычитанием не удастся. Остающаяся после вычитания часть регулярной составляющей может значительно превышать случайную составляющую $s_{cl}(t)$. Ввиду этого фильтр Калмана — Бьюси, рассчитанный на нулевую регулярную составляющую, дает на выходе расходящийся результат (ошибка фильтрации со временем не убывает, а растет). Это относится в равной мере и к фильтру Винера, который эквивалентен фильтру Калмана — Бьюси и использует тот же критерий оптимальности.

Указанный недостаток в фильтре Калмана — Бьюси можно устранить, расширив до нужных пределов вектор состояния входного процесса (представляя, в частности, процесс степенным рядом, как это было показано в примере 2 § 7.6). Однако с расширением вектора состояния возрастает дисперсия ошибки фильтрации — эффективность фильтра резко падает. При ограниченной размерности вектора состояния существует ряд методов повышения устойчивости фильтра, которые позволяют эффективно решить задачу фильтрации лишь в частных случаях.

Принципиальный недостаток фильтров Винера и Калмана состоит в том, что для фильтрации нецентрированных процессов, у которых $s_p(t) \neq 0$, критерий минимума дисперсии ошибки фильтрации σ_e^2 уже непригоден, так как ошибка $\varepsilon = \varepsilon_p + \varepsilon_{сл}$ кроме параметра $\sigma_e^2 = M[\varepsilon_{сл}^2] = M[(\varepsilon - \varepsilon_p)^2]$ включает еще и параметр $M[\varepsilon] = \varepsilon_p$, на который никаких ограничений не накладывается. При этом необходимо обратить внимание на то обстоятельство, что влияние ширины полосы пропускания фильтра Δf на среднее квадратическое значение случайной составляющей ошибки σ_e и регулярную (динамическую) составляющую ошибки ε_p прямо противоположно: с уменьшением полосы Δf величина σ_e уменьшается, а ε_p растет (при узкой полосе наряду с шумами сглаживается (искажается) и действительный вид фильтруемого процесса).

Ввиду изложенного при фильтрации нецентрированных случайных процессов был введен критерий минимума суммарной квадратической ошибки

$$\varepsilon_{\Sigma} = \sqrt{\varepsilon_p^2 + k^2 \sigma_e^2}, \quad (7.75)$$

где k — коэффициент Лагранжа. Минимум суммарной ошибки достигается при некоторой оптимальной полосе пропускания фильтра Δf_{opt} , когда динамическая ошибка $\varepsilon_p \neq 0$ (рис. 7.12). Таким образом, при фильтрации нецентрированных случайных сигналов оптимальная оценка принадлежит к классу смещенных.

Следует, однако, отметить, что критерий минимума суммарной ошибки (7.75) имеет принципиальный недостаток — неопределенность коэффициента k . Указанный недостаток можно исключить, если использовать критерий максимума вероятности попадания ошибки фильтрации в заданную область $[-\alpha, \beta]$: $P = P(-\alpha \leq \varepsilon \leq \beta)$. Этот критерий учитывает как дисперсию случайной составляющей ошибки фильтрации, так и регулярную составляющую ошибки: вероятность P повышается с уменьшением σ_e^2 и ε_p .

Поскольку каждая из этих составляющих зависит от параметров фильтра, в частности его импульсной характеристики $h(t, t')$, указанный критерий можно определить в виде сложной функции

$$P = P\{-\alpha \leq \varepsilon \leq \beta; \varepsilon_p[h(t, t')], \sigma_e^2[h(t, t')]\} \rightarrow \max. \quad (7.76)$$

Фильтр, оптимальный по критерию максимума вероятности попадания ошибки фильтрации в заданную область (7.76), называют вероятностным. Его теория развита в работах Н. И. Андреева, Б. Е. Рудницкого и наиболее полно Б. Г. Мельникова.

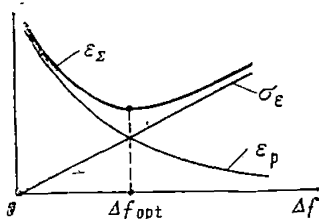


Рис. 7.12. Зависимость динамической ошибки ε_p , среднее квадратического значения случайной ошибки σ_e и суммарной квадратической ошибки фильтрации ε_{Σ} от ширины полосы фильтра ($k=1$)

В частном случае центрированных процессов ($\varepsilon_p = 0$) критерий (7.76) сводится к критерию минимума дисперсии ошибки фильтрации, а сам фильтр сводится к фильтрам Винера или Калмана, описанным выше. Таким образом, вероятностный фильтр является обобщением указанных фильтров.

Чтобы найти оптимальную характеристику одномерного вероятностного фильтра $h_0(t, t')$, применим вариационный метод, положив $h(t, t') = h_0(t, t') = \mu \kappa(t, t')$, и правило дифференцирования сложной функции

$$\frac{\partial P}{\partial \sigma_\varepsilon^2} \frac{d}{d\mu} \sigma_\varepsilon^2 [h(t, t')]_{\mu=0} + \frac{\partial P}{\partial \varepsilon_p} \frac{d}{d\mu} \varepsilon_p [h(t, t')]_{\mu=0} = 0$$

или

$$\frac{d}{d\mu} \sigma_\varepsilon^2 [h(t, t')]_{\mu=0} - 2\psi(t) \frac{d}{d\mu} \varepsilon_p [h(t, t')]_{\mu=0} = 0, \quad (7.77)$$

где

$$\psi(t) = -\frac{1}{2} \frac{\partial P / \partial \varepsilon_p(t)}{\partial P / \partial \sigma_\varepsilon^2} \quad (7.78)$$

— коэффициент связи, учитывающий относительное влияние σ_ε^2 и ε_p на вероятность P (здесь знак минус учитывает отмеченный выше противоположный характер влияния характеристик фильтра на σ_ε^2 и ε_p и, следовательно, на вероятность P , а коэффициент $1/2$ введен в связи с тем, что σ_ε^2 есть среднее от квадрата $\varepsilon_{сл}$, а ε_p входит в первой степени).

Представляет интерес частный случай, наиболее часто встречающийся на практике, когда ошибка фильтрации ε имеет нормальное распределение, т. е.

$$\begin{aligned} P = P(-\alpha \leq \varepsilon \leq \beta) &= \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \int_{-\alpha}^{\beta} e^{-\frac{(\varepsilon - \varepsilon_p)^2}{2\sigma_\varepsilon^2}} d\varepsilon = \\ &= \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \int_{-\alpha - \varepsilon_p}^{\beta - \varepsilon_p} e^{-\frac{\varepsilon_{сл}^2}{2\sigma_\varepsilon^2}} d\varepsilon_{сл}. \end{aligned} \quad (7.79)$$

Тогда

$$\begin{aligned} \frac{\partial P}{\partial \varepsilon_p} &= \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \frac{1}{\sigma_\varepsilon^2} \int_{-\alpha}^{\beta} (\varepsilon - \varepsilon_p) e^{-\frac{(\varepsilon - \varepsilon_p)^2}{2\sigma_\varepsilon^2}} d\varepsilon = \\ &= \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \frac{1}{\sigma_\varepsilon^2} \int_{-\alpha - \varepsilon_p}^{\beta - \varepsilon_p} \varepsilon_{сл} e^{-\frac{\varepsilon_{сл}^2}{2\sigma_\varepsilon^2}} d\varepsilon_{сл}; \end{aligned}$$

$$\frac{\partial P}{\partial \sigma_{\varepsilon}^2} = \frac{1}{V 2\pi \sigma_{\varepsilon}^2} \frac{1}{2 \sigma_{\varepsilon}^2} \int_{-\alpha - \varepsilon_p}^{\beta - \varepsilon_p} \left(\frac{\varepsilon_{\text{сл}}^2}{\sigma_{\varepsilon}^2} - 1 \right) e^{-\frac{\varepsilon_{\text{сл}}^2}{2\sigma_{\varepsilon}^2}} d\varepsilon_{\text{сл}}.$$

Отсюда, отбросив индекс «сл» у переменной интегрирования, находим согласно формуле (7.78) значение коэффициента связи

$$\psi(t) = \frac{\int_{-\alpha - \varepsilon_p}^{\beta - \varepsilon_p} \varepsilon e^{-\frac{\varepsilon^2}{2\sigma_{\varepsilon}^2}} d\varepsilon}{\int_{-\alpha - \varepsilon_p}^{\beta - \varepsilon_p} \left(1 - \frac{\varepsilon^2}{\sigma_{\varepsilon}^2} \right) e^{-\frac{\varepsilon^2}{2\sigma_{\varepsilon}^2}} d\varepsilon}. \quad (7.80)$$

Аналогично определяется коэффициент связи при любом другом распределении ошибки фильтрации.

Переходя к определению оптимальной характеристики вероятностного фильтра для нестационарного нецентрированного процесса вариационным методом, заметим, что первое слагаемое уравнения (7.77) полностью соответствует формуле (7.10) и в несколько измененных обозначениях имеет вид

$$\begin{aligned} \frac{d}{d\mu} \sigma_{\varepsilon}^2 [h(t, t')]_{\mu=0} = \\ = 2 \int_0^t \left\{ \int_0^t h(t, t'') M[x_{\text{сл}}(t'') x_{\text{сл}}(t')] dt'' - M[s_{\text{сл}}(t) x_{\text{сл}}(t')] \right\} \kappa(t, t') dt', \end{aligned} \quad (7.81)$$

где $M[x_{\text{сл}}(t'') x_{\text{сл}}(t')]$ — функция корреляции центрированного процесса $x_{\text{сл}} = x(t) - s_p(t) = s_{\text{сл}}(t) + n(t)$, поскольку дисперсия σ^2 не зависит от регулярной составляющей $s_p(t)$ и ошибки $\varepsilon_p(t)$, а $M[s_{\text{сл}}(t) x_{\text{сл}}(t')]$ — функция взаимной корреляции центрированных сигнала и смеси, равная $M[s_{\text{сл}}(t) s_{\text{сл}}(t')]$; когда сигнал и шум не коррелированы. Для вычисления второго слагаемого в уравнении (7.77) запишем выражение для оценки регулярной составляющей процесса с учетом характеристики $h(t, t') = h_0(t, t') + \mu \kappa(t, t')$:

$$\hat{s}_p(t) = \int_0^t [h_0(t, t') + \mu \kappa(t, t')] s_p(t') dt',$$

откуда и находим

$$\begin{aligned} \frac{d}{d\mu} \varepsilon_p [h(t, t')]_{\mu=0} = \\ = \frac{d}{d\mu} \{ \hat{s}_p[t, h(t, t')] - s_p(t) \}_{\mu=0} = \int_0^t \kappa(t, t') s_p(t') dt'. \end{aligned} \quad (7.82)$$

Подставив соотношения (7.81) и (7.82) в формулу (7.77), получим

$$\begin{aligned} 2 \int_0^t \int_0^t h(t, t'') M[\dot{x}(t'') x(t')] dt'' - M[s(t) x(t')] - \\ - \psi(t) s_p(t') \} \kappa(t, t') dt' = 0. \end{aligned}$$

При переменном пределе внешний интеграл равен нулю тогда и только тогда, когда равно нулю само подынтегральное выражение. Приравняв последнее нулю и сократив его на $x(t, t')$, в общем случае не равное нулю, получим интегральное уравнение для оптимальной характеристики одномерного вероятностного фильтра, представленной в неявном виде:

$$\int_0^t h(t, t'') M[x(t'') x(t')] dt'' = M[s(t) x(t')] + \psi(t) s_p(t'). \quad (7.83)$$

Нетрудно видеть, что в отсутствие регулярной составляющей (при $s_p(t) = 0$) оно обращается в интегральное уравнение Винера — Хопфа (7.11).

Поскольку правая часть уравнения (7.83) представлена в виде суммы, можно, используя принцип декомпозиции, представить в виде суммы и левую часть. Тогда от уравнения (7.83) приходим к системе двух интегральных уравнений

$$\int_0^t h_1(t, t'') M[x(t'') x(t')] dt'' = M[s(t) x(t')], \quad (7.84)$$

$$\int_0^t h_2(t, t'') M[x(t'') x(t')] dt'' = \psi(t) s_p(t'), \quad (7.85)$$

где значения интегралов к текущему моменту t необходимо вычислить для всех $0 \leq t' \leq t$.

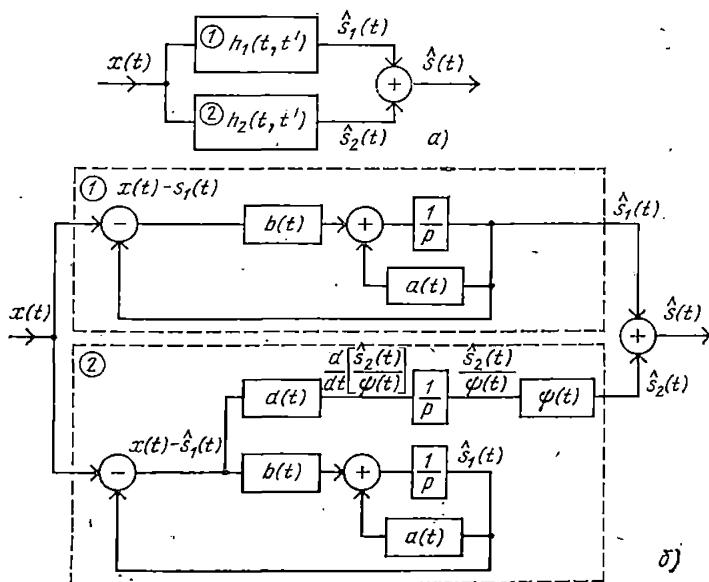


Рис. 7.13. Представление вероятностного фильтра в виде параллельного соединения двух фильтров при интегральном (а) и дифференциальном (б) описании

Таким образом, искомый фильтр представляется в виде параллельно соединенных двух фильтров (рис. 7.13,а):

$$h(t, t') = h_1(t, t') + h_2(t, t'). \quad (7.86)$$

Выходные сигналы этих фильтров

$$\hat{s}_1(t) = \int_0^t h_1(t, t') x(t') dt', \quad (7.87)$$

$$\hat{s}_2(t) = \int_0^t h_2(t, t') x(t') dt', \quad (7.88)$$

суммируясь, дают оценку процесса в целом

$$\hat{s}(t) = \hat{s}_1(t) + \hat{s}_2(t). \quad (7.89)$$

Уравнение первого фильтра (7.84) является уравнением Винера — Хопфа, которое при переходе от интегрального описания к дифференциальному дает уравнение одномерного фильтра Калмана — Бьюси

$$\frac{d}{dt} \hat{s}_1(t) = a(t) \hat{s}_1(t) + b(t) [x(t) - \hat{s}_1(t)], \quad s_1(0) = s_{10}, \quad (7.90)$$

повторяющее соотношение (7.42). Здесь обозначено

$$b(t) = h_1(t, t). \quad (7.91)$$

Чтобы найти дифференциальное уравнение второго из двух фильтров, продифференцируем формулы (7.85) и (7.88):

$$\frac{d}{dt} \hat{s}_2(t) = \int_0^t \frac{d}{dt} h_2(t, t') x(t') dt' + h_2(t, t) x(t); \quad (7.92)$$

$$\begin{aligned} & \int_0^t \frac{d}{dt} h_2(t, t'') M[x(t'') x(t')] dt'' + h_2(t, t) M[s(t) x(t)] = \\ & = \frac{d}{dt} \psi(t) s_p(t'). \end{aligned} \quad (7.93)$$

Здесь использовано правило Лейбница (7.37), причем в последнем случае, как и при дифференцировании зависимости (7.36), учтено, что $M[x(t) x(t')] = M[s(t) x(t')]$. Подставив в формулу (7.93) вместо $M[s(t) x(t')]$ и $s_p(t')$ левые части выражений (7.84) и (7.85) и положив

$$d(t) = h_2(t, t) / \psi(t), \quad (7.94)$$

получим

$$\begin{aligned} & \int_0^t \left[\frac{d}{dt} h_2(t, t'') + \psi(t) d(t) h_1(t, t'') - \right. \\ & \left. - \frac{d}{dt} \psi(t) \frac{h_2(t, t'')}{\psi(t)} \right] M[x(t'') x(t')] dt'' = 0. \end{aligned}$$

Приравняв нулю подынтегральное выражение этого интеграла с переменными пределами и сократив его на $M[x(t'')x(t')] \neq 0$, находим

$$\frac{d}{dt} h_2(t, t'') = \frac{d}{dt} \psi(t) \frac{h_2(t, t'')}{\psi(t)} - \psi(t) d(t) h_1(t, t'').$$

Подставим правую часть этого равенства в формулу (7.92):

$$\begin{aligned} \frac{d}{dt} \hat{s}_2(t) &= \frac{d}{dt} \psi(t) \frac{1}{\psi(t)} \int_0^t h_2(t, t') x(t') dt' - \\ &- \psi(t) d(t) \int_0^t h_1(t, t') x(t') dt' + \psi(t) d(t) x(t) = \\ &= \frac{d}{dt} \psi(t) \frac{\hat{s}_2(t)}{\psi(t)} + \psi(t) d(t) [x(t) - \hat{s}_1(t)], \end{aligned}$$

где учтены зависимости (7.87), (7.88) и (7.94). Продифференцируем теперь тождество $\hat{s}_2(t) = \psi(t) [\hat{s}_2(t)/\psi(t)]$:

$$\frac{d}{dt} \hat{s}_2(t) = \frac{d}{dt} \psi(t) \frac{\hat{s}_2(t)}{\psi(t)} + \psi(t) \frac{d}{dt} \left[\frac{\hat{s}_2(t)}{\psi(t)} \right].$$

Приравняв правые части двух последних равенств, после приведения подобных и сокращения на $\psi(t) \neq 0$ окончательно получим дифференциальное уравнение второго фильтра:

$$\frac{d}{dt} \left[\frac{\hat{s}_2(t)}{\psi(t)} \right] = d(t) [x(t) - \hat{s}_1(t)], \quad \hat{s}_2(0) = s_{20}. \quad (7.95)$$

Таким образом, для формирования оценки $\hat{s}(t)$ во втором фильтре разность $x(t) - \hat{s}_1(t)$ умножается на коэффициент $d(t)$, интегрируется, затем полученный результат умножается на $\psi(t)$. Как следует из формулы (7.95), второй фильтр должен также вырабатывать и оценку $\hat{s}_1(t)$, т. е. содержать в своем составе фильтр Калмана — Бьюси (рис. 7.13,б).

Рассмотренная двухканальная схема одномерного вероятностного фильтра в дифференциальной форме структурно избыточна, так как содержит в обеих ветвях два совершенно одинаковых фильтра с общим входом. Поэтому можно отказаться от представления фильтра в виде двух параллельных каналов и ограничиться лишь одной нижней ветвью, содержащей замкнутый фильтр Калмана — Бьюси (обведен штриховой линией), вырабатывающий оценку $\hat{s}_1(t)$, и разомкнутую ветвь формирования оценки $\hat{s}_2(t)$, соединив их выходы через сумматор (рис. 7.14,а). Этот фильтр имеет обратную связь только по одной составляющей выходного сигнала, а именно $\hat{s}_1(t)$. В ряде случаев желательно замкнуть обратную связь по полной реакции фильтра $\hat{s}(t) = \hat{s}_1(t) + \hat{s}_2(t)$. Так как это равносильно дополнительному вычитанию составляющей $\hat{s}_2(t)$ из входной смеси $x(t)$, для сохранения исходного состояния на выходе схемы вычитания необходимо вновь ввести эту составляющую (рис. 7.14,б). Тогда оба канала формирования оценки

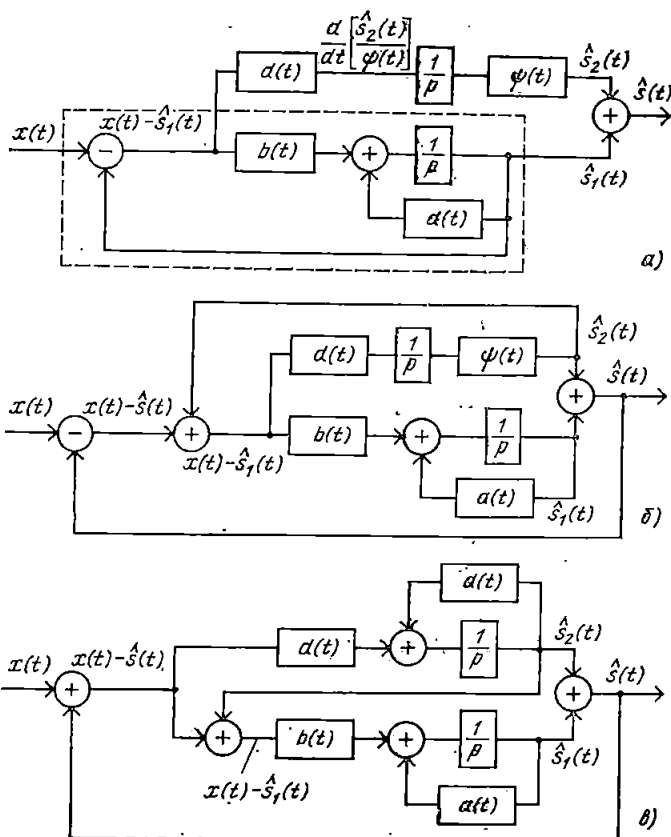


Рис. 7.14. Схема одномерного вероятностного фильтра с обратной связью по неполной (а) и полной (б и в) реакции

$\hat{s}_1(t)$ и $\hat{s}_2(t)$ оказываются взаимно зависимыми и дифференциальные уравнения (7.90) и (7.95), описывающие вероятностный фильтр с обратной связью по полной реакции, принимают вид

$$\frac{d}{dt} \hat{s}_1(t) = a(t) \hat{s}_1(t) + b(t) [x(t) - \hat{s}_1(t) + \hat{s}_2(t)], \quad (7.96)$$

$$\frac{d}{dt} \left[\frac{\hat{s}_2(t)}{\psi(t)} \right] = d(t) [x(t) - \hat{s}_1(t) + \hat{s}_2(t)]. \quad (7.97)$$

Последнее уравнение можно привести к виду

$$\frac{d}{dt} \left[\frac{\hat{s}_2(t)}{\psi(t)} \right] = (t) [x(t) - \hat{s}_1(t)] + d(t) \hat{s}_2(t), \quad (7.98)$$

при котором структура фильтра несколько изменяется (рис. 7.14, в). В схеме вероятностного фильтра, замкнутого по полной реакции, уже нельзя обособить фильтр Калмана — Бьюси, хотя

структура его калмановской ветви сохраняется. Для последней схемы характерно, что в ней четко выделяется формирующий фильтр и во второй, некалмановской ветви, состоящей из интегратора, сумматора и безынерционных усилителей.

В вероятностном фильтре через обе его ветви протекает как регулярная, так и случайная составляющие. Поэтому, в частности, оценку регулярной составляющей $s_p(t)$ с учетом обозначения $h_2(t, t') = \psi(t) d(t')$, вытекающего из формулы (7.94), можно представить в виде суммы составляющих оценки в каждой из двух ветвей:

$$\hat{s}_p(t) = \hat{s}_{p1}(t) + \hat{s}_{p2}(t), \quad (7.99)$$

где

$$\hat{s}_{p1}(t) = \int_0^t h_1(t, t') s_p(t') dt', \quad (7.100)$$

$$\hat{s}_{p2}(t) = \int_0^t h_2(t, t') s_p(t') dt' = \psi(t) \int_0^t d(t') s_p(t') dt'. \quad (7.101)$$

Аналогичные представления справедливы и для случайной составляющей:

$$\hat{s}_{сл}(t) = \hat{s}_{сл1}(t) + \hat{s}_{сл2}(t), \quad (7.102)$$

где

$$\hat{s}_{сл1}(t) = \int_0^t h_1(t, t') s_{сл}(t') dt', \quad (7.103)$$

$$\hat{s}_{сл2}(t) = \int_0^t h_2(t, t') s_{сл}(t') dt'. \quad (7.104)$$

Характеристики этих ветвей одномерного вероятностного фильтра в интегральной форме определяются из решения интегральных уравнений (7.84) и (7.85), а коэффициент связи $\psi(t)$ задан формулами (7.78) или (7.80). Параметры $a(t)$ и $b(t)$ дифференциального вероятностного фильтра также известны и определяются формулами (7.45) и (7.55) с учетом уравнения (7.60) для фильтра Калмана — Бьюси. Остается определить коэффициент $d(t)$.

Коэффициент $d(t)$ наиболее просто вычисляется, если входная смесь $x(t) = x_p(t) = s_p(t) + n(t)$ не содержит случайной составляющей $s_{сл}(t)$. Для вычислений воспользуемся уравнением (7.85) с учетом обозначения $h_2(t, t') = \psi(t) d(t')$ и корреляционной функции центрированного входного процесса, который в данном случае представляет собой белый шум $n(t)$, $M[x(t)x(t')] = M[n(t)n(t')] = = 0,5 N_0(t) \delta(t-t')$. Тогда после сокращения на $\psi(t)$ получим

$$\int_0^t d(t') M[n(t)n(t')] dt' = s_p(t) \quad \text{или} \quad \frac{N_0(t)}{2} \int_0^t d(t') \delta(t-t') dt' = s_p(t),$$

откуда с учетом фильтрующего свойства дельта-функции находим

$$d(t) = 2s_p(t)/N_0(t), \quad (7.105)$$

где $N_0(t)$ — спектральная плотность шума, равномерная по частоте и переменная во времени (нестационарный процесс).

В общем случае вычисление $d(t)$ очень сильно усложняется. Однако, имея в виду, что на вторую ветвь воздействует разность $x(t) - \hat{s}_1(t)$, регулярная составляющая которой равна $s_p(t) - \hat{s}_{p1}(t)$, можно от частного соотношения (7.105) перейти к более общему

$$d(t) = 2[s_p(t) - \hat{s}_{p1}(t)]/N_0(t), \quad (7.106)$$

которое доказывается совершенно строго.

