

Воронежский государственный технический университет

А. А. Зибров, С. В. Ветров, А. И. Климов, В. И. Юдин, Ю. Г.
Пастернак

РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ

Учебное пособие

Воронеж 2002

ББК 32.84

P15
УДК 621.396.4

А. А. Зибров, С. В. Ветров, А. И. Климов, В. И. Юдин, Ю. Г. Пастернак.
Радиотехнические системы. Учебное пособие. Воронеж: Изд-во ВГТУ,
2002. 189 с.

Учебное пособие содержит общие сведения о радиотехнических системах. Рассмотрены основы статистической теории радиотехнических систем, в т.ч. алгоритмы и устройства оптимального обнаружения, различения, разрешения и оценки параметров радиосигналов. Даны основы построения радиолокационных систем, в т.ч. рассмотрены принципы радиолокации, методы и устройства определения местоположения и параметров движения объектов, основы радиотеплолокации и оптической локации, затронуты вопросы радиоэлектронной борьбы и обеспечения помехозащищенности радиотехнических систем. Изложены методы и средства адаптивной обработки сигналов в радиоэлектронных системах различного назначения.

Для студентов специальности 200800 «Проектирование и технология радиоэлектронных средств» технических вузов.

Табл. 2 . Ил. 95. Библиогр.: 21 назв.

Рецензенты: Воронежское НКТБ «Феррит»;
д-р техн. наук, проф. А. П. Ярыгин

© Зибров А. А., Ветров С. В., Климов А. И., Юдин
В. И., Пастернак Ю. Г., 2002

© Оформление. Издательство Воронежского государственного технического университета,

Предисловие

В подготовке инженера-радиоконструктора дисциплина «Основы радиоэлектроники и связи» занимает особое место. В рамках данной дисциплины студентам специальности 200800 «Проектирование и технология радиоэлектронных средств» освещаются вопросы системного подхода к схемотехническому проектированию радиоэлектронной аппаратуры различного назначения.

Учитывая наметившуюся тенденцию интеграции схемотехнического, системного и конструкторско-технологического проектирования радиоэлектронных устройств и систем, во вторую часть лекционного курса дисциплины «Основы радиоэлектроники и связи», охватывающую вопросы системного проектирования радиоэлектронных средств, включены основные положения практически всех общетехнических и специальных дисциплин, предусмотренных Государственным образовательным стандартом РФ по специальности 200700 «Радиотехника».

Учебное пособие состоит из шестнадцати глав. В первой главе содержатся общие сведения о радиотехнических системах. В главах II-V рассмотрены основы статистической теории радиотехнических систем, в т.ч. алгоритмы и устройства оптимального обнаружения, различения, разрешения и оценки параметров радиосигналов. В главах VI-XII даны основы построения радиолокационных систем, в т.ч. рассмотрены принципы радиолокации, методы и устройства определения местоположения и параметров движения объектов, основы радиотеплолокации и оптической локации, а также затронуты вопросы радиоэлектронной борьбы и обеспечения помехозащищенности радиотехнических систем. Главы XIII-XVI посвящены вопросам адаптивной обработки сигналов в радиоэлектронных системах различного назначения.

Учебное пособие является результатом совместной деятельности сотрудников кафедры радиотехнических систем Воронежского института Министерства внутренних дел РФ и кафедры радиоэлектронных устройств и систем Воронежского государственного технического университета; между авторами работа была распределена следующим образом: А. А. Зибров - главы I,III-IV, С. В. Ветров - глава II, А. И. Климов - предисловие и главы VII-XII, В. И. Юдин и Ю. Г. Пастернак – главы XIII- XVI.

Авторы глубоко признательны И. В. Лазареву, И. М. Бушкову и В. А. Козьмину за полезные рекомендации по подготовке рукописи; а также Л. М. Карповой и Е. Б. Куракиной за оказанную помощь в оформлении материалов пособия.

I. ВВЕДЕНИЕ В ТЕОРИЮ РАДИОТЕХНИЧЕСКИХ СИСТЕМ

§ 1.1. Определение и классификация радиотехнических систем

Радиотехнические системы (РТС) - это информационно-управляющие технические системы, осуществляющие извлечение, передачу или разрушение информации с помощью радиоволн (электромагнитных волн, входящих в радиодиапазон).

Отличительный признак РТС - наличие одного или нескольких радиоканалов, состоящих из источников радиоволн, являющихся носителями информации, среды распространения радиоволн и приемника, извлекающего информацию путем соответствующей обработки радиоволн.

Радиоволны, несущие информацию, называют радиосигналом. В РТС сигнал принимается на фоне помех, поэтому он искажается. Приемник, в котором искажения сообщения в определенном смысле минимальны, называют оптимальным. Но, как правило, для снижения стоимости РТС используют квазиоптимальные приемники.

По назначению информации РТС подразделяются на системы передачи, извлечения и разрушения информации, а также системы радиоуправления.

В свою очередь, каждая из этих групп имеет свои разновидности, отличающиеся функциональным назначением системы:

1. СПИ включают в себя системы радиосвязи, телеметрии, передачи команд, сигналов радиовещания и телевидения.

2. Системы извлечения информации включают в себя радиолокационные (РЛС) и радионавигационные (РНС) системы, системы радиоастрономии, радионаблюдения поверхности Земли или других планет, радиоразведки технических средств противника.

3. Системы разрушения информации (радиопротиводействия) включают в себя различные разновидности систем, создающих условия, в которых работа радиосистем противника становится невозможной.

4. Системы радиоуправления включают в себя множество типов систем, управляющих различными объектами с помощью радиосигналов.

Более подробная классификация РЛС и СПИ будет дана в соответствующих разделах данного курса.

§ 1.2. Общая модель радиосистемы

В общем виде структурная схема любой РТС имеет вид, показанный на рис. 1.1.



Рис. 1.1. Структурная схема радиотехнической системы

Отправитель, имеющий информацию от первичного источника в виде значений конкретных физических величин (уровня и высоты звука в радиовещании, интенсивности и цвета элемента изображения в телевидении и т.д.), с помощью преобразователя “сообщение-волна” взаимно однозначно отображает ее в значениях параметров радиоволн (амплитуде, частоте, фазе и т.п.), посылаемых в канал распространения (эфир). Названное преобразование может быть продуктом осознанных действий отправителя, как в СПИ, когда оно выполняется модулятором и передатчиком. Но возможен вариант, когда сообщение “управляет” параметрами радиоволн независимо от воли отправителя, например, в РЛС, где координаты лоцируемой цели автоматически преобразуются во время запаздывания и направления прихода отраженных от нее радиоволн.

Взаимно однозначная связь параметров электромагнитной волны с передаваемым сообщением позволяет на приемной стороне применить обратное преобразование “волна-сообщение”, придав принятой информации конкретную форму, необходимую получателю. Вышеназванный преобразователь обычно состоит из антенной системы, приемника, демодулятора и др.

Кроме радиоволн, несущих полезную информацию, на преобразователь “волна-сообщение” реальной РТС воздействуют и помехи различной природы. В канале распространения передаваемые электромагнитные колебания искажаются помехами: атмосферными (грозовыми разрядами, изменениями физических свойств атмосферы), индустриальными (работой различных электроустановок), межсистемными (создаваемыми другими РТС), преднамеренными (умышленно излучаемыми противником). Кроме того, помехи возникают и на самой приемной стороне, так как процессу преобразования радиоволны в сообщение всегда сопутствуют шумы антенно-фидерного тракта и внутренние шумы приемника.

§ 1.3. Эффективность радиотехнических систем, их тактические и технические характеристики

При проектировании РТС стремятся к получению оптимальных результатов при заданных условиях. Наиболее общим показателем качества РТС является ее эффективность $\Theta = C_{\text{Э}}/C_0$, где $C_{\text{Э}}$ — стоимостное выражение выходного эффекта; C_0 — общие затраты на создание системы.

Величину $C_{\text{Э}}$ трудно определить, поэтому используют целый ряд тактико-технических характеристик (ТТХ) РТС.

К тактическим относят характеристики РТС, определяющие ее возможности при практическом использовании. Наиболее важными из них являются:

1. Зона действия — область пространства, в которой РТС надежно выполняет функции, определенные ее назначением. Так, например, для радиолокационной станции зоной действия является область пространства, в которой объекты с заданными характеристиками отражения и рассеяния радиоволн обнаруживаются с вероятностью не ниже заданной при определенном максимальном значении вероятности ложной тревоги, а также имеется возможность достоверного измерения основных характеристик целей (траектории движения, скорости, габаритных размеров и т.д.).

2. Дальность действия — расстояние, на котором обеспечивается получение заданных показателей РТС. Чаще всего максимальная дальность действия системы зависит от допустимой погрешности при измерении параметров радиосигнала. Иногда зона действия ограничена и со стороны минимальных значений. В этом случае система характеризуется минимальной R_{min} и максимальной R_{max} дальностью действия.

3. Время обзора пространства (поиска цели) — время, необходимое для однократного обзора заданной зоны действия. Данный показатель характерен для РЛС. Выбор времени обзора РЛС связан с маневренностью наблюдаемых или управляемых объектов, объемом пространства обзора, уровнем сигнала и помех, и рядом ТТХ.

4. Количество измеряемых пространственно-временных параметров цели и точность измерения. Данный показатель также характерен для РЛС. Пространственно-временные параметры цели — это, чаще всего, ее координаты и скорость. Точность системы, характеризующаяся погрешностями при измерении параметров цели, определяется применяемым методом измерения, несовершенством аппаратуры, влиянием внешних условий и радиопомех.

5. Разрешающая способность — способность РТС отдельно измерять параметры близко расположенных целей (координат, угла места, скорости и т.д.) или отдельно управлять их движением. Количественно ее принято оценивать минимальной разностью значений измеряемых параметров двух или более сигналов, при которой они еще не воспринимаются системой как один сигнал. Для ряда РЛС разрешающая способность является основной характеристикой. Различают разрешающую способность по дальности, угловым координатам, радиальной и угловой составляющей скорости. В СПИ и радионавигации понятие разрешающей способности часто связывают с возможностью

разделения сигнала, несущего полезную информацию, и различных паразитных сигналов (помех), подобных по форме полезному, но не содержащих полезной информации. Подобный подход применим также к извещателям и системам передачи информации.

6. Пропускная способность (быстродействие) — максимальное количество информации, которое может быть передано или извлечено РТС за единицу времени. По смыслу - это скорость передачи информации. Техническая скорость передачи информации $R=1/T$ [бод] - количество символов за секунду, где T - длительность одного символа. Для РЛС и РНС пропускную способность связывают с числом объектов, обслуживаемых системой одновременно или в единицу времени, или с вероятностью обслуживания заданного числа объектов.

7. Помехозащищенность РТС — ее способность надежно выполнять заданные функции в условиях воздействия непреднамеренных и преднамеренных помех. Помехозащищенность определяется двумя основными понятиями: помехоустойчивость и скрытность работы системы.

Помехоустойчивость — способность РТС противостоять мешающему действию помех. Она количественно оценивается вероятностью ошибки при приеме или отношением сигнала к помехе на входе приемника, при котором вероятность ошибки не превосходит заданной величины.

Скрытность характеризует трудность обнаружения факта работы РТС и измерения основных параметров ее радиосигнала, а значит, и трудность создания специально организованных помех.

8. Надежность — способность РТС сохранять рабочие характеристики в заданных пределах при определенных условиях эксплуатации. Это определение полностью соответствует универсальному определению надежности по ГОСТу 27002-82.

В зависимости от причин, вызывающих отказы, различают следующие разновидности надежности: 1) аппаратную, связанную с состоянием аппаратуры; 2) программную, обусловленную состоянием программ используемых вычислительных устройств; 3) функциональную, т.е. надежность выполнения отдельных функций. В этом смысле помехозащищенность также можно отнести к функциональной надежности РТС.

Кроме перечисленных выше, к тактическим характеристикам РТС можно отнести ее экономические показатели, массу и габариты.

Технические характеристики РТС — параметры, непосредственно определяющие ее тактические характеристики. К ним относятся:

1) метод обзора (поиска) и измерения пространственно-временных параметров объекта (для РЛС);

2) рабочие частоты, стабильность, мощность, вид модуляции, ширина спектра излучаемых колебаний;

3) коэффициент усиления, форма и ширина диаграммы направленности антенны;

4) чувствительность и полоса пропускания приемного устройства;

- 5) вид и параметры устройств отображения и съема информации;
- 6) габариты и масса устройств, составляющих систему, потребляемая ими от источников питания энергия;
- 7) количество каналов передачи информации и их разрядность (для СПИ).

В дальнейшем взаимосвязь тактических и технических характеристик [1, 2, 3] будет рассмотрена для конкретных типов РТС.

Требования к РТС определяются ее назначением. Неоправданное завышение требований приводит к усложнению системы и, как следствие, к снижению экономичности и надежности.

II. ВВЕДЕНИЕ В СТАТИСТИЧЕСКУЮ ТЕОРИЮ РАДИОТЕХНИЧЕСКИХ СИСТЕМ

§ 2.1. Модели радиосигналов и помех в радиосистемах

Физические явления, колебания, процессы, отображающие сообщение (информацию) и используемые для его переноса, называют сигналами. Так как в радиотехнике сообщения передают посредством радиоволн, то есть колебаний электромагнитного поля на частоте, входящей в радиодиапазон, то это поле и можно назвать радиосигналом. Далее будем его называть просто сигналом.

В общем случае сигнал в РТС является пространственно-временным, задаваемым зависимостью $u(t, \vec{r})$, где \vec{r} — радиус вектор рассматриваемой точки пространства. В некоторых источниках принято обозначать сигнал $s(t, \vec{r})$ [1].

Статистическая теория РТС определяет оптимальные (наилучшие в заданном смысле) способы использования пространственных и временных свойств принятого сигнала $u(t, \vec{r})$ с целью получения наиболее достоверного воспроизведения информации на приемной стороне.

В дальнейшем, для упрощения изложения, будем полагать, что структура приемной антенной системы задана заранее. Поэтому зависимостями принимаемого сигнала от пространственных координат \vec{r} будем пренебрегать. В результате наблюдаемый сигнал следует рассматривать как текущую функцию времени $u(t)$.

Сигнал в точке приема является не только функцией времени, но и функцией целого ряда случайных параметров: $a \cdot u(t, \alpha_1, \alpha_2, \dots, \alpha_n)$, где a — полезный параметр; $\alpha_i, i \in [1, n]$ — паразитные параметры. Обозначим $\alpha_1^n = (\alpha_1, \alpha_2, \dots, \alpha_n)$ для краткости записи выражения в скобках. Данное обозначение показывает, что имеется набор $\alpha_1, \alpha_2, \dots, \alpha_n$ из n случайных паразитных параметров α_i , имеющих плотность вероятности $W(\alpha_i)$, где $i=1, \dots, n$.

В задачах РТС используются различные модели сигналов, описанные в курсе РТЦС [4].

Вопросу оптимального выбора сигналов и способов кодирования в статистической теории РТС принадлежит исключительное место, поскольку от правильности решения данной проблемы существенно зависят тактические и технические характеристики разрабатываемых РТС. Данный вопрос оптимизации сигналов неразрывно связан с теорией оптимального приема.

Помехи — это некоторое вредное поле $n(t, \vec{r})$, являющееся вероятностным (случайным) процессом, и некоторым образом взаимодействующее с сигналом $u(t, \vec{r})$. Результатом является поле $x(t, \vec{r}) = f[u(t, \vec{r}); n(t, \vec{r})]$. В силу сказанного выше, будем игнорировать пространственные свойства сигналов и помех. Поэтому результирующее поле, сигнал и помеха будут функциями времени $x(t) = f[u(t); n(t)]$. Сигнал, кроме того, может быть функцией информационных “а” и паразитных α_1^n параметров, поэтому, $x(t) = f[a \cdot u(t, \alpha_1^n); n(t)]$.

Способы взаимодействия сигнала и помехи (вид функции f) делятся на два типа: 1) аддитивный, когда принимаемый сигнал является результатом сложения полезного сигнала и помехи $x(t) = u(t) + n(t)$; 2) мультипликативный, когда помеха перемножается с полезным сигналом $x(t) = B(t) \cdot u(t)$. Мультипликативная помеха проявляет себя только при наличии полезного сигнала, например, при замирании сигнала на коротких волнах. В качестве аддитивной помехи чаще всего выступает собственный шум радиоприемного устройства (РПУ). Он имеет практически равномерный спектр [4, 5]:

$$N(f) = N_0, \quad 0 \leq f \leq F_B; \quad N_0 = \frac{\sigma_{\text{ш}}^2}{F_B}, \quad (2.1)$$

где $\sigma_{\text{ш}}^2$ — дисперсия шума, имеющая физический смысл мощности шума на нагрузке в 1 Ом, а N_0 — односторонняя спектральная плотность мощности шума, измеряемая Вт/Гц.

Полагают, что собственный шум РПУ — нормальный (гауссовский) стационарный случайный процесс с нулевым средним $\bar{n} = 0$, плотность вероятности которого определяется как [4, 5]:

$$W(n) = \frac{1}{\sqrt{2\pi}\sigma_{\text{ш}}} \exp\left\{-\frac{n^2}{2\sigma_{\text{ш}}^2}\right\}. \quad (2.2)$$

Если $F_B \rightarrow \infty$, то получаем “белый шум” - нормальный стационарный дельта-коррелированный процесс, который и является моделью помехи в большинстве рассматриваемых далее задач. “Белый шум” является неосуществимой в реальности абстракцией, т.к. его дисперсия (средняя мощность) бесконечна. Однако реальные РТС имеют ограниченную полосу пропускания, в пределах которой спектр шума оказывается

практически постоянным. В результате неравномерность спектра реального шума вне полосы пропускания РТС почти не будет сказываться на ее выходе. Следовательно, если ширина полосы входного шума много больше ширины полосы пропускания РТС, то возможно применять модель “белого шума”.

В общем случае, случайный процесс описывается плотностью вероятности $W(n_1, n_2, \dots, n_m)$ своих отсчетов n_1, \dots, n_m . В случае ограниченного по полосе белого шума $n(t)$ его можно представить согласно теореме Котельникова отсчетами с интервалом дискретизации по времени $\Delta t = 1/2F_B$. Если длительность процесса T , то число отсчетов $m = T/\Delta t$, а m -мерная плотность вероятности ограниченного по полосе шума $n(t)$ с равномерной N_0 [1, 4, 5]:

$$W(n_1^m) = \left[\frac{1}{\sqrt{2\pi\sigma_{\text{ш}}}} \right]^m \cdot \exp \left\{ -\frac{1}{2\sigma_{\text{ш}}^2} \sum_{i=1}^m n_i^2 \right\}; \quad n_i = n(i \cdot \Delta t), \quad i=1, \dots, m. \quad (2.3)$$

Если воспользоваться следствием из теоремы Котельникова $\int_0^T f^2(t) dt = \frac{1}{2F_B} \sum_{k=1}^m f_k^2$, где $f(t)$ — сигнал, а $f^2(t)$ — мощность сигнала, то перейдя в (2.3) от экспоненты суммы к экспоненте интеграла, получим функционал плотности вероятности процесса $n(t)$:

$$W[n(t)] = k \cdot \exp \left[-\frac{1}{N_0} \int_0^T n^2(t) dt \right], \quad (2.4)$$

где множитель $k = \left[\frac{1}{\sqrt{2\pi\sigma_{\text{ш}}}} \right]^m$ не зависит от реализации белого шума $n(t)$.

Выражение (2.4) показывает, что чем выше мощность реализации $n(t)$, тем меньше вероятность ее появления.

§ 2.2. Прием сигналов как статистическая задача проверки гипотез

В РТС сигнал принимают на фоне помех, поэтому он искажается. Вероятностный характер наблюдаемого колебания $x(t)$ приводит к тому, что любое устройство извлечения информации не застраховано от ошибок, то есть выносит решения с возможными ошибками. Радиоприемник (РПУ), для которого искажения информации в определенном смысле минимальны, называется в этом смысле оптимальным. При выбранном критерии оптимальности и заданных условиях приема указанный приемник должен обеспечивать минимальные искажения сообщения.

Очевидно, что реальное РПУ не может обеспечивать искажений информации меньших оптимального РПУ.

Для типовых РТС можно сформулировать пять основных задач оптимального приема:

1. Обнаружение сигнала. Пусть неизвестен факт наличия или отсутствия полезного сигнала на $u(t)$ входе РПУ. Принимаемое колебание можно описать как $x(t)=a \cdot u(t)+n(t)$; $t \in [0, T]$, где a — полезный параметр, несущий информацию о наличии или отсутствии полезного сигнала, то есть a — случайная величина, принимающая значения:

1) $a=a_0=0$ с вероятностью $P(a_0)$ — полезного сигнала нет;

2) $a=a_1=1$ с вероятностью $P(a_1)$ — полезный сигнал есть; $P(a_0)+P(a_1)=1$.

По результатам наблюдения реализации $x(t)$ на интервале времени $t \in [0, T]$ требуется вынести решение о наличии или отсутствии полезного сигнала. Эта задача характерна для РЛС, извещателей и цифровых СПИ.

2. Различение сигналов. Пусть на входе РПУ может присутствовать один из двух полезных сигналов: $u_1(t, \alpha_{11}^n)$ и $u_2(t, \alpha_{12}^n)$. Тогда принимаемая реализация имеет вид

$x(t)=a \cdot u_1(t, \alpha_{11}^n)+(1-a) \cdot u_2(t, \alpha_{12}^n)+n(t)$; $t \in [0, T]$, где полезный параметр принимает значения $a=a_0=0$ или $a=a_1=1$ и $P(a_0)+P(a_1)=1$. Требуется, наблюдая $x(t)$ на интервале $t \in [0, T]$, определить, какой из передаваемых сигналов содержится в принятом колебании. Данная задача характерна для СПИ, телеметрических систем, телеграфии.

Очевидно, что если один из сигналов нулевой, то задача различения трансформируется в задачу обнаружения. Поэтому в данном пособии некоторые разделы теории различения и обнаружения будут рассматриваться совместно.

3. Оценка параметров сигналов. Пусть какой-либо параметр α_i сигнала $u(t, \alpha_i^n)$ является полезным (информативным) и имеет плотность вероятности $W(\alpha_i)$, так что принимаемая реализация имеет вид $x(t)=u(t, \alpha_i^n)+n(t)$; $t \in [0, T]$. Требуется, проведя анализ реализации $x(t)$ на интервале $t \in [0, T]$, выработать оценку α_i^* измеряемого параметра α_i . Эта задача характерна для РЛС, РНС, телеметрии, СПИ и т.д.

4. Фильтрация сообщений. Пусть параметр α_i сигнала $u(t, \alpha_i^n)$ является случайным процессом, содержащим информацию, т.е. $\alpha_i=\alpha_i(t)$ — полезное сообщение. Требуется, располагая сведениями о помехе, наилучшим образом выделить информацию процесс $\alpha_i(t)$. Задача характерна для следящих систем. Если α_i на интервале наблюдения $t \in [0, T]$ не меняется со временем, то задача фильтрации переходит в задачу оценки параметров.

5. Разрешение сигналов. Пусть на входе РПУ действуют, возможно налагаясь друг на друга, два полезных сигнала $u_1(t, \alpha_{11}^n)$ и $u_2(t, \alpha_{12}^n)$, так что наблюдаемая реализация $x(t)=a_1 \cdot u_1(t, \alpha_{11}^n)+a_2 \cdot u_2(t,$

α_{12}^n) $+n(t)$, где a_1 и a_2 — полезные параметры сигналов u_1 и u_2 , принимающие значения $a_1=\{0, 1\}$, $a_2=\{0, 1\}$ и являющиеся независимыми. Параметры сигналов α_{11}^n и α_{12}^n могут отличаться друг от друга незначительно. Рассматриваемый в данном пособии частный случай задачи разрешения заключается в совместном обнаружении сигналов и измерении их параметров. Данная задача характерна для РЛС и РНС.

В результате решения всех перечисленных выше задач нужно для выбранного критерия найти математическое правило (алгоритм) извлечения информации и определить по нему структурную схему оптимального РПУ.

Качественные показатели данных алгоритмов носят вероятностный характер. Разрабатывая конкретное устройство оптимального приема следует стремиться так выбрать стратегию (алгоритм) его работы, чтобы вредные последствия выносимых им неправильных решений были минимальными. Такая стратегия и построенное на ее основе РПУ называются оптимальными.

Решение задач оптимального приема базируется на методах математической статистики — теории проверки статистических гипотез. Более подробно применение теории проверки гипотез при решении задач оптимального приема рассмотрим на примере обнаружения и различения сигналов, решив задачу различения, т.к. она является наиболее общей и при определенных условиях, указанных выше, переходит в задачу обнаружения.

Основные элементы задачи проверки гипотез видны из следующей модели:



Рис. 2.1. Структурная схема задачи проверки гипотез

Источник сообщения создает выходную величину “а”, принимающую одно из двух значений: a_0 или a_1 , соответствующие гипотезам H_2 и H_1 , где H_2 — гипотеза, что на входе РПУ существует сигнал u_2 , т.е. $a=a_0=0$ с вероятностью $P(a_0)$, а H_1 — гипотеза, что на входе РПУ есть сигнал u_1 , то есть $a=a_1=1$ с вероятностью $P(a_1)$.

Вероятностный механизм перехода отражает действие на полезную величину “а” мешающей случайной величины — помехи. В соответствии с определенным (априори известным) вероятностным законом для каждой из гипотез на выходе этого блока генерируется случайная величина x . Она наблюдается, и по определенному правилу выносятся решение H_2^* или H_1^* в пользу гипотезы H_2 или H_1 соответственно. Это **двухальтернативная** задача.

Задача проверки гипотез сводится к установлению в определенном смысле наилучшего правила принятия решения. Это означает, что пространство наблюдения Y должно быть оптимальным образом разбито на две части.

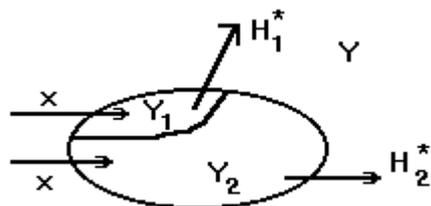


Рис. 2.2. Разбиение пространства наблюдения и принятие решений по результатам наблюдений

Если x попадает в область Y_2 , то выносится решение H_2^* , в противном случае — H_1^* . Оптимальное разбиение пространства Y осуществляется в соответствии с определенными критериями оптимальности.

Рассмотрим задачу проверки двух гипотез. В ней нужно определить: что справедливо — H_2 или H_1 . При этом возможны два исхода, когда решение H_i^* , $i=1, 2$ выносится в пользу гипотезы H_j , $j=1, 2$ правильно, т.е. $i=j$ ($H_2 \rightarrow H_2^*$ или $H_1 \rightarrow H_1^*$), и два исхода, когда решение H_i^* , $i=1, 2$ выносится в пользу гипотезы H_j , $j=1, 2$ неправильно, т.е. $i \neq j$ ($H_1 \rightarrow H_2^*$ или $H_2 \rightarrow H_1^*$). Первые два исхода соответствуют правильному решению двухальтернативной задачи проверки гипотез, а два последних — ошибочному. В соответствии с характером решаемой задачи они могут называться по-разному. Два последних исхода характеризуются вероятностями ошибочного решения $P_{\text{л}}$, а два первых — вероятностями правильного решения $P_{\text{п}}$ и для $i=1, 2$; $j=1, 2$:

$$P(H_j^* / H_i) = P_{\text{л}} = \int_{Y_j} W(x / H_i) dx = 1 - \int_{Y_i} W(x / H_i) dx, \quad i \neq j, \quad (2.5)$$

$$P(H_i^* / H_i) = P_{\text{п}} = \int_{Y_i} W(x / H_i) dx. \quad (2.6)$$

При справедливости каждой из гипотез H_i правильное и ошибочное решения являются противоположными:

$P(H_i^* / H_i) + P(H_j^* / H_i) = 1$, $i=1, 2$, $j=1, 2$, $j \neq i$. При различении двух сигналов $P(H_2^* / H_1)$ и $P(H_1^* / H_2)$ представляют собой вероятности трансформации передаваемых сигналов друг в друга. Как видно из (2.5),

меняя границу между областями, можно получить любое соотношение между $P(H_2^* / H_1)$ и $P(H_1^* / H_2)$.

§ 2.3. Отношение правдоподобия. Функция правдоподобия

Все перечисленные выше задачи оптимального приема могут быть решены на основе **правила отношения правдоподобия**. При дискретной обработке правило принятия решения имеет вид:

$$\Lambda(x) = \frac{W(x/H_1)}{W(x/H_2)} \stackrel{H_1^*}{>} \stackrel{H_2^*}{<} \Lambda_0. \quad (2.7)$$

Здесь $\Lambda(x)$ — отношение правдоподобия, которое показывает насколько, исходя из наблюдаемых данных, правдоподобнее предположение о справедливости одной гипотезы по сравнению с другой. Если в результате наблюдения получен результат x , а $\Lambda(x) > \Lambda_0$, то нужно выносить решение H_1^* в пользу гипотезы H_1 , в противном случае — решение H_2^* .

В выражении (2.7) Λ_0 называется порогом принятия решения и имеет смысл отношения правдоподобия, соответствующего случаю попадания результата наблюдения x на границу между областями Y_1 и Y_2 . Величина Λ_0 зависит от выбранного критерия оптимальности.

В выражении (2.7) числитель и знаменатель — это не плотности вероятности, а отдельные их значения, полученные в результате наблюдения величины x . Если в качестве независимой переменной в числителе и знаменателе считать номер гипотезы или значение полезного параметра “а”, то числитель и знаменатель называются **функциями правдоподобия**. В случае непрерывной обработки в отношении правдоподобия (2.7) функции правдоподобия необходимо заменить **функционалами правдоподобия** $L_1=W(x(t)/H_1)$ и $L_2=W(x(t)/H_2)$. Функция (функционал) правдоподобия показывает степень соответствия результата x гипотезе H_i .

В простейшем случае для детерминированных сигналов вид функции (функционала) правдоподобия определяется статистикой случайного процесса $n(t)$.

На основании правила отношения правдоподобия (2.7) можно построить приемник распознавания гипотез, состоящий из вычислителя $\Lambda(x)$, на вход которого подается наблюдаемый процесс $x(t)$ и схемы сравнения $\Lambda(x)$ с Λ_0 , выносящей решения H_i^* . Структура вычислителя $\Lambda(x)$ зависит от вида отношения правдоподобия (2.7), то есть от характера решаемой задачи и априорно известных плотностей вероятности $W(x/H_1)$ и $W(x/H_2)$. Его можно построить с использованием цифровых или

аналоговых устройств. Мы будем рассматривать аналоговые оптимальные приемники, которые в ряде случаев оказываются гораздо проще.

§ 2.4. Статистические критерии оптимальности при решении задач приема и обработки сигналов

Решение задачи проверки гипотез начинается с выбора и обоснования критерия оптимальности. Он зависит от условий решения задачи и назначения РТС, в которой будет использоваться синтезированное устройство распознавания гипотез.

Из (2.5) видно, что изменение границы, приводящее к увеличению $P(H_2^*/H_1)$, уменьшает $P(H_1^*/H_2)$ и наоборот. Поэтому изменением границы нельзя добиться одновременно уменьшения $P(H_2^*/H_1)$ и $P(H_1^*/H_2)$, хотя это был бы наилучший вариант. В этом случае изменением границы между областями можно минимизировать лишь линейную комбинацию $P(H_2^*/H_1)$ и $P(H_1^*/H_2)$. В соответствии с этим вводят несколько критериев оптимальности:

1) **Критерий минимального среднего риска:**
 $\bar{R} = P(a_0) \cdot P(H_1^*/H_2) \cdot R_2 + P(a_1) \cdot P(H_2^*/H_1) \cdot R_1 = \min$, где R_1 и R_2 — стоимости ошибок (стоимость риска, весовые коэффициенты, веса), а $P(a_0)$ и $P(a_1)$ — априорные вероятности. Например, для систем оповещения и сигнализации, при решении задачи обнаружения сигнала u_1 (гипотеза H_1 — проникновение на объект), в качестве R_1 можно использовать стоимость ценностей, которые могут быть похищены в случае необнаруженного проникновения, а в качестве R_2 — стоимость одного выезда дежурной группы.

2) **Критерий минимальной вероятности полной ошибки (критерий Котельникова, или “идеального наблюдателя”):**
 $\bar{P}_{ош} = P(a_0) \cdot P(H_1^*/H_2) + P(a_1) \cdot P(H_2^*/H_1) = \min$. Этот критерий соответствует задачам, возникающим в СПИ, когда одинаково опасны ошибки обоих видов, поэтому $R_1 = R_2$.

3) **Критерий минимальной “взвешенной ошибки”:**
 $\bar{r} = r_2 \cdot P(H_1^*/H_2) + r_1 \cdot P(H_2^*/H_1) = \min$, где r_1 , r_2 — весовые коэффициенты. Критерий соответствует задачам РЛС, где неизвестны $P(a_0)$, $P(a_1)$.

4) **Критерий Неймана-Пирсона:** $P(H_1^*/H_2) = \text{const}$, $P(H_2^*/H_1) = \min$. Данный критерий, например, используется в РЛС. В этом случае решается задача обнаружения полезного сигнала u_1 , сигнал u_2 является нулевым, $P(H_1^*/H_2)$ называется вероятностью ложной тревоги, а $P(H_2^*/H_1)$ — вероятностью пропуска цели.

В третьем критерии нужно знать r_1 и r_2 , а в четвертом — только $P(H_1^*/H_2)$, поэтому чаще используют критерий Неймана-Пирсона.

Наиболее общим является критерий минимальной “взвешенной ошибки”. От него возможен переход к критерию минимального среднего риска, если взять $r_{л}=P(a_0) \cdot R_2$ и $r_{п}=P(a_1) \cdot R_1$.

Любой критерий приводит к алгоритму принятия решения (2.7), причем различными будут только значения порога Λ_0 . Для этого необходимо привести критерии к виду $P(H_1^* / H_2) + \Lambda_0 \cdot P(H_2^* / H_1) = \min$, подставить выражения (2.5), перейти от суммы интегралов к интегралу суммы и выбрать область интегрирования, обеспечивающую необходимое значение данного интеграла.

Таким образом, можно сделать вывод, что приемники распознавания гипотез, оптимальные по любому из рассмотренных выше критериев, должны выполнять одни и те же действия: вычислять отношение правдоподобия (2.7) и сравнивать его с порогом Λ_0 . От конкретного критерия зависит лишь значение порога Λ_0 и поэтому приемник, наилучший по одному критерию трансформируется в оптимальный по другому критерию простым изменением порога Λ_0 .

III. ОБНАРУЖЕНИЕ И РАЗЛИЧЕНИЕ СИГНАЛОВ

§ 3.1. Алгоритмы оптимального обнаружения детерминированных сигналов

Формулировка задачи приведена в § 2.2. Единственным неизвестным параметром сигнала $u(t)$ является параметр “а”, поэтому существуют две гипотезы — “сигнал не обнаружен” $H_0: a=a_0=0, P(a_0)$ и “сигнал обнаружен” $H_1: a=a_1=1, P(a_1)$. Требуется определить структуру оптимального обнаружителя сигнала $u(t)$. Решение задачи осуществляется на основе правила отношения правдоподобия (2.7).

Представим принятое колебание $x(t)$ на интервале $t \in [0; T]$ выборками в соответствии с теоремой Котельникова. Тогда при отсутствии полезного сигнала в $x(t)$, т.е. когда $u(t)=0$ и $x(t)=n(t)$, выборки x_k этого колебания $x(t)$ будут представлять собой выборки шума $n(t)$: $x_1^m = n_1^m$. Так как n_1^m — нормальный процесс, то и x_1^m также нормальный случайный процесс. Тогда в случае $a=a_0=0$ согласно (2.3) m -мерная плотность вероятности процесса $x(t)$:

$$W(x_1^m) = \left[\frac{1}{\sqrt{2\pi\sigma_{ш}}} \right]^m \cdot \exp \left\{ -\frac{1}{2\sigma_{ш}^2} \sum_{k=1}^m x_k^2 \right\}, \quad k=1, \dots, m. \quad (3.1)$$

Аналогично, в случае $a=a_1=1$ выборки x_k колебания $x(t)$ представляют собой выборки смеси сигнала и шума, т.е. $x_k = u_k + n_k$ и

$$W(x_1^m) = \left[\frac{1}{\sqrt{2\pi\sigma_{\text{ш}}}} \right]^m \cdot \exp \left\{ -\frac{1}{2\sigma_{\text{ш}}^2} \sum_{k=1}^m (x_k - u_k)^2 \right\}, \quad k=1, \dots, m. \quad (3.2)$$

Отсчеты u_k здесь играют роль ненулевого математического ожидания случайного процесса x_k , т.к. при сложении независимых случайных величин их математические ожидания складываются.

Подставив в полученные выражения результаты наблюдений x_1^m , получим функции правдоподобия:

$$L_{a0}(x) = \left[\frac{1}{\sqrt{2\pi\sigma_{\text{ш}}}} \right]^m \cdot \exp \left\{ -\frac{1}{2\sigma_{\text{ш}}^2} \sum_{k=1}^m x_k^2 \right\}; \quad k=1, \dots, m. \quad (3.3)$$

$$L_{a1}(x) = \left[\frac{1}{\sqrt{2\pi\sigma_{\text{ш}}}} \right]^m \cdot \exp \left\{ -\frac{1}{2\sigma_{\text{ш}}^2} \sum_{k=1}^m (x_k - u_k)^2 \right\}. \quad (3.4)$$

Воспользовавшись так же, как и в § 2.1, следствием из теоремы Котельникова (из (3.3) и (3.4)), найдем функционалы правдоподобия:

$$L_{a0} = k \cdot \exp \left\{ -\frac{1}{N_0} \cdot \int_0^T x^2(t) dt \right\}; \quad (3.5)$$

$$L_{a1} = k \cdot \exp \left\{ -\frac{1}{N_0} \cdot \int_0^T [x(t) - u(t)]^2 dt \right\}; \quad (3.6)$$

Тогда отношение правдоподобия имеет вид:

$$\Lambda(x) = \frac{L_{a1}}{L_{a0}} = \exp \left\{ -\frac{1}{N_0} \cdot \int_0^T \left\{ [x(t) - u(t)]^2 - x^2(t) \right\} dt \right\}.$$

Преобразовав данное выражение, получаем:

$$\Lambda(x) = \exp \left\{ -\frac{E}{N_0} + \frac{2Z(T)}{N_0} \right\}, \quad (3.7)$$

$$E = \int_0^T u^2(t) dt, \quad Z(T) = \int_0^T x(t)u(t) dt, \quad (3.8)$$

где E — энергия полезного сигнала, а $Z(T)$ — корреляционный интеграл. По смыслу $Z(T)$ — взаимнокорреляционная функция (ВКФ) процессов $x(t)$ и $u(t)$, показывающая их взаимную энергию. Ее можно рассматривать и как свертку двух функций.

Таким образом, оптимальное правило решения при обнаружении детерминированного сигнала $u(t)$ имеет вид:

$$\exp\left\{-\frac{E}{N_0} + \frac{2Z(T)}{N_0}\right\} \underset{H_0^*}{\overset{H_1^*}{>}} \Lambda_0. \quad (3.9)$$

Непосредственная оценка значения левой части выражения (3.9) в приемном устройстве является довольно сложной технической задачей. Упростим это выражение, прологарифмировав обе его части:

$$Z(T) \underset{H_0^*}{\overset{H_1^*}{>}} Z_0, \quad (3.10)$$

где Z_0 — порог, определяемый по формуле:

$$Z_0 = \frac{N_0}{2} \left[\ln \Lambda_0 + \frac{E}{N_0} \right]. \quad (3.11)$$

Если функция $Z(T)$ описывает напряжение на входе порогового устройства в момент времени $t=T$, то Z_0 — порог по напряжению.

§ 3.2. Структуры оптимальных приемников обнаружения детерминированных сигналов

Алгоритм оптимального обнаружения детерминированного сигнала (3.10) и формулы (3.8) и (3.11) служат основой для построения приемников обнаружения. В зависимости от того, как рассматривать $Z(T)$ — как ВКФ или как свертку двух функций — возможны два варианта схем приема: 1) корреляционный приемник; 2) приемник с согласованным фильтром (фильтровый приемник). Рассмотрим их подробнее.

Корреляционный приемник. Из формул (3.8), (3.10) видно, как надо обработать входное воздействие $x(t)$. Поэтому структурная схема данного приемника имеет следующий вид (см. рис. 3.1)

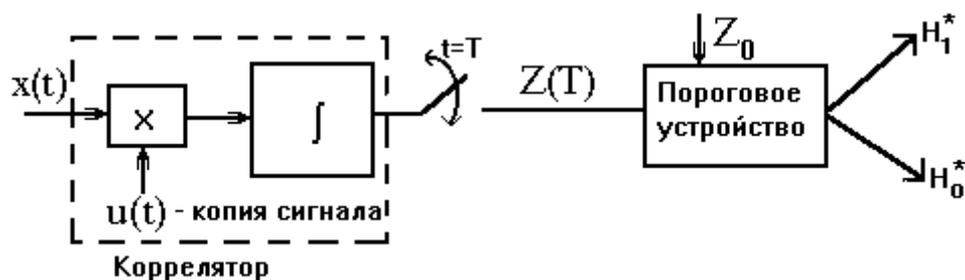


Рис. 3.1. Структурная схема корреляционного приемника обнаружения детерминированного сигнала

Выходное напряжение интегратора $Z(t)$ в момент времени $t=T$ при помощи ключа подается на пороговое устройство (ПУ), где сравнивается с порогом Z_0 . По итогам сравнения результата интегрирования $Z(T)$ с порогом Z_0 возможны два исхода:

- 1) если $Z(T) > Z_0$, то выносится решение H_1^* (сигнал обнаружен);
- 2) если $Z(T) < Z_0$, то выносится решение H_0^* (сигнал не обнаружен).

Последовательно включенные перемножитель и интегратор будем называть коррелятором, который в данном случае и является устройством вычисления $\Lambda(x)$. Принцип работы коррелятора, его реакция на различные полезные сигналы и предельное отношение сигнал/шум на выходе q^2 известны из курса РТЦС [4]. Напомним только, что если шума нет $n(t)=0$ и $x(t)=u(t)$, то напряжение на выходе интегратора $Z(t)$ в момент времени $t=T$ равно энергии E сигнала $u(t)$ (3.8). Если $n(t) \neq 0$, то в $Z(t)$ появится случайная составляющая.

В результате воздействия помех большой интенсивности возможны ошибки: при наличии полезного сигнала на входе приемника $Z(T)$ может оказаться меньше Z_0 , т.е. будет пропуск сигнала и, наоборот, при отсутствии полезного сигнала на входе $Z(T)$ может оказаться больше Z_0 , т.е. будет ложная тревога.

Рассмотрим второй вариант устройства оптимальной обработки сигналов — приемник с **согласованным фильтром**. Согласованным называют линейный фильтр, максимизирующий отношение сигнал/шум на выходе. Принцип работы согласованного фильтра (СФ), его реакция на различные полезные сигналы и q^2 известны из курса РТЦС. Напомним, что импульсная характеристика СФ с точностью до константы представляет собой зеркальное отображение полезного сигнала относительно линии $t=t_0/2$.

$$h(t) = c \cdot u(t_0 - t), \quad c = \text{const}, \quad (3.12)$$

где t_0 — полное время задержки фильтра.

Амплитудно-частотная характеристика СФ $|K_0(\omega)|$ с точностью до константы совпадает по форме с огибающей спектра сигнала $g(\omega)$, а его комплексный коэффициент передачи определяется по формуле:

$$K_0(\omega) = c \cdot e^{-j\omega t_0} \cdot g(\omega). \quad (3.13)$$

Предельные отношения сигнал/шум на выходе СФ и коррелятора одинаковы:

$$q^2 = \frac{2E}{N_0}. \quad (3.14)$$

В момент времени $t=t_0=T$ на выходе СФ напряжение $Z(T)$ и отношение сигнал/шум такие же, что и на выходе коррелятора. Поэтому в схеме любого приемника коррелятор можно заменить на СФ. При этом следует учитывать ряд дополнительных условий, о которых будет сказано ниже.

Таким образом, структурная схема приемника обнаружения с СФ для полностью известного сигнала имеет следующий вид (рис. 3.2):

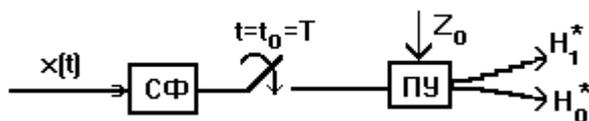


Рис. 3.2. Структурная схема фильтрового приемника обнаружения детерминированного сигнала

Процедура вынесения решения о справедливости одной из гипотез в фильтровом приемнике такая же, как и в корреляционном. Как видно из сравнения рис. 3.1 и 3.2 коррелятор и СФ в данной ситуации оказались взаимозаменяемыми, однако у каждого из этих устройств имеются особенности, определяющие предпочтительность использования одного из них в конкретных условиях. О некоторых из них, известных из курса РТЦС напомним ниже, а другие сформулируем в главе, посвященной разрешению сигналов.

И коррелятор и СФ неинвариантны к начальной фазе сигнала β . В случае неизвестной β нужно поставить после СФ амплитудный детектор (АД), который выделит огибающую $u_{\text{вых}}(t)$, позволит выбрать ее максимум при $t=T$ и сравнить его с Z_0 . Однако при этом уменьшается отношение сигнал/шум из-за нелинейности детектора.

Если входной сигнал СФ изменит временное положение, то и выходной сигнал изменит временное положение. Поэтому СФ является инвариантным ко времени действия сигнала. Коррелятор же неинвариантен ко времени прихода сигнала. При неизвестном времени прихода сигнала это позволяет делать фильтровые приемники одноканальными, а при корреляционном приеме необходимо иметь несколько каналов, каждый из которых согласован с определенным временным положением сигнала.

К существенному недостатку СФ можно отнести зависимость его структуры от полезного сигнала. При смене полезного сигнала необходимо заменить и СФ, т.к. его $h(t)$ и $K_0(\omega)$ зависят от сигнала и его спектра. У коррелятора же при смене полезного сигнала достаточно только изменить его копию, подаваемую на перемножитель.

Иногда бывает трудно спроектировать СФ, $h(t)$ которого имеет необходимую форму (3.12). Поэтому часто вместо СФ применяют квазисогласованные фильтры. Напомним, что квазисогласованным называется фильтр, форма АЧХ которого задана заранее, а максимальное

отношение сигнал/шум на выходе достигается подбором полосы пропускания.

В § 2.2 были перечислены вероятностные характеристики различных исходов двухальтернативной задачи проверки гипотез, частным случаем которой является рассматриваемая задача обнаружения. Две из этих вероятностных характеристик являются основными качественными показателями оптимального приемника обнаружения: вероятность обнаружения $P_0 = P(H_1^* / H_1)$ и вероятность ложной тревоги $P_{\text{л}} = P(H_1^* / H_0)$.

Вероятность обнаружения P_0 — отношение числа обнаружений полезного сигнала, поступившего на вход приемника к общему числу поступлений этого сигнала: $P_0 = \lim_{M \rightarrow \infty} (M_{\text{пр}} / M_0)$, где M_0 — общее число сравнений $Z(T)$ с порогом Z_0 , $M_{\text{пр}}$ — количество превышений $Z(T)$ порога Z_0 при наличии полезного сигнала на входе приемника.

Вероятность ложной тревоги $P_{\text{л}}$ — отношение числа обнаружений полезного сигнала, когда данный сигнал на вход приемника не поступал, к общему числу случаев отсутствия полезного сигнала на входе: $P_{\text{л}} = \lim_{M_0} (M_{\text{л}} / M_0)$, где $M_{\text{л}}$ — количество превышений $Z(T)$ порога Z_0 при отсутствии полезного сигнала на входе приемника.

Основными характеристиками при обнаружении сигнала являются: рабочая характеристика (РХ) и характеристика обнаружения (ХО). Рабочая характеристика — это зависимость P_0 от $P_{\text{л}}$ при $q = \text{const}$. Семейство РХ при различных значениях q приведено на рис. 3.3а. Если полезного сигнала нет, то $q=0$ и РХ имеет вид прямой. При увеличении мощности сигнала (т.е. при возрастании q) P_0 , при фиксированной $P_{\text{л}}$, тоже увеличивается. Характеристика обнаружения — это зависимость P_0 от q при фиксированной вероятности ложной тревоги $P_{\text{л}} = \text{const}$. Семейство ХО при различных значениях $P_{\text{л}}$ приведено на рис. 3.3б. Видно, что с уменьшением порогового напряжения Z_0 (т.е. с увеличением $P_{\text{л}}$) P_0 тоже возрастает.

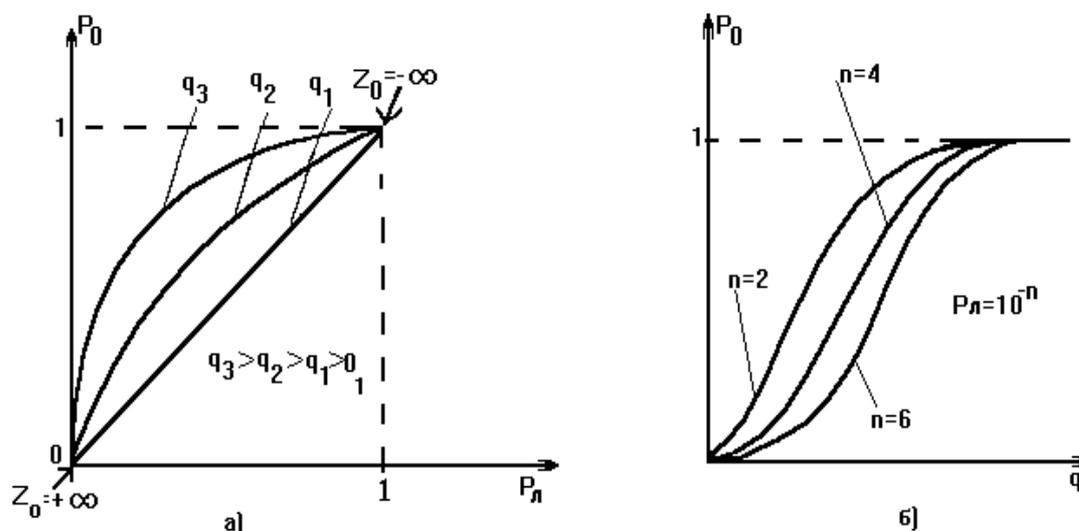


Рис. 3.3 Семейство рабочих характеристик (а) и

характеристик обнаружения (б)

Зная распределение сигналов и помех на входе порогового устройства приемника, ХО и РХ можно построить аналитически. Этот вопрос подробно рассмотрен в [1]. Вероятности P_0 и P_1 для случая детерминированного сигнала описываются при помощи табличного интеграла вероятности $\Phi(x)$ следующим образом:

$$P_1 = \frac{1}{2} - \Phi(\eta_0); \quad P_0 = \frac{1}{2} - \Phi(\eta_0 - q), \quad q^2 = \frac{2E}{N_0}, \quad (3.15)$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt. \quad (3.16)$$

Зная выражения для P_0 и P_1 , можно производить расчет обнаружителя в соответствии с принятым критерием оптимальности. Так, при использовании критерия Неймана-Пирсона, когда известны P_0 и P_1 , которые должен обеспечить обнаружитель, можно определить минимальное отношение сигнал/шум $q_{m\ i}^2$ и соответствующую ему минимальную энергию полезного сигнала E_{min} еще достаточные для обеспечения P_0 и P_1 . Далее по E_{min} можно определить и соответствующий ей порог обнаружения Z_0 для заданной N_0 (3.11).

§ 3.3. Оптимальное обнаружение сигналов со случайной начальной фазой

Полезный сигнал со случайной начальной фазой β имеет вид $u(t, \alpha_1^n) = u(t, \beta)$, $t \in [0, T]$, где β — случайная величина, имеющая, как правило, плотность вероятности $W(\beta) = 1/2\pi$; $\beta \in [-\pi; \pi]$. Входное воздействие приемника имеет вид $x(t) = a \cdot u(t, \beta) + n(t)$; $t \in [0, T]$, где $a \in \{a_0=0; a_1=1\}$. Требуется определить структуру оптимального обнаружителя, что делается с помощью правила отношения правдоподобия (2.7), где $\Lambda(x)$ описывается выражением (3.7). В нашем случае «примешивается» случайная величина β . Например, для прямоугольного радиоимпульса со случайной β :

$$Z(T) = \int_0^T x(t) \cdot u(t, \beta) dt = \int_0^T x(t) \cdot A(t) \cos(\omega_0 t + \varphi(t) + \beta) dt = Z(T, \beta),$$

где $Z(T, \beta)$ — случайная величина. Тогда отношение правдоподобия $\Lambda(x)$ тоже становится случайной величиной:

$$\Lambda(x) = e^{-\frac{E}{N_0} + \frac{2Z(T, \beta)}{N_0}} \equiv \Lambda(x, \beta). \quad (3.17)$$

Чтобы исключить случайную величину β нужно усреднить $\Lambda(x, \beta)$ по всем значениям β , то есть найти его математическое ожидание $M\{\Lambda(x, \beta)\} = \overline{\Lambda(x, \beta)}$. Но сначала преобразуем $Z(T, \beta)$:

$$Z(T, \beta) = \cos \beta \int_0^T x(t) A(t) \cos[\omega_0 t + \varphi(t)] dt - \sin \beta \int_0^T x(t) A(t) \sin[\omega_0 t + \varphi(t)] dt.$$

Таким образом, на входе порогового устройства должно формироваться напряжение:

$$Z(T, \beta) = Z_1 \cos \beta - Z_2 \sin \beta, \quad (3.18)$$

$$Z_1 = \int_0^T x(t) A(t) \cos(\omega_0 t + \varphi(t)) dt, \quad Z_2 = \int_0^T x(t) A(t) \sin(\omega_0 t + \varphi(t)) dt. \quad (3.19)$$

Введем огибающую корреляционного интеграла

$$Z = \sqrt{Z_1^2 + Z_2^2}. \quad (3.20)$$

Тогда из (3.18) можно представить $Z(T, \beta) = Z[(Z_1/Z) \cos \beta - (Z_2/Z) \sin \beta]$. Если поделить (3.20) на Z и возвести в квадрат, то $(Z_1/Z)^2 + (Z_2/Z)^2 = 1$, причем видно, что $|Z_1/Z| \geq 0$ и $|Z_2/Z| \geq 0$. Поэтому огибающие $|Z_1/Z|$ и $|Z_2/Z|$ можно рассматривать как синус и косинус произвольного угла θ : $Z_1/Z = \cos \theta$, $Z_2/Z = \sin \theta$.

Отсюда $Z(T, \beta) = Z[\cos \beta \cos \theta - \sin \beta \sin \theta] = Z \cdot \cos(\beta + \theta)$.

Подставив это выражение в (3.17), получим отношение правдоподобия

$$\Lambda(x, \beta) = \exp \left\{ -\frac{E}{N_0} + \frac{2Z \cos(\beta + \theta)}{N_0} \right\}. \quad (3.21)$$

Среднее от $\Lambda(x, \beta)$ по всем значениям β находим как:

$$\begin{aligned} \overline{\Lambda(x, \beta)} &= \int_{-\pi}^{\pi} \Lambda(x, \beta) W(\beta) d\beta = \int_{-\pi}^{\pi} \frac{1}{2\pi} \exp \left\{ -\frac{E}{N_0} \right\} \exp \left\{ -\frac{2}{N_0} Z \cos(\beta + \theta) \right\} d\beta = \\ &= \exp \left\{ -\frac{E}{N_0} \right\} \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp \left\{ \frac{2}{N_0} Z \cos(\beta + \theta) \right\} d\beta. \end{aligned}$$

Используя модифицированную функцию Бесселя первого рода нулевого порядка, $I_0(x)$, получим

$$\bar{\Lambda}(x, \beta) = \exp\left\{-\frac{E}{N_0}\right\} \cdot I_0\left(\frac{2Z}{N_0}\right), \quad (3.22)$$

$$I_0(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\{x \cdot \cos(\theta + \theta_0)\} d\theta = \sum_{k=0}^{\infty} \frac{(x/2)^{2k}}{(k!)^2}. \quad (3.23)$$

Таким образом, оптимальное правило решения имеет вид:

$$\bar{\Lambda}(x, \beta) \underset{H_0^*}{\overset{H_1^*}{>}} \Lambda_0. \quad (3.24)$$

Так как $I_0(x)$ — функция монотонная, то можно найти обратную ей однозначную функцию и использовать правило:

$$Z \underset{H_0^*}{\overset{H_1^*}{>}} Z_0, \quad (3.25)$$

где Z описывается выражениями (3.19) и (3.20).

Данный алгоритм служит основой для построения приемников обнаружения сигналов со случайной начальной фазой. Как и в случае с полностью известным сигналом, здесь возможны две схемы приема: корреляционная и фильтровая. Очевидно, что из правила (3.25) и (3.20), (3.19) следует структурная схема корреляционного приемника, приведенная на рис. 3.4, включающая в себя два квадратурных канала.

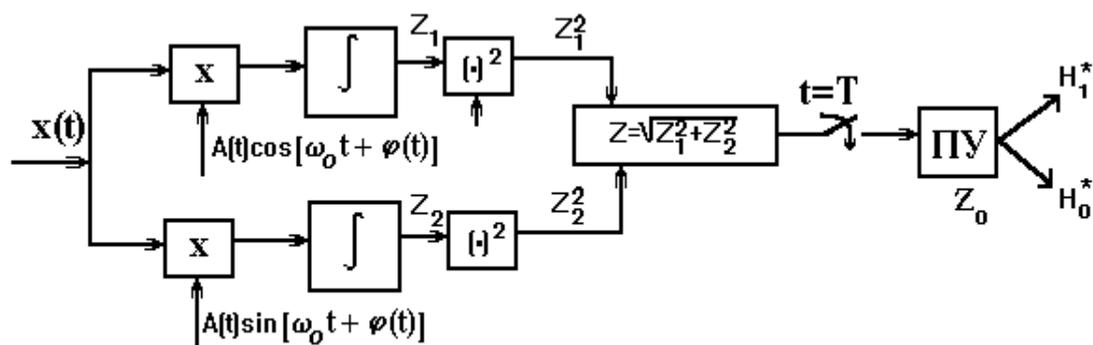


Рис.

3.4. Структурная схема оптимального корреляционного обнаружителя сигналов со случайной начальной фазой

Наличие двух квадратурных каналов приема позволяет исключить пропуски сигнала, обусловленные случайным характером его начальной фазы. Если при изменении начальной фазы отклик одного из каналов увеличивается, то отклик второго уменьшается, поэтому Z остается

постоянной и от β не зависит. Наличие двух каналов приводит к ухудшению качества приема по сравнению с приемником для полностью известного сигнала, так как в силу независимости шумов в квадратурных каналах они суммируются, уменьшая отношение сигнал/шум на входе порогового устройства.