Таким образом, для вычисления коэффициента $d(t)$ необходимо точно знать регулярную составляющую входного процесса $s_p(t)$. Однако это требование противоречиво: если регулярная составляющая точно известна, то ее можно вычесть, и для фильтрации процесса использовать фильтр Калмана — Бьюси, где необходимость в определении коэффициента $d(t)$ полностью отпадает. На самом деле регулярная составляющая почти никогда точно не известна, хотя приближенное представление о ней имеется или может быть получено перед началом фильтрации. Точно так же не определен вначале и коэффициент связи $\psi(t)$. Однако все эти трудности преодолеваются благодаря замечательному свойству вероятностного фильтра, которое рассматривается ниже.

Чтобы определить начальное значение коэффициента связи, подставим в формулу (7.99) оценку (7.101):

$$\hat{s}_p(t) = \hat{s}_{p1}(t) + \psi(t) \int_0^t d(t') s_p(t') dt',$$

откуда находим

$$\psi(t) = [\hat{s}_p(t) - \hat{s}_{p1}(t)] / \int_0^t d(t') s_p(t') dt'. \quad (7.107)$$

Так как динамическая ошибка $\varepsilon_p(t) = \hat{s}_p(t) - s_p(t)$ обычно мала, в качестве начального значения примем такое $\psi^0(t)$, при котором оценку можно считать несмещенной, т. е. $\varepsilon_p(t) = 0$ или $\hat{s}_p(t) = s_p(t)$. Тогда

$$\psi^0(t) = \frac{s_p(t) - \hat{s}_{p1}(t)}{\int_0^t d(t') s_p(t') dt'}$$

или с учетом формулы (7.106)

$$\psi^0(t) = \frac{s_p(t) - \hat{s}_{p1}(t)}{2 \int_0^t \{s_p(t') [s_p(t') - \hat{s}_{p1}(t')] / N_0(t')\} dt'} \quad (7.108)$$

Вернемся к частному случаю, когда входная смесь $x(t) = s_p(t) + n(t)$ не содержит случайной составляющей. При этом надобность в фильтрации этой составляющей отпадает и $\hat{s}_{p1}(t) = 0$. Тогда выходной сигнал фильтра

$$\hat{s}(t) = \hat{s}_2(t) = \psi(t) \int_0^t d(t') x(t') dt'$$

или после подстановки соотношения (7.105) и замены $\psi(t)$ на $\psi^0(t)$ с учетом $\hat{s}_{p1}(t) = 0$

$$\hat{s}(t) = s_p(t) \int_0^t [s_p(t')/N_0(t')] x(t') dt' \bigg/ \int_0^t [s_p^2(t')/N_0(t')] dt'. \quad (7.109)$$

Предположим теперь, что при вычислении коэффициентов $\psi^0(t)$ и $d(t)$, значения которых подставлены в формулу (7.109), вместо истинного значения регулярной составляющей $s_p(t)$ подставлено ее приближенное значение $s_p^0(t)$, но такое, что оно изменяется пропорционально $s_p(t)$, т. е. $s_p(t) = k s_p^0(t)$. Такая замена не вызовет изменения оценки $\hat{s}(t)$, т. е. не приведет к появлению дополнительной ошибки фильтрации, так как коэффициенты k^2 в числителе и знаменателе формулы (7.109) сократятся. В общем случае, когда коэффициенты определяются полными формулами (7.106) и (7.108), появятся коэффициенты $1/k$ у слагаемого $\hat{s}_{p1}(t)$, которое много меньше первого слагаемого $s_p(t)$, независимость $\hat{s}(t)$ от выбора расчетного значения $s_p^0(t)$ (коэффициента k) сохраняется с высокой степенью приближения. Условие же $s_p(t) = k s_p^0(t)$ легко выполняется, если изменять коэффициенты усиления $\psi(t)$ и $d(t)$ в соответствии с изменением оценки регулярной составляющей $\hat{s}_p(t)$.

Итак, для определения оптимальных характеристик вероятностного фильтра при условии самонастройки не требуется точно знать истинный характер регулярной составляющей — в этом достоинство вероятностного фильтра. Устранение же влияния априорного незнания регулярной составляющей в фильтре Калмана — Бьюси путем расширения вектора состояния (т. е. включение параметров регулярной составляющей в число неизвестных) влечет за собой значительное снижение точности фильтрации.

В заключение без выводов дадим матричное описание многомерного вероятностного фильтра.

Представим интегральное уравнение многомерного вероятностного фильтра в матричной форме

$$\int_0^t \mathbf{H}(t, t'') \mathbf{M}[\mathbf{x}(t'') \mathbf{x}^T(t')] dt'' = \mathbf{M}[\mathbf{s}(t) \mathbf{x}^T(t')] + \psi(t) \mathbf{s}_p^T(t'), \quad (7.110)$$

где

$$\mathbf{H}(t, t'') = \mathbf{H}_1(t, t'') + \psi(t) \mathbf{d}^T(t'') \quad (7.111)$$

— матричная весовая функция двух параллельных ветвей фильтра;

$$\psi(t) = \frac{1}{2} \left[\frac{\partial P(t)}{\partial \mathbf{R}(t)} \right]^{-1} \frac{\partial P(t)}{\partial \hat{\mathbf{s}}_p^T(t)} \quad (7.112)$$

— вектор весовых коэффициентов;

$$d^T(t) = [s_p(t) - C(t) \hat{s}_{p1}(t)]^T N^{-1}_0(t) \quad (7.113)$$

— транспонированный вектор коэффициентов ветви регулярной составляющей. Остальные обозначения прежние (см. § 7.7).

В пространстве состояний многомерный вероятностный фильтр описывается двумя векторно-матричными уравнениями

$$\frac{d}{dt} \hat{s}_1(t) = A(t) \hat{s}_1(t) + B(t) [x(t) - C(t) \hat{s}_1(t)], \quad \hat{s}_1(0) = 0, \quad (7.114)$$

$$\frac{d}{dt} [\Psi^T(t) \hat{s}_2(t)] = d^T(t) [x(t) - C(t) \hat{s}_1(t)], \quad (7.115)$$

где второе уравнение скалярное (одномерное). Нетрудно установить сходство приведенных формул с аналитическим описанием одномерного фильтра.

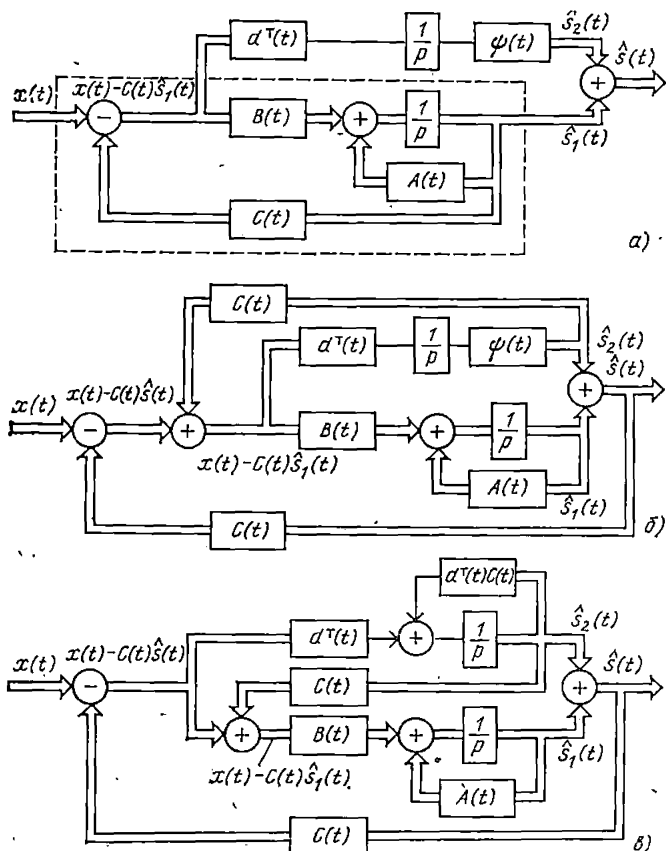


Рис. 7.15. Матричная схема многомерного вероятностного фильтра с обратной связью по неполной (а) и полной (б и в) реакции

Матричная структура многомерного вероятностного фильтра во всех его модификациях (рис. 7.15) имеет почти полное сходство со структурой одномерного фильтра (рис. 7.14). Отличие вызвано только включением матрицы $C(t)$ в цепь обратной связи. Следует также отметить, что цепь интегратора некалмановской ветви фильтра одномерная, что соответствует скалярному равенству (7.115) (умножение вектора на транспонированный вектор слева дает скаляр).

Аналогично можно определить структуру вероятностного фильтра-экстраполятора. Интегральное описание вероятностного фильтра-экстраполятора (как и фильтра Винера) не отличается от описания обычного фильтра, за исключением того, что в матричном интегральном уравнении (7.110) вместо текущего времени t нужно подставить упреждение $t+t_3$:

$$\int_0^{t+t_3} H(t+t_3, t'') M[x(t'') x^T(t')] dt'' = M[s(t+t_3) x^T(t')] + \psi(t+t_3) s^T(t'). \quad (7.116)$$

Особенность дифференциального описания вероятностного фильтра (как и калмановского) состоит в том, что он содержит цепь обратной связи и эта цепь должна быть замкнута по текущим значениям сравниваемых сигналов (на входе фильтра текущее значение), в то время как на выходе выдается оценка будущего значения. Нужно также иметь в виду, что предсказание относится только к случайной составляющей сигнала $s_{сл}(t)$. С уче-

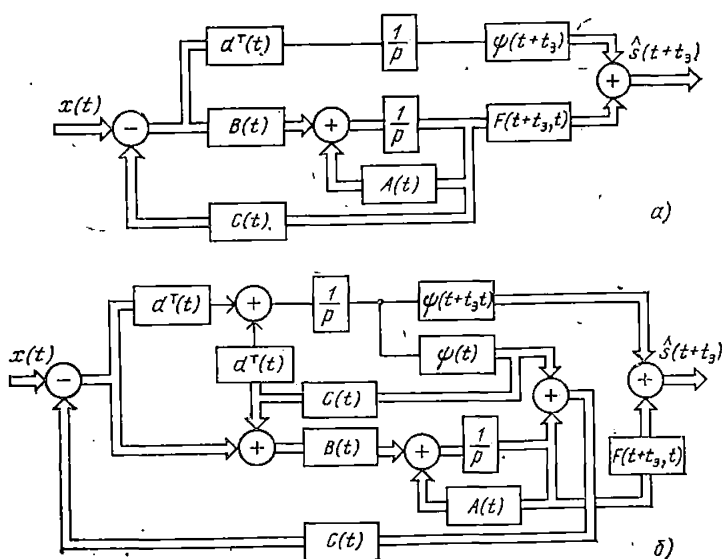


Рис. 7.16. Матричная схема многомерного вероятностного фильтра-экстраполятора с обратной связью по неполной (а) и полной (б) реакции

том этих обстоятельств из интегрального уравнения (7.116) можно получить дифференциальное уравнение многомерного вероятностного фильтра-экстраполятора с обратной связью по неполной реакции, матричная структура которого приведена на рис. 7.16,а (ср. с рис. 7.15,а). Здесь связь между текущим и будущим значениями выходного сигнала калмановской ветви фильтра выражается матрицей Коши $F(t+t_0, t)$, определяемой из дифференциального уравнения

$$\frac{d}{dt} F(t, t') = A(t) F(t, t'), \quad F(t, t) = I, \quad (7.117)$$

где I — единичная матрица. От этой схемы можно перейти к структуре вероятностного фильтра-экстраполятора, замкнутого по полной реакции (рис. 7.16,б).

Глава 8. НЕЛИНЕЙНАЯ ФИЛЬТРАЦИЯ И ИЗМЕРЕНИЕ ПАРАМЕТРОВ СИГНАЛА

8.1. НЕЛИНЕЙНАЯ ФИЛЬТРАЦИЯ ПАРАМЕТРОВ СИГНАЛА

Измеряемые радиотехническими системами параметры положения и движения объектов в пространстве — координаты и их производные — в общем случае представляют собой случайные процессы. Случайными процессами являются и параметры принимаемого от этих объектов сигнала, связанные с координатами и их производными: временной и частотный сдвиги, амплитудные и фазовые соотношения.

Случайный процесс $v_s(t)$ под воздействием помех (шумов) искажается, что при определенных условиях можно трактовать как наложение на него шумовой составляющей процесса $v_n(t)$. Назначением измерителя в составе измерительной системы является формирование оценки $\hat{v}(t)$ процесса $v(t) = v_s(t) + v_n(t)$. Эта оценка должна быть максимально близкой согласно некоторому критерию к истинному значению процесса $v_s(t)$. Таким образом, измеритель радиотехнической системы выполняет роль сглаживающего фильтра для процесса $v(t)$, закодированного в сигнале. Измеритель постоянного параметра v является частным примером такого фильтра при $v_s(t) = \text{const}$.

Особенность измеряемых процессов в радиосистемах состоит в том, что они закодированы в сигнале нелинейно. Действительно, все измеряемые параметры сигнала, характеризующие положение и движение объектов в пространстве, относятся к числу не-

энергетических и проявляются через фазовую модуляцию высокочастотного сигнала (волны)

$$\varphi(t) = 2\pi\mu_{\text{ЭК}}v_s(t), \quad (8.1)$$

где $\mu_{\text{ЭК}}$ — частота эквивалентной гармонической огибающей. При описании сигнала обобщенной комплексной огибающей

$$S[v_s(t)] = A(t) e^{j\varphi(t)} = A(t) e^{j2\pi\mu_{\text{ЭК}}v_s(t)}, \quad (8.2)$$

измеряемый процесс входит только в показатель экспоненты, поэтому неэнергетический параметр не оказывает влияния на амплитудный множитель $A(t)$. Соотношение (8.2) как раз указывает на существенно нелинейную связь между сигналом и измеряемым процессом. Ввиду этого измеритель процесса $v_s(t)$, закодированного во входном сигнале как функции времени и пространства $x(t, r) = s[t, r; v_s(t)] + n(t, r)$, следует рассматривать как нелинейный фильтр, сглаживающий искажения процесса $v_s(t, r)$, вызванные шумовой составляющей входного сигнала (волны) $v_n(t, r)$.

Строгое решение задачи нелинейной фильтрации связано с большими математическими трудностями и получено лишь для ограниченного числа простейших ситуаций. Поэтому используют приближенное решение, с достаточной степенью точности справедливое при точных измерениях, которые как раз и представляет основной интерес при синтезе радиосистем. Дело в том, что высокая точность измерений (как и высокое качество обнаружения) может быть обеспечена лишь при достаточно большом отношении сигнал-шум. Однако известно, что при наложении гауссовских шумов с дисперсией σ_n^2 на высокочастотный квазигармонический сигнал его амплитуда A (огибающая) и фаза φ в линейном тракте приема при достаточно большом отношении сигнал-шум $a = A/\sigma_n > 3 \dots 5$ также имеют гауссовское распределение, хотя и являются нелинейными функциями входного сигнала (рис. 8.1). Это обстоятельство можно трактовать как непосредственное наложение входных гауссовских шумов на амплитуду и фазу сигнала, т. е. считать шумы амплитуды и фазы высокочастотного сигнала аддитивными, что соответствует линейному взаимодействию сигнала и шума. Так как измеряемый процесс $v_s(t)$ пропорционален

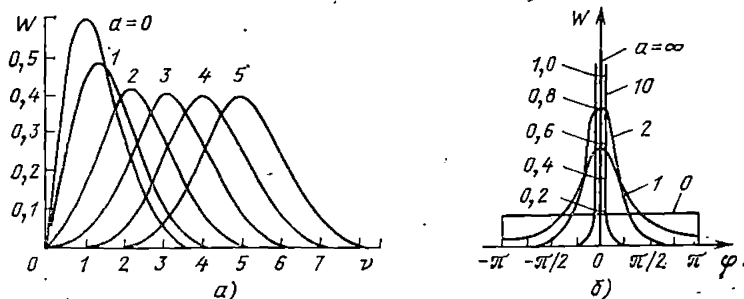


Рис. 8.1 Распределение амплитуды (а) и фазы (б) синусоидального сигнала с шумами в зависимости от отношения сигнал-шум $a = A/\sigma_n$

текущей фазе (8.1), взаимодействие сигнала и гауссовского шума можно трактовать как непосредственное наложение гауссовских шумов $v_n(t)$ на процесс $v_s(t)$. Тем самым задача точного измерения изменяющегося во времени параметра сигнала сводится к линейной фильтрации процесса $v(t)$.

Оптимальный линейный замкнутый фильтр для входного процесса $v(t)$ в соответствии с изложенными в гл. 7 положениями можно представить состоящим из двух частей: компаратора и формирователя оценки (рис. 8.2, а). Формирователь оценки, используя механизм обратной связи через компаратор, формирует оценку процесса $\hat{v}(t)$ по любому из ранее описанных алгоритмов и имеет соответствующую структуру.

В действительности на вход фильтра поступает не сам процесс $v(t)$, а смесь сигнала и шума $x[t, \tau; v(t)] = s[t, \tau; v_s(t)] + n(t, \tau)$, являющаяся в общем случае функцией времени и координат раскрыва антенны, которая (как отмечалось) связана с фильтруемым процессом нелинейно. Поэтому непосредственное формирование рассогласования $v(t) - \hat{v}(t)$ путем вычитания в компараторе невозможно. Необходимо произвести нелинейное преобразование входной смеси $x(t, \tau)$ в линейную разность $v(t) - \hat{v}(t)$ — сигнал рассогласования по измеряемому параметру. Это преобразование осуществляется дискриминатором. Поэтому нелинейный замкнутый фильтр для процесса $v(t)$ можно представить в виде двух частей: линейной и нелинейной (рис. 8.2, б). Линейную часть по-прежнему представляет формирователь оценки; а нелинейную — дискриминатор, через который замыкается цепь обратной связи. Если рассматривать процессы только в самой цепи, отвлекаясь от входного воздействия, то нетрудно заметить, что они в линейном и нелинейном фильтрах полностью совпадают.

К синтезу оптимального нелинейного фильтра применим принцип декомпозиции: линейную и нелинейную части можно оптимизировать независимо. Оптимальная структура линейной части — формирователя оценки процесса — уже определена в гл. 7 для различных ситуаций. Задача оптимального дискриминатора — формирование из входного процесса $x(t, \tau)$ сигнала рассогласования $v(t) - \hat{v}(t)$ с наименьшей погрешностью.

При отыскании структуры оптимального дискриминатора следует иметь в виду, что сигнал в радиосистемах обычно представ-

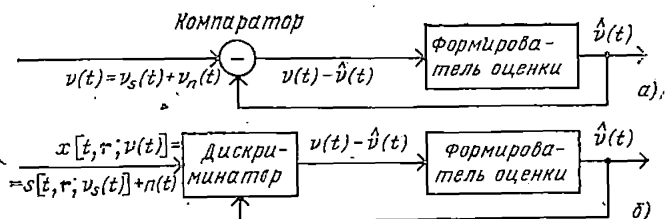


Рис. 8.2 Обобщенная схема замкнутого фильтра

ляет собой последовательность коротких импульсов. Это в равной мере относится и к непрерывным периодическим последовательностям, которые после прохождения через линейную цепь оптимального приемника принимают вид коротких импульсов. В пределах такого импульса любой измеряемый параметр сигнала можно считать практически постоянным. Таким образом, фильтруемый процесс, закодированный в сигнале, почти всегда представлен серией дискретных отсчетов этого процесса, а фильтрация сводится к нахождению некоторой плавной функции времени, в среднем максимально близкой к этим отсчетам с учетом корреляционной связи между ними. Так как оцениваемый процесс нелинейно закодирован в высокочастотном сигнале, дискретные отсчеты его формируются в результате выполнения операции оценки постоянного параметра сигнала в каждом импульсе (оценка в точке), в оптимальном случае — по методу максимального правдоподобия.

Для оценки малого постоянного рассогласования $v - v^0$ по измеряемому параметру методом максимального правдоподобия воспользуемся огибающей корреляционного частного сигнала (импульса) $Z(v - v^0)$, которую разложим в ряд Тейлора относительно некоторого опорного значения v^0 , достаточно близкого к истинному значению параметра v_s . Ввиду малости отклонений ограничимся тремя членами ряда:

$$Z(v - v^0) \approx Z(v^0) + Z'(v^0)(v - v^0) + Z''(v^0)(v - v^0)^2/2 \quad (8.3)$$

и составим уравнение правдоподобия, взяв от огибающей (8.3) первую производную по измеряемому параметру и приравняв ее нулю:

$$Z'(v - v^0) \approx Z'(v^0) + Z''(v^0)(v - v^0) = 0. \quad (8.4)$$

Отсюда находим оптимальную оценку малого рассогласования

$$v' = v - v^0 = -Z'(v^0)/Z''(v^0). \quad (8.5)$$

Оценка (8.5) обращает в нуль первую производную огибающей корреляционного интеграла в точке v , которая заранее неизвестна. Дискриминатор удобнее строить таким образом, чтобы на его выходе сигнал рассогласования обращался в нуль в опорной точке v^0 (т. е. в точке его настройки), а в точке оценки v первая производная была отлична от нуля. Заменив в формуле (8.5) v^0 на v и соответственно изменив знак первой производной, находим оптимальную оценку постоянного рассогласования в дискриминаторе

$$v' = v - v^0 = Z'(v)/Z''(v). \quad (8.6)$$

Формула (8.6) отображает алгоритм работы и структуру оптимального дискриминатора, вырабатывающего оценку рассогласования (оценку в точке) с минимальной погрешностью (дисперсией).

Нетрудно видеть, что если в качестве опорного значения ν^0 в дискриминатор ввести оценку $\hat{\nu}$ от формирователя оценки, то на его выходе вырабатывается отсчет разности $\nu - \hat{\nu}$, которая, изменяясь с течением времени, дает текущее значение сигнала рассогласования $\nu(t) - \hat{\nu}(t)$, как и на выходе компаратора линейного фильтра. При этом дискретный, импульсный характер входных сигналов дискриминатора обычно не сказывается на его выходном сигнале благодаря фильтру нижних частот (ФНЧ), входящему в состав дискриминатора. Действительно, спектр импульсных выборок процесса отличается от спектра самого процесса только наличием дополнительных спектральных полос в области частот, кратных частоте повторения импульсов. Эти полосы подавляются ФНЧ, если ширина спектра процесса значительно меньше частоты повторения импульсов, что обычно выполняется на практике.

Свойства оптимального дискриминатора определяются зависимостью его выходного сигнала (8.6) от рассогласования по измеряемому параметру при условии, что шум отсутствует, т. е. $\nu_n = 0$ и $Z(\nu) = Z_s(\nu)$:

$$D(\nu - \nu^0) = Z'_s(\nu) / Z''_s(\nu). \quad (8.7)$$

Формирование характеристики оптимального дискриминатора иллюстрируется рис. 8.3, где штриховыми линиями показан ход кривых при параболической аппроксимации огибающей корреляционного интеграла

$$Z_{\text{пар}}(\nu - \nu^0) = k[1 - a(\nu - \nu^0)^2/2], \quad (8.8)$$

где $a = -Z''_s(\nu^0)/k$ — абсолютное значение второй производной огибающей корреляционного интеграла в точке настройки (крутизна характеристики дискриминатора). Параболическая аппроксимация обеспечивает линейность характеристики дискриминатора в бесконечных пределах. Действительно, $Z'_{\text{пар}}(\nu - \nu^0) = -ka(\nu - \nu^0)$ и $Z''_{\text{пар}}(\nu - \nu^0) = -ka$, откуда следует, что выходной сигнал такого дискриминатора

$$\begin{aligned} D_{\text{пар}}(\nu - \nu^0) &= \\ &= Z'_{\text{пар}}(\nu - \nu^0) / Z''_{\text{пар}}(\nu - \nu^0) = \\ &= \nu - \nu^0 \end{aligned} \quad (8.9)$$

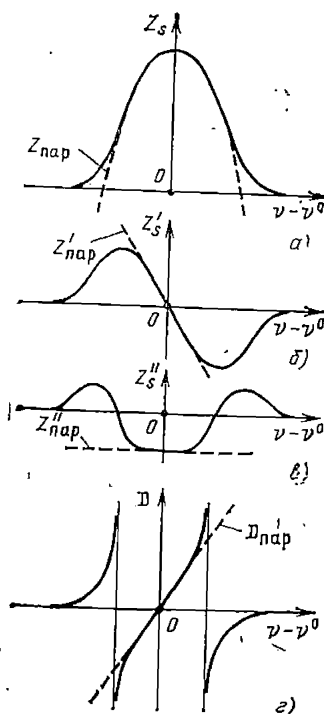


Рис. 8.3. Огибающая корреляционного интеграла (а), ее первая (б) и вторая (в) производные в отсутствие шумов и характеристика оптимального дискриминатора (г)

равен сигналу рассогласования $v' = v - v^0$ при любых v . Однако реальную огибающую корреляционного интеграла при больших рассогласованиях v' нельзя представить параболой, ветви которой уходят в отрицательную область (огibaющая — существенно положительная функция). Поэтому параболическую аппроксимацию огибающей и, следовательно, линейную характеристику дискриминатора можно реализовать лишь для малых рассогласований. Приблизненно размер рабочего участка характеристики дискриминатора, где ее с заданной точностью можно считать линейной, равен интервалу корреляции сигнала по измеряемому параметру $\Delta v = 1/\Delta\mu_{эф}$, где $\Delta\mu_{эф}$ — эффективная протяженность обобщенного спектра.

Благодаря нормировке — делению Z' на Z'' — выходной сигнал дискриминатора является функцией только самого рассогласования по измеряемому параметру независимо от амплитуды принимаемого сигнала (в частности, у параболической аппроксимации огибающей корреляционного интеграла (8.9) амплитудный коэффициент k входит в числитель и знаменатель отношения). Так как амплитуда сигналов на входе приемника изменяется в широком диапазоне, выполнение нормировки является сложной технической задачей. Необходимо отметить, что реальная характеристика дискриминатора в отличие от изображенной на рис. 8.3,2 не уходит в бесконечность при $Z''_s = 0$, поскольку коэффициент передачи дискриминатора всегда ограничен. Необходимо также отметить еще одну особенность операций в дискриминаторе: на вход его поступает высокочастотный сигнал или электромагнитная волна, а на выходе формируется видеосигнал рассогласования, полярность (знак) которого зависит от стороны рассогласования. Для сохранения знака выходного сигнала используется синхронный детектор, на который подается опорное напряжение из основного канала приема, где формируется сам корреляционный интеграл.

8.2. СЛЕДЯЩИЙ И РАЗОМКНУТЫЙ ИЗМЕРИТЕЛИ

Рассмотренная выше структурная схема нелинейного фильтра не позволяет ответить на два вопроса, связанных с практической реализацией фильтра: как обеспечивается настройка характеристики дискриминатора на опорное значение параметра, снимаемое с выхода формирователя оценки, и как устанавливается начальное опорное значение в дискриминаторе, когда оценка еще отсутствует, чтобы попасть в пределы линейного рабочего участка его характеристики?

Ответить на первый вопрос можно, только зная конкретную структуру дискриминатора. Дискриминатор, как и любое другое устройство радиосистемы, на которое подается смесь сигнала и шума, выполняется в виде фильтра или коррелятора. Свойства коррелятора полностью определяются видом опорного сигнала.

Для настройки дискриминатора типа коррелятора необходимо синтезировать опорный сигнал требуемого вида, соответствующий параметр которого в качестве опорного задается формирователем оценки. Свойства дискриминатора типа фильтра определяются его структурой, включающей элемент настройки на заданное формирование оценки опорное значение параметра, поэтому необходимо устройство, обеспечивающее эту настройку. Таким образом, в цепи обратной связи между формирователем оценки и дискриминатором необходимо иметь вспомогательный элемент — синтезатор. Синтезатор обеспечивает (синтезирует) требуемую характеристику дискриминатора, формируя опорный сигнал коррелятора или производя настройку фильтра в соответствии с данными формирователя оценки. Синтезатор выполняет важную практически, но несущественную с информационной точки зрения операцию перекодирования выходного сигнала формирователя оценки в опорную характеристику дискриминатора — опорный сигнал коррелятора или импульсную характеристику фильтра. Поэтому к нему предъявляется единственное требование — не вносить дополнительных ошибок.

Оценка параметра и, следовательно, настройка дискриминатора непрерывно изменяется вслед за изменениями самого измеряемого параметра входного сигнала. Поэтому замкнутые нелинейные фильтры получили название *слеящих измерителей* (устаревшее название — системы автоматического сопровождения). Вследствие инерционности замкнутой цепи слежения оценка повторяет не все изменения отслеживаемого параметра, искаженного шумами, а только плавные, закономерные, что эквивалентно непосредственной фильтрации процесса. Благодаря перестройке в соответствии с закономерными изменениями параметра входного сигнала работа дискриминатора протекает на линейном участке его характеристики. В результате на выходе дискриминатора вырабатывается сигнал рассогласования, равный разности $v(t) - \hat{v}(t)$, как и в компараторе линейного фильтра. Тем самым слеящий измеритель, нелинейный по отношению к самому сигналу, подобен линейному фильтру для процесса $v(t)$, закодированного во входном сигнале.

Для решения второго вопроса рассмотренная ранее схема нелинейного фильтра дополняется схемой поиска и захвата (СПЗ). Таким образом, слеящий измеритель состоит из кольца слежения, включающего дискриминатор, формирователь оценки и синтезатор, и подключенной параллельно дискриминатору СПЗ, включающей обнаружитель, реле захвата и схему поиска (рис. 8.4).

Слеящий измеритель работает в двух режимах: режиме поиска (вспомогательный) и режиме слежения (основной).

В режиме поиска кольцо слежения разомкнуто — дискриминатор расстроен относительно неизвестного значения параметра — и сигнал рассогласования на его выходе отсутствует. Схема поиска принудительно изменяет опорное значение v^0 обнаружителя

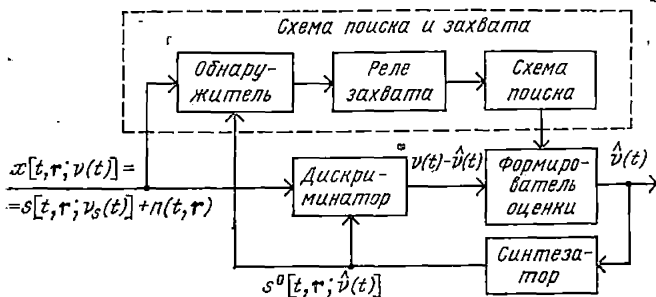


Рис. 8.4. Схема слеящего измерителя процесса, закодированного в сигнале

СПЗ и дискриминатора. Обычно перестройка производится по линейному закону, для чего от схемы поиска СПЗ на последний интегратор формирователя оценки подается постоянное напряжение. Пределы перестройки устанавливаются на основе априорных данных о возможном значении отслеживаемого параметра; если таких данных нет, перестройка производится по всему диапазону его возможных значений.

Обнаружитель, на который параллельно дискриминатору поступает входная смесь сигнала и шума, в оптимальном случае вырабатывает огибающую корреляционного интеграла $Z(v-v^0)$. В процессе принудительной перестройки опорное значение v^0 в некоторый момент времени оказывается близким к фактическому значению параметра v входного сигнала. В этот момент огибающая корреляционного интеграла $Z(v-v^0)$ на выходе обнаружителя близка к максимальной. В случае превышения заданного порога обнаружения реле захвата срабатывает и прекращает работу схемы поиска — снимается постоянное напряжение, подводимое ко входу интегратора в формирователе оценки. Поскольку в этот момент рассогласование $v-v^0$ мало, входной сигнал попадает на рабочий участок характеристики дискриминатора, кольцо слежения замыкается, и через некоторое время (время установления) следающий измеритель переходит в основной режим — режим точного слежения. В режиме слежения в качестве опорного значения v^0 используется оценка $\hat{v}$ и дискриминатор вырабатывает сигнал рассогласования $v(t)-\hat{v}(t)$. В случае длительного пропадания входного сигнала или ошибочного захвата реле захвата отпускает и вновь начинается режим поиска.

Наряду с замкнутыми следающими измерителями в радиосистемах широко используют *разомкнутые измерители*, предназначенные для оценки текущих значений измеряемых параметров сигнала в процессе обзора заданной зоны пространства. Их особенность состоит в том, что измеренные от обзора к обзору значения параметров никакому сглаживанию не подвергаются, причем измеряются параметры сигналов не одного объекта, как в следающих измерителях, а всех объектов, находящихся в зоне обзора.

Для оценки текущего значения параметра сигнала в каждом данном цикле обзора можно использовать дискриминаторы, если в них установить опорные значения параметров, достаточно близкие к измеряемым. Однако, поскольку измеряемый параметр может принимать любые значения в пределах диапазона измерений, разомкнутый измеритель должен состоять из большого числа дискриминаторов, у которых линейные участки характеристик вплотную примыкают друг к другу, перекрывая весь диапазон измерений, т. е. измеритель должен быть многоканальным и иметь на выходе анализатор для стыковки показаний на границе смежных каналов. Основная трудность реализации таких измерителей возникает при выполнении операций нормировки, т. е. формировании отношения Z'/Z'' . Во-первых, нормировка в двух соседних элементах дальности должна выполняться практически мгновенно. Во-вторых, она должна быть идеальной, иначе точность измерения параметров $v = v^0_i + v'_i$ в дискриминаторе i -го канала разомкнутого измерителя, настроенного на опорное значение v^0_i , резко снижается из-за колебаний измеренного рассогласования v'_i с изменением интенсивности принимаемого сигнала. В следящих измерителях инерционность и неидеальность нормировки, выполняемой обычно с помощью простых схем автоматической регулировки усиления (АРУ), проявляется значительно слабее, поскольку слежение производится за медленно изменяющимся сигналом единственного объекта и нормальному режиму слежения соответствует нулевое рассогласование v' .

Многоканальные измерители можно заменить последовательным одноканальным в виде дискриминатора, настройка которого на опорное значение изменяется во времени. Однако при этом возрастает время измерений (период обзора) во столько раз, сколько каналов заменяет такой измеритель. Иногда перестройка связана со значительными энергетическими затратами, а сложность выполнения операции нормировки в одноканальном измерителе не уменьшается.

Реализация разомкнутых измерителей существенно упрощается, если в процессе обзора некоторые измеряемые параметры преобразуются естественно или принудительно во временное положение принимаемого сигнала. Так, дальность естественным образом преобразуется во временной интервал за счет конечной и постоянной скорости распространения радиоволн, а измерение угловой координаты также можно свести к измерению временного интервала, если антенный луч принудительно перемещать с постоянной угловой скоростью. При этом для измерений во всем диапазоне можно использовать единственный согласованный фильтр, и отсчет параметра производить по моменту наступления максимума огибающей корреляционного интеграла — выходного сигнала оптимального приемника.

Если при этом съем данных производится оператором с экрана индикатора, у которого имеется линейная развертка по измеряемому параметру и нанесена шкала, оператор, сравнивая положение

ние центра отметки сигнала с делениями шкалы, выдает оценку параметра с достаточно высокой точностью. При этом радиосистема имеет единственный канал обнаружения и измерения.

При автоматическом измерении параметра, преобразованного во временной интервал, для реализации высокой точности измерений необходимо иметь устройство фиксации центра (максимума) выходного сигнала в дополнение к каналу обнаружения. В этом устройстве огибающая корреляционного интеграла дифференцируется и в момент перехода ее через нуль вырабатывается короткий импульс отсчета. Серьезным достоинством такого измерителя является то, что в принципе не требуется нормировка, так как момент перехода через нуль не зависит от интенсивности входного сигнала. Однако скорость перехода через нуль (крутизна) зависит от интенсивности сигнала, а это может вызвать дополнительную погрешность в реальной схеме, вырабатывающей импульс отсчета, которая срабатывает тем быстрее, чем больше крутизна. Указанный способ можно непосредственно использовать в разомкнутом измерителе дальности (рис. 8.5,а). Однако при измерении угловой координаты по сигналу, представленному пакетом импульсов, формирование производной путем дифференцирования огибающей выходного пакета невозможно, так как са-

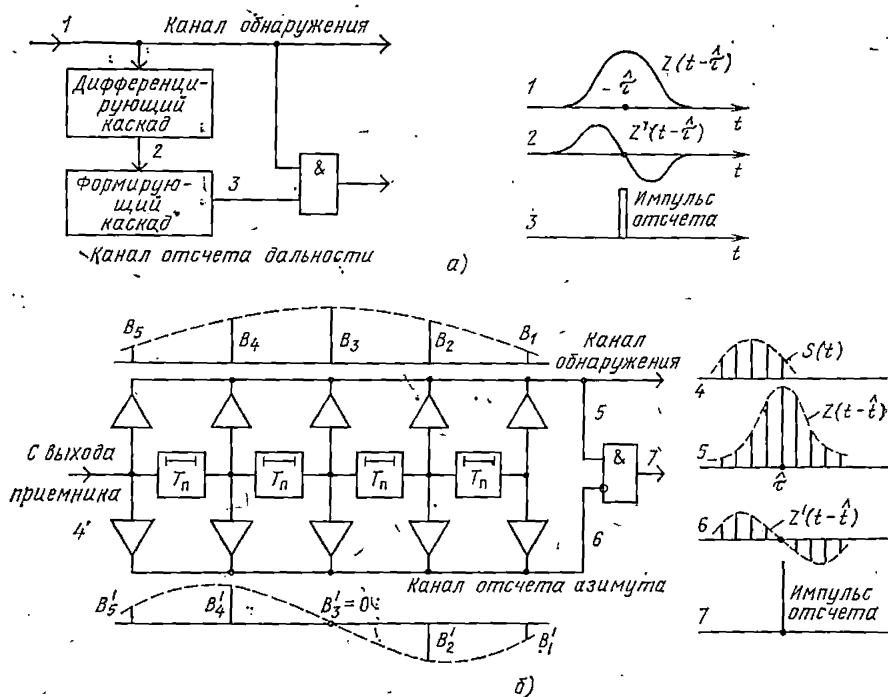


Рис. 8.5. Каналы обнаружения и формирования импульсов отсчета дальности (а) и азимута (б)

мой огибающей реально не существует. Для получения производной огибающей выходного пакета используют гребенчатый фильтр с нечетной весовой функцией в отличие от интегратора импульсов в канале обнаружения, имеющего четную, весовую функцию (рис. 8.5,б).

К сожалению, к отсчету времени можно свести только измерение дальности и одной из двух угловых координат. Для измерения второй угловой координаты в процессе обзора необходим более сложный многоканальный разомкнутый измеритель. Поэтому радиосистемы с трехкоординатными разомкнутыми измерителями (дальность — азимут — угол места) по сложности и стоимости значительно превосходят двухкоординатные (дальность — азимут).

Тип и структура разомкнутых измерителей в решающей степени определяются конкретной схемой радиосистемы и видом используемого в ней оконечного устройства. Следящие же измерители, наоборот, сами определяют схему построения радиосистемы.

8.3. ОПТИМАЛЬНЫЙ ДИСКРИМИНАТОР

Оптимальный дискриминатор вырабатывает оценку (8.6) малых рассогласований между текущим и опорным значениями измеряемого параметра сигнала в каждом импульсе. Для получения сигнала рассогласования (8.6) из входной смеси в дискриминаторе необходимо сформировать первую и вторую производные огибающей корреляционного интеграла, а в обнаружителе СПЗ саму огибающую.

При определении структуры оптимального дискриминатора (и обнаружителя) ограничимся случаем симметричного энергетического спектра $|S(\mu)|^2$, при котором точность измерений максимальна, а огибающая корреляционного интеграла и обе ее производные вещественные. Тогда в оптимальном следящем измерителе из входной смеси сигнала и шума, описываемой огибающей как функцией измеряемого параметра $X(v) = U_0 S(v_s) + N$, на выходе трех параллельных каналов, настроенных на опорное значение v^0 , получаем три вещественные функции рассогласования, описываемые либо через обобщенную огибающую сигнала $S(v)$, либо через ее спектр $S(\mu)$:

$$\begin{aligned} Z(v - v^0) &= \frac{1}{2} \int_{-\infty}^{\infty} X(v'' - v) S^*(v'' - v^0) dv'' = \\ &= \frac{1}{2} \int_{-\infty}^{\infty} X^*(\mu) S(\mu) e^{j 2\pi \mu (v - v^0)} d\mu; \end{aligned} \quad (8.10)$$

$$\begin{aligned} Z'(v - v^0) &= \frac{1}{2} \int_{-\infty}^{\infty} X(v'' - v) \frac{d}{dv^0} S^*(v'' - v^0) dv'' = \\ &= j \frac{2\pi}{2} \int_{-\infty}^{\infty} \mu X^*(\mu) S(\mu) e^{j 2\pi \mu (v - v^0)} d\mu, \end{aligned} \quad (8.11)$$

$$Z''(v-v^0) = \frac{1}{2} \int_{-\infty}^{\infty} X(v''-v) \frac{d^2}{dv^0{}^2} S^*(v''-v^0) dv'' =$$

$$= -\frac{(2\pi)^2}{2} \int_{-\infty}^{\infty} \mu^2 X^*(\mu) S(\mu) e^{j2\pi\mu(v-v^0)} d\mu. \quad (8.12)$$

Операции (8.10)—(8.12) и (8.6) полностью определяют структуру оптимального дискриминатора (и обнаружителя) следящего измерителя (рис. 8.6). Эти операции можно реализовать в фильтрах или корреляторах, настраиваемых синтезатором на опорное значение параметра v^0 . Поскольку выходные сигналы всех трех фильтров являются высокочастотными колебаниями, для выделения огибающих с сохранением знака первой производной на выходе этого канала включают синхронный детектор, в остальных двух — обычный детектор.

Свойства фильтров, заложенные в их структуре, описываются функциями передачи

$$H_0(\mu) = kS^*(\mu), H_1(\mu) = j2\pi k\mu S^*(\mu), H_2(\mu) =$$

$$= -(2\pi)^2 k\mu^2 S^*(\mu) \quad (8.13)$$

или, что равносильно, импульсными характеристиками

$$H_0(v) = kS^*(v), H_1(v) = k \frac{d}{dv} S^*(v), H_2(v) = k \frac{d^2}{dv^2} S^*(v) \quad (8.14)$$

для каналов огибающей корреляционного интеграла и его первой и второй производных соответственно. В этом случае синтезатор изменяет только параметр настройки фильтров v^0 . Корреляторы всех трех каналов совершенно идентичны, а их свойства целиком определяются опорными сигналами вида (8.14), поступающими от синтезатора, соответствующий параметр которых равен опорному значению v^0 . В режиме поиска опорное значение измеряемого па-

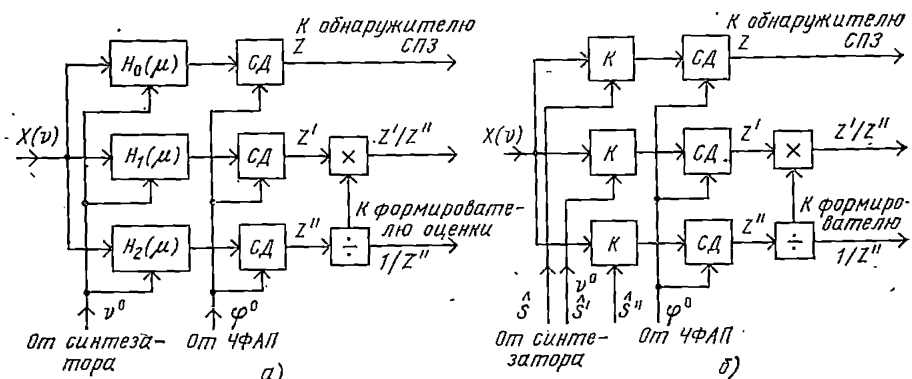


Рис. 8.6. Схемы оптимального дискриминатора (и входного каскада обнаружителя)

параметра задается (перестраивается) схемой поиска, а в режиме слежения оно повторяет текущую оценку параметра: $v^0(t) = \hat{v}(t)$.

Опорное напряжение синхронных детекторов формируется специальной схемой частотно-фазовой автоподстройки (ЧФАП) синфазно с принимаемым сигналом. Сами корреляторы могут выполнять функции синхронных детекторов, если несущие их опорных сигналов совпадают по частоте и фазе с несущей входного сигнала. Однако обычно приемники строятся по супергетеродинной схеме, где основное усиление и обработка сигналов производятся в УПЧ (что для пояснения основной идеи не отражено на схеме), поэтому функции формирования опорных и детектирования выходных сигналов корреляторов разделяются: опорные сигналы корреляторов подаются на частоте гетеродина, а опорные напряжения синхронных детекторов на промежуточной частоте.

Операция нормировки (8.6) обычно выполняется в усилительном канале первой производной, коэффициент передачи которого изменяется обратно пропорционально второй производной (на схеме рис. 8.6 это обозначено операцией деления).

Оптимальный дискриминатор наряду с сигналом рассогласования Z'/Z'' должен выдать данные для настройки формирователя оценки на параметры входного процесса. В частности, для обработки параметра $b = \sigma_e^2 / q_v$ следящего измерителя типа фильтра Калмана необходимо оценить спектральную плотность $q_v = \sigma_v^2 / \Delta f$ шумовой составляющей оцениваемого процесса v_n , ширина спектра которой Δf во много раз больше ширины полосы пропускания самого следящего измерителя, ввиду чего составляющую v_n можно аппроксимировать белым шумом. Дисперсия оценки рассогласования в дискриминаторе $\sigma_v^2 = 1/l'' = 1/Z''$ и, следовательно, спектральная плотность входного шума q_v в соответствии с формулой (6.32) определяется средним значением второй производной логарифма отношения правдоподобия l'' или огибающей корреляционного интеграла Z'' . Поэтому кроме рассогласования Z'/Z'' с выхода дискриминатора в формирователь оценки необходимо выдать величину Z'' , являющуюся мерой его точности.

Наконец, сигналы, пропорциональные огибающей корреляционного интеграла Z в каждом импульсе, подаются на обнаружитель СПЗ, который представляет собой гребенчатый фильтр, интегрирующий эти импульсы. Если выходной сигнал этого фильтра превышает порог срабатывания реле захвата, следящий измеритель переходит из режима поиска в режим слежения.

Для определения основных свойств оптимального дискриминатора, частично уже отмеченных выше, воспользуемся представлением функции (8.10) — (8.12) через обобщенный энергетический спектр $|S(\mu)|^2$, с учетом симметрии которого и формулы Эйлера в отсутствие шумов получим

$$Z_s(v - v^0) = k U_0 \int_0^\infty |S(\mu)|^2 \cos 2\pi\mu(v - v^0) d\mu, \dots$$

$$Z'_s(v-v^0) = -2\pi kU_0 \int_0^\infty \mu |S(\mu)|^2 \sin 2\pi\mu(v-v^0) d\mu,$$

$$Z''_s(v-v^0) = -(2\pi)^2 kU_0 \int_0^\infty \mu^2 |S(\mu)|^2 \cos 2\pi\mu(v-v^0) d\mu, \quad (8.15)$$

где интегралы берутся только по области положительных частот.

Из анализа полученных формул непосредственно вытекают основные свойства характеристики оптимального дискриминатора (8.7).

1. Нечетная симметрия и линейность при малых рассогласованиях: положив в формулах (8.15) $\sin x \approx x$ и $\cos x \approx 1$, убеждаемся, что Z'_s отличается от Z''_s только множителем $v-v^0$, откуда их отношение $D(v-v^0) \approx v-v^0$.

2. Нормировка: отношение Z'_s/Z''_s не зависит от амплитуды входного сигнала U_0 , так как числитель и знаменатель содержат общий множитель U_0 .

3. Фазирование, обеспечивающее вещественный характер сигнала рассогласования: все три функции (8.15) вещественные, а соответствующие им высокочастотные сигналы находятся в фазе или противофазе, что позволяет по соотношению фаз этих сигналов определить знак (сторону) рассогласования. Комплексная функция не может быть мерой рассогласования, так как к комплексным числам понятия больше или меньше неприменимы.

4. Оценка точности: величина Z''_s , как отмечалось, характеризует интенсивность помеховой составляющей входного процесса следящего измерителя.

Сохранение указанных свойств дискриминаторов позволяет упростить аппаратуру, практически не снижая точности измерений. Так, для оценки точности нет необходимости вычислять Z''_s , поскольку интенсивность (мощность) помеховой составляющей входного процесса можно измерить непосредственно на выходе дискриминатора (свойство 4). Учитывая также, что огибающая корреляционного интеграла и ее вторая производная при малых рассогласованиях ведут себя практически одинаково и пропорциональны амплитуде принимаемого сигнала, только имеют противоположный знак, можно вообще отказаться от вычислений второй производной, используя для нормировки (свойство 2) вместо отношения Z'/Z'' отношение $-Z'/Z$. При этом

$$D(v-v^0) = -k_1 Z'_s(v-v^0)/Z_s(v-v^0), \quad (8.16)$$

где коэффициент k_1 компенсирует различие $|Z''_s|$ и Z_s при нулевой расстройке. Такой дискриминатор (с учетом только что указанных особенностей работы следящих измерителей в реальных условиях) обладает всеми свойствами оптимального дискриминатора, поэтому назовем его *модифицированным оптимальным*.

Модифицированный оптимальный дискриминатор имеет целый ряд преимуществ перед описанным выше оптимальным дискриминатором:

тором, перестраивающим фильтр или вырабатывающим опорный сигнал коррелятора.

Обычно опорный канал вместе со схемой АРУ относят к общему приемному тракту радиосистемы, а собственно дискриминатором считают только измерительный канал. Приемные каналы открываются селектором только на время прихода очередного импульса с помощью стробирующего импульса, также формируемого синтезатором. Благодаря селектору простейший накопитель (конденсатор) на его выходе выполняет роль гребенчатого фильтра-интегратора в обнаружителе СПЗ. При этом схема АРУ работает только по сигналу отслеживаемого объекта, выполняя регулировку усиления по среднему значению сигнала — для сглаживания шумовых всплесков (инерционная АРУ). Следует, однако, отметить, что нормировка с помощью схемы АРУ далеко не идеальна. Поэтому иногда дополнительно используют более совершенные схемы нормировки. Опорное напряжение синхронных детекторов вырабатывается схемой ЧФАП также по сигналу основного канала.

8.4. ПОДОПТИМАЛЬНЫЕ ДИСКРИМИНАТОРЫ

На практике широко применяют подоптимальные дискриминаторы, которые проще оптимальных в реализации, но лишь незначительно уступают им по своим параметрам. Независимо от характера операций над входными сигналами выходные сигналы этих дискриминаторов должны удовлетворять условиям нормировки, фазирования и нечетной симметрии. Выходной сигнал всех подоптимальных дискриминаторов формируется в результате сравнения сигналов, различающихся характером зависимости от измеряемого параметра. В зависимости от способа сравнения сигналов подоптимальные дискриминаторы подразделяют на две большие группы:

линейные, или аддитивные, основанные на сравнении разности или суммы разделяющихся сигналов;

нелинейные, или мультипликативные, использующие для сравнения отношение или произведение сигналов.