На основе правила (3.25) можно построить и фильтровый приемник обнаружения сигнала с неизвестной β . Фильтр можно строго согласовать только для одного значения начальной фазы β , прежде всего, когда $\beta=0$. Так как неинвариантность СФ к β устраняется при помощи амплитудного детектора (АД), то приемник будет состоять из СФ, АД и порогового устройства, т.е. будет гораздо проще. Инвариантность по отношению к β достигается ценой некоторого ухудшения качественных показателей приемника, что обуславливается влиянием амплитудного детектора, так как он является нелинейным устройством и ухудшает отношение сигнал/шум на выходе приемника. ХО при неизвестной β похожа на ХО детерминированного сигнала, но незнание β при $P_d = \text{const}$ и $q^2 = \text{const}$ будет приводить к снижению P_0 .

В [4] показано, что похожее на (3.25) правило обнаружения применяется и при обнаружении сигнала со случайными амплитудой B и начальной фазой β : $u(t, \alpha_1^n) = u(t, \beta, B) = B \cdot A(t) \cos(\omega_0 t + \varphi(t) + \beta)$, где β и B — случайные величины.

В данном правиле (3.25) изменится только значение порога обнаружения Z_0 , причем от закона распределения амплитуды будет зависеть только величина Z_0 . Поэтому структурная схема обнаружителя сигналов со случайными β и B будет такой же, как и в случае со случайной β и фиксированной амплитудой, изменится только значение порога Z_0 .

ХО сигналов со случайными β и B будет похожа на ХО при случайной β , но флуктуация амплитуды, при больших q^2 будет приводить к существенному ухудшению P_0 , а при малых q^2 — к увеличению P_0 . Однако область малых q^2 практического значения не имеет из-за неприемлемо низких значений P_0 . Различие между ХО сигнала со случайной β и ХО сигнала со случайными амплитудой и β тем существеннее, чем больше дисперсия амплитуды.

§ 3.4. Оптимальное обнаружение пачек радиоимпульсов

В радиолокации часто применяют сигналы, представляющие собой пачку импульсов. Ее можно рассматривать как одиночный импульсный сигнал сложной формы. Поэтому при обнаружении пачки импульсов можно воспользоваться результатами, полученными выше. Напомним, что пачка называется когерентной, если начальная фаза первого импульса может быть любой, а изменение фазы от импульса к импульсу является закономерным. Если фаза первого импульса неизвестна, то когерентную пачку далее будем называть случайной. Если начальная фаза первого импульса известна, то когерентная пачка полностью известна. В

моноимпульсной РЛС принимаемый сигнал имеет вид пачки радиоимпульсов, если время облучения цели равно нескольким периодам излучения импульсов T_n . В СПИ примером пачки радиоимпульсов является фазоманипулированный (ФМН) сигнал, его период повторения импульсов T_n равен длительности одного импульса τ_c .

Полностью известную когерентную пачку можно рассматривать как модификацию детерминированного сигнала, а случайную пачку — как модификацию сигнала со случайной начальной фазой. Поэтому для полностью известной пачки справедливы все выводы §§ 3.1 и 3.2, а для случайной когерентной пачки — § 3.3. При этом энергия полезного сигнала определяется как суммарная энергия всех ее импульсов, входящих

в пачку, $E = \sum_{i=1}^N E_i$.

Как и ранее, возможны две схемы приема: корреляционная и фильтровая. Структурная схема корреляционных приемников для обработки когерентных пачек имеют такой же вид, что и для одиночных сигналов (рис. 3.1 — для полностью известной пачки и рис. 3.4 — для случайной пачки), только копии должны иметь вид ожидаемой пачки импульсов. Структура схемы СФ зависит от вида сигнала и для пачки существенно отличается от структуры СФ для одиночных сигналов. Все вопросы, связанные с СФ для пачек, подробно изложены в курсе РТЦС. Напомним только, что он состоит из фильтра, согласованного с одиночным импульсом, и накопителя. В качестве последнего часто используется схема, состоящая из линий задержки, фазовращателя и сумматора [4]. Существуют и другие схемы реализации накопителей [1].

Кроме рассмотренных выше корреляционного и фильтрового приема возможно их совместное использование — **корреляционно-фильтровая обработка когерентных пачек**. Напряжение на входе порогового устройства в момент времени $t=T$ описывается выражением (3.8), где $u(t)$ — пачка радиоимпульсов. Представим $u(t)=u_1(t) \cdot u_2(t)$, где $u_1(t)$ — последовательность видеоимпульсов, $u_2(t)$ — радиоимпульс длительностью T (рис. 3.5). Тогда

$$Z(T) = \int_0^T [x(t) \cdot u_1(t)] \cdot u_2(t) dt.$$

И $Z(T)$ можно рассматривать как свертку двух функций: $x'(t)=x(t) \cdot u_1(t)$ и $u_2(t)$.

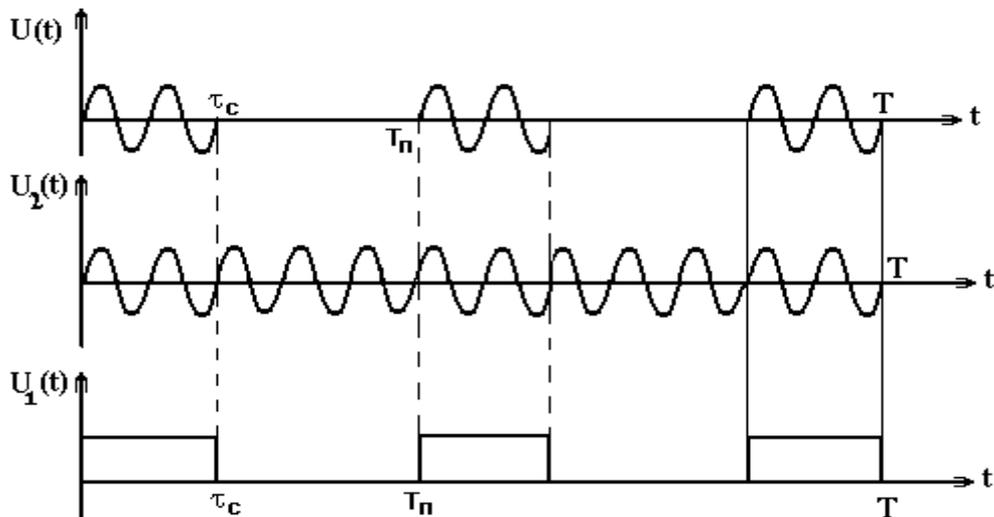


Рис. 3.5. Представление когерентной пачки радиоимпульсов $u(t)$ в виде произведения сигналов $u_1(t)$ и $u_2(t)$

Импульсную характеристику вида $u_2(t)$ имеет фильтр, согласованный с радиоимпульсом длительностью T (см. рис. 3.5), то есть возможен еще один вариант оптимальной обработки пачек — корреляционно-фильтровой. Преимущество этого метода по сравнению с СФ состоит в том, что отсутствует накопитель, имеющий довольно сложную структуру, а по сравнению с коррелятором — в том, что видеосигнал $u_1(t)$ проще формировать, чем радиосигнал $u(t)$. Структурная схема такого корреляционно-фильтрового приемника обнаружения приведена на рис. 3.6.

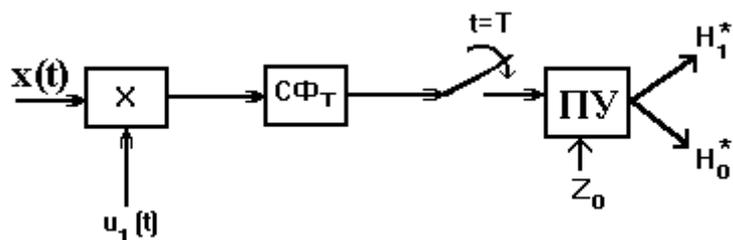


Рис. 3.6. Структурная схема корреляционно-фильтрового приемника обнаружения когерентной пачки радиоимпульсов

Рассмотрим две важные схемы приемников обнаружения пачек импульсов — схемы с когерентным и некогерентным накоплением радиоимпульсов. **Когерентное (додетекторное) накопление** (рис. 3.7) используется, если пачка имеет закономерную структуру, а **некогерентное (последетекторное) накопление** (рис. 3.8) — если пачка некогерентная. СФ в первом случае состоит из фильтра $СФ_1$, согласованного с одиночным импульсом, и накопителя.

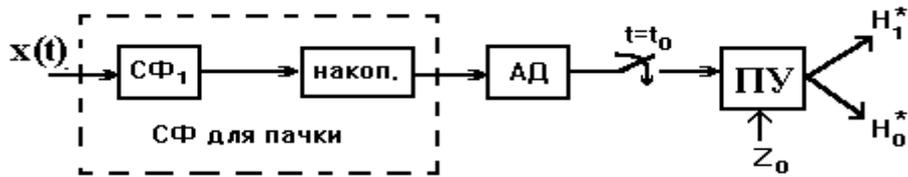


Рис. 3.7. Структурная схема приемника обнаружения с когерентным накоплением импульсов



Рис. 3.8. Структурная схема приемника обнаружения с некогерентным накоплением импульсов

Если энергия отдельных импульсов E_1 в пачке одинакова, то можно перейти к отношению сигнал/шум на один импульс E_1/N_0 . Причем минимальное отношение E_1/N_0 , при котором обеспечиваются заданные характеристики обнаружения, существенно зависит от числа импульсов N в пачке (рис. 3.9) [3].

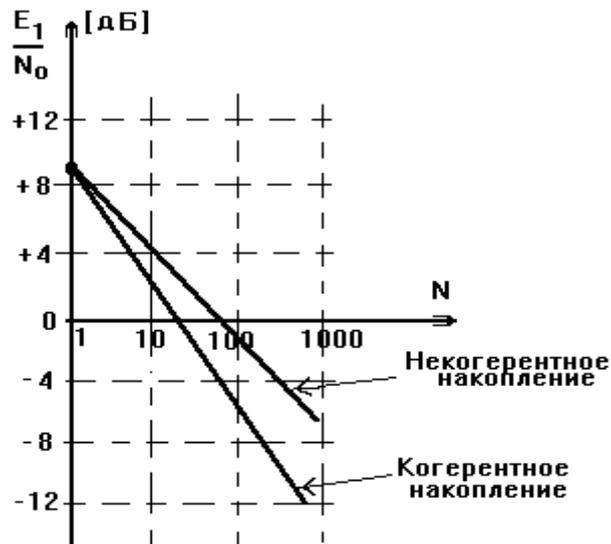


Рис. 3.9. Зависимость минимально-допустимого (порогового) отношения сигнал/шум на один импульс от числа импульсов в пачке

Как видно из рис. 3.9, для обеспечения заданных P_0 и P_d при когерентном накоплении требуется энергия сигнала меньшая, чем при некогерентном накоплении.

§ 3.5. Оптимальное различение детерминированных сигналов

Задача различения сигналов характерна для систем передачи информации (СПИ). Она сформулирована в разделе § 2.2. Рассмотрим

случай различения детерминированных сигналов $u(t, \alpha_{11}^n) = u_1(t)$ и $u(t, \alpha_{12}^n) = u_2(t)$ сигналов (когерентный прием). Принимаемая реализация имеет вид $x(t) = au_1(t) + (1-a)u_2(t) + n(t)$, $t \in [0; T]$, $a = \{a_0=0, P(a_0); a_1=1, P(a_1)\}$, $P(a_0) + P(a_1) = 1$, где $P(a_0)$, $P(a_1)$ — априорные вероятности передачи сигналов $u_2(t)$ и $u_1(t)$ соответственно. Требуется определить структуру оптимального приемника различения двух сигналов. Для СПИ применим критерий Котельникова, который требует знания $P(a_0)$ и $P(a_1)$. В этом случае порог принятия решения $\Lambda_0 = P(a_0)/P(a_1)$. Синтез производится на основе правила отношения правдоподобия (2.7). Рассмотрим случай непрерывной реализации $x(t)$. Из (3.6) видно, что функционалы правдоподобия по двум гипотезам H_i , $i=1, 2$ после преобразования квадрата разности будут иметь вид:

$$L_i(x) = k \cdot e^{-\frac{E_i}{N_0}} \cdot e^J \cdot \exp\left\{\frac{2}{N_0} \int_0^T x(t)u_i(t) dt\right\}, \quad (3.26)$$

$$E_i = \int_0^T u_i^2(t) dt; \quad J = \frac{1}{N_0} \int_0^T x^2(t) dt, \quad i=1, 2. \quad (3.27)$$

Тогда правило отношения правдоподобия имеет вид:

$$\Lambda(x) = \frac{L_1(x)}{L_2(x)} = e^{-\frac{(E_1 - E_2)}{N_0}} \exp\left\{\frac{2}{N_0} \int_0^T x(t) [u_1(t) - u_2(t)] dt\right\} \underset{H_2^*}{\overset{H_1^*}{>}} \Lambda_0. \quad (3.28)$$

Прологарифмировав это выражение, получим алгоритм различения сигналов $u_1(t)$ и $u_2(t)$:

$$Z_1(T) - Z_2(T) \underset{H_0^*}{\overset{H_1^*}{>}} Z_0, \quad Z_i(T) = \int_0^T x(t)u_i(t) dt, \quad i=1, 2, \quad (3.29)$$

$$Z_0 = \frac{N_0}{2} \left[\ln \Lambda_0 + \frac{E_1 - E_2}{N_0} \right]. \quad (3.30)$$

Как всегда, данный алгоритм реализуется при помощи двух схем приема: корреляционной (рис. 3.10) или фильтровой.

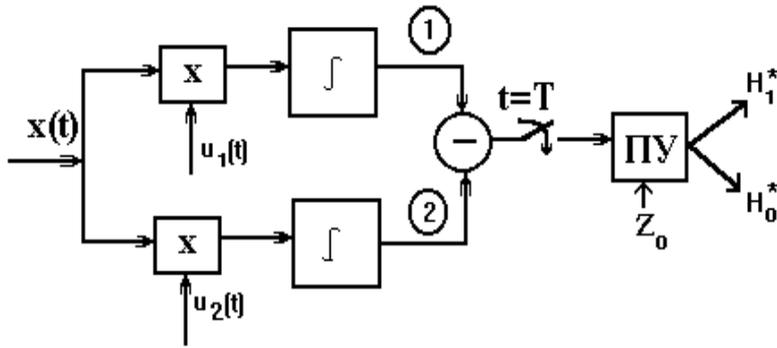


Рис. 3.10. Корреляционная схема устройства различения сигналов $u_1(t)$ и $u_2(t)$

В каждом из двух каналов приемника (рис. 3.10) определяется степень схожести наблюдаемой реализации $x(t)$ с сигналами $u_1(t)$ или $u_2(t)$. Полученные результаты сравниваются в вычитающем устройстве, и выносятся соответствующее решение. Если напряжение в точке 1 (рис. 3.10) больше, чем напряжение в точке 2, то справедлива гипотеза H_1 , в противном случае — гипотеза H_2 .

Схема фильтрового приемника похожа, но в каждом из каналов нужно вместо корреляторов поставить по СФ. Один из них должен быть согласован с сигналом $u_1(t)$, а другой — с $u_2(t)$.

Обычно двоичные СПИ являются симметричными. В них $P(a_0)=P(a_1)=0,5$, $E_1=E_2=E$. Тогда $\Lambda_0=1$ и алгоритм различения сигналов имеет вид:

$$\mathbb{I}_1 - Z_2 \begin{matrix} H_1^* \\ - > \\ < 0 \\ H_2^* \end{matrix} \cdot \quad (3.31)$$

Для примера рассмотрим случай двоичной СПИ с активной паузой. В ней используются полезные сигналы: $u_1(t)=A \cos \omega_0 t$, $t \in [0, T]$, $u_2(t)=-u_1(t)$ (рис 3.11а). В случае, когда помехи нет, $n(t)=0$, напряжения на выходах каналов различных схем приема имеют форму, показанную на рис. 3.11б для корреляционной схемы и на рис. 3.11в для фильтровой.

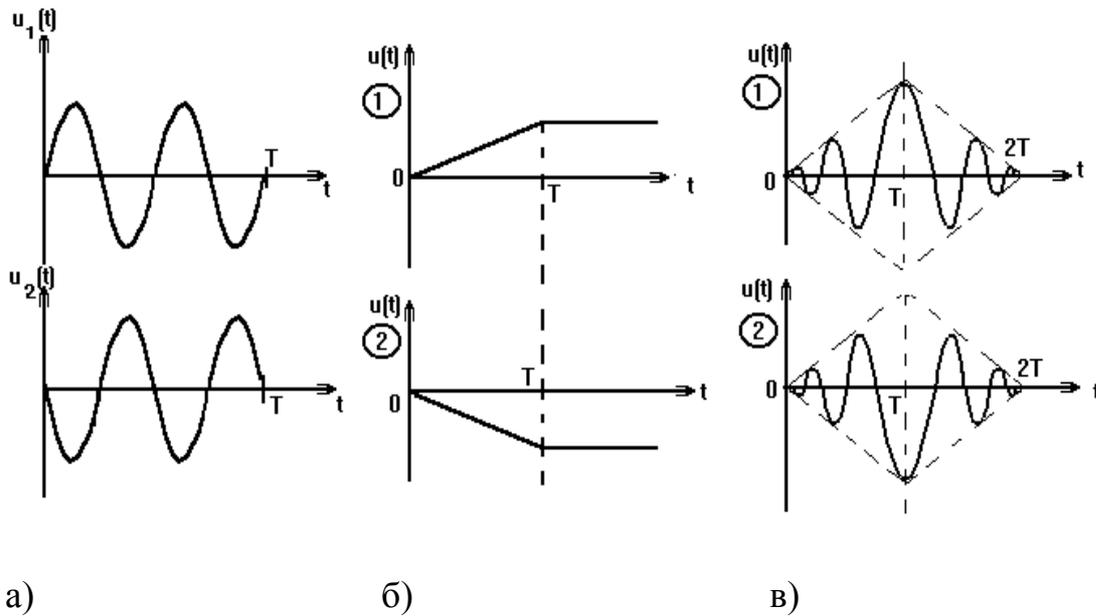


Рис. 3.11. Отклики каналов двоичной СПИ с активной паузой при использовании корреляционной (б) или фильтровой (в) схем различения сигналов

Как видно из рис. 3.11, при таких выбранных сигналах $u_1(t)$ и $u_2(t)$ выходные сигналы каналов 1 и 2 максимально отличаются друг от друга при $t=T$. Тем не менее при наличии помех могут возникнуть ошибки. Если использовать сигналы $u_1(t)$ и $u_2(t)$ со сдвигом по фазе менее 180° , то помехоустойчивость будет хуже, так как сигналы 1 и 2 в момент времени $t=T$ станут по абсолютной величине меньше. Если $u_1(t)=u_2(t)$, то сигналы 1 и 2 будут одинаковыми. Такие сигналы будут неразличимы.

При использовании противоположных сигналов $u_1(t)$ и $u_2(t)$, схема приемника упрощается и остается один канал (рис. 3.1, 3.2).

В двоичных симметричных системах, в которых $P(a_0)=P(a_1)$ и $P(H_1^* / H_2) = P(H_2^* / H_1)$, вероятность ошибки $P_{\text{ош}} = P(H_1^* / H_2) = P(H_2^* / H_1)$ описывается выражением [1]:

$$P_{\text{ош}} = 0,5 - \Phi \left(\sqrt{\frac{q^2}{2} (1 - r_{\text{и}})} \right), \quad r_{\text{и}} = \frac{1}{E_0} \int_0^T u_1(t) u_2(t) dt, \quad (3.32)$$

где $\Phi(x)$ — табличный интеграл вероятности (3.16);

$r_{\text{и}}$ — коэффициент взаимной корреляции, характеризующий схожесть сигналов $u_1(t)$ и $u_2(t)$, принимающий значения $-1 \leq r_{\text{и}} \leq +1$.

Если сигналы $u_1(t)$ и $u_2(t)$ противоположные, то $r_{\text{и}} = -1$, а значит $P_{\text{ош}} = \min$. Если $u_1(t) = u_2(t)$, то $r_{\text{и}} = +1$ и $P_{\text{ош}} = 0,5$, то есть сигналы неразличимы. Если $r_{\text{и}} = 0$, то $P_{\text{ош}}$ больше, чем для противоположных сигналов, но меньше, чем для неразличимых. Зависимость $P_{\text{ош}}$ от $r_{\text{и}}$ для различных значений q^2 дана на рис. 3.12. Как видно из него, $P_{\text{ош}}$ падает с уменьшением $r_{\text{и}}$ и с ростом q^2 .

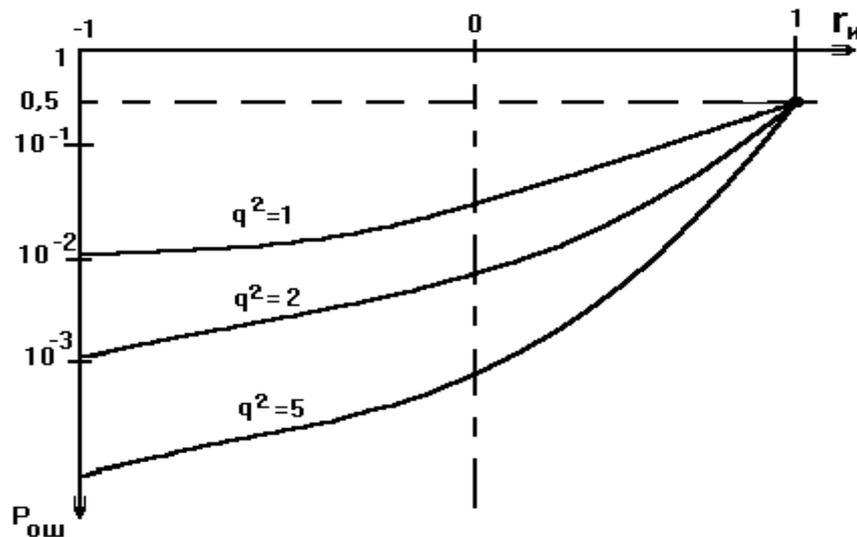


Рис. 3.12. Зависимость вероятности ошибки $P_{ош}$ от коэффициента взаимной корреляции r_u для различных значений отношения сигнал/шум q^2

На рис. 3.13 сплошной линией приведена зависимость $P_{ош}$ от q^2 для различных видов манипуляции несущего колебания, при помощи которых формируются полезные сигналы $u_1(t)$ и $u_2(t)$. Как видно из рисунка, наилучшие результаты дает фазовая манипуляция (ФМН) на 0 и π , для которой $r_u = -1$. Второй по помехоустойчивости является частотная манипуляция (ЧМН), для которой рассмотрен случай $r_u = 0$. Наихудшая помехоустойчивость наблюдается при амплитудной манипуляции (АМН).

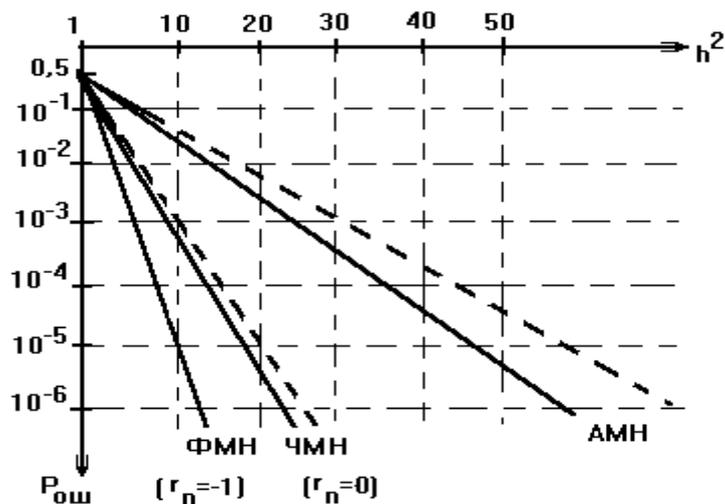


Рис. 3.13. Зависимость вероятности ошибки $P_{ош}$ от q^2 для различных типов полезных сигналов при когерентном и некогерентном приеме

————— когерентный прием
 - - - - - некогерентный прием

Используя методы и приемы, описанные в §§ 3.1-3.3, можно определить правило различения сигналов со случайной начальной фазой β

(некогерентный прием): $u_i(t) = A_i \cos[\omega_i t + \varphi_i(t) + \beta]$, $t \in [0, T]$, $i = 1, 2$. Решение данной задачи подробно описано в [1]. Для сравнения, зависимость $P_{\text{ош}}$ от q^2 для различных типов полезных сигналов при некогерентном приеме показана на рис. 3.13 пунктиром. Как видно из рисунка при ЧМН и АМН помехоустойчивость некогерентного приема хуже, чем когерентного. ФМН-сигналы в схеме некогерентного приема неразличимы. В общем случае наибольшую помехоустойчивость имеет фазовая манипуляция при когерентном приеме.

IV. ОПТИМАЛЬНАЯ ОЦЕНКА ПАРАМЕТРОВ СИГНАЛОВ

§ 4.1. Задача оптимальной оценки параметров сигналов

Часто пользователю для извлечения нужных сведений необходимо измерить (оценить) значения тех или иных параметров сигнала. Термин “оценка” применяется для обозначения как самого процесса измерения параметра, так и его результата. Полученные в результате оценки значения параметров не обязательно точно воспроизведут истинные, так как в реальных условиях полезный сигнал поступает на вход РПУ в смеси с помехами и содержит не только несущие нужную информацию полезные (информационные) параметры, но и параметры, не содержащие нужной информации (неинформационные). В целях общности все параметры будем обозначать символом α , а их оценку — величиной α^* .

Так как α неизвестна, то она будет рассматриваться как случайная величина. Значение α^* вырабатывается в условиях воздействия помех, поэтому α^* также является случайной величиной. Формулировка задачи оценки параметров дана в § 2.2. Априорная плотность вероятности $W(\alpha)$ может быть как известной, так и неизвестной наблюдателю.

Решение задачи оценки параметров предполагает предварительный выбор критерия оптимальности. В результате решения должна быть получена структура оптимального измерителя и определена дисперсия оценки. Методы решения этой задачи базируются на методах математической статистики. Оценка может производиться как по выборкам входного процесса $x(t)$, так и по непрерывным реализациям. В последнем случае оценка будет более точной.

К измеряемым параметрам в РТС относят время прихода сигнала, его длительность, частоту, начальную фазу и т.д. Оценка параметров сигналов будет производиться в предположении, что полезный сигнал на входе РПУ обязательно присутствует. Все параметры сигнала делятся на энергетические и неэнергетические. К энергетическим относят те, от которых зависит отношение сигнал/шум $q^2 = q^2(\alpha)$. Это амплитуда и длительность сигнала. Остальные параметры (частота, фаза и т.п.) относят к неэнергетическим. Ниже мы будем рассматривать только неэнергетические параметры.

Если α принимает дискретные значения из конечного множества размерностью M , то задачу оценки дискретного параметра можно рассматривать как различение M сигналов. Очевидно, что от данного

случая возможен переход и к измерению непрерывного параметра [1]. В силу этой взаимосвязи в первую очередь будем рассматривать только оценку непрерывного параметра.

Процесс оценки параметров сопровождается ошибками. Ошибки делятся на грубые промахи, систематические и случайные. Если есть источник систематической ошибки, то есть смещение оценки $\overline{\alpha^*} - \alpha \neq 0$, то оценка называется смещенной, в противном случае — несмещенной. Ниже мы будем рассматривать только случайные ошибки одного параметра сигнала, обусловленные действием помех.

Для характеристики качества измерителя будем использовать дисперсию оценки $D[\alpha] = \overline{(\alpha^* - \overline{\alpha^*})^2}$. Она является основным качественным показателем измерителя, который всегда стремятся уменьшить: $D[\alpha] = \sigma_\alpha^2 = D_\alpha = \min$. В этом случае измеритель и оценка называются измерителем и оценкой с минимальной дисперсией. В последней формуле приведены различные обозначения дисперсии.

Доказано, что дисперсия любой оценки при некоторых условиях не может быть ниже границы, определяемой неравенством Рао-Крамера, которое для несмещенных оценок имеет вид:

$$D_\alpha \geq \frac{1}{M \left[\frac{\partial^2}{\partial \alpha^2} \ln W(x_1^m / \alpha) \right]}, \quad (4.1)$$

где $W(x_1^m / \alpha)$ — условная плотность вероятности колебания x_1^m .

Несмещенную оценку, имеющую значение дисперсии, равное нижней границе этого неравенства, называют эффективной.

Эффективностью оценки параметра называют отношение дисперсии эффективной оценки к дисперсии рассматриваемой оценки.

При оценке параметров возможны две ситуации: 1) априорная плотность вероятности $W(\alpha)$ известна; 2) $W(\alpha)$ неизвестна. Ниже мы будем рассматривать только первый случай.

При оценке параметров с известной априорной $W(\alpha)$ используется байесовский подход. Он основан на минимизации среднего риска ошибки измерения:

$$\bar{r}[\alpha, \alpha^*] = \int_A \int_{A^*} r[\alpha, \alpha^*] \cdot W[\alpha, \alpha^*] d\alpha d\alpha^* = \min, \quad (4.2)$$

где A и A^* — области изменения α и α^* . Далее полагаем их от $-\infty$ до ∞ ; $r[\alpha, \alpha^*]$ — функция стоимости (риска, потерь).

Тогда $\bar{r}[\alpha, \alpha^*]$ — математическое ожидание риска $r[\alpha, \alpha^*]$, получаемого усреднением по α и α^* .

Если средний риск $\bar{r}[\alpha, \alpha^*]$ минимален, то измерение является наилучшим. Оценка, минимизирующая $\bar{r}[\alpha, \alpha^*]$, называется оптимальной байесовской оценкой. На основе байесовского критерия (4.2) при решении конкретных задач можно строить различные правила оценки, которые рассматриваются в следующем разделе.

Функцию $r[\alpha, \alpha^*]$ выбирают на основе интуитивных соображений, но так, чтобы $r[\alpha, \alpha^*] = \min$ при $\alpha = \alpha^*$ и являлась четной функцией. Наиболее часто используются следующие виды функций стоимости:

- 1) простая: $r[\alpha, \alpha^*] = c - \delta(\alpha - \alpha^*)$;
- 2) квадратичная: $r[\alpha, \alpha^*] = (\alpha - \alpha^*)^2$;
- 3) прямоугольная: (ступенчатая) $r[\alpha, \alpha^*] = 0$, при $|\alpha - \alpha^*| \leq \varepsilon$, $r[\alpha, \alpha^*] = \text{const}$, при $|\alpha - \alpha^*| > \varepsilon$;
- 4) линейная: $r[\alpha, \alpha^*] = k|\alpha - \alpha^*|$.

§ 4.2. Байесовская оценка неэнергетических параметров сигнала

Рассмотрим на примере задачи оптимального измерения неэнергетических параметров сигнала, как на основе байесовского подхода формируются правила оценки.

Пусть параметр α_i полезного сигнала $u(t, \alpha_i^n)$ является измеряемой случайной величиной с априори известной плотностью вероятности $W(\alpha_i)$. Требуется с минимальной погрешностью восстановить значение параметра α_i по принятой реализации $x(t)$, где $x(t) = u(t, \alpha_i^n) + n(t)$, $t \in [0, T]$.

Рассмотрим ситуацию, когда все параметры, кроме полезного, являются неслучайными: $x(t) = u(t, \alpha) + n(t)$, $t \in [0, T]$, где α — полезный параметр. Оценка α^* вырабатывается на основе наблюдений непрерывной реализации $x(t)$ и зависит от $x(t)$, то есть $\alpha^* = \alpha^*(x)$. Таким образом, решение задачи оценки параметров сводится к определению вида $\alpha^* = \alpha^*(x)$. Изложенный ниже подход к оценке параметров справедлив для случая $q^2 \gg 1$, где $q^2 = 2E/N_0$.

Для решения задачи используем байесовский подход (4.2). Если $\alpha^* = \alpha^*(x)$, то, в соответствии с преобразованием двумерной случайной величины, получим: $W(\alpha, \alpha^*(x)) d\alpha d\alpha^*(x) = W(\alpha, x) d\alpha dx$, а двумерная плотность вероятности $W(\alpha, x) = W(\alpha/x) \cdot W(x)$, где $W(\alpha/x)$ — апостериорная плотность вероятности. Тогда:

$$\begin{aligned} \bar{r}[\alpha, \alpha^*(x)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r[\alpha, \alpha^*(x)] \cdot W(\alpha/x) \cdot W(x) d\alpha dx = \\ &= \int_{-\infty}^{\infty} W(x) \int_{-\infty}^{\infty} r[\alpha, \alpha^*(x)] \cdot W(\alpha/x) d\alpha dx = \min. \end{aligned}$$

Из данного выражения видно, что $\bar{r}[\alpha, \alpha^*(x)]$ будет минимальным, если при любом значении величины x оценка $\alpha^*(x)$ выбрана таким образом, что имеет место минимум внутреннего интеграла. Значит, от критерия минимального среднего риска можно перейти к **критерию минимального условного среднего риска**:

$$\bar{r}[\alpha^*(x) / x] = \int_{-\infty}^{\infty} r[\alpha, \alpha^*(x)] \cdot W(\alpha / x) d\alpha = \min. \quad (4.3)$$

Использование различных функций стоимости $r[\alpha, \alpha^*(x)]$ может привести к различным результатам измерений по критериям (4.2) и (4.3). Но, как правило, на практике роль выбора $r[\alpha, \alpha^*(x)]$ несущественна. Если апостериорная плотность вероятности $W(\alpha/x)$ несимметричная, но точность измерения параметра высока, то $W(\alpha/x)$ имеет вид острого пика и различие результатов измерений, произведенных на основе разных $r[\alpha, \alpha^*(x)]$ будет незначительно. Для симметричных $W(\alpha/x)$ и симметричных неубывающих функций потерь $r[\alpha, \alpha^*(x)]$ все байесовские оценки не зависят от выбора $r[\alpha, \alpha^*(x)]$.

Рассмотрим случай выбора квадратичной функции стоимости. Из рис. 4.1 видно, что для неудачно выбранной оценки α^* , минимум условного среднего риска не достигается. Чтобы он был обеспечен, значение α^* нужно выбрать соответствующим абсциссе максимума кривой $W(\alpha/x)$.

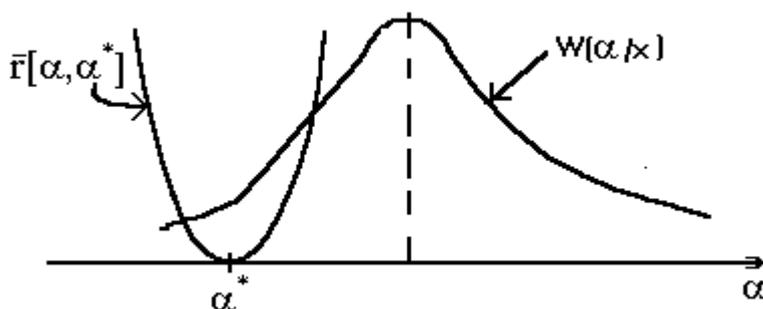


Рис. 4.1. Выбор оценки α^ на основании критерия минимального условного среднего риска и квадратичной функции стоимости $r[\alpha, \alpha^*]$*

Математически задача минимизации условного среднего риска решается путем нахождения его экстремума:

$$\frac{d}{d\alpha^*} \int_{-\infty}^{\infty} (\alpha - \alpha^*)^2 W(\alpha / x) d\alpha = 0 \Big|_{\alpha^* = \alpha_{опт}}$$

Преобразовав квадрат разности и продифференцировав левую часть данного выражения, учитывая, что $\int_{-\infty}^{\infty} W(\alpha/x) d\alpha = 1$, получим:

$$-2 \int_{-\infty}^{\infty} \alpha W(\alpha/x) d\alpha + 2\alpha^* \int_{-\infty}^{\infty} W(\alpha/x) d\alpha = -2 \int_{-\infty}^{\infty} \alpha W(\alpha/x) d\alpha + 2\alpha^* = 0 \quad \alpha^* = \alpha_{opt}^*$$

Из последнего выражения находим:

$$\alpha_{opt}^* = \int_{-\infty}^{\infty} \alpha W(\alpha/x) d\alpha = M[\alpha/x]. \quad (4.4)$$

Таким образом, оптимальная оценка соответствует математическому ожиданию измеряемого параметра $\alpha_{opt}^* = M[\alpha/x]$. Если кривая $W(\alpha/x)$ симметричная, то α_{opt}^* выбирается равной моде (абсциссе максимума):

$$\alpha_{opt}^* = M_0[\alpha/x]. \quad (4.5)$$

Чаще всего $W(\alpha/x)$ действительно симметричная, поэтому последнее правило широко применяют на практике. Его называют **принципом максимума апостериорной вероятности**.

Условная плотность вероятности $W(\alpha/x)$ имеет фундаментальное значение в теории радиосвязи — если она известна, то не надо иметь никаких других сведений для определения α_{opt}^* . Если помеха $n(t)$ статистически задана и известна априорная $W(\alpha)$, построение апостериорной $W(\alpha/x)$ в принципе возможно.

Часто при определении $W(\alpha/x)$ встречаются значительные трудности. К более простым результатам можно прийти, если априорная плотность вероятности $W(\alpha)$ “широкая-равномерная”, то есть в диапазоне возможных значений параметра $\alpha \in [\alpha_{min}, \alpha_{max}]$ она постоянна: $W(\alpha) = 1/(\alpha_{max} - \alpha_{min})$.

Известно, что двумерная плотность вероятности $W(\alpha, x) = W(\alpha/x)W(x) = W(x/\alpha)W(\alpha)$. Из данного выражения получается формула Байеса: $W(\alpha/x) = W(x/\alpha) \cdot W(\alpha) / W(x)$, где $W(x/\alpha)$ — условное априорное распределение x , которое обычно известно. Исходя из формулы Байеса и условия нормировки

$$\int_{-\infty}^{\infty} W(\alpha/x) d\alpha = \frac{1}{W(x)} \int_{-\infty}^{\infty} W(x/\alpha) \cdot W(\alpha) d\alpha,$$

получаем:

$$W(\alpha/x) = K \cdot W(x/\alpha) \cdot W(\alpha), \quad (4.6)$$

где K — коэффициент, не зависящий от α и не влияющий на положение моды распределения $W(\alpha/x)$. Выражение (4.6) позволяет построить $W(\alpha/x)$ по $W(x/\alpha)$ и априорной $W(\alpha)$. Так как в рассматриваемом случае $W(\alpha)$ равномерная, то из (4.6) видно, что моды кривых $W(\alpha/x)$ и $W(x/\alpha)$ совпадают, а значит, от принципа максимума апостериорного распределения можно перейти к **принципу максимума плотности вероятности $W(x/\alpha)$** .

Отметим, что $W(x/\alpha)$ можно рассматривать как функцию от “условия” α при фиксированной наблюдаемой реализации [1]. Если обратиться к определению функции (функционала) правдоподобия (см. § 2.3), то станет очевидным, что $W(x/\alpha)$ — это функция (функционал) правдоподобия $L(x/\alpha)$. Если речь идет о реализациях, то говорят о функционале правдоподобия, а если о дискретных наблюдениях — то о функции правдоподобия. Таким образом, чтобы построить $W(\alpha/x)$ достаточно перемножить $W(\alpha)$ и функцию (функционал) правдоподобия и при “широком-равномерном” распределении $W(\alpha)$ приходим к **принципу максимума функции (функционала) правдоподобия**

$$\frac{\partial}{\partial \alpha} L(x/\alpha) = 0 \Big|_{\alpha = \alpha_{\text{опт}}^*} \quad (4.7)$$

Максимум функции (функционала) правдоподобия достигается в том случае, если $u(t, \alpha)$ в наибольшей степени совпадает с наблюдаемой реализацией $x(t)$. Для решения задачи по правилу (4.7) нужно знать $W(x/\alpha)$; по результату наблюдения x' определить $W(x'/\alpha)$, продифференцировать по α , приравнять к нулю и найти соответствующее $\alpha_{\text{опт}}^*$.

Функционал (функция) правдоподобия монотонно связан с корреляционным интегралом. Поэтому можно получить еще один принцип для нахождения оптимальной оценки — **максимума корреляционного интеграла**.

$$\frac{\partial}{\partial \alpha} Z(x/\alpha) = 0 \Big|_{\alpha = \alpha_{\text{опт}}^*} \quad (4.8)$$

Использование данного правила для построения устройств оптимальной оценки параметра будет рассмотрено в следующем разделе.

§ 4.3. Устройства байесовской оценки неэнергетических параметров сигнала

Синтез устройств оптимальной оценки параметров по правилу максимума корреляционного интеграла (4.8) производится с учетом результатов, полученных в §§ 3.1, 3.2, 3.5. При этом будем считать, что постановка задачи оценки позволяет применять принцип (4.8).

Из (3.26) следует, что для смеси нормального белого шума и сигнала $u(t, \alpha)$ функционал правдоподобия имеет вид:

$$L(x / \alpha) = k \cdot e^J \cdot e^{-\frac{E}{N_0}} \cdot \exp \left\{ \frac{2}{N_0} \cdot \int_0^T [x(t)u(t, \alpha)] dt \right\}, \quad (4.9)$$

где E (энергия сигнала $u(t, \alpha)$) и J описываются выражениями (3.27).

Для получения α_{opt}^* в соответствии с выбранным критерием (4.8) нужно иметь множество значений $Z(x/\alpha)$ в зависимости от значений α . Поэтому схемы оптимальных измерителей должны быть многоканальными. Как всегда, возможны две схемы устройств оптимальной оценки параметров. Корреляционная схема приведена на рис. 4.2.

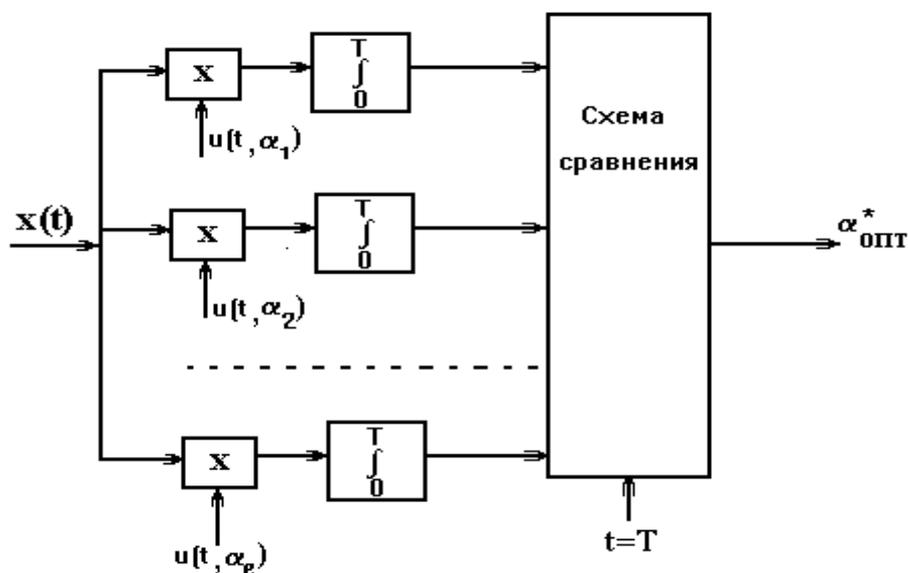


Рис. 4.2. Корреляционная схема устройства оптимальной оценки параметра α

В каждом из каналов вычисляется степень схожести реализации $x(t)$ и копии сигнала с параметром α_i . Решение выносится по максимальному отклику канала. Значение параметра α_i копии сигнала в канале, давшем наибольший отклик, и есть искомая оценка параметра α_{opt}^* . Число каналов l определяется по формуле: $l = (\alpha_{max} - \alpha_{min}) / \Delta\alpha$, где $\Delta\alpha$ — интервал неопределенности по координате α (будет подробнее рассмотрен в главах V, IX) [15].

Схема фильтрового приемника аналогична приведенной на рис. 4.2, только вместо корреляторов устанавливаются фильтры, каждый из которых согласован со своим вариантом полезного сигнала $u(t, \alpha_i)$, $i=1, 2, \dots, l$.

Иногда используются одноканальные, перестраиваемые во времени, схемы оценки параметра. В них параметр копии сигнала α меняется во времени до достижения максимального отклика коррелятора. Данные схемы проще многоканальных, но требуют существенно большего времени анализа, а это определяет и существенно большую длительность наблюдения сигнала. Многоканальные схемы обычно называют параллельными схемами анализа, а одноканальные — последовательными.

На основе приближенного решения уравнения (4.8)

$$\frac{\partial}{\partial \alpha} Z(x / \alpha) = \int_0^T x(t) \frac{\partial}{\partial \alpha} u(t, \alpha) dt \approx \int_0^T x(t) \cdot u(t, \alpha + \Delta \alpha) dt - \int_0^T x(t) \cdot u(t, \alpha - \Delta \alpha) dt = 0 \Big|_{\alpha = \alpha_{opt}^*},$$

можно получить схемы дифференциального типа (рис. 4.3):

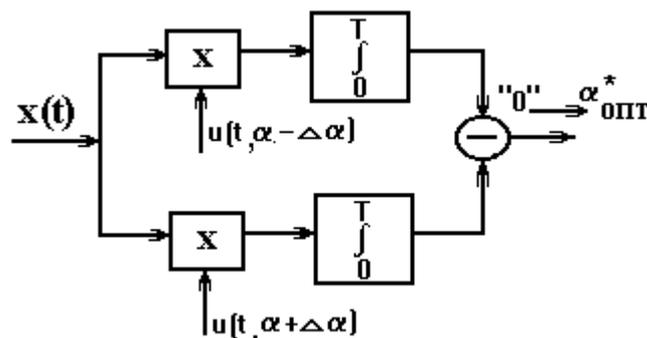


Рис. 4.3. Схема дифференциального типа устройства оптимальной оценки параметра α

Здесь у опорных сигналов α изменяется во времени до получения нулевого сигнала на выходе. Соответствующее ему значение α и есть оптимальная оценка α_{opt}^* .

Основной характеристикой измерителя является дисперсия оценки $\sigma_\alpha^2 = D[\alpha]$. Оптимальный измеритель минимизирует дисперсию: $\sigma_\alpha^2 = \sigma_{min}^2$. Определим дисперсию оценки для рассмотренного случая. Если $q^2 \gg 1$, то можно воспользоваться известной формулой

$$\sigma_\alpha^2 = - \frac{N_0}{2 \cdot Z_c''(\alpha_0, \alpha_0)}, \quad Z_c(\alpha, \alpha_0) = \int_0^T u(t, \alpha_0) u(t, \alpha) dt, \quad (4.10)$$

где $Z_c(\alpha, \alpha_0)$ — автокорреляционная функция (АКФ) по параметру α (сигнальная функция);

α_0 — истинное значение параметра α у принимаемого сигнала;

$Z_c''(\alpha_0, \alpha_0)$ — значение второй производной от сигнальной функции по параметру α в точке $\alpha = \alpha_0$.

Как видим, дисперсия уменьшается с уменьшением спектральной плотности мощности шума N_0 и с увеличением второй производной сигнальной функции $Z_c(\alpha, \alpha_0)$ в точке α_0 . Значение $Z_c''(\alpha_0, \alpha_0)$ тем больше, чем уже главный лепесток АКФ (область высоких значений $Z_c(\alpha, \alpha_0)$ в районе ее главного максимума).

§ 4.4 Оптимальная оценка начальной фазы и смещения частоты сигнала

Рассмотренная выше многоканальная схема является универсальной, но в некоторых случаях могут быть получены более простые схемы. Рассмотрим для примера оценку начальной фазы β радиоимпульса. Будем считать, что постановка задачи позволяет воспользоваться принципом максимума корреляционного интеграла (4.8). Тогда:

$$\frac{\partial}{\partial \beta} Z(x/\beta) = \int_0^T x(t) \cdot A(t) \cdot \frac{\partial}{\partial \beta} \cos[\omega_0 t + \varphi(t) + \beta] dt = 0 \Big|_{\beta = \beta_{\text{опт}}^*}.$$

Ниже индекс “опт” писать не будем, подразумевая его. Продифференцировав последнее выражение, получаем, что максимум $Z(x/\beta)$ будет наблюдаться при

$$\int_0^T x(t) A(t) \sin[\omega_0 t + \varphi(t) + \beta^*] dt = 0.$$

Преобразовав синус суммы по правилу $\sin(\alpha + \beta) = \sin\alpha \cdot \cos\beta + \cos\alpha \cdot \sin\beta$ и решив полученное выражение относительно β^* , получим

$$\beta^* = -\arctg \left\{ \frac{\int_0^T x(t) A(t) \sin[\omega_0 t + \varphi(t)] dt}{\int_0^T x(t) A(t) \cos[\omega_0 t + \varphi(t)] dt} \right\}. \quad (4.11)$$

Пусть $A(t) = 1$, $\varphi(t) = 0$ тогда схема, построенная на основе правила (4.11), будет иметь вид, показанный на рис. 4.4.

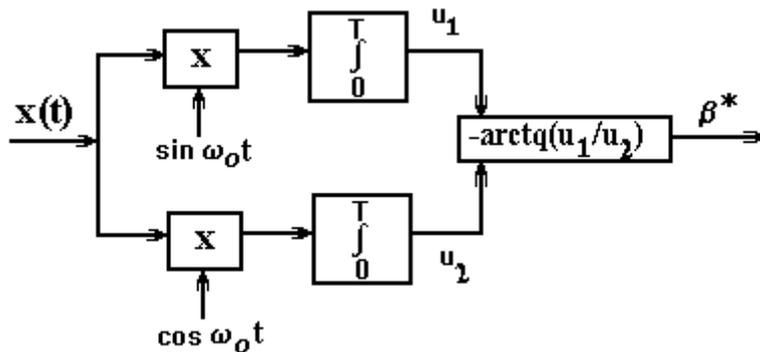


Рис. 4.4. Схема устройства оптимальной оценки начальной фазы β радиоимпульса

Определим дисперсию оценки σ_β^2 , получаемую в данном случае. АКФ по параметру β имеет вид $Z_c(\beta, \beta_0) \approx E \cos(\beta - \beta_0)$, а $Z_c''(\beta, \beta_0) = -E \cos(\beta - \beta_0)$, тогда $Z_c''(\beta_0, \beta_0) = -E$ и из (4.10) получаем $\sigma_\beta^2 = N_0/2E = 1/q^2$. Видно, что дисперсия оценки начальной фазы сигнала не зависит от вида АМН и ФМН и определяется только отношением сигнал/шум на выходе оптимального устройства.

Рассмотрим второй пример — оценку смещения частоты радиоимпульса Ω , которая возникает обычно из-за доплеровского сдвига частоты. Сигнал с неизвестной частотой имеет и неизвестную начальную фазу: $u(t, \Omega, \beta) = A(t) \cos[(\omega - \Omega)t + \beta]$, где β и Ω — случайные величины. Из результатов, полученных при решении задачи обнаружения сигнала со случайной начальной фазой (3.22), известно, что в рассматриваемом случае функционал правдоподобия имеет вид:

$$\bar{L}(x / \Omega) = \text{const} \cdot I_0\left(\frac{2Z(x / \Omega)}{N_0}\right).$$

Максимально правдоподобная оценка соответствует максимуму $\bar{L}(x / \Omega)$. Вследствие монотонности функции Бесселя $I_0(x)$ это соответствует максимуму функции $Z(x / \Omega)$. Простейшим вариантом вычисления $Z(x / \Omega)$ является последовательное включение СФ и детектора огибающей. Континуум значений $Z(x / \Omega)$ можно вычислить с помощью многоканальной фильтровой схемы. В ней используются СФ с непрерывными частотными полосами, которые перекрывают весь априорный диапазон частот. Число каналов l определяется априорным диапазоном доплеровского сдвига и интервалом неопределенности по частоте $\Delta\omega$. Обычно $\Delta\omega \leq \sigma_\Omega$. Дисперсия оценки в данном случае имеет вид:

$$\sigma_{\Omega}^2 = \frac{1}{q^2 \gamma^2},$$

где значение параметра γ зависит от длительности сигнала. Для одиночного радиоимпульса с длительностью $\tau_{\text{и}}$ дисперсия оценки смещения частоты сигнала $\sigma_{\Omega}^2 = 12 / (q^2 \cdot \tau_{\text{и}}^2)$. Следовательно, дисперсия оценки частоты уменьшается с увеличением длительности сигнала и отношения сигнал/шум.

V. ОСНОВЫ ТЕОРИИ РАЗРЕШЕНИЯ СИГНАЛОВ

§ 5.1. Разрешение сигналов

Часто в зоне действия РЛС оказываются не один, а несколько объектов. Поэтому, кроме обнаружения объектов необходимо их разделять (разрешать) (формулировка задачи приведена в § 2.2). В РТС говорят о разрешении по дальности (времени), скорости (частоте), угловым координатам и т.д. Ниже мы будем рассматривать только разрешение по дальности и скорости.

Разрешающая способность является основной характеристикой систем разрешения сигналов (объектов). Количественно разрешающая способность по некоторому параметру α характеризуется минимальной разностью значений данного параметра $\delta\alpha = \alpha_1 - \alpha_2$ у двух еще разделяемых сигналов (объектов), имеющих одинаковые прочие параметры. Например, разрешающая способность по дальности δR равна минимальной разности дальностей двух разрешаемых точечных объектов, расположенных на одном направлении. В теории обнаружения форма сигналов никакого значения не имеет. При разрешении сигналов она имеет определяющее значение.

В радиолокации дальность до цели R и радиальная составляющая скорости цели V_r определяются по времени распространения зондирующего сигнала от РЛС до объекта и обратно τ_0 и доплеровскому сдвигу частот $\Delta\Omega$. Поэтому можно говорить о разрешении по времени τ и частоте Ω .

§ 5.2 Двумерная автокорреляционная функция

Рассмотрим вопрос совместного разрешения сигналов по дальности и скорости. Для этого нужно знать двумерную автокорреляционную функцию (ДАФ). Рассмотрим два комплексных сигнала $\dot{U}(t)$ и $\dot{U}_1(t - \tau)$, отличающихся по задержке τ и частоте Ω :

$$\begin{aligned} \dot{U}(t) &= \dot{A}(t) \exp(j\omega_0 t); \quad \dot{A}(t) = A(t) \exp \mathcal{F}[\varphi(t) + \beta(t)]; \\ \dot{U}_1(t - \tau) &= \dot{A}(t - \tau) \exp \mathcal{F}[(\omega_0 - \Omega)(t - \tau)]; \end{aligned} \quad (5.1)$$

$$\dot{A}(t - \tau) = A(t - \tau) \cdot \exp \left[\alpha(t - \tau) + \beta(t - \tau) \right]$$

Например, в РЛС первый сигнал является излученным, а второй — принимаемым. В качестве меры различения таких сигналов Вудвордом было предложено использовать средний квадрат разности их значений:

$$\Delta^2(\tau, \Omega) = \frac{1}{2E} \int_{-\infty}^{\infty} |\dot{U}(t) - \dot{U}_1(t - \tau)|^2 dt . \quad (5.2)$$

Если комплексные сигналы представить в алгебраической форме, то можно получить:

$$\Delta^2(\tau, \Omega) = 2 \left[1 - \frac{1}{2E} \operatorname{Re} \left\{ \int_{-\infty}^{\infty} \dot{U}(t) \cdot \dot{U}_1^*(t - \tau) dt \right\} \right]. \quad (5.3)$$

Из данной формулы можно выделить часть, зависящую от τ и Ω и называемую **двумерной автокорреляционной функцией (ДАФ)**:

$$\dot{R}(\tau, \Omega) = \int_{-\infty}^{\infty} \frac{1}{2E} \cdot \dot{U}(t) \cdot \dot{U}^*(t - \tau) dt . \quad (5.4)$$

Из (5.4) следует, что если $\tau=0$ и $\Omega=0$, то $\dot{R}(\tau, \Omega) = 1$ и $\Delta^2(\tau, \Omega) = 0$, то есть сигналы неразличимы. Во всех других случаях $\Delta^2(\tau, \Omega) > 0$. Как видно из (5.3), $\Delta^2(\tau, \Omega)$ определяется функцией $\dot{R}(\tau, \Omega)$.

На практике важно знать модульное значение этой функции $|\dot{R}(\tau, \Omega)|$. Подставив (5.1) в (5.4), взяв модуль и, учитывая, что задержка τ является функцией только от дальности R , получим:

$$|\dot{R}(\tau, \Omega)| = \frac{1}{2E} \cdot \left| \int_{-\infty}^{\infty} \dot{A}(t) \cdot \dot{A}^*(t - \tau) \cdot e^{j\Omega t} dt \right|. \quad (5.5)$$

В литературе $|\dot{R}(\tau, \Omega)|$ часто называют **частотно-временной функцией рассогласования**.