Линейные подоптимальные дискриминаторы. Здесь вместо первой производной огибающей корреляционного интеграла используют конечную разность двух корреляционных интегралов; алгебраическую (вычитание амплитуд выходных синфазных сигналов, разнесенных по самому измеряемому параметру ν) или геометрическую (вычитание векторов сдвинутых по фазе сигналов, разнесенных по обобщенной частоте μ и имеющих одинаковую амплитуду). В соответствии с этим линейные подоптимальные дискриминаторы подразделяют на амплитудные и фазовые.

В амплитудных линейных дискриминаторах вместо производной огибающей корреляционного интеграла как предела отноше-

$$Z'(\nu - \nu^0) = \lim_{\Delta \nu_0 \rightarrow 0} \frac{1}{\Delta \nu_0} \left[Z\left(\nu - \nu^0 + \frac{\Delta \nu_0}{2}\right) - Z\left(\nu - \nu^0 - \frac{\Delta \nu_0}{2}\right) \right]$$

используют конечную разность (обратную)

$$U_{\Delta} = Z \left(v - v^0 - \frac{\Delta v_0}{2} \right) - Z \left(v - v^0 + \frac{\Delta v_0}{2} \right) \quad (8.17)$$

при фиксированной расстройке приемных каналов Δv_0 , симметричной относительно опорного значения v^0 , т. е. точки перехода характеристики дискриминатора через нуль. Отсюда в соответствии с формулой (8.16) находим характеристику амплитудного дискриминатора

$$D(v - v^0) = k_2 \frac{Z_s(v - v^0 - \Delta v_0/2) - Z_s(v - v^0 + \Delta v_0/2)}{Z_s(v - v^0)} \quad (8.18)$$

С уменьшением расстройки Δv_0 характеристика (8.18) при $k_2 = = k_1/\Delta v_0$ приближается к оптимальной (8.16), однако разностный сигнал убывает и все большую роль играют шумы. С увеличением Δv_0 , когда $\Delta v_0/2 > v_{\text{прг}}$ (абсцисса точки перегиба огибающей корреляционного интеграла $v_{\text{прг}}$ находится из условия $Z''_s(v' = v_{\text{прг}}) = 0$), характеристика дискриминатора при малых рассогласованиях становится нелинейной и разностный сигнал также убывает. Поэтому оптимальное смещение выбирают вблизи точки перегиба ($\Delta v_0/2 \leq v_{\text{прг}}$), когда потерями в точности по сравнению с оптимальным дискриминатором практически можно пренебречь.

Чтобы реализовать характеристику (8.18) и получить выходной сигнал $v - v^0 \approx k_2 U_{\Delta} / Z$, необходимо из входной смеси сформировать три огибающих корреляционного интеграла, отличающихся только опорным значением (v^0 , $v^0 + \Delta v_0/2$, $v^0 - \Delta v_0/2$): огибающая $Z(v - v^0)$ формируется в опорном канале, а огибающие $Z(v - v^0 - \Delta v_0/2)$ и $Z(v - v^0 + \Delta v_0/2)$ подаются на схему вычитания (компаратор) дискриминатора (рис. 8.8, а). Выходное колебание опорного канала $Z(v - v^0)$ используется для нормировки с помощью схемы АРУ и для создания опорного напряжения синхронного детектора (СД) дискриминатора путем отслеживания его частоты и фазы с помощью схемы ЧФАП, а также подается на обнаружитель СПЗ. Если же корреляционный интеграл выдается в виде непрерывной функции времени (в измерителе дальности),

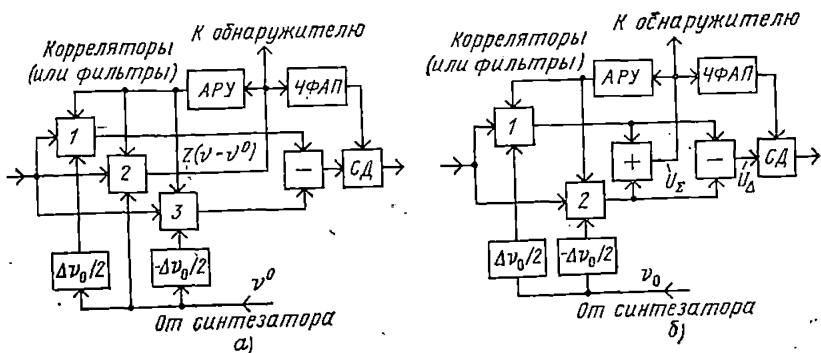


Рис. 8.8. Схемы амплитудных подоптимальных дискриминаторов

то все три значения можно получить в виде отсчетов, взятых в три момента времени, на выходе общего тракта приема.

Рассмотренный выше амплитудный дискриминатор называют разностным, он отличается от оптимального только в измерительном канале. Существуют еще более простые амплитудные дискриминаторы, называемые суммарно-разностными, отличающиеся от оптимального и в опорном канале, где вместо огибающей $Z(v-v^0)$ используют сумму $U_{\Sigma} = Z(v-v^0-\Delta v_0/2) + Z(v-v^0+\Delta v_0/2)$ тех же сигналов, из которых формируется разностный сигнал (рис. 8.8,б). Характеристика суммарно-разностного амплитудного дискриминатора

$$D(v-v^0) = k_3 \frac{Z_s(v-v^0-\Delta v_0/2) - Z_s(v-v^0+\Delta v_0/2)}{Z_s(v-v^0-\Delta v_0/2) + Z_s(v-v^0+\Delta v_0/2)}, \quad (8.19)$$

а его выходной сигнал при малых рассогласованиях пропорционален расстройке: $k_3 U_{\Delta} / U_{\Sigma} \approx v-v^0$. Отношение сигнал-шум в суммарном канале, являющемся в этом случае опорным каналом приема, меньше, чем в опорном канале разностного дискриминатора, и убывает с увеличением разнosa Δv_0 . Это влечет за собой снижение дальности действия системы. Поэтому разнос в суммарно-разностных дискриминаторах выбирают меньше оптимального по точности — как компромисс между снижением дальности действия и точности измерений.

Фазовые линейные дискриминаторы используют там, где возможно разделение обобщенного энергетического спектра $|S(\mu)|^2$ на две одинаковые части вида $|S_1(\mu)|^2$, которые симметричны относительно своих фазовых центров $\Delta\mu_0/2$ и $-\Delta\mu_0/2$, разнесенных на интервал $\Delta\mu_0 \geq \Delta\mu/2$, где $\Delta\mu$ — общая ширина спектра. Таким образом,

$$|S(\mu)|^2 = |S_1(\mu-\Delta\mu_0/2)|^2 + |S_1(\mu+\Delta\mu_0/2)|^2. \quad (8.20)$$

Корреляционный интеграл является обратным преобразованием Фурье от энергетического спектра: $Z_s(v-v^0) = \mathcal{F}^{-1}\{|S(\mu)|^2\}$ и $Z_{s1}(v-v^0) = \mathcal{F}^{-1}\{|S_1(\mu)|^2\}$. Так как спектры $|S(\mu)|^2$ и $|S_1(\mu)|^2$ — симметричные и вещественные функции $Z_s(v-v^0)$ и $Z_{s1}(v-v^0)$ также симметричные и вещественные, т. е. представляют огибающую корреляционных интегралов в отсутствие шумов. Из равенства (8.20) с учетом аддитивности преобразований Фурье и теоремы смещения получим

$$\begin{aligned} Z_s(v-v^0) &= \mathcal{F}^{-1}\{|S_1(\mu-\Delta\mu_0/2)|^2\} + \mathcal{F}^{-1}\{|S_1(\mu+\Delta\mu_0/2)|^2\} = \\ &= Z_{s1}(v-v^0) \exp[j\pi\Delta\mu_0(v-v^0)] + Z_{s1}(v-v^0) \exp[-j\pi\Delta\mu_0 \times \\ &\times (v-v^0)] = Z_{s1}(v-v^0) + Z_{s1}^*(v-v^0) = U_{s\Sigma}, \end{aligned} \quad (8.21)$$

где $Z_{s1}(v-v^0) = Z_{s1}(v-v^0) \exp[j\pi\Delta\mu_0(v-v^0)]$; $Z_{s1}^*(v-v^0) = Z_{s1}(v-v^0) \exp[-j\pi\Delta\mu_0(v-v^0)]$ — комплексно-сопряженные слагаемые, полученные раздельной оптимальной фильтрацией двух сигналов, которые соответствуют двум половинам обобщенного спектра в отсутствие шума; суммарный сигнал $U_{s\Sigma}$ вещественный и равен огибающей корреляционного интеграла $Z_s(v-v^0)$.

Первая производная огибающей корреляционного интеграла (8.21) по измеряемому параметру содержит две составляющие — амплитудную (действительная) и фазовую (мнимая):

$$Z'_s(v-v^0) = Z'_{s1}(v-v^0) \exp[j\pi\Delta\mu_0(v-v^0)] - Z'_{s1}(v-v^0) \exp[-j\pi\Delta\mu_0(v-v^0)] + j\pi\Delta\mu_0 \{Z_{s1}(v-v^0) \exp[j\pi\Delta\mu_0(v-v^0)] - Z_{s1}(v-v^0) \exp[-j\pi\Delta\mu_0(v-v^0)]\} = 2Z'_{s1}(v-v^0) \cos \pi\Delta\mu_0(v-v^0) + j\pi\Delta\mu_0 \{Z_{s1}(v-v^0) - Z^*_{s1}(v-v^0)\}. \quad (8.22)$$

Фазовая составляющая содержит разностный сигнал

$$U_s = Z_{s1}(v-v^0) - Z^*_{s1}(v-v^0), \quad (8.23)$$

который сдвинут относительно суммарного на $\pi/2$ (суммарный и разностный сигнал образуют диагонали ромба, образованного равными по модулю векторами $Z_{s1}(v-v^0)$ и $Z^*_{s2}(v-v^0)$, составляющими между собой угол $\Delta\varphi = 2\pi\Delta\mu_0(v-v^0)$, т. е. являются геометрическими суммой и разностью).

Подставляя формулы (8.21) и (8.22) в (8.16) и пренебрегая амплитудной составляющей производной огибающей корреляционного интеграла (8.22), получаем характеристику фазового суммарно-разностного подоптимального дискриминатора:

$$D(v-v^0) = -j k_1 \pi \Delta\mu_0 \times \\ \times \frac{\exp[j\pi\Delta\mu_0(v-v^0)] - \exp[-j\pi\Delta\mu_0(v-v^0)]}{\exp[j\pi\Delta\mu_0(v-v^0)] + \exp[-j\pi\Delta\mu_0(v-v^0)]} = k_4 \operatorname{tg} \pi\Delta\mu_0(v-v^0), \quad (8.24)$$

где было учтено, что сумма экспонент дает удвоенный косинус, а разность — удвоенный синус, умноженный на мнимую единицу, и $k_4 = \pi\Delta\mu_0 k_1$. Из формулы (8.24), в частности, следует, что при строгой симметрии одинаковых полуспектров амплитудные соотношения не играют никакой роли. Реальная характеристика дискриминатора в отличие от идеальной (8.24) при больших рассогласованиях имеет спадающие участки, так как спадание до нуля суммарного сигнала (входящего в знаменатель) невозможно скомпенсировать конечным усилением разностного сигнала.

Структура фазового суммарно-разностного дискриминатора определяется (при малых рассогласованиях) зависимостью

$$v-v^0 \approx -j k_1 \Delta\mu_0 (U_\Delta/U_\Sigma) = k_4 (U_\Delta/U_\Sigma) \exp(-j\pi/2),$$

где множитель $-j = \exp(-j\pi/2)$ указывает на необходимость сдвига фазы разностного сигнала относительно суммарного на $-\pi/2$ перед синхронным детектированием (операция фазирования). Таким образом, суммарно-разностный фазовый дискриминатор содержит отдельные каналы оптимальной фильтрации сигналов двух полуспектров, суммарно-разностный мост, фазовращатель и синхронный детектор, а также схемы АРУ и ЧФАП (рис. 8.9). При этом амплитудная составляющая производной (8.22), ортогональная фазовой составляющей, подавляется с помощью синхронного детектора вместе с ортогональной составляющей шу-

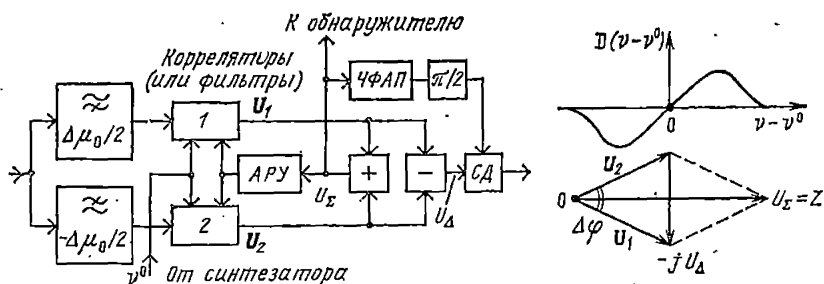


Рис. 8.9. Схема фазового подоптимального суммарно-разностного дискриминатора

ма. Максимально возможный проигрыш в точности у фазового дискриминатора составляет 13%.

Необходимо отметить важную особенность линейных дискриминаторов (как амплитудных, так и фазовых), что в силу линейности схем сумму и разность сигналов можно образовать еще на входе приемно-усилительных трактов суммарного и разностного каналов. Вследствие этого резко уменьшается влияние неидентичности амплитудных и фазовых характеристик обоих каналов на выходной сигнал дискриминатора.

Нелинейные подоптимальные дискриминаторы. В основу сравнения сигналов в нелинейных (мультипликативных) дискриминаторах положено отношение сравниваемых сигналов, которое удовлетворяет условию нормировки. Однако условие нечетной симметрии не удовлетворяется и при равенстве сигналов характеристика проходит через единицу, а не через нуль. Поэтому берут логарифм отношения, удовлетворяющий условию нечетной симметрии. Нелинейные подоптимальные дискриминаторы также подразделяют на амплитудные и фазовые. В амплитудных сравниваются сигналы, разнесенные по параметру ν , в фазовых — по обобщенной частоте μ .

В **амплитудных дискриминаторах** сравниваемые сигналы $U_1 = Z(\nu - \nu^0 - \Delta\nu_0/2)$ и $U_2 = Z(\nu - \nu^0 + \Delta\nu_0/2)$ можно представить в виде линейных функций малых отклонений $\nu - \nu^0$ от параметра настройки ν^0 :

$$U_{1,2} = Z\left(\frac{\Delta\nu_0}{2}\right) \pm Z'\left(\frac{\Delta\nu_0}{2}\right)(\nu - \nu^0) = \\ = Z\left(\frac{\Delta\nu_0}{2}\right) \left[1 \pm \frac{Z'(\Delta\nu_0/2)}{Z(\Delta\nu_0/2)}(\nu - \nu^0) \right] \approx Z\left(\frac{\Delta\nu_0}{2}\right) \exp[\pm k_5(\nu - \nu^0)],$$

где учтено, что $1 \pm a \approx \exp(\pm a)$ при $a \ll 1$, а $k_5 = 2Z'_s(\Delta\nu_0/2)/Z_s(\Delta\nu_0/2)$. Логарифм отношения

$$\ln(U_1/U_2) = \ln U_1 - \ln U_2 \approx k_5(\nu - \nu^0) \quad (8.25)$$

определяет структуру нелинейного амплитудного дискриминатора (рис. 8.10, а). Сравнимые сигналы проходят через оптимальные

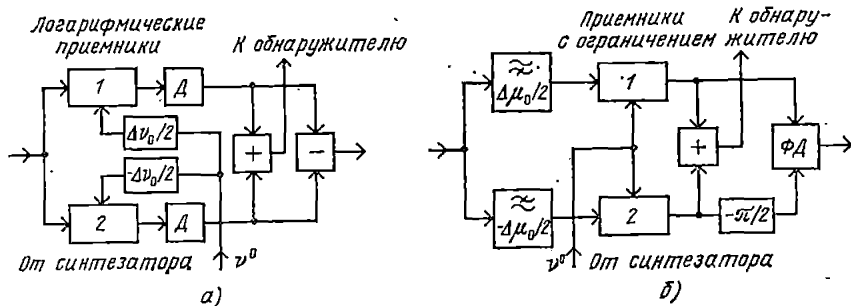


Рис. 8.10. Схемы нелинейных дискриминаторов

приемные каналы с логарифмической амплитудной характеристикой, детектируются и поступают на схему вычитания. Фактически амплитудная характеристика для очень малых сигналов делается линейной, в противном случае она должна быть отрицательной, что нереально. Для нормировки и расширения динамического диапазона приемных каналов ставить схему АРУ нет необходимости; ее роль выполняет сама логарифмическая характеристика. Опорный сигнал для амплитудного детектора не требуется, однако для работы обнаружителя СПЗ формируется также суммарный сигнал двух каналов.

Характеристика нелинейного (логарифмического) амплитудного дискриминатора

$$D(v-v^0) = \ln(U_{1s}/U_{2s}) = \ln Z_s(v-v^0 - \Delta v_0/2) - \ln Z_s(v-v^0 + \Delta v_0/2) \quad (8.26)$$

линейна при малых рассогласованиях, как это следует из выражения (8.25).

В фазовых нелинейных дискриминаторах, как и в линейных, сравниваются сигналы двух полуспектров, имеющие одинаковую амплитуду, но различающиеся по фазе на $\Delta\varphi = 2\pi\Delta\mu_0(v-v^0)$ из-за разности $\Delta\mu_0$ фазовых центров двух симметричных полуспектров: $U_{1,2} = Z_1(v-v^0) \exp(\pm j\Delta\varphi/2)$. После логарифмирования отношения этих двух сигналов и умножения на $-j$ для выполнения условия фазирования находим зависимость выходного сигнала дискриминатора от рассогласования $-j \ln(U_1/U_2) = \Delta\varphi = 2\pi\Delta\mu_0(v-v^0)$, которая определяет характеристику нелинейного фазового дискриминатора

$$D(v-v^0) = -j \ln(U_{1s}/U_{2s}) = 2\pi\Delta\mu_0(v-v^0) \quad (8.27)$$

и его структуру. Таким образом, измерение рассогласования сводится к определению разности фаз двух сравниваемых сигналов, причем характеристика (8.27) линейна на всем интервале рассогласования. Практически разность фаз определяют с помощью фазового детектора с ограничителями амплитуды на входе. Тогда характеристика (8.27) принимает вид

$$D(v-v^0) = \sin 2\pi\Delta\mu_0(v-v^0). \quad (8.28)$$

Если спектр сплошной, эта характеристика однозначна; если же полуспектры сосредоточены лишь около своих фазовых центров $\Delta\mu_0/2$ и $-\Delta\mu_0/2$, то многозначна.

При указанных условиях структурная схема нелинейного фазового дискриминатора (рис. 8.10,б) отличается от линейного тем, что на выходе двух высокочастотных каналов включают амплитудные ограничители, сигналы которых подаются на фазовый детектор с поворотом фазы $\pi/2$ на одном из входов. Ограничители выполняют операцию нормировки, поэтому надобность в схеме АРУ также отпадает. Отсутствует необходимость и в опорном сигнале, однако суммарный сигнал формируется для работы обнаружителя. Фазовые нелинейные дискриминаторы в отличие от амплитудных широко применяют при измерении угловых координат ввиду их относительной простоты и высокой точности измерений при большом разnose антенн (обобщенный параметр $\Delta\mu_0$).

В нелинейных дискриминаторах в отличие от линейных нельзя размещать приемные каналы после схемы сравнения. Поэтому все амплитудные и фазовые различия и нестабильности непосредственно влияют на выходной сигнал дискриминатора. В частности, различие коэффициентов усиления двух каналов в амплитудных дискриминаторах или набегов фазы в фазовых вызывает паразитное смещение нуля их характеристики. В результате точность измерений ухудшается. Поэтому нелинейные дискриминаторы почти не используются в следящих измерителях, к точности измерения которых предъявляются наиболее высокие требования. Нелинейные дискриминаторы обеспечивают высокую точность лишь при достаточно большом входном отношении сигнал-шум (в каждом импульсе), так как при малых отношениях сигнал-шум в них происходит подавление сигнала шумом (повышается относительный уровень шума и искажается сам сигнал). Ввиду этого чувствительность и дальность действия систем с нелинейными дискриминаторами при одинаковых энергетических затратах ниже, чем с линейными. Однако при отсутствии энергетических ограничений нелинейные измерители не уступают по точности линейным, если приняты специальные меры по устранению неидентичности амплитудных и фазовых характеристик приемных каналов.

Использование того или иного типа дискриминаторов, оптимальных и подоптимальных, зависит от вида измеряемого параметра. Поэтому более подробное описание их структуры возможно лишь при рассмотрении конкретных видов следящих измерителей.

При независимых каналах измерений нескольких параметров каждый дискриминатор можно рассматривать отдельно, но учитывать, что все они должны быть одновременно настроены на параметры общего входного сигнала. Если же каналы зависимы, необходимо рассматривать многомерный дискриминатор в целом, с несколькими выходами.

8.5. СЛЕДЯЩИЕ ИЗМЕРИТЕЛИ ДАЛЬНОСТИ

Следящий измеритель дальности (СИД) предназначен для формирования оценки случайного процесса — текущей дальности объекта $R(t)$ по данным измерений времени запаздывания сигнала $\tau(t) = 2R(t)/c$ в каждом периоде посылок. СИД включает кольцо слежения, состоящее из временного дискриминатора, формирователя оценки и синтезатора задержки, а также схему поиска и захвата (СПЗ) и другие элементы РЛС, в состав которой он входит.

Работу следящего измерителя дальности рассмотрим во взаимодействии с основными элементами приемопередающего тракта (рис. 8.11). Под воздействием запросных импульсов (ЗИ), поступающих от синхронизатора станции, передатчик через антенну излучает зондирующие колебания. Эти колебания после отражения или переизлучения их объектом образуют ответный сигнал, который воспринимается антенной и через антенный переключатель (АП), обеспечивающий работу приемника и передатчика на общую антенну, поступает на вход приемника РЛС. Ответные импульсы (ОИ) очень слабы и искажены шумами, поэтому их временное положение относительно запросных импульсов испытывает случайные колебания, что приводит к большим погрешностям измерения дальности. Задачей следящего измерителя дальности является формирование следящих импульсов (СИ), временное положение которых отображает лишь плавное, закономерное изменение дальности и почти не изменяется под воздействием хаотических шумов. Тем самым СИД выполняет функции сглаживающего фильтра для функции $\tau(t) = 2R(t)/c$. Отсчет дальности по положению следящих импульсов относительно запросных обеспечивает высокую точность измерений.

Ответный импульс с выхода приемника непосредственно или через селектор дальности поступает на вход временного дискриминатора, опорным сигналом которого является следящий им-

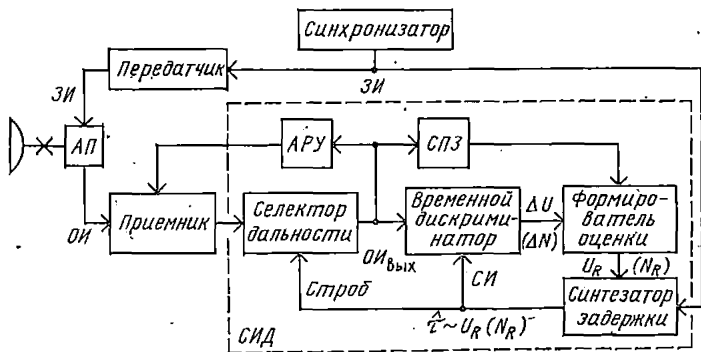


Рис. 8.11. Следящий измеритель дальности в составе радиосистемы, работающий по отраженному или ответному сигналу

пульс. Временной дискриминатор выдает на вход формирователя оценки сигнал рассогласования в аналоговом (ΔU) или цифровом (ΔN) виде. Для нормировки сигнала рассогласования с помощью схемы АРУ только по сигналу отслеживаемого объекта селектор дальности открывается следящим импульсом, обычно несколько расширенным (стробом). Под воздействием сигнала рассогласования в режиме слежения формирователь оценки вырабатывает оценку процесса $\hat{t}(t)$ в виде напряжения U_R или кода N_R , которые управляют синтезатором задержки. Синтезатор вырабатывает следящий импульс, задерживая запросный импульс синхронизатора на время $\hat{t}$, пропорциональное U_R (или N_R). Следящий импульс через дискриминатор замыкает кольцо слежения. Благодаря действию обратной связи рассогласование между ответным и следящим импульсами в установившемся режиме в среднем равно нулю. Оценка снимается в виде напряжения U_R (или кода N_R) на выходе формирователя оценки или же непосредственно в виде временного интервала $\hat{t}$ между запросным и следящим импульсами.

Режиму слежения (измерения) предшествует режим поиска (обнаружения), в котором кольцо слежения разомкнуто, так как ответный и следящий импульсы не перекрываются. В режиме поиска программное устройство схемы поиска СПЗ воздействует на формирователь оценки, который линейно изменяет U_R (или N_R), тем самым медленно смещая по времени t^0 следящий импульс на выходе синтезатора задержки вместе со стробом селектора дальности (перестройка по дальности).

Когда в процессе медленной перестройки по дальности ответные импульсы попадают в строб, они через селектор дальности проходят на обнаружитель СПЗ, представляющий собой накопитель (в простейшем случае конденсатор). Так как селектор дальности и накопитель образуют гребенчатый фильтр, ответные импульсы интегрируются. Когда накопленный сигнал превысит порог срабатывания реле захвата, поиск прекращается. При этом рассогласование между ответным и следящим импульсами мало, поэтому в дискриминаторе выделяется сигнал рассогласования и кольцо слежения замыкается. Происходит захват сигнала, и СИД переходит в режим слежения — основной режим работы.

При описании процессов в СИД умышленно опущены вопросы, связанные с формированием опорного напряжения синхронного детектора с помощью схемы ЧФАП, которые будут рассмотрены при описании следящих измерителей скорости. Эти вопросы приобретают особую важность при использовании сложных сигналов, поэтому мы и ограничились простыми импульсными сигналами.

Поскольку структура формирователя оценки почти не зависит от конкретного вида следящего измерителя, рассмотрим только особенности временного дискриминатора и структуру синтезатора задержки в аналоговом и цифровом вариантах.

Временные дискриминаторы выполняются только в виде корреляторов, так как фильтры инвариантны к времени прихода сигнала. В соответствии с обобщенными характеристиками дискриминаторов можно построить временные дискриминаторы различных типов. Опишем лишь некоторые из них.

В принципе наиболее просто реализовать разностный временной дискриминатор, характеристика которого в соответствии с общей формулой (8.16) имеет вид

$$D(\tau - \tau^0) = k_2 \frac{Z_s(\tau - \tau^0 - \Delta\tau_0/2) - Z_s(\tau - \tau^0 + \Delta\tau_0/2)}{Z_s(\tau - \tau^0)}. \quad (8.29)$$

Поскольку в данном случае измеряемым параметром является время, выходной импульс оптимального приемника ОИ_{вых} повторяет корреляционный интеграл $Z(t - \tau)$ во времени. Следовательно, для формирования сигнала рассогласования временного дискриминатора с характеристикой (8.29) достаточно получить отсчеты выходного импульса $Z(t - \tau)$ в три момента времени, отсчитываемых от начала текущего периода: $\tau^0 - \Delta\tau_0/2$, τ^0 и $\tau^0 + \Delta\tau_0/2$, а затем образовать разность двух крайних отсчетов и разделить ее на средний отсчет. Эти отсчеты можно получить в аналого-цифровом преобразователе (АЦП) напряжение — код с помощью трех коротких (наносекундных) импульсов, вырабатываемых синтезатором задержки. Операции вычитания и деления выполняются цифровыми схемами. В рассмотренном цифровом временном дискриминаторе операция нормировки выполняется идеально (дополнительно к схеме АРУ), а инструментальная погрешность определяется только ошибками квантования и округления при выполнении операции деления. Недостаток указанного дискриминатора — высокие требования к быстродействию его элементов.

Наиболее широко применяют аналоговые временные дискриминаторы, приближенно реализующие характеристику (8.29), в которых для отсчета используют не короткие импульсы, а сравнительно широкие импульсы вилки, причем нормировка производится только схемой АРУ приемника, т. е. неидеально. Процессы в дискриминаторе (и приемопередатчике) в каждом периоде посылок протекают следующим образом (рис. 8.12).

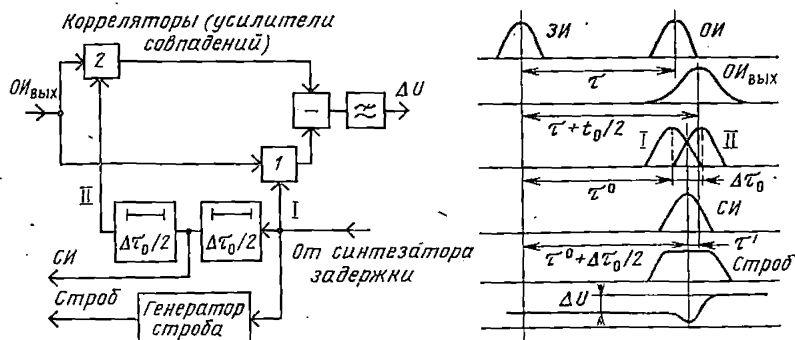


Рис. 8.12. Схема временного дискриминатора и процессы в нем

На вход приемника поступает ответный импульс (ОИ) спустя время τ относительно запросного импульса (ЗИ). Ответный импульс на выходе оптимального приемника (фильтра) ОИ_{вых}, повторяющий по форме корреляционный интеграл, как было показано ранее, запаздывает относительно входного на время $t_0/2$ (относительно его середины) или t_0 (относительно его начала). Этот импульс поступает на два коррелятора; опорным сигналом первого является импульс вилки I, поступающий с выхода синтезатора задержки с запаздыванием τ^0 относительно ЗИ, а у второго импульс вилки II, дополнительно задержанный на Δt_0 . Середине импульсов вилки соответствует положение $\tau^0 + \Delta t_0/2$ следящего импульса (СИ). Одновременно с импульсами вилки формируется широкий строб для пропускания выходных сигналов приемника на схему АРУ и обнаружитель СПЗ (селекция по дальности). Напряжения с выходов корреляторов подаются на схему вычитания и ФНЧ, реально представляющих собой цепь заряда-разряда конденсатора. Степень заряда или разряда определяется площадью перекрытия ОИ_{вых} с импульсами вилки: часть ОИ_{вых}, совпадающая с импульсом I, идет на разряд конденсатора, а совпадающая с импульсом II — на заряд.

Таким образом, напряжение ΔU на выходе дискриминатора определяется не разностью отсчетов сигнала в разнесенных точках, как это требуется согласно формуле (8.29), а разностью площадей перекрытия, что несколько изменяет характеристику дискриминатора. При малых рассогласованиях $\tau' = (\tau + t_0/2) - (\tau^0 + t_0/2) = \tau - \tau^0$ и $\Delta t_0 = t_0$ между ОИ_{вых} и серединой импульсов вилки характеристика линейна, а знак зависит от стороны рассогласования (в данном случае $\tau^0 < \tau$ и $\Delta U > 0$). В установившемся режиме слежения среднее значение рассогласования $\tau' = \tau - \tau^0$ равно нулю и текущее положение следящего импульса $\tau^0(t)$ становится оценкой времени запаздывания $\hat{\tau}(t) = M[\tau(t)] = \tau^0(t)$, т. е. дальности.

Для формирования следящих импульсов (точнее, импульсов вилки I), задержанных относительно запросных на время τ^0 , пропорциональное показаниям U_R (или N_R) формирователя оценки, используются аналоговые (или цифровые) синтезаторы задержки. Простейшие импульсные синтезаторы задержки работают следующим образом (рис. 8.13).

В аналоговом импульсном синтезаторе задержки с поступлением очередного запросного импульса вырабатывается линейно нарастающее напряжение $U = kt$ с помощью специального генератора. Благодаря диоду, закрытому напряжением U_R , которое поступает с выхода аналогового формирователя, ток через нагрузку генератора вначале не протекает. Нагрузкой является трансформатор входной цепи заторможенного генератора импульсов (например, блокинг-генератора). Как только линейно нарастающее напряжение достигнет значения U_R , диод открывается, пропуская ток через трансформатор. Генератор импульсов срабатывает и выдает СИ, задержанный относительно ЗИ на время $\tau^0 = U_R/k$, что и тре-

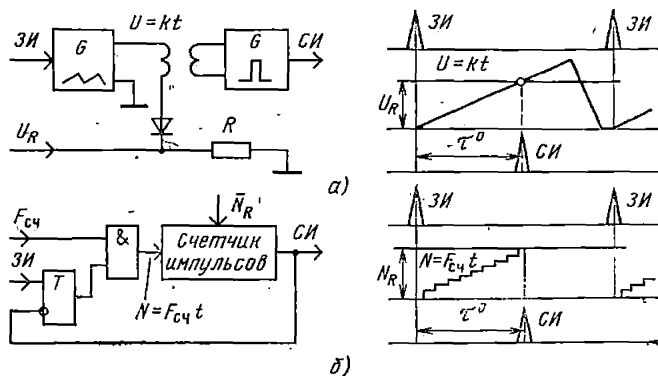


Рис. 8.13. Схемы синтезаторов задержки, управляемые напряжением (а) и кодом (б)

бовалось. К следующему периоду посылок линейно нарастающее напряжение спадает до нуля, и все повторяется в приведенной последовательности.

В цифровом варианте импульсного синтезатора генератор линейно нарастающего напряжения заменен счетчиком импульсов. Показания счетчика $N = F_{сч}t$ с приходом ЗИ линейно нарастают благодаря поступлению на его вход счетных импульсов частоты $F_{сч}$ через вентиль, открываемый управляющим триггером. Перед началом работы счетчика в нем устанавливается в дополнительном коде число $\bar{N}_R$, поступающее от цифрового формирователя оценки. Так как добавление $\bar{N}_R$ равносильно его вычитанию, с поступлением на вход $N = N_R$ импульсов счетчик устанавливается в нулевое состояние, а на выходе его старшего разряда появляется импульс переполнения. Этот импульс и является СИ, так как отстоит от ЗИ на время $\tau^0 = N_R/F_{сч}$. Одновременно он поступает на инверсный вход управляющего триггера, прекращая поступление счетных импульсов на вход счетчика, и через специальную схему (не показанную на рисунке) вновь вводит число $\bar{N}_R$. Синтезатор подготовлен к новому циклу.

Цифровой синтезатор задержки сохраняет линейность и высокую точность при сколь угодно большой задержке (дальности отслеживаемого объекта). Точность его определяется стабильностью и значением частоты $F_{сч}$, задаваемой высокостабильным генератором, параметры которого выбираются при проектировании. Точность же аналогового импульсного синтезатора очень низкая, особенно при больших задержках, так как вырабатываемое в них напряжение лишь приближенно можно считать линейным (это — начальный участок экспоненты), а колебания напряжения вызывают значительный разброс времени срабатывания генератора СИ. Такие синтезаторы применяют лишь при измерении сравнительно небольших расстояний — не более нескольких десятков километров.

В следящих измерителях больших дальностей используют более сложные синтезаторы задержки, в которых следящие импульсы, как и запросные, формируются из синусоидальных колебаний. При этом следящие импульсы формируются из колебаний частоты ω_0 , которые сдвинуты по фазе на φ° , что соответствует сдвигу сформированных следящих импульсов относительно запросных на $\tau^\circ = \varphi^\circ / \omega$. Сдвиг фазы выполняется с помощью фазовращателя, механически перестраиваемого электродвигателем, либо управляемого генератора синусоидальных колебаний. В первом случае скорость вращения двигателя, а во втором частота колебаний пропорциональны приложенному к ним напряжению. Поэтому применительно к фазе φ° или параметру τ° выполняется операция интегрирования, т. е. эти каскады выполняют также роль последнего интегратора в формирователе оценки СИД. Для обеспечения высокой точности эти синтезаторы работают на частоте, во много раз превышающей частоту повторения импульсов. Однозначность отсчета обеспечивается делением этой частоты до частоты повторения, для чего кроме делителей требуются схемы привязки, исключаящие паразитные сдвиги в делителях. В результате все устройство значительно усложняется.

8.6. СЛЕДЯЩИЕ ИЗМЕРИТЕЛИ СКОРОСТИ

Следящие измерители скорости (СИС) предназначены для отслеживания (оценки), с точностью до некоторого постоянного значения, текущей частоты принимаемого сигнала, которая содержит доплеровский сдвиг F , пропорциональный радиальной скорости. Частота вырабатываемых СИС колебаний повторяет сравнительно медленные, закономерные изменения частоты входного сигнала, а случайные флуктуации, вызываемые шумами, сглаживаются. Таким образом, следящий измеритель скорости эквивалентен узкополосному частотному фильтру, медленно перестраиваемому в такт с изменением доплеровского сдвига частоты входного сигнала $F(t)$. Структура СИС определяется двумя его особенностями.

1. СИС используется одновременно для поддержания постоянной частоты (и фазы) сигнала в УПЧ основного канала приема измерительной системы. Благодаря этому оптимальная фильтрация сигнала производится на постоянной промежуточной частоте, что особенно важно при использовании сигналов сложной формы, для которых трудно создать перестраиваемый согласованный фильтр. Ввиду этого частота вырабатываемого СИС колебания $f_\Gamma(t) = f(t) \pm f_0$ отличается от оценки $\hat{f}(t)$ текущей частоты входного сигнала на постоянную величину f_0 , а перестройка приемника по частоте производится гетеродинированием принимаемых сигналов. Частота гетеродинного напряжения f_Γ изменяется в такт с частотой входного сигнала с помощью схемы автоподстройки, что и обеспечивает постоянство промежуточной частоты f_0 .

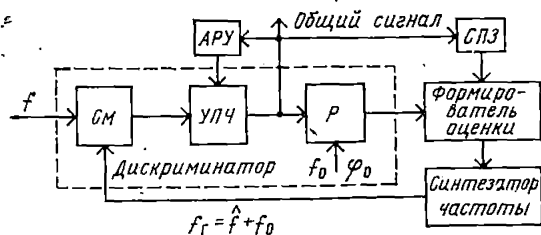


Рис. 8.14 Схема следящего измерителя радиальной скорости

Отмеченная особенность СИС отражается в его структурной схеме (рис. 8.14). Используется разнесенный дискриминатор в виде перестраиваемого смесителя (СМ) с УПЧ и неперестраиваемого различителя (Р), настроенного на постоянное опорное значение промежуточной частоты f_0 . В УПЧ с постоянной настройкой производится основное усиление и фильтрация сигнала (а также нормировка с помощью схемы АРУ, общей для всех следящих измерителей системы). Схема поиска и захвата, также входящая в общий канал, дополнительно к поиску по дальности производит поиск по частоте путем перестройки синтезатора частоты через формирователь оценки.

При описании функциональной схемы СИС удобнее пользоваться не комплексной огибающей, а реальным высокочастотным сигналом. Для простоты описания процессов в схеме предположим, что принимаемый сигнал, как и вырабатываемое синтезатором частоты гетеродинное напряжение, являются гармоническими колебаниями

$$u_c(t) = U_c \cos \varphi_c(t), \quad u_r(t) = U_r \cos \varphi_r(t),$$

где $\varphi_c(t) = 2\pi \int_0^t f(t') dt' + \varphi_{0c}$, $\varphi_r(t) = 2\pi \int_0^t f_r(t') dt' + \varphi_{0r}$ — текущие фазы колебаний на частоте входного сигнала $f(t)$ и гетеродина $f_r(t)$ соответственно. На выходе смесителя при $f_r > f$ образуются колебания суммарной и разностной частот

$$u_c u_r = (k U_c U_r / 2) \{ \cos [\varphi_r(t) + \varphi_c(t)] + \cos (\varphi_r(t) - \varphi_c(t)) \},$$

из которых в УПЧ получаем сигнал промежуточной (разностной) частоты

$$u_p(t) = U_p \cos \varphi_p(t),$$

где $U_p = k U_c U_r / 2$ — амплитуда разностного сигнала; $\varphi_p(t) = 2\pi \int_0^t f_p(t') dt' + \varphi_{0p}$ — его текущая фаза; $f_p = f_r - f$ — разностная (промежуточная) частота; $\varphi_{0p} = \varphi_{0r} - \varphi_{0c}$ — начальная фаза разностного сигнала.

2. В СИС используют два типа различителей: частотный и фазовый. На выходе частотного или фазового различителя соответственно имеем

$$\Delta \bar{U}, \Delta N = \begin{cases} k(f_p - f_0), \\ k(\varphi_p - \varphi_0), \end{cases}$$

где f_0 — частота настройки частотного различителя; φ_0 — фаза опорного сигнала фазового различителя. В соответствии с типом различителя СИС бывают с частотной (ЧАП) или фазовой (ФАП) автоподстройкой. Установившемся режиму $\Delta U = \Delta N = 0$ в следящих измерителях скорости с ЧАП соответствует равенство промежуточной и переходной частоты: $M[f_p] = f_0$ и $f_{\Gamma} = f + f_0$, а в следящих измерителях скорости с ФАП — равенство фаз $\varphi_p = \varphi_0$, которое также достигается при $M[f_p] = f_0$ и $f_{\Gamma} = f + f_0$, так как равенство текущих фаз двух колебаний возможно только при равенстве их частот. Поскольку $f(t) = f_{изл} + F(t)$, оценка текущей скорости — доплеровского сдвига частоты $F(t) = f(t) - f_{изл}$, где $f_{изл}$ — частота излучаемых колебаний.

Таким образом, следящие измерители скорости с ЧАП и с ФАП в определенном смысле эквивалентны. Основное различие их состоит в том, что рассогласование в фазовом различителе является результатом набега фазы, возникающего вследствие разности частот $f_p - f_0$, т. е. пропорционально интегралу от этой разности. Ввиду этого фазовый различитель применительно к измеряемому параметру — частоте — эквивалентен последовательному соединению частотного различителя и интегратора.

Учитывая обе особенности СИС, приходим к выводу, что по сравнению со структурной схемой следящего измерителя общего вида в СИС отсутствует элементарное разделение функций. Так, смеситель совместно с настроенным частотным различителем является перестраиваемым частотным дискриминатором, а смеситель совместно с настроенным фазовым различителем выполняет функции перестраиваемого частотного различителя и интегратора, структурно входящего в формирователь оценки.

Важной особенностью схемы ФАП является эквивалентность ее узкополосному фильтру, выделяющему несущую сигнала с точностью до фазы, что позволяет использовать выходной сигнал схемы ФАП в качестве опорного для синхронных детекторов.

Аналоговые частотные различители выполняются в виде фильтра с фиксированной переходной частотой (частотой настройки на нуль его характеристики), а фазовые являются корреляторами, на которые подается опорное напряжение заданной частоты и фазы.

Наиболее широко в СИС используется амплитудно-частотный дискриминатор, имеющий характеристику разностного амплитудного дискриминатора общего вида

$$D(f_p - f_0) = k_2 \frac{Z_s [f_p - (f_0 + \Delta f_0/2)] - Z_s [f_p - (f_0 - \Delta f_0/2)]}{Z_s (f_p - f_0)}, \quad (8.30)$$

где f_0 — частота настройки различителя; Δf_0 — разнос по частоте настройки двух резонансных фильтров (рис. 8.15, а).

В соответствии с формулой (8.30) амплитудно-частотный различитель включает два резонансных фильтра, симметрично построенных на $\Delta f_0/2$ относительно переходной частоты f_0 , на выходе которых стоят детекторы (Д), и схему вычитания. При этом предполагается, что нормировка производится в общем приемном

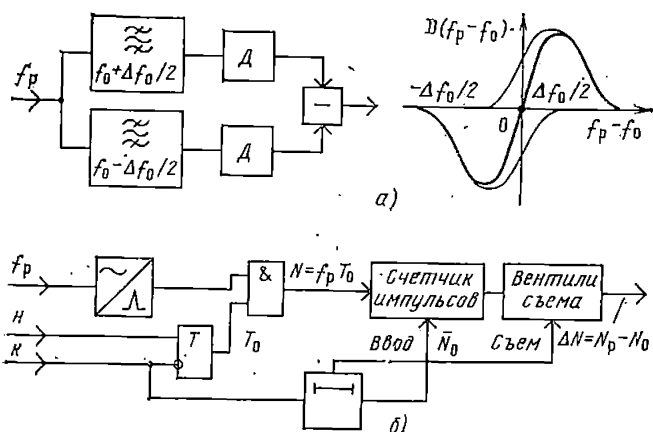


Рис. 8.15. Аналоговый (а) и цифровой (б) частотные различители

тракте, настроенном на частоту f_0 , с помощью схем АРУ. В принципе детектирование необходимо производить с помощью синхронных детекторов, однако ввиду узкополосности входных фильтров обеспечивается высокое отношение сигнал-шум и почти без потерь можно ограничиться амплитудными детекторами. Оптимальный разнос Δf_0 соответствует пересечению АЧХ двух фильтров в точках перегиба, где их крутизна максимальна.

В фазочастотном аналоговом различителе, представляющем собой фазовый детектор, сигнал подается на один вход через резонансный фильтр, настроенный на частоту f_0 , а на другой — через фазовращатель на $\pi/2$. По своим характеристикам он приближается к амплитудно-частотному, хотя его структура не вытекает из описанных выше схем дискриминаторов общего вида.

Цифровые частотные различители строятся на базе частотомера или периодомера. Цифровой частотный различитель — частотомер — подсчитывает число импульсов разностной частоты $N = f_p T_0$, поступающих на счетчик в течение мерного интервала времени T_0 , определяемого импульсами начала (н) и конца (к) этого интервала (рис. 8.15, б). Переходной частоте различителя f_0 соответствует число $N_0 = f_0 T_0$, которое перед каждым новым циклом счета устанавливается в счетчике в дополнительном коде. Число, выдаваемое на выход различителя, $\Delta N = N - N_0 = T_0 (f_p - f_0)$ в каждом цикле счета пропорционально разности текущей частоты f_p в УПЧ и переходной частоты различителя f_0 . В конце счета задержанный импульс конца интервала производит считывание показаний частотного различителя ΔN и установку исходного числа N_0 . Положительные ΔN выдаются в прямом коде, отрицательные — в дополнительном. Переходную частоту $f_0 = N_0 / T_0$ в цифровом частотном различителе в отличие от аналогового можно сделать высокостабильной путем формирования импульсов начала и конца интервала, определяющих мерный интервал T_0 , из колебаний стабилизированного эталона частоты.

Однако при ограниченном T_0 погрешность такого частотного различителя, определяемая дискретностью отсчета частоты $\delta f = 1/T_0$, может оказаться недопустимо большой, ввиду чего используют цифровые частотные различители — периодомеры. Периодомер значительно сложнее частотомера, но обладает повышенной точностью, особенно при измерении низких частот.

Фазовый аналоговый различитель состоит из фазового детектора (ФД), на который кроме входного сигнала частоты f_p и фазы φ_p через фазовращатель на $\pi/2$ подается опорное напряжение частоты f_0 и фазы φ_0 (рис. 8.16, а). Сдвиг фазы на $\pi/2$ в отличие от коррелятора канала обнаружения, который при отсутствии рассогласования между входным и опорным сигналами дает максимальный выходной сигнал, обеспечивает нулевой сигнал рассогласования. Фазовый аналоговый различитель проще частотного и обладает в принципе более высокой точностью. Стабильность его переходной частоты определяется стабильностью общего для всей станции стабилизированного задающего генератора, в то время как переходная частота аналогового частотного различителя определяется параметрами контуров, которые испытывают случайные изменения под воздействием внешних условий.

Однако измерителям с ФАП присущи специфические погрешности, вызываемые периодичностью характеристики фазового дискриминатора. При большом отношении сигнал-шум рабочей областью является линейная часть характеристики дискриминатора, которую с помощью специальных мер можно расширить до 2π . Однако при малых отношениях сигнал-шум случайные рассогласования выходят за пределы линейного участка характеристики и линейная аппроксимация следящего измерителя становится неприемлемой. Действительно, при случайных отклонениях фазы входного сигнала от фазы опорных колебаний, превышающих по абсолютному значению π , следящая система может перейти в новую точку устойчивого равновесия, отстоящую от прежней на 2π . Это эквивалентно соответствующему скачку фазы гетеродин-

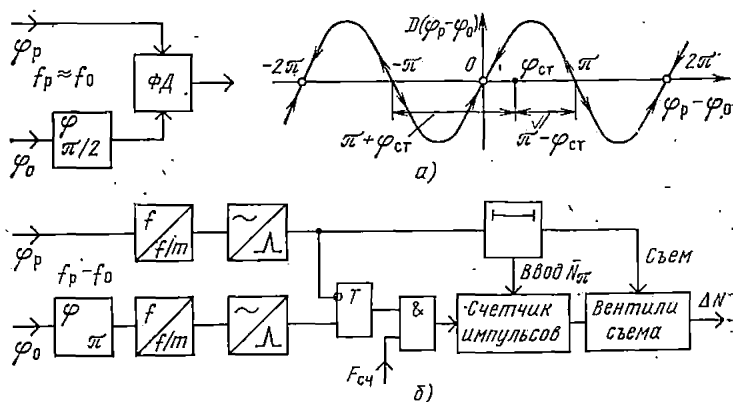


Рис. 8.16. Аналоговый (а) и цифровой (б) фазовые различители

ного напряжения, вырабатываемого синтезатором частоты. Частота его колебаний испытывает случайные изменения.

Если статистическая погрешность системы отсутствует, то среднее число выбросов фазы, выходящих за пределы $-\pi$ и $+\pi$, одинаково, т. е. положительные и отрицательные скачки фазы гетеродинного напряжения равновероятны. Поэтому средний уход частоты выходного напряжения следящего измерителя скорости с ФАП равен нулю независимо от отношения сигнал-шум. При наличии статической погрешности по фазе $\varphi_{ст}$ (см. рис. 8.16,а) перескок в новую точку устойчивого состояния системы при положительных и отрицательных отклонениях фазы будет происходить с неодинаковой частотой. При положительных отклонениях скачку фазы достаточно превысить величину $|\pi - \varphi_{ст}|$, при отрицательных отклонениях — величину $|\pi + \varphi_{ст}|$. Такая неравномерность перескоков фазы приводит к смещению средней частоты гетеродинного напряжения по отношению к частоте входного сигнала, в результате точность отслеживания (оценки) доплеровской частоты и радиальной скорости снижается. Погрешность зависит от статического смещения фазы и отношения сигнал-шум. При больших $\varphi_{ст}$ смещение частоты гетеродинных колебаний значительно даже при больших отношениях сигнал-шум.

Системы с ЧАП не имеют ошибок, связанных с перескоками фазы, но не обеспечивают высокой точности. Для того чтобы реализовать высокую точность измерений и исключить влияние перескоков фазы, в СИС применяют комбинацию ЧАП и ФАП — схему ЧФАП. В этой схеме кольцо ЧАП обеспечивает предварительную грубую подстройку частоты, благодаря которой большие отклонения фазы исключаются, и стоящее за схемой ЧАП кольцо ФАП производит точное отслеживание частоты в отсутствие перескоков фазы.