Рассмотрим связь $\dot{R}(\tau, \Omega)$ и $|\dot{R}(\tau, \Omega)|$ с откликом устройства оптимальной обработки, например, коррелятора. Данный отклик в момент времени $t=T$ описывается ВКФ $Z(T)$ (3.8). Будем полагать, что шумов нет ($n(t) \equiv 0$), а принимаемый $u_1(t-\tau)$ и опорный (он же излучаемый) сигналы $u(t)$ отличаются по времени τ и частоте Ω . Тогда в общем случае, не накладывая ограничения на время наблюдения сигнала T , из (3.8) и (5.1), воспользовавшись формулой $\operatorname{Re} \left\{ \frac{1}{2} \operatorname{Re} \left\{ \dot{A} \cdot \dot{B} + \dot{A} \cdot \dot{B}^* \right\} \right\}$, получим:

$$Z(\tau, \Omega) = \frac{1}{2} \operatorname{Re} \left\{ \exp(-j(\omega_0 - \Omega)\tau) \cdot \int_{-\infty}^{\infty} \dot{A}(t) \cdot \dot{A}(t - \tau) \cdot \exp(j(2\omega_0 - \Omega)t) dt + \right. \\ \left. + \frac{1}{2} \operatorname{Re} \left\{ \exp(j(\omega_0 - \Omega)\tau) \cdot \int_{-\infty}^{\infty} \dot{A}(t) \cdot A^*(t - \tau) \cdot \exp(j\Omega t) dt \right\} \right\}.$$

Первый интеграл содержит быстро осциллирующий множитель $\exp(j(2\omega_0 - \Omega)t)$, поэтому им можно пренебречь по сравнению со вторым интегралом и сигнал на выходе коррелятора:

$$Z(\tau, \Omega) = \frac{1}{2} \operatorname{Re} \left\{ \exp[j(\omega_0 - \Omega)\tau] \cdot \int_{-\infty}^{\infty} \dot{A}(t) \cdot A^*(t - \tau) \cdot \exp(j\Omega t) dt \right\}.$$

Если в выражении для $Z(\tau, \Omega)$ отбросить знак “ $\operatorname{Re}\{\dots\}$ ”, то получим комплексный сигнал $\dot{Z}(\tau, \Omega)$ на выходе коррелятора:

$$\dot{Z}(\tau, \Omega) = \frac{1}{2} \left[\int_{-\infty}^{\infty} \dot{A}(t) \cdot A^*(t - \tau) \cdot \exp(j\Omega t) dt \right] \cdot \exp[j(\omega_0 - \Omega)\tau]. \quad (5.6)$$

Найдя модуль (огibaющую) сигнала на выходе коррелятора и, сравнив его с величиной $|\dot{R}(\tau, \Omega)|$ (5.5), получим:

$$|\dot{Z}(\tau, \Omega)| = \left| \int_{-\infty}^{\infty} \dot{A}(t) \cdot A^*(t - \tau) \exp(j\Omega t) dt \right| = E |\dot{R}(\tau, \Omega)|. \quad (5.7)$$

Видно, что $|\dot{R}(\tau, \Omega)|$ представляет собой огibaющую сигнала на выходе оптимального устройства (СФ или коррелятора). Эта функция показывает относительную величину отклика коррелятора или СФ на сигнал, сдвинутый по времени τ и частоте Ω относительно сигнала, оптимального для этих устройств, и, тем самым, количественно описывает разрешающую способность по времени и частоте. Огibaющие сигнала на выходе СФ при расстройках по времени и частоте показаны на рис. 5.1.

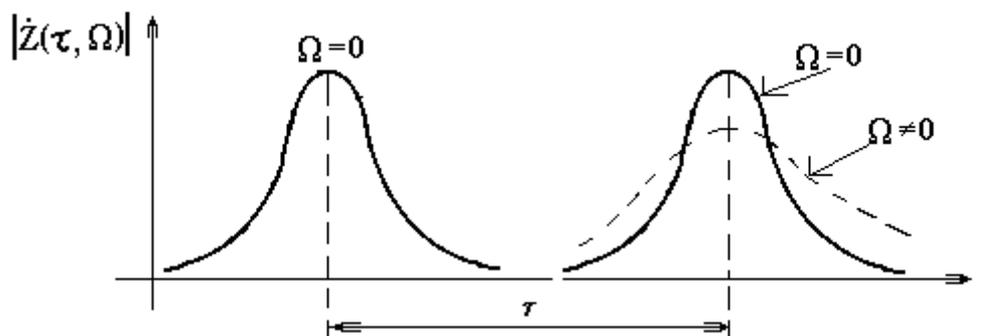


Рис. 5.1. Огibaющие сигнала на выходе СФ

Таким образом, вид сигнала на выходе СФ определяет разрешающую способность по времени и частоте.

§ 5.3. Функция неопределенности и ее свойства.

Принцип неопределенности в РЛС

Более фундаментальным, чем ДАФ и модуль ДАФ, является понятие **функции неопределенности (ФН)**. ФН — это квадрат частотно-временной функции рассогласования $|\dot{R}(\tau, \Omega)|^2$. ФН представляет собой уравнение поверхности над координатной плоскостью Ω, τ , причем, из (5.5) следует, что $|\dot{R}(0,0)|^2 = 1$. Тело, образованное поверхностью неопределенности над плоскостью τ, Ω , называется **телом неопределенности**.

Эквивалентной площадью неопределенности называют площадь основания цилиндра с единичной высотой и объемом, равным объему всего тела неопределенности (рис. 5.2). Она служит совместной мерой разрешения по τ и Ω , обозначаемой $\delta(\tau, \Omega)$, и описывается формулой:

$$\delta(\tau, \Omega) = \frac{1}{|\dot{R}(0,0)|^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\dot{R}(\tau, \Omega)|^2 d\tau d\Omega . \quad (5.8)$$

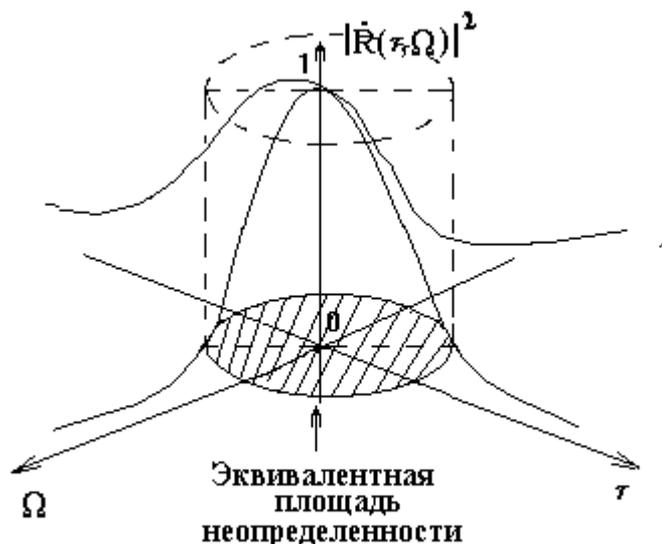


Рис. 5.2. Эквивалентная площадь неопределенности

Если рассогласование по времени τ и частоте Ω у сигналов такие, что точка с данными координатами попадает внутрь эквивалентной площади неопределенности, то сигналы считаются неразличимыми. Этот вариант соответствует области сильной корреляции (область больших

значений $|\dot{R}(\tau, \Omega)|^2$). Остальная область называется областью слабой корреляции, в которой сигналы различимы. Можно показать, что:

$$\int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} |\dot{R}(\tau, \Omega)|^2 d\tau d\Omega = 2\pi . \quad (5.9)$$

Это соотношение описывает принцип **неопределенности в РЛС**. В соответствии с ним, объем тела неопределенности не зависит от вида сигнала (ФМН, ЧМН и т.д.) и является строго постоянным.

Для получения высоких разрешающих способностей по τ и Ω необходимо использовать такие виды сигнала, которые отвечают следующим условиям в отношении ФН:

- 1) $|\dot{R}(\tau, \Omega)|^2 \approx 1$ только в небольшой окрестности точки ($\tau=0, \Omega=0$);
- 2) $|\dot{R}(\tau, \Omega)|^2 \ll 1$ во всех других областях плоскости τ, Ω .

Однако из соотношения неопределенности (5.9) следует, что всякое сужение области сильной корреляции в окрестностях точки ($\tau=0, \Omega=0$) ведет к увеличению значений ФН в других областях и даже к появлению в них новых областей высокой корреляции. Из-за последнего, в частности, может возникнуть неоднозначность при измерении параметров сигнала (объекта).

На практике часто определяют интервалы неопределенности по времени $\tau_{\text{ЭК}}$ и частоте $\Omega_{\text{ЭК}}$ (величины, определяющие эффективную протяженность сигналов по времени и частоте), (рис. 5.3):

$$\tau_{\text{ЭК}} = \frac{1}{|\dot{R}(0,0)|^2} \int_{-\infty}^{\infty} |\dot{R}(\tau, 0)|^2 d\tau , \quad \Omega_{\text{ЭК}} = \frac{1}{|\dot{R}(0,0)|^2} \int_{-\infty}^{\infty} |\dot{R}(0, \Omega)|^2 d\Omega . \quad (5.10)$$

Интервалом неопределенности по времени в общем случае называют размер основания прямоугольника с единичной высотой и площадью, равной площади сечения тела неопределенности плоскостью, проходящей по координатной оси $\Omega=0$ и перпендикулярной плоскости τ, Ω . Аналогично формулируется определение интервала неопределенности по частоте, но сечение тела неопределенности производится по оси $\tau=0$.

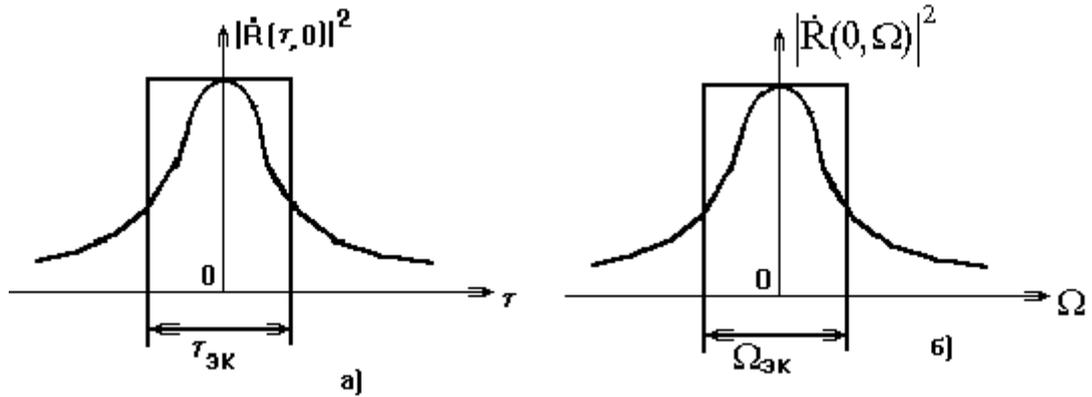


Рис. 5.3. Интервалы неопределенности по времени $\tau_{\text{эк}}$ (а) и частоте $\Omega_{\text{эк}}$ (б)

Найдем интервалы неопределенности по времени и частоте для одиночного радиоимпульса $\dot{U}(t) = \dot{A}(t) \exp\{j\omega_0 t\}$. Пусть $A(t)=1$ при $t \in [0, T]$, тогда энергия сигнала $E = A^2 T / 2 = T/2$. Подставляя $\dot{U}(t)$ в (5.5), легко найти $|\dot{R}(\tau, \Omega)|$. Подставив $|\dot{R}(\tau, \Omega)|$ в (5.10), получаем:

$$\tau_{\text{эк}} = \int_{-\infty}^{\infty} |\dot{R}(\tau, 0)|^2 d\tau = \frac{2}{3} T, \quad \Omega_{\text{эк}} = \int_{-\infty}^{\infty} |\dot{R}(0, \Omega)|^2 d\Omega = \frac{2\pi}{T}. \quad (5.11)$$

Как видно, $\tau_{\text{эк}}$ и $\Omega_{\text{эк}}$ жестко связаны друг с другом. Уменьшение $\tau_{\text{эк}}$ приводит к увеличению $\Omega_{\text{эк}}$ и наоборот. То же самое касается и разрешающей способности по дальности и скорости.

Радиоимпульс является простым сигналом. Напомним, что простыми называют сигналы, у которых база $B = FT \approx 1$. У них B не меняется с изменением F или T , т.к. $F \approx 1/T$. Получить сигналы с $B \gg 1$ можно, промодулировав или проманипулировав по определенному закону простые сигналы по частоте или фазе [4].

Обобщая все вышеизложенное, можно сказать:

1. Разрешение сигналов улучшается, если протяженность главного лепестка ФН и высота боковых лепестков ФН по соответствующей координате уменьшается. Под главным лепестком ФН здесь понимается область высоких значений ФН в окрестностях точки $(\tau=0, \Omega=0)$, а под боковыми лепестками — области достаточно высоких значений ФН вне окрестности точки $(\tau=0, \Omega=0)$.

2. Неоднозначность разрешения уменьшается, если растет расстояние между выбросами ФН по соответствующей координате.

3. Ширина главного лепестка ФН вдоль оси задержки τ пропорциональна $1/F$, а вдоль оси частот — $1/T$, где F и T — соответственно — ширина спектра и длительность сигнала, то есть для получения высоких разрешающих способностей по обеим координатам нужно одновременно увеличивать F и T .

§ 5.4. Широкополосные сигналы и их особенности

Анализ ФН простого сигнала показал его ограниченные возможности. Это вызвано тем, что ширина спектра сигнала F жестко связана с его длительностью T : $F \approx 1/T$. Очевидно, что для обеспечения высоких разрешающих способностей и по дальности, и по скорости необходим сигнал, ширина спектра которого $F \gg 1/T$, то есть $FT \gg 1$. Сигналы, имеющие базу $B \gg 1$ называют широкополосными (ШПС) или сложными. Примерами ШПС являются сигналы с частотной или фазовой манипуляцией или модуляцией — ФМН, ЧМН, с линейной частотной модуляцией (ЛЧМ-сигналы) и т.д. Причем, чем меньше период модуляции или манипуляции τ_c , тем больше B . Отклики СФ и коррелятора на одинаковый сигнал различны по форме, но в момент времени $t=T$ одинаковы по величине. Для СФ:

$$u_{\text{вых}}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{g}_{\text{вых}}(\omega) \cdot e^{j\omega t} d\omega, \quad \dot{g}_{\text{вых}}(\omega) = \dot{g}(\omega) \cdot \dot{K}_0(\omega),$$

где $\dot{g}_{\text{вых}}(\omega)$ — спектр сигнала на выходе СФ;

$\dot{K}_0(\omega)$ — комплексный коэффициент передачи СФ (3.13).

Тогда выходной сигнал СФ определяется как

$$u_{\text{вых}}(t) = c \cdot \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} |\dot{g}(\omega)|^2 \cdot e^{j\omega(t-T)} \cdot d\omega. \quad (5.12)$$

Как видно, длительность выходного сигнала СФ определяется энергетическим спектром сигнала на входе $|\dot{g}(\omega)|^2$. Для белого шума (с полосой $-\infty < \omega < \infty$) энергетический спектр $|\dot{g}(\omega)|^2 = \frac{N_0}{2}$, тогда $u_{\text{вых}}(t) = c \cdot N_0 \cdot \delta(t-T)/2$, где $\delta(t-T)$ — дельта-функция. Из данного выражения видно, что для сигналов, имеющих бесконечно широкий спектр, длительность выходного сигнала СФ $\tau_{\text{эк}}$ бесконечно мала. В реальных случаях у сигналов F не бесконечна и $\tau_{\text{эк}} \approx 1/F$. Поэтому можно определить коэффициент сжатия сигнала по времени:

$$K_{\text{сж}} = \frac{T}{\tau_{\text{эк}}} \approx F \cdot T = B. \quad (5.13)$$

То есть СФ сжимает входной сигнал по времени примерно в B раз. Поэтому СФ для ШПС еще называют фильтром сжатия по времени.

Если на вход коррелятора подать ШПС длительностью T , например ФМН-сигнал, состоящий из N радиоимпульсов длительностью τ_c , то из (3.8) видно, что на выходе коррелятора будет видеосигнал длительностью T [1]. В данном случае ширина спектра входного сигнала $F_{\text{вх}} \approx 1/\tau_c$, а ширина спектра выходного сигнала $F \approx 1/T$. Так как $\tau_c = T/N$, а база ФМН-сигнала $B=N$, то $F_{\text{вх}} \approx N/T = NF = BF$, то есть F уменьшилась по сравнению с $F_{\text{вх}}$ в B раз. Обобщив вышесказанное, можно сделать вывод, что для ШПС коррелятор является устройством сжатия по частоте. Комбинируя корреляторы и СФ, можно получить одновременное сжатие сигнала и по времени, и по частоте, т.е. улучшить разрешающую способность по дальности и скорости.

Увеличивая T и F , можно сузить площадь главного лепестка ФН. Например, ее проекции в плоскости τ , Ω , для ФМН-сигнала, состоящего из N радиоимпульсов длительностью τ_c , показана на рис. 5.4.

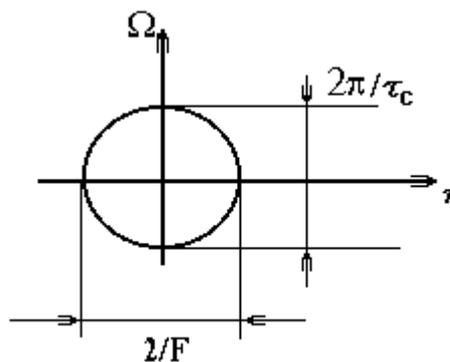


Рис. 5.4. Интервалы неопределенности по времени и частоте ФМН-сигнала, состоящего из N радиоимпульсов длительностью τ_c

Сложные сигналы обладают следующими особенностями, обуславливающими их широкое применение в РТС:

1. Использование ШПС позволяет увеличивать разрешающую способность по дальности и (или) скорости, не снижая дальность действия РЛС. Введение внутриимпульсной модуляции (или манипуляции) позволяет, не меняя энергию сигнала, увеличить F и, следовательно, увеличить разрешающую способность по дальности (см. рис. 5.4). Аналогично при уменьшении периода модуляции (манипуляции) τ_c увеличивается разрешающая способность по частоте (скорости) без уменьшения энергии сигнала.

2. ШПС позволяет осуществлять скрытую передачу сообщений (энергетическая скрытность). Для ШПС на выходе СФ:

$$q^2 = \frac{2E}{N_0} = 2 \frac{P_C T}{P_{\text{ШФ}} / F} = 2FT \cdot \left(\frac{P_C}{P_{\text{ШФ}}} \right) = B \cdot h_{\text{вх}}^2, \quad (5.14)$$

где $h_{\text{вх}}^2 = P_C/P_{\text{ШФ}}$ — отношение сигнал/шум на входе приемника;

F — ширина спектра сигнала (полоса пропускания приемника);

$P_{\text{ШФ}}$ — мощность шума в полосе F .

Видно, что при увеличении F можно достичь сколь угодно малой величины $h_{\text{ВХ}}^2$ при сохранении на выходе требуемого значения q^2 . Данное свойство ШПС поясняет рис. 5.5.

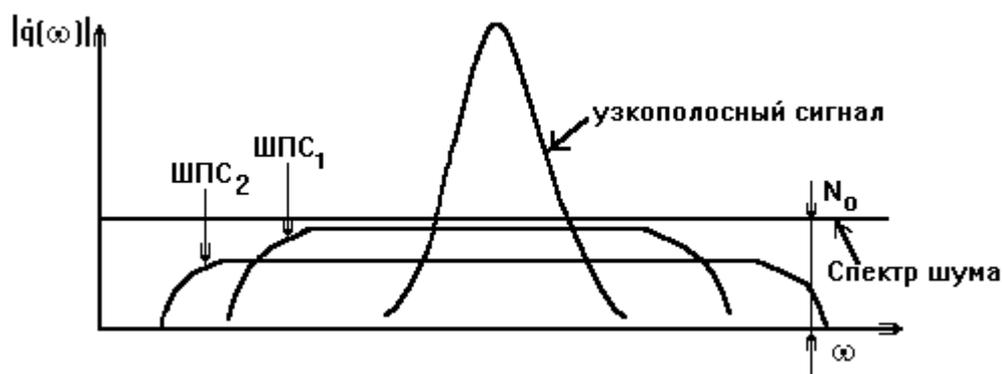


Рис. 5.5. Спектры различных сигналов и белого шума

Пусть энергия ШПС (а значит, и значение отношение сигнал/шум на выходе оптимального РПУ) такая же, как и у узкополосного сигнала (рис. 5.5), но она распределена по более широкой полосе частот, так что на каждой частоте величина спектральных составляющих ШПС меньше N_0 . Поэтому, не зная структуру ШПС, обнаружить его на фоне шума труднее, чем узкополосный сигнал.

3. ШПС позволяют повысить помехоустойчивость СПИ по отношению к сосредоточенным по спектру помехам (прицельным помехам), что поясняет рис. 5.5 (если считать узкополосный сигнал помехой). Из него видно, что относительная доля искаженных помехой спектральных составляющих сигнала больше у ШПС₁, чем у ШПС₂. Поэтому в последнем случае помеху можно отфильтровать, потеряв при этом меньшую долю энергии сигнала, чем в случае ШПС₁.

4. ШПС позволяют обеспечить работу многих РТС в одной и той же полосе частот (что реализуется в системах с разделением каналов по форме сигналов).

VI. Общие сведения о радиолокации

§ 6.1. Основные понятия и определения

Радиолокацией называется область радиотехники, объединяющая методы и средства обнаружения, измерения координат и параметров движения, а также опознавания, определения свойств и характеристик различных объектов (радиолокационных целей), основанные на использовании радиоволн, отражаемых, переизлучаемых или излучаемых этими объектами.

Процесс обнаружения объектов, измерения их координат и параметров движения называют радиолокационным наблюдением, а

используемые для этого РТС – радиолокационными станциями (РЛС) или радиолокаторами. В иностранной литературе для радиолокации и РЛС принято название радар (radar – radio detection and ranging – радиообнаружение и определение расстояния) [1].

С учетом различной природы радиосигналов, несущих информацию об объекте радиолокационного наблюдения, радиолокацию подразделяют на активную, полуактивную, активную с активным ответом и пассивную.

При активной радиолокации информацию об объекте несет радиосигнал, отраженный от объекта (называемый радиолокационным) при его облучении зондирующим (прямым) радиосигналом РЛС. В состав активной РЛС входят передатчик и приемник, а также передающая и приемная антенны (часто используется одна приемопередающая антенна).

При полуактивной радиолокации информацию об объекте также несет отраженный от него радиосигнал, но источник облучающих радиоволн удален от приемника РЛС и может работать независимо от него.

Важным достоинством активных и полуактивных РЛС является возможность обнаружения объектов, не излучающих радиоволны.

При активной радиолокации с активным ответом информацию об объекте несет радиосигнал РЛС, ретранслируемый (переизлучаемый) специальным приемопередатчиком (ответчиком), размещенным на объекте. Так как ответный сигнал может иметь большую мощность, чем отраженный при обычной активной радиолокации, становится возможным существенно увеличить дальность действия и помехозащищенность системы. Ответный сигнал может нести какую-либо дополнительную информацию и используется, например, для опознавания объекта. При пассивной радиолокации информация об объектах извлекается в результате приема и обработки сигналов, излучаемых установленными на них радиоэлектронными средствами, радиоволн, излучаемых создаваемыми объектом ионизированными областями окружающей среды, а также радиотеплового излучения объектов.

§ 6.2. Классификация РЛС, их тактико-технические характеристики

Радиолокационные станции принято классифицировать следующим образом:

- 1) по происхождению принимаемого радиосигнала различают активные, полуактивные и пассивные РЛС;
- 2) по используемому диапазону радиоволн – РЛС декаметрового, метрового, дециметрового, сантиметрового и миллиметрового диапазонов;
- 3) по виду зондирующего сигнала – РЛС с непрерывным и импульсным излучением;
- 4) по количеству каналов излучения и приема сигналов – одно- и многоканальные;

- 5) по количеству измеряемых координат – одно-, двух- и трехкоординатные;
- 6) по месту установки – наземные, корабельные, самолетные и спутниковые РЛС;
- 7) по функциональному назначению – РЛС измерения скорости автомобилей, РЛС предупреждения столкновений и т. д.

Совокупность основных параметров РЛС составляет ее тактико-техническую характеристику (ТТХ).

Тактическими называют характеристики РЛС, определяющие ее функциональные возможности при практическом, в том числе военном применении.

Основными тактическими характеристиками РЛС являются:

- 1) зона действия или зона обзора;
- 2) время обзора;
- 3) определяемые параметры объекта и точность их измерения;
- 4) разрешающая способность;
- 5) пропускная способность;
- 6) помехозащищенность;
- 7) надежность.

Важными параметрами РЛС, влияющими на общую оценку качества, являются экономические показатели, а также масса и габариты входящих в ее состав устройств.

Основными техническими характеристиками РЛС, непосредственно определяющими ее тактические характеристики, являются:

- 1) метод обзора (поиска) и измерения координат и параметров движения объекта;
- 2) рабочие частоты, их стабильность, мощность, вид модуляции (манипуляции) и ширина спектра излучаемых электромагнитных колебаний;
- 3) коэффициенты усиления, форма и ширина диаграмм направленности антенн;
- 4) чувствительность и полоса пропускания приемного устройства;
- 5) вид и параметры устройств отображения и считывания информации;
- 6) габариты и масса устройств, входящих в состав РЛС;
- 7) мощность, потребляемая от источников питания. Подробное содержание конкретных ТТХ приведено в § 1.3.

§ 6.3. Основные сведения о радиотехнических методах определения координат и их производных

Мгновенное пространственное положение объекта в общем случае описывается тремя координатами x_i ($i=1; 2; 3$) в какой-либо системе координат. Параметры движения объекта являются производными

координат по времени $\frac{\partial^n x_i}{\partial t^n}$, $n=1; 2; \dots$, количество которых зависит от сложности траектории и характера движения объекта. Обычно при радиолокации объектов определяют, кроме собственно координат x_i их первые и вторые производные, т.е. скорость $V_i = \frac{\partial x_i}{\partial t}$ и ускорение $a_i = \frac{\partial^2 x_i}{\partial t^2}$ (как правило, для центра тяжести объекта). При этом удобно использовать сферические координаты, а именно: дальность R – расстояние между РЛС и объектом, азимут α – угол в горизонтальной плоскости между направлением на объект и выбранным направлением отсчета N (например, направлением на север) и угол места β – угол между направлением на объект и его проекцией на горизонтальную плоскость (см. рис. 6.1.):

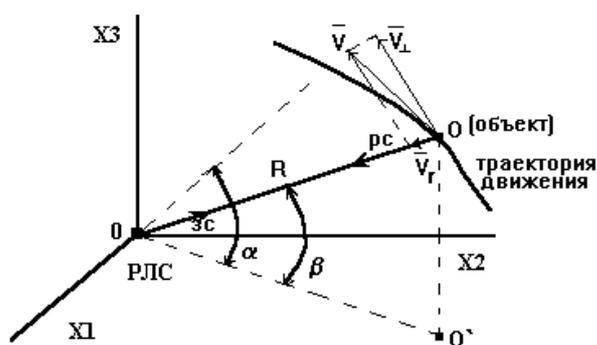


Рис. 6.1. Определение параметров движения объекта

Радиотехнические методы определения координат объектов и их производных основаны на свойстве радиоволн распространяться в однородной изотропной среде прямолинейно с постоянной скоростью, зависящей от диэлектрических и магнитных свойств среды ($c \approx 3 \cdot 10^8$ м/с в свободном пространстве (вакууме)) и заключаются в оценке и сравнении параметров радиолокационного сигнала, возникающего при отражении от объекта, с параметрами зондирующего сигнала, излучаемого передающей антенной РЛС.

Так, при активной радиолокации по измеренному времени τ распространения радиоволн от РЛС до объекта и обратно можно рассчитать дальность объекта R :

$$R = c\tau/2; \quad (6.1)$$

по доплеровскому смещению частоты отраженного сигнала $F_v = 2f_0 V_r/c = 2V_r/\lambda_0$ (f_0 и λ_0 – несущая частота и длина волны зондирующего излучения) – радиальную скорость движения объекта:

$$V_r = F_v c / (2f_0) = F_v \lambda_0 / 2; \quad (6.2)$$

по направлению прихода отраженного сигнала – угловые координаты (азимут α и угол места β) с использованием направленных свойств антенны РЛС.

Для иллюстрации использования радиотехнических методов определения координат объектов, взаимодействия компонент РЛС и взаимосвязи технических и тактических характеристик РЛС рассмотрим в упрощенном виде структуру и принцип действия широко используемой РЛС кругового обзора [1, 6].

§ 6.4. Устройство, принцип действия и основные характеристики РЛС кругового обзора

РЛС кругового обзора (РЛС КО) предназначена для обнаружения, опознавания, измерения дальности и азимута целей. Упрощенная структурная схема некогерентной импульсной РЛС КО приведена на рис. 6.2, а.

Рассмотрим структуру РЛС КО, состав и назначение ее основных блоков. На структурной схеме РЛС КО (рис. 6.2, а) обозначены: С – синхронизатор; ПРД – передатчик; АП – антенный переключатель; А – приемопередающая антенна; УВА и ДПА – устройство вращения и датчик положения антенны; ПРМ – приемник; ИКО – индикатор кругового обзора.

На экране ИКО (рис. 6.2, б) показаны: ШОА – шкала отсчета азимута; МКД – масштабные кольца дальности; ЛРД – линия развертки дальности; Ц1, Ц2 – отметки целей 1 и 2.

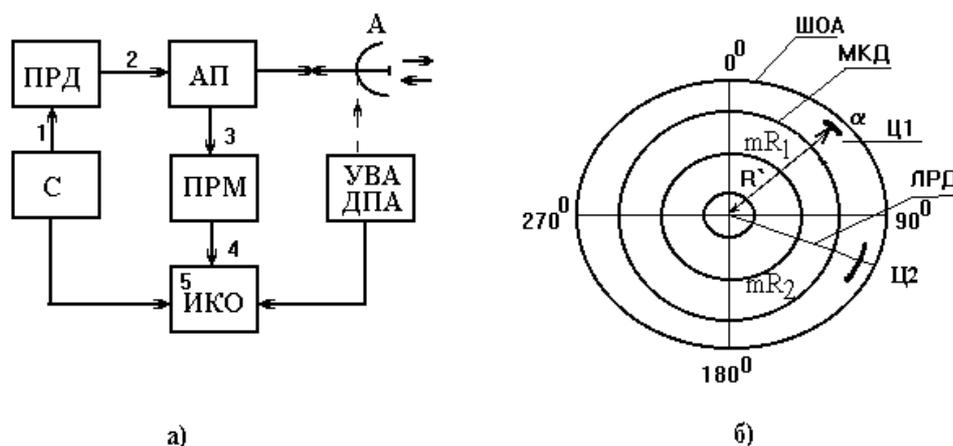


Рис. 6.2. Структура РЛС КО (а) и экран ИКО (б)

Синхронизатор содержит генератор синусоидальных колебаний со стабилизированной частотой и формирователь импульсов управления импульсным модулятором передатчика и устройством радиальной

развертки, формирующим линию развертки дальности на экране электронно-лучевой трубки (ЭЛТ) индикатора кругового обзора.

Передачик состоит из импульсного модулятора и мощного генератора несущих колебаний (либо маломощного задающего генератора и усилителя мощности); он формирует радиоимпульсы с заданными длительностью, периодом повторения и мощностью.

Антенный переключатель поочередно подключает антенну к выходу передатчика во время излучения зондирующих импульсов и ко входу приемника во время пауз между ними.

Антенна служит для излучения зондирующих радиоволн и приема радиоволн, отраженных от целей. Устройство вращения антенны (электродвигатель с редуктором) обеспечивает электромеханическое сканирование диаграммы направленности (ДН) в азимутальной плоскости путем вращения антенны с заданной угловой скоростью, определяющей требуемое время кругового обзора. Датчик положения антенны вырабатывает сигналы синхронизации поворота антенны с круговым поворотом линии развертки дальности на экране ЭЛТ индикатора кругового обзора.

Приемник обычно выполняется по супергетеродинной схеме и содержит усилитель высокой частоты, преобразователь частоты, усилитель промежуточной частоты, устройства автоматической регулировки усиления и подстройки частоты гетеродина и детектор огибающей сигналов, отраженных от целей. На выходе приемника устанавливается устройство селекции движущихся целей, подавляющее сигналы, отраженные от неподвижных объектов (пассивные помехи).

Индикатор кругового обзора состоит из видеоусилителя, электронно-лучевой трубки, устройств формирования масштабных колец дальности и меток шкалы отсчета азимута, а также устройств радиальной и круговой развертки электронного луча и служит для визуальной индикации сигналов обнаруженных целей и сигналов их опознавания, а также для определения дальностей и азимутов целей.

Принцип действия РЛС КО иллюстрируется временными диаграммами напряжений и токов в различных точках структурной схемы РЛС, рис. 6.3.

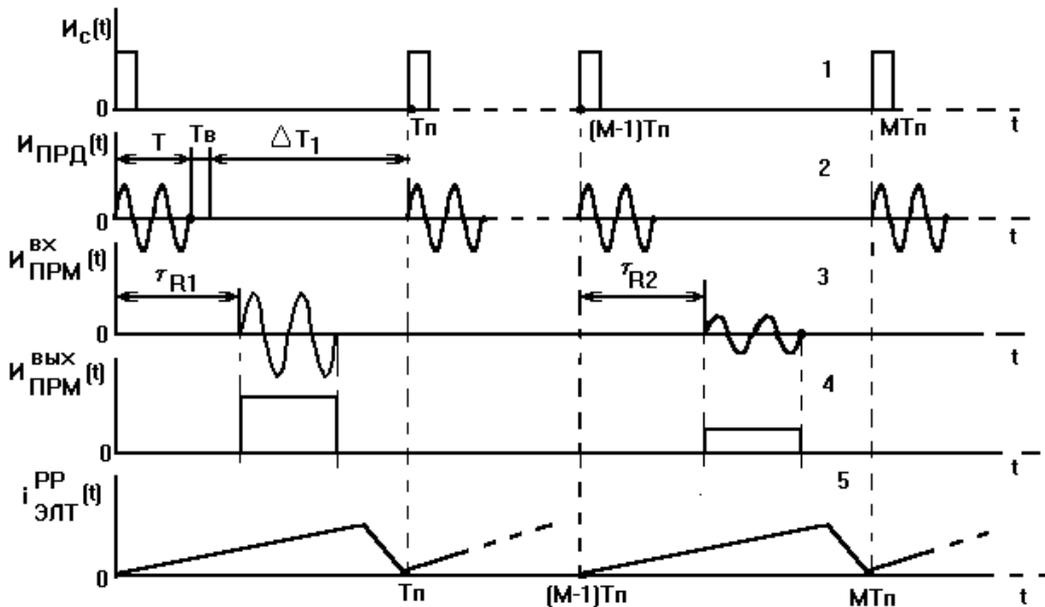


Рис. 1.3. Временные диаграммы, поясняющие работу РЛС КО

РЛС КО действует следующим образом. Управляющие импульсы синхронизатора периодически запускают импульсный модулятор, в результате чего на выходе передатчика формируются радиоимпульсы с длительностью T и периодом повторения $T_{\text{п}}$, поступающие по общей фидерной линии через антенный переключатель на вход приемопередающей антенны и излучаемые ею в окружающее пространство. Через некоторое время восстановления $T_{\text{в}}$ после окончания каждого зондирующего импульса АП отключает передатчик от антенны и подключает к ней приемник. В течение интервалов времени

$$\Delta T_i = (T_{\text{п}} - (T + T_{\text{в}}))_i, \quad i=1; 2; 3; \dots \quad (6.3)$$

происходит прием и обработка радиолокационных сигналов, отраженных от целей, находящихся в текущем (по азимуту) угловом поле зрения антенны, ограниченном шириной ее ДН в горизонтальной и вертикальной плоскостях ($\alpha_{\text{А}}$ и $\beta_{\text{А}}$).

В электронно-лучевой трубке ИКО при этом происходит радиально-круговая развертка электронного луча (ток в катушках схемы радиальной развертки обозначен как $i_{\text{ЭЛТ}}^{\text{PP}}(t)$), в результате чего на ее экране появляется линия развертки дальности (ЛРД), вращающаяся синхронно с вращением антенны, а также изображаются масштабные кольца дальности (МКД) и шкала отсчета азимута (ШОА).

В течение времени ΔT_i на выходе приемника появляются видеоимпульсы, управляющие током луча ЭЛТ так, что на ЛРД высвечиваются яркие отметки целей (Ц1 и Ц2 на рис. 6.2, б), удаленные от центра экрана ЭЛТ на расстояния $mR_{1,2}$, пропорциональные дальностям до целей (m – масштабный коэффициент).

Дальности до целей R' определяются с помощью МКД, а азимуты α – с помощью ШОА. Протяженность ярких отметок по R и α (имеющих вид светящихся дуг) зависит от разрешающей способности РЛС

δR и $\delta \alpha$ и размеров целей по R и α . Длительность зондирующего импульса T и время восстановления АП T_v ограничивают минимальную дальность обнаружения целей (мертвую зону РЛС) величиной

$$R_{\text{мин}} \geq \frac{c(T+T_v)}{2}. \quad (6.4)$$

Кратко рассмотрим взаимосвязь технических и тактических характеристик РЛС. Так, максимальная дальность $R_{\text{макс}}$ непосредственно зависит от энергии зондирующего импульса $E_{\text{и}}=P_{\text{и}}T$, пороговой мощности радиолокационного сигнала (чувствительности приемника) $P_{\text{смин}}$, коэффициента различимости ИКО (отношения сигнал/шум, достаточного для надежного обнаружения отметки цели на экране), а также от коэффициента усиления антенны РЛС. Разрешающая способность по дальности связана с шириной спектра и формой сигнала, а разрешающая способность по азимуту – с шириной ДН антенны РЛС [1].

Кроме названных блоков, в состав РЛС КО может входить специализированный компьютер, производящий вторичную обработку информации о целях (например, вычисляющий траектории их движения) [1].

VII. Физические основы радиолокации

§ 7.1. Объекты радиолокационного наблюдения и формирование отраженных сигналов

Любой объект, электрофизические параметры которого (диэлектрическая или магнитная проницаемости, электропроводность) отличаются от параметров окружающей его среды так, что на границе “среда” – “поверхность” возникает скачок волнового сопротивления, является неоднородностью для падающих на него (зондирующих) радиоволн, вызывающей их дифракцию (рассеяние) [1, 7]. Существенно, что при этом появляются отраженные вторичные радиоволны, достигающие приемной антенны РЛС (называемые **отраженным** или **радиолокационным сигналом**).

Вторичное излучение может быть зеркально отраженным, диффузно рассеянным или резонансным в зависимости от свойств поверхности и размеров объекта. **Зеркальное отражение** имеет место при облучении “гладких” поверхностей, размеры которых многократно превышают длину волны λ падающей радиоволны, а размеры шероховатостей не превосходят $h_0=\lambda/(8\cos\Theta)$, где Θ – угол падения (критерий Рэлея) [7, 8]. При этом углы падения и отражения равны по величине, а направления падающей, отраженной и преломленной волн лежат в одной плоскости. **Диффузное рассеяние** (ненаправленное излучение) возникает при облучении больших поверхностей с шероховатостями, имеющими высоту $h>h_0$. **Резонансное излучение** возникает при размерах облучаемых объектов, кратных нечетному количеству полуволн. При этом

интенсивность вторичного излучения зависит от ориентации большей оси объекта относительно вектора напряженности электрического поля падающей волны.

Особенно интересные явления имеют место при рассеянии волн пространственно-периодическими структурами (решетками проводов, металлическими гребенками и т. п.), период которых близок к длине волны: это может быть полное незеркальное отражение (угол падения не равен углу отражения); автоколлимационное отражение (навстречу падающей волне), в том числе полное; боковое рассеяние с минимальным (близким к 0) уровнем зеркального отражения и др.

Цель считается **малоразмерной** (точечной), если ее размеры много меньше элементов разрешения по дальности и угловым координатам. Для РЛС с разрешающей способностью по дальности δR , по азимуту $\delta\alpha$ и углу места $\delta\beta$ элементом разрешения по дальности является разрешаемое расстояние δR , а по азимуту и углу места – разрешаемые угловые расстояния $\delta R_\alpha = R\delta\alpha$ и $\delta R_\beta = R\delta\beta$. Часто используют понятие элементарной площади $\delta S = \delta R_\alpha \delta R_\beta$ и элементарного (разрешаемого) объема РЛС $\delta V = \delta R \delta S$. Названные параметры используются при анализе рассеяния радиоволн **распределенными объектами**, представляющими собой совокупность множества объектов, расположенных на какой-либо поверхности или заполняющих некоторый объем с размерами, большими элементов разрешения РЛС (δS , δV). Примерами распределенных объектов служат дождевые облака, участки земной поверхности и др. В этих случаях радиолокационный сигнал является результатом интерференции сигналов элементарных рассеивателей, попавших в пределы элемента разрешения РЛС.

§ 7.2. Эффективная площадь рассеяния цели

Эффективной площадью рассеяния (ЭПР) цели $\sigma_{ц}$ называется площадь поперечного сечения такого воображаемого объекта, который, равномерно (изотропно) рассеивая падающие на него радиоволны, в месте приема создает такую же плотность потока мощности, что и реальная цель. Если плотность потока мощности зондирующего излучения РЛС вблизи цели Π_1 , то при ЭПР цели $\sigma_{ц}$ мощность изотропно рассеянного ею излучения $P_1 = \sigma_{ц} \Pi_1$. При расстоянии R между целью и РЛС плотность потока мощности излучения Π_2 , падающего на раскрыт ее приемной антенны [1]

$$\Pi_2 = \frac{P_1}{4\pi R^2} = \frac{\sigma_{ц} \Pi_1}{4\pi R^2}, \quad (7.1)$$

следовательно, ЭПР цели определяется выражением

$$\sigma_{ц} = 4\pi R^2 \frac{\Pi_2}{\Pi_1}. \quad (7.2)$$

Из формулы (7.2) следует, что ЭПР $\sigma_{ц}$ имеет размерность площади. Величина $\sigma_{ц}$ не зависит ни от интенсивности излучения, ни от расстояния, так как при увеличении Π_1 пропорционально изменяется Π_2 и отношение Π_2/Π_1 остается постоянным, а при увеличении R это отношение изменяется обратно пропорционально R^2 .

Воспользовавшись известным из электродинамики соотношением $\Pi=E^2/W$, связывающим плотность потока мощности с действующим значением напряженности электрического поля E и волновым сопротивлением среды W , ЭПР можно рассчитать и по известным напряженностям электрического поля E_1 в районе цели и E_2 в раскрыве приемной антенны РЛС:

$$\sigma_{ц} = 4\pi R^2 \frac{E_2^2}{E_1^2}.$$

Данное выражение справедливо при совпадении поляризации падающих и отраженных волн. В общем случае следует учитывать поляризационные характеристики антенны РЛС, свойства среды распространения радиоволн и параметры целей.

Реальные объекты радиолокационного наблюдения (автомобиль, самолет и др.) имеют сложную форму и рассчитать их ЭПР затруднительно. Расчет возможен только для некоторых простейших объектов – плоского металлического листа, металлического и диэлектрического шаров, зеркального уголкового отражателя, полуволнового вибратора, участка диффузно-рассеивающей поверхности и некоторых групповых и распределенных целей. В большинстве случаев ЭПР реальных объектов определяется экспериментально.

§ 7.3. ЭПР простейших объектов

ЭПР плоского идеального проводящего листа. Если идеальный проводящий лист прямоугольной формы с размерами $a, b \gg \lambda$; $a, b \ll R$ расположен перпендикулярно направлению облучения, его ЭПР рассчитывается по формуле [1]

$$\sigma_{л} = \frac{4\pi S^2}{\lambda^2} = \frac{4\pi(ab)^2}{\lambda^2}. \quad (7.3)$$

ЭПР листа из металла с конечной проводимостью при тех же размерах $\sigma'_{л} = \sigma_{л} |\dot{\rho}|^2$, где $\dot{\rho}$ – комплексный коэффициент отражения металла [9].

Как видно из (7.3), при нормальном облучении идеально проводящего листа его ЭПР может существенно превосходить

геометрическую площадь S : например, при $\lambda=10$ см ЭПР листа площадью $S=1$ м² $\sigma_{\text{л}}=1256$ м², что в несколько раз превышает ЭПР транспортного самолета. Вместе с тем, даже при малом отклонении направления облучения от нормали к плоскости листа его ЭПР резко снижается.

В ряде случаев возникает потребность в устройствах, формирующих интенсивные отраженные сигналы при облучении с различных направлений. К числу таких устройств относятся зеркальные уголкового отражатели.

ЭПР зеркального уголкового отражателя (ЗУО). Конструктивно ЗУО представляет собой сочетание трех взаимно перпендикулярных зеркально отражающих граней (рис. 7.1.):

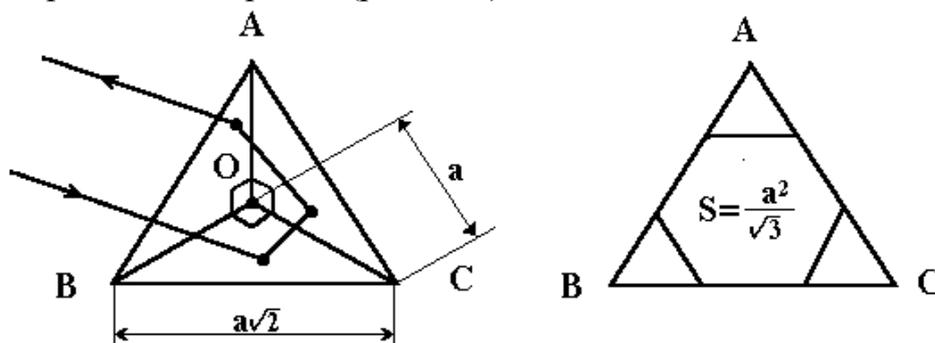


Рис. 2.1. Конструкция ЗУО

Если направление облучения находится вблизи оси симметрии, проходящей через вершину и центр плоскости раскрытия (ABC) – в пределах шестиугольника, вписанного во внешний контур ABC отражателя (ABC – равносторонний треугольник, $|AB|=|BC|=|CA|=a\sqrt{2}$, a – размеры "теневого" граней – $|OA|=|OB|=|OC|=a$), то благодаря трехкратному переотражению волны от граней ЗУО возникает интенсивное отражение навстречу облучению. С учетом площади шестиугольника $S=a^2/\sqrt{3}$ из (7.2) для расчета ЭПР ЗУО можно получить следующее выражение [1]: $\sigma_{\text{зуо}}=4\pi a^4/(3\lambda^2)$.

Сравним ЭПР плоского квадратного металлического листа и ЗУО с одинаковым размером a . Пусть $a=1$ м, $\lambda=10$ см, тогда $\sigma_{\text{л}}=1256$ м², $\sigma_{\text{зуо}}=419$ м². Как видно, $\sigma_{\text{л}} > \sigma_{\text{зуо}}$, однако уголкового отражатель в отличие от листа сохраняет большое значение ЭПР в пределах достаточно существенного телесного угла, тогда как ЭПР листа резко снижается при незначительных отклонениях направлении облучения от нормали.

ЭПР шара. ЭПР шара в общем случае зависит от соотношения между его диаметром и длиной облучающей волны, а также от электрофизических параметров материала, из которого выполнен шар. ЭПР шара, имеющего гладкую идеально проводящую поверхность, при радиусе $r_{\text{ш}} \gg \lambda$ равна площади его поперечного сечения и не зависит от длины волны и направления облучения: $\sigma_{\text{ш}} = \pi r_{\text{ш}}^2 = S_{\text{ш}}$, поэтому металлические шары с $r_{\text{ш}} \gg \lambda$ часто используют в качестве эталонных отражателей при измерении ЭПР реальных объектов путем сравнения интенсивности отраженных волн.

При $r_{ш}/\lambda \leq 2$ шар начинает вести себя подобно вибратору и функция $f(r_{ш}/\lambda) = \sigma_{ш}/S_{ш}$ становится осциллирующей; при $r_{ш} \approx \lambda/2$ $\sigma_{ш} \approx 4S_{ш}$ ЭПР малого шара ($r_{ш} \leq \lambda/2\pi$) определяется формулой Рэлея [1] $\sigma_{ш} = 4,4 \cdot 10^4 r_{ш}^4 / \lambda^4$ и, как видно, существенно зависит от λ (на практике последний случай встречается при рассеянии радиоволн дождевыми каплями).

Отражение от шара – частный случай рассеяния ЭМВ криволинейной поверхностью. В общем случае отражения от идеально проводящей поверхности ее ЭПР определяется как $\sigma = \pi r_1 r_2$, где $r_1, r_2 \gg \lambda$ – главные радиусы кривизны в такой точке поверхности, от которой отраженные волны распространяются навстречу падающим. Такую точку принято называть “блестящей”.

ЭПР линейного электрического вибратора. Многие объекты радиолокационного наблюдения содержат различные металлические стержни, которые принципиально могут играть роль электрических вибраторов (не говоря уже о самих вибраторных антеннах).

Получим выражение для ЭПР линейного электрического вибратора, облучаемого параллельно поляризованной волной по направлению, составляющему угол α с нормалью к оси вибратора. В этом случае в вибраторе возбуждается электрический ток $I = E_1 h_d \frac{\cos \alpha}{Z_{вх}}$, где h_d – действующая высота, $Z_{вх}$ – входное сопротивление вибратора. Вследствие протекания тока возникает вторичное излучение, модуль напряженности электрического поля которого в дальней зоне на расстоянии $R \gg \lambda$ определяется выражением [1, 10]:

$$E_2 = 60\pi \frac{I h_d \cos \alpha}{R \lambda} = 60\pi \frac{E_1 h_d^2 \cdot \cos^2 \alpha}{R \lambda Z_{вх}}; \text{ тогда}$$

$$\sigma_v = 4\pi R^2 (E_2^2 / E_1^2) = 4\pi^3 3600 h_d^2 \cos^4 \alpha / (\lambda^2 Z_{вх}^2). \quad (7.4)$$

Для полуволнового вибратора (длина которого $l = \lambda/2$) $h_d = \lambda/\pi$ и $Z_{вх} = R_{\Sigma} \approx 73$ Ом, поэтому $\sigma_v = 0,86 \lambda^2 \cos^4 \alpha$. ЭПР максимальна, когда вектор \vec{E}_1 падающей волны параллелен оси вибратора ($\alpha = 0$) и равна нулю, когда \vec{E}_1 перпендикулярен оси.

Множество произвольно ориентированных в пространстве вибраторов создает неполяризованное вторичное излучение, всегда принимаемое антенной РЛС, поэтому оно используется, например, для радиолокационной маскировки летательных аппаратов.

Подробный анализ ЭПР и особенностей радиолокационного наблюдения сложных, групповых, объемно- и поверхностно-распределенных объектов приведен в работах [1, 9, 15]. Отметим только, что дальность обнаружения таких объектов оценивается с учетом их усредненных по всем направлениям облучения значениям ЭПР $\bar{\sigma}$ и возможных флуктуаций параметров отраженных сигналов.

VIII. Дальность действия радиотехнических систем

§ 8.1. Дальность действия радиолиний

Дальностью действия РТС называется максимальное расстояние, на котором мощность принимаемого радиосигнала на входе приемника РТС снижается до минимально допустимого (порогового) уровня, еще достаточного для выполнения системой ее основных функций с требуемым качеством; фактически она определяется дальностью действия используемой радиолинии, включающей в себя передатчик, приемник и антенно-фидерные устройства. Таким образом, расчет дальности действия различных РТС сводится к расчету дальности действия одной из трех радиолиний: радиолинии связи, радиолинии с активным ответом и радиолинии с пассивным ответом.

Дальность действия радиолинии связи. Рассмотрим идеализированную модель радиолинии, состоящей из передатчика ПРД и приемника ПРМ радиосигнала с соответствующими антеннами $A_{и}$ и $A_{п}$ и фидерными устройствами (рис. 8.1). При этом будем считать, что в фидерных линиях передатчика и приемника обеспечен режим бегущей волны (коэффициент стоячей волны (КСВ) равен 1), потерь энергии в линиях нет (КПД=1); распространение радиоволн происходит в свободном пространстве, главные максимумы диаграмм направленности (ДН) антенн находятся на одном направлении и антенны согласованы по поляризации.

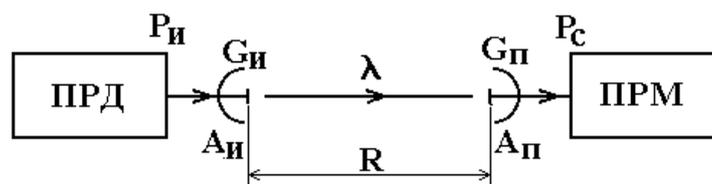


Рис.8.1. Модель радиолинии связи

Введем обозначения: R – расстояние между антеннами передатчика $A_{и}$ и приемника $A_{п}$, λ – длина волны, $P_{и}$ – мощность сигнала на выходе передатчика, $P_{с}$ – мощность сигнала на входе приемника, $G_{и}$ и $G_{п}$ – коэффициенты усиления передающей ($A_{и}$) и приемной ($A_{п}$) антенн. Выражение, определяющее дальность действия радиолинии, получим следующим образом.

Плотность потока мощности излучения в месте расположения приемной антенны [10]

$$\Pi = \frac{P_{и} G_{и}}{4\pi R^2}; \quad (8.1)$$

мощность сигнала на выходе приемной антенны (в данном случае – на входе приемника) с учетом ее эффективной площади раскрыва $S_э$ [10] $P_{с} = \Pi S_э = \Pi G_{п} \lambda^2 / 4\pi$; с учетом (8.1) пороговая мощность сигнала на входе приемника при максимальной дальности $R = R_{\text{макс}}$

$$P_{\text{смин}} = \frac{P_{\text{и}} G_{\text{и}} G_{\text{п}} \lambda^2}{(4\pi)^2 R_{\text{макс}}^2}; \quad (8.2)$$

Из (8.2) получим выражение, определяющее дальность действия идеализированной радиолинии связи в свободном пространстве:

$$R_{\text{макс}} = \sqrt{\frac{P_{\text{и}} G_{\text{и}} G_{\text{п}} \lambda^2}{(4\pi)^2 P_{\text{смин}}}}; \quad (8.3)$$

Обычно формулу (8.3) используют для грубой оценки дальности действия радиолиний.

Применительно к реальной радиолинии связи СВЧ-диапазона мощность сигнала на входе приемника более точно можно рассчитать по следующей формуле:

$$P_c = \frac{P_u \cdot e^{-L} \cdot G_u \cdot \eta_u \cdot (1 - \rho_u^2) \cdot G_n \cdot \eta_n \cdot (1 - \rho_n^2) \cdot \lambda^2}{(4\pi)^2 R^2} \quad (8.4)$$

где $\eta_{\text{и}}$, $\eta_{\text{п}}$ – коэффициенты полезного действия фидерных линий передатчика и приемника;

$\rho_{\text{и}}$ и $\rho_{\text{п}}$ – модули коэффициентов отражения по напряжению, рассчитываемые с учетом известных коэффициентов стоячей волны ($K_{\text{СВ}}$) для фидерных линий по формуле [11] $\rho = (K_{\text{СВ}} - 1) / (K_{\text{СВ}} + 1)$;

e^{-L} – множитель ослабления радиоволн на трассе распространения;

$L = \alpha R$ – коэффициент суммарного ослабления;

α – удельный коэффициент ослабления (1/м). Обычно коэффициент α , учитывающий поглощение и рассеяние радиоволн на трассе распространения, приводится в литературе в дБ/км; в этом случае в формуле (8.4) следует принять $L = 0,23\alpha R$ [9]. Заметим, что выражение (8.4) не учитывает влияния подстилающей поверхности на ДН антенн, деполяризацию и рефракцию радиоволн.

Дальность действия радиолинии с активным ответом.

Радиолиния с активным ответом состоит из двух радиолиний связи – линии запроса и линии ответа (рис. 8.2). На рис. 8.2 обозначены: ПРДЗ и ПРМЗ – передатчик и приемник запроса, ПРДО и ПРМО – передатчик и приемник ответа.

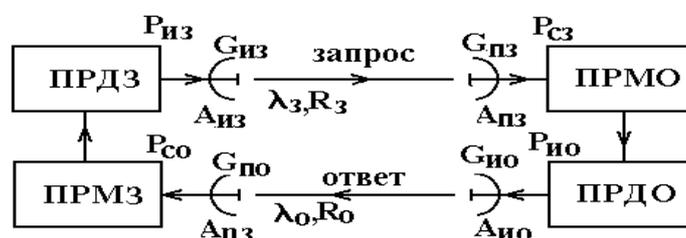


Рис. 3.2. Модель радиолинии с активным ответом

Дальность действия линий запроса $R_{\text{змакс}}$ и $R_{\text{смакс}}$ ответа можно рассчитать по формуле (8.3), присвоив параметрам $P_{\text{и}}$, $P_{\text{с}}$, $G_{\text{и}}$, $G_{\text{п}}$ и λ соответствующие дополнительные индексы (см. рис. 8.2) – “з” (запрос) и “о” (ответ):

$$R_{\text{змакс}} = \sqrt{\frac{P_{\text{из}} G_{\text{пз}} G_{\text{пз}} \lambda_{\text{з}}^2}{(4\pi)^2 P_{\text{сзмин}}}}; \quad R_{\text{омакс}} = \sqrt{\frac{P_{\text{ио}} G_{\text{по}} G_{\text{по}} \lambda_{\text{о}}^2}{(4\pi)^2 P_{\text{сомин}}}}. \quad (8.5)$$

В общем случае $R_{\text{змакс}} \neq R_{\text{омакс}}$ и дальность действия радиолинии с активным ответом $R_{\text{макс}}$ ограничена меньшей из $R_{\text{змакс}}$ и $R_{\text{омакс}}$. Обычно радиолинии стремятся сбалансировать, обеспечив $R_{\text{змакс}} = R_{\text{омакс}}$ при равенстве надежностей каналов запроса и ответа. В частном случае, когда в запросчике и ответчике для передачи приема используется по одной антенне, а длины волн радиосигналов запроса и ответа близки ($\lambda_{\text{з}} \approx \lambda_{\text{о}}$), $G_{\text{из}} \approx G_{\text{по}}$ и $G_{\text{пз}} \approx G_{\text{ио}}$, поэтому $G_{\text{из}} G_{\text{пз}} \approx G_{\text{ио}} G_{\text{по}}$, и тогда условие баланса линии можно записать в виде $\frac{P_{\text{из}}}{P_{\text{сзмин}}} \approx \frac{P_{\text{ио}}}{P_{\text{сомин}}}$. Примером радиолинии с активным ответом служит радиолиния опознавания летательных аппаратов (“свой” – “чужой”).

Дальность действия радиолинии с пассивным ответом. В данной радиолинии (например, ‘РЛС – объект’ (цель)) запросным можно считать зондирующий сигнал РЛС, а ответным – радиолокационный сигнал (отраженный объектом (целью) с ЭПР $\sigma_{\text{ц}}$), рис. 8.3.

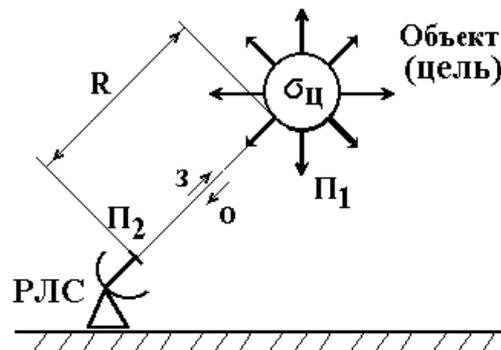


Рис. 8.3. Модель радиолинии с пассивным ответом

Используем те же обозначения, что и при анализе дальности действия радиолинии связи в свободном пространстве, считая, что РЛС имеет две антенны – передающую и приемную.