Цифровой фазовый различитель основан на преобразовании разности фаз между колебаниями разностной и опорной частот $\Delta\varphi = \varphi_p - \varphi_0$ во временной интервал с последующим заполнением его счетными импульсами, поступающими с частотой $F_{сч}$. Число ΔN , снимаемое со счетчика, пропорционально разности фаз, т. е. цифровой фазовый дискриминатор в отличие от аналогового имеет линейную характеристику на интервале однозначности $-\pi, \pi$.

Синусоидальные колебания разностной и опорной частот проходят через делители частоты, преобразуются в остроконечные импульсы, которые с помощью управляющего триггера и вентиля определяют время прохождения счетных импульсов на счетчик (рис. 8.16,б). Благодаря делителям частоты период сравниваемых колебаний возрастает в m раз и во столько же раз уменьшается ошибка, вносимая дискретностью отсчета фазы $\delta\varphi = 2\pi f_0 / m F_{сч}$. При этом временной интервал $m\Delta\varphi / 2\pi f_0$, соответствующий сдвигу фазы $\Delta\varphi$ на частоте f_0/m , представляется числом $\Delta N = (m\Delta\varphi / 2\pi f_0) F_{сч}$ на выходе счетчика.

Для того чтобы учесть знак разности фаз и обеспечить нормальную работу схемы вблизи $\Delta\varphi = 0$ (наиболее ответственный

участок характеристики дискриминатора), производят сдвиг фазы опорного колебания на $-\pi$, что соответствует упреждению опорного импульса на полпериода, т. е. на $m/2f_0$, при нулевом рассогласовании по фазе между сравниваемыми колебаниями. В течение этого интервала времени на счетчике, открываемом импульсами опорного канала, установится число $N_\pi = F_{сч}m/2f_0$, а не нуль, как требуется. Чтобы скомпенсировать его, в счетчике перед началом счета устанавливается число N_π (в дополнительном коде). Благодаря этому положительное рассогласование ΔN снимается с фазового различителя в прямом коде, а отрицательное — в дополнительном.

Линейность характеристики на всем интервале однозначности $-\pi, \pi$ является преимуществом цифрового фазового дискриминатора перед аналоговым. Однако неоднозначность и перескоки фазы при сильных возмущениях сохраняют свою силу.

Синтезаторы частоты также выполняются в аналоговом и цифровом вариантах. В качестве аналогового синтезатора частоты обычно используется автогенератор синусоидальных колебаний (гетеродин), управляемый по частоте напряжением с помощью реактивной лампы или полупроводникового диода с управляемой емкостью $p-n$ -перехода. Зависимость частоты гетеродина f_H от приложенного напряжения линейна лишь в небольшом интервале относительно исходной f_{H0} . В этих пределах синтезатор частоты можно считать безынерционным передаточным звеном с постоянным коэффициентом передачи k , т. е. частота колебаний

$$f_H = f_{H0} + kU_V,$$

где U_V — управляющее напряжение, которое поступает с выхода аналогового формирователя, пропорциональное радиальной скорости объекта, а kU_V — измеренное (отслеженное) значение доплеровской частоты.

Задачей цифровых синтезаторов частоты является формирование гетеродинных синусоидальных колебаний частоты

$$f_H = f_{H0} + N_V \delta f,$$

где N_V — управляющее число, пропорциональное скорости, снимаемое с выходного регистра цифрового формирователя оценки; δf — дискретность отсчета частоты. Существует несколько видов цифровых синтезаторов частоты. Однако наиболее точным из них является синтезатор частоты, который представляет собой обычную аналоговую схему ФАП с перестраиваемым делителем частоты в цепи обратной связи, управляемым выходным числом формирователя оценки N_V (рис. 8.17). Благодаря замкнутой цепи регулирования фазовый детектор — управляющий каскад — управляемый генератор (ФД — УК — УГ) через смеситель (См) и перестраиваемый делитель частоты в установившемся режиме частота колебаний на выходе перестраиваемого делителя поддерживается равной частоте $f_2 = M\delta f$ одного из двух выходных напряжений генератора эталонных частот. Следовательно, частота

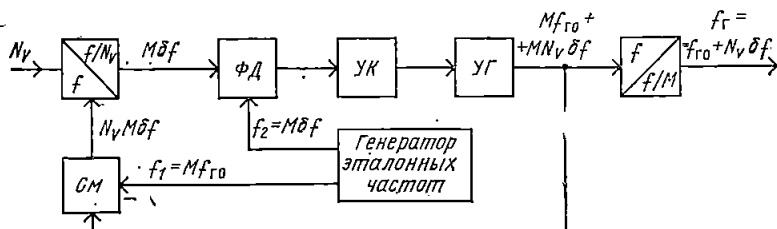


Рис. 8.17. Схема цифрового синтезатора частоты с делителем частоты в цепи обратной связи

колебаний на входе перестраиваемого делителя с коэффициентом деления $1:N_V$ равна $N_V M \delta f$, а на выходе управляемого генератора $M f_{Г0} + M N_V \delta f$, что и на первом входе смесителя, так как на его второй вход подается опорное напряжение частоты $f_1 = M f_{Г0}$ с другого выхода генератора эталонных частот. Делитель частоты с постоянным коэффициентом деления $1:M$ выдает гетеродинное колебание частоты $f_Г = f_{Г0} + N_V \delta f$, что и требуется.

Стабильность частоты выходных колебаний замкнутого синтезатора частоты определяется стабильностью генератора эталонных частот, который как более высокочастотная система стабилизирует частоту колебаний управляемого генератора. Таким образом, в цифровых синтезаторах частоты реализуется новое качество благодаря разделению функций перестройки и стабилизации частоты между различными элементами схемы: перестройка достигается выбором одной из фиксированных частот при смене числа N_V , а высокая стабильность сменных фиксированных частот обеспечивается благодаря автоподстройке под частоту непереключаемого эталонного генератора. В аналоговых системах противоречие между требованиями легкой управляемости и высокой стабильности частоты вырабатываемых колебаний принципиально неразрешимо: частота управляемого генератора одинаково легко изменяется под воздействием как управляющего напряжения, так и случайных шумов. Замкнутая аналоговая система в конечном счете сама выбирает любые возмущения, однако для этого она должна иметь достаточно широкую эффективную полосу, пропускающую шумовые изменения частоты и, следовательно, снижающую точность аналогового СИС. Цифровым СИС свойственны ошибки квантования частоты, которые, однако, проектировщиком могут быть сведены к допустимому значению.

Для снижения ошибок, вызываемых квантованием, шаг частоты δf берется достаточно малым — единицы или доли герц. Непосредственно на столь низких частотах фазовый детектор работает плохо: постоянная времени его фильтра, которая должна быть равна нескольким периодам сравниваемых колебаний, может достигать нескольких секунд, ввиду чего время установления процессов в кольце ФАП при смене числа N_V получается недопустимо большим. Поэтому частоту опорных колебаний фазового

детектора $f_2 = M\delta f$ берут во много раз больше шага частоты δf , во столько же раз завышают и частоту f_T по сравнению с f_{T0} , а на выходе синтезатора частоты ставят делитель частоты с фиксированным коэффициентом деления $1/M$.

Перестраиваемый делитель частоты, являющийся основным элементом синтезатора, обычно выполняется на счетчике-делителе, на вход которого поступают короткие импульсы, сформированные из синусоидальных колебаний частоты $MN_V\delta f$. Счетчик-делитель представляет собой обычный двоичный счетчик, в котором перед началом счета вводится управляющее число N_V (в дополнительном коде $\bar{N}_V$) с помощью специальной схемы ввода. С поступлением на счетчик N_V входных импульсов во всех разрядах счетчика устанавливается нулевое состояние $N_V + \bar{N}_V = 0$, а на выходе появится импульс переполнения. Этот импульс через схему ввода вновь запишет в счетчик число $\bar{N}_V$, и следующий импульс переполнения вновь появится после поступления N_V очередных импульсов. Следовательно, частота следования импульсов переполнения на выходе счетчика в N_V раз меньше, чем на входе, что и требуется. При этом фазовый детектор можно заменить временным различителем, не нарушая принципа работы схемы. Обычно коэффициент деления счетчика-делителя ограничен. Для изменения частоты в широких пределах используются синтезаторы частоты с двумя и более счетчиками-делителями: одним счетчиком управляют младшие разряды числа N_V , другим — старшие.

8.7. СЛЕДЯЩИЕ ИЗМЕРИТЕЛИ НАПРАВЛЕНИЯ

Следящие измерители направления (СИН) определяют текущие угловые координаты объекта путем непрерывного отслеживания направления прихода волны $u_x(t)$, $u_y(t)$ опорным направлением антенны $u^0_x(t)$, $u^0_y(t)$. В режиме слежения опорное направление отслеживает лишь плавные, закономерные изменения угловых координат, а случайные возмущения подавляются; при этом положение опорного направления становится оценкой направления на объект: $u^0_x(t) = \hat{u}_x(t)$, $u^0_y(t) = \hat{u}_y(t)$. Поскольку угловых координат две, используют два контура слежения, которые почти всегда независимы. Следящий измеритель направления включает угловой дискриминатор, формирователь оценки и синтезатор поворота, а также схему поиска по направлению. Обнаружитель с реле захвата СПЗ общий для всех следящих измерителей, входящих в данную систему.

Для захвата сигнала объекта следящей измерительной системой необходим поиск в четырехмерном пространстве параметров сигнала, т. е. по дальности, радиальной скорости и обоим угловым координатам. Так как на это потребовались бы недопустимо большие затраты времени, всемерно сокращают диапазон поиска,

используя априорную информацию. Эта информация поступает либо от специальных обзорных систем, либо от самой измерительной системы, работающей первоначально в режиме обзора. Пониженные требования к точности обзорных систем и отсутствие необходимости поиска по дальности (а иногда и скорости) значительно сокращают затраты времени на обзор всей зоны.

Управление положением опорного направления выполняется с помощью синтезатора поворота либо механическим разворотом всей антенны, либо электрическим смещением антенных лучей относительно неподвижной решетки. В соответствии с этим сами синтезаторы поворота подразделяют на механические и электронные. Механический синтезатор поворота — электродвигатель антенного привода — одновременно выполняет и роль последнего интегратора в схеме формирователя оценки, так как его угловая скорость пропорциональна приложенному напряжению, а сам угол — интегралу от него. Электронные синтезаторы поворота управляют фазовращателями антенной решетки, электрически смещающими опорное направление. Они подразделяются на аналоговые и цифровые: первые под воздействием выходного сигнала формирователя оценки вырабатывают управляющие напряжения фазовращателей, вторые — управляющий код, разный для фазовращателей различных элементов решетки. Следовательно, выходное напряжение или код формирователя оценки преобразуются в управляющие напряжения или коды для каждого фазовращателя.

Основным элементом СИН является угловой дискриминатор. Он охватывает почти все элементы приемного тракта системы, включая антенну. Угловые дискриминаторы могут быть любого вида, свойственного дискриминаторам вообще: пространственного фильтра или пространственного коррелятора, оптимального или подоптимального. В последнем случае они подразделяются на амплитудные или фазовые, линейные (аддитивные) или нелинейные (мультипликативные). Кроме того, по способу образования сигнала рассогласования (параллельный или последовательный) подоптимальные угловые дискриминаторы разделяют на многоканальные (моноимпульсные) и одноканальные (сканирующие). В многоканальных дискриминаторах сигнал углового рассогласования вырабатывается в результате одновременного сравнения сигналов нескольких каналов приема с приемом, даже одного импульса (откуда они и получили название моноимпульсных). В одноканальных системах сканирующий антенный луч, последовательно занимая различные положения относительно опорного направления, позволяет получить сигнал рассогласования путём сравнения амплитуды сигналов в различные моменты времени. Для получения сигнала рассогласования требуется время не меньше, одного периода сканирования. Широкое распространение получили одноканальные СИН с коническим сканированием, при котором максимум антенного луча описывает в пространстве ко-

ническую поверхность относительно опорного (равносигнального) направления.

В СИН с независимыми каналами слежения по каждой из двух угловых координат в принципе достаточно рассмотреть одномерный угловой дискриминатор. Структура одномерного модифицированного оптимального углового дискриминатора в любом из двух каналов, измеряющих u_x или u_y , согласно общей формуле (8.16) определяется его характеристикой

$$D(u_{x,y} - u_{x,y}^0) = -K_1 Z'_s(u_{x,y} - u_{x,y}^0) / Z_s(u_{x,y} - u_{x,y}^0). \quad (8.31)$$

Входная смесь $X(v_{x,y}) = U_0 G(v_{x,y}) \exp(j2\pi v_{x,y} u_{x,y}) + N$ в каждой точке раскрыва v_x или v_y содержит сигнал, зависящий от направления прихода волны $u_{x,y}$ и неизвестного амплитудного множителя U_0 , и шум N , пересчитанный на вход антенны, который полагаем независимым от координаты раскрыва и направления приема (не коррелированный в пространстве). Если система настроена на направление $u_{x,y}^0$, то при симметричной функции раскрыва огибающая пространственного корреляционного интеграла имеет вид

$$\begin{aligned} Z(u_{x,y} - u_{x,y}^0) &= \\ &= \int_{-\infty}^{\infty} X(v_{x,y}) G^*(v_{x,y}) \exp[j2\pi v_{x,y} (u_{x,y} - u_{x,y}^0)] dv_{x,y}, \end{aligned} \quad (8.32)$$

а ее производная

$$\begin{aligned} Z'(u_{x,y} - u_{x,y}^0) &= \\ &= \int_{-\infty}^{\infty} X'(v_{x,y}) \{j2\pi v_{x,y} G^*(v_{x,y}) \exp[j2\pi v_{x,y} (u_{x,y} - u_{x,y}^0)]\} dv_{x,y}. \end{aligned} \quad (8.33)$$

Таким образом, для формирования сигнала углового рассогласования необходимо иметь опорный канал с пространственной функцией передачи

$$H_0(v_{x,y}) = G^*(v_{x,y}) \exp(-j2\pi v_{x,y} u_{x,y}^0) \quad (8.34)$$

и измерительный канал с функцией передачи

$$H_1(v_{x,y}) = j2\pi K_1 v_{x,y} G^*(v_{x,y}) \exp(-j2\pi v_{x,y} u_{x,y}^0). \quad (8.35)$$

Оптимальный угловой дискриминатор можно построить только на основе ФАР, т. е. при дискретном отсчете координаты $v_{x,y,i}$. При вещественной и симметричной функции раскрыва ФАР коэффициент передачи для i -го элемента ФАР в опорном канале

$$B_i = B(v_{x,y,i}) = G_i = B_{-i}, \quad (8.36)$$

в измерительном канале (с учетом коэффициента K_1 формулы (8.31))

$$B'_i = B'(v_{x,y,i}) = -2\pi K_1 v_{x,y,i} G_i = -B'_{-i}, \quad (8.37)$$

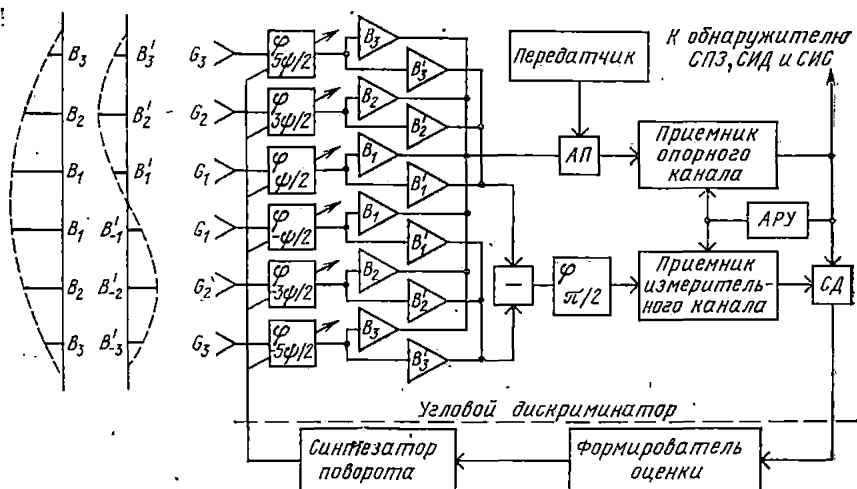


Рис. 8.18. Одномерный оптимальный угловой дискриминатор в составе слеящего измерителя направления

т. е. имеет нечетную симметрию. Кроме того, учитывая множитель $j = \exp(j\pi/2)$ в формуле (8.35), в измерительном канале необходимо поставить фазовращатель на $\pi/2$. Настройка обоих каналов, опорного и измерительного, на заданное направление $\alpha_{x,y}$ достигается изменением крутизны линейного по раскрыву набега фазы с помощью перестраиваемых фазовращателей.

Из полученных соотношений непосредственно вытекает структура одномерного оптимального углового дискриминатора и СИН в целом. В качестве примера на рис. 8.18 показан угловой дискриминатор типа пространственного фильтра, в котором для ввода отрицательных весовых коэффициентов измерительного канала используется вычитающее устройство, а в опорном приемном канале через антенный переключатель (АП) подключен передатчик системы. Приемники опорного и измерительного каналов одностипны, содержат преобразователи частоты, секторы по дальности и схему ЧФАП (не показанные на схеме). В них производится оптимальная временная фильтрация сигналов и нормировка с помощью схемы АРУ по сигналу опорного канала. Опорный канал представляет собой уже рассмотренный в гл. 4 оптимальный пространственно-временной фильтр. Дополнительные отводы с весовыми коэффициентами B'_i и последующие цепи образуют измерительный (разностный) канал, в котором сигнал опорного канала используется в качестве опорного напряжения синхронного детектора (СД). Это позволяет определить знак углового рассогласования по полярности выходного сигнала СД, а величину рассогласования — по его абсолютному значению. Выходной сигнал СД обычно имеет вид коротких импульсов, из которых благодаря стробированию приемников (селекции) по дальности выделяется постоянная составляющая, являющаяся входным воздей-

ствием аналогового формирователя оценки, а перед цифровым формирователем оценки выходной сигнал СД преобразуется в код.

Отличие углового дискриминатора типа пространственного коррелятора состоит только в том, что фазовращатели, через которые проходят входные сигналы элементов ФАР, заменены смесителями, опорные сигналы которых подаются от блока фазовращателей, входящих в этом случае в состав синтезатора поворота.

Совершенно аналогично можно построить на ФАР подоптимальные линейные угловые дискриминаторы, которые по своим характеристикам уступают оптимальным дискриминаторам, в то время как аппаратура не упрощается. Поэтому подоптимальные угловые дискриминаторы целесообразно создавать только на основе зеркальных антенн, когда оптимальную структуру реализовать невозможно. Подоптимальные одномерные линейные угловые дискриминаторы на зеркальных антеннах (амплитудные и фазовые) также содержат опорный и измерительный каналы, на которые сигналы, воспринимаемые облучателями антенны, поступают через суммарно-разностный мост (рис. 8.19). В угловых дискриминаторах используют суммарно-разностные мосты различных типов (на рисунке изображен кольцевой).

Амплитудные разностные угловые дискриминаторы (рис. 8.19,а) строятся на двухзеркальных антеннах Кассегрена, а суммарно-разностные (рис. 8.19,б) — на двух- и однозеркальных (по-

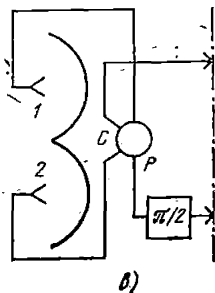
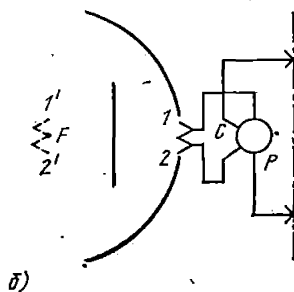
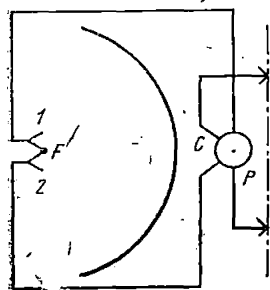
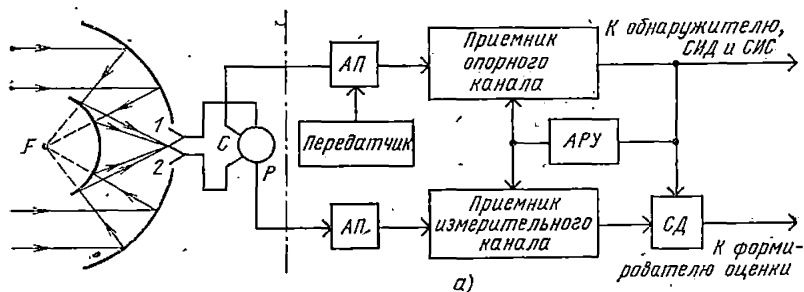


Рис. 8.19. Схемы одномерного подоптимального разностного амплитудного углового дискриминатора (а) и входные части суммарно-разностного амплитудного (б) и фазового (в) угловых дискриминаторов на зеркальных антеннах

казана только входная часть, отделенная на рис. 8.19,а штрих-пунктирной линией). Двухзеркальные антенны удобны тем, что максимально сокращена протяженность фидеров перед суммарно-разностным мостом, а следовательно, исключаются погрешности, вносимые неидентичностью их характеристик. Кроме того, уменьшается продольный размер антенны, который у длиннофокусных однозеркальных антенн достаточно велик. Плоский контррефлектор суммарно-разностной схемы как бы выносит облучатели в то место относительно основного зеркала, где у однозеркальной антенны располагаются реальные излучатели. Затенение основного зеркала контррефлектором можно практически полностью исключить, выбрав специальную конструкцию обоих зеркал. Максимальная точность суммарно-разностного дискриминатора достигается при сравнительно большом разnose лучей, когда суммарный сигнал резко снижается. В разностном угловом дискриминаторе этот недостаток устраняется благодаря фокусирующим свойствам двухзеркальной антенны Кассегрена, сохраняющим максимальным суммарный сигнал (двухзеркальная антенна с плоским контррефлектором, эквивалентная однозеркальной, этим свойством не обладает).

Фазовый линейный угловой дискриминатор отличается от амплитудного только тем, что антенна состоит из двух одинаковых зеркал, разнесенных на расстояние b , диаграммы направленности которых (в дальней зоне) совпадают, а также дополнительным фазовращателем на $\pi/2$ (рис. 8.19,в). Всем подоптимальным дискриминаторам на зеркальных антеннах присущ общий недостаток — невозможность оптимального весового суммирования сигналов по раскрыву. Однако возникающие при этом потери в отношении сигнал-шум весьма незначительны.

Обычно угловые дискриминаторы являются двумерными, они измеряют рассогласование по двум угловым координатам. Оптимальный двумерный дискриминатор на кондуктивных плоских ФАР содержит опорный сигнал, являющийся согласованным пространственно-временным фильтром или коррелятором, описанным в гл. 4. Двумерный дискриминатор в отличие от одномерного в канале каждого элемента ФАР имеет по два дополнительных отвода с весовыми коэффициентами B_{ij} и B'_{ij} , которые объединяются в два измерительных канала — для каждой из двух угловых координат. Опорный канал является общим, а управление фазовращателями производится двумерным синтезатором поворота, формирующим требуемое распределение фаз по обеим координатам раскрыва ФАР.

В двумерных дискриминаторах на зеркальных антеннах (рис. 8.20), как правило, используют четыре входных канала — по числу облучателей. Сигналы этих каналов объединяют в три выходных канала (опорный и два измерительных) с помощью схемы линейного преобразования (СЛП). Структура СЛП зависит от расположения облучателей — рядного или крестообразного. Схе-

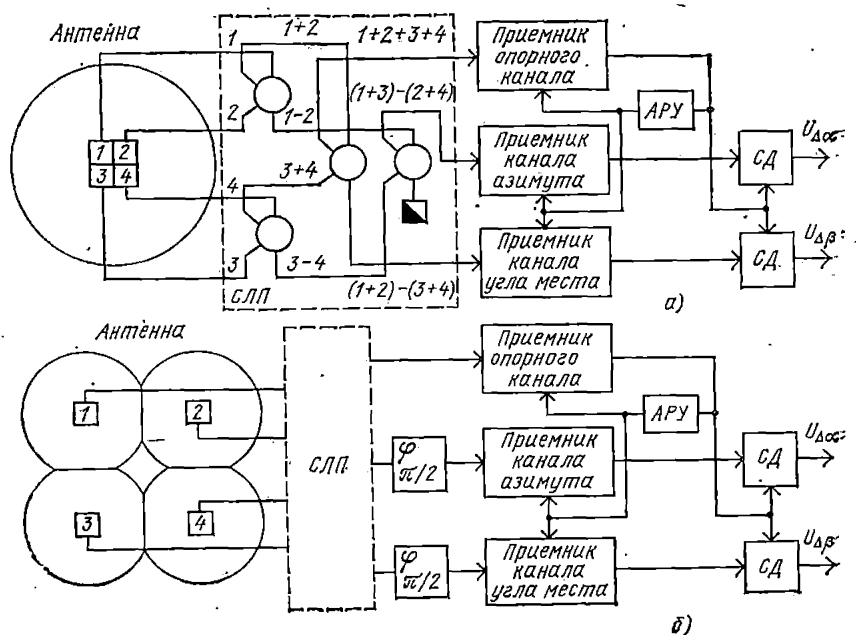


Рис. 8.20. Схемы двумерного подоптимального суммарно-разностного углового дискриминатора на зеркальных антеннах

мы линейного преобразования дискриминаторов с рядным расположением облучателей содержат четыре суммарно-разностных моста, причем разностный выход последнего моста в обеих схемах подключают к поглощающей нагрузке. В ней происходит объединение сигналов по рядам (правому и левому, верхнему и нижнему) с последующим вычитанием, причем в одном из каналов разность сумм получается из суммы разностей. Фазовые дискриминаторы отличаются от амплитудных только антенной и наличием фазовращателей на $\pi/2$.

Зеркальные антенны двумерных фазовых дискриминаторов по конструкции, использованию площади раскрыва и форме диаграммы направленности уступают антеннам амплитудных угловых дискриминаторов, поэтому фазовые линейные дискриминаторы используют значительно реже амплитудных. Однако амплитудные угловые дискриминаторы с разнесенными облучателями не обеспечивают согласования характеристических сопротивлений волноводов и открытого пространства, в результате ухудшается коэффициент бегущей волны, возрастают потери, снижается точность. Поэтому используют однорупорные моноимпульсные облучатели с возбуждением нескольких типов волн (мод), которые заменяют обычный четырехрупорный облучатель вместе со схемой линейного преобразования.

При использовании линейно-поляризованных волн применяют рупоры на волноводах квадратного сечения, размеры которых допускают существование четырех низших типов волн: H_{10} , H_{20} , H_{11} , E_{11} (рис. 8.21). Волна типа H_{10} наводится в рупоре вертикально поляризованной волной, приходящей с любого направления в пределах основного лепестка диаграммы направленности антенны; она образует опорный сигнал E_0 . Волны H_{20} , H_{11} , E_{11} наводятся в рупоре только при несимметричном его возбуждении, т. е. волной, приходящей под углом к оси зеркала. Колебание H_{20} возбуждается волной, приходящей от смещенного по азимуту объекта, так как фокусируемый зеркалом максимум облучения (дифракционное пятно) смещается при этом из фокуса (центра рупора) вправо или влево. Наводимое поле тем интенсивнее, чем больше азимутальное смещение, а полярность волны H_{20} по отношению к волне H_{10} (знак) зависит от стороны смещения. В направлении оси зеркала условия для возбуждения несимметричной волны H_{20} отсутствуют, так как максимум возбуждения рупора располагается строго в центре. Таким образом, волна H_{20} дает сигнал E_α , который в функции от угла рассогласования образует характеристику углового дискриминатора по азимуту. К сожалению, при вертикальной поляризации не существует такого типа волны, который обеспечивает аналогичную зависимость по углу места. Поэтому сигнал рассогласования по углу места β получают как сумму волн двух типов H_{11} и E_{11} .

Многомодовый облучатель для вертикально поляризованной волны включает волновод с тремя секциями и рупор. В ближайшей к рупору секции квадратного сечения существуют все четыре типа волн, в широком прямоугольном сечении — волны H_{11} и

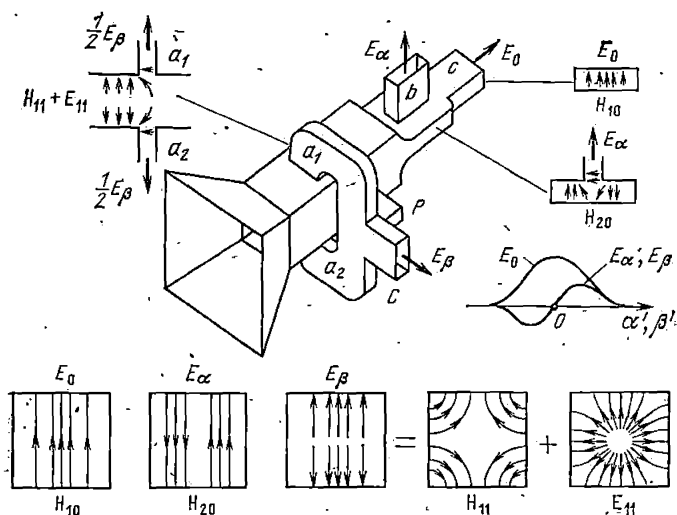


Рис. 8.21. Схема многомодового моноимпульсного облучателя и структура волн в нем

E_{11} , а в секции с узким прямоугольным сечением, являющейся отводом c для суммарного сигнала E_0 , распространяется только волна H_{10} . В отводе b второй секции существуют условия для отбора мощности волны H_{20} , образующей разностный сигнал азимута E_α . В отводах a_1 и a_2 квадратной секции волновода отбирается разностный сигнал угла места E_β , наводимый в них волнами H_{11} и E_{11} . Для сохранения симметрии поля в квадратном волноводе используется не один, а два отвода, синфазные сигналы которых объединяются в суммарном отводе C дополнительного суммарно-разностного моста (разностный отвод P этого моста подключают к поглощающей нагрузке). Многомодовые моноимпульсные облучатели отличаются предельной простотой конструкции. Однако к точности их изготовления предъявляются весьма высокие требования, так как малейшая асимметрия вызывает паразитные сигналы рассогласования.

Серьезным недостатком рассмотренных схем угловых дискриминаторов является использование далеко не идеальной нормировки с помощью схемы АРУ. Для реализации идеальной нормировки используется схема преобразования амплитудных соотношений в фазовые, которая устанавливается непосредственно на выходе суммарно-разностного моста (рис. 8.22). Схема амплитудного дискриминатора содержит два фазовращателя, сдвигающих фазу на $\pi/2$, и два сумматора, с помощью которых из суммарного U_Σ и разностного U_Δ сигналов формируются два одинаковых по амплитуде напряжения $U_1 = U_\Sigma + jU_\Delta$ и $U_2 = U_\Delta + jU_\Sigma$. Фазовый сдвиг между U_1 и U_2 , равный $\omega = \pi/2 - \varphi$, однозначно связан с отношением U_Δ/U_Σ . Поэтому измеряемая разность

$$\varphi = \pi/2 - \omega = 2 \arctg(U_\Delta/U_\Sigma) \quad (8.38)$$

является монотонной функцией характеристики амплитудного дискриминатора $D(u_{x,y} - u_{0x,y}) = U_\Delta/U_\Sigma$, линейной при малых рассогласованиях. Полученную разность фаз можно преобразовать затем в код в цифровом фазометре.

Нетрудно показать, что в фазовом угловом дискриминаторе разность фаз (8.38) равна разности фаз сигналов на выходе антенны $\Delta\varphi = 2\pi b(u_{x,y} - u_{0x,y})/\lambda$, а сама преобразующая схема не-

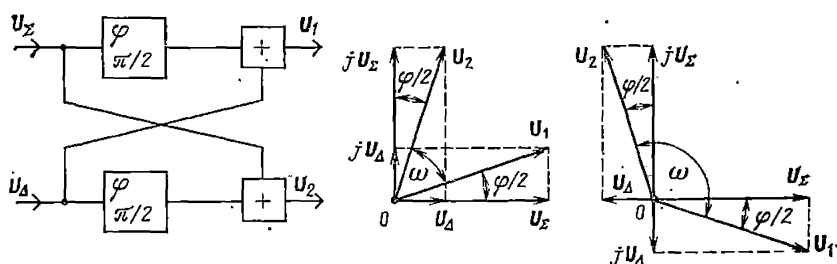


Рис. 8.22. Преобразование амплитудных соотношений суммарного и разностного сигналов в фазовые

сколько видоизменяется, так как разностный сигнал сдвинут на $\pi/2$ относительно суммарного. К сожалению, преобразование полученной разности фаз в код возможно только на сравнительно низкой частоте, после преобразования частоты. Период колебаний этой частоты и взаимный временной сдвиг сравниваемых сигналов U_1 и U_2 в импульсных системах оказывается во много раз больше длительности импульса. Поэтому непосредственное преобразование частоты импульсных сигналов невозможно. Необходимо в каждом из двух отводов схемы преобразования ставить схемы ЧФАП, вырабатывающие непрерывные колебания, которые совпадают по фазе с высокочастотными импульсами U_1 и U_2 . Эти колебания уже можно переносить на низкую частоту для точного измерения разности фаз цифровым фазометром. Однако аппаратура при этом существенно усложняется.

Нелинейные амплитудные угловые дискриминаторы в соответствии с общей структурой нелинейных дискриминаторов содержат в канале каждого облучателя приемники с логарифмической амплитудной характеристикой (для малых сигналов линейной) и схему вычитания (рис. 8.23,а). Нелинейные фазовые угловые дискриминаторы определяют разность фаз сравниваемых сигналов с помощью фазометров, а для нормировки используется глубокое безынерционное (не вызывающее фазовых сдвигов) ограничение амплитуды (рис. 8.23,б). Реализовать идентичные приемники с ограничением значительно проще, чем логарифмические, поэтому нелинейные фазовые угловые дискриминаторы нашли более широкое применение, чем амплитудные, в отличие от линейных.

В целом нелинейные угловые дискриминаторы по своим показателям уступают линейным. Во-первых, они работоспособны только при достаточно большом отношении сигнал-шум (при слабых сигналах в амплитудных дискриминаторах характеристика отличается от логарифмической, а в фазовых отсутствует ограничение амплитуды). Во-вторых (что более важно), сравнение сигналов происходит не на входе, а на выходе приемных устройств. Поэтому неидентичность характеристик приемных каналов вызывает паразитное смещение опорного направления; в линейных дискриминаторах это смещение отсутствует, так как сигналы сравниваются еще до прохождения через приемный тракт.

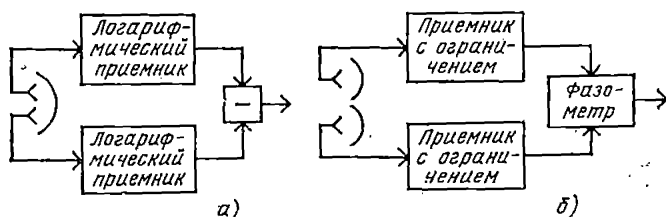


Рис. 8.23. Схемы нелинейных подоптимальных угловых дискриминаторов: амплитудного логарифмического (а) и фазового интерферометрического (б)

В-третьих, в двумерных нелинейных угловых дискриминаторах необходимо иметь четыре приемных канала — по два на каждую координату, в то время как в линейных требуется только три или меньше (в линейных системах с коническим сканированием используется всего один приемный канал).

В заключение кратко остановимся на синтезаторах поворота. Механические синтезаторы поворота — электродвигатели привода антенны — не требуют специального рассмотрения. Аналоговый электронный синтезатор поворота, в котором переменные фазовые сдвиги достигаются путем изменения частоты колебаний, был описан в гл. 4 (см. рис. 4.4, в). Возможны и другие виды аналоговых электронных синтезаторов поворота, например волновод с отводами, разность фаз между выходными сигналами которых изменяется при изменении ширины прямоугольного волновода или воздействии магнитным полем на феррит, помещенный внутрь этого волновода. Более подробного рассмотрения требуют цифровые синтезаторы поворота.

Для ввода фазового сдвига φ_{ij}^0 в канал ij -го элемента ФАР синтезатор поворота содержит специализированную цифровую вычислительную машину (СЦВМ), преобразующую коды формирования оценки, пропорциональные отслеживаемым угловым координатам $u_x^0 = \hat{u}_x$ и $u_y^0 = \hat{u}_y$, в управляющие коды в соответствии с формулами (4.15) и (4.12). Так как в цифровых фазовращателях выдается квантованное значение фазы с шагом $\delta\varphi$ (обычно $\delta\varphi = \pi/8 = 22,5^\circ$), СЦВМ выбирает в канале каждого элемента такие квантованные значения фазы, которые наиболее близки к требуемым при настройке на направление u_x^0, u_y^0 . Если угловой дискриминатор выполнен в виде пространственного коррелятора, цифровой синтезатор поворота включает также управляемые цифровые фазовращатели, включенные в цепь гетеродинного сигнала каждого элемента ФАР. Эти фазовращатели подобны цифровым фазовращателям, непосредственно входящим в тракт принимаемых сигналов дискриминаторов типа пространственного фильтра.

Различают проходные и отражательные цифровые фазовращатели. Ограничимся рассмотрением цифровых проходных фазовращателей, коммутируемых *pin*-диодами (рис. 8.24). Звенья

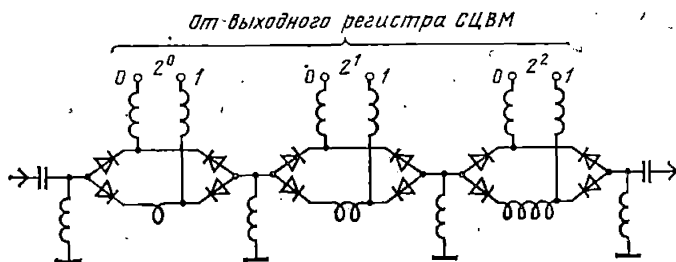


Рис. 8.24. Схемы цифровых фазовращателей проходного типа, коммутируемые *pin*-диодами

такого фазовращателя (показано три звена) имеют две ветви: верхнюю и нижнюю. Длина верхней ветви l_0 одинакова у всех звеньев (включая длину соединительных проводов), а длина нижних ветвей $l_0 + 2^i \delta l$ отличается от первой на величину, пропорциональную значению единицы i -го разряда управляющего двоичного кода ($i=0, 1, \dots, n-1$). Ветви каждого звена подключены к выходным триггерам СЦВМ соответствующих разрядов: верхняя ветвь к инверсному выходу (0), нижняя к прямому (1). В состоянии 0 триггер данного разряда подает напряжение прямого смещения на *pin*-диод верхней ветви соответствующего звена, в состоянии 1 проводящей становится нижняя ветвь. Благодаря этому суммарный путь, проходимый СВЧ сигналом через всю цепочку, равен $nl_0 + N\delta l$, где N — управляющее число, n — число разрядов (звеньев). Следовательно, и фазовый сдвиг, получаемый проходящим через цепочку сигналом, с точностью до постоянной nl_0 пропорционален управляющему числу N . Так, при $N=5$ (101) сигнал пройдет через нижнюю ветвь первого и третьего звеньев и верхнюю ветвь второго звена, получив суммарную задержку $(l_0 + \delta l) + l_0 + (l_0 + 4\delta l) = 3l_0 + 5\delta l$. Конденсаторы и дроссели в схеме служат для разделения цепей СВЧ сигнала и управляющего напряжения. Несколько сложнее протекают процессы в отражательных фазовращателях, где каждое звено, также управляемое через *pin*-диоды, содержит либо суммарно-разностный мост, либо рециркулятор. Все эти фазовращатели обладают высоким быстродействием и не требуют большой мощности для управления.

8.8. ВЗАИМОДЕЙСТВИЕ СЛЕДЯЩИХ ИЗМЕРИТЕЛЕЙ НЕСКОЛЬКИХ ПАРАМЕТРОВ СИГНАЛА

Взаимодействие каналов слежения по дальности и скорости при сигналах сложной формы. Для обеспечения линейности временных дискриминаторов при сравнительно малом отношении сигнал-шум в каждом импульсе необходимо синхронное детектирование его входного (или выходного) сигнала. Опорное напряжение синхронного детектора формируется в следящем измерителе скорости, который играет роль схемы ЧФАП. Таким образом, оптимальные условия функционирования СИД обеспечиваются каналом СИС.

Схема ЧФАП эквивалентна узкополосному фильтру, выделяющему несущую входного сигнала и подавляющему боковые спектральные составляющие и шумы. Однако никакой линейный фильтр не в состоянии выделить несущую, если она отсутствует в спектре входного сигнала. Поэтому при работе по сложным сигналам, у которых несущая полностью или почти полностью подавлена, непосредственно использовать описанную схему СИС (ЧФАП) нельзя. Такими сигналами являются, в частности, псевдошумовые фазоманипулированные сигналы, включая непрерыв-

ные M -последовательности. Необходимо во входном сигнале СИС восстановить несущую путем демодуляции сигнала, т. е. произвести свертку спектра, описанную в § 6.12. Для работы демодулятора необходимо опорное напряжение, совпадающее по форме и временному положению с входным сигналом. Совпадение опорного напряжения по времени с входным достигается с помощью СИД, т. е. условия функционирования СИС обеспечиваются каналом СИД. Таким образом, в общем случае СИД и СИС взаимно связаны между собой через общий входной сигнал даже в отсутствие корреляции между временным и частотным сдвигами. При этом принимаемый сигнал на вход рассмотренной схемы СИС нужно подать после демодулятора несущей, а на вход СИД — после синхронного детектора (т. е. временной различитель СИД в этом случае оперирует с видеосигналами).

Для ввода СИД и СИС в режим слежения необходимо, чтобы расстройки по времени и частоте попали в узкий пик тела неопределенности псевдошумовых сигналов, т. е. необходим совместный поиск (перестройка) измерителей по времени и частоте. Траектория поиска на плоскости время-частота имеет вид плотного телевизионного раstra, в котором смещение строк не может превышать ширины пика тела неопределенности во избежание пропуска сигнала. В результате на поиск сигналов сложной формы затрачивается значительно больше времени. Зона поиска и время поиска существенно сокращаются при наличии априорных данных о дальности и скорости объекта.

Таким образом, использование сложных сигналов вносит специфику в структуру следящих измерителей и их функционирование. Эта специфика наиболее ярко проявляется при использовании непрерывных фазоманипулированных сигналов типа M -последовательности. При наличии демодулятора на входе СИС, восстанавливающего непрерывную несущую M -последовательности, сама схема СИС не имеет особенностей. Однако схема СИД существенно изменяется, что обусловлено самим способом формирования и свойствами M -последовательности (см. § 6.12). Напомним, что для формирования M -последовательности длиной $N=2^n-1$ используется генератор кода в виде n -разрядного регистра сдвига с обратной связью через сумматор по $\text{mod } 2$ с выходного каскада регистра и одного или нескольких промежуточных каскадов, работой которого управляют тактовые импульсы (рис. 6.22). Затем высокочастотная несущая, вырабатываемая задающим генератором передатчика, в модуляторе манипулируется по фазе в соответствии с выходным кодом (видеосигналом), снимаемым с одного из каскадов регистра генератора кода. С помощью точно такого же генератора формируется и следящий код в тракте приема. В корреляторе обнаружителя для получения огибающей корреляционного интеграла в качестве опорного используется точно такой же код, как и у входного сигнала. Однако для получения нечетной характеристики временного дискриминатора опорный код должен отличаться от входного. В оптимальном дискримина-

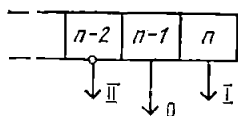
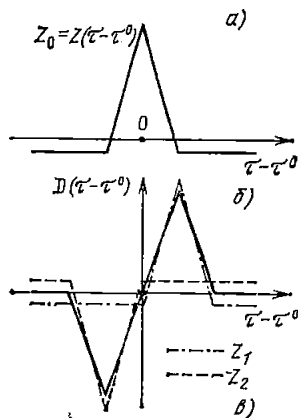


Рис. 8.25. Формирование опорных и выходных сигналов M -последовательности:

a — съем опорных сигналов; b — выходной сигнал обнаружителя; $в$ — характеристика временного дискриминатора

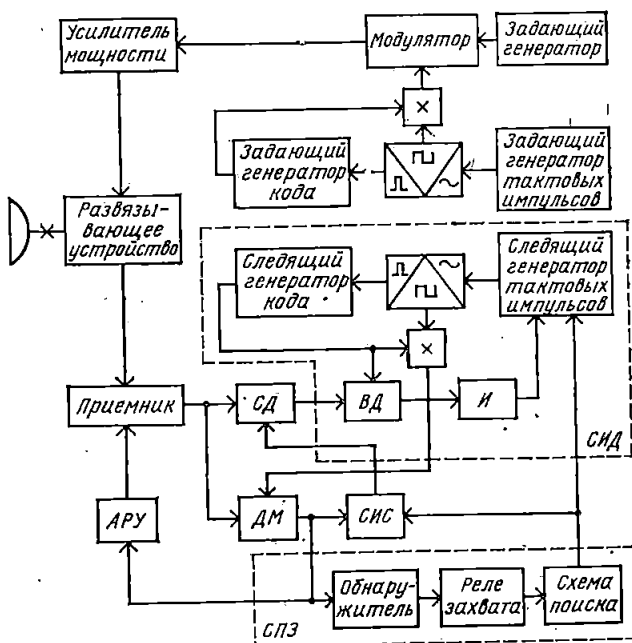


торе согласно общей теории измерителей в качестве опорного сигнала коррелятора используют производную огибающей входного сигнала, что в данном случае нельзя реализовать, так как M -последовательность — разрывная функция. Поэтому используется подоптимальный разностный временной дискриминатор с характеристикой (8.29), подобный описанному выше (см. рис. 8.8, а). В данном случае отличие состоит в том, что опорные сигналы, различающиеся сдвигом, снимаются не с отводов линии задержки, а с выхода смежных каскадов регистра генератора следящего кода (рис. 8.25).

Опорный сигнал обнаружителя O снимается с предпоследнего каскада регистра генератора кода (рис. 8.25, а). На выходе обнаружителя формируется напряжение, зависимость которого от расстройки повторяет огибающую корреляционного интеграла $Z_0 = Z(\tau - \tau^0)$ (рис. 8.25, б). Опорные сигналы вилки, опережающий I и отстающий II , снимаются с соседних каскадов (последний в противофазе). В результате в корреляторах временного дискриминатора образуются два противофазных напряжения, описываемые огибающими $Z_1 = Z(\tau - \tau^0 - \tau_c)$ и $Z_2 = Z(\tau - \tau^0 + \tau_c)$, где τ_c — длительность такта, которые в сумме образуют нечетно-симметричную характеристику дискриминатора (рис. 8.25, в).

Примерно такой же вид имеет характеристика временного дискриминатора, когда излучаемый сигнал модулируется сложным кодом, являющимся результатом перемножения видеосигнала M -последовательности и меандра, период которого равен одному такту (эта операция эквивалентна суммированию по $\text{mod } 2$ кодов M -последовательности и меандра). В этом случае временной дискриминатор представлен единственным коррелятором, опорным напряжением которого является M -последовательность, а выходным — нечетная функция взаимокорреляции простого и сложного кодов. На коррелятор обнаружителя в качестве опорного подается сложный код, что дает на выходе функцию корреляции входного и опорного сложных кодов. В излучаемом сигнале сложного кода под воздействием меандра одна половина каждого такта высокочастотного заполнения находится в одной фазе, другая в противоположной, в результате несущая подавляется полностью, в то время как в M -последовательности сохраняется остаток несущей порядка $1/N$.

Для усвоения изложенных идей рассмотрим в качестве приме-



ра функциональную схему системы непрерывного излучения, использующей сложный код (рис. 8.26). В системе можно выделить четыре основные части: тракт передатчика, следящий измеритель дальности (СИД), следящий измеритель скорости (СИС) и схему поиска и захвата (СПЗ), входящие в приемный тракт (ввиду отсутствия особенностей схема СИС не раскрывается). Так как системы непрерывного излучения используют сигнал, переизлученный на другой частоте, при работе на общую антенну включают устройство, развязывающее по частоте тракты передачи и приема. Процессы в данной системе протекают следующим обра-

Принятые колебания параллельно поступают на синхронный детектор (СД) и демодулятор (ДМ). С выхода СД видеосигнал сложного кода поступает в тракт СИД, а с выхода ДМ восста-

новленная несущая проходит в тракт СИС, на обнаружитель СПЗ и схему АРУ для нормировки.

Следящий измеритель дальности содержит временной дискриминатор (ВД), набор интеграторов (И), которые вместе со следящим генератором тактовых импульсов образуют формирователь оценки, и синтезатор задержки, состоящий из тех же элементов, что и канал формирования кода передающего тракта. Частота синусоидальных колебаний следящего генератора тактовых импульсов отличается от номинальной (т. е. частоты задающего) на величину, пропорциональную приложенному напряжению. Следовательно, фаза этих колебаний, а также временное положение тактовых импульсов и всего следящего кода пропорционально интегралу от приложенного напряжения. Тем самым этот генератор выполняет роль последнего (а иногда и единственного) интегратора в схеме формирователя оценки. Временной дискриминатор в результате умножения входного сигнала сложного кода на простой опорный выдает сигнал, пропорциональный рассогласованию по времени между входным и опорным сигналами.

В режиме точного слежения временное положение следящего кода относительно излучаемого является оценкой текущей дальности. Чтобы получить ее, специальная схема (не показанная на рис. 8.26) вырабатывает импульсы отсчета, когда генераторы кода (задающий и следящий) находятся в определенном состоянии (например, единичном во всех разрядах регистра). Интервал между импульсами отсчета, снимаемыми с этих двух генераторов, и фиксирует оценку текущей дальности. Этот интервал можно преобразовать в число путем заполнения его счетными импульсами.

Отказ от измерения дальности или скорости не упрощает схему, поскольку для функционирования хотя бы одного канала нужен второй. Более того, схема остается прежней, даже если не измеряется ни один из этих параметров, а вся система используется для синхронизации связи между объектом и станцией. В этом случае на излучаемый сигнал, форма которого известна, с помощью фазовой манипуляции накладывается неизвестное получателю сообщение, выделяемое затем из принятого сигнала компенсацией известного кода.

Совмещенные дискриминаторы взаимно зависимых параметров. При совместном отслеживании нескольких взаимно зависимых параметров сигнала вместо единственного сигнала рассогласования $\nu' = Z'/Z''$, определяемого алгоритмом работы оптимального дискриминатора (8.6), необходимо вычислить совокупность (вектор) сигналов рассогласования по всем измеряемым параметрам

$$\nu' = [Z'']^{-1} Z', \quad (8.39)$$

где $[Z'']^{-1}$ — матрица, обратная матрице вторых смешанных частных производных Z'' , а Z' — вектор первых частных производных огибающей корреляционного интеграла по зависимым из-

меряемым параметрам, число которых обычно не превышает двух (дальность и радиальная скорость или две угловые координаты)...