Плотность потока мощности зондирующего излучения РЛС вблизи объекта (8.1) $\Pi_1 = P_{\text{и}} G_{\text{и}} / 4\pi R^2$; мощность, перехватываемая целью (см. § 7.2), $P_1 = \Pi_1 \sigma_{\text{ц}}$; по определению ЭПР, вся эта мощность изотропно рассеивается целью и плотность потока мощности излучения, отраженного обратно к РЛС, вблизи ее приемной антенны

$$\Pi_2 = \frac{P_1}{4\pi R^2} = \frac{P_{\text{и}} G_{\text{и}} \sigma_{\text{ц}}}{(4\pi)^2 R^2}; \quad (8.6)$$

мощность сигнала на входе приемника РЛС

$$P_c = \Pi_2 S_3 = \frac{P_i G_i \sigma_c}{(4\pi)^2 R^4} \cdot \frac{G_p \lambda^2}{4\pi}; \quad (8.7)$$

тогда при $P_c = P_{\text{смин}}$ дальность действия $R_{\text{макс}}$ определяется выражением

$$R_{\text{макс}} = \sqrt[4]{\frac{P_i G_i G_p \lambda^2 \sigma_c}{(4\pi)^3 P_{\text{смин}}}}. \quad (8.8)$$

Это выражение называется **основным уравнением радиолокации** или **уравнением дальности РЛС в свободном пространстве** и отражает связь дальности действия РЛС с ее основными параметрами и ЭПР цели.

Обобщенное уравнение дальности радиолокационного наблюдения в свободном пространстве. При расчете дальности обнаружения какой-либо цели используют среднее значение ее ЭПР ($\sigma_c = \bar{\sigma}_c$), а возможные флуктуации ЭПР и параметров радиолокационного сигнала (например, амплитуды при изменении состояния атмосферы, наличии осадков и т. д.) учитывают при выборе модели принимаемого сигнала (вида флуктуации амплитуды и фазы). Пороговая мощность сигнала на входе приемника РЛС рассчитывается с учетом порогового отношения сигнал/шум $q_{\text{мин}}$, которое можно рассчитать или определить графически по характеристике обнаружения – зависимости $P_o = f(q)$ при $P_d = \text{const}$ (§ 3.2) с учетом требуемых значений вероятности обнаружения P_o и ложной тревоги P_d . В случае действия на входе приемника помехи с равномерной односторонней спектральной плотностью мощности (СПМ) N_o

$$q_{\text{мин}}^2 = \frac{2E_c}{N_o} = \frac{2P_c T}{N_o}; \quad (8.9)$$

где T – длительность принимаемого сигнала; тогда

$$P_{\text{смин}} = \frac{q_{\text{мин}}^2 N_o}{2T}. \quad (8.10)$$

Потери энергии в фидерных трактах приемника и передатчика, а также собственно в приемнике, в том числе возникающие из-за неоптимальности обработки сигнала, учитываются введением общего коэффициента потерь $K > 1$, на который следует умножить $P_{\text{смин}}$ в уравнении (8.8).

С учетом сказанного, основное уравнение радиолокации (8.8) можно записать в следующем виде

$$R_{\max} = \sqrt[4]{\frac{P_{\text{и}} G_{\text{и}} G_{\text{п}} \lambda^2 \sigma_{\text{ц}}}{(4\pi)^3 P_{\text{смин}}}} = \sqrt[4]{\frac{2P_{\text{и}} T G_{\text{и}} G_{\text{п}} \lambda^2 \sigma_{\text{ц}}}{(4\pi)^3 q_{\text{мин}}^2 N_{\text{o}} K}} = \sqrt[4]{\frac{2E_{\text{и}} G_{\text{и}} G_{\text{п}} \lambda^2 \sigma_{\text{ц}}}{(4\pi)^3 q_{\text{смин}}^2 N_{\text{o}} K}}; \quad (8.11)$$

где $E_{\text{и}}=P_{\text{и}}T$ – энергия зондирующего импульса РЛС.

Уравнение дальности в форме (8.11) называется **обобщенным уравнением дальности** или **обобщенным уравнением радиолокации**.

В случае, когда учитывается только мешающее действие шумов антенны $P_{\text{ША}}$ и собственных шумов приемника с приведенной ко входу мощностью шумов $P_{\text{шп}}$, суммарная мощность на входе приемника

$P_{\text{ш}} = P_{\text{ША}} + P_{\text{шп}}$. Мощность шумов антенны с известной температурой $T_{\text{А}}$ в полосе пропускания приемника Δf [1,11] $P_{\text{ША}} = kT_{\text{А}}\Delta f$, где $k=1,38 \cdot 10^{-23}$ Дж/К – постоянная Больцмана.

Часто при расчетах СПМ шума N_{o} ее выражают через шумовую температуру $T_{\text{ш}} = T_{\text{А}} + T_{\text{o}}(k_{\text{ш}} - 1)$; где $k_{\text{ш}} = 1 + P_{\text{шп}}/P_{\text{ША}}$ – коэффициент шума приемника; $T_{\text{o}} = 290$ К; тогда [1]

$$P_{\text{ш}} = kT_{\text{А}}\Delta f + kT_{\text{o}}\Delta f(k_{\text{ш}} - 1) = k\Delta f[T_{\text{А}} + T_{\text{o}}(k_{\text{ш}} - 1)] = k\Delta fT_{\text{ш}}; \quad (8.12)$$

и в предположении равномерности мощности шума в полосе Δf его СПМ

$N_{\text{o}} = \frac{P_{\text{ш}}}{\Delta f} = kT_{\text{ш}}$ и (8.11) можно записать следующим образом:

$$R_{\max} = \sqrt[4]{\frac{2E_{\text{и}} G_{\text{и}} G_{\text{п}} \lambda^2 \sigma_{\text{ц}}}{(4\pi)^3 q_{\text{мин}}^2 kT_{\text{ш}} K}}. \quad (8.13)$$

Таким образом, дальность действия РЛС в свободном пространстве в наибольшей степени зависит от энергии импульса излучения $E_{\text{и}}$ и от коэффициентов усиления передающей $G_{\text{и}}$ и приемной $G_{\text{п}}$ антенн, а также от ЭПР объекта $\sigma_{\text{ц}}$.

§ 8.2. Влияние условий распространения радиоволн на дальность действия РТС

Существенное влияние на дальность действия РТС оказывают поглощение, рефракция, дифракция и деполяризация радиоволн на трассе распространения, подробно рассмотренные, например, в работах [7, 8, 12]. В данном разделе обратим внимание лишь на наиболее важные последствия названных явлений.

Распространение радиоволн в тропосфере (нижнем слое атмосферы) сопровождается их затуханием, главным образом, вследствие резонансного поглощения молекулами кислорода и воды, а также рассеяния гидрометеорами (дождь, туман, снег и т.п.) и различными

аэрозолями (частицами пыли, дыма и др.). Затухание радиоволн в соответствии с законом Бугера описывается выражением [1, 7]

$$P_{cз} = P_c e^{-L} = P_c e^{-\alpha R}, \quad (8.14)$$

где $P_{cз}$ – мощность излучения в точке приема с учетом ослабления;

P_c – мощность излучения без учета ослабления (см. 8.2);

$L = \alpha R$ – коэффициент суммарного ослабления;

α (1/м) – удельный коэффициент ослабления;

R – протяженность трассы.

Для радиолиний с пассивным ответом, используемых в РЛС, с учетом двукратного прохождения радиоволн выражение (8.14) записывается в виде

$$P_{cз} = P_c e^{-2\alpha R}. \quad (8.15)$$

В справочной литературе α часто приводится в дБ/км, при этом (8.15) представляется как $P_{cз} = P_c \exp(-0,46\alpha R)$ [1].

В результате несложных рассуждений для расчета максимальной дальности действия РЛС с учетом затухания радиоволн (8.14) можно получить трансцендентное уравнение вида [1]

$$R_{максз} = R_{макс} \cdot \exp(-\alpha R_{максз}), \quad (8.16)$$

где $R_{макс}$ – дальность действия РЛС в свободном пространстве.

Уравнение (8.16) относительно $R_{максз}$ можно решить численным методом, либо графически, после преобразования к виду

$$\alpha R_{максз} = 2 \left(\frac{1}{\gamma} \right) \ln \left(\frac{1}{\gamma} \right), \quad (8.17)$$

где $\gamma = R_{максз} / R_{макс}$.

По графику зависимости $\gamma = f(\alpha R_{максз})$ приведенному, например, в [1], по заданным α и $R_{макс}$ можно определить γ и, следовательно, $R_{максз} = \gamma R_{макс}$.

Рефракция радиоволн в атмосфере (криволинейная траектория их распространения в неоднородной среде) может приводить как к уменьшению $R_{максз}$, так и к ее увеличению.

Так, например, при отсутствии рефракции радиоволн дальность действия многих радиолиний вследствие кривизны поверхности Земли ограничена дальностью прямой видимости (дальностью горизонта) $R_r \approx 3,57 \sqrt{h_1 + h_2}$ – км; $h_1, 2$ (м) – высоты подъема передающей и приемной антенн. Это характерно для радиолиний систем связи и локации

низкорасположенных объектов в диапазонах электромагнитных волн от метрового вплоть до оптического.

В большинстве случаев в тропосфере имеет место положительная рефракция радиоволн (траектория радиолуча отклоняется вверх от поверхности Земли), приводящая к увеличению дальности горизонта до $R_r \approx 4,12 \sqrt{h_1 + \sqrt{h_2}}$, и, соответственно, дальности действия РТС [1, 7].

Влияние подстилающей поверхности на дальность действия радиолучей необходимо рассматривать в каждом случае с учетом рабочего диапазона частот и комплексной диэлектрической проницаемости почвы, воды и др. [9].

§ 8.3. Влияние отражения радиоволн от земной поверхности на дальность действия РЛС

Во многих реальных ситуациях при использовании различных радиолучей высоты расположения передающих и приемных антенн (или объектов радиолокационного наблюдения) таковы, что даже в случае применения остронаправленных антенн в пункт приема приходят не только прямые радиоволны, но и рассеянные (отраженные) земной поверхностью. В результате их интерференции величина суммарной напряженности электрического поля, а, значит, и мощности сигнала на входе приемника может оказаться меньше порогового уровня ($P_c < P_{смин.}$).

Рассмотрим общий случай радиолокационного наблюдения низколетящих целей (НЛЦ), используя рис. 8.4:

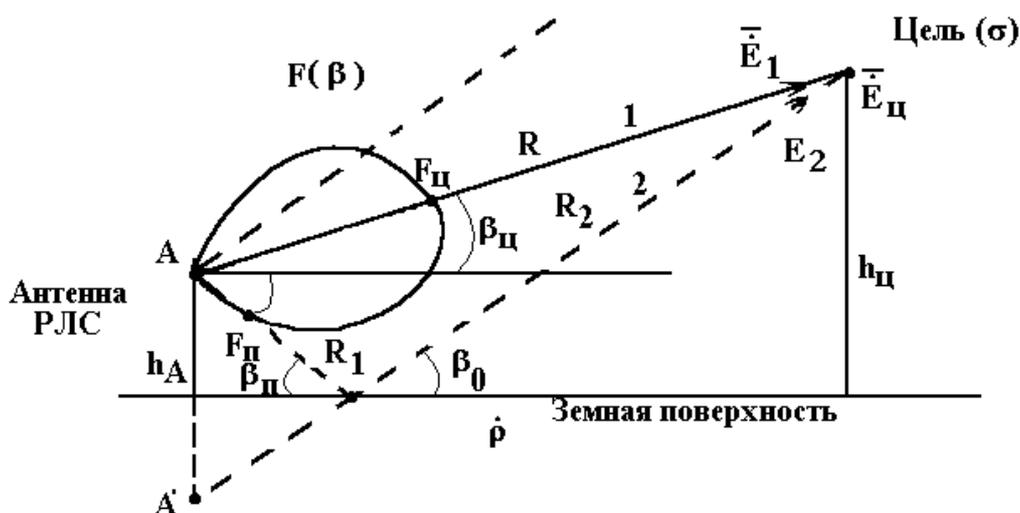


Рис. 8.4. Влияние отражения радиоволн на дальность действия РЛС

Для приближенного анализа воспользуемся лучевой моделью радиолучей, что допустимо, если высоты $h_A, h_Ц > \lambda$ и $h_A, h_Ц \ll R$ [7]. В этом случае можно считать, что напряженность электрического поля вблизи цели определяется скалярной суммой $\vec{E}_ц = \vec{E}_1 + \vec{E}_2$, причем луч 2 исходит от воображаемого источника A' (зеркального отображения антенны РЛС относительно земной поверхности). Считая лучи 1 и 2 параллельными,

примем, что угол места цели $\beta_{ц}$ и углы $\beta_{п}$ и β_0 равны ($\beta_{ц} = \beta_{п} = \beta_0$), разность фаз прямой и отраженной волн, обусловленная разностью расстояний $\Delta R = R_1 + R_2 - R = 2h_A \sin \beta_{ц}$ составляет величину $\Delta \varphi = 2\pi \Delta R / \lambda = (4\pi h_A / \lambda) \sin \beta_{ц}$. Для общности предположим, что РЛС и цель неподвижны, направление максимума ДН приемопередающей антенны РЛС (по напряженности поля) $F(\beta)$ в вертикальной плоскости не совпадает с направлением на цель; значения ДН в направлении цели и земной поверхности обозначим $F_{ц}$ и $F_{п}$; мощность передатчика РЛС – $P_{и}$, коэффициент усиления антенны – G ; ЭПР цели – σ , комплексный коэффициент отражения земной поверхности – $\dot{\rho}$. Запишем выражения для напряженностей полей \dot{E}_1 и \dot{E}_2 с учетом перечисленных параметров [10]:

$$\dot{E}_1 = \frac{\sqrt{60P_{и}G}}{R} e^{-jkR} F_{ц}; \quad \dot{E}_2 = \frac{\sqrt{60P_{и}G}}{R_1 + R_2} e^{-jk(R_1+R_2)} \dot{\rho} F_{п};$$

тогда $\dot{E}_{ц}$ описывается следующим выражением (промежуточные выкладки опустим):

$$\dot{E}_{ц} = \dot{E}_1 + \dot{E}_2 = \dot{E}_1 \left[1 + \rho_{об} e^{-j(\varphi + \Delta\varphi)} \right] = \dot{E}_1 \Phi_{ц}, \quad (8.18)$$

где $\rho_{об} = \frac{R}{R + \Delta R} \frac{F_{п}}{F_{ц}} \dot{\rho}$ – обобщенный коэффициент отражения, учитывающий как свойства земной поверхности, так и форму ДН антенны РЛС; ρ , φ – модуль и аргумент $\dot{\rho}$.

Действующее значение напряженности поля $E_{ц}$, полученное из (8.18), равно $E_{ц} = E_1 \Phi_{ц}$, где

$\Phi_{ц} = \sqrt{1 + \rho_{об}^2 + 2\rho_{об} \cos(\varphi + \Delta\varphi)} = \sqrt{1 + \rho_{об}^2 + 2\rho_{об} \cos \left[\varphi + (4\pi h_A / \lambda) \sin \beta_{ц} \right]}$ – интерференционный множитель.

Можно показать, что мощность радиолокационного сигнала на входе приемника РЛС с учетом $\Phi_{ц}$ описывается выражением

$$P_c = \frac{P_{и} G^2 F_{ц}^4 \sigma}{(4\pi)^3 R^4} \Phi_{ц}^4, \quad (8.19)$$

из которого дальность действия РЛС с учетом интерференции волн определяется как

$$R_{макс0} = \sqrt[4]{\frac{P_{и} G^2 F_{ц}^4 \lambda^2 \sigma}{(4\pi)^3 P_{смин}}} \Phi_{ц} = R_{макс} \Phi_{ц}, \quad (8.20)$$

где $R_{макс}$ – дальность действия РЛС без учета влияния Земли (см. (8.8)). Если ДН антенны РЛС направлена точно на цель, $F_{ц}=1$.

Необходимо помнить, что **при движении цели** ее угол места $\beta_{ц}$ и высота $h_{ц}$ могут изменяться, следовательно, в выражениях (8.19) и (8.20) подразумеваются $F_{ц}=f(\beta_{ц})$, $\Phi_{ц}=f(\beta_{ц}, h_{ц})$.

При изменении угла места цели оказывается, что $\Phi_{ц}(\beta_{ц\text{мин}})=1-\rho_{об}$, $\Phi_{ц}(\beta_{ц\text{макс}})=1+\rho_{об}$. Таким образом, поскольку $\rho_{об\text{макс}}=1$, дальность радиолокационного наблюдения низкорасположенных целей может изменяться от 0 до удвоенного значения $R_{\text{макс}}$, рассчитанного без учета влияния Земли. Соответственно, диаграмма видимости (ДВ) РЛС в вертикальной плоскости с учетом интерференционного множителя становится многолепестковой, даже если антенна РЛС не обладает направленностью в указанной плоскости. Если цель находится на направлении, совпадающем с направлением какого-либо минимума ДВ, то она не будет обнаружена.

Рассмотрим конкретный пример. Пусть антенна РЛС не обладает направленностью в вертикальной плоскости т.е. $F_{ц}=1$, угол места цели мал, тогда $\sin\beta_{ц}\approx h_{ц}/R$, $|\rho|=1$, $\varphi=\pi$, и интерференционный множитель записывается в виде $\Phi_{ц}\approx 4\pi h_{ц}h_{ц}/(\lambda R)$ [9]. Из выражения (8.20) получим формулу для расчета дальности обнаружения НЛЦ:

$$R_{\text{макс}} = \sqrt[8]{\frac{4\pi P_{и} G^2 h_{А}^4 h_{ц}^4 \sigma}{P_{\text{смин}} \lambda^2}}. \quad (8.21)$$

Расстояние от РЛС до цели, при котором она считается низколетящей, определяется неравенством $R \geq 10\pi h_{ц}h_{ц}/\lambda$ [9].

Из (8.21) следует, что для увеличения дальности обнаружения НЛЦ вдвое мощность передатчика РЛС необходимо увеличить в $2^8=256$ раз. Детальный анализ ситуаций, возникающих при радиолокационном наблюдении НЛЦ, приведен, например, в работе [9].

IX. Поиск объектов в зоне действия РЛС

§ 9.1. Радиолокационный обзор пространства

В РЛС различного назначения поиск сигналов, отраженных от объектов, предшествует процедуре измерения их параметров. Рассмотрим общий пример, включающий в себя элементы поиска, присущие всем РТС – поиск сигналов в рабочей зоне РЛС.

Размеры рабочей зоны (зоны обзора) обычно заданы предельными значениями измеряемых координат и скоростей объекта, т.е. дальности ($R_{\text{мин}} - R_{\text{макс}}$), азимута ($\alpha_{\text{мин}} - \alpha_{\text{макс}}$), угла места ($\beta_{\text{мин}} - \beta_{\text{макс}}$) и радиальной скорости ($V_{\text{гмин}} - V_{\text{гмакс}}$). Величины этих интервалов можно выразить через количество содержащихся в них элементов разрешения по дальности δR , азимуту $\delta\alpha$, углу места $\delta\beta$ и радиальной скорости δV_r :

$$N_R = \frac{R_{\max} - R_{\min}}{\delta R}; N_\alpha = \frac{\alpha_{\max} - \alpha_{\min}}{\delta \alpha}; N_\beta = \frac{\beta_{\max} - \beta_{\min}}{\delta \beta}; N_V = \frac{V_{\max} - V_{\min}}{\delta V_r}; \quad (9.1)$$

При обзоре проверяется наличие объекта в каждом элементе разрешения по методу, выбор которого зависит от назначения и ТТХ РТС.

Эффективность метода обзора характеризуется средним временем до обнаружения объекта и средним временем между соседними ложными тревогами (усредненной частотой ложных тревог).

Обзор элементов разрешения может быть **последовательным** во времени или **параллельным** (одновременным). Иногда применяют комбинированный обзор – параллельно-последовательный.

При параллельном обзоре поиск сигнала происходит одновременно во всех элементах разрешения и обнаружение объекта происходит сразу при его появлении в зоне обзора; однако для этого требуется весьма сложное оборудование.

Обзор по дальности происходит при распространении сигнала до объекта и обратно. При максимальной дальности объекта время задержки приема сигнала $\tau_{R_{\max}} = 2R_{\max}/c$. За это время при определенной ориентации ДН антенны РЛС должны быть просмотрены все N_R элементов разрешения по дальности.

Обзор по радиальной скорости может быть параллельным (при наличии в приемнике N_V частотных каналов (фильтров) или последовательным – при наличии одного перестраиваемого в диапазоне ($F_{V_{\min}} - F_{V_{\max}}$) канала с полосой пропускания, не превышающей требуемого элемента разрешения по частоте δF_V при последовательном обзоре (§§ 4.4, 6.1).

Последовательный обзор по дальности и скорости обычно сопряжен с энергетическими потерями и используется при сравнительно высоких отношениях сигнал/шум на входе приемника (например, в РЛС с активным ответом).

Обзор по угловым координатам также может быть параллельным, последовательным или параллельно-последовательным. При параллельном обзоре в РЛС должно быть $N_{\alpha\beta} = N_\alpha \cdot N_\beta$ угловых каналов, т.е. необходима антенна с $N_{\alpha\beta}$ -лучевой ДН, перекрывающей всю зону обзора, и приемник с таким же количеством каналов. Однако при широких секторах обзора и высокой требуемой разрешающей способности по угловым координатам $N_{\alpha\beta}$ становится слишком большим и реализация такой системы затруднена. В этом случае производится одноканальный последовательный обзор (или параллельно-последовательный) со сканированием ДН по всей зоне. Естественно, что за простоту технической реализации при этом приходится платить резким увеличением времени однократного обзора всей зоны, которое не должно быть менее величины

$$T_0 \geq \tau_{R_{\max}} \cdot N_{\alpha\beta} = \frac{2 \cdot R_{\max}}{c} N_{\alpha\beta} \quad (8.2)$$

Обнаружение слабых сигналов требует накопления их энергии, что заставляет увеличивать T_0 : например, если в импульсной РЛС принимается пакет из N импульсов, T_0 возрастет в N раз, т.к. период T_{Π} их повторения задан условием однозначного измерения дальности $T_{\Pi} \geq \tau_{R_{\max}}$ [1].

Упомянутые выше ограничения могут быть смягчены при программируемом по априорным данным о вероятном направлении на объект или адаптивном обзоре. При адаптивном последовательном обзоре по результатам анализа этапов, предшествующих текущему, может изменяться очередность и время обзора различных элементов разрешения, а также энергия излучения. При параллельном адаптивном обзоре автоматически увеличивается энергия излучения РЛС, если при работе в обычном (дежурном) режиме ожидается появление объекта в рабочей зоне РЛС.

§ 8.2. Методы последовательного обзора пространства

Чаще всего в РЛС используют одноканальный последовательный обзор со сканированием ДН антенны в пределах интересующего участка пространства. Основным параметром при этом является время **обзора заданного сектора**, связанное с дальностью действия, точностью и разрешающей способностью РЛС. Так, из основного уравнения радиолокации (8.8) следует, что увеличение дальности действия при ограничении энергии зондирующего сигнала и чувствительности приемника возможно лишь за счет увеличения коэффициентов усиления антенн РЛС, а следовательно, за счет сужения их ДН. Повышение разрешающей способности по угловым координатам также достигается сужением ДН антенн, что приводит к увеличению требуемого времени обзора T_0 заданного объема пространства.

Обозначим сектор обзора в горизонтальной плоскости $\Theta_{\alpha} = \alpha_{\max} - \alpha_{\min}$, в вертикальной - $\Theta_{\beta} = \beta_{\max} - \beta_{\min}$, ширину ДН по уровню половинной мощности, соответственно, α_A и β_A . Считая, что $\delta\alpha \approx \alpha_A$ и $\delta\beta \approx \beta_A$, неравенство (9.2) можно записать в следующем виде:

$$T_0 \geq \frac{2 \cdot R_{\max}}{c} \cdot \frac{\Theta_{\alpha}}{\alpha_A} \cdot \frac{\Theta_{\beta}}{\beta_A} \quad (9.3)$$

Если в процессе обзора при нахождении ДН антенны в каком-либо разрешаемом угловом положении необходимо принять не один, а N импульсов отраженного сигнала, скорость сканирования ДН должна быть снижена в N раз и тогда время обзора увеличится до величины, определяемой неравенством:

$$T_{0N} \geq T_0 \cdot N = \frac{2 \cdot R_{\max}}{c} \cdot \frac{\Theta_\alpha}{\alpha_A} \cdot \frac{\Theta_\beta}{\beta_A} \cdot N \quad (9.4)$$

При этом считается, что сканирование ДН равномерно и условия просмотра всех элементов разрешения одинаковы.

Наибольшее распространение в РЛС получили следующие методы равномерного обзора пространства: круговой, секторный, винтовой, спиральный, конический и строчно-кадровый (растровый).

Круговой и секторный обзор. Ширина ДН в вертикальной плоскости β_A полностью перекрывает заданный сектор Θ_β . При вращении ДН в горизонтальной плоскости (например, путем вращения самой антенны вокруг вертикальной оси или при однокоординатном сканировании ДН кольцевой ФАР) происходит последовательный просмотр сектора $\Theta_\alpha = 2 \cdot \pi$ рад., т.е. **круговой обзор**. Если $\Theta_\alpha < 2 \cdot \pi$, то обзор называется **секторным**. Время (период) кругового обзора

$$T_{\text{ок}} = \frac{2 \cdot \pi}{\Omega_A}, \text{ секторного обзора} - T_{\text{ос}} = \frac{\Theta_\alpha}{\Omega_A} \cdot k_c, \quad \Omega_A - \text{угловая скорость}$$

сканирования ДН (вращения антенны); k_c – коэффициент потерь времени при изменении направления отклонения ДН (при "обратном" ходе). Для определения $T_{\text{ос}}$ и $T_{\text{ок}}$ рассчитаем угловую скорость Ω_A . При ширине ДН α_A и количестве импульсов N , принятых за время поворота ДН на угол

$$\alpha_A \quad \Omega_A \leq \frac{\alpha_A}{\tau_{R_{\max}}} \cdot N; \text{ используя условия однозначного измерения дальности}$$

$\tau_{R_{\max}} \leq T_{\text{п}}$, можно считать, что $\Omega_A \leq \alpha_A F_{\text{п}} / N$ и тогда время кругового обзора должно быть не менее

$$T_{\text{ок}} \geq \frac{2\pi \cdot N}{\alpha_A \cdot F_{\text{п}}}, \quad (9.5)$$

секторного –

$$T_{\text{ос}} \geq \frac{\Theta_\alpha}{\alpha_A} \cdot \frac{N \cdot k_c}{F_{\text{п}}}. \quad (9.6)$$

В одноканальных обзорных РЛС при использовании антенны с однолучевой ДН при круговом или секторном обзоре определяются дальность и азимут объекта.

При необходимости измерения третьей координаты – угла места объекта – применяют рассматриваемые ниже методы обзора.

Винтовой обзор. При винтовом обзоре ДН антенны сканирует в пространстве по винтовой траектории, представляющей собой комбинацию кругового вращения по азимуту с угловой скоростью Ω_α и подъема по углу места со скоростью Ω_β (рис. 9.1). Обратный ход по β обычно происходит очень быстро. Важное значение имеет не только

согласование скорости вращения Ω_α с частотой повторения импульсов F_Π и шириной ДН α_A , но и обеспечение перекрытия ДН на смежных витках при отклонении по углу места.

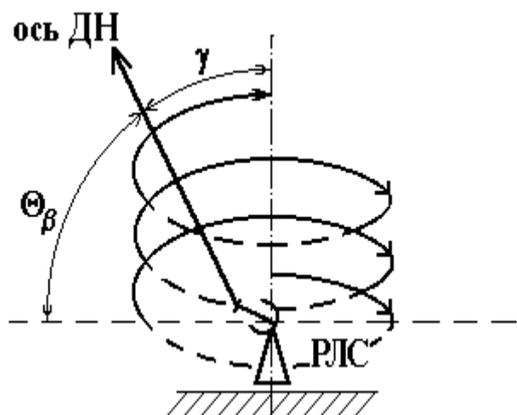


Рис. 9.1. Винтовой обзор

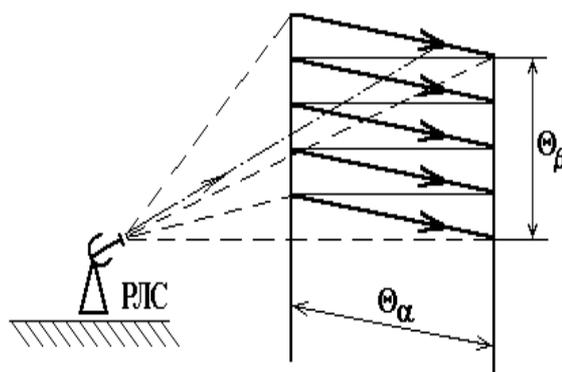


Рис. 9.2. Строчно-кадровый обзор

Для исключения пропусков объектов обеспечивается двойное перекрытие ДН: за один поворот по азимуту ДН смещается по углу места на величину $\Delta\Theta = \beta_A/2$. Поскольку на один оборот ДН по азимуту требуется время $t_1 = 2\pi/\Omega_\alpha$, скорость отклонения по углу места с учетом двойного перекрытия

$$\Omega_\beta = \frac{\Delta\Theta}{t_1} \cdot \frac{\beta_A}{2 \cdot t_1} = \beta_A \frac{\Omega_\alpha}{4\pi} \quad (9.7)$$

и при требуемом секторе Θ_β по углу места время винтового обзора составляет

$$T_{\text{об}} = \frac{\Theta_\beta}{\Omega_\beta} = \frac{4\pi \cdot \Theta_\beta}{\beta_A \cdot \Omega_\alpha} \quad (9.8)$$

Подставив в (9.8) значение $\Omega_\alpha \leq \alpha_A F_\Pi / N$, получим

$$T_{\text{об}} \geq \frac{4\pi \cdot \Theta_\beta \cdot N}{\alpha_A \cdot \beta_A \cdot F_\Pi} \quad (9.9)$$

Из (9.9) следует, при узких ДН антенн и широком секторе обзора информация об объекте поступает с большой дискретностью, что, в первую очередь, приводит к снижению точности измерения параметров движения объекта.

Спиральный обзор. Как и винтовой, спиральный обзор представляет собой комбинацию вращения ДН с одновременным изменением угла γ между осью вращения и осью ДН (рис. 9.1). От

винтового отличается лишь тем, что угол $\gamma \leq 45^0$ (при винтовом обзоре $\gamma \geq 45^0$).

Пусть секторы обзора в горизонтальной и вертикальной плоскостях равны, т.е. $\Theta_\alpha = \Theta_\beta = \Theta$ (зона обзора ограничена конической поверхностью); Ω_α и Ω_β – угловые скорости вращения вдоль витка спирали и, соответственно, отклонения от оси вращения. При двойном перекрытии витков спирали в случае симметричной ДН ($\alpha_A = \beta_A$)

$$T_{\text{осн}} \geq \frac{2\pi \cdot \Theta \cdot N \cdot \sin \gamma}{\alpha_A^2 \cdot F_{\text{п}}} \quad (9.10)$$

Конический обзор. Представляет собой частный случай спирального, угол γ постоянен и обычно выбирается не превышающим $\alpha_A/2$, т.е. сектор обзора $\Theta = \alpha_A + 2\gamma$. Поскольку обзор происходит за один оборот ДН,

$$T_{\text{окон}} \geq \frac{2\pi \cdot N \cdot \sin \gamma}{\alpha_A \cdot F_{\text{п}}} \quad (9.11)$$

Конический обзор используется в устройствах точного автоматического измерения угловых координат равносигнальным методом и автоматического слежения за объектом (целью) (см. §10.4).

Строчно-кадровый (растровый) обзор. ДН антенны сканирует подобно электронному лучу в приемной телевизионной трубке (кинескопе) по строкам и кадрам (рис. 9.2).

Время обзора с учетом (9.2) определяется неравенством:

$$T_{\text{оск}} \geq \frac{2 \cdot \Theta_\alpha \cdot \Theta_\beta \cdot N}{\alpha_A \cdot \beta_A \cdot F_{\text{п}}} \quad (9.12)$$

Сканирование ДН может быть как чисто механическим (путем изменения положения антенны или ее элементов) или чисто электронным (фазовым, коммутационным или частотным), так и комбинированным, сочетающим, например, механическое по азимуту и электронное по углу места. Предельная скорость обзора при механическом сканировании ограничена инерционностью электроприводов, при электронном – быстродействием фазовращателей, переключателей и управляемых генераторов СВЧ. Так как время переключения состояния ферритовых и р-и-п-диодных фазовращателей СВЧ в среднем составляет десятые доли – единицы микросекунд, время переориентации ДН фазированной антенной решетки (ФАР) в соседнее положение происходит практически мгновенно по сравнению с механическим перемещением ДН. Тем не менее, время одноканального обзора в конечном итоге ограничено условием (9.2) независимо от реализации метода обзора. Поэтому существенное

сокращение времени поиска может быть достигнуто лишь при использовании параллельного (многоканального) обзора.

§ 9.3. Параллельный и управляемый обзор пространства

Во многих современных обзорных РЛС высокая точность и разрешающая способность в сочетании с большой дальностью действия и малым временем обзора обеспечивается, например, комбинацией последовательного обзора по азимуту и параллельного (многоканального) по углу места. Для этого используются антенны с многолучевыми ДН в вертикальной плоскости, причем количество лучей (перекрывающихся по заданному уровню главных лепестков ДН) должно быть $N_{\beta} = \Theta_{\beta} / \beta_{\Delta 1}$, а ширина каждого луча $\beta_{\Delta 1}$ выбирается с учетом требуемой разрешающей способности и точности РЛС по углу места. Многолучевые ДН могут быть сформированы в зеркальных и линзовых антеннах с соответствующим количеством облучателей, смещенных относительно фокуса параболического зеркала или линзы, в гибридных антеннах (с облучателем в виде ФАР), а также в антенных решетках, в том числе ФАР. Для обеспечения кругового обзора плоские ФАР объединяются, например, в пирамидальные конструкции [14].

В РЛС с ФАР возможен эффективный управляемый и адаптивный поиск объектов, при котором очередность просмотра различных направлений устанавливается в зависимости от результатов предшествующего зондирования рабочей зоны.

Оптимальным поиском считают такой, при котором достигается минимальное среднее время поиска объекта (среднее время пребывания необнаруженного объекта в зоне обзора).

При **адаптивном поиске** управление обзором может производиться, например, с помощью двухпорогового последовательного обнаружителя, в котором для каждого положения ДН антенны вычисляется сумма $\Lambda_{\Sigma} = \sum_{i=1}^M P_i \cdot \Lambda_i$ коэффициентов правдоподобия Λ_i для M элементов разрешения S_i ; P_i - вероятности нахождения объекта в элементе S_i . ДН перемещается в соседнее положение, если Λ_{Σ} либо превышает верхний порог $\Lambda_{\text{ов}}$ (наличие объекта в S_i), либо не достигает нижнего порога $\Lambda_{\text{он}}$ за заданное время (отсутствие объекта в S_i) [1].

Х. Измерение дальности, радиальной скорости и угловых координат объектов

§ 10.1. Методы и устройства измерения дальности объектов

Как показано в § 6.3, радиотехнические методы измерения дальности основаны либо на непосредственном измерении времени задержки (запаздывания) τ приема радиосигнала, отраженного от объекта, относительно момента излучения зондирующего радиосигнала, либо на

измерении разности фаз или частот этих сигналов, пропорциональной задержке τ . Соответственно, существуют три основных метода измерения дальности: импульсный, фазовый и частотный. Рассмотрим кратко каждый из методов.

Импульсный метод. Благодаря простоте технической реализации широко применяется в импульсных РЛС (дальность объекта рассчитывается по формуле (6.1), позволяет однозначно измерять дальность; обеспечивает достижение высокой разрешающей способности δR . Недостатки – невозможность измерения малых дальностей (см. 6.3), необходимость высокой импульсной мощности излучения в РЛС с простыми зондирующими сигналами. Последний недостаток устраняется при использовании широкополосных зондирующих сигналов и применении в приемниках РЛС согласованных фильтров, сжимающих ШПС по времени (см. разд. 1, § 5.4).

Фазовый метод. Дальность определяется по измеренной разности фаз отраженного от объекта $u_2(t)$ и зондирующего $u_1(t)$ сигналов, представляющих собой непрерывные гармонические колебания вида $u_1(t) = U_1 \cos(\omega_0 t + \varphi)$ и $u_2(t) = U_2 \cos(\omega_0(t + \tau) + \varphi + \varphi_r)$, где φ_r – сдвиг фазы при отражении от объекта. Разность фаз $u_2(t)$ и $u_1(t)$ составляет величину $\Delta\varphi = \omega_0 \tau + \varphi_r$, измерив которую, можно определить время задержки $\tau = (\Delta\varphi - \varphi_r) / \omega_0$ и рассчитать дальность объекта:

$$R = c\tau/2 = c(\Delta\varphi - \varphi_r) / (2\omega_0) = \lambda(\Delta\varphi - \varphi_r) / (4\pi). \quad (10.1)$$

Однако разность фаз $\Delta\varphi$ может многократно превышать 2π радиан, т.е. $\Delta\varphi = 2\pi n + \Delta\Phi$, $n = 0; 1; 2$, где $\Delta\Phi$ – показания фазометра. Поэтому дальность нельзя определить однозначно. Кроме того, неизвестный сдвиг фаз φ_r вносит дополнительную ошибку в результат измерения.

Для однозначного измерения дальности и исключения влияния φ_r на точность измерения можно использовать **двухчастотный фазовый метод**: в качестве зондирующих излучать два сигнала с близкими частотами $\omega_{01} \approx \omega_{02}$, а дальность объекта определять по измеренной разности фаз отраженных сигналов $\Delta\varphi_{21} = \Delta\varphi_2 - \Delta\varphi_1 = 2\pi m + \Delta\Phi_2 - 2\pi n - \Delta\Phi_1$; при $\omega_{01} \approx \omega_{02}$ можно принять $n = m$ и $\varphi_{r1} = \varphi_{r2}$, тогда

$$\Delta\varphi_{21} = \tau(\omega_{02} - \omega_{01}) = \frac{2R}{c}(\omega_{02} - \omega_{01}) = \frac{2R}{c}\omega_p$$

и дальность объекта R однозначно определяется выражением

$$R = c \frac{\Delta\varphi_{21}}{2\omega_p} = \lambda_p \frac{\Delta\varphi_{21}}{4\pi}, \quad (10.2)$$

где $\lambda_p = 2\pi c / \omega_p$ – длина волны сигнала биений (колебаний с разностной частотой ω_p); $\lambda_p \gg \lambda_{01}; \lambda_{02}$.

Таким образом, в этом случае шкалу фазометра можно проградуировать непосредственно в единицах дальности. Погрешность измерения дальности при известной дисперсии оценки фазы фазометром δ_φ^2 можно рассчитать как $\delta_R = \frac{\delta_\varphi \lambda_p}{4\pi}$ [15].

Важным достоинством фазового метода измерения дальности является наименьшая инструментальная погрешность; наиболее существенный недостаток – невозможность измерения дальности многих объектов из-за отсутствия разрешения по дальности. Второй недостаток – высокие требования к развязке передающего и приемного каналов.

Частотный метод. Дальность определяется по измеренной частоте биений зондирующего и отраженного от объекта сигналов, пропорциональной τ . Применение частотной модуляции (ЧМ) зондирующего сигнала позволяет реализовать дальномер с непрерывным излучением, обеспечивающий высокую δR и точность измерения дальности. Частотный метод удобен и тем, что позволяет одновременно измерять радиальную компоненту V_r скорости движения объекта.

Рассмотрим случай, когда объект либо неподвижен, либо движется со скоростью $V_r \ll c$. Допустим, что частота f_3 зондирующего сигнала непрерывно возрастает по линейному закону со скоростью $\gamma = df_3(t)/dt$, тогда приращение частоты f_3 за время $\tau = 2R/c$ распространения сигнала от дальномера до объекта и обратно составит $F_R = \gamma\tau = 2\gamma R/c$ и, измерив F_R , можно определить дальность объекта $R = cF_R/(2\gamma)$. В реальных устройствах непрерывное линейное изменение частоты практически неосуществимо, поэтому обычно в частотных дальномерах используют симметричную (СЛЧМ), несимметричную (НЛЧМ) линейную ЧМ, либо гармоническую ЧМ с ограниченной девиацией частоты.

Рассмотрим принцип измерения дальности неподвижного объекта дальномером с СЛЧМ, упрощенная структурная схема приведена на рис. 10.1.

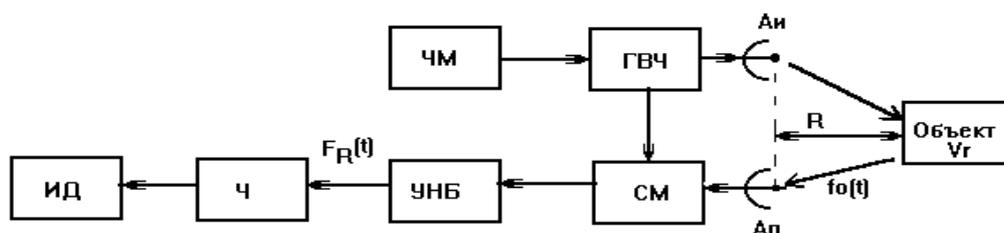


Рис. 10.1. Структурная схема частотного дальномера

В состав дальномера (рис. 10.1) входят: передающая и приемная антенны $A_{и}$ и $A_{п}$, генератор высокой частоты (ГВЧ) одновременно выполняющий функции когерентного гетеродина, частотный модулятор ЧМ, смеситель СМ, усилитель напряжения биений (сигнала разностной частоты), частотомер Ч и индикатор дальности ИД. Обозначим среднее значение несущей частоты f_0 , девиацию Δf , период модуляции T_m ;

допустим, что $\tau=2R/c \ll T_M$, тогда $F_R \gg F_M = 1/T_M$. Временные диаграммы, поясняющие принцип измерения, приведены на рис. 10.2.

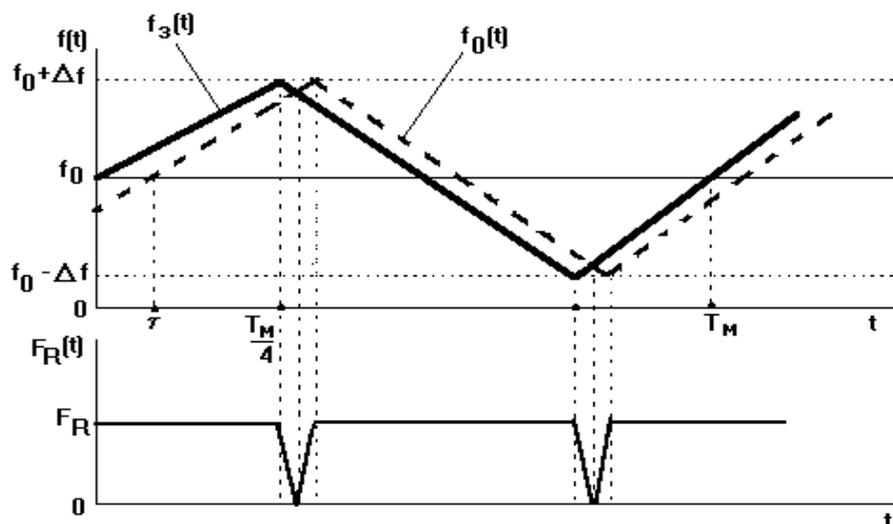


Рис. 10.2. Временные диаграммы, поясняющие принцип измерения дальности

Для большей наглядности будем считать, что при $t=0$ $f_3(t)=f_0$ и на интервале $t \in [0; T_M/4]$ $f_3(t)=f_0+4\Delta f t/T_M=f_0+\gamma t/T_M$. Тогда за время задержки принятого сигнала, отраженного от объекта, приращение f_3 составит $F_R=\gamma\tau=4\Delta f\tau/T_M=8\Delta f R/(cT_M)$. Таким образом, на указанном интервале $F_R(t)=F_R=\text{const}$ и прямо пропорциональна дальности объекта R . На самом деле облучение объекта обычно происходит в течение произвольного интервала времени $\Delta t > T_M$, поэтому в зависимости $F_R(t)$ появляются нулевые провалы (рис. 10.2) и средняя частота сигнала биений на выходе смесителя $F_{Rcp} \neq F_R$. Однако при $\tau \ll T_M$ можно считать, что $F_{Rcp} \approx F_R$ и, измерив F_R , можно определить дальность объекта R :

$$R = \frac{cF_R T_M}{8\Delta f} = \frac{cF_R}{8\Delta f F_M}. \quad (10.3)$$

В простейших дальномерах в качестве частотомера используется счетчик импульсов (числа биений n_R за период модуляции T_M) $n_R=F_R T_M=8\Delta f R/c$, поэтому дальность измеряется с дискретом ΔR , определяемым из выражения $n_R \pm 1 = 8\Delta f (R \pm \Delta R)/c$ [1]:

$$\Delta R = c/(8\Delta f). \quad (10.4)$$

Как видно из (10.4), уменьшения ΔR и, следовательно, увеличения точности измерения дальности, можно достичь путем увеличения девиации частоты Δf . По ряду причин это не всегда возможно и для повышения точности измерений применяют усложненные системы, например, с двойной ЧМ. Главная причина возникновения ошибки ΔR кроется в дискретности спектра периодических ЧМ сигналов. Спектр

биений оказывается тоже дискретным, состоящим из компонент с частотами $F_k = kF_m$, $k=1; 2; 3; \dots$. Поэтому при измерении фактически определяется частота F_k , ближайшая к F_R , связанной с дальностью R выражением (10.3). Таким образом, минимальная дальность, измеряемая частотным дальномером, определяется минимальной частотой спектра биений $F_k = F_m$: $R_{\min} = c / (8\Delta f)$ [1].

Разрешающая способность δR частотного дальмера определяется шириной спектра используемого ЧМ сигнала, который в первом приближении можно принять равным удвоенной девиации частоты $2\Delta f$; таким образом, $\delta R = c / (2\Delta f)$ [1].

§ 10.2. Методы и устройства измерения радиальной скорости движения объектов

В § 6.3 показано, что величину радиальной компоненты V_r скорости движения объекта легко определить, например, по измеренному доплеровскому смещению частоты сигнала, отраженного от объекта, причем как в непрерывном, так и в импульсном режиме работы РЛС. Простейший доплеровский измеритель скорости (ДИС) можно построить по схеме, аналогичной схеме частотного дальмера, описанной выше (рис. 10.1), исключив из нее частотный модулятор и заменив индикатор дальности индикатором скорости. Нетрудно показать, что на выходе смесителя при $V_r \neq 0$ возникнут колебания биений с разностной частотой $F_v = 4\pi f_0 V_r / c$, пропорциональной V_r [1, 15].

Очевидно, что такой ДИС, во-первых, имеет невысокую чувствительность, во-вторых, не позволяет определять знак V_r . Для устранения этих недостатков можно использовать супергетеродинный приемник с детектором нулевых биений [1].

Основные достоинства непрерывных ДИС: простота конструкции, отсутствие мертвой зоны, возможность селекции целей по скорости, не критичность к медленным изменениям несущей частоты передатчика. Недостатки непрерывных ДИС: необходимость использования отдельных антенн; низкая помехоустойчивость; высокие требования к кратковременной стабильности частоты генераторов; вторичный эффект Доплера при малых расстояниях до объекта – биения между отдельными компонентами спектра отраженного сигнала, затрудняющие измерение скорости [15]; невозможность измерения скоростей многих объектов.

Импульсные ДИС имеют следующие преимущества в сравнении с непрерывными: возможность одновременного измерения дальности объектов, применения совмещенной антенны. Недостатками импульсных ДИС являются неоднозначность измерения скорости, появление “слепых” скоростей, наличие “мертвой зоны”, сложность конструкции.

ДИС неследящего типа, реализующие оптимальные алгоритмы измерения, могут быть построены на основе многоканальных схем, содержащих набор из N_v СФ или квазисогласованных фильтров (см. § 5.4), а также одноканальных схем с перестройкой частоты вспомогательного

гетеродина в диапазоне ΔF_v . В следящих измерителях скорости используются частотные или фазовые дискриминаторы и корреляционные устройства [13, 15].

На основе эффекта Доплера работает множество современных радиоволновых извещателей систем охранной сигнализации, реагирующих на появление движущихся объектов в охраняемой зоне, а также портативные РЛС СВЧ диапазона, измеряющие скорость движения автомобилей.

§ 10.3. Совместное измерение дальности и радиальной скорости частотным методом

С помощью частотных дальномеров можно измерять и радиальную компоненту V_r скорости движущихся объектов благодаря появлению доплеровского сдвига частоты отраженного сигнала F_v . Для этого в дальномере нужно иметь два частотомера для измерения разностных частот F_r и F_v , пропорциональных дальности и скорости. Временные диаграммы, поясняющие принцип совместного измерения дальности и скорости, для случая $F_{\text{мин}} \geq F_{\text{макс}}$ приведены на рис. 10.3.

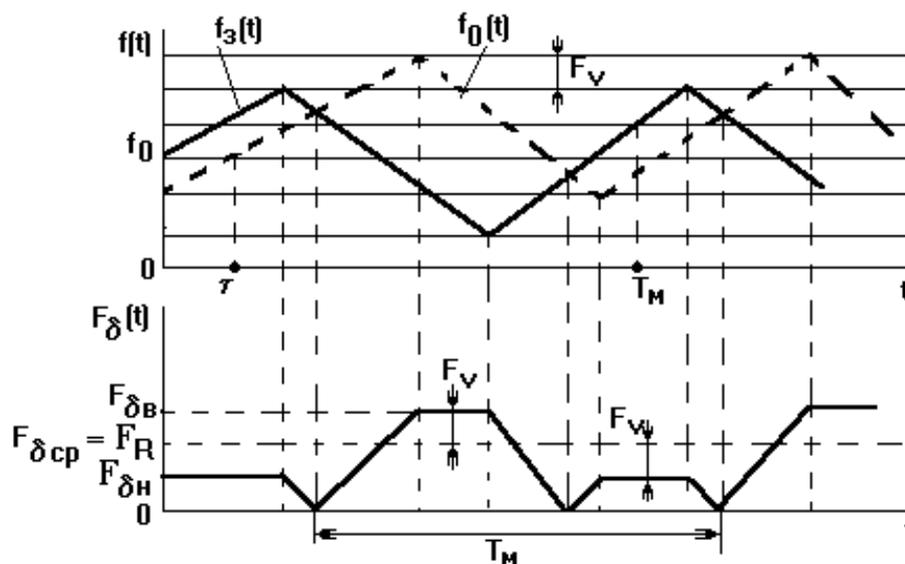


Рис. 10.3. Совместное измерение дальности и скорости

Как видно, в течение одного полупериода модуляции частота биений $F_{\text{бн}} = F_R - F_v$, в течение второго — $F_{\text{бв}} = F_R + F_v$ ($F_R = 8\Delta f_R / (cT_M)$, $F_v = 2f_v V_r / c \approx 2f_0 V_r / c$); соответственно, F_R и F_v определяются выражениями:

$$F_R = F_{\delta \text{ср}} = \frac{F_{\delta \text{н}}}{2}, \quad F_v = \frac{F_{\delta \text{в}} - F_{\delta \text{н}}}{2}, \quad (10.5)$$

в которых $F_{\text{бв}}$, $F_{\text{бн}}$ и $F_{\text{бср}}$ — высокая, низкая и средняя частоты биений (без учета провалов в $F_{\delta}(t)$) [1].

Таким образом, для измерения дальности и радиальной скорости объекта в схеме дальномера необходимо использовать два счетчика.

Однозначное измерение R и V_r обеспечивается при $F_{\text{мин.}} > F_{\text{макс.}}$, следовательно, параметры зондирующего сигнала должны удовлетворять неравенству

$$\frac{4\Delta f}{f_0 T_m} > \frac{V_{\text{макс.}}}{R_{\text{мин.}}} \quad (10.6)$$

Основными причинами погрешностей измерения R и V_r являются: паразитная амплитудная модуляция, искажающая огибающую сигнала биений, а также нестабильность девиации частоты Δf и частоты модуляции F_m . Линейность изменения частоты не имеет большого значения при измерении R и V_r одного объекта, однако при измерении R и V_r многих объектов она должна быть высокой [1].

§ 10.4. Методы и устройства радиопеленгования

Установление местоположения, направления и траектории движения объекта в РЛС, РНС и РТС радиоэлектронной разведки основывается на **радиопеленговании** – определении направления прихода радиосигналов, отраженных или излучаемых объектом с помощью устройств, называемых **радиопеленгаторами (РП)**. Радиопеленгом применительно к наземным объектам принято называть угол между северным направлением и линией пеленга РП – объект. Простейший РП содержит две антенны, разнесенные на некоторое расстояние d , называемое базой РП, а также приемник и индикатор угла прихода радиоволн. С помощью такого РП можно измерить, например, лишь азимут α наземного объекта; для установления местоположения объекта нужно использовать по крайней мере два удаленных друг от друга РП. Координаты объекта при этом определяются по карте как координаты точки пересечения линий пеленга ЛП 1, 2 (угломерный метод), см. рис. 10.4:

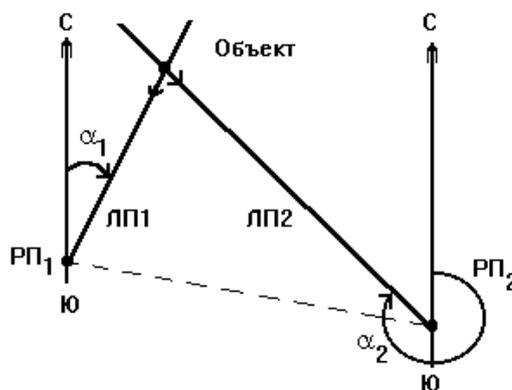


Рис. 10.4. Определение местоположения объекта

Существуют и более совершенные РП, позволяющие определять местоположение объекта из одной точки пространства (например, РП с широкобазисными антеннами, РП с двумя парами антенн с ортогональными базами, дальномерные РП) [1, 17].

Широкое распространение получили два метода радиопеленгования: фазовый и амплитудный, а также их разновидности.

Фазовый метод заключается в использовании зависимости разности фаз сигналов, принимаемых двумя антеннами, от угла прихода радиоволн (рис. 10.5).

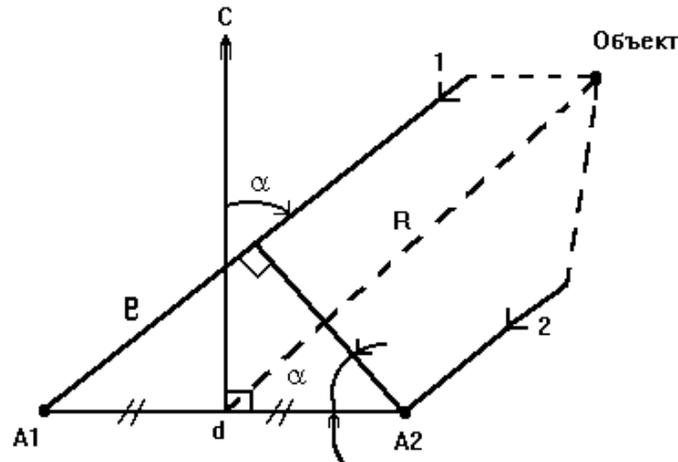


Рис. 10.5. Фазовый метод радиопеленгования

При расстоянии до объекта (например, радиостанции) $R \gg d, \lambda$ лучи 1 и 2 практически параллельны и разность фаз φ сигналов на выходах антенн A1 и A2 определяется их разностью хода l : $\varphi = 2\pi l / \lambda = (2\pi d / \lambda) \sin \alpha$. Следовательно, по измеренной величине φ можно рассчитать радиопеленг (азимут) α по формуле

$$\alpha = \arcsin \left[\frac{\varphi \lambda}{2\pi d} \right], \quad (10.7)$$

при этом надо учитывать, что однозначное измерение угла α возможно только при $d \leq \lambda$. Один из возможных вариантов построения структурной схемы фазового РП показан на рис. 10.6, на котором обозначены: СМ – смеситель, Г – гетеродин, А – антенны, УПЧ – усилитель промежуточной частоты, АО – амплитудный ограничитель, ФД – фазовый детектор.

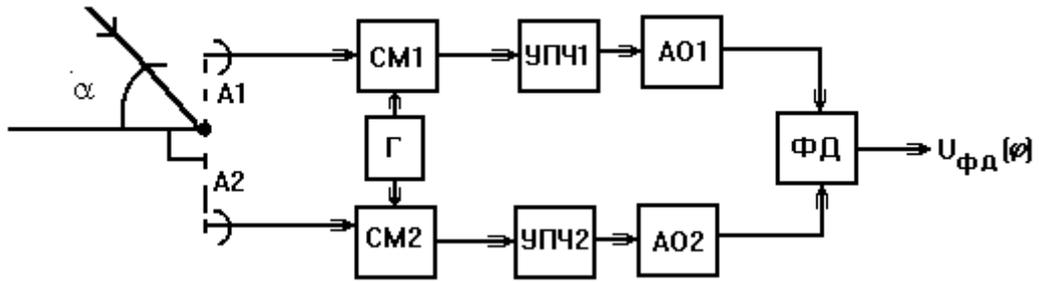


Рис. 10.6. Структурная схема фазового РП

Амплитудные ограничители введены в схему с целью устранения влияния флуктуаций амплитуд сигналов на выходах антенн (можно также использовать автоматическую регулировку усиления в УПЧ); в качестве ФД целесообразно использовать квадратурный балансный фазовый детектор, что позволит определять не только величину, но и знак угла α (при использовании простейшего ФД в один из каналов надо ввести фазовращатель на $\pi/2$ для придания нечетности зависимости $U_{\text{ФД}}$ от φ). В режиме квадратичного детектирования напряжение на выходе ФД можно представить [16] в виде $U_{\text{ФД}}(\varphi) = K_{\text{ФД}} U_0 \sin \varphi = K_{\text{ФД}} U_0 \sin(2\pi(d/\lambda)\sin \alpha) = U_{\text{ФД}}(\alpha)$, где $K_{\text{ФД}}$ – коэффициент передачи ФД; $U_0 = \text{const}$ – напряжение ограничения. При малых $\alpha < 10^\circ$

$$U_{\text{ФД}}(\alpha) \approx K_{\text{ФД}} U_0 \cdot \frac{2\pi d}{\lambda} \alpha. \quad (10.11)$$

Зависимость нормированного напряжения на выходе ФД $F(\alpha)$, называемая пеленгационной характеристикой (ПХ) РП, описывается выражением [1]:

$$F(\alpha) = \frac{U_{\text{ФД}}(\alpha)}{K_{\text{ФД}} U_0} = 2\pi \frac{d}{\lambda} \alpha = S_\alpha \cdot \alpha, \quad (10.12)$$

где параметр S_α – крутизна ПХ [1]:

$$S_\alpha = \left. \frac{dF(\alpha)}{d\alpha} \right|_{\alpha=0} = \frac{2\pi d}{\lambda}. \quad (10.13)$$

Очевидно, что при увеличении d/λ возрастает S_α и уменьшается угол нечувствительности $\alpha_{\text{мин}}$ РП (при котором отношение $(с/\ш)_{\text{ФД}} \approx 1$), однако возможность увеличения S_α за счет расширения базы РП d ограничена условием однозначности радиопеленгования $d/\lambda \leq 1$.

Простейшие фазовые РП с ненаправленными антеннами не обеспечивают угловое разрешение и имеют небольшую дальность

действия. Для достижения высокой РС $\delta\alpha$ и увеличения дальности действия необходимы остронаправленные антенны.

Фазовые РП можно применять для измерения и азимута α , и угла места β источника радиосигнала, если использовать две пары антенн с ортогональными базами, расположенными в горизонтальной плоскости.

Амплитудные методы основаны на зависимости амплитуды сигнала в виде суммы или разности выходных сигналов антенн от угла прихода радиоволн (обусловленной сдвигом фаз φ). Рассмотрим пять разновидностей этого метода.

Метод максимума – радиопеленгование по максимуму амплитуды суммарного сигнала. С учетом фазовых сдвигов $\pm\varphi/2$ относительно фазы $\varphi=0$ в центре базы РП представим напряжения на выходах антенн в виде $u_{1,2}(t)=U_{1,2}\sin(\omega t \pm \varphi/2)$; при расстоянии до источника сигнала $R \gg \lambda$ можно принять $U_1=U_2=U$ [1], тогда амплитуда U_c суммарного напряжения

$$u_c(t) = u_1(t) + u_2(t) = 2U \cos \varphi / 2 \sin \omega t = U_c(\varphi) \sin \omega t, \quad (10.14)$$

где $U_c(\varphi) = 2U \cos[\pi d / \lambda \sin \alpha] = U_c(\alpha)$, действительно зависит от α . Таким образом, вращая антенную систему РП в горизонтальной плоскости, можно измерить пеленг α , зафиксировав максимум $U_c(\alpha)$. Очевидные недостатки метода максимума: малая крутизна ПХ в области небольших углов α и, соответственно, большая погрешность измерения α ; невозможность определения знака α (ввиду четности функции \cos); зависимость $U_c(\alpha)$ в реальных РП от неизвестной амплитуды U .

Метод минимума – радиопеленгование по минимуму амплитуды U_p разностного сигнала

$$u_p(t) = u_1(t) - u_2(t) = 2U \sin \varphi / 2 \cos \omega t = U_p(\varphi) \cos \omega t, \quad (10.15)$$

где $U_p(\varphi) = 2U \sin[(\pi d / \lambda) \sin \alpha] = U_p(\alpha)$ – свободен от двух первых недостатков метода максимума, однако имеет свой специфический недостаток: при точной ориентации антенн РП на источник сигнала d в идеальном случае полностью идентичных каналов приема $U_p(\alpha=0)=0$, поэтому наблюдение сигнала на экране индикатора и тем более измерение дальности объекта практически невозможно.

Наиболее широко используется **суммарно-разностный** метод радиопеленгования, основанный на измерении отношения разностного (10.15) и суммарного (10.14) напряжений:

$$F(\alpha) = \frac{U_p(\alpha)}{U_c(\alpha)} = \operatorname{tg}[\sin(\pi d / \lambda) \sin \alpha]. \quad (10.16)$$

При этом исключается влияние изменяющихся амплитуд входных сигналов, определяется сторона уклонения от перпендикуляра к базе РП

(т.к. tg – нечетная функция), обеспечивается высокая точность пеленгования, поскольку в рабочей области ($\alpha \approx 0$) крутизна ПХ $S_\alpha = \pi d / \lambda$ может быть высокой. Наличие суммарного сигнала позволяет наблюдать отметку объекта (цели) на экране индикатора и измерять дальность объекта.

Метод анализа огибающей (разновидность метода максимума), часто применяемый в РЛС, основан на использовании остронаправленной антенны. В случае малоразмерных целей при сканировании ДН амплитуда сигнала на входе приемника РЛС изменяется пропорционально текущему значению ДН $F_A(\alpha)$ в направлении цели (рис. 10.7).

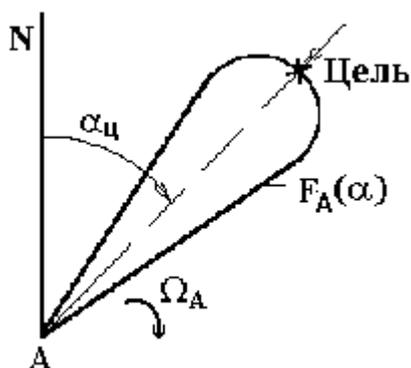


Рис. 10.7. Пеленгование методом анализа огибающей ДН

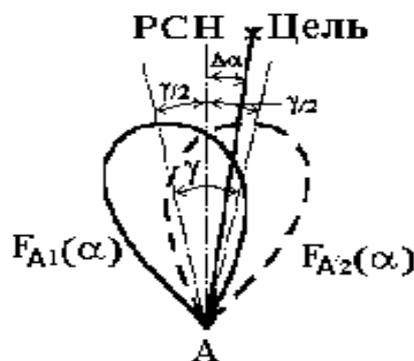


Рис. 10.8. Пеленгование равносигнальным методом

При измерении напряжения огибающей принимаемого сигнала фиксируют угловое положение максимума ДН, определяющее пеленг цели α_c .

Потенциальная точность измерения пеленга определяется дисперсией $\sigma_\alpha^2 = \left(\frac{q^2 \pi d_A}{\lambda} \right)^2$, разрешающая способность $\delta\alpha \approx 1,5 d_A / \lambda$, где q^2 – отношение сигнал/шум, d_A – размер раскрыва антенны в соответствующей плоскости [1].

Равносигнальный метод (разновидность метода минимума) основан на сравнении амплитуд принимаемых сигналов в двух положениях ДН (рис. 10.8).

Если направление на цель совпадает с линией, проходящей через точку пересечения ДН $F_{A1}(\alpha)$ и $F_{A2}(\alpha)$, называемой равносигнальным направлением (РСН), то амплитуды сигналов на выходе антенны, соответствующие первому и второму положению ДН, равны и разностный сигнал

$$U_p(\Delta\alpha) = U_p \left[F_A \left(\frac{\gamma}{2} + \Delta\alpha \right) - F_A \left(\frac{\gamma}{2} - \Delta\alpha \right) \right] = 0.$$

При рассогласовании $\Delta\alpha$ между направлением на цель и РСН $U_p(\Delta\alpha) \neq 0$, и по его величине и знаку определяют $\Delta\alpha$.

Угол смещения ДН относительно РСН γ выбирают так, чтобы они пересекались в области крутого спада $F_A(\alpha)$, при этом увеличивается точность пеленгования.

Сравнение сигналов может осуществляться последовательно в двух положениях одной и той же ДН с помощью одноканальной схемы или одновременно – при наличии двух пересекающихся ДН с помощью двухканальной схемы.

При последовательном сравнении в случае измерения двух угловых координат ДН вращают вокруг оси, совпадающей с РСН, ДН описывает в пространстве конус. Такой метод называется методом конического сканирования. Он широко применяется в **следающих измерителях** направления (СИН), в которых производится непрерывное совмещение опорного направления антенны измерителя $[\alpha_0(t), \beta_0(t)]$ с направлением прихода волны $[\alpha(t), \beta(t)]$, отраженной от цели. Ввиду простоты СИН с коническим сканированием используется в РЛС с автосопровождением целей [11, 13].

Более совершенны **моноимпульсные** измерители угловых координат, в которых сравнение амплитуд или фаз сигналов одновременно принятых по двум или нескольким каналам происходит в течение времени, равного длительности одного принятого импульса. Соответственно используемому методу радиопеленгации, моноимпульсные системы могут быть фазовыми, амплитудными и суммарно-разностными. Шире распространены суммарно-разностные системы, обеспечивающие высокую стабильность РСН и точность определения угловых координат; они могут применяться и для поиска целей и для одновременного измерения координат нескольких целей [1, 11, 13].

XI. Радиоэлектронная борьба и защита РТС от активных помех

§ 11.1. Радиоэлектронное подавление и защита от активных помех

Под **радиоэлектронной борьбой (РЭБ)** понимают комплекс взаимосвязанных по цели, месту, задачам и времени мероприятий и действий войск по выявлению систем управления войсками и оружием противника, их радиоэлектронному подавлению, защите своих систем управления войсками и оружием, а также противодействию радиоэлектронной разведке противника [17].

Радиоэлектронное подавление (РЭП) производится путем создания активных помех, снижающих эффективность работы РЭС противника. Технические комплексы РЭП называются постановщиками помех (ПП). Активные помехи, излучаемые ПП, по характеру воздействия подразделяются на **маскирующие** и **имитационные**. Маскирующие помехи, создавая мешающий фон, затрудняют обнаружение полезных сигналов, их различение и оценку параметров; они могут быть случайными (шумовыми) или детерминированными, непрерывными или импульсными [1, 9].

Непрерывная шумовая помеха универсальна, так как может использоваться для подавления РТС с любым типом сигнала. В зависимости от ширины спектра сигнала F_c и помехи $F_{ш}$ различают **прицельные** ($F_{ш} \approx F_c$) и **заградительные** ($F_{ш} \gg F_c$) помехи. При одинаковом мешающем действии прицельные помехи могут иметь меньшую мощность, чем заградительные, но для их постановки нужно иметь данные о спектре подавляемого сигнала.

Не менее действенными являются **имитационные помехи (ИП)**. ИП должны быть схожи с сигналом подавляемой РТС, отличаясь от него значениями информационных параметров. Типичный пример – ИП для радиодальномера, представляющие собой принятую и переизлученную ПП с некоторой задержкой последовательность зондирующих импульсов РЛС. Изменение задержки имитирует ложное движение цели.

К сказанному выше следует добавить, что мощные активные помехи вызывают перегрузку приемника РТС, вследствие чего прием полезных сигналов может оказаться невозможным.

Действие активных помех приводит к необходимости увеличения пороговой мощности полезного сигнала $P_{смин}$ на входе приемника РЛС и, соответственно, к снижению дальности действия РЛС. Рассмотрим общий случай, когда ПП создает шумовую заградительную помеху, причем цель и ПП не совмещены, т.е. разнесены в пространстве. Обозначим P_i и $P_{ип}$ – мощности передатчиков РЛС и ПП, G_i , G_p , $G_{ип}$ – коэффициенты усиления передающей и приемной антенн РЛС и антенны ПП, λ – длина волны излучения РЛС, σ_c – ЭПР цели, F_{Ai} , F_{Ap} – значения ДН передающей антенны РЛС в направлении на цель и приемной антенны в направлении на ПП (по напряженности поля). Примем, что ДН передающей антенны ПП ориентирована на РЛС. Пусть P_p и $P_{смин}^p = P_{смин} + m_{pp} P_p$ – мощность помехи и пороговая мощность сигнала на входе приемника РЛС в условиях РЭП; $m_{pp} = q_{мин}^2 (F_c / F_{ш})$ – коэффициент различимости сигнала на фоне помехи [1]. Используя формулы § 8.1, можно получить следующее выражение для расчета дальности действия подавляемой РЛС:

$$R_{\max} = \sqrt[4]{\frac{P_i G_i G_n F_{Ai}^2 F_{An}^2 \lambda^2 \sigma_c}{(4\pi)^3 \left[P_{смин} + m_{pp} \frac{P_{ип} G_{ип} G_n F_{An}^2 \lambda^2}{(4\pi)^2 R_n^2} \right]}} \quad (11.1)$$

В случае совмещенной помехи (ПП находится на цели) в выражении (11.1) P_p следует заменить на R_{\max} и из полученного уравнения выразить R_{\max} .

Формула (11.1) может быть использована также для оценки дальности действия радиоволновых охранных извещателей в условиях непреднамеренных или организованных помех.

Подробный анализ эффективности РЭП применительно к РЛС приведен в работах [1, 9].

Большинство методов защиты от активных помех основано на различных способах их селекции: пространственной, поляризационной, амплитудной, временной и частотной [1, 9].

Пространственная селекция обеспечивается применением остронаправленных антенн с низким уровнем боковых лепестков ДН и адаптивных антенн.

Поляризационная селекция основана на различии поляризации излучения источников полезного сигнала и помех. Осуществляется путем использования антенн с низким уровнем кросс-поляризации и поляризационных фильтров.

Амплитудная селекция производится путем применения различных систем АРУ, амплитудных ограничителей, а также усилителей с расширенным динамическим диапазоном.

Временная селекция осуществляется путем запираания входа приемника РТС на время отсутствия полезного сигнала и использования систем временных АРУ.

При частотной селекции используется различие в структуре спектров сигнала и помехи и их несущих частот. Для подавления помех в приемниках РТС применяются различные аналоговые и цифровые фильтры, в том числе адаптивные.

Первостепенной мерой защиты от РЭП является обеспечение скрытности работы РТС, достигаемое сокращением времени работы, применением остронаправленных антенн, использованием специфических особенностей атмосферного распространения радиоволн, а также применением ШПС. Одним из радикальных способов повышения помехозащищенности РТС является комплексирование систем, работающих на основе различных физических принципов в разных диапазонах электромагнитных волн.

§ 11.2. Защита РЛС от пассивных помех

Пассивными помехами называются сигналы, отраженные мешающими объектами при их облучении зондирующими сигналами РЛС. Пассивные помехи могут подавлять и маскировать сигналы, отраженные от наблюдаемой цели. В качестве примера создания преднамеренных пассивных помех можно привести противорадиолокационную маскировку летательных аппаратов путем разбрасывания большого количества полосок из фольги или металлизированной полимерной пленки с длиной, близкой к половине длины волны зондирующего сигнала РЛС.

Методы защиты РЛС от пассивных помех основаны на различии параметров сигналов, отраженных целями и мешающими объектами, обусловленными их размерами, пространственным положением, различной способностью рассеивать электромагнитные волны. На практике часто применяют поляризационную селекцию радиолокационных сигналов, а также селекцию движущихся целей (СДЦ) [1, 9].

СДЦ на основе эффекта Доплера. Принцип действия доплеровской СДЦ основан на использовании различия радиальных скоростей движения цели и источников пассивных помех, и, соответственно, доплеровских смещений частот полезного и мешающего сигналов.

Наиболее просто производится сравнение частот непрерывных зондирующего и радиолокационного сигналов. Но, как показано выше, в РЛС чаще применяются периодические импульсные зондирующие сигналы, позволяющие обеспечить высокую разрешающую способность и точность измерения дальности цели. Однако в этом случае СДЦ сопряжена с появлением "слепых" скоростей, т. е. таких, при которых подавляется сигнал, отраженный от цели. Устранить это нежелательное явление можно, используя изменение частоты повторения зондирующих импульсов или излучение на двух и более несущих частотах [1].

Существуют системы СДЦ с внутренней и внешней когерентностью [1]. В системах с внутренней когерентностью опорные колебания (когерентные с излучаемыми), необходимые для выделения доплеровского приращения частоты отраженных от цели импульсных сигналов, формируются в самой системе СДЦ. В системах с внешней когерентностью в качестве опорных используются сигналы, отраженные от неподвижных объектов, находящихся в мгновенном угловом поле зрения антенны РЛС в пределах объема разрешения, в котором находится и движущаяся цель.

Ввиду большей эффективности подавления помех чаще применяются устройства СДЦ с внутренней когерентностью [1]. В современных РЛС сигналы, отраженные от неподвижных объектов, подавляются с помощью специальных компенсаторов, устанавливаемых на выходе приемника [1]. Рассмотрим простейший пример СДЦ с помощью череспериодного компенсатора, в котором из сигналов, принимаемых в текущий момент времени, вычитаются сигналы, задержанные на время, равное периоду T_n повторения зондирующих импульсов РЛС. Структурная схема компенсатора показана на рис. 11.1.

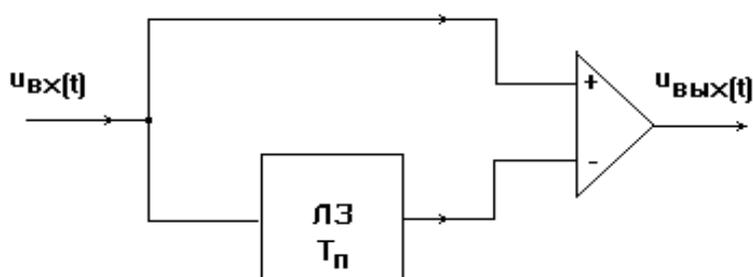


Рис. 11.1. Структурная схема компенсатора

В данном устройстве сигналы, отраженные от неподвижных объектов, компенсируются, так как в течение интервала времени T_n их амплитуды не изменяются; сигналы, отраженные от движущихся объектов, имеют амплитуду, изменяющуюся с доплеровской частотой F_D и разностное напряжение $u_{\text{вых}}(t)$, размах которого определяется набегом

фазы за T_n $\Delta\varphi = 2\pi F_n T_n$ [1]. Таким образом, череспериодный компенсатор является гребенчатым фильтром с АЧХ $|K(f)| = 2|\sin \pi f T_n|$, показанной на рис. 11.2.

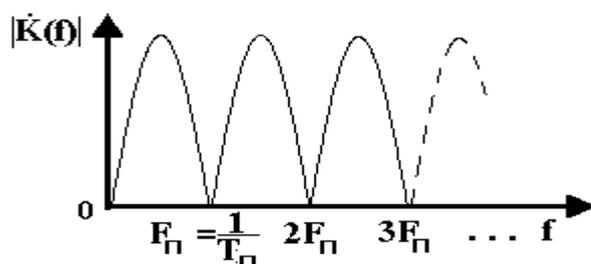


Рис. 11.2. АЧХ компенсатора

Фильтр обеспечивает полное подавление сигналов неподвижных объектов, так как их спектральные компоненты имеют частоты $f_n = nF_n$. Однако, если цель имеет радиальную скорость $V_{rn} = \frac{\lambda n F_n}{2}$, где λ – длина волны зондирующего излучения, $n=1; 2; 3; \dots$, то фильтр подавит и полезный сигнал, т. е. соответствующие скорости V_{rn} – ”слепые” – такие, при которых РЛС не обнаруживает цель.

В реальных РЛС вследствие конечности времени облучения цели принимаемый сигнал оказывается не бесконечной периодической последовательностью, а пакетом импульсов. Поэтому из-за ”расплывания” спектральных линий череспериодный компенсатор не обеспечивает полное подавление помех от неподвижных объектов. Для повышения эффективности СДЦ часто используют двукратное вычитание мешающих сигналов путем последовательного соединения гребенчатых фильтров. Поскольку аналоговые фильтры СДЦ сложны в реализации и дороги, в настоящее время используются цифровые режекторные фильтры подавления помех, а также фильтры, в которых в качестве линий задержки применяют приборы с зарядовой связью [1].

§ 11.3. Системы радиоэлектронной разведки

Радиоэлектронной разведкой (РЭР) принято называть получение информации о противнике с помощью специальных радиоэлектронных средств [17]. Средства РЭР могут быть как пассивными (неизлучающими), так и активными. В первом случае информация добывается в результате обзора пространства, обнаружения и приема сигналов источников электромагнитного излучения, во втором случае – путем зондирования пространства и приема сигналов, отраженных объектами разведки. РЭР классифицируется по следующим признакам: целевое назначение; место базирования средств разведки; характер добываемой информации; используемые принципы и технические средства.

По последнему признаку выделяют следующие виды РЭР: радиоразведка, радиотехническая, радиолокационная, лазерная, инфракрасная, телевизионная и фоторазведка [17]. Техническую основу РЭР составляют: пассивные устройства поиска, перехвата и анализа сигналов и пеленгования источников электромагнитных излучений; активные радиолокационные и лазерные устройства наблюдения, обнаружения и распознавания объектов; оптико-электронные приборы; средства регистрации (индикации) и передачи добываемой информации; аппаратура обработки поступающей разведывательной информации, в том числе компьютеры.

Рассмотрим подробнее первый вид РЭР – **радиоразведку (РР)**. Назначение РР – обнаружение, перехват радиопередач и пеленгование источников радиосигналов, обработка и анализ получаемой информации с целью вскрытия ее содержания и определения местонахождения радиостанций.

Объектами РР являются работающие средства радиосвязи. В состав комплекса технических средств РР, называемого **системой РР**, входят: радиоприемники и антенны, радиопеленгаторы, панорамные приемники и анализаторы спектра частот, компьютеры, выходные устройства для выделения и перехвата сигналов многоканальных систем радиосвязи, оконечные устройства слухового контроля; аппаратура рассекречивания зашифрованных и кодированных радиопередач, средства управления, связи и передачи добываемой информации. Обработка и анализ разведывательной информации – первостепенная задача, возникающая при организации РЭБ. Принципы построения современных радиоприемных устройств, используемых в системах РР (со сжатием сигналов, с акустическими преобразователями и т.д.) подробно изложены, например, в работах [1, 16, 17].

ХII. Основы построения радиотеплолокационных и оптических локационных систем

§ 12.1. Физические основы радиотеплолокации

При радиотеплолокации для получения информации об объектах (целях) используется их естественное электромагнитное излучение теплового происхождения. Технические средства радиотеплолокации называют радиотеплолокаторами (РТЛ) или радиометрами. Тепловое излучение нагретых тел может проявляться в очень широком интервале длин электромагнитных волн – от десятых долей микрометра до единиц сантиметров; РТЛ работают в диапазонах сантиметровых, миллиметровых и субмиллиметровых волн [9].

Тепловое излучение различных объектов характеризуется следующими энергетическими параметрами:

- 1) спектральной яркостью V_f (Вт/(ср. м² Гц));
- 2) спектральной плотностью потока излучения Π_f (Вт/(м² Гц));
- 3) эффективной яркостной температурой $T_{яэ}$ (К).

Идеальным тепловым источником является так называемое абсолютно черное тело (АЧТ), полностью поглощающее падающее на него электромагнитное излучение на всех частотах. Для излучения сантиметрового и миллиметрового диапазонов близкими по свойствам к АЧТ являются, например, лесные массивы и специальные противорадиолокационные покрытия [15].

Спектральная плотность потока теплового излучения АЧТ в диапазонах частот и температур, характерных для РТЛ, определяется законом Рэлея-Джинса [1]: $\Pi_f^{AT} = 2\pi f^2 kT / c^2 = 2\pi kT / \lambda^2$, где $k=1,38 \cdot 10^{-23}$ Дж/К – постоянная Больцмана; T – абсолютная температура АЧТ.

Это выражение можно использовать и для определения Π_f реальных объектов, не полностью поглощающих падающее на них электромагнитное излучение, если вместо T подставить: $T_{яэ} = \alpha T + \gamma T_{ср}$, где α – коэффициент поглощения, γ – коэффициент отражения внешнего излучения, причем по закону сохранения энергии $\alpha + \gamma = 1$ (для непрозрачных тел); T – температура объекта; $T_{ср}$ – эквивалентная температура окружающей среды, обусловленная внешним излучением. Форма ДН теплового излучения в первую очередь зависит от шероховатости поверхности объекта. Эффективность различения двух объектов (например, человека на фоне стены здания, самолета на фоне неба и т.д.) по их тепловому излучению характеризует температурный контраст $\Delta T = T_{яэ1} - T_{яэ2}$.

В радиотеплолокации для оценки мощности радиотеплового сигнала обычно используют температуру антенны T_A [1]. При наблюдении протяженной цели, площадь $S_{ц}$ поперечного сечения которой превышает площадь сечения главного лепестка ДН антенны T_A принимают приблизительно равной $T_{яэ}$ цели; в случае малоразмерной цели $T_A \approx T_{яэ} \Omega_{ц} / \Omega_A$, где $\Omega_{ц} \approx S_{ц} / R^2$ – угловой размер цели, R – дальность; $\Omega_A = 4\pi / D_A$ – телесный угол, охватывающий главный лепесток ДН антенны с коэффициентом направленного действия D_A [1].

§ 12.2. Обнаружение радиотепловых сигналов. Принципы построения РТЛ

Задача обнаружения радиотеплового сигнала может быть сведена к задаче обнаружения случайного радиосигнала при использовании модели белого шума, так как Π_f теплового излучения в полосе пропускания Δf приемников РТЛ оказывается практически равномерной [1, 16]. Таким образом, смесь $x(t)$ на входе приемника на интервале наблюдения $t \in [0; T]$, как и ранее, представляется в виде $x(t) = au(t) + n(t)$, $a = \{1; 0\}$; где $u(t)$ – белый шум, соответствующий радиотепловому сигналу, $n(t)$ – собственный шум приемника.

Мощности сигнала и шума определяются их дисперсиями: $P_c = \sigma_c^2 = kT_A \Delta f$, $P_{ш} = \sigma_{ш}^2 = kT_{ш} \Delta f$, где $T_{ш}$ – шумовая температура приемника [15].

Оптимальный обнаружитель может быть синтезирован на основе правила отношения правдоподобия (2.7). В конечном итоге (2.7) приводится к следующему виду [1, 16]:

$$E = \int_0^T x^2(t) dt \underset{H_0}{\overset{H_1}{>}} Z_0, \quad (12.1)$$

где E – энергия смеси $x(t)$, накопленная за время наблюдения T , $Z_0 = f(\Lambda_0, \sigma_c^2, \sigma^2)$ – пороговое напряжение обнаружения. Отношение сигнал/шум на выходе обнаружителя при слабом сигнале [15]:

$$q^2 \approx \frac{(m_{сш} - m)^2}{\sigma^2},$$

где $m_{сш}$ и m – математические ожидания выходного напряжения (тока) при наличии и отсутствии сигнала [15].

Структурная схема обнаружителя радиотеплового сигнала, реализующая правило (12.1), приведена на рис. 12.1. Такой обнаружитель называется энергетическим.

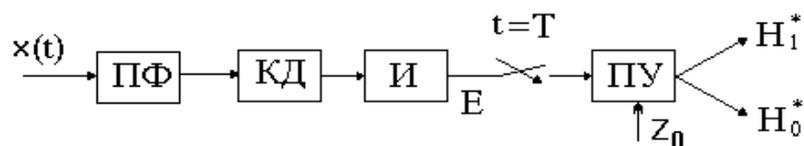


Рис. 12.1. Структурная схема энергетического обнаружителя

Входная смесь поступает на полосовой фильтр ПФ, полоса пропускания которого Δf приближенно равна ширине спектра F сигнала $u(t)$; после квадратичного детектора КД сигнал подается на вход интегратора И, в качестве которого можно использовать ФНЧ с частотой среза АЧХ $F_B = 1/T$, где T – эффективная длительность радиотеплового сигнала, зависящая от скорости сканирования ДН антенны РТЛ, углового размера цели Ω_c и ширины ДН Ω_A [1].

На самом деле P_f объекта и мешающих объектов, создающих тепловой фон, неизвестны, поэтому в реальных РТЛ часто используется контрастный метод обнаружения, состоящий в сравнении разности напряжений интегратора, пропорциональной разности $\Delta E = E_1 - E_2$ энергий сигналов при переключении ДН в соседнее положение (1→2), с порогом Z_0 , зависящим от ожидаемого температурного контраста ΔT . Пороговый контраст, обнаруживаемый с заданными вероятностями P_0 и P_n , определяется выражением [1]: $\Delta T_n = T_n [\Phi^{-1}(1 - P_n) + \Phi^{-1}(P_0)] / (0,7 \sqrt{F / F_B})$, где $T_n = T_A + T_{ш}$ – полная температура источника теплового излучения (цели). При этом правило обнаружения записывается в виде $\Delta T \geq \Delta T_n$ (условие яркостного контраста). Интересно, что это условие применительно к протяженным объектам не связано с их дальностью R [1]. Для

малоразмерных объектов $\Delta T \geq \Delta T_{\text{п}} \frac{\Omega_A}{\Omega_{\text{ц}}}$ и максимальная дальность обнаружения – дальность действия РТЛ – определяется выражением [1]:

$$R_{\text{макс}} = \sqrt{\frac{\Delta T}{\Delta T_n} \frac{S_{\text{ц}} D_A}{4\pi}}. \quad (12.2)$$

В РТЛ, как видно из сказанного выше, на результат обнаружения сильно влияют шумы приемника. С целью их подавления в реальных РТЛ используются методы компенсации, корреляционной обработки и модуляции принимаемых сигналов. Кратко рассмотрим структуру, принципы действия, достоинства и недостатки соответствующих РТЛ. Компенсационный РТЛ в одном из вариантов строится по схеме, показанной на рис. 12.2 [15].

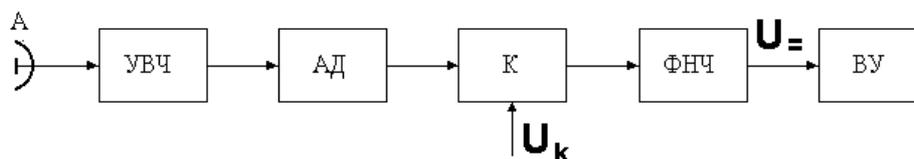
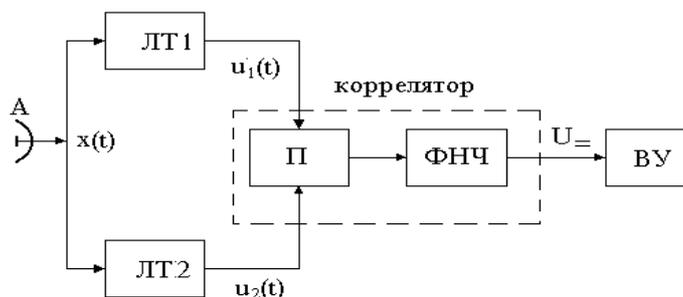


Рис. 12.2. Структурная схема компенсационного РТЛ

РТЛ (рис. 12.2) содержит приемную антенну А, усилитель высокой частоты УВЧ, амплитудный детектор АД, компенсатор К, ФНЧ и выходное (индикаторное) устройство ВУ. В компенсаторе, выполняющем функции порогового устройства, постоянная составляющая напряжения шумов приемника U_k вычитается из выходного напряжения АД. Поэтому в отсутствие радиотеплового сигнала напряжение $U_{\text{н}}$ на выходе ФНЧ (пропорциональное $\Delta m = m_{\text{сш}} - m$) близко к нулю, а при наличии сигнала ($\Delta m > 0$) $U_{\text{н}}$ не равно нулю, что и указывает на его обнаружение соответственно правилу оптимального обнаружения (12.1). Достоинства компенсационного РТЛ: высокая чувствительность и конструктивная простота; главный недостаток – неустойчивость компенсации собственных шумов.

Корреляционный РТЛ строится по схеме, показанной на рис. 12.3; он содержит антенну А, два линейных тракта ЛТ1, 2 (содержащих УВЧ, преобразователи частоты и усилители промежуточной частоты), перемножитель П, ФНЧ и ВУ [1]:



$$u_1(t) = x_1(t) + u_{ш1}(t); \quad u_2(t) = x_2(t) + u_{ш2}(t)$$

Рис. 12.3. Структурная схема корреляционного РТЛ

В таком РТЛ $U_{\Sigma} = K \int_0^{\tau} u_1(t) u_2(t) dt$, где K – коэффициент

пропорциональности, τ – постоянная времени ФНЧ; $u_{ш1, 2}(t)$ – напряжения шумов ЛТ1, 2. Вследствие статистической независимости $u_{ш1, 2}(t)$ и принимаемого радиотеплового сигнала уровень собственных шумов на выходе интегратора (ФНЧ) минимален; в схеме реализуется оптимальное правило обнаружения (12.1). Недостаток корреляционного РТЛ – зависимость U_{Σ} от изменения фазовых сдвигов сигналов в ЛТ1, 2.

Наиболее распространенными являются модуляционные РТЛ. Простейший модуляционный РТЛ содержит антенну А, модулятор М, генератор сигнала модуляции ГСМ, УВЧ, АД, узкополосный усилитель УУ, синхронный детектор СД, ФНЧ и ВУ, рис. 12.4 [15].

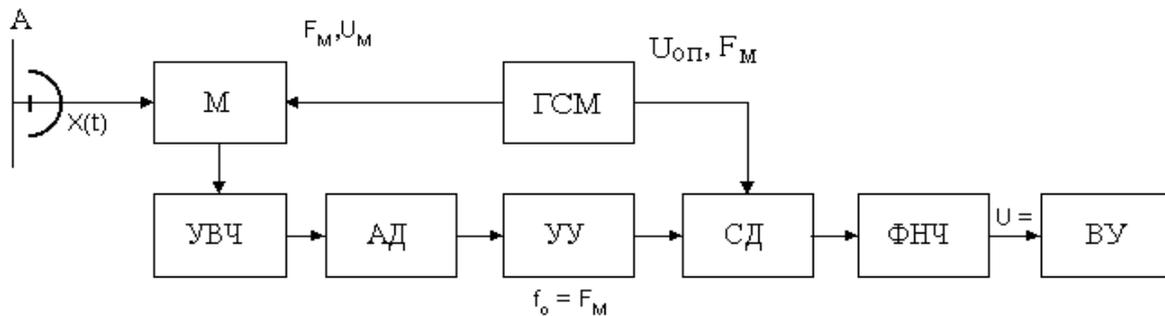


Рис. 12.4. Структурная схема модуляционного РТЛ

В данном РТЛ входной радиотепловой сигнал подвергается амплитудной модуляции; за счет последующего синхронного детектирования удается резко ослабить влияние собственных шумов РТЛ. Достоинство описанного модуляционного РТЛ – большая, чем у компенсационного, временная стабильность параметров; недостаток – меньшая чувствительность, которую можно повысить, используя супергетеродинный прием.

Очевидно, что РТЛ имеют ряд преимуществ по сравнению с РЛС: скрытность работы, обусловленная отсутствием собственного зондирующего излучения, компактность, низкое энергопотребление. Основной недостаток РТЛ – значительное время накопления энергии для обнаружения объекта из-за низкого уровня мощности теплового излучения, а также невозможность определения дальности объекта одним РТЛ из-за случайности параметров радиотепловых сигналов. Для измерения дальности нужны как минимум два РТЛ: используется

триангуляционный метод – дальности до объекта R_1 и R_2 определяются по известной базе РТЛ и измеренным пеленгам α_1 и α_2 , (см. рис. 12.4).

§ 12.3. Физические основы оптической локации

Оптической локацией называется обнаружение и определение местоположения объекта с помощью оптического излучения (в диапазоне длин электромагнитных волн от десятков до десятых долей микрометра). Так как в качестве источников излучения обычно используются лазеры, под оптической ниже будем подразумевать лазерную локацию (ЛЛ). Технические средства ЛЛ, аналогично радиолокационным, называются лазерными локационными системами или станциями (ЛЛС), либо просто лазерными локаторами. ЛЛ и радиолокация имеют много общего: объекты наблюдения, физические основы, способы определения координат и параметров движения объектов. Важнейшее отличие оптических ЛЛ – использование в качестве источников зондирующего излучения лазеров, – генераторов электромагнитных колебаний с чрезвычайно высокой пространственно-временной когерентностью. Благодаря этому свойству излучение лазерных передатчиков характеризуется высокой стабильностью частоты ($\Delta f/f$ достигает $10^{-10} \dots 10^{-11}$), “чистотой” спектра (монохроматичностью), спектральной плотностью мощности и направленностью (угловая расходимость излучения – аналог ширины ДН – у газовых лазеров составляет единицы – десятые доли миллирадиан и легко может быть снижена на порядок) [18]. Перечисленные особенности лазерного излучения (с учетом диапазона абсолютных значений частот колебаний – приблизительно от $3 \cdot 10^{13}$ до $5 \cdot 10^{14}$ Гц) позволяют реализовать ЛЛС с точностью измерения координат и параметров движения объектов, недостижимой в РЛС. Кроме того, за счет возможности формирования сверхкоротких импульсов излучения (длительностью порядка единиц – десятков наносекунд) при использовании твердотельных лазеров в режиме модуляции добротности разрешающая способность ЛЛС по дальности достигает десятых долей метра. Еще одним важным качеством ЛЛС является возможность получения изображений объектов, удаленных на сотни – тысячи метров.

С учетом сказанного выше можно выделить следующие специфические особенности ЛЛ [18 – 20]:

1. Пространственная распределенность большинства объектов (целей) для ЛЛС вследствие малой длительности зондирующих импульсов и узости ДН передающих и приемных антенн.
2. Зависимость параметров ЛЛС от условий распространения излучения в атмосфере и от наличия в зоне действия ЛЛС осадков и других рассеивателей – дымов, пыли, различных аэрозолей и т.п.
3. Большое время обзора рабочей зоны ЛЛС и сложность наведения оптических антенн на объекты наблюдения.

4. Ограниченная эффективность обнаружения локационных сигналов – как из-за внешних (фоновых) и внутренних шумов, так и из-за квантовой природы самих сигналов.
5. Использование дискретного способа регистрации оптических сигналов малой мощности путем счета отдельных фотоэлектронов (фотонов).

Таким образом, ЛЛС обладают следующими достоинствами: позволяют с большей, чем РЛС, эффективностью решать задачи локации; характеризуются большей помехозащищенностью и повышенной электромагнитной совместимостью. Наиболее существенными недостатками ЛЛС являются: сильная зависимость дальности действия от состава и состояния атмосферы, наличия и интенсивности осадков; сложность селекции полезного сигнала на фоне помехи обратного рассеяния; сложность наведения на цель и большее, нежели в РЛС, время обзора.

Лазерная локация подразделяется на **активную** и **полуактивную**. **Активная ЛЛ** производится путем облучения объекта (цели), приема и обработки отраженного от него оптического излучения. В состав активной ЛЛС (рис. 12.5) входят: лазерный передатчик ЛП, содержащий лазер 1 и устройство модуляции излучения, а также передающую оптическую систему 2 с устройством сканирования излучения; оптический приемник ОП, содержащий приемную оптическую систему 3 (которая может быть совмещена с передающей) и фотоприемник 4; выходное (индикаторное) устройство 5.

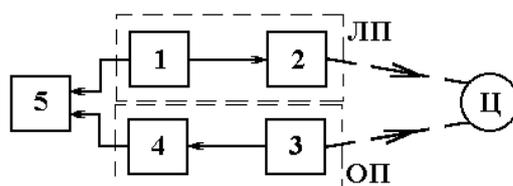


Рис. 12.5. Упрощенная структурная схема ЛЛС

При **полуактивной ЛЛ** лазерный передатчик и оптический приемник размещаются в различных пунктах.

Методы измерения дальности, скорости и угловых координат объектов аналогичны используемым в РЛС.

§ 12.4. Дальность действия ЛЛС

В зависимости от соотношения размеров поперечного сечения зондирующего лазерного луча и цели диапазон дальностей разбивается на два поддиапазона:

- 1) $R_1 < R_r$ (R_r – расстояние, на котором площадь поперечного сечения луча приблизительно равна площади поперечного сечения цели) – локация протяженных целей;
- 2) $R_2 > R_r$ – локация малоразмерных (точечных) целей.

На интервале $R_1 < R_r$ мощность излучения, достигающего цели, практически постоянна; на интервале $R_2 > R_r$ она зависит от расстояния. При ЛЛ различных объектов может иметь место как зеркальное отражение, так и диффузное рассеяние зондирующего излучения. Рассмотрим худший случай, когда на поверхности цели нет "блестящих" точек и отраженный от нее сигнал образуется за счет диффузного рассеяния. Дальность действия импульсных ЛЛС в свободном пространстве при этом определяется следующими выражениями [18]:

$$R_{\text{макс1}} = \sqrt{\frac{P_{\text{и}} G k_0 S_A \rho_D(\alpha)}{\pi P_{\text{смин}}}}, \text{ при } R_1 < R_r ; \quad (12.3)$$

где $P_{\text{и}}$ – импульсная мощность излучения лазера;

G – коэффициент усиления передающей оптической системы (антенны);

$k_0 = k_1 k_2$ – общий коэффициент пропускания излучения передающей (k_1) и приемной (k_2) оптических систем;

S_A – площадь объектива приемной оптической системы;

$\rho_D(\alpha)$ – коэффициент диффузного отражения цели в угловом направлении α по отношению к нормали к ее поверхности;

$P_{\text{смин}}$ – пороговая мощность сигнала на входе фотоприемника.

$$R_{\text{макс2}} = \sqrt[4]{\frac{P_{\text{и}} G k_0 S_A \sigma_{\text{ц}}}{\pi^2 \Theta_{\text{л}}^2 P_{\text{смин}}}}, \text{ при } R_2 > R_r; \quad (12.4)$$

где $\sigma_{\text{ц}}$ – ЭПР малоразмерной цели;

$\Theta_{\text{л}}$ – угловая расходимость излучения лазера (рад.).

При работе ЛЛС на **атмосферных** трассах необходимо учитывать ослабление излучения при его распространении до цели и обратно. В этом случае [14]:

$$1) \text{ при } R_1 < R_r: R_{\text{макс1}}^{\text{атм}} = R_{\text{макс1}} \cdot \tau_{\text{ал}};$$

$$2) \text{ при } R_2 > R_r: R_{\text{макс2}}^{\text{атм}} = R_{\text{макс2}} \cdot \sqrt{\tau_{\text{ал}}};$$

где $\tau_{\text{ал}}$ – спектральный коэффициент одностороннего пропускания атмосферы на длине волны λ зондирующего излучения; его значение может быть рассчитано или взято из справочника [18, 20].

При ЛЛ целей, поверхности которых обладают свойством зеркального отражения или имеют "блестящие" точки, например целей, снабженных ЗУО или решетками ЗУО (катафотами), дальность обнаружения (при равных прочих условиях) существенно превышает дальность обнаружения диффузно-рассеивающих целей [18].

В ЛЛ ЛЛС чаще всего используются твердотельные лазеры – на рубине ($\lambda \approx 0,69$ мкм), стекле, легированном неодимом (1,06 мкм), иттрий-алюминиевом гранате (1,06 мкм); газовые лазеры – на углекислом газе (10,6 мкм) и полупроводниковые лазеры (≈ 1 мкм) [18 – 20].

В качестве фотодетекторов в ОП ЛЛС используются фотоэлектронные умножители, р-і-п и лавинные фотодиоды, фоторезисторы. ОП могут быть как прямого детектирования (энергетические), в том числе на поднесущей частоте, так и гетеродинного или гомодинного типов. Оптимальные обнаружители оптических сигналов строятся, как правило, на основе когерентных оптических корреляторов или согласованных фильтров [1, 18].

В настоящее время ЛЛС применяются в космической и военной технике, в геодезии, метеорологии и других областях науки и техники [1, 18 – 20]. Перспективы применения ЛЛС в системах охранной сигнализации обусловлены возможностью организации охраны протяженных (до нескольких километров) периметров, определения размеров, скорости и направления движения объекта-нарушителя.

ХІІІ. АДАПТИВНЫЕ СИСТЕМЫ

13.1. Определения и свойства

В последние годы в результате исследований по «адаптивным системам» [21] появились различные адаптивные автоматы, свойства которых в некотором смысле напоминают определенные свойства живых систем и биологических адаптивных процессов. Ниже приводятся некоторые значения слова «адаптация»:

1. Действие, процесс приспособления.
2. Состояние приспособляемого; приспособление.
3. *Биологическое:*

а) любое изменение в структуре или функции организма или любой из его частей в результате естественного отбора, с помощью которого организм становится более приспособленным для выживания и размножения в окружающей его среде;

б) видоизменение формы или структуры в соответствии с изменением окружающей среды.

4. *Физиологическое.* Ослабление отклика органов сенсорных рецепторов» таких, как: зрение, осязание, температура, обоняние, слух и боль - на изменяющиеся, постоянно воздействующие условия окружающей среды.

5. *Офтальмологическое.* Регулирование зрачком количества света, поступающего в глаз.

6. *Социальное.* Медленное, обычно неосознанное, изменение индивидуальной и социальной деятельности в процессе приспособления к культурной среде.

Можно заметить, что приведенные определения даны главным образом в терминах биологической адаптации к окружающей среде. Такие же определения в некоторой степени подходят и для «искусственных», или созданных человеком, адаптивных систем, которые в основном

рассматриваются в этой главе. Адаптивный автомат представляет собой систему, структура которой изменяется или приспособливается таким образом, чтобы его поведение или функционирование улучшалось (в соответствии с некоторым подходящим критерием) в результате взаимодействия с окружающей его средой. Простым примером автомата для автоматической адаптивной системы является автоматическая регулировка усиления (АРУ), применяемая в радио- и телевизионных приемниках. Функция этой системы — уменьшение чувствительности приемника при увеличении среднего уровня входного сигнала. Таким образом, приемник может адаптироваться к широкому диапазону уровней входных сигналов и формировать значительно более узкий диапазон уровней выходных сигналов.

Цель настоящей главы — описать некоторые основные принципы адаптации; изложить методы разработки, функциональные свойства, области применения адаптивных систем простейших видов и способы их физической реализации. Рассматриваемые системы включают в себя системы, созданные прежде всего для адаптивного управления и адаптивной обработки сигналов. Как правило, такие системы обладают некоторыми или всеми перечисленными ниже свойствами:

1. Могут адаптироваться (самооптимизироваться) при изменении (нестационарном) условий окружающей среды и требований к системе.

2. Они могут обучаться для осуществления заданного вида фильтрации и выполнения задачи принятия решения. Системы с такими свойствами можно автоматически синтезировать через обучение. Адаптивные системы можно в некотором смысле «запрограммировать» процессом обучения.

3. Они не требуют тщательно разработанных методов синтеза, обычно необходимых для неадаптивных систем. Наоборот, их можно считать «самоорганизующимися».

4. Они могут экстраполировать модель поведения для функционирования в новых условиях после обучения на конечном и часто небольшом числе обучающих сигналов или ситуации.

5. Они могут в некоторой степени восстанавливаться, т. е. адаптироваться к определяемым внутренним дефектам.

6. Их можно рассматривать как нелинейные системы с изменяющимися во времени параметрами.

7. Их сложнее анализировать, чем неадаптивные системы, но они позволяют значительно увеличить область функционирования системы, когда параметры входного сигнала не известны или изменяются во времени.

13.2. Области применения адаптивных устройств и систем

Достигнутый за последнее время прогресс в разработке и производстве микросхем привел к созданию очень компактных, экономичных и надежных устройств обработки сигналов, конкурирующих

с биологическими нейронными системами по размерам и, очевидно, превосходящих биологические системы по быстродействию. В результате этого значительно расширилась область их применения во всех видах цифровой обработки сигналов, в том числе адаптивной обработки. В настоящее время адаптивные системы применяются в таких областях, как связь, радиолокация, гидролокация, сейсмология, проектирование механических систем, навигация и биомедицинская электроника.

13.3. Общие свойства адаптивных устройств и систем

Основным свойством адаптивной системы является изменяющееся во времени функционирование с саморегулированием. Необходимость такого функционирования очевидна из следующих рассуждений. Если разработчик проектирует «неизменяемую» систему, которую он считает оптимальной, то это означает, что разработчик предвидит все возможные условия на ее входе, по меньшей мере в статистическом смысле, и рассчитывает, что система будет работать при каждом из этих условий. Далее разработчик выбирает критерий, по которому должно оцениваться функционирование, например среднее число ошибок между выходным сигналом реальной системы и выходным сигналом некоторой выбранной модели или «идеальной» системы. Наконец, разработчик выбирает систему, которая оказывается лучшей в соответствии с установленным критерием функционирования, обычно из некоторого априорно ограниченного класса (например, из класса линейных систем).

Однако во многих случаях весь диапазон входных условий может быть не известен точно даже в статистическом смысле или условия могут время от времени изменяться. Тогда адаптивная система, которая, используя регулярный процесс поиска, постоянно ищет оптимум в пределах допустимого класса возможностей, имеет преимущества по сравнению с неизменяемой системой.

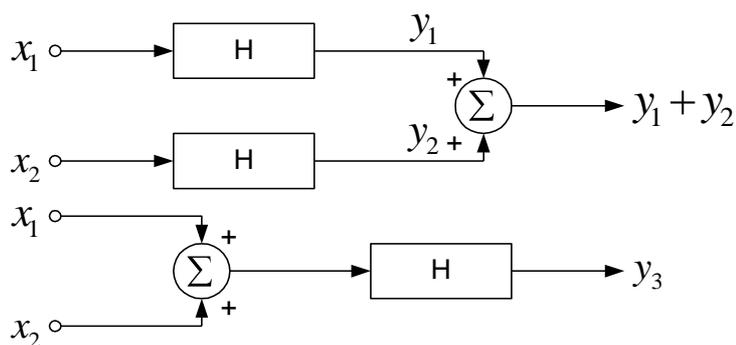


Рис. 13.1. Если Н – линейная система, то входной сигнал $y_3 = y_2 + y_1$. Если Н – адаптивная система, то y_3 в общем случае отличен от $y_2 + y_1$.

Адаптивные системы по своей природе должны быть изменяющимися во времени и нелинейными. Их свойства зависят, помимо всего прочего, от входных сигналов. Если на вход подается сигнал x_1 , то адаптивная система настроится на него и сформирует выходной сигнал — назовем его y_1 . Если на вход подается другой сигнал x_2 , то система наст-

роится на этот второй сигнал и снова сформирует выходной сигнал — назовем его y_2 . В общем случае структура и процессы коррекции адаптивной системы будут различными для двух различных входных сигналов.

Если на вход адаптивной системы подается сумма двух сигналов, то она настраивается на этот новый входной сигнал, в общем случае не равный сумме выходных сигналов $y_2 + y_1$, которые соответствовали бы входным сигналам x_1, x_2 . В этом случае, как показано на рис. 13.1, не выполняется принцип суперпозиции, имеющий место в линейных системах. Если сигнал подается на вход адаптивной системы для определения свойств по ее отклику, то система адаптируется к этому определенному входному сигналу и тем самым изменяет собственную структуру. Таким образом, адаптивные системы по существу трудно описать в обычных представлениях.

Нельзя сказать, что адаптивные системы принадлежат к абсолютно четкому подмножеству нелинейных систем. Однако присущие им две особенности в общем случае отличают их от других видов нелинейных систем. Во-первых, адаптивные системы являются регулируемыми, и процессы их регулирования зависят от усредненных в ограниченном интервале времени характеристик сигнала, а не от мгновенного значения сигналов или мгновенных значений внутренних состояний системы. Во-вторых, процессы регулирования адаптивных систем целенаправленно изменяются для того, чтобы оптимизировать заданные параметры функционирования.

Некоторые виды адаптивных систем становятся линейными системами, если их структура после адаптации остается постоянной. Их можно назвать линейными адаптивными системами, описать математически и в общем случае легче разрабатывать, чем другие виды адаптивных систем.

13.4. Адаптация с обратной связью и без обратной связи

В литературе предложено несколько способов классификации адаптивных систем. Лучше всего ее начать с разделения на адаптацию без обратной связи и с обратной связью. Процесс адаптации без обратной связи состоит из измерений характеристик входного сигнала или окружающей среды, введения этой информации в формулу или вычислительный алгоритм и использования результатов для регулирования адаптивной системы. При адаптации с обратной связью, кроме того, автоматически вносятся коррекции и с целью оптимизации параметров функционирования системы определяется их влияние на выходной сигнал. Этот процесс можно назвать адаптацией с функциональной обратной связью.

Принципы адаптации без обратной связи и с обратной связью проиллюстрированы на рис. 13.2 и 13.3. В обоих случаях удобно рассматривать адаптивный процесс так, как если бы его проводил вручную

оператор или «наблюдатель». На рис. 13.2,а и 13.3,а показано, как наблюдатель корректирует органы управления устройства обработки, наблюдая за дисплеем, регистрирующим параметры по заранее выбранному критерию функционирования. В системе без обратной связи таким критерием являются некоторые характеристики входного сигнала и, возможно, другие данные, а в системе с обратной связью — кроме того, функция, выходного сигнала. Коррекция в системе на рис. 13.3 осуществляется даже в случае, когда оператор не знает, что находится внутри устройства обработки или каковы функции органов управления. Оператор не обрабатывает входной сигнал, он только управляет коррекциями устройства обработки для того, чтобы поддерживать функционирование оптимизированным в соответствии с заранее выбранным критерием.

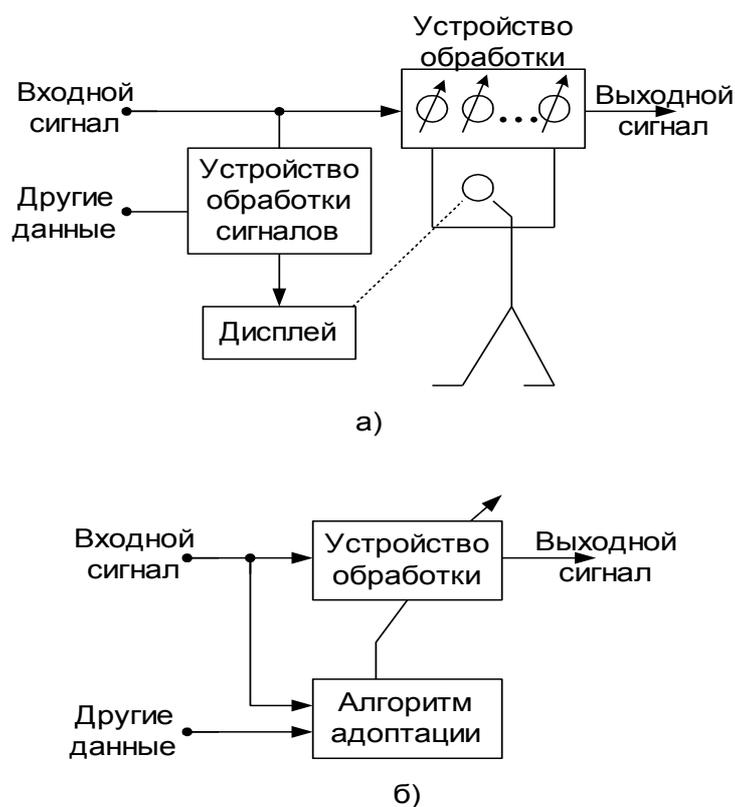


Рис. 13.2. Адаптация без обратной связи:
 а) общее представление; б) эквивалентная схема

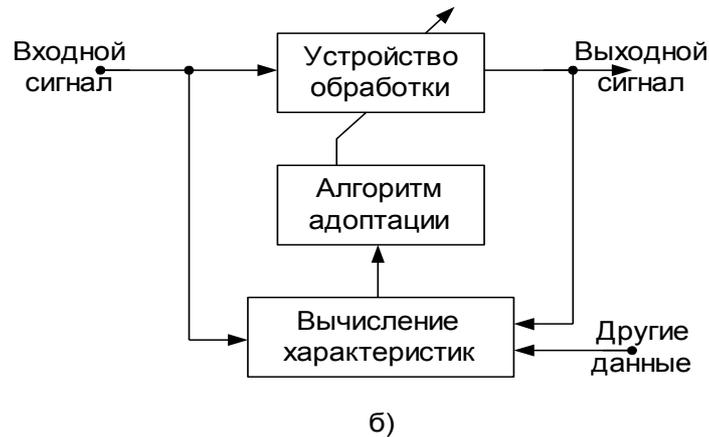
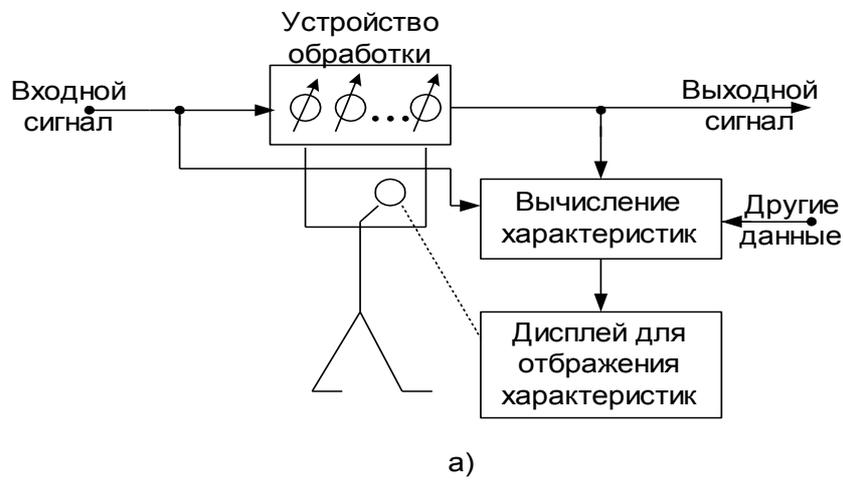


Рис. 13.3. Адаптация с обратной связью

Таким образом, функция оператора—только наблюдение, а в реальных автоматических адаптивных системах оператора заменяют вычислительные или «адаптивные» алгоритмы, как показано на рис. 13.2, б и 13.3, б. «Другими данными» для адаптивной системы на этих рисунках могут быть данные об окружающей среде или для системы с обратной связью — необходимый вид выходного сигнала.

В процессе разработки адаптивного алгоритма выбор адаптации без обратной связи или с обратной связью определяется многими факторами. Прежде всего учитывается вид входных сигналов и сигналов, отражающих параметры функционирования. Кроме того, количество вычислений и тип ЭВМ, требуемые для реализации алгоритмов адаптации без обратной связи или с обратной связью, как правило, различны. Для некоторых алгоритмов необходимо использовать универсальную цифровую ЭВМ, в то время как реализация других алгоритмов намного экономичнее на специализированных вычислителях или других устройствах. Некоторые из этих структурных подходов обсуждаются в последующих главах. Общие принципы выбора разработать трудно, но можно выделить ряд достоинств и недостатков алгоритмов адаптации с обратной связью, составляющих основное содержание этой и последующих глав.

Достоинством алгоритмов адаптации с обратной связью является их работоспособность во многих приложениях, где аналитических методов синтеза либо не существует, либо они неизвестны; например, где помимо среднеквадратических критериев используют критерии ошибки, где системы являются нелинейными или изменяются во времени, где сигналы являются нестационарными и т. д. Кроме того, адаптацию с обратной связью можно эффективно использовать в случаях, когда физические величины компонентов системы являются переменными или известны не точно. В этом случае адаптация с обратной связью обеспечит выбор наилучших величин компонентов. При частичном повреждении системы механизм адаптации, постоянно следящий за функционированием, оптимизирует ее функционирование с помощью коррекции и переоптимизации неповрежденной части. В результате при использовании функциональной обратной связи зачастую можно повысить надежность системы.