В частности, при совместном отслеживании задержки τ (дальности) и частотного сдвига F (радиальной скорости), которые при использовании ЛЧМ сигнала коррелированы между собой, по аналогии с формулами (6.69) имеем

$$\mathbf{Z}'' = \begin{bmatrix} Z''_{\tau\tau} & Z''_{\tau F} \\ Z''_{F\tau} & Z''_{FF} \end{bmatrix}, \quad \mathbf{Z}' = \begin{bmatrix} Z'_{\tau} \\ Z'_{F} \end{bmatrix}.$$

В соответствии с правилом (6.64) находим обратную матрицу

$$\begin{aligned} [\mathbf{Z}'']^{-1} &= \frac{1}{Z''_{\tau\tau} Z''_{FF} - (Z''_{\tau F})^2} \begin{bmatrix} Z''_{FF} & -Z''_{F\tau} \\ -Z''_{\tau F} & Z''_{\tau\tau} \end{bmatrix} = \\ &= \frac{1}{1 - \rho_{\tau F}^2} \begin{bmatrix} \frac{1}{Z''_{\tau\tau}} & -\frac{\rho_{\tau F}^2}{Z''_{\tau F}} \\ -\frac{\rho_{\tau F}^2}{Z''_{\tau F}} & \frac{1}{Z''_{FF}} \end{bmatrix}, \end{aligned}$$

где $\rho_{\tau F}^2 = (Z''_{\tau F})^2 / Z''_{\tau\tau} Z''_{FF}$ — квадрат коэффициента время-частотной корреляции ($Z''_{\tau F} = Z''_{F\tau}$). При этом вектор рассогласования (8.39) принимает вид

$$\mathbf{v}' = \begin{bmatrix} \tau' \\ F' \end{bmatrix} = \frac{1}{1 - \rho_{\tau F}^2} \begin{bmatrix} \frac{Z'_{\tau}}{Z''_{\tau\tau}} & -\frac{\rho_{\tau F}^2 Z'_{F}}{Z''_{\tau F}} \\ -\frac{\rho_{\tau F}^2 Z'_{\tau}}{Z''_{\tau F}} & \frac{Z'_{F}}{Z''_{FF}} \end{bmatrix}.$$

Этому вектору соответствует алгоритм работы совмещенного дискриминатора по времени и частоте:

$$\begin{aligned} \tau' &= \frac{1}{1 - \rho_{\tau F}^2} \left(\frac{\partial Z / \partial \tau'}{\partial^2 Z / \partial \tau'^2} - \rho_{\tau F}^2 \frac{\partial Z / \partial F'}{\partial^2 Z / \partial \tau' \partial F'} \right), \\ F' &= \frac{1}{1 - \rho_{\tau F}^2} \left(\frac{\partial Z / \partial F'}{\partial^2 Z / \partial F'^2} - \rho_{\tau F}^2 \frac{\partial Z / \partial \tau'}{\partial^2 Z / \partial \tau' \partial F'} \right). \end{aligned} \quad (8.40)$$

Таким образом, в оптимальном дискриминаторе двух коррелированных параметров для получения сигналов рассогласования необходимо сформировать две первые частные производные и три вторые смешанные частные производные огибающей корреляционного интеграла, вычислить по ним коэффициент корреляции, а затем воспользоваться алгоритмом (8.40).

Число операций над входной смесью резко сокращается при использовании модифицированных оптимальных дискриминаторов, не уступающих по точности исходным оптимальным дискриминаторам. Алгоритм работы модифицированных оптимальных

дискриминаторов коррелированных параметров находим делением и умножением формулы (8.40) на огибающую корреляционную интеграла Z :

$$\begin{aligned}\tau' &= \frac{1}{1 - \rho_{\tau F}^2} \left(\frac{Z'_\tau}{Z} \frac{Z}{Z''_{\tau\tau}} - \rho_{\tau F}^2 \frac{Z'_F}{Z} \frac{Z}{Z''_{\tau F}} \right) = \\ &= K_0 \left(-K_\tau \frac{Z'_\tau}{Z} + K_{\tau F} \frac{Z'_F}{Z} \right), \\ F' &= \frac{1}{1 - \rho_{\tau F}^2} \left(\frac{Z'_F}{Z} \frac{Z}{Z''_{FF}} - \rho_{\tau F}^2 \frac{Z'_\tau}{Z} \frac{Z}{Z''_{\tau F}} \right) = \\ &= K_0 \left(-K_F \frac{Z'_F}{Z} + K_{\tau F} \frac{Z'_\tau}{Z} \right),\end{aligned}\tag{8.41}$$

где $K_0 = 1/(1 - \rho_{\tau F}^2)$, $K_\tau = -Z/Z''_{\tau\tau}$, $K_F = -Z/Z''_{FF}$, $K_{\tau F} = -\rho_{\tau F}^2 Z/Z''_{\tau F}$ — коэффициенты, вычисляемые заранее по известной форме сигнала. Знак минус учитывает различие знаков огибающей корреляционного интеграла и ее вторых производных в пределах рабочей зоны характеристики дискриминатора.

В соответствии с алгоритмом (8.41) строится схема модифицированного оптимального совмещенного дискриминатора по времени и частоте (рис. 8.27, а). Она включает в себя рассмотренные ранее временной и частотный модифицированные оптимальные дискриминаторы (ВД и ЧД), формирующие сигналы рассогласования Z'_τ/Z и Z'_F/Z соответственно, где нормировка производится по огибающей Z , формируемой в обнаружителе (в опорном канале). Следовательно, такой дискриминатор отличается от совокупности независимых дискриминаторов только включением устройств перекрестного суммирования их выходных сигналов с относительными весами, определяемыми степенью корреляции.

Включение весового множителя $K_0 = 1/(1 - \rho_{\tau F}^2) \gg 1$ объясняется поведением функции неопределенности ЛЧМ сигнала (рис. 8.27, б). Геометрические места точек, где сигналы рассогласова-

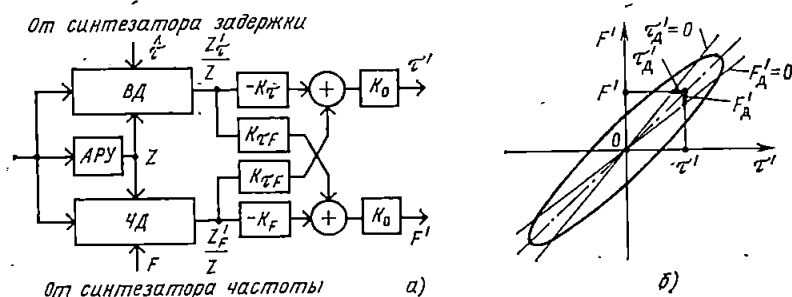


Рис. 8.27. Модифицированный оптимальный совмещенный дискриминатор времени и частоты:

а — функциональная схема; б — формирование сигнала рассогласования

ния временного и частотного дискриминаторов $\tau'_d = Z'_\tau / Z$ и $F'_d = Z'_F / Z$ равны нулю, не совпадают и проходят по разные стороны от оси (хребта) функции неопределенности. Поскольку рассогласования не могут выходить за пределы основного лепестка функции неопределенности, они меньше истинных рассогласований τ' и F' , поэтому их и усиливают в $K_0 \gg 1$ раз.

Значительно слабее проявляются корреляционные связи в совмещенных угловых дискриминаторах, поскольку (как было показано в гл. 6) коэффициент корреляции, вызванный асимметрией двумерного энергетического спектра (в данном случае пространственного), невелик. Вводимая асимметрия спектра окупается упрощением аппаратуры. Наглядным примером этого служит совмещенный суммарно-разностный угловой дискриминатор, амплитудный в азимутальной плоскости и фазовый в угломестной (рис. 8.28, а). Антенна дискриминатора состоит из двух зеркал, фазовые центры которых по вертикали разнесены, диаграммы направленности в вертикальной плоскости совпадают, а в горизонтальной раздвинуты на угол Δ за счет небольшого смещения облучателей относительно вертикальной оси. По объему аппаратуры этот двумерный дискриминатор совпадает с одномерным фазовым дискриминатором (рис. 8.5), только содержит на выходе не один, а два синхронных детектора, опорные напряжения которых сдвинуты на $\pi/2$.

Если аппроксимировать диаграммы направленности обоих зеркал гауссовской кривой с одинаковым углом раствора θ на уровне 0,46, то при смещении угла прихода волны относительно опорного направления на угол α по азимуту и угол β по углу места комплексные амплитуды сравниваемых сигналов принимают вид

$$\begin{aligned} U_1 &= U_0 \exp \left[-2\pi \frac{(\alpha - \Delta/2)^2 + \beta^2}{\theta^2} \right] \exp \left[j\pi \frac{b}{\lambda} \sin \beta \right], \\ U_2 &= U_0 \exp \left[-2\pi \frac{(\alpha + \Delta/2)^2 + \beta^2}{\theta^2} \right] \exp \left[-j\pi \frac{b}{\lambda} \sin \beta \right], \end{aligned} \quad (8.42)$$

т. е. сигналы отличаются по амплитуде и имеют разнос фаз. $\Delta\varphi = 2\pi b \sin \beta / \lambda$ (рис. 8.28, б). При отсутствии рассогласования по

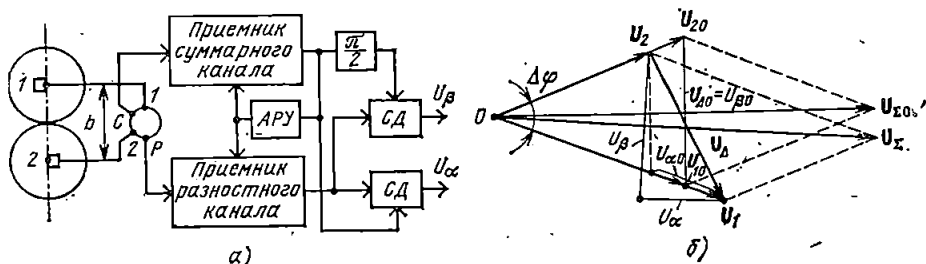


Рис. 8.28. Совмещенный суммарно-разностный угловой дискриминатор, амплитудный по азимуту и фазовый по углу места:

а — функциональная схема; б — формирование сигнала рассогласования

азимуту ($\alpha=0$) сравниваемые сигналы имеют одинаковую амплитуду, а суммарный U_{Σ} и разностный U_{Δ} сигналы взаимно ортогональны (индекс 0 на рис. 8.28,б). Поэтому сигнал рассогласования $U_{\beta 0}$ вырабатывается только на выходе синхронного детектора угла места. Наоборот, при отсутствии рассогласования по углу места ($\beta=0$) сравниваемые сигналы отличаются только по амплитуде и разностный сигнал $U_{\alpha 0}$, по фазе совпадающий или противоположный суммарному, возникает только на выходе синхронного детектора азимута.

В общем же случае ($\alpha \neq 0, \beta \neq 0$) выходные сигналы U_{α} и U_{β} зависят от рассогласований и по азимуту, и по углу места каждый, т. е. взаимно связаны. Однако, как это можно видеть из сравнения сигналов U_{α} и U_{β} при совместном рассогласовании с сигналами $U_{\alpha 0}$ и $U_{\beta 0}$ при рассогласовании лишь по одному из двух углов, влияние взаимной корреляции незначительно, а выигрыш в аппаратуре заметный.

РЕКОМЕНДУЕМЫЙ СПИСОК ЛИТЕРАТУРЫ

- Амиантов И. Н. Избранные вопросы статистической теории связи. — М.: Сов. радио, 1971. — 416 с.
- Астанин Л. Ю., Костылев А. А. Сверхширокополосные радиолокационные измерители. — М.: Воениздат, 1983. — 221 с.
- Бакулев П. А., Степин В. М. Методы и устройства селекции движущихся целей. — М.: Радио и связь, 1986. — 288 с.
- Бендат Дж., Пирсол А. Применение корреляционного и спектрального анализа: Пер. с англ. — М.: Мир, 1983. — 312 с.
- Буренин Н. И. Радиолокационные станции с синтезированной антенной. — М.: Сов. радио, 1972. — 160 с.
- Бычков С. И., Лукьянов Д. П., Бакаляр Г. В. Лазерный гироскоп. — М.: Сов. радио, 1975. — 424 с.
- Василенко Г. И. Теория восстановления сигналов. — М.: Сов. радио, 1979. — 272 с.
- Васильев К. К. Прием сигналов при мультипликативных помехах. — Саратов: Саратов. ун-т, 1983. — 128 с.
- Виницкий А. С. Автономные радиосистемы: Учебное пособие для вузов. — М.: Радио и связь, 1986. — 336 с.
- Вопросы статистической теории радиолокации в 2 т./П. А. Бакут, И. А. Большаков, Б. М. Герасимов и др.; Под ред. Г. П. Тартаковского. — М.: Сов. радио, 1963—1964. Т. 1. — 1963. — 424 с.; Т. 2. — 1964. — 1080 с.
- Вудворд Ф. М. Теория вероятностей и теория информации с применениями в радиолокации: Пер. с англ./Под ред. Г. С. Горелика. — М.: Сов. радио, 1955. — 128 с.
- Голд Б., Рэйдер Ч. Цифровая обработка сигналов: Пер. с англ./Под ред. А. М. Трахтмана. — М.: Сов. радио, 1973. — 368 с.
- Гудмен Дж. Введение в Фурье-оптику: Пер. с англ. — М.: Мир, 1970. — 364 с.
- Драбкин А. Л., Зузенко В. Л., Кислов А. Г. Антенно-фидерные устройства. — 2-е изд., перераб. и доп. — М.: Сов. радио, 1974. — 536 с.
- Журавлев А. К., Лукошкин А. П., Поддубный С. С. Обработка сигналов в адаптивных антенных решетках. — Л.: ЛГУ, 1983. — 239 с.
- Зверев В. А. Радиооптика. — М.: Сов. радио, 1975. — 304 с.
- Ибрагимов И. А., Хаслинский Р. З. Асимптотическая теория оценивания. — М.: Наука, 1979. — 528 с.

- Казаков И. Е. Статистическая теория систем управления в пространстве состояний. — М.: Наука, 1975. — 432 с.
- Канарейкин Д. Б., Потехин В. А., Павлов Н. Ф. Поляризация радиолокационных сигналов/Под ред. В. Е. Дулевича. — М.: Сов. радио, 1966. — 440 с.
- Коронкевич В. П., Соболев В. С., Дубищев Ю. А. Лазерная интерферометрия. — Новосибирск: Наука, 1983. — 213 с.
- Коростелев А. А. Когерентные оптические системы. — Л.: ЛМИ, 1980. — 62 с.
- Коростелев А. А., Мельник Ю. А., Касаткин А. С. Методы измерения координат объектов и обработки радиолокационных сигналов. — М.: Воениздат, 1968. — 241 с.
- Коростелев А. А., Родионов Ю. В. Радиотехнические координаторы. Ч. 4. Пространственно-временная теория оптимального приема радиосигналов. — Л.: ЛМИ, 1979. — 75 с.
- Космические траекторные измерения/Под ред. П. А. Агаджанова, В. Е. Дулевича, А. А. Коростелева. — М.: Сов. радио, 1969. — 504 с.
- Костылев А. А., Степанов М. Г. Смещенные оценки и метод регуляризации в радиотехнических задачах. — М.: Воениздат, 1984. — 83 с.
- Котельников В. А. Теория потенциальной помехоустойчивости. — М.: Госэнергоиздат, 1956.
- Кузьмин С. З. Основы теории цифровой обработки радиолокационной информации. — М.: Сов. радио, 1974. — 432 с.
- Кузьмин С. З. Основы проектирования систем цифровой обработки радиолокационной информации. — М.: Радио и связь, 1986. — 348 с.
- Кук Ч., Бернфельд М. Радиолокационные сигналы: Пер. с англ./Под ред. В. С. Кельзона. — М.: Сов. радио, 1971. — 568 с.
- Лазерные измерительные системы/А. С. Батраков, М. М. Бутусов, Г. П. Гречка и др.; Под ред. Д. П. Лукьянова. — М.: Радио и связь, 1981. — 456 с.
- Лихарев В. А. Цифровые методы и устройства в радиолокации. — М.: Сов. радио, 1973. — 456 с.
- Мельников Б. Г. Статистический синтез радиотехнических измерителей навигационных параметров. — М.: Воениздат, 1982. — 198 с.
- Оводенко А. А. Робастные локационные устройства. — Л.: ЛГУ, 1981. — 182 с.
- Пространственно-временная обработка сигналов/И. Я. Кремер, А. И. Кремер, В. М. Петров и др.; Под ред. И. Я. Кремера. — М.: Радио и связь, 1984. — 224 с.
- Радиолокационные устройства/В. В. Васин, О. В. Власов, В. В. Григорин-Рябов и др.; Под ред. В. В. Григорина-Рябова. — М.: Сов. радио, 1970. — 680 с.
- Радиолокационные станции бокового обзора/А. П. Реутов, Б. А. Михайлов, Г. С. Кондратенков и др.; Под ред. А. П. Реутова. — М.: Сов. радио, 1970. — 360 с.
- Радиотехнические системы/Ю. М. Казаринов, Ю. А. Коломенский, Ю. К. Пестов и др.; Под ред. Ю. М. Казаринова. — М.: Сов. радио, 1968. — 496 с.
- Родс Д. Р. Введение в моноимпульсную радиолокацию: Пер. с англ. — М.: Сов. радио, 1960. — 160 с.
- Современная радиолокация. Пер. с англ./Под ред. Ю. Б. Кобзарева. — М.: Сов. радио, 1969. — 704 с.

Справочник по радиолокации в 4 т./Под ред. М. Сколника: Пер. с англ./ Под ред. К. Н. Трофимова. — М.: Сов. радио, 1976—1979. Т. 1. — 1976. — 456 с.; Т. 2. — 1977. — 408 с.; Т. 3. — 1978. — 528 с.; Т. 4. — 1979. — 376 с.

Стратонович Р. Л. Принципы адаптивного приема. — М.: Сов. радио, 1973. — 144 с.

Теоретические основы радиолокации/А. А. Коростелев, Н. Ф. Клюев, Ю. А. Мельник и др.; Под ред. В. Е. Дулевича. — 2-е изд., перераб. и доп. — М.: Сов. радио, 1978. — 608 с.

Теоретические основы радиолокации/Я. Д. Ширман, В. Н. Голиков, И. Н. Бусыгин; Под ред. Я. Д. Ширмана. — М.: Сов. радио, 1970. — 560 с.

Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. — М.: Наука, 1979. — 288 с.

Тихонов В. И. Статистическая радиотехника. — М.: Радио и связь, 1982. — 624 с.

Тихонов В. И. Нелинейные преобразования случайных процессов. — М.: Радио и связь, 1986. — 296 с.

Фалькович С. Е., Хомяков Э. Н. Статистическая теория измерительных радиосистем. — М.: Радио и связь, 1981. — 288 с.

Царьков Н. М. Многоканальные радиолокационные измерители. — М.: Сов. радио, 1980. — 190 с.

Ширман Я. Д. Разрешение и сжатие сигналов. — М.: Сов. радио, 1974. — 360 с.

Ширман Я. Д., Манжос В. Н. Теория и техника обработки сигналов на фоне помех. — М.: Радио и связь, 1981. — 416 с.

Шкирятов В. В. Радионавигационные системы и устройства. — М.: Радио и связь, 1984. — 160 с.

ОГЛАВЛЕНИЕ

Предисловие	3
Глава 1. ОБЩИЕ СВЕДЕНИЯ ОБ ОБЗОРНО-ИЗМЕРИТЕЛЬНЫХ СИСТЕМАХ	5
1.1. Назначение обзорно-измерительных систем	5
1.2. Классификация радиосистем	6
1.3. Принципы измерения относительных параметров — координат и их первых производных	10
1.4. Методы измерения относительных координат и составляющих вектора скорости объекта	16
1.5. Обобщенная схема и основные характеристики радиосистем	18
1.6. Принципы и методы измерения абсолютных параметров движения	22
Глава 2. ОПТИМАЛЬНЫЙ ПРИЕМ СИГНАЛОВ ИЗВЕСТНОЙ ФОРМЫ	27
2.1. Оценка качества приема сигналов	27
2.2. Оптимальный приемник	28
2.3. Логарифм отношения правдоподобия и корреляционный интеграл временных сигналов	31
2.4. Оптимальный приемник временных сигналов	36
2.5. Примеры согласованных фильтров	41
2.6. Когерентный и некогерентный обнаружители сигналов известной формы	47
2.7. Энергетический потенциал и дальность действия радиосистем	52
Глава 3. ПРОСТРАНСТВЕННО-ВРЕМЕННОЕ ОПИСАНИЕ ПРОЦЕССОВ ПРИ ОПТИМАЛЬНОМ ПРИЕМЕ	54
3.1. Электромагнитная волна как пространственно-временной сигнал	54
3.2. Сигнал как временная входная характеристика радиосистемы	56
3.3. Апертура антенны как пространственная входная характеристика радиосистемы	64
3.4. Принимаемый сигнал как функция времени и пространства	72
3.5. Выходной сигнал оптимального приемника как пространственно-временной корреляционный интеграл	75
Глава 4. ОПТИМАЛЬНЫЙ ПРИЕМНИК КАК СОГЛАСОВАННЫЙ ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ ФИЛЬТР ИЛИ КОРРЕЛЯТОР	80
4.1. Оптимальная структура пространственно-временного приемного устройства	80
4.2. Пространственный тракт согласованного приемника радиосистем с антенными решетками	86
4.3. Подоптимальные структуры пространственного тракта радиосистем	93
Глава 5. ВЗАИМОСВЯЗЬ И ПРЕОБРАЗОВАНИЕ ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ХАРАКТЕРИСТИК В РАДИОСИСТЕМАХ	101
5.1. Физический смысл пространственных частот	101
5.2. Пространственно-временная эквивалентность	105

5.3. Пространственно-временные преобразования и оптическая обработка сигналов	111
5.4. Когерентные оптические процессоры	119
5.5. Голографические методы	128
5.6. Поляризованные сигналы, их спектры и законы прохождения через радиотракт	136
Глава 6. ПОТЕНЦИАЛЬНАЯ ТОЧНОСТЬ И РАЗРЕШАЮЩАЯ СПОСОБНОСТЬ РАДИОСИСТЕМ	145
6.1. Многомерная пространственно-временная функция корреляции и частные функции корреляции	145
6.2. Обобщенная функция корреляции	148
6.3. Потенциальная точность измерений фазы, амплитуды и произвольного неэнергетического параметра сигнала	150
6.4. Оптимальная по точности форма огибающей	157
6.5. Разрешающая способность по одному параметру и выбор формы обобщенной огибающей	164
6.6. Сложная огибающая и оптимальное сжатие сигнала	169
6.7. Неоптимальное сжатие огибающей и уточнение понятия разрешающей способности	173
6.8. Потенциальная точность измерений совокупности неизвестных параметров сигнала	178
6.9. Потенциальная точность совместного измерения угловых координат	184
6.10. Потенциальная точность совместного измерения дальности и радиальной скорости	187
6.11. Совместное разрешение по двум параметрам. Разрешающая способность по направлению	191
6.12. Совместное разрешение по дальности и радиальной скорости. Принцип неопределенности	199
6.13. Обобщенное соотношение неопределенности. Особенности проявления угловой скорости	208
Глава 7. ОПТИМАЛЬНАЯ ЛИНЕЙНАЯ ФИЛЬТРАЦИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ	215
7.1. Фильтрация и экстраполяция случайных процессов	215
7.2. Фильтрация и экстраполяция непрерывных процессов. Интегральное уравнение Винера—Хопфа	219
7.3. Фильтрация и экстраполяция дискретных отсчетов процесса	222
7.4. Расчет параметров фильтра Винера	227
7.5. Дифференциальное уравнение одномерного фильтра Калмана—Бьюси	231
7.6. Расчет параметров одномерного фильтра Калмана—Бьюси	234
7.7. Векторно-матричное дифференциальное уравнение многомерного фильтра Калмана—Бьюси	239
7.8. Вероятностный фильтр	247
Глава 8. НЕЛИНЕЙНАЯ ФИЛЬТРАЦИЯ И ИЗМЕРЕНИЕ ПАРАМЕТРОВ СИГНАЛА	260
8.1. Нелинейная фильтрация параметров сигнала	260
8.2. Следящий и разомкнутый измерители	265
8.3. Оптимальный дискриминатор	270
8.4. Подоптимальные дискриминаторы	275
8.5. Следящие измерители дальности	282
8.6. Следящие измерители скорости	287
8.7. Следящие измерители направления	295
8.8. Взаимодействие следящих измерителей нескольких параметров сигнала	306
Рекомендуемый список литературы	315

ПОПРАВКА

В схемах на рис. 7.14,в и 7.15,в на выходе верхнего интегратора необходимо включить каскад $\psi(t)$.

Учебное пособие

Аркадий Алексеевич Коростелев

ПРОСТРАНСТВЕННО-ВРЕМЕННАЯ ТЕОРИЯ РАДИОСИСТЕМ

Заведующая редакцией *Г. И. Козырева* Редактор *Т. М. Толмачева*
Художественный редактор *Н. С. Шенин* Переплет художника *Ю. В. Архангельского*
Технический редактор *Т. Н. Зыкина* Корректор *Н. М. Давыдова*

ИБ № 623

Сдано в набор 1.08.86	Подписано в печать 8.01.87	Т-23573
Формат 60×90/16	Бумага тип. № 2	Гарнитура литературная
Усл. печ. л. 20,0	Усл. кр.-отт. 20,0	Уч.-изд. л. 21,64
Зак. № 94	Цена 1 р.	Тираж 7500 экз.
Изд. № 20390		

Издательство «Радио и связь». 101000 Москва, Почтамт, а/я 693

Московская типография № 5 ВГО «Союзучетиздат». 101000 Москва, ул. Кирова, д. 40