Однако процессу адаптации с обратной связью присущи и недостатки. В некоторых случаях нет одного оптимума, и тогда автоматическая оптимизация становится неопределенным процессом. В других случаях в системе управления с обратной связью процесс адаптации может быть неустойчивым, т. е. расходящимся, а не сходящимся. Несмотря на это, функциональная обратная связь представляет собой мощный, широко применяемый метод реализации адаптации. Большая часть описываемых в данной и последующих главах процессов адаптации является процессами с функциональной обратной связью.

13.5. Приложения алгоритмов адаптации с обратной связью

Рассмотрим теперь кратко некоторые приложения алгоритмов с функциональной обратной связью. Начнем с более конкретного представления процесса функциональной обратной связи (см. рис. 13.3, б) на рис. 13.4. Обозначим через x входной сигнал и примем в качестве «требуемого отклика» сигнал d , положим, что он представляет собой требуемый выходной сигнал адаптивной системы. В рассматриваемом случае сигнал d является «другими данными» на рис. 13.3, б.

Сигнал ошибки ε — это разность между требуемым и действительными выходными сигналами адаптивной системы. С помощью минимизации некоторого параметра сигнала ошибки адаптивный алгоритм позволяет изменять характеристики отклика, тем самым замыкая петлю функциональной обратной связи. Покажем в общем виде, как работает схема, приведенная на рис. 13.4, в реальных условиях.

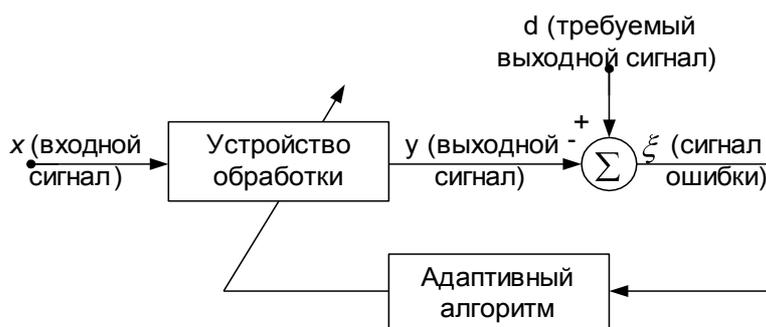


Рис. 13.4. Сигналы при адаптации с обратной связью

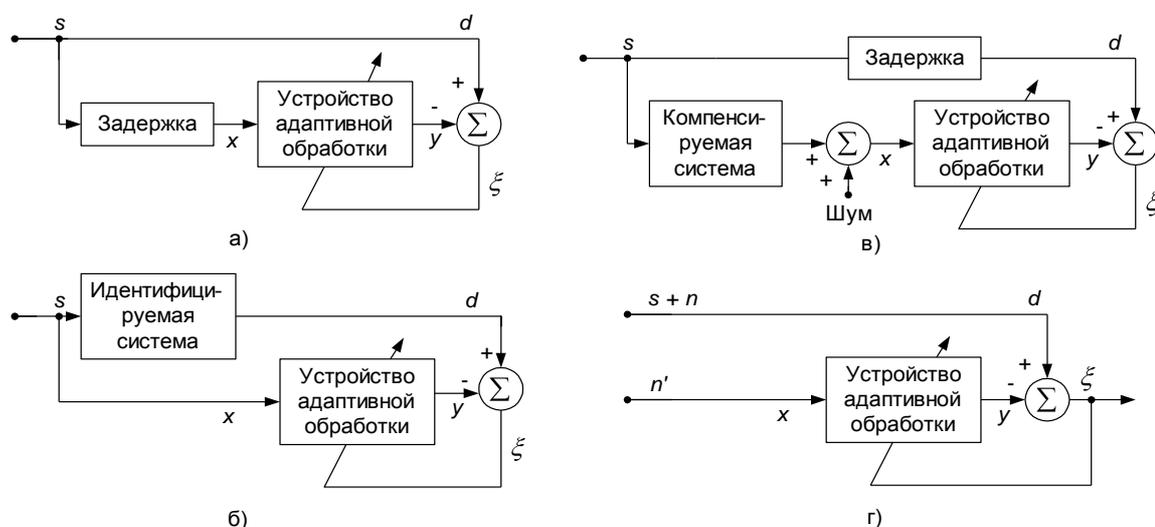


Рис. 13.5. Примеры применения схемы, приведенной на рис. 13.4., для предсказания (а), идентификация (моделирования) (б), выравнивая (компенсации) характеристик (в), подавления помехи (г).

На рис. 13.5 дано несколько примеров применения адаптивных систем. Отметим, что схема на рис. 13.4, иллюстрирующая основной принцип процесса адаптации с обратной связью, немного упрощена и является составной частью всех схем на рис. 13.5, а способ получения требуемого сигнала d зависит от конкретной прикладной задачи.

Простейшей из четырех схем, по-видимому, является представленная на рис. 13.5, а схема применения адаптивной системы в качестве устройства предсказания. Полезным сигналом является входной сигнал s , поступающий после задержки в устройство адаптивной обработки,

которое должно стремиться «предсказать» текущий входной сигнал для того, чтобы сигнал u компенсировал сигнал d и сводил ε к нулю.

Таким же простым для понимания является представленный на рис. 1.5,б пример применения адаптивной системы в задаче идентификации систем. Здесь широкополосный сигнал s служит входным сигналом для устройства адаптивной обработки, а также для неизвестной «идентифицируемой» системы (термин, заимствованный из литературы по управлению). Для уменьшения ε устройство адаптивной обработки стремится воспроизвести передаточную функцию идентифицируемой системы. После адаптации идентифицируемая система является «идентифицированной» в том смысле, что ее передаточную функцию можно задать, по существу, в виде передаточной функции устройства адаптивной обработки. Адаптивную идентификацию или моделирование системы можно использовать для моделирования медленноменяющейся системы, в которой имеются входные и выходные сигналы, например при изучении вибраций механических систем.

Применение адаптивной системы в качестве устройства выравнивания (компенсации) рассматривается на рис. 13.5, б. В этом случае устройство адаптивной обработки стремится восстановить задержанный сигнал s . Считается, что сигнал преобразуется в медленноменяющейся компенсируемой системе и содержит аддитивный шум. Задержка на рис. 13.5,б предназначена для компенсации задержки сигнала при прохождении через компенсируемую систему и устройство адаптивной обработки. Адаптивное выравнивание можно использовать для исключения влияния преобразователей, каналов связи и некоторых других систем или для формирования характеристики, обратной характеристике некоторой неизвестной компенсируемой системы. Кроме того, возможно применение таких систем при разработке цифровых фильтров, а также в задачах адаптивного управления и т. д.

Наконец, на рис. 13.5,г показано устройство адаптивной обработки в схеме подавления помехи. Здесь сигнал s искажен аддитивным шумом n , и, кроме того, имеется искаженный, но коррелированный шум n' . В этом случае назначение устройства адаптивной обработки — сформировать выходной сигнал u , близкий к шуму n , так, чтобы общий выходной сигнал ε приближался к сигналу s . Далее будет показано, что при определенных, весьма распространенных условиях оптимальным является устройство адаптивной обработки, минимизирующее среднее квадратическое значение сигнала ε .

13.6. Простой пример адаптивной системы

Теперь уже ясно, что адаптивная обработка сигналов — очень общее и фундаментальное понятие, которое подразумевает использование самонастраивающихся устройств обработки с изменяющимися во времени параметрами. Считают, что функционирование этих систем является целенаправленным, полезным, а иногда даже «интеллектуальным» в некотором смысле. Прежде чем завершить вводную

главу, для более наглядной иллюстрации процесса адаптации в целом и адаптации с обратной связью, в частности, приведем всего один конкретный пример такой системы. Будем использовать схему «предсказания», представленную на рис. 13.5,а, имея в виду, что она является лишь одним частным примером очень широкого и общего класса систем.

На рис. 13.6 схема адаптивного устройства предсказания рис. 13.5,а воспроизведена с условными обозначениями, используемыми далее по всему тексту. Символы x_k, y_k, ε_k представляет собой k -е элементы временных последовательностей x, y, ε . Обычно полагают, что временные последовательности получены в виде отсчетов непрерывных сигналов. Например, $x_k = x(kT)$ и т. д., где T —временной шаг, или интервал между отсчетами.

Символ z^{-M} обозначает фиксированную задержку на M временных шагов, поэтому на рис. 13.6 выходной сигнал z^{-M} обозначен через x_{k-M} . Символ H_k представляет собой передаточную характеристику устройства адаптивной обработки — адаптивного фильтра. Подробное описание H_k приведено в последующих главах, а здесь отметим, что индекс k говорит о том, что передаточная характеристика изменяется в каждой точке отсчета.

Таким образом, на рис. 1.6 показано, что входной сигнал x_k задерживается, затем в результате фильтрации преобразуется в сигнал y_k , выделяемый из x_k . для получения сигнала ошибки ε_k .

Передаточная функция H_k корректируется таким образом, чтобы поддерживать наименьшее среднеквадратическое значение ошибки. В этом случае устройство обработки всегда использует предыдущие значения x для предсказания текущего значения x , в то же время используя ε для коррекции H , и таким образом включает в себя процесс функциональной обратной связи.

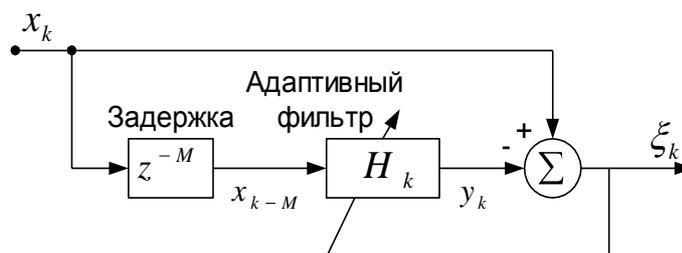


Рис. 13.6. Пример адаптивной системы: адаптивное устройство предсказания

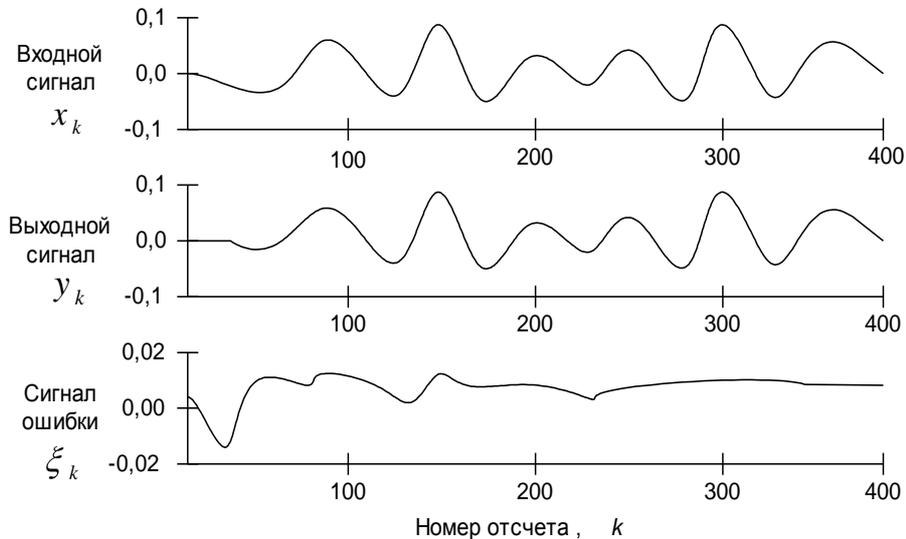


Рис. 13.7. Сигналы адаптивного устройства предсказания, приведенного на рис. 13.6. По мере того, как система обучается для предсказания входного сигнала x , выходной сигнал y приближается к x и сигнал ошибки ε стремится к нулю

На рис. 13.7 приведен пример функционирования адаптивного устройства предсказания. Здесь, а также в последующих примерах сигналы построены на основе цифровых данных соединением точек отсчетов прямыми линиями. Входной сигнал x является ограниченным по частоте случайным сигналом, в этом случае $M=1$. Отметим, что с увеличением k и по мере того, как адаптивная система постепенно обучается предсказанию, y_k приближается к x_k и амплитуда ε_k уменьшается.

XIV. АДАПТИВНЫЙ ЛИНЕЙНЫЙ СУММАТОР

14.1. Общее описание адаптивного линейного сумматора

Адаптивный линейный сумматор, или нерекурсивный адаптивный фильтр, является фундаментальным понятием в адаптивной обработке сигналов. В том или ином виде он присутствует в большинстве адаптивных фильтров и систем и является единственным наиболее важным элементом «обучения» систем и устройств адаптивной обработки в целом.

Структура адаптивного линейного сумматора является нерекурсивной, он представляет собой нерекурсивный фильтр с изменяющимися во времени параметрами, так что принцип его действия очень прост, а функционирование, способы адаптации, а также реализация в различных схемах вполне доступны для понимания. Более того, известны такие приложения, в которых функционирование адаптивного линейного сумматора является в некотором смысле «наилучшим».

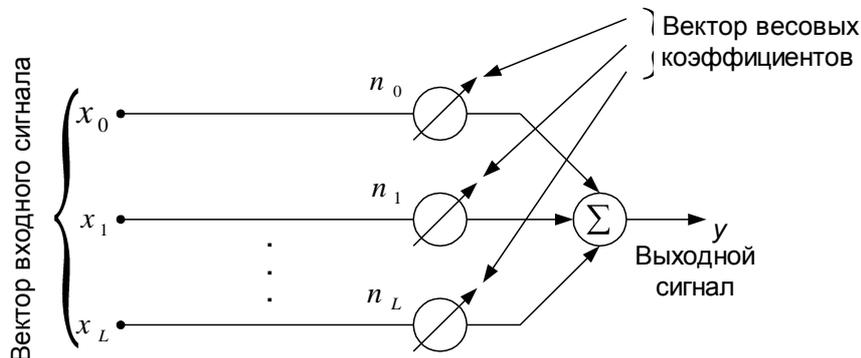


Рис. 14.1. Общий вид адаптивного сумматора

Схема адаптивного линейного сумматора в общем виде показана на рис. 14.1. В этой схеме имеются вектор входного сигнала с компонентами x_0, x_1, \dots, x_L , соответствующее множество регулируемых весовых коэффициентов $\omega_0, \omega_1, \dots, \omega_L$, устройство суммирования и выходной сигнал y . Процесс регулирования или адаптации весовых коэффициентов называют «весовой коррекцией», коррекцией коэффициента передачи или процессом адаптации. Сумматор называют линейным, поскольку для некоторого заданного набора весовых коэффициентов выходной сигнал представляет собой линейную комбинацию компонентов входного сигнала. Однако в процессе коррекции весовых коэффициентов последние также являются функциями компонентов входного сигнала и выходной сигнал уже не является линейной функцией входного. Таким образом, в соответствии с общим принципом, изложенным в гл. 13, правило функционирования адаптивного линейного сумматора становится нелинейным.

14.2. Векторы входного сигнала и весовых коэффициентов

Существуют две важные физические интерпретации элементов вектора входного сигнала (см. рис. 14.1). В первой их можно рассматривать как *одновременно действующие* входные сигналы от $L+1$ *различных* источников. Примерами такой интерпретации могут служить адаптивная антенна и адаптивная акустическая система обнаружения, в которых каждый вход соединен с отдельным чувствительным элементом.

Кроме того, элементы x_0, \dots, x_L , можно рассматривать как $L+1$ *последовательных* отсчетов сигнала *одного источника*. Примером такой интерпретации является устройство адаптивной обработки, приведенное в гл. 13 на рис. 13.6.

Будем называть эти две интерпретации системой с многими входами и системой с одним входом. Для обоих случаев принято обозначать входные векторы по-разному, а именно:

$$\text{Для многих входов} \quad X_k = [x_{0k} \ x_{1k} \ \dots \ x_{Lk}]^T \quad (14.1)$$

Для одного входа $X_k = [x_k x_{k-1} \dots x_{k-L}]^T$

(14.2)

В этих выражениях T — знак транспонирования, поэтому в обоих случаях X_k — фактически вектор-столбец. Индекс k используется для обозначения времени. Таким образом, в системе с многими входами все элементы получены на k -м временном отсчете, тогда как в системе с одним входом элементы являются последовательными отсчетами, взятыми в моменты $k, k-1, \dots, k-L$.

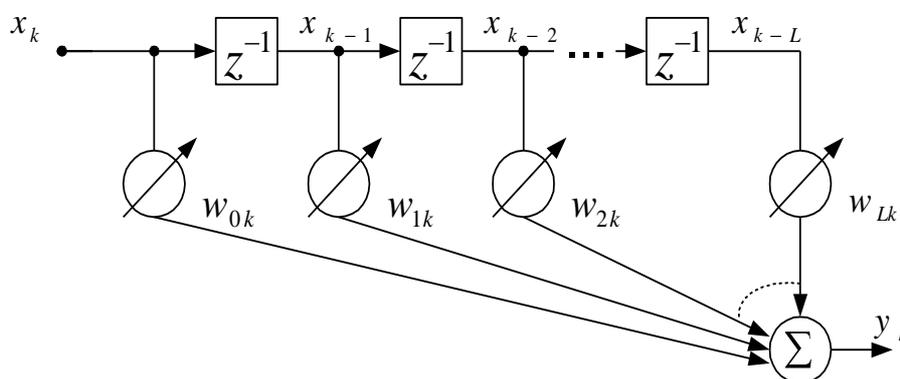


Рис. 14.2. Схема адаптивного линейного сумматора в виде адаптивного трансверсального фильтра с одним входом

В системе с одним входом устройство адаптивной обработки можно реализовать в виде адаптивного линейного сумматора с элементами задержки (рис. 14.2). Такую структуру называют адаптивным трансверсальным фильтром. Заметим, что весовые коэффициенты имеют второй индекс k , который добавлен для того, чтобы показать их зависимость от времени в явном виде. Адаптивный трансверсальный фильтр является временной (в противоположность пространственной) формой нерекурсивного адаптивного фильтра и широко применяется при адаптивном моделировании и адаптивной обработке сигналов. Большинство адаптивных систем, рассматриваемых в последующих главах, основаны на использовании адаптивного трансверсального фильтра.

В некоторых системах с многими входами необходимо вводить *весовой коэффициент смещения*, тогда к сумме y_k прибавляется переменное смещение. При необходимости его удобно вводить, полагая первый входной элемент x_{0k} в (14.1) постоянным и равным единице (или какой-либо другой постоянной величине), как показано на рис. 14.3. В

системах с одним входом, как правило, не требуется вводить весовой коэффициент смещения.

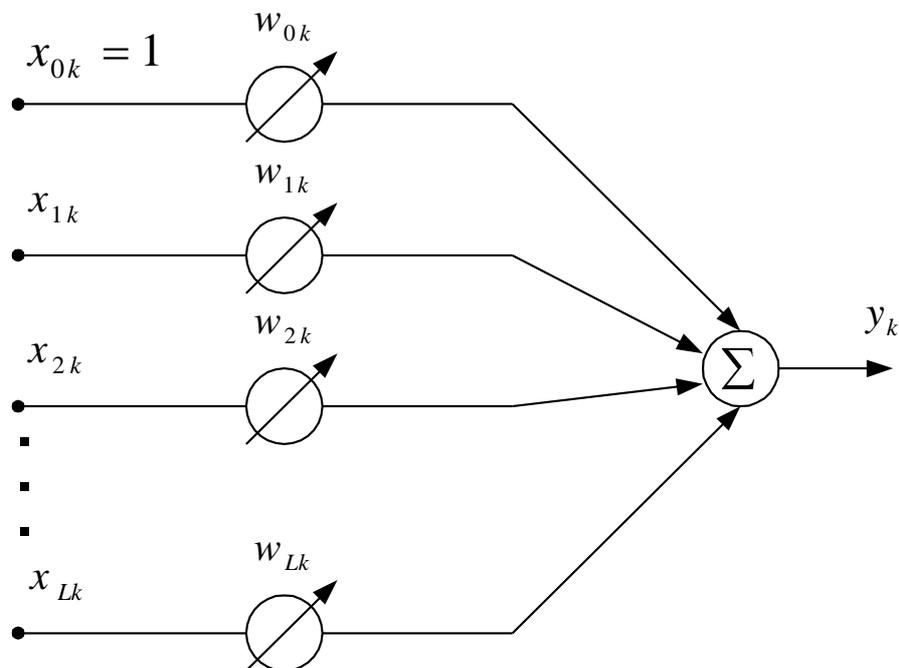


Рис. 14.3. Схема адаптивного линейного сумматора с многими входами и с весовым коэффициентом смещения ω_{0k}

Используя обозначения входного сигнала (14.1) и (14.2), получаем следующие выражения для выходных сигналов в схемах на рис. 14.2 и 14.3:

для одного входа
$$y_k = \sum_{l=0}^L \omega_{lk} x_{k-l}$$
 (14.3)

для многих входов
$$y_k = \sum_{l=0}^L \omega_{lk} x_{lk}$$
 (14.4)

Если в (14.4) x_{0k} тождественно равен единице, то, как отмечено выше, ω_{0k} становится весовым коэффициентом смещения. По аналогии с (14.1) и (14.2) вектор весовых коэффициентов

$$W_k = [\omega_{0k}, \omega_{k1}, \dots, \omega_{kL}]^T$$

(14.5)

Тогда, используя векторные обозначения, можно записать (14.3) и (14.4) одним соотношением

$$y_k = X_k^T W_k = W_k^T X_k$$

(14.6)

Теперь можно рассмотреть принцип работы адаптивного линейного сумматора и изучить влияние зависимости вектора W_k от индекса времени k .

14.3. Полезный отклик и сигнал ошибки

Адаптивный линейный сумматор можно использовать в системах как с обратной, так и без обратной связи. Как отмечалось при рассмотрении рис. 13.2, процесс коррекции вектора весовых коэффициентов в системе без обратной связи в явном виде не зависит от выходного сигнала и определяется только входным сигналом и состоянием окружающей среды.

Однако в системе с обратной связью, показанной на рис. 13.3, вектор весовых коэффициентов зависит от выходного сигнала, а также от других данных. Вообще в адаптивном линейном сумматоре другими данными являются полезный отклик или обучающий сигнал. В учебном пособии обсуждаются прежде всего системы с функциональной обратной связью, поэтому необходимо четкое понимание существа этих сигналов.

В процессе адаптации с функциональной обратной связью вектор весовых коэффициентов линейного сумматора корректируется таким образом, чтобы выходной сигнал y_k , имел наилучшее приближение к полезному отклику. Для этого выходной сигнал сравнивается с полезным откликом, формируется сигнал ошибки и затем корректируется или оптимизируется вектор весовых коэффициентов, минимизирующий сигнал ошибки. В большинстве практических случаев процесс адаптации направлен на минимизацию среднеквадратического значения, или средней мощности сигнала ошибки. Оптимизация по этому критерию, как в адаптивных, так и в неадаптивных системах давно и широко применяется и имеет много достоинств.

На рис. 14.4 показан способ получения сигнала ошибки в системе с многими входами за счет введения полезного отклика. Для формирования сигнала ошибка ε_k выходной сигнал y_k просто вычитается из полезного сигнала d_k .

Источник сигнала полезного отклика d_k определяется конкретным применением адаптивного сумматора. Пока будем считать, что такой сигнал уже имеется. Более подробно вопрос его формирования рассматривается в последующих главах. Однако отметим, что зачастую для того, чтобы найти подходящий сигнал, требуется значительная изобретательность, поскольку если бы в действительности полезный сигнал имелся, то адаптивная система была бы не нужна.

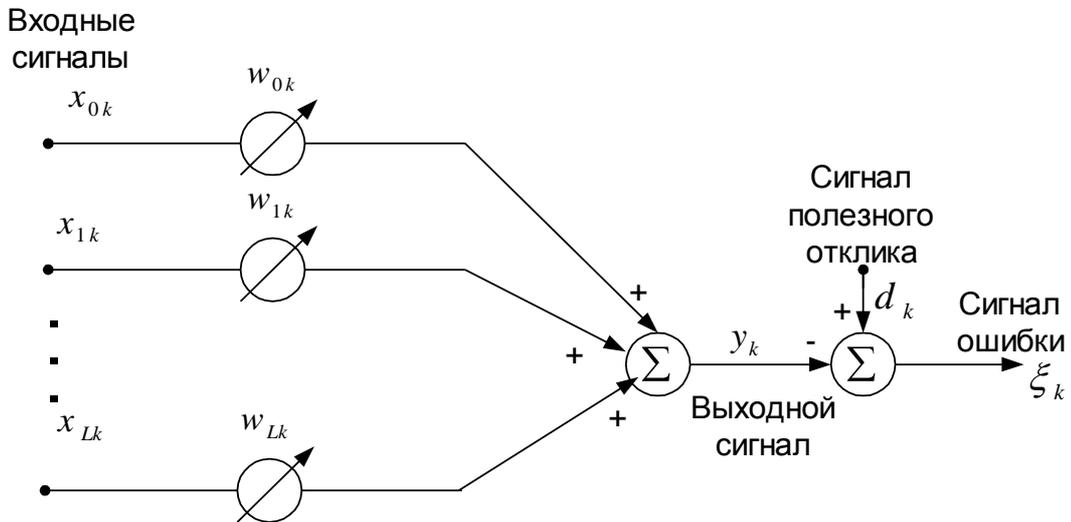


Рис. 14.4. Сигналы полезного отклика и ошибки в адаптивном линейном сумматоре с многими входами

Рассмотрим теперь только что введенную рабочую функцию.

14.4 Рабочая функция

Из рис. 14.4 следует, что сигнал ошибки с временным индексом k

$$\varepsilon_k = d_k - y_k$$

(14.7)

Подставляя (14.6) в это выражение, получаем

$$\varepsilon_k = d_k - X^T_k W = d_k - W^T X_k$$

(14.8)

Здесь для удобства у вектора весовых коэффициентов опущен индекс k , поскольку в данном случае коррекция весовых коэффициентов не рассматривается. Чтобы получить мгновенное квадратичное значение сигнала ошибки, возведем в квадрат выражение (14.8):

$$\varepsilon^2_k = d^2_k - W^T X^T_k X_k W - 2d_k W X^T_k$$

(14.9)

Положим, что ε_k , d_k и X_k стационарны в статистическом смысле, и найдем математическое ожидание функции (14.9) по k :

$$E[\varepsilon^2_k] = E[d^2_k] - W^T E[X^T_k X_k] W - 2E[d_k X^T_k] W$$

(14.10)

Отметим, что математическое ожидание суммы равно сумме математических ожиданий, но математическое ожидание произведений равно

произведению математических ожиданий только тогда, когда случайные величины статистически независимы. В общем случае сигналы x_k и d_k не являются независимыми.

Удобнее функцию СКО представить следующим образом.

Пусть R — квадратная матрица

$$R = E[X_k X_k^T] = E \begin{bmatrix} x_{0k}^2 & x_{0k}x_{1k} & x_{0k}x_{2k} & \dots & x_{0k}x_{Lk} \\ x_{1k}x_{0k} & x_{1k}^2 & x_{1k}x_{2k} & \dots & x_{1k}x_{Lk} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{Lk}x_{0k} & x_{Lk}x_{1k} & x_{Lk}x_{2k} & \dots & x_{Lk}^2 \end{bmatrix} \quad (14.11)$$

Эта матрица называется корреляционной матрицей входного сигнала. Элементы, расположенные на главной диагонали, равны среднеквадратическим значениям входных компонентов, а остальные элементы — значениям взаимокорреляционной функции входных компонентов. Таким образом, пусть P —вектор-столбец

$$P = E[X_k X_k^T] = E[d_k x_{0k} d_k x_{1k} \dots d_k x_{Lk}] \quad (14.12)$$

Этот вектор представляет собой множество значений взаимокорреляционной функции отсчетов полезного отклика и отсчетов входного сигнала. Если X_k и d_k —стационарны, то все элементы, как R , так и P , являются постоянными статистиками второго порядка. Отметим, что в (14.11) и (14.12) X_k представлено для системы с многими входами, но так же легко можно было бы воспользоваться представлением для системы с одним входом.

Обозначим теперь СКО в (14.10) через ξ и запишем ее с помощью (14.11) и (14.12):

$$\text{СКО} \underline{\Delta} \xi = E[\varepsilon_k^2] = E[d_k^2] + W^T R W - 2P^T W \quad (14.13)$$

Из этого выражения видно, что если отсчеты входного сигнала и полезного отклика — стационарные случайные величины, то СКО точно совпадает с квадратичной функцией компонентов вектора весовых коэффициентов W , т. е. если раскрыть (14.13), то элементы W входят в (14.12) только в первой и второй степенях.

На рис. 14.5 показан фрагмент характерной двумерной функции СКО. По вертикальной оси откладываются значения СКО, по горизонтальным осям — значения весовых коэффициентов. Построенный таким способом график квадратичной функции ошибки (поверхность, образованная графиком рабочей функции) является параболоидом (или гиперпараболоидом, если число весовых коэффициентов больше двух). Он должен быть вогнутым и направленным вверх; в противном случае при некоторых значениях весовых коэффициентов значение СКО было бы отрицательным, что невозможно для реальных физических сигналов. При

ξ , равной в (14.13) константе, сечение поверхности имеет эллиптическую форму. Проекция «нижней» точки графика на плоскость векторов весовых коэффициентов представляет собой вектор оптимальных весовых коэффициентов W^* и соответствует точке минимального значения СКО. Квадратическая функция ошибки имеет только один глобальный оптимум, локальных минимумов у такой функции не существует.

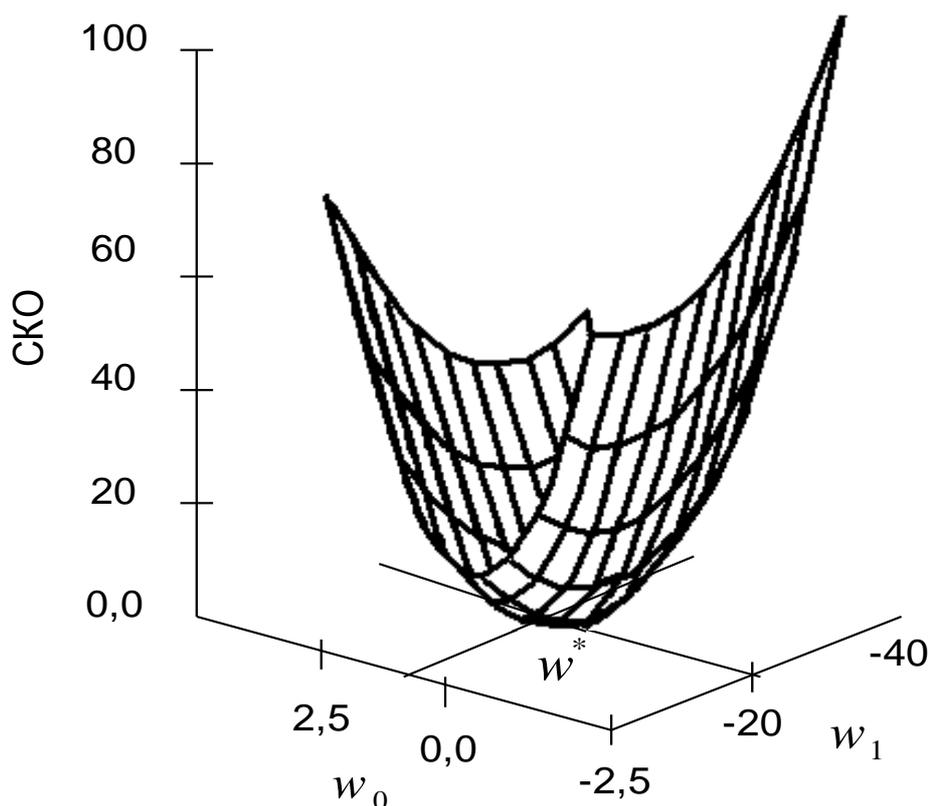


Рис. 14.5. Фрагмент графика двумерной квадратичной рабочей функции. В данном примере оптимальным является вектор $W^* = (0,65 - 2,1)$, а минимальное значение СКО равно нулю

14.5. Минимальная среднеквадратическая ошибка и градиент

Во многих полезных для практики способах адаптации поиск вектора весовых коэффициентов, соответствующего минимуму рабочей функции, осуществляется градиентными методами. Градиент функции СКО, обозначаемый $\nabla(\xi)$ или просто ∇ , можно получить дифференцированием функции (14.13), при этом вектор-столбец

$$\nabla_{\Delta} \frac{d\xi}{dW} = \left[\frac{d\xi}{dw_0} \quad \frac{d\xi}{dw_1} \quad \dots \quad \frac{d\xi}{dw_L} \right]^T$$

(14.14)

$$\nabla = 2RW - 2P$$

(14.15)

Величины R и P определяются по (14.11) и (14.12). Это выражение получено дифференцированием функции (14.13) по каждому из

компонентов вектора весовых коэффициентов. Дифференцирование члена $W^T R W$ можно осуществить дифференцированием произведения $(W^T)(R W)$.

Для нахождения минимального значения СКО полагаем, что вектор весовых коэффициентов W равен оптимальному W^* , градиент которого равен нулю:

$$\nabla = 0 = 2R W^* - 2P \quad (14.16)$$

Полагая, что R является неособенной матрицей, из (14.16) находим вектор W^* , иногда называемый винеровским вектором весовых коэффициентов:

$$\nabla = R^{-1} P \quad (14.17)$$

Это равенство является уравнением Винера—Хопфа, записанным в матричной форме. Подставляя теперь (14.17) в (14.13), получаем минимальное значение СКО:

$$\xi_{\min} = E[d^2_k] + W^{*T} R W^* - 2P^T W^* = E[d^2_k] + [R^{-1} P]^T R R^{-1} P - 2P^T R^{-1} P \quad (14.18)$$

Упростим полученный результат, используя следующие три свойства, которые полезны при рассмотрении рабочей функции СКО:

1. Для любой квадратной матрицы существует единичная матрица:

$$A A^{-1} = I$$

2. Транспонирование произведения матриц:

$$[A B]^T = B^T A^T$$

3. Симметричность корреляционной матрицы входного сигнала:

$$R^T = R;$$

$$[R^{-1}]^T = R^{-1}$$

[см. (14.11)].

В соответствии с этими свойствами (14.18) принимает вид

$$\xi_{\min} = E[d^2_k] - P^T R^{-1} P = E[d^2_k] - P^T W^* \quad (14.19)$$

Теперь для того чтобы пояснить введенные понятия квадратичной поверхности, градиента и СКО, рассмотрим пример.

14.6. Пример анализа рабочей функции

На рис. 14.6 приведен простой пример адаптивного линейного сумматора с одним входом и двумя весовыми коэффициентами. Входной и полезный сигналы представляют собой отсчеты синусоиды с частотой, равной N отсчетам за период. Для того чтобы не все отсчеты были равны нулю, полагаем $N > 2$. Здесь не обсуждается способ получения этих сигналов, а рассматривается только рабочая функция и ее свойства.

Для нахождения рабочей функции, т. е. ξ в (14.13), необходимо вычислить математическое ожидание произведений сигналов в (14.11) и (14.12). Отметим, что для системы с одним входом необходимо изменить индексы при x в соответствии с (14.2).

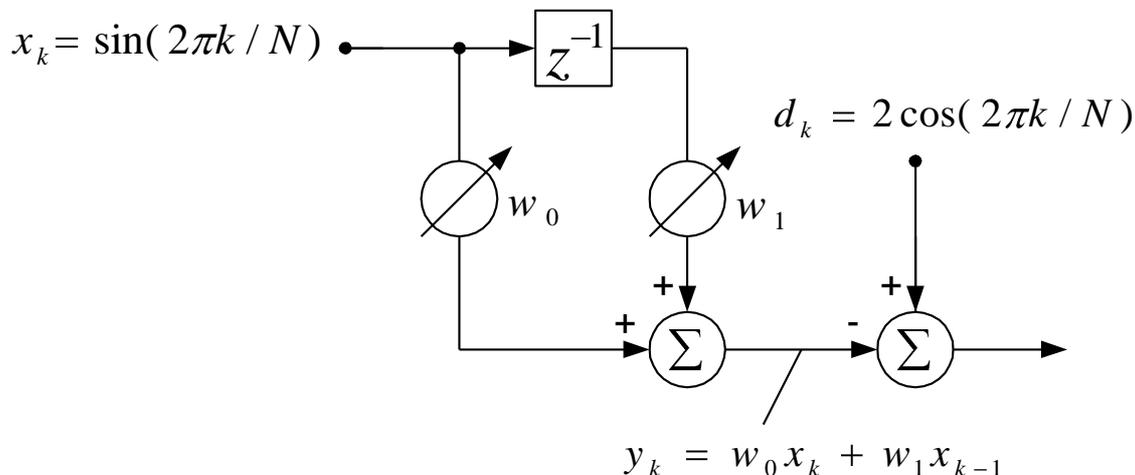


Рис. 14.6. Пример адаптивного линейного сумматора с двумя весовыми коэффициентами

Для любого произведения синусоидальных функций математическое ожидание произведения можно найти усреднением этого произведения за один или более периодов. Таким образом,

$$E[x_k x_{k-n}] = \frac{1}{N} \sum_{k=1}^N \sin \frac{2\pi k}{N} \sin \frac{2\pi(k-n)}{N} = 0.5 \cos \frac{2\pi n}{N} \quad n=0, \quad 1; \quad (14.20)$$

$$E[d_k x_{k-n}] = \frac{2}{N} \sum_{k=1}^N \cos \frac{2\pi k}{N} \sin \frac{2\pi(k-n)}{N} = -\sin \frac{2\pi n}{N} \quad n=0, \quad 1; \quad (14.21)$$

Отметим, что $E[x_k^2] = E[x_{k-1}^2]$, поскольку усреднение осуществляется по индексу k .

На основании этих результатов можно получить выражения для корреляционных матриц входного сигнала R (14.11) и вектора P (14.12) в рассматриваемом примере системы с одним входом и двумя весовыми коэффициентами:

$$R = E \begin{bmatrix} x_k^2 & x_k x_{k-1} \\ x_{k-1} x_k & x_{k-1}^2 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \cos \frac{2\pi}{N} \\ 0.5 \cos \frac{2\pi}{N} & 0.5 \end{bmatrix} \quad (14.22)$$

$$P = E[d_k x_k, d_k x_{k-1}]^T = [0 \quad -\sin \frac{2\pi}{N}]^T$$

(14.23)

Так же, как и в (14.20) и (14.21), получаем, что $E[d_k^2] = 2$. Подставляя полученные результаты в (14.13), находим функцию СКО ошибки для нашего примера:

$$\begin{aligned} \xi &= E[d_k^2] + W^T R W - 2P^T W = 2 + 0.5[\omega_0 \omega_1] \begin{bmatrix} 1 & \cos \frac{2\pi}{N} \\ \cos \frac{2\pi}{N} & 1 \end{bmatrix} \begin{bmatrix} \omega_0 \\ \omega_1 \end{bmatrix} - 2[0 \quad -\sin \frac{2\pi}{N}] \begin{bmatrix} \omega_0 \\ \omega_1 \end{bmatrix} = \\ &= 0.5(\omega_0^2 + \omega_1^2) + \omega_0 \omega_1 \cos \frac{2\pi}{N} - 2\omega_1 \sin \frac{2\pi}{N} + 2 \end{aligned}$$

4.24)

График этой функции для $N=5$ приведен на рис. 14.5. Отметим, что зависимость является квадратичной по ω_0 и ω_1 имеет единственный глобальный минимум. Подставив (14.22) и (14.23) в (14.15), найдем вектор градиента для любой точки ω_0, ω_1 :

$$\begin{aligned} \nabla &= 2RW - 2P = \\ &= \begin{bmatrix} 1 & \cos \frac{2\pi}{N} \\ \cos \frac{2\pi}{N} & 1 \end{bmatrix} \begin{bmatrix} \omega_0 \\ \omega_1 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ -\sin \frac{2\pi}{N} \end{bmatrix} = \\ &= \begin{bmatrix} \omega_0 + \omega_1 \cos \frac{2\pi}{N} \\ \omega_0 \cos \frac{2\pi}{N} + \omega_1 + 2 \sin \frac{2\pi}{N} \end{bmatrix} \end{aligned}$$

(14.25)

В этом примере винеровский вектор весовых коэффициентов можно найти формально из (14.17), вычислив обратную матрицу R или приравняв ∇ нулю в (14.25). Конечно, обе операции эквивалентны, и в обоих случаях в результате имеем

$$W^* = [2 \operatorname{ctg} \frac{2\pi}{N} - 2 \operatorname{cosec} \frac{2\pi}{N}]^T$$

(14.26)

Напомним, что ранее принято $N > 2$, поэтому здесь ω_0^* и ω_1^* всегда конечны.

Наконец, подставляя (14.23) и (14.26) в (14.19), находим для данного примера минимальное значение СКО

$$\xi_{\min} = E[d_k^2] - P^T W^* = 2 - \begin{bmatrix} 0 & -\sin \frac{2\pi}{N} \end{bmatrix} \begin{bmatrix} 2 \operatorname{ctg} \frac{2\pi}{N} \\ -2 \operatorname{cosec} \frac{2\pi}{N} \end{bmatrix} = 0$$

(14.27)

С первого взгляда полученный результат может показаться удивительным, поскольку для любого значения N весовые коэффициенты в схеме на рис. 14.6 можно скорректировать так, чтобы свести ξ_k к нулю. Сам по себе элемент задержки может изменить x_k с синусоидальной на косинусоидальную функцию только при $N=4$, т. е. только при задержке, равной четверти периода. В этом случае из (14.26) следует, что $\omega^*_0 = 0$ и $\omega^*_1 = -2$. Однако в адаптивном линейном сумматоре с задержкой и двумя весовыми коэффициентами всегда возможен такой сдвиг x_k , при котором для любого $N > 2$ можно получить соответствующую косинусоидальную функцию.

14.7. Альтернативное представление градиента

Поскольку СКО является квадратичной функцией W , достигающей своего минимального значения при $W=W^*$, можно записать

$$\xi = \xi_{\min} + (W - W^*)^T R (W - W^*)$$

(14.28)

Покажем, что это выражение справедливо. Отметим, что в общем случае $(A - B)^T = A^T - B^T$, раскроем скобки в (14.28) и найдем

$$\xi = \xi_{\min} + W^{*T} R W^* + W^T R W - W^T R W^* - W^{*T} R W$$

(14.29)

Каждый член в (14.29) является скалярной величиной и поэтому равен своему транспонированному значению. Следовательно, последние два члена равны друг другу. В результате подставив вместо ξ_{\min} выражение (14.19), запишем

$$\xi = E[d_k^2] - P^T W^* + W^{*T} R W^* + W^T R W - 2W^T R W^*$$

(14.30)

Подставляя вместо W^* выражение (14.17) и имея в виду, что R — симметрическая матрица, получаем

$$\begin{aligned} \xi &= E[d_k^2] - P^T R^{-1} P + P^T R^{-1} R R^{-1} P + W^T R W - 2W^T R R^{-1} P = \\ &= E[d_k^2] + W^T R W - 2W^T P = E[d_k^2] + W^T R W - 2P^T W \end{aligned}$$

(14.31)

Этот результат соответствует (14.13) и тем самым доказывает справедливость выражения (14.28).

Квадратичную форму в (14.28) можно привести к более удобному виду, если ввести вектор отклонения весовых коэффициентов

$$V = W - W^* = [v_0 v_1 \dots v_L]^T$$

(14.32)

Соответственно выражение (14.28) принимает вид

$$\xi = \xi_{\min} + V^T R V$$

(14.33)

Вектор V — отклонение вектора весовых коэффициентов от винеровского оптимального вектора весовых коэффициентов. Любое отклонение W от W^* вызывает в соответствии с квадратичной формой $V^T R V$ увеличение СКО.

Чтобы при всех возможных $V\xi$; было неотрицательным, необходимо выполнение для всех V условия $V^T R V \geq 0$. Если $V^T R V > 0$ для всех $V \neq 0$, то говорят, что матрица R — положительно определенная. Если $V^T R V = 0$ для всех или некоторого конечного множества векторов V , то говорят, что матрица R — положительно полуопределенная. В практических случаях R почти всегда является положительно определенной, но иногда может быть и положительно полуопределенной. Условия положительной определенности и положительной полуопределенности обычно рассматриваются в теории матриц.

Градиент СКО относительно V получаем дифференцированием функции (14.33):

$$\frac{d\xi}{dV} = \left[\frac{d\xi}{dv_0} \frac{d\xi}{dv_1} \dots \frac{d\xi}{dv_L} \right] = 2RV$$

(14.34)

Этот градиент такой же, как и в формуле (14.15), так как W и V отличаются только на константу. Таким образом,

$$\nabla = \frac{d\xi}{dW} = \frac{d\xi}{dV} = 2RV = 2(RW - P)$$

(14.35)

Выражение (14.35) будет использовано при синтезе и анализе различных адаптивных алгоритмов.

14.8. Декорреляция сигнала ошибки и элементов входного сигнала

При $W=W^*$ имеет место полезное и важное статистическое соотношение между сигналом ошибки и компонентами вектора входного сигнала. В соответствии с (14.8)

$$\varepsilon_k = d_k - X_k^T W$$

(14.36)

Умножим на X_k обе части этого равенства. Поскольку каждый член является скалярной величиной, его можно умножать на X_k как слева, так и справа. Тогда

$$\varepsilon_k X_k = d_k X_k - X^k X^T_k W$$

(14.37)

Далее находим математическое ожидание функции (14.37):

$$E[\varepsilon_k X_k] = P - RW$$

(14.38)

Наконец, пусть W равен оптимальному значению (14.17), при этом

$$E[\varepsilon_k X_k]_w = w^* = P - P = 0$$

(14.39)

Этот результат совпадает с хорошо известным результатом винеровской теории фильтрации: когда импульсный отклик фильтра оптимизирован, сигнал ошибки не коррелирован (ортогонален) с входными сигналами, взятыми с весовыми коэффициентами.

XV. ПОИСК РАБОЧЕЙ ФУНКЦИИ

Как было показано выше, для адаптивного линейного сумматора функция СКО является квадратичной, если входные сигналы и требуемый отклик стационарны в статистическом смысле. Во многих представляющих интерес случаях параметры этой квадратичной функции неизвестны и нет ее аналитического описания. Однако, усредняя квадрат сигнала ошибки за некоторый период времени, можно измерить или оценить положение точек на квадратичной поверхности. Задача состоит в том, чтобы разработать аналитические методы или алгоритмы, позволяющие осуществлять поиск параметров рабочей функции и находить оптимальный вектор весовых коэффициентов только по данным измерения или оценки. В большинстве практических методов поиск параметров не проводится, а оптимальное или близкое к нему решение находится введением контрольных расстроек.

15.1. Методы поиска параметров рабочей функции

Эта глава посвящена разработке алгоритмов для двух широко известных методов поиска параметров рабочей функции: Ньютона и наискорейшего спуска. Для определения направления, на котором расположен минимум функции, используются градиентные оценки. Поэтому их называют методами спуска. В дальнейшем будет показано, что эти методы можно применять, в частности, к квадратичной рабочей функции, а также к рабочим функциям других видов.

Метод Ньютона является фундаментальным с точки зрения математического описания, хотя зачастую его трудно применить на практике. Это - метод градиентного поиска, при котором на каждом шаге процесса поиска или на каждом цикле итерации, определение которого

будет дано ниже, изменяются все компоненты вектора весовых коэффициентов. При условии, что функция является квадратичной, эти изменения всегда имеют направление на минимум рабочей функции.

Метод наискорейшего спуска прост в применении и имеет большое значение для решения широкого круга прикладных задач. Это - метод градиентного поиска, при котором также на каждом шаге или цикле итерации изменяются все компоненты вектора весовых коэффициентов. Однако в этом случае все изменения осуществляются в направлении отрицательного градиента рабочей функции. Таким образом, они не обязательно имеют направление на минимум, поскольку, отрицательный градиент направлен к минимуму только тогда, когда его начало расположено на одной из главных осей.

15.2. Основные принципы методов градиентного поиска

Для введения основных понятий методов градиентного поиска, в том

числе понятия рекурсивного алгоритма и сходимости, рассмотрим

сначала простейший случай, когда имеется только один весовой

коэффициент. Для этого случая, имеющего ограниченное практическое

значение, все методы градиентного поиска сводятся к одному.

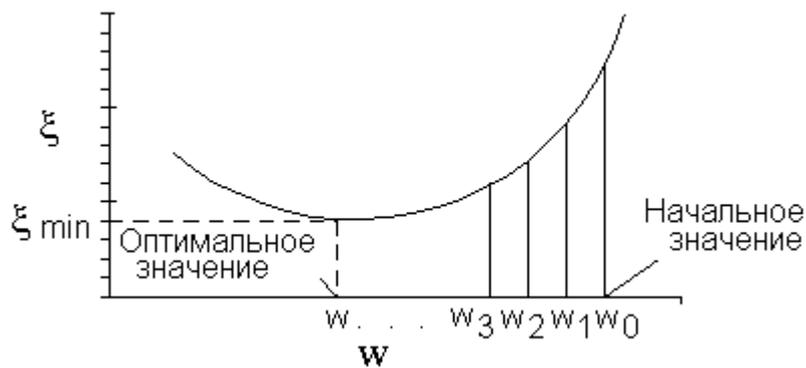


Рис. 15.1. Иллюстрация процесса градиентного поиска для рабочей функции одной переменной

График рабочей функции одного весового коэффициента (одной переменной), которая является параболой, показан на рис. 15.1. Эта функция определяется как

$$\xi = \xi_{\min} + \lambda (\omega - \omega^*)^2 \quad (15.1)$$

Отметим, что для случая одной переменной собственное значение λ равно r_{00} . Первая производная

$$\frac{d\xi}{d\omega} = 2\lambda (\omega - \omega^*) \quad (15.2)$$

Вторая производная

$$\frac{d^2\xi}{d\omega^2} = 2\lambda \quad (15.3)$$

является постоянной для всей кривой.

Задача состоит в том, чтобы найти такой весовой коэффициент ω^* , при котором минимизируется значение СКО. Полагая рабочую функцию неизвестной, начнем с произвольного значения ω_0 и измерим наклон кривой в этой точке. Далее выберем новое значение ω_1 , равное начальному значению ω_0 , плюс приращение, пропорциональное наклону с обратным знаком¹. Затем при измерении наклона кривой в точке ω_1 точно так же получается еще одно новое значение ω_2 . Этот процесс повторяется до тех пор, пока не будет достигнуто оптимальное значение ω^* .

Значение, полученное измерением наклона кривой рабочей функции в расположенных с дискретными интервалами точках $\omega_0, \omega_1, \omega_2, \dots$, называется градиентной оценкой. Метод проведения измерений и его точность

обсуждается в гл. 5. Для задачи, рассматриваемой в этой главе, предположим, что имеется точное значение градиента. Отметим, что использование отрицательного градиента необходимо для «движения вниз» по кривой.

Отметим, что здесь, в отличие от гл. 3, индекс используется для обозначения номера итерации, а не номера весового коэффициента.

15.3. Простой алгоритм градиентного поиска

Повторный, или итеративный, процесс градиентного поиска, описанный выше для случая с одним весовым коэффициентом, алгебраически можно представить в виде

$$\omega_{k+1} = \omega_k + \mu(-\nabla_k) \quad , \quad (15.4)$$

где k — номер шага или итерации. Таким образом, ω_k является текущим значением, в то время как ω_{k+1} — новым значением. Через ∇_k обозначен градиент при $\omega = \omega_k$. Параметр μ представляет собой константу, от которой зависит устойчивость и скорость сходимости; вопрос его выбора обсуждается ниже в этой главе.

Для случая с одним весовым коэффициентом из (15.2) получаем

$$\nabla_k = \left. \frac{d\xi}{d\omega} \right|_{\omega=\omega_k} = 2\lambda(\omega_k - \omega^*) \quad (15.5)$$

Динамическое, или мгновенное, состояние процесса итерации от начального значения ω_0 до оптимального решения ω^* можно проанализировать с помощью уравнения, получаемого при подстановке (15.5) в (15.4):

$$\omega_{k+1} = \omega_k - 2\mu\lambda(\omega_k - \omega^*) \quad . \quad (15.6)$$

Меняя местами члены уравнения, получаем

$$\omega_{k+1} = (-2\mu\lambda)\omega_k + 2\mu\lambda\omega^* \quad (15.7)$$

Это уравнение является линейным однородным разностным уравнением первого порядка с постоянными коэффициентами. Его можно решить методом индукции на основе нескольких первых итераций. Относительно начального значения ω_0 первые три итерации уравнения (15.7) дают

$$\omega_1 = (-2\mu\lambda)\omega_0 + 2\mu\lambda\omega^* \quad (15.8)$$

$$\omega_2 = (1 - 2\mu\lambda)^2 \omega_0 + 2\mu\lambda\omega^* (1 - 2\mu\lambda) + 1 \quad (15.9)$$

$$\omega_3 = (1 - 2\mu\lambda)^3 \omega_0 + 2\mu\lambda\omega^* (1 - 2\mu\lambda)^2 + (1 - 2\mu\lambda) + 1 \quad (15.10)$$

С учетом этих результатов можно сделать обобщение для k-й итерации:

$$\omega_k = (1 - 2\mu\lambda)^k \omega_0 + 2\mu\lambda\omega^* \sum_{n=0}^{k-1} (1 - 2\mu\lambda)^n \quad (15.11)$$

$$\omega_k = (1 - 2\mu\lambda)^k \omega_0 + 2\mu\lambda\omega^* \frac{1 - (1 - 2\mu\lambda)^k}{1 - (1 - 2\mu\lambda)} \quad (15.12)$$

$$\omega_k = \omega^* + (1 - 2\mu\lambda)^k (\omega_0 - \omega^*) \quad (15.13)$$

Этот результат дает в явном виде значение ω_k для любой точки в процессе поиска и тем самым является «решением» алгоритма градиентного поиска.

15.4. Устойчивость и скорость сходимости

Величина $r = 1 - 2\mu\lambda$ в (15.13) называется знаменателем геометрической прогрессии, так как является отношением соседних членов геометрической суммы в (15.11). Очевидно, для итеративного процесса с одним весовым коэффициентом величина r является определяющей. Равенство (15.13) будет «устойчивым» тогда и

$$|r| = |1 - 2\mu\lambda| < 1 \quad \text{только тогда, когда} \quad (15.14)$$

Это условие можно представить также в виде

$$1/\lambda > \mu > 0. \quad (15.15)$$

Если выполняется условие (15.14) или (15.15), т. е. если алгоритм (15.13) является устойчивым, то очевидно, что он является сходящимся к оптимальному решению:

$$\lim_{k \rightarrow \infty} \omega_k = \omega^* \quad (15.16)$$

Скорость сходимости также зависит от знаменателя геометрической прогрессии.

На рис. 15.2 изображены типичные зависимости, которые имеют место в процессе коррекции, при различных значениях знаменателя геометрической прогрессии r . Кривые не имеют физического смысла и получены простым

соединением ряда точек, представляющих собой дискретные значения ω_k . Отметим, что если абсолютное значение $r < 1$, то скорость сходимости растет с уменьшением r , достигая своего максимума при $r = 0$, когда оптимальное решение достигается за один шаг. Кроме того, при положительных значениях $r < 1$ нет колебаний мгновенных значений весового коэффициента, а при отрицательных — мгновенные значения весового коэффициента неоптимальны и сходятся к ω^* по правилу затухающего колебания. В первом случае говорят, что процесс является недорегулированным, во втором — с перерегулированием. При $r = 0$ процесс эквивалентен методу Ньютона (рассматриваемому ниже) и говорят, что он является критическим. Если $r \geq 1$, то в соответствии с (15.14) процесс является неустойчивым и расходящимся.

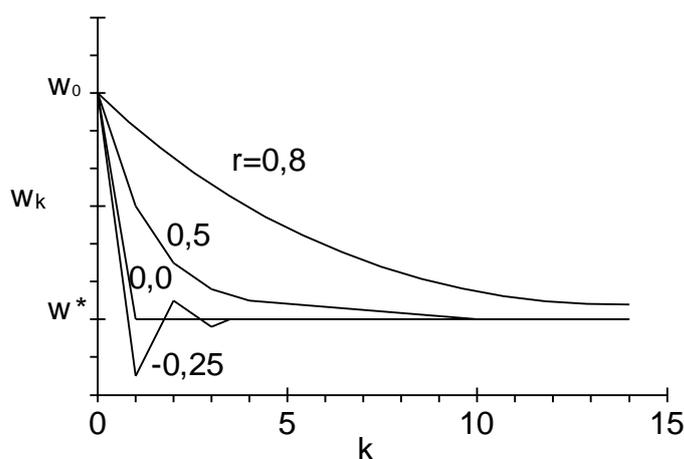


Рис. 15.2. Процесс коррекции весовых коэффициентов при различных значениях знаменателя геометрической прогрессии r .

При $r = 0$ (метод Ньютона) ω^* достигается за одну итерацию

Влияние выбора параметра μ на r и характер итеративного процесса в системе с одним весовым коэффициентом показано в табл. 15.1.

Таблица 15.1

Характер итеративного процесса	Параметр μ	Значение r
Устойчивый (сходящийся)	$0 < \mu < 1/\lambda$	$ r < 1$
Ндорегулированный	$0 < \mu < 1/2\lambda$,	$1 > r > 0$
Критический	$\mu = 1/2\lambda$	$r = 0$
С перерегулированием	$1/2\lambda < \mu < 1/\lambda$	$0 > r > -1$
Неустойчивый (расходящийся)	$\mu \geq 1/\lambda$ ИЛИ $\mu \leq 0$	$ r > 1$

15.5. Обучающая кривая

Зависимость СКО от изменений весового коэффициента в процессе коррекции можно вывести исходя из (15.1). Если допустить по определению, что ξ_k — значение СКО для фиксированного весового коэффициента ω_k , то с учетом (15.1)

$$\xi_k = \xi_{\min} + \lambda (\omega_k - \omega^*)^2 \quad (15.17)$$

Подставляя в эту формулу выражение (15.13), получаем

$$\xi_k = \xi_{\min} + \lambda (\omega_0 - \omega^*)^2 (-2\mu\lambda)^k \quad (15.18)$$

Очевидно, поскольку ω_k стремится к ω^* по закону геометрической прогрессии, СКО также стремится к ξ_{\min} по закону геометрической прогрессии. Следовательно, в (15.18) знаменатель геометрической прогрессии значений СКО

$$r_{\text{СКО}} = r^2 = (-2\mu\lambda)^2 \quad (15.19)$$

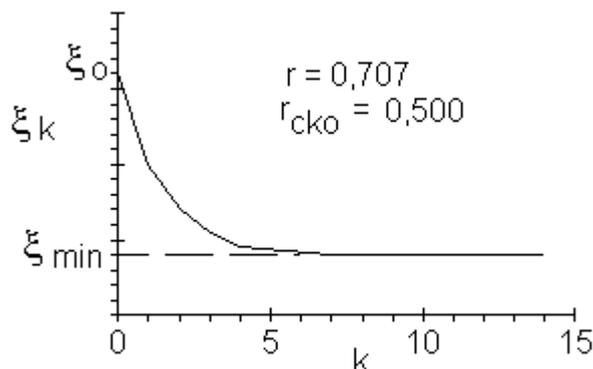


Рис. 15.3. Обучающая кривая—график зависимости СКО ξ_k от k

На рис. 15.3 показано приближение СКО от начального ξ_0 к оптимальному значению ξ_{\min} для системы с одним весовым коэффициентом. В приведенном примере $r_{\text{СКО}}=0,5$, что соответствует $r=0,707$. Как и прежде, кривая не имеет физического смысла в промежутках между целыми k . Она построена простым соединением точек, соответствующих мгновенным дискретным значениям ошибки. Кривую называют обучающей, и она показывает, как в итеративном процессе происходит уменьшение СКО.

15.6. Градиентный поиск методом Ньютона

Выше было показано, что процесс градиентного поиска для одной переменной является критическим при $\gamma=0$, т. е. когда

$$\gamma = 1 - 2\mu\lambda = 0, \mu = 1/2\lambda. \quad (15.20)$$

В этом случае процесс за один шаг сходится к квадратичным функциям СКО и алгоритм его реализации называется методом Ньютона, так как он связан с элементарными вычислениями при нахождении корней полинома. Рассмотрим теперь вопрос применения этого метода к функциям одного весового коэффициента (одной переменной) и затем распространим его на рабочую функцию многих переменных.

Метод Ньютона является прежде всего методом нахождения нулей функции, допустим функции $f(\omega)$, т. е. методом решения уравнения $f(\omega)=0$. Он состоит в том, что задается начальное значение ω_0 , а затем для вычисления следующей оценки ω_1 находится первая производная $f'(\omega)$. Как показано на рис. 15.4, ω_1 находится как пересечение касательной в точке $f(\omega_0)$ с осью ω . Таким образом, из рис. 15.4

$$f'(\omega_0) = f(\omega_0) / (\omega_0 - \omega_1)$$

или

$$\omega_1 = \omega_0 - f(\omega_0) / f'(\omega_0). \quad (15.21)$$

Следующую точку, ω_2 , вычисляют, используя в качестве начальной точки ω_1 , и т. д. В общем случае

$$\omega_{k+1} = \omega_k - f(\omega_k) / f'(\omega_k); k=0, 1, \dots \quad (15.22)$$

Очевидно, сходимость метода Ньютона зависит от выбора начального значения ω_0 и от вида функции $f(\omega)$, но известно, что для широкого класса функций он обладает быстрой сходимостью.

Равенство (15.22) называют *непрерывной формой* метода Ньютона, так как в явном виде использована непрерывная функция $f(\omega)$ и ее производная.

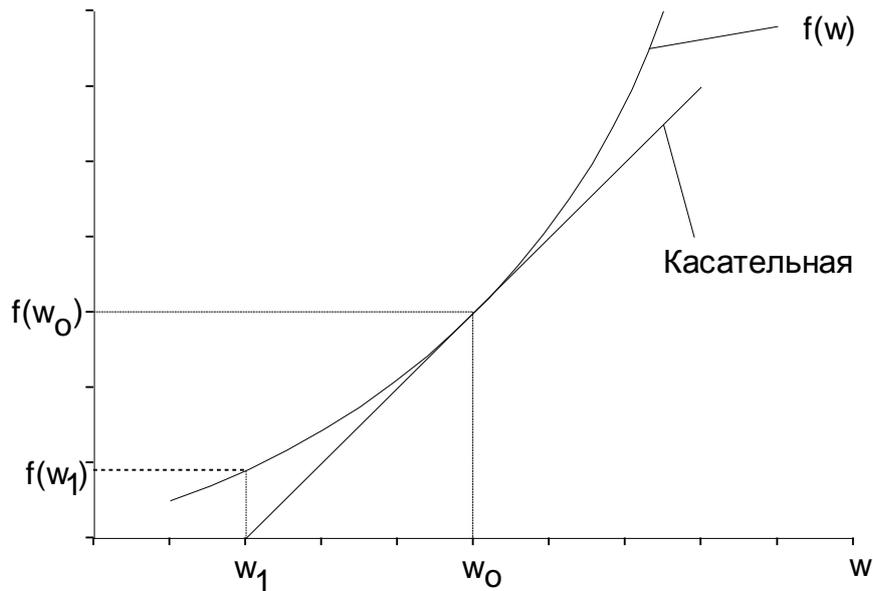


Рис. 15.4. Метод Ньютона при нахождении нуля функции $f(\omega_0)$ состоит в переходе от ω_0 к ω_1 по касательной к $f(\omega)$

Кроме того, существует и применяется *дискретная форма* этого метода, когда необходимо вычислить $f'(\omega)$. Полагая, что функция $f(\omega)$ известна, или ее можно точно вычислить, на основании формулы обратной разности можно определить

$$f'(\omega_k) \approx \frac{f(\omega_k) - f(\omega_{k-1})}{\omega_k - \omega_{k-1}} \quad (15.23)$$

Подставляя эту приближенную формулу в (15.22), получаем выражение для дискретной формы метода Ньютона:

$$\omega_{k+1} = \omega_k - \frac{f(\omega_k)(\omega_k - \omega_{k+1})}{f(\omega_k) - f(\omega_{k-1})}; k = 0, 1, \dots \quad (15.24)$$

Отметим, что в (15.24), а также в (15.22) на каждой итерации необходимо проверять, не обращается ли знаменатель в нуль. Пока будем пользоваться непрерывной формой метода Ньютона.

Как и выше, поиск минимума рабочей функции методом Ньютона необходимо начать с уравнения вида $f(\omega)=0$. В общем случае этим уравнением является уравнение $\xi=0$ или $\nabla=0$, так как требуется найти минимум $\xi(\omega)$. Таким образом, для рабочей функции одной переменной

$$f(\omega) = \xi'(\omega) \quad (15.25)$$

Поэтому, например, непрерывная форма (15.22) принимает вид

$$\omega_{k+1} = \omega_k - \frac{\xi'(\omega_k)}{\xi''(\omega_k)}; k = 0, 1, \dots$$

(15.26)

Отметим, что дискретная форма алгоритма в соответствии с (15.23) теперь представляет собой аппроксимацию второй производной функции ξ .

Если рабочая функция является квадратичной, как это было во всех предыдущих рассуждениях, то применение метода Ньютона приводит, как показано на рис. 15.2, к решению за один шаг. Подстановка (15.2) в (15.26) при $k=0$ дает

$$\omega_1 = \omega_0 - 2\lambda (\omega_0 - \omega^*) \cdot 2\lambda = \omega^*. \quad (15.27)$$

Таким образом, метод Ньютона прост для применения в случае одной переменной, когда рабочая функция является квадратичной и определена для всех значений ω .

Возможны два случая, когда метод Ньютона усложняется. Во-первых, бывает необходимо вычислить ξ' и ξ'' в (15.26), когда точно не известна функция $\xi(\omega)$. Во-вторых, рабочая функция может не быть квадратичной. До сих пор рассматривался только такой адаптивный линейный сумматор, для которого ξ —квадратичная функция, но ниже вводятся другие адаптивные структуры, имеющие неквадратичные рабочие функции. Пример такой системы приведен на рис. 15.5, где изображен график рабочей функции для конкретного рекурсивного адаптивного фильтра.

Отметим, что даже если рабочая функция является неквадратичной, то при начальном весовом коэффициенте $\omega=0$ метод Ньютона приводит почти к оптимальной точке $\omega^*=0,448$ всего лишь после четырех итераций. Однако при некоторых начальных значениях весового коэффициента для данного примера метод Ньютона не приводит к оптимуму.

15.7. Метод Ньютона для многомерного пространства

Выше показано, что если имеется один весовой коэффициент и рабочая функция является квадратичной, то методом Ньютона оптимальный весовой коэффициент ω^* находится за один шаг. Расширим понятие метода Ньютона на случай с многими весовыми коэффициентами, определив его как метод, который приводит к оптимальной квадратичной рабочей функции за один шаг.

Напомним, что оптимальный вектор весовых коэффициентов задается соотношением

$$\mathbf{W}^* = \mathbf{R}^{-1} \mathbf{P}, \quad (15.28)$$

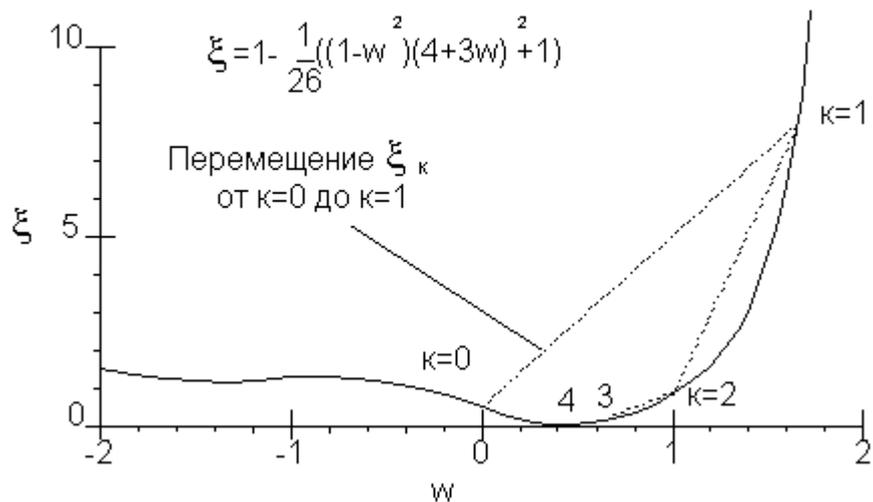


Рис. 15.5. Аппроксимация методом Ньютона для неквадратичной рабочей функции с начальным значением $\omega=0$ и вектор градиента

$$\nabla = 2\mathbf{R}\mathbf{W} - 2\mathbf{P}. \quad (15.29)$$

Можно умножить обе части равенства (15.29) слева на $1/2\mathbf{R}^{-1}$ и затем на основании этих двух равенств получить

$$\mathbf{W}^* = \mathbf{W} - 1/2\mathbf{R}^{-1}\nabla. \quad (15.30)$$

Запишем этот результат в виде адаптивного алгоритма¹

$$\mathbf{W}_{k+1} = \mathbf{W}_k - 1/2\mathbf{R}^{-1}\nabla_k. \quad (15.31)$$

Индекс k вектора градиента означает, что градиент находится на шаге k , когда вектор весовых коэффициентов равен \mathbf{W}_k .

Таким образом, равенство (15.31) описывает метод Ньютона для многих переменных. Если функция ошибки является квадратичной, то этот метод, так же, как и (15.30), приводит к оптимальному решению за один шаг. На рис. 15.6 проиллюстрирована квадратичная функция с двумя весовыми коэффициентами. В этом «идеальном» случае значения весовых коэффициентов переходят от любых начальных $\mathbf{W}_0 = (\omega_{00}, \omega_{10})$ к оптимальным $\mathbf{W}^* = (\omega_{0}^*, \omega_{1}^*)$ за один шаг.

Как следует из рис. 15.6 и равенства (15.31), в методе Ньютона шаги коррекции осуществляются не в направлении градиента. Для этого нужно, чтобы направление изменения весовых коэффициентов на рис. 15.6 было перпендикулярно каждой кривой.

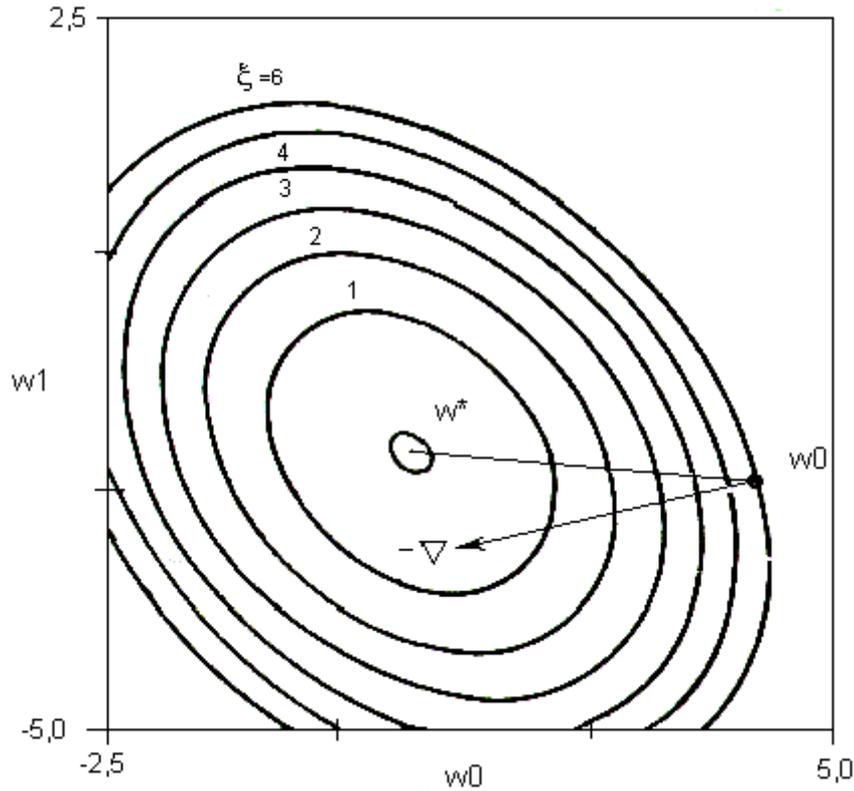


Рис. 15.6. Иллюстрация метода Ньютона для $\mu=1$ и двух весовых коэффициентов. Квадратичная рабочая функция.

А это возможно только тогда, когда \mathbf{W}_0 соответствует точке на одной из главных осей.

Заметим, что можно обобщить метод Ньютона, если для (15.31) снова ввести константу μ , ранее введенную в (15.4), и определяющую скорость сходимости. Если (15.31) представить в виде

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \mu \mathbf{R}^{-1} \nabla_k. \tag{15.32}$$

то при $\mu=1/2$ получаем формулу алгоритма, приводящего к оптимальному решению за один шаг. Во всех других случаях можно выбирать любое другое значение параметра μ , в пределах области устойчивости, как это следует из приведенного ниже соотношения (15.35)

$$0 < \mu < 1. \tag{15.33}$$

Для квадратичной рабочей функции можно вычислить (15.32), подставляя в него выражение для градиента (15.29) и затем (15.28):

$$\mathbf{W}_{k+1} = (1 - 2\mu) \mathbf{W}_k + 2\mu \mathbf{W}^*. \tag{15.34}$$

Теперь, имея равенство вида (15.7), можно методом индукции найти решение

аналогично тому, как из (15.7) получено (15.13). Для данного случая соответствующее решение

$$\mathbf{W}_k = \mathbf{W}^* + (1-2\mu)^k (\mathbf{W}_0 - \mathbf{W}^*). \quad (15.35)$$

Чтобы проверить правильность этого решения, заметим, что при $\mathbf{W}_1 = \mathbf{W}^*$ в результате имеем $\mu=1/2$, что соответствует алгоритму поиска решения за один шаг, а при выполнении условия (15.33) $\mathbf{W}_\infty = \mathbf{W}^*$.

15.8. Градиентный поиск методом наискорейшего спуска

Второй важный метод поиска, который обсуждается в этой главе, называют методом наискорейшего спуска, потому что здесь в отличие от метода Ньютона на каждом шаге весовые коэффициенты корректируются по направлению градиента. На рис. 15.7 приведен пример, в котором использована та же квадратичная рабочая функция, что и на рис. 15.6. Однако в отличие от примера на рис. 15.6, где сходимость к оптимальному решению достигается за один шаг, здесь используется малый шаг, чтобы показать траекторию наискорейшего спуска.

Сходимость за один шаг является достоинством при численном анализе, когда желательно уменьшить число итераций, необходимых для нахождения оптимума рабочей функции.

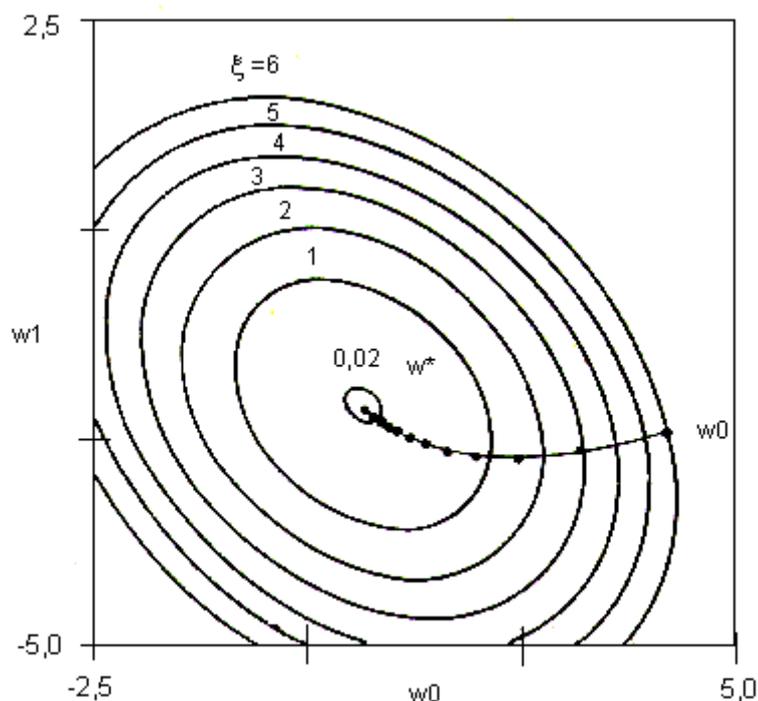


Рис. 15.7. Иллюстрация метода наискорейшего спуска для системы с двумя весовыми коэффициентами. Показана та же квадратичная рабочая функция, что на рис. 15.6, но для $\mu=0,3$

Однако для разработчика адаптивной системы такая сходимость вообще является слишком быстрой и в действительности нежелательной. При численном анализе можно полагать, что функция, для которой необходимо

осуществить поиск оптимума, задана, а во многих практических приложениях адаптивных систем рабочая функция неизвестна, и ее надо измерить и приближенно вычислить на основе случайных входных данных. При медленной адаптации имеет место процесс фильтрации, который снижает влияние шума, связанного с измерением градиента. Поэтому метод Ньютона не так полезен при разработке практических алгоритмов, как некоторые другие, из которых метод наискорейшего спуска оказался наиболее широко применимым.

Из определения ясно, что метод наискорейшего спуска выражается в виде следующего алгоритма, в котором параметр μ является константой, определяющей размер шага, с размерностью, обратной мощности сигнала:

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu(-\nabla_k). \quad (15.36)$$

Напомним, что (15.4) представляет собой одномерный вариант соотношения (15.36). Для определения характера процесса, возникающего при использовании этого алгоритма для поиска оптимума квадратичной рабочей функции, подставим в (15.36) соотношение для градиента (15.29) и затем (15.28). При этом

$$\mathbf{W}_{k+1} = \mathbf{W}_k - 2\mu\mathbf{R}\mathbf{V}_k = \mathbf{W}_k + 2\mu\mathbf{R}(\mathbf{W}^* - \mathbf{W}_k). \quad (15.37)$$

После преобразований

$$\mathbf{W}_{k+1} = (\mathbf{I} - 2\mu\mathbf{R})\mathbf{W}_k + 2\mu\mathbf{R}\mathbf{W}^*. \quad (15.38)$$

Решение этого уравнения усложняется тем, что различные компоненты вектора \mathbf{W}_k взаимосвязаны между собой. Матрица \mathbf{R} в общем случае не диагональная, а поскольку матрица \mathbf{W}_k в (15.38) содержит член $2\mu\mathbf{R}$, то она также является не диагональной. Для понимания отличия этого случая от предыдущего можно сравнить (15.38) с уравнением (15.34), соответствующим методу Ньютона.

Однако решить уравнение (15.38) можно, если привести его к системе координат, образованной главными осями. Для этого сначала, как и в (3.37), введем смещение $\mathbf{V} = \mathbf{W} - \mathbf{W}^*$. Тогда (15.38) принимает вид

$$\mathbf{V}_{k+1} = (\mathbf{I} - 2\mu\mathbf{R})\mathbf{V}_k. \quad (15.39)$$

Затем, используя соотношения (3.13) и (3.38), приведем уравнение к главным осям, т. е. учитывая $\mathbf{V} = \mathbf{Q}\mathbf{V}^*$, получаем

$$\mathbf{Q}\mathbf{V}^*_{k+1} = (\mathbf{I} - 2\mu\mathbf{R})\mathbf{Q}\mathbf{V}^*_k. \quad (15.40)$$

Умножив обе части этого уравнения слева на \mathbf{Q}^{-1} , найдем

$$\mathbf{V}^*_{k+1} = \mathbf{Q}^{-1}(\mathbf{I} - 2\mu\mathbf{R})\mathbf{Q}\mathbf{V}^*_k = (\mathbf{Q}^{-1}\mathbf{I}\mathbf{Q} - 2\mu\mathbf{Q}^{-1}\mathbf{R}\mathbf{Q})\mathbf{Q}\mathbf{V}^*_k = (\mathbf{I} - 2\mu\mathbf{\Lambda})\mathbf{V}^*_k. \quad (15.41)$$

Теперь, как и в (3.4), матрица собственных значений $\mathbf{\Lambda}$ является диагональной, поэтому (15.41) представляет собой множество $L+1$ уравнений вида (15.7). Отсюда ясно, что в системе координат, образованной главными осями, компоненты вектора \mathbf{W}_k не являются взаимосвязанными. Более того, методом индукции находим решение (15.41):

$$\mathbf{V}^*_k = (\mathbf{I} - 2\mu\mathbf{\Lambda})^k \mathbf{V}^*_0. \quad (15.42)$$

Из (15.42) следует, что алгоритм наискорейшего спуска является устойчивым и сходящимся, если

$$\lim_{k \rightarrow \infty} \left(-2\mu\lambda \right)^k = 0. \quad (15.43)$$

Поскольку произведение двух диагональных матриц равно матрице, составленной из произведений соответствующих элементов, (15.43) можно записать в виде

$$\begin{bmatrix} \lim_{k \rightarrow \infty} \left(-2\mu\lambda_0 \right)^k & & & \\ & \lim_{k \rightarrow \infty} \left(-2\mu\lambda_1 \right)^k & & \\ & & \ddots & \\ & & & \lim_{k \rightarrow \infty} \left(-2\mu\lambda_L \right)^k \end{bmatrix} = 0 \quad (15.44)$$

Такая запись показывает, что условие сходимости выполняется, если параметр μ выбран так, чтобы

$$0 < \mu < 1/\lambda_{\max}. \quad (15.45)$$

где λ_{\max} — максимальное собственное значение матрицы \mathbf{R} . Соотношение (15.45) является необходимым и достаточным условием сходимости алгоритма наискорейшего спуска для квадратичной рабочей функции. Если это условие выполняется, то

$$\lim_{k \rightarrow \infty} V'_k = 0 \quad (15.46)$$

Подставляя в (15.46) $\mathbf{V} = \mathbf{Q}^{-1}\mathbf{V} = \mathbf{Q}^{-1}(\mathbf{W} - \mathbf{W}^*)$ и осуществляя обратное преобразование к исходной системе координат, находим

$$\lim_{k \rightarrow \infty} \mathbf{W}_k = \mathbf{W}^* \quad (15.47)$$

Таким образом, в общем случае метод наискорейшего спуска является устойчивым и сходящимся тогда и только тогда, когда выполняется условие (15.45).

На рис. 15.8 представлен еще один пример применения алгоритма наискорейшего спуска для двумерной квадратичной рабочей функции. Здесь же показаны главные оси v_0 и v_1 . В соответствии с (15.42) сходимость определяется независимо по каждой главной оси. Скорость сходимости по каждой оси зависит от соответствующего знаменателя геометрической прогрессии. Как следует из (15.44), эти знаменатели

$$\begin{aligned} r_0 &= 1 - 2\mu\lambda_0, \\ r_1 &= 1 - 2\mu\lambda_1, \\ &\vdots \\ r_L &= 1 - 2\mu\lambda_L. \end{aligned} \quad (15.48)$$

Это означает, что по мере продвижения итеративного процесса последовательность проекций вектора \mathbf{W}_k на каждую главную ось является чисто геометрической и определяется знаменателем, который задается соответствующим собственным значением. Последовательность проекций вектора \mathbf{W}_k в исходной системе координат складывается из суммы геометрических прогрессий и поэтому является более сложной

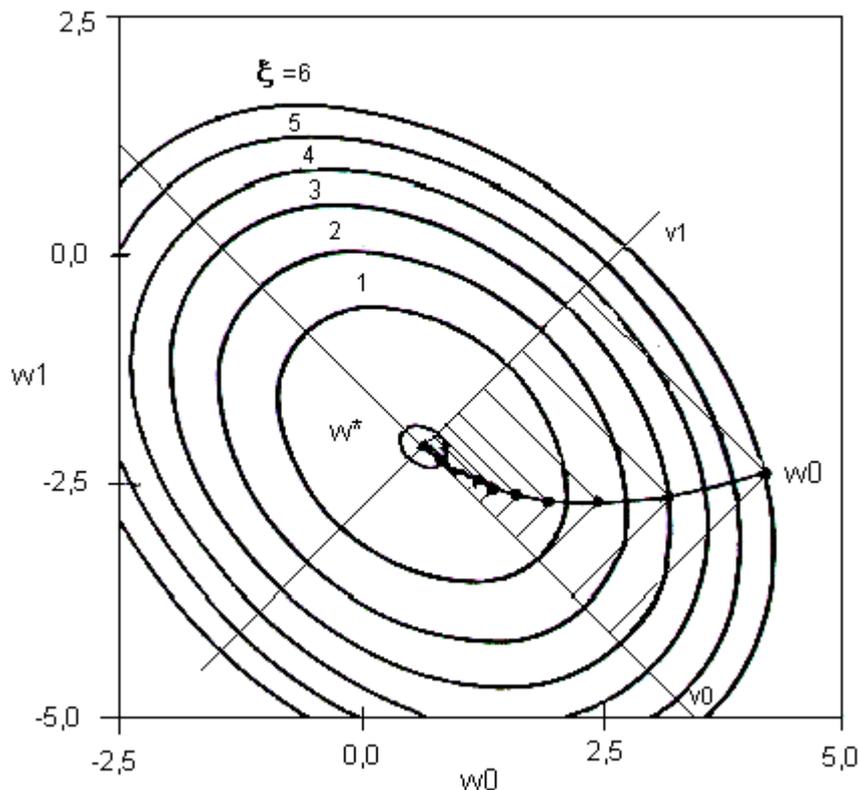


Рис. 15.8. Применение алгоритма наискорейшего спуска в системе с двумя весовыми коэффициентами. В соответствии с (15.48) по осям v_0 и v_1 знаменатели геометрических прогрессий являются постоянными

Итеративный процесс в исходной системе координат можно описать, выразив (15.42) через вектор \mathbf{W}_k . Умножим сначала обе части уравнения (15.42) слева на \mathbf{Q} :

$$\mathbf{Q}\mathbf{V}_k = \mathbf{Q}(\mathbf{I} - 2\mu\mathbf{A})^k \mathbf{V}_0. \quad (15.49)$$

Используя подстановку $\mathbf{V} = \mathbf{Q}^{-1}(\mathbf{W} - \mathbf{W}^*)$, получаем

$$\mathbf{W}_k = \mathbf{W}^* + \mathbf{Q}(\mathbf{I} - 2\mu\mathbf{A})^k \mathbf{Q}^{-1}(\mathbf{W} - \mathbf{W}^*). \quad (15.50)$$

Далее воспользуемся следующим соотношением:

$$(\mathbf{Q}\mathbf{A}\mathbf{Q}^{-1})^k = \mathbf{Q}\mathbf{A}\mathbf{Q}^{-1} \mathbf{Q}\mathbf{A}\mathbf{Q}^{-1} \dots \mathbf{Q}\mathbf{A}\mathbf{Q}^{-1} = \mathbf{Q}\mathbf{A}^k \mathbf{Q}^{-1}, \quad (15.51)$$

где \mathbf{A} — любая матрица, для которой существуют эти произведения. Подставляя в (15.50) $\mathbf{A} = \mathbf{I} - 2\mu\mathbf{A}$ и выражение (3.5), имеем

$$\mathbf{W}_k = \mathbf{W}^* + (\mathbf{Q}\mathbf{I}\mathbf{Q}^{-1} - 2\mu\mathbf{Q}\mathbf{A}\mathbf{Q}^{-1})^k (\mathbf{W}_0 - \mathbf{W}^*) =$$

$$= \mathbf{W}^* + (\mathbf{I} - 2\mu\mathbf{R})^k (\mathbf{W}_0 - \mathbf{W}^*), \quad (15.52)$$

что представляет собой решение разностного уравнения для алгоритма наискорейшего спуска в исходной системе координат.

В заключение отметим, что в алгоритме наискорейшего спуска коррекция весовых коэффициентов всегда направлена по градиенту рабочей функции. В общем виде алгоритм определяется выражением (15.36); другой записью алгоритма, или решением разностного уравнения, является соотношение (15.52). Как было показано, алгоритм становится устойчивым при выполнении условия (15.45).

15.9. Сравнение обучающих кривых

Полезно сравнить методы Ньютона и наискорейшего спуска на основе сравнения обучающих кривых. Для вывода формул, описывающих обе обучающие кривые, запишем выражения для квадратичной функции СКО

$$\xi = \xi_{\min} + \mathbf{V}^T \mathbf{R} \mathbf{V}. \quad (15.53)$$

Подставим сюда полученное решение разностного уравнения для метода Ньютона (15.35), преобразованное в соответствии с (3.37):

$$\xi_k = \xi_{\min} + (1-2\mu)^{2k} \mathbf{V}_0^T \mathbf{R} \mathbf{V}_0. \quad (15.54)$$

Это соотношение описывает простую геометрическую прогрессию, знаменатель которой

$$r_{\text{СКО}} = r^2 = (1-2\mu)^2. \quad (15.55)$$

На рис. 15.9 приведена характерная обучающая кривая для метода Ньютона, отображающая зависимость СКО от числа итераций. Эта кривая — график обычной экспоненциальной функции с единственной постоянной времени.

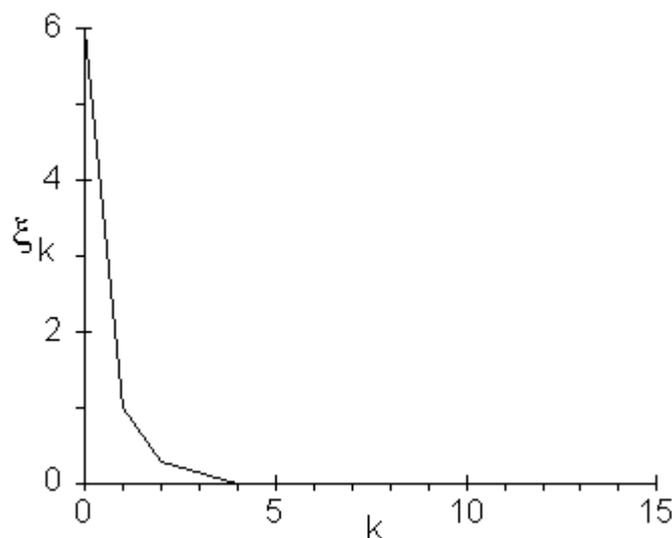


Рис. 15.9. Обучающая кривая для метода Ньютона, примененного в системе с многомерной квадратичной рабочей функцией, графики сечений которой приведены на рис. 15.6 при $\mu=0,3$

Аналогично выводится формула обучающей кривой для метода наискорейшего спуска. Для этого случая запишем выражение рабочей функции (15.53), полученное в (3.35) в системе координат, образованной главными осями:

$$\xi_k = \xi_{\min} + \mathbf{V}^T \Lambda \mathbf{V} \quad (15.56)$$

При подстановке в (15.56) решения разностного уравнения (15.42) выражение принимает вид

$$\begin{aligned} \xi_k &= \xi_{\min} + [(\mathbf{I} - 2\mu\Lambda)^k \mathbf{V}_0]^T \Lambda (\mathbf{I} - 2\mu\Lambda)^k \mathbf{V}_0 \\ &= \xi_{\min} + \mathbf{V}_0^T [(\mathbf{I} - 2\mu\Lambda)^k]^T \Lambda (\mathbf{I} - 2\mu\Lambda)^k \mathbf{V}_0 \end{aligned} \quad (15.57)$$

Здесь матрицы $(\mathbf{I} - 2\mu\Lambda)$ и Λ — диагональные, а произведение диагональных матриц коммутативно. Поэтому

$$\xi_k = \xi_{\min} + \mathbf{V}_0^T (\mathbf{I} - 2\mu\Lambda)^{2k} \Lambda \mathbf{V}_0; \quad (15.58)$$

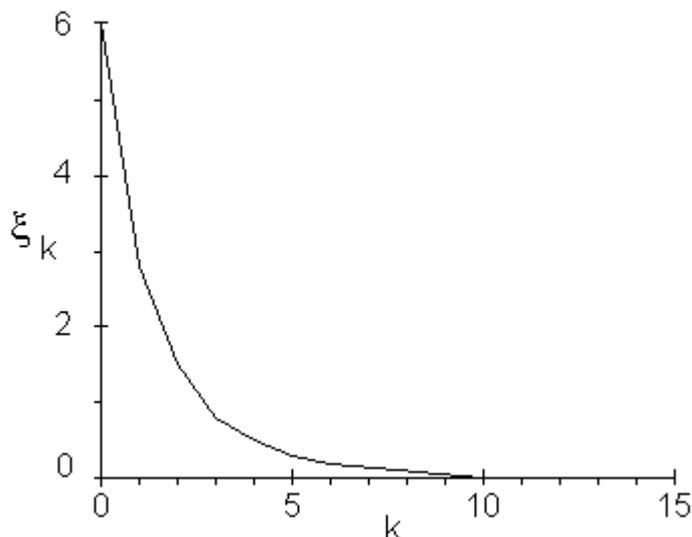
$$\xi_k = \xi_{\min} + \sum_{n=0}^L v_{on}' \lambda_n (-2\mu\lambda_n)^{2k}. \quad (15.59)$$

Итак, ясно, что обучающая кривая метода наискорейшего спуска складывается из суммы убывающих геометрических прогрессий со знаменателями вида

$$(r_{\text{СКО}})_n = r_n^2 = (1 - 2\mu\lambda_n)^2. \quad (15.60)$$

Обучающая кривая для метода наискорейшего спуска, примененного в системе с квадратичной рабочей функцией, приведена на рис. 15.10. Эта кривая является суммой экспонент, число которых равно числу весовых коэффициентов.

Сравнение рис. 15.9 и 15.10 показывает, что при одинаковых параметрах ρ , и при прочих равных условиях метод Ньютона обладает более быстрой



сходимостью, чем метод наискорейшего спуска. В общем случае это действительно так, потому что в методе Ньютона для нахождения прямого пути к минимуму функции ошибки ξ_{\min} используется информация, содержащаяся в элементах матрицы **R**.

Рис. 15.10. Обучающая кривая для метода наискорейшего спуска, примененного в системе с многомерной квадратичной рабочей функцией.

Графики сечений приведены на рис. 15.7 при $\mu=0.3$

Как показано на рис. 15.7 и 15.8, обычно алгоритм наискорейшего спуска реализует более длинный путь к ошибке ξ_{\min} .

С другой точки зрения можно считать, что методы Ньютона и наискорейшего спуска отражают процессы в системе с обратной связью, т. е. являются примерами реализации введенной выше функциональной обратной связи.

XVI. ВВЕДЕНИЕ В АДАПТИВНЫЕ АНТЕННЫЕ РЕШЕТКИ

И АДАПТИВНОЕ ФОРМИРОВАНИЕ ЛУЧЕЙ

В данной заключительной главе описан принцип адаптивной коррекции сигналов в пространственной области. Для достижения пространственной избирательности необходимо осуществлять прием сигналов на антенную решетку, имеющую два или более независимых пространственно разнесенных элементов. В зависимости от конкретного приложения элементами антенны могут быть диполи, предназначенные для приема электромагнитных сигналов, гидрофоны, помещенные в океане для приема акустических сигналов, сейсмометры, или геофоны, погруженные в землю для прослушивания сейсмических сигналов, или другие виды датчиков. Для приема волновых сигналов существует много типов таких элементов, которые можно устанавливать в пространственную решетку для обеспечения приема с изменяющейся в зависимости от направления чувствительностью.

Здесь основное внимание уделяется адаптивным приемным решеткам.

Однако интерес представляют и адаптивные передающие решетки.

В одном практическом приложении [21] адаптивной фильтрации в пространственной области применялась решетка датчиков на груди беременной женщины, формирующих эталонные сигналы группы адаптивных фильтров для коррекции, суммирования и выделения ЭКГ плода из входного сигнала, снимаемого с датчика на животе. В результате такого использования адаптивной решетки подавлялась помеха в виде ЭКГ матери.

Адаптивная антенна или адаптивная система формирования лучей состоит из множества пространственно разнесенных элементов, соединенных с одноканальным или многоканальным адаптивным устройством обработки сигналов. Вполне естественно рассмотреть адаптивные решетки сразу для всех видов сигналов (электромагнитных, акустических, сейсмических). Добавление пространственного измерения в обработку сигналов открывает новые возможности, приводящие к широкой области необычных приложений и алгоритмов. Все их рассмотреть невозможно, поскольку происходит быстрое развитие технологии. Поэтому внимание сконцентрировано на нескольких основных направлениях.

Адаптивные приемные решетки можно применять для уменьшения или исключения направленных помех с помощью адаптивного подавления или адаптивного формирования нулей, что приводит к улучшению отношения сигнал-шум. На основе одних видов адаптивных алгоритмов адаптивные приемные решетки могут «самонастраиваться», т. е. автоматически перестраиваться в сторону сигнала при не известном заранее направлении его прихода и отделять этот сигнал от направленных помех, пока направления их прихода отличны от направления прихода сигнала. Другие адаптивные алгоритмы позволяют определять направление прихода сигнала на фоне помех. С помощью одного вида алгоритмов можно выделять слабые сигналы на фоне сильных, когда они приходят под разными углами. Можно сделать так, чтобы адаптивные приемные решетки были чувствительными к ближайшим сигналам и нечувствительными к дальним, и наоборот. Кроме того, они могут быть чувствительными к редко передаваемым сигналам и нечувствительными к часто передаваемым или стационарным сигналам и наоборот. Число возможностей очень велико. Рассмотрение начнем с приемных решеток, в

которых можно формировать нули диаграммы направленности в направления помех.

16.1. Подавление боковых лепестков

Простейшим видом адаптивной антенны является устройство подавления боковых лепестков, впервые предложенное Хауэллзом в конце 1950-х годов и в дальнейшем разработанное Хауэллзом и Аппельбаумом. На рис. 16.1 приведена схема такого устройства, в котором используются два ненаправленных элемента (с одинаковой чувствительностью во всех направлениях) — один «входной», другой «эталонный». Положим, что одновременно действуют один сигнал и одна помеха. Как входной, так и эталонный направленные элементы принимают сигнал и помеху. Поскольку эти элементы пространственно разнесены, сигналы на их выходах не одинаковы, но представляют собой связанные временные функции.

Важным является случай, когда помеха намного мощней сигнала. Весовые коэффициенты адаптивного фильтра почти полностью определяются помехой. После адаптации выходной сигнал адаптивного фильтра содержит составляющую помехи, близкую к составляющей помехи, действующей на входном ненаправленном элементе. Следовательно, выходной сигнал системы почти не содержит помехи, но обязательно содержит сигнал. Выходной сигнал равен разности составляющей сигнала на входе и прошедшей через адаптивный фильтр составляющей сигнала на эталонном входе. Поскольку без подробных априорных сведений о характере и направлениях прихода помехи и сигнала нельзя предсказать, какой будет передаточная функция адаптивного фильтра после адаптации, в общем случае ожидаемое значение выходного сигнала является неопределенным. Однако при мощной помехе схема на рис. 16.1 приводит к повышению отношения сигнал-помеха на выходе.

В практическом случае сигнал каждого ненаправленного элемента обычно подается на отдельное приемное устройство для усиления, фильтрации и выделения сигнала. Сами приемные устройства и антенны вносят во входные сигналы независимые шумы, которые будем называть шумами приемника. Эти шумы оказывают значительное влияние на работу адаптивных антенн.

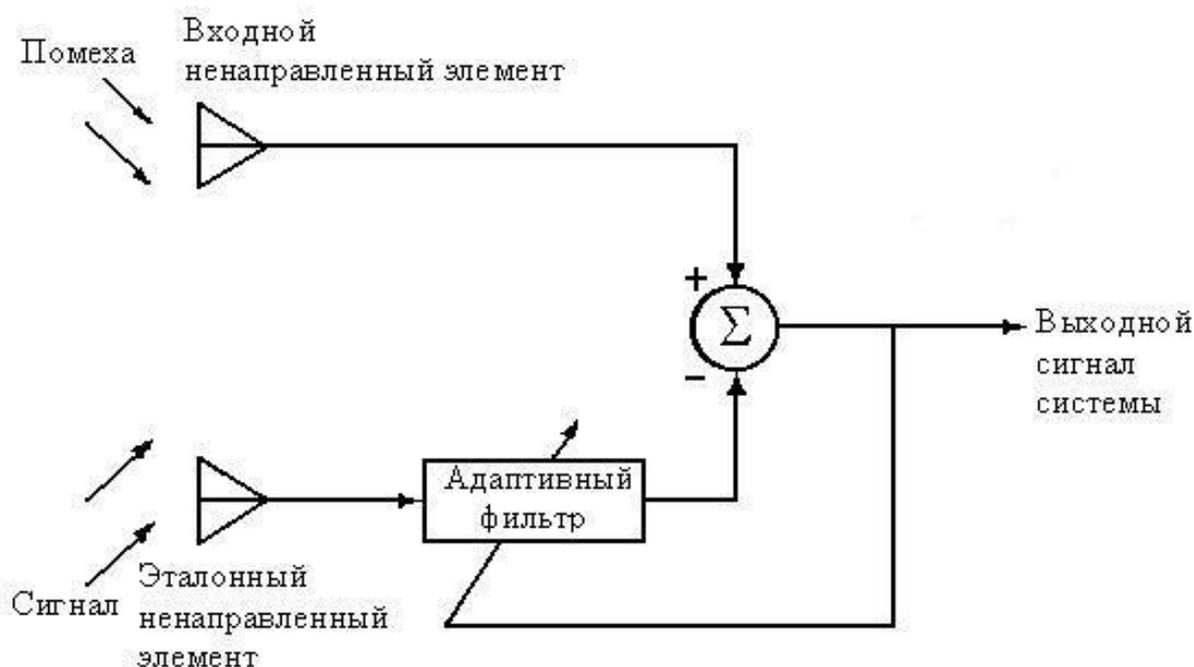


Рис. 16.1. Адаптивное подавление боковых лепестков в

двухэлементной решетке

Чтобы в системе на рис. 16.1 подавлялась направленная помеха, но не подавлялся полезный сигнал, помеха должна быть мощной, а сигнал слабым по сравнению с шумом приемника. Эти условия часто выполняются на практике.

На рис. 16.2 приведена схема адаптивного подавления боковых лепестков, в которой приемники на выходах обоих ненаправленных элементов вносят шумы. Ненаправленные элементы разнесены на расстояние l . При этом полагают, что принимается единственный сигнал, поступающий под углом θ_0 . Составляющая сигнала на входном ненаправленном элементе

$$\text{входной сигнал} = C \cos k\omega \quad (16.1)$$

где C – постоянная амплитуда; k – номер отсчета; ω измеряется в радианах. Фазовые фронты, перпендикулярные направлению прихода сигнала, проходят через двухэлементную решетку так, как показано на рис. 16.2.

Если c – скорость распространения, то данный фазовый фронт приходит на эталонный ненаправленный элемент раньше, чем на входной, на число временных шагов

$$\delta_0 = l \cdot \sin \theta_0 / cT = 2\pi l \cdot \sin \theta_0 / \lambda_0 \omega_0 \quad (16.2)$$

где T – временной шаг, а λ_0 – длина волны. Отсюда составляющая сигнала на эталонном ненаправленном элементе

$$\text{эталонный сигнал} = C \cos[(k + \delta_0)\omega_0] \quad (16.3)$$

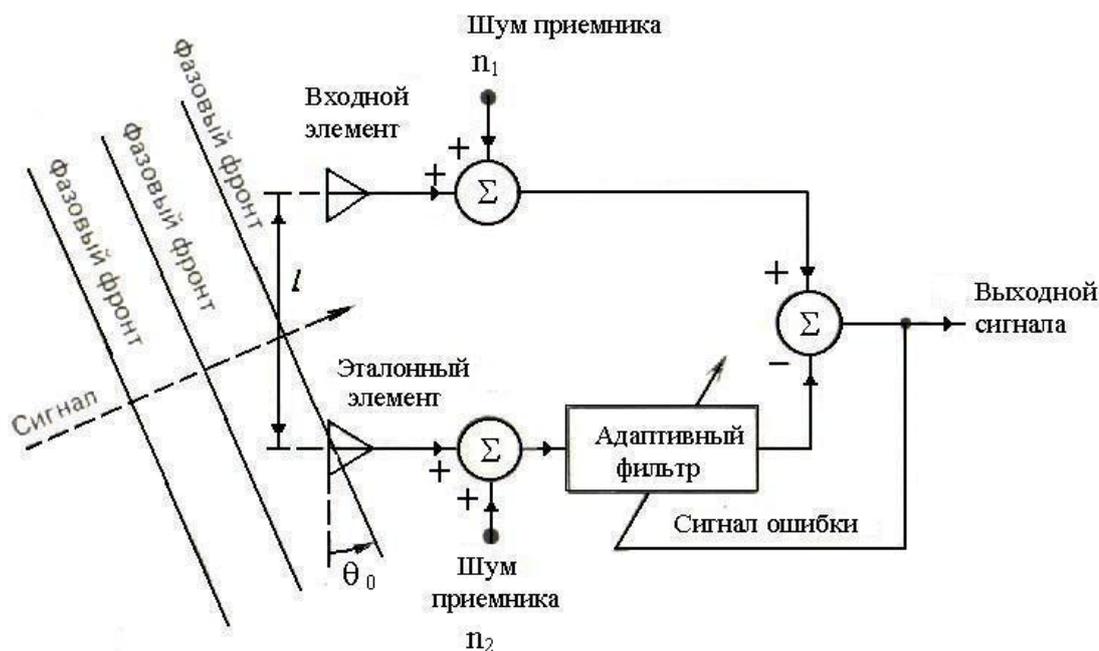


Рис. 16.2. Схема устройства подавления боковых лепестков при приеме одного сигнала при слабом относительно помех полезном входном сигнале

При отсутствии шума приемника весовые коэффициенты адаптивного фильтра настраиваются так, что сигнал подавляется полностью. Однако следует показать, что при наличии этого шума и мощном по сравнению с ним сигнале подавления сигнала не происходит, а адаптивное устройство подавления боковых лепестков функционирует удовлетворительно. Проведем анализ работы устройства при действии на входе одного сигнала и наличии шума приемника. Входной и эталонный сигналы определяются, соответственно, выражениями (16.1) и (16.3). Для анализа рассмотрим схему приведенного на рис. 16.3,а узкополосного устройства по давлению боковых лепестков.

Включение в тракты прохождения сигнала настроенных полосовых фильтров соответствует настроенной узкополосной приемной системе.

Будем считать, что каждый из этих фильтров имеет единичный коэффициент передачи на центральной частоте полосы пропускания приемника ω_0 . Для регулирования амплитуды и фазы сигнала на частоте ω_0 адаптивный фильтр должен иметь только синфазный и квадратурный весовые коэффициенты.

Предположим, что шумы приемника n_1 и n_2 в схеме на рис.16.3,а являются независимыми, в том числе относительно составляющих сигнала, а σ_r мощность каждого из них в полосе пропускания.

Тогда из (16.3) вектор автокорреляционной функции эталонного сигнала

$$\mathbf{p}_{xx} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma_r^2 + C^2/2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (16.4)$$

Отметим, что $\varphi_{xx}(1) = 0$, так как является корреляционной функцией сигнала (16.3) и его квадратурной составляющей, а среднее значение их произведения равно нулю.

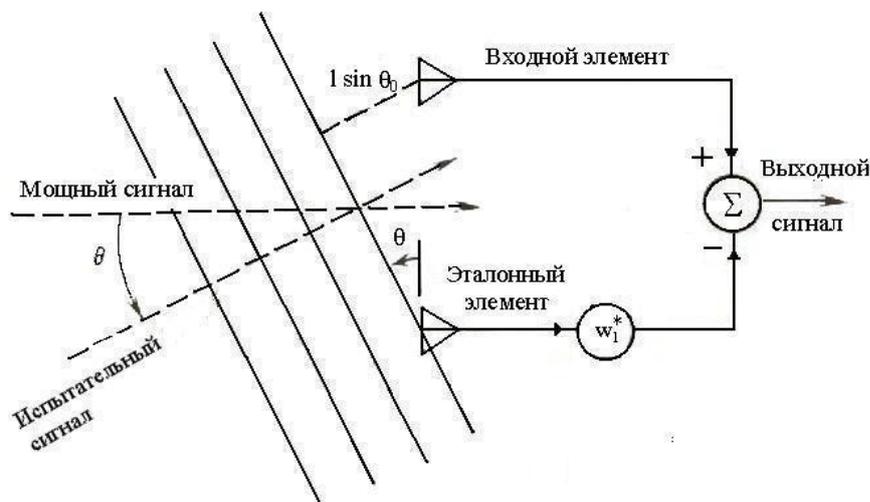
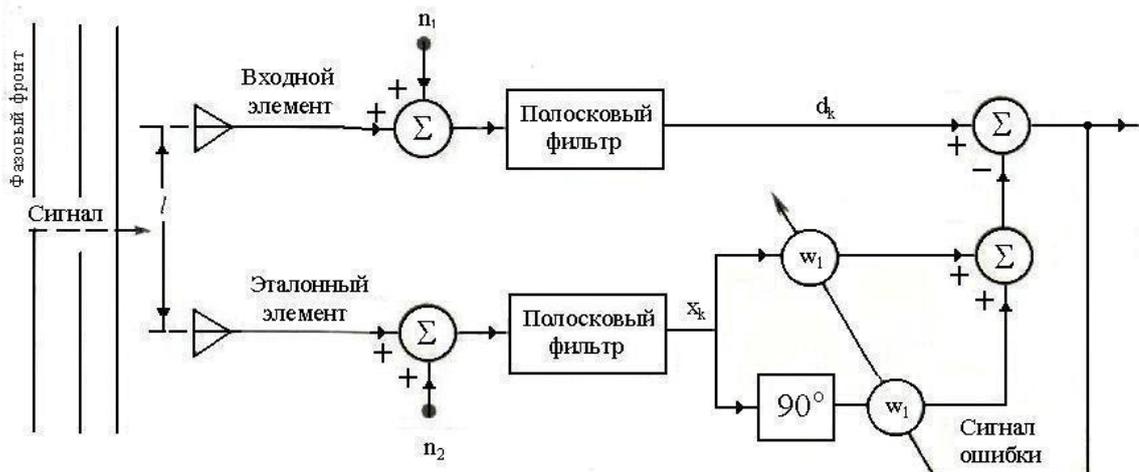


Рис. 16.3. Схема настроенного устройства подавления

боковых лепестков для направления приема сигнала под углом $\theta = \theta_0$ и ее упрощенный вариант для определения ДН устройства подавления боковых лепестков после его адаптации

Аналогично (16.4) находится вектор взаимокорреляционной функции для произвольного угла прихода θ_0

$$P = \begin{bmatrix} \varphi_{dx}(0) \\ \varphi_{dx}(1) \end{bmatrix} = \begin{bmatrix} E[C \cos k \omega_0 C \cos(k + \delta_0) \omega_0] \\ E[C \cos k \omega_0 C \sin(k + \delta_0) \omega_0] \end{bmatrix} = \quad (16.5)$$

$$= \frac{C^2}{2} \begin{bmatrix} \cos \delta_0 \omega_0 \\ \sin \delta_0 \omega_0 \end{bmatrix}$$

На основе этих выражений находим оптимальный вектор весовых коэффициентов для схемы на рис. 16.3,а:

$$\begin{bmatrix} \varpi_1^* \\ \varpi_2^* \end{bmatrix} = \frac{C^2}{2\sigma_r^2 + C^2} \begin{bmatrix} \cos \delta_0 \omega_0 \\ \sin \delta_0 \omega_0 \end{bmatrix}. \quad (16.6)$$

Пусть для рассматриваемой схемы $\theta_0=0$, тогда $\delta_0=0$, и оптимальный вектор весовых коэффициентов

$$\begin{bmatrix} \varpi_1^* \\ \varpi_2^* \end{bmatrix}_{\theta_0=0} = \begin{bmatrix} \frac{C^2}{2\sigma_r^2 + C^2} \\ 0 \end{bmatrix}. \quad (16.7)$$

В этом случае можно определить отношение сигнал-шум на входе как отношение мощностей входного или эталонного сигналов к мощности шума приемника в полосе пропускания:

$$ОСШ = \left(\frac{C^2}{2} \right) \sigma_r^2 \quad (16.8)$$

Соответственно из (16.7) получаем

$$\varpi_1^* = \frac{ОСШ}{1 + ОСШ}, \quad (16.9)$$

$$\varpi_2^* = 0.$$

Из (16.1) и (16.7) можно найти выражение для выходного сигнала:

$$\begin{aligned} \text{выходной сигнал} &= C \cos k \omega_0 - \varpi_1^* C \cos k \omega_0 = \\ &= \frac{2\sigma_r^2 C}{2\sigma_r^2 + C} \cos k \omega_0. \end{aligned} \quad (16.10)$$

Для мощного входного сигнала, т. е. при

$$C^2/2\sigma_r^2 \gg 1, \quad (16.11)$$

выходной сигнал можно записать в виде

$$\text{выходной сигнал} = \frac{2\sigma_r^2}{C} \cos k\omega_0 \quad (16.12)$$

Из приведенного выше анализа видно, что при слабом, по сравнению с шумом приемника в эталонном канале, входном сигнале отношение сигнал-шум мало, значение весового коэффициента ϖ_1^* почти равно нулю, выходной сигнал приближенно совпадает с входным, т. е. сигнал почти не подавляется. Однако при мощном, по сравнению с шумом приемника, входном сигнале имеем высокое отношение сигнал-шум, близкое к 1 значение весового коэффициента и небольшой уровень выходного сигнала, примерно равный уровню входного сигнала, деленному на отношение сигнал-шум. В этом случае амплитуда выходного сигнала, как следует из (16.12), обратно пропорциональна амплитуде входного сигнала, и происходит значительное подавление сигнала.

Если угол прихода сигнала не равен нулю, то имеет место соотношение (16.5). Весовой коэффициент ϖ_2 не стремится к нулю, а ϖ_1 принимает другое значение, однако амплитуда выходного сигнала определяется аналогично выражением (16.10).

Кроме того, представляет интерес определение чувствительности приемной системы на рис. 16.3,а для всех углов при $\theta_0=0$. Пусть после адаптации к сигналу, приходящему под нулевым углом, весовые коэффициенты имеют значения, соответствующие (16.9). На рис. 16,б приведена упрощенная схема, удобная для определения зависимости чувствительности от угла прихода. Подобную зависимость для решетки после адаптации обычно называют диаграммой направленности (ДН).

Положим, что контрольный сигнал на частоте ω_0 приходит под углом θ . Пусть, как и в (16.2), $\delta=(1 \sin \theta)/cT$, а соответствующая составляющая контрольного сигнала на входном элементе равна $\cos \omega_0$ при $C=1$, тогда в соответствии с рис. 16.6, б составляющая на эталонном элементе равна

$$\cos[(k+\delta)\omega_0]; \quad (16.13)$$

$$\begin{aligned} \text{выходной сигнал} &= \cos k\omega_0 - \varpi_1^* \cos(k+\delta)\omega_0 = \\ &= A \cos(k\omega_0 - \alpha), \end{aligned} \quad (16.14)$$

где

$$A = [1 + \varpi_1^{*2} - 2\varpi_1^* \cos(\delta\omega_0)]^{1/2} \quad (16.15)$$

и

$$\alpha = \text{tg}^{-1} \frac{\varpi_1^* \sin(\delta\omega_0)}{1 - \varpi_1^* \cos(\delta\omega_0)}. \quad (16.16)$$

Пусть по определению коэффициент передачи решетки по амплитуде есть отношение амплитуд выходного и входного сигналов. Поскольку амплитуда контрольного входного сигнала $\cos k\omega_0$ равна единице,

$$\begin{aligned} \text{коэффициент передачи решетки по амплитуде} = \\ = \frac{\text{амплитуда выходного сигнала}}{\text{амплитуда входного сигнала}} = A \text{ в выражении (16.14)}. \end{aligned} \quad (16.17)$$

Пусть, кроме того, по определению коэффициент передачи решетки по мощности равен квадрату упомянутого выше коэффициента, т. е.

$$\begin{aligned} \text{коэффициент передачи решетки по мощности} \stackrel{\Delta}{=} \\ \stackrel{\Delta}{=} \left(\frac{\text{амплитуда выходного сигнала}}{\text{амплитуда входного сигнала}} \right)^2 = A^2 = 1 + \varpi_1^{*2} - 2\varpi_1^* \cos(\delta\omega_0). \end{aligned} \quad (16.18)$$

Значение этого коэффициента в децибелах

$$10 \cdot \lg A^2 \stackrel{\Delta}{=} 20 \cdot \lg A. \quad (16.19)$$

Пусть теперь по определению фаза выходного сигнала решетки есть разность фаз выходного и входного сигналов, тогда из (16.14) имеем

$$\text{фаза выходного сигнала решетки} \stackrel{\Delta}{=} \alpha \text{ в выражении (13.16)} \quad (16.20)$$

Для рассмотрения (16.18) и (16.20) напомним, что ϖ_1^* - функция отношения сигнал-шум (16.9) и первоначальный входной сигнал, относительно которого проведена адаптация решетки, представляется выражением (16.1) и имеет направление прихода с $\theta=0$. Коэффициент передачи решетки принимает минимальное значение при $\theta=0$, что равносильно пространственной режекции в направлении прихода первоначального сигнала. Фактически адаптивный процесс стремится к подавлению входного сигнала. При этом чем мощнее сигнал, тем выше уровень режекции. По мере роста отношения сигнал-шум (которое больше единицы) происходит быстрое увеличение уровня адаптивной режекции, и, как это следует из (16.12), коэффициент передачи становится обратно пропорциональным уровню сигнала.

На рис. 16.4 построена зависимость коэффициента передачи решетки по мощности от угла θ в полярных координатах, в предположении, что угол прихода первоначального сигнала $\theta_0=0$. Отношение сигнал-шум на входе равно

6 дБ для внешней кривой, 10 дБ для средней и 100 дБ для внутренней. Основной сигнал приходит под углом $\theta=0^\circ$.

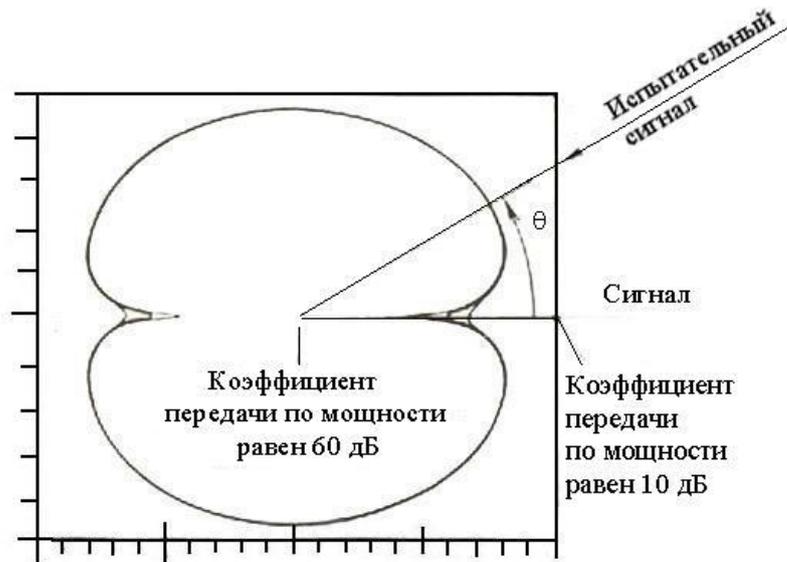


Рис. 16.4. Диаграмма направленности двухэлементного устройства подавления боковых лепестков при $l\omega_0 T = c\pi$

Приведенные кривые построены в соответствии с (16.9) и (16.18) для различных значений отношения сигнал-шум. Для этих кривых апертура решетки $l = c\pi T / \omega_0 = \lambda_0 / 2$. Отметим, что это значение l в некотором смысле является оптимальным, так как приводит к двум различным областям режекции и максимальному значению в (16.18).

Полагая $\theta = 0$, из (16.18); и (16.9) находим коэффициент передачи решетки по мощности для направления режекции:

$$\begin{aligned}
 (\text{Коэффициент передачи решетки по мощности})_{\theta=0} &= (1 - \varpi_1^*)^2 = \\
 &= \frac{1}{(1 + \text{ОСШ})^2}
 \end{aligned}
 \tag{16.21}$$

На рис. 16.5 приведена зависимость минимального коэффициента передачи решетки по мощности от отношения сигнал-шум. Фактически, как видно из рис. 16.5, адаптивное устройство подавления боковых лепестков производит оценку мощности. Мощные сигналы подавляются, в то время как слабые сигналы, приходящие по другим направлениям, пропускаются по существу с единичным коэффициентом передачи. Снова отметим, что «мощный» и «слабый» сигналы определяются относительно уровня шума приемника.

Рассматриваемое устройство нормально работает тогда, когда полезные входные сигналы являются слабыми, а помехи—мощными. Однако иногда полезный сигнал может быть достаточно мощным по сравнению с шумом приемника и при этом подавляться. Для того чтобы предотвратить такое подавление, для увеличения шума приемника можно специально прибавить к сигналу эталонного элемента независимый шум. Однако такой метод имеет недостаток, который заключается в том, что увеличивается шум системы в целом и в конечном итоге увеличивается шум в выходном сигнале. Более эффективным способом адаптации для этого случая является применение упрощенного алгоритма наименьших квадратов.

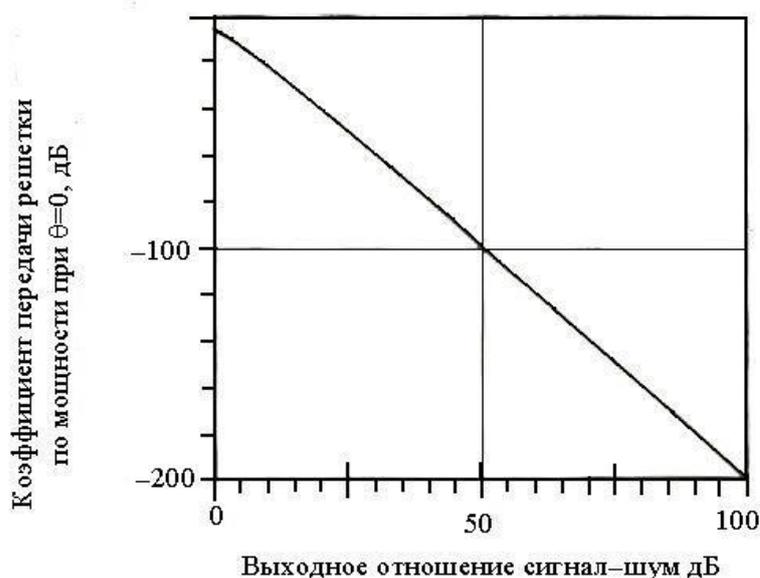


Рис. 16.5. Зависимость коэффициента передачи по мощности двухэлементного устройства подавления боковых лепестков по направлению режекции от отношения сигнал-шум на входе

Вывод этого алгоритма начнем с самого алгоритма наименьших квадратов. Из (6.3) для данного алгоритма можно записать

$$\begin{aligned} W_{k+1} &= W_k + 2\mu \varepsilon_k X_k = W_k + 2\mu X_k (d_k - X_k^T W_k) = \\ &= (I - 2\mu X_k X_k^T) W_k + 2\mu X_k d_k \end{aligned} \quad (16.22)$$

Полагая, что X_k и W_k некоррелированы, можно найти математическое ожидание от (16.22):

$$E[W_{k+1}] = (I - 2\mu R) E[W_k] + 2\mu P. \quad (16.23)$$

Математическое ожидание вектора весовых коэффициентов после адаптации для μ , соответствующего области устойчивости,

$$\lim_{k \rightarrow \infty} E[W_k] = W^* = R^{-1} P. \quad (16.24)$$

На рис. 16.6,а приведена схема адаптивного фильтра, работающего по алгоритму наименьших квадратов. К входному сигналу x_k прибавляется белый шум мощностью σ^2 . Для такого случая (16.23) принимает вид

$$E[W_{k+1}] = (I - 2\mu(R + \sigma^2 I))E[W_k] + 2\mu P, \quad (16.25)$$

а математическое ожидание вектора

$$\lim_{k \rightarrow \infty} E[W_k] = (R + \sigma^2 I)^{-1} P. \quad (16.26)$$

Упрощенный алгоритм наименьших квадратов зададим следующим образом:

$$W_{k+1} = \gamma W_k + 2\mu \varepsilon_k X_k, \quad (16.27)$$

где параметр γ — положительная константа, принимающая значения в интервале $1 > \gamma > 0$. (16.28)

Если $\mu=0$ в обычном алгоритме наименьших квадратов (16.22), то вектор весовых коэффициентов остается постоянным в течение неограниченного времени. Однако в упрощенном алгоритме даже если γ чуть меньше единицы, при $\mu=0$ вектор весовых коэффициентов постепенно, по закону геометрической прогрессии, стремится к нулю.

Влияние параметра γ на оптимальное решение можно определить следующим образом. Аналогичное (16.22) уравнение для упрощенного алгоритма наименьших квадратов имеет вид

$$\begin{aligned} W_{k+1} &= \gamma W_k + 2\mu \varepsilon_k X_k = \gamma W_k (d_k - X_k^T W_k) = \\ &= (\gamma I - 2\gamma X_k X_k^T) W_k + 2\mu X_k d_k \end{aligned} \quad (16.29)$$

Здесь также X_k и W_k — некоррелированы. Математическое ожидание для выражения (16.29)

$$\begin{aligned} E[W_{k+1}] &= (\gamma I - 2\mu R) E[W_k] + 2\mu P = \\ &= (I - 2\mu(R + \frac{1-\gamma}{2\mu} I)) E[W_k] + 2\mu P. \end{aligned} \quad (16.30)$$

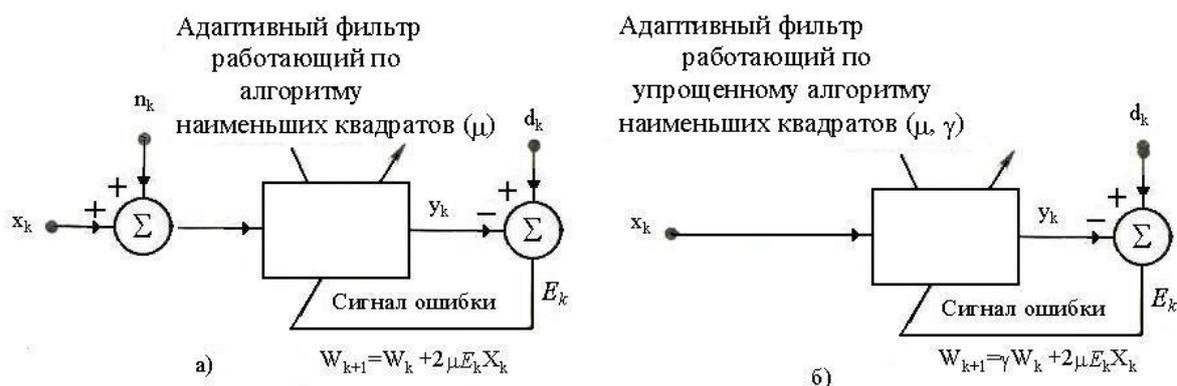


Рис. 16.6. Схемы эквивалентных адаптивных фильтров, работающих по алгоритму наименьших квадратов: а- фильтр с шумом на входе; б— эквивалентный фильтр, работающий по упрощенному алгоритму наименьших квадратов

Из сравнения (16.30) и (16.25) видно, что для упрощенного алгоритма можно получить такое же оптимальное решение, что и для схемы с шумом на рис. 16.6,а. Для определения влияния заданной мощности дополнительного шума на входе σ параметр γ выбирается следующим образом:

$$(1-\gamma)/2\mu = \sigma^2$$

или

$$\gamma = 1 - 2\mu\sigma^2. \quad (16.31)$$

Из (16.26) и (16.31) для упрощенного алгоритма наименьших квадратов находим оптимальное решение

$$\lim_{k \rightarrow \infty} E[W_k] = \left(R + \frac{1-\gamma}{2\mu} I \right)^{-1} P. \quad (16.32)$$

Этот алгоритм можно использовать в устройстве подавления боковых лепестков для установления эквивалентного уровня мощности шума на входе, хотя это и приведет к увеличению шума на выходе. Мощность входного сигнала может быть большой или малой по сравнению с этой эквивалентной мощностью шума, и их отношение определяется по (16.21), а уровень приема или степени подавления - по рис. 16.4 и 16.5. Эквивалентная мощность шума на входе

$$\sigma_r^2 + \sigma^2 = \sigma_r^2 + (1-\gamma)/2\mu, \quad (16.33)$$

где σ_r^2 — действительная мощность шума приемника.

При высоком уровне шума приемника, приводящем к значительному смещению оптимальных значений весовых коэффициентов, можно принять

$\gamma > 1$, при этом эквивалентная мощность будет меньше σ_r^2 . Однако при $\gamma > 1$ и $\mu = 0$ адаптивный алгоритм становится неустойчивым. Процесс адаптации следует по возможности сохранять устойчивым. Рассматриваемый алгоритм можно применять при $\gamma > 1$. Если этот алгоритм является устойчивым и используется в схеме с шумом приемника то математическое ожидание оптимального решения по аналогии с (16.32)

$$\lim_{k \rightarrow \infty} E[W_k] = \left[R + \left(\frac{1-\gamma}{2\mu} + \sigma_r^2 \right) I \right]^{-1} P, \quad (16.34)$$

Таким образом, рассматриваемый алгоритм приводит к такому же оптимальному вектору весовых коэффициентов, как и в случае шума на входе, но не требует добавления шума на входе, т. е. не приводит к дополнительному шуму на выходе или в значениях весовых коэффициентов.

На рис. 16.7 приведены некоторые ДН устройств подавления боковых лепестков, использующих этот алгоритм. Они рассчитаны так же, как на рис. 16.4, за исключением того, что отношение сигнал-шум в (16.8) в этом случае равно отношению мощности сигнала к эквивалентной мощности шума на входе $\sigma_r^2 + \sigma^2$. Кривые на рис. 16.7 построены в предположении, что мощность шума приемника равна единице, а мощность входного сигнала — десяти, т. е. отношение сигнал-шум составляет 10 дБ. Для регулирования порога приема или режекции сигнала изменяется значение γ . Полученный здесь результат соответствует почти ненаправленной ДН, что означает отсутствие режекции входного сигнала.

Рассмотрим теперь практический вариант решетки, когда сигнал и помехи действуют одновременно. Положим, что имеется одна узкополосная помеха, не коррелированная с узкополосным сигналом. Тогда можно дополнить матрицы R и векторы P для обоих сигналов, которые определяются выражениями (16.4) и (16.5). Пусть в направлении $\theta = 0$ мощность сигнала помехи σ_j^2 , а мощность более слабого полезного сигнала σ_s^2 угол прихода θ_0 . Тогда в соответствии с (16.4) и (16.5) имеем

$$(\sigma_r^2 + \sigma_s^2 + \sigma_j^2) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varpi_1^* \\ \varpi_2^* \end{bmatrix} = \begin{bmatrix} \sigma_j^2 + \sigma_s^2 \cos(\delta_0 \omega_0) \\ \sigma_s^2 \sin(\delta_0 \omega_0) \end{bmatrix}, \quad (16.35)$$

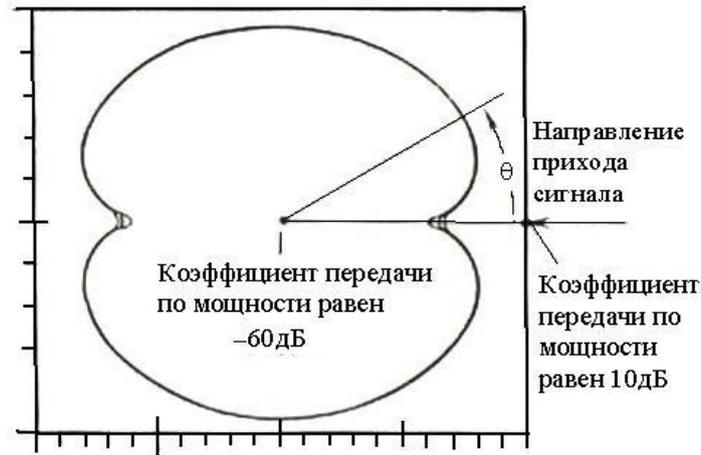


Рис. 16.7. Диаграммы направленности (зависимости коэффициента передачи по мощности в децибелах от угла θ) двухэлементного устройства подавления боковых лепестков после адаптации по упрощенному алгоритму наименьших квадратов при $l = \lambda_0/2$.

Сигнал приходит под углом $\theta=0^\circ$. Отношение сигнал-шум на входе равно 10 дБ. Для внутренней, средней и внешней кривых значения σ/σ_r^2 составляют 0,5; 1 и 1,5 где $\delta_0=(l \sin \theta_0)/cT$ (16.2). Оптимальные весовые коэффициенты

$$\begin{aligned} \varpi_1^* &= \frac{\sigma_j^2 + \sigma_s^2 \cos \delta_0 \omega_0}{\sigma_r^2 + \sigma_s^2 + \sigma_j^2}; \\ \varpi_2^* &= \frac{\sigma_s^2 \sin \delta_0 \omega_0}{\sigma_r^2 + \sigma_s^2 + \sigma_j^2}. \end{aligned} \quad (16.36)$$

Если теперь считать, что весовые коэффициенты оптимальны и не меняются, а входной контрольный сигнал $\cos k\omega_0$ с единичной амплитудой приходит под некоторым другим углом θ , тогда аналогично (16.14)

$$\begin{aligned} \text{выходной контрольный сигнал} &= \cos k\omega_0 - \varpi_1^* \cos[(k + \delta)\omega_0] - \\ &- \varpi_2^* \sin[(k + \delta)\omega_0] = A \cos(k\omega_0 - \alpha) \end{aligned} \quad (16.37)$$

где $\delta=(l \sin \theta)/cT$ (16.2). Таким образом, аналогично (16.18)

$$\begin{aligned} \text{коэффициент передачи решетки по мощности} &= A^2 = 1 + \varpi_1^{*2} + \varpi_2^{*2} - \\ &- 2(\varpi_1^* \cos \delta\omega_0 + \varpi_2^* \sin \delta\omega_0). \end{aligned} \quad (16.38)$$

На рис. 16.8 показана ДН для рассматриваемого случая. Здесь помеха в 10 раз больше по мощности сигнала, приходящего под углом $\theta_0=30^\circ$, и сказывается влияние сигнала, которое заключается в небольшом смещении обоих направлений режекции в сторону направления прихода этого сигнала.

Еще один практический случай относится к действию множества помех. Для обработки нескольких одновременно действующих помех нужно иметь не менее двух эталонных ненаправленных элементов. На рис. 16.9 приведена система с двумя пространственно разнесенными эталонными ненаправленными элементами. Отметим, что сигнал ошибки формируется здесь вычитанием из выходного сигнала обеих помех, поступающих с эталонных каналов.

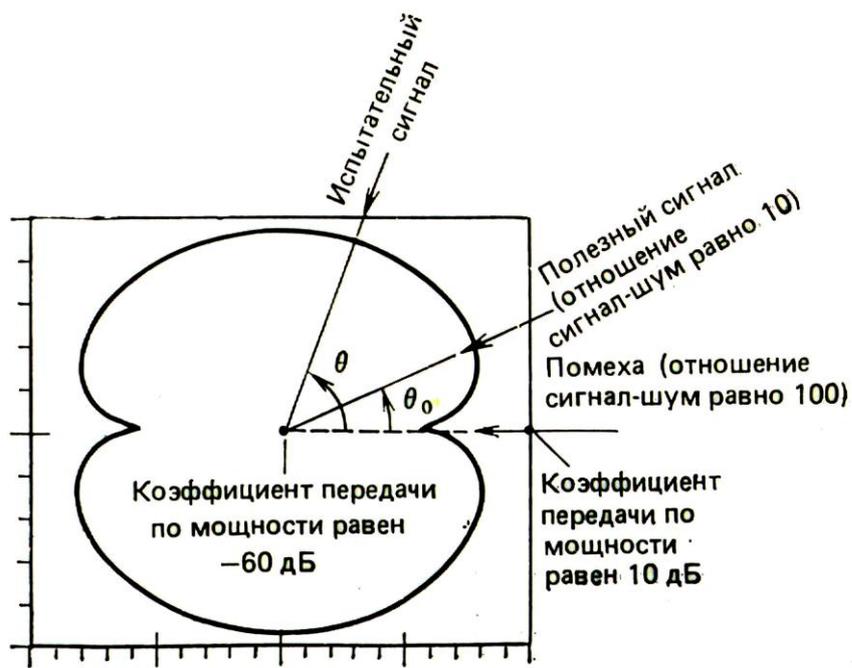


Рис. 16.8. Диаграмма направленности для сигнала с помехой при $l = \lambda_0/2$.

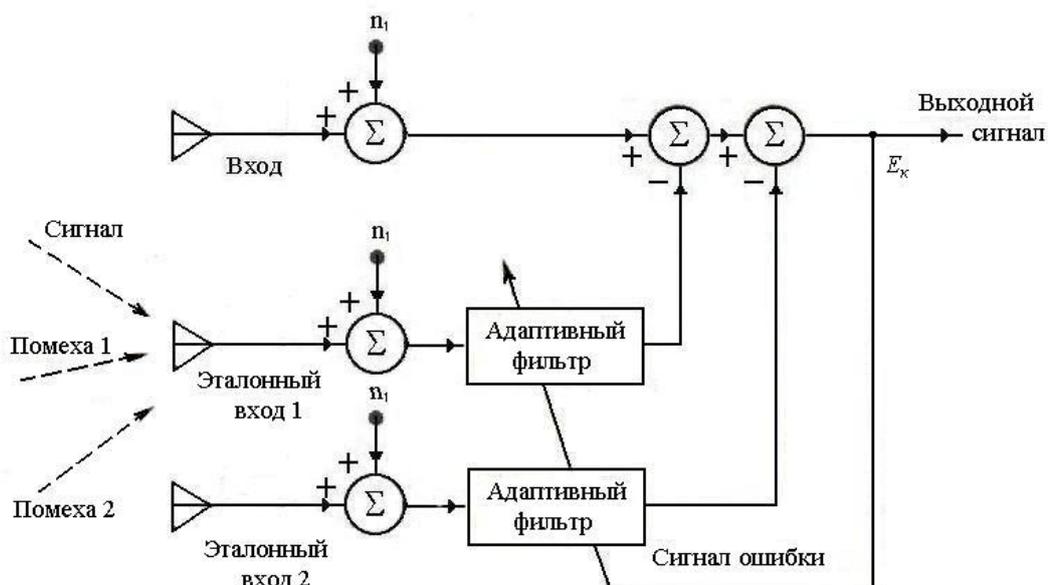


Рис. 16.9. Устройство подавления помех с двумя эталонными ненаправленными элементами

Поскольку эталонные ненаправленные элементы пространственно разнесены, система на рис. 16.9 может формировать в своей ДН два отдельных направления режекции. Пусть, как и в (16.2), $\delta_i = (l_i \sin \theta) / cT$ задержка, вносимая в сигнал с углом прихода θ i -м направленным элементом решетки, где l_i — расстояние между входным и эталонным элементами, а ϖ_{i1} и ϖ_{i2} ($i=1, 2$) - весовые коэффициенты фильтров. Тогда (16.27) для этого случая принимает вид

$$\begin{aligned} \text{выходной сигнал} = \cos k\omega_0 - \varpi_{11}^* \cos[(k + \delta_1)\omega_0] - \varpi_{12}^* \sin[(k + \delta_1)\omega_0] - \\ - \varpi_{21}^* \cos[(k + \delta_2)\omega_0] - \varpi_{22}^* \sin[(k + \delta_2)\omega_0]. \end{aligned} \quad (16.39)$$

На рис. 16.10 приведена ДН, построенная в соответствии с (16.39), для следующих данных: помеха 1: $\theta=60^\circ, l_1=\lambda_0/4$; помеха 2: $\theta=30^\circ, l_2=\lambda_0/8$; $\varpi_{11}^* = -0,48$; $\varpi_{12}^* = -0,87$; $\varpi_{21}^* = 1,7$; $\varpi_{22}^* = 1$.

Из рис. 16.10 видно, что в диаграмме сформированы два различных направления режекции под углами прихода помех 30 и 60° и два симметричных направления режекции под углами 150 и 120° соответственно. Отметим также, что из-за наличия двух помех диаграмма искажена относительно круговой сильней, чем в предыдущих примерах.

Если рассчитывать на большее число помех, то необходимо включить большее число эталонных элементов, по крайней мере по одному для каждой помехи.



Рис. 16.10. Диаграмма направленности для системы, показанной на рис. 16.9 при мощных помехах, приходящих под углами $\theta=30^\circ$ и $\theta=60^\circ$

Если число помех превышает число эталонных элементов, то адаптивный алгоритм приводит к минимизации выходной мощности, как, например, в схеме

на рис. 16.8. В противном случае алгоритм наименьших квадратов легко формирует нужное число направлений режекции в ДН.

16.2. Формирование лучей по пилот-сигналу

Из предыдущего материала следует, что адаптивное устройство подавления помех формирует провалы в ДН. Еще один вид адаптивного формирователя лучей основан на алгоритме с пилот-сигналом. В отличие от формирователя лучей Хауэллза–Аппельбаума, который сначала имеет ненаправленную диаграмму, а в результате адаптивного процесса уменьшает чувствительность в направлении мощных сигналов (которые считаются помехами), адаптивный формирователь лучей с пилот-сигналом формирует луч в заданном направлении приема и использует адаптивный процесс для поддержания этого луча при одновременном формировании направлений режекции для подавления помех, приходящих не по направлению приема. Этот процесс подавления определяется направлением прихода и уровнем мощности помех.

В процессе адаптации системы введенный пилот-сигнал моделирует принятый по направлению приема (выбранного оператором) сигнал. Этот же пилот-сигнал используют в качестве полезного отклика адаптивного устройства обработки, подключенного к элементам антенной решетки. По пилот-сигналу осуществляется обучение адаптивного формирователя лучей так, что его ДН имеет основной лепесток в заданном направлении приема, а также провалы, соответствующие направлениям прихода помех, отличным от направления приема. Таким образом, адаптация решетки проводится без формирования основного лепестка в направлении и с шириной полосы пропускания, определяемыми пилот-сигналом. В то же время в решетке режектируются не коррелированные с пилот-сигналом помехи и шум, приходящие по направлениям вне основного лепестка. Все эти свойства достигаются в наилучшем с точки зрения минимума СКО смысле.

Грифитс и Фрост предложили алгоритмы работы адаптивных антенн по результатам, аналогичные алгоритму с пилот-сигналом, но более простые в практической реализации и в некоторых случаях обладающие лучшими характеристиками. Во многих приложениях алгоритм с пилотсигналом заменяется на алгоритм Гриффитса и Фроста но для разработки других алгоритмов основой явился алгоритм с пилот-сигналом.

Многие решетки датчиков являются линейными в том смысле, что (ненаправленные) элементы антенны размещаются вдоль одной линии, или плоскими с размещением элементов в одной плоскости. Часто по такой схеме строятся антенные решетки. На рис. 16.11 приведен пример обычной приемной линейной антенной решетки. Антенна на рис. 16.11,а состоит из семи изотропных элементов, разнесенных друг от друга вдоль одной линии на расстояние $\lambda_0/2$, где λ_0 — длина волны центральной частоты решетки ω_0 . Принятые сигналы суммируются и образуют выходной сигнал решетки.

Диаграмма направленности, т. е. относительная чувствительность отклика на сигналы разных направлений, построена на рис. 16.11, а для углов $-\pi/2 < \theta < \pi/2$ и частоты ω_0 . Эта диаграмма является симметричной относительно направлений с $\theta=0^\circ$ и $\theta=90^\circ$, а основной лепесток расположен в центре при $\theta=0^\circ$. Наибольший по амплитуде боковой лепесток с $\theta=24^\circ$ имеет максимальную чувствительность на 12,5 дБ ниже максимальной чувствительности главного лепестка. Диаграмма направленности, построенная для отличных от ω_0 частот, имеет другой вид.

На рис. 16.11,б приведена схема той же решетки, однако в этом случае перед суммированием выходной сигнал каждого элемента задерживается. Результирующая ДН имеет теперь главный лепесток в направлении под углом ψ радиан, который аналогично (16.2) имеет вид

$$\psi = \sin^{-1} \frac{\lambda_0 \delta \omega_0}{2\pi d T} = \sin^{-1} \frac{c \delta}{d}, \quad (16.40)$$

где ω_0 —нормированная частота принятого сигнала; λ_0 —длина волны на частоте ω_0 ; δ —разность задержек выходных сигналов соседних элементов, число отсчетов; d —расстояние между элементами антенны, м; c —скорость распространения сигнала, равная $\lambda_0 \omega_0 / 2\pi T$ м/с; T —временной шаг, с. Чувствительность максимальна для угла ψ , поскольку сигналы, принятые от источника плоской волны, приходят под таким углом и вследствие задержек (рис. 16.11,б) находятся в фазе друг с другом, что приводит к максимальному выходному сигналу.

В рассматриваемом примере $d=\lambda_0/2$, $\delta=(0,8131/\omega_0)$, и, следовательно, $\psi=\sin^{-1}(\delta\omega_0/\pi)=15^\circ$.

Существует много возможных схем фазированных решеток. На рис 16.12, а показана одна такая схема, в которой каждый из выходных сигналов элементов антенны и задержанный на четверть периода частоты ω_0 (т. е. на 90° или $\pi T/2\omega_0$) параллельно умножаются на два весовых коэффициента.

Выходной сигнал является их суммой, но поскольку значения всех весовых коэффициентов равны единице, ДН на частоте ω_0 такая же по симметрии,

что и на рис. 16.11, а. Для иллюстрации на рис. 16.12, а штриховой линией показано направление прихода синусоидального сигнала помехи с частотой ω_0 . Направление прихода этой помехи (45°) таково, что она принимается по одному из боковых лепестков ДН с чувствительностью, только на 17 дБ меньшей чувствительности в направлении главного лепестка при $\theta=0^\circ$.

При других значениях весовых коэффициентов ДН на частоте ω_0 становится такой, какая показана на рис. 16.12,б. В этом случае главный лепесток остается почти неизменным и таким же, как на рис. 16.11, а и 16.12, а, в то время как боковой лепесток, по которому приходила синусоидальная

помеха в схеме на рис. 16.12, а, смещен таким образом, что в направлении помехи образовался провал. При этом чувствительность в направлении помехи на 77 дБ ниже чувствительности в направлении главного лепестка, что соответствует увеличению режекции помехи на 60 дБ.

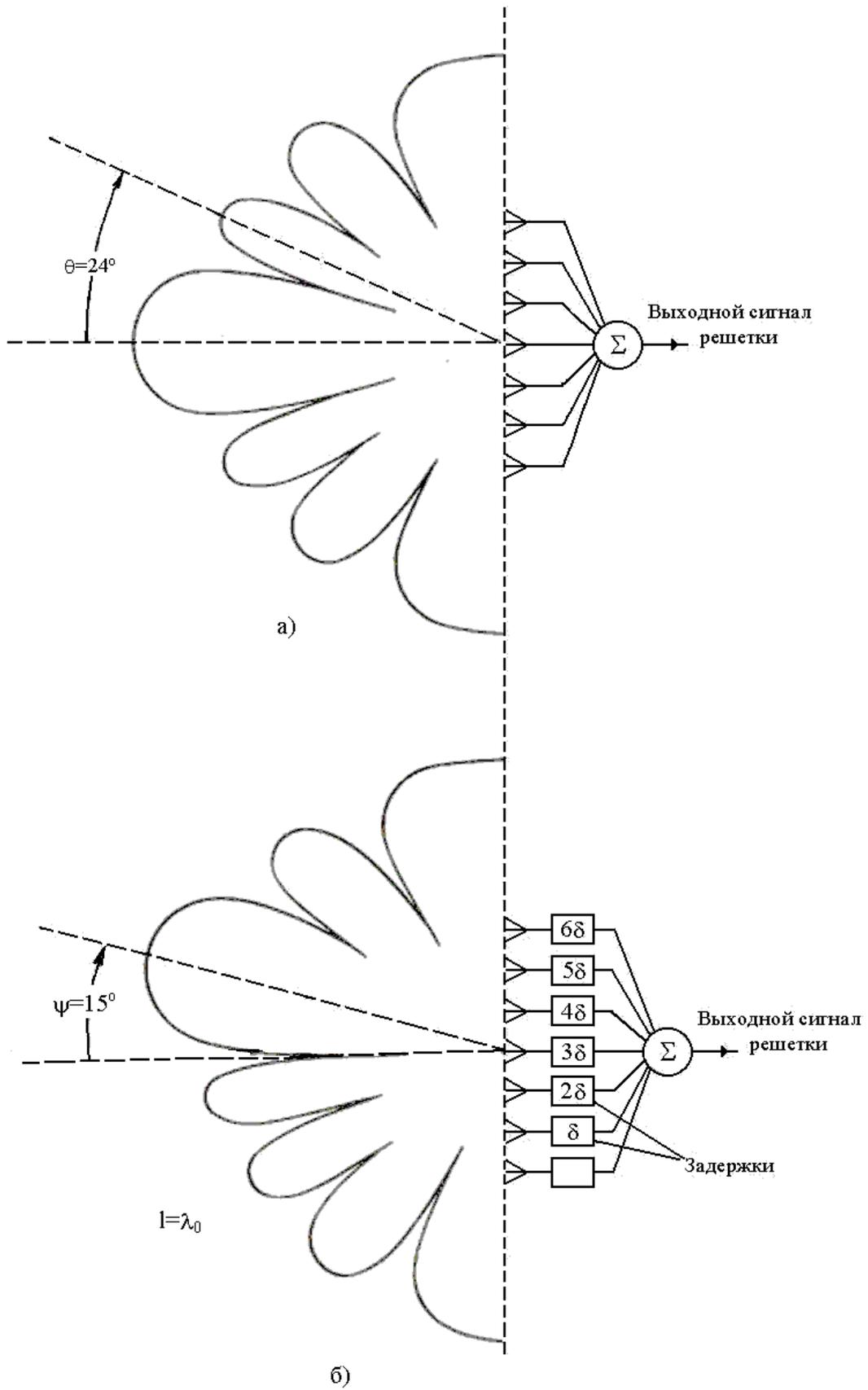


Рис. 16.11. Диаграмма направленности линейной решетки:
 а—простая решетка; б—решетка с задержками

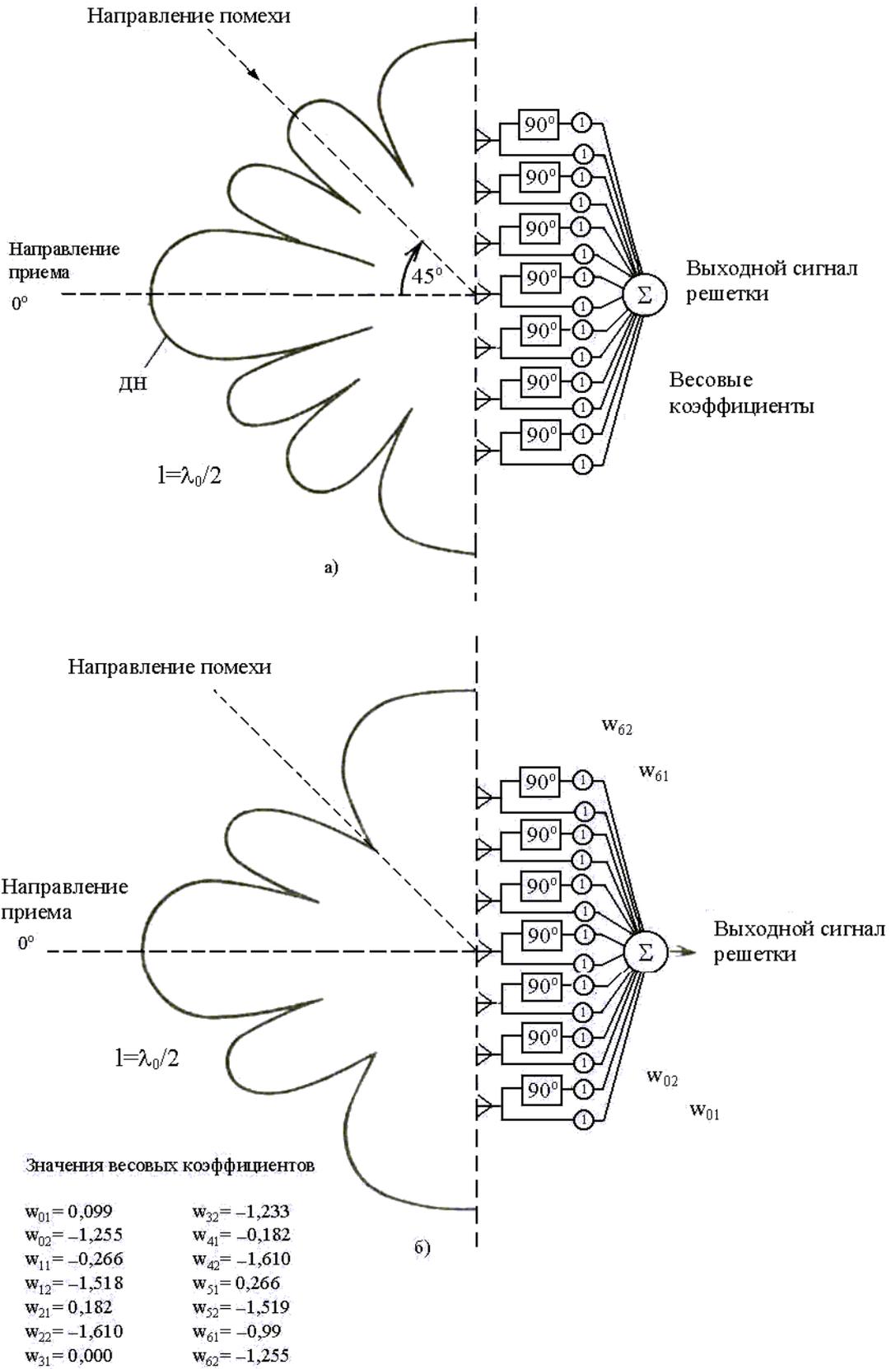


Рис. 16.12. Диаграмма направленности линейной решетки:
 а—схема с равномерным взвешиванием; б—схема с взвешиванием для
 подавления шума

Теперь чтобы показать, что существует (и его можно вычислить) множество весовых коэффициентов, при которых сигнал поступающий по заданному направлению, принимается, а помеха, приходящая по другим направлениям, режектируются, рассмотрим простой пример, приведенный на рис. 16.13. Сигнал, приходящий по заданному направлению с $\theta=0^\circ$, назовем пилот-сигналом $p_k=P \sin k\omega_0$. Пусть $n_k=N \sin k\omega_0$ —помеха, принимаемая решеткой под углом $\theta=\pi/6$ рад/с. В данном примере полагаем, что сигнал и помеха точно имеют одну и ту же частоту ω_0 и находятся в фазе в точке, расположенной посередине между элементами антенной решетки. Два одинаковых ненаправленных элемента разнесены на расстояние $\lambda_0/2$. Сигналы, принимаемые каждым элементом, разбиваются на прямую и задержанную на $\pi/2$ с составляющие, каждая из которых умножается на переменный весовой коэффициент. Далее все четыре взвешенных сигнала суммируются, и формируется выходной сигнал вида

$$A \sin (k\omega_0 + \varphi)$$

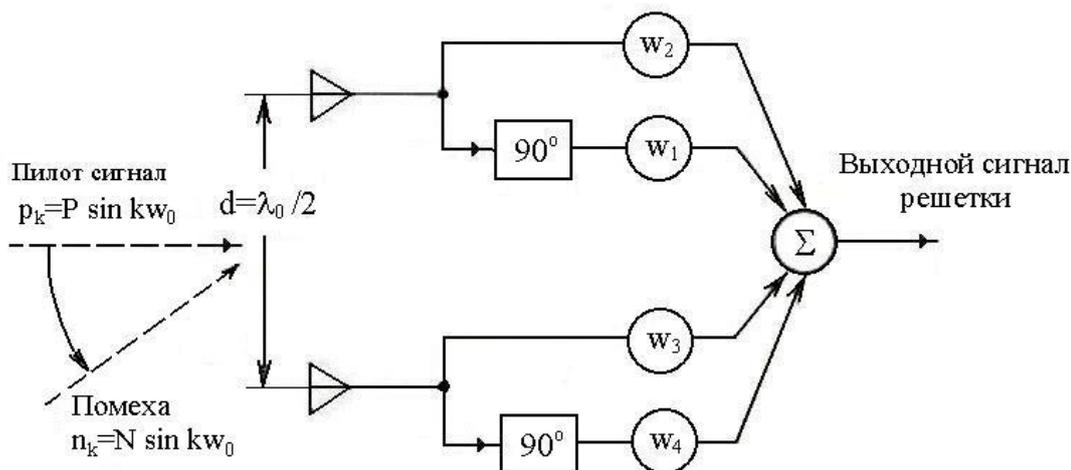


Рис. 16.13. Схема с пилот-сигналом, приходящим под углом $\theta=0^\circ$, и помехой, приходящей под углом $\theta=\pi/6$

При этом существует ряд решений, при которых выходной сигнал равен p_k . Однако, если считать, что решетка предназначена для режекции помехи, то выходной сигнал решетки не должен зависеть от амплитуды и фазы помехи. Этому ограничению удовлетворяет единственное множество весовых коэффициентов, которое находится следующим образом. При наличии пилот-сигнала выходной сигнал решетки имеет вид

$$P[(w_1+w_3) \sin k\omega_0 + (w_2+w_4)\sin(k\omega_0-\pi/2)]. \quad (16.41)$$

Этот выходной сигнал будет равен требуемому $p_k=\sin k\omega_0$ (т.е. пилот-сигналу) при условии, что

$$w_1+w_3=1; w_2+w_4=0 \quad (16.42)$$

По отношению к точке, расположенной в середине между элементами антенны, относительные задержки помехи на обоих элементах составляют $\pm(d/2c)\sin(\theta) = \pm d/4c = \pm \pi T/4\omega_0$, что соответствует сдвигам фаз $\pm\pi/4$ для частоты ω_0 . Тогда для помехи на входе, приходящей под углом $\theta = \pi/6$, выходной сигнал решетки имеет вид

$$N[\varpi_1 \sin(k\omega_0 - \pi/4) + \varpi_2 \sin(k\omega_0 - 3\pi/4)] + \varpi_3 \sin(k\omega_0 + \pi/4) + \varpi_4 \sin(k\omega_0 - \pi/4)]. \quad (16.43)$$

Этот сигнал будет равен нулю при условии, что

$$\varpi_1 + \varpi_4 = 0; \quad \varpi_2 - \varpi_3 = 0. \quad (16.44)$$

Таким образом, решая одновременно (16.42) и (16.44), можно найти множество весовых коэффициентов, удовлетворяющее требованиям к откликам на сигнал и помеху. Это множество равно

$$\varpi_1 = 1/2, \quad \varpi_2 = 1/2, \quad \varpi_3 = 1/2, \quad \varpi_4 = -1/2; \quad (16.45)$$

При этих весовых коэффициентах решетка обладает требуемыми свойствами, которые заключаются в том, что она принимает сигнал по требуемому направлению и режектирует помеху с частотой сигнала ω_0 по другим направлениям.

Описанный выше способ вычисления весовых коэффициентов является больше иллюстративным, чем практическим. Его можно использовать только при небольшом числе источников направленных помех, когда они являются монохроматическими, и направления их прихода известны априори. На практике устройства обработки должны быть такими, чтобы не требовалось подробной информации о числе и характере помех. Таковым является рассматриваемое далее адаптивное устройство, в котором по рекурсивному правилу находятся решения ряда уравнений, в результате чего достигается минимальная СКО между пилот-сигналом и общим выходным сигналом решетки.

16.3. Пространственные схемы

Прежде чем перейти к рассмотрению способов адаптивной фильтрации и обработки сигналов, которые можно использовать в адаптивных решетках, необходимо рассмотреть несколько видов схем пространственных решеток. На рис. 16.14 приведена схема адаптивной решетки для обработки узкополосных сигналов. Здесь каждый отдельный элемент антенны соединен с умножителем на переменный весовой коэффициент и с задержкой на четверть периода, выход которой, в свою очередь, соединен с еще одним умножителем на переменный

весовой коэффициент. Как показано на рисунке, взвешенные сигналы суммируются. Таким образом, сигнал каждого элемента, который полагаем монохроматическим или узкополосным, умножается на комплексный коэффициент передачи $Ae^{j\varphi}$. Фиксируя значения обоих весовых коэффициентов, можно выбрать любые фазовый угол $\varphi = -\text{tg}^{-1}(\omega_2/\omega_1)$ амплитуду комплексного коэффициента передачи $A = (\omega_1^2 + \omega_2^2)^{1/2}$ из области значений, ограниченной только значениями этих двух весовых коэффициентов. Таким образом, оба весовых коэффициента и задержка (на фазовый угол 90°) полностью обеспечивают перестраиваемую линейную обработку узкополосных сигналов, принятых каждым отдельным элементом антенны.

Решетка на схеме рис. 16.14 реализует общий способ сложения сигналов элементов антенны в перестраиваемой линейной структуре, когда сигналы и помехи являются узкополосными. Отметим, что такой же общности (для узкополосных сигналов) можно достичь даже в том случае, когда задержки не соответствуют точно фазовому сдвигу $\pi/2$ на центральной частоте ω_0 . Выбор сдвигов фаз, близких к $\pi/2$, необходим для обеспечения малых значений весовых коэффициентов, но принципиально не является обязательным. При необходимости принимать сигналы в некотором диапазоне частот каждую из задержек в схеме на рис. 16.14 можно заменить адаптивным трансверсальным фильтром (рис.16.15).

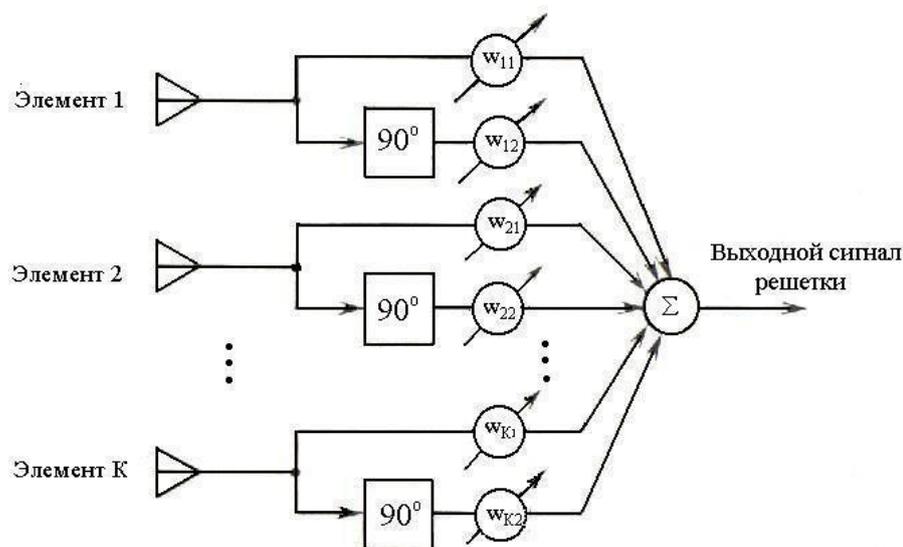


Рис. 16.14. Схема адаптивной решетки для приема узкополосных сигналов

Такая линия задержки с отводами позволяет перестраивать коэффициент передачи и фазу для ряда частот в заданном диапазоне. При очень небольшом временном шаге Δ эта схема приближается к идеальному фильтру, который

позволяет регулировать коэффициент передачи и фазу на каждой частоте в полосе пропускания.

Следующим шагом является разработка такого алгоритма адаптации, который позволял бы перестраивать весовые коэффициенты и добиваться требуемой пространственной и частотной фильтрации. Этот алгоритм должен обеспечивать заданный коэффициент передачи решетки в заданном направлении приема при одновременной режекции помех.

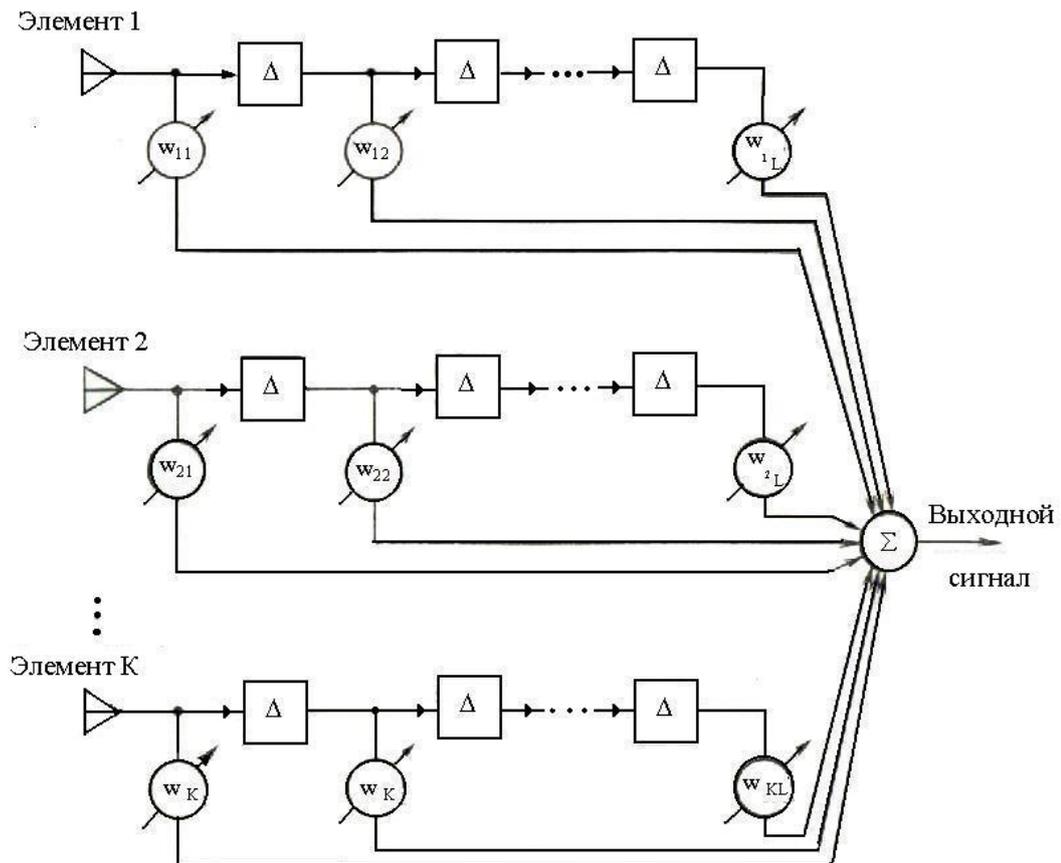


Рис. 16.15. Схема адаптивной решетки для приема широкополосных сигналов

Если принятые элементами адаптивной решетки сигналы состоят из суммы составляющих сигнала и шума, то сигнал воспроизводится (а шум подавляется) наилучшим в среднеквадратическом смысле образом тогда, когда полезным откликом адаптивного устройства обработки является сам сигнал. Однако обычно при адаптации сигнал не известен. Если бы сигнал был известен, то не нужны были бы приемники и приемная решетка.

В рассматриваемых здесь адаптивных антенных системах сигнал полезного отклика получают искусственным введением описанного выше пилот-сигнала, который полностью известен и обычно формируется в приемнике. Пилот-сигнал формируется таким образом, что его спектральные и пространственные характеристики соответствуют входному сигналу. В

некоторых случаях эти характеристики могут быть известны априори, но вообще они являются оценками параметров входного сигнала.

Адаптация с пилот-сигналом приводит к тому, что решетка формирует луч в направлении пилот-сигнала, который имеет, по существу, плоскую спектральную характеристику и линейный фазовый сдвиг в полосе пилот-сигнала. Более того, попадающие на адаптивную решетку направленные помехи приводят к тому, что в их полосах пропускания и их направлениях формируются провалы.

Обычно при введении пилот-сигнала использование выходного сигнала устройства формирования лучей становится неэффективным. Для исключения этого недостатка разработаны специальные алгоритмы адаптации. В одном из них попеременно осуществляется адаптация по пилот-сигналу для формирования основного луча, а затем, при отключенном пилот-сигнале, по естественным входным сигналам — для подавления помех. В другом алгоритме производится адаптация по входному сигналу в течение всего времени, но его реализация требует больших аппаратурных затрат.

16.4. Адаптивные алгоритмы

Схема на рис. 16.16 иллюстрирует способ получения пилот-сигнала, который передается с антенны, расположенной на некотором расстоянии от решетки, в необходимом направлении приема. На рис. 16.17 приведена схема получения местного пилот-сигнала, впервые предложенная Дж. Хоффом. На входные устройства обработки на рис. 16.17 подаются либо действительные входные сигналы элементов антенны (в режиме А), либо множество задержанных сигналов, полученных от генератора пилот-сигнала (в режиме Р). Фильтры $\delta_1, \dots, \delta_k$ (при одинаковых элементах решетки это идеальные задержки) выбраны так, что множество входных сигналов идентично сигналам, которые принимались бы решеткой от расположенного на расстоянии источника пилот-сигнала с плоской волной в требуемом направлении приема, которое соответствует основному лепестку ДН на приеме.

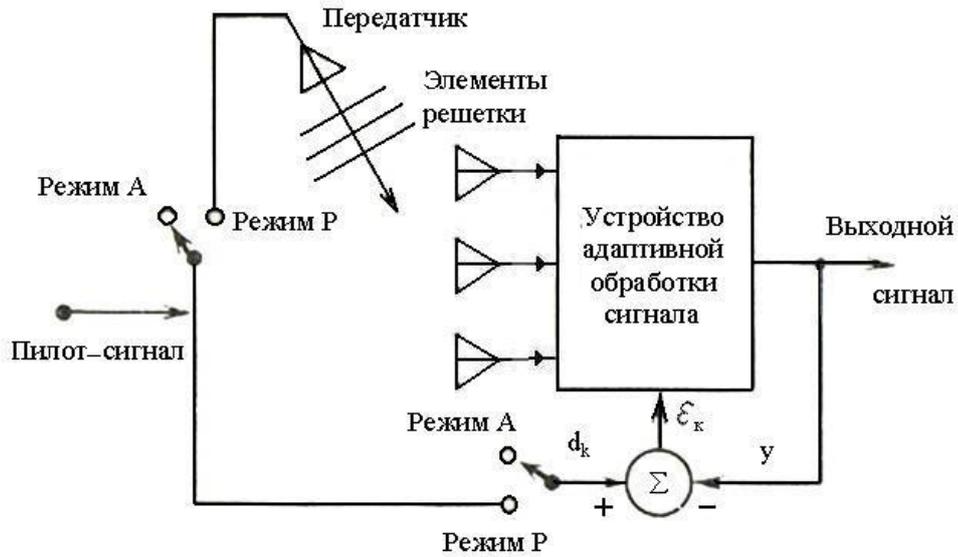


Рис. 16.16. Схема адаптации по внешнему пилот-сигналу с чередованием двух режимов: Р—при наличии пилот-сигнала; А—при отключенном пилот-сигнале

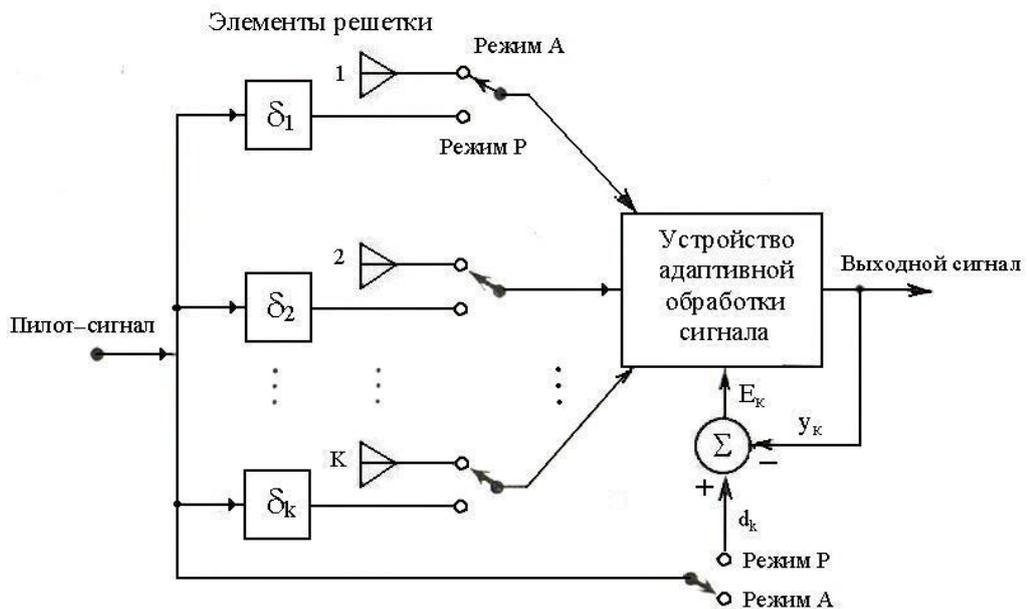


Рис. 16.17. Схема адаптации по внутреннему пилот-сигналу с чередованием двух режимов: Р—при наличии пилот-сигнала; А—при отключенном пилот-сигнале

Во время адаптации в режиме Р входные сигналы адаптивного устройства обработки получают из пилот-сигнала, а его полезным откликом является собственно пилот-сигнал. Например, при синусоидальном пилот-сигнале с

частотой ω_0 адаптация весовых коэффициентов с целью минимизации СКО приводит к тому, что коэффициент передачи антенной решетки в направлении приема имеет заданные амплитуду и фазовый сдвиг на частоте ω_0 .

Во время адаптации в режиме А все сигналы, подаваемые на адаптивное устройство обработки, принимаются элементами антенны из реального поля с шумом. В этом режиме процесс адаптации приводит к подавлению всех принятых сигналов, так как полезный отклик равен нулю. Однако непрерывное функционирование в режиме А приводит к тому, что значения всех весовых коэффициентов стремятся к нулю, и система отключается. Тем не менее при частом чередовании режимов Р и А в течение адаптации в каждом из них происходят лишь небольшие изменения вектора весовых коэффициентов, и можно приближенно поддерживать луч с единичным коэффициентом передачи в заданном направлении приема и, помимо этого, приблизительно минимизировать мощность принимаемой помехи.

В качестве пилот-сигнала можно взять сумму нескольких синусоидальных сигналов с различными частотами, так что при адаптации в режиме Р коэффициент передачи и фаза антенны имеют заданные для каждой из этих частот значения в направлении приема. Более того, если суммируются вместе несколько пилот-сигналов, соответствующих различным направлениям, то при адаптации в режиме Р можно поддерживать заданный коэффициент передачи решетки одновременно для различных углов и частот. Это свойство позволяет некоторым образом регулировать ширину полосы и ширину лучей по различным направлениям приема. В адаптивном режиме с двумя режимами приближенно осуществляется минимизация среднеквадратического значения (общей мощности) всех принятых элементами антенны сигналов, не коррелированных с пилот-сигналами; при этом коэффициент передачи и фаза луча приближенно равны заданным значениям для частот и направлений, определяемых составляющими пилот-сигнала. При адаптации с двумя режимами формирование и поддержание луча производится в режиме Р, а подавление в среднеквадратическом смысле (в соответствии с характеристиками пилот-сигнала) — в режиме А. В режиме Р из-за того, что устройство обработки соединено с генератором пилот-сигнала, прием сигнала невозможен. Следовательно, прием осуществляется только в режиме А. Этот недостаток устраняется в системе на рис. 16.18, в которой можно одновременно реализовать режимы Р и А. В этой системе пилот-сигналы и принятые сигналы подаются на адаптивное устройство обработки точно так же, как описано выше. Полезным откликом этого устройства является пилот-сигнал. Второе вспомогательное устройство обработки работает по реальному выходному сигналу решетки, но не реализует адаптивный процесс. В сигналах на его входе пилот-сигнал не содержится.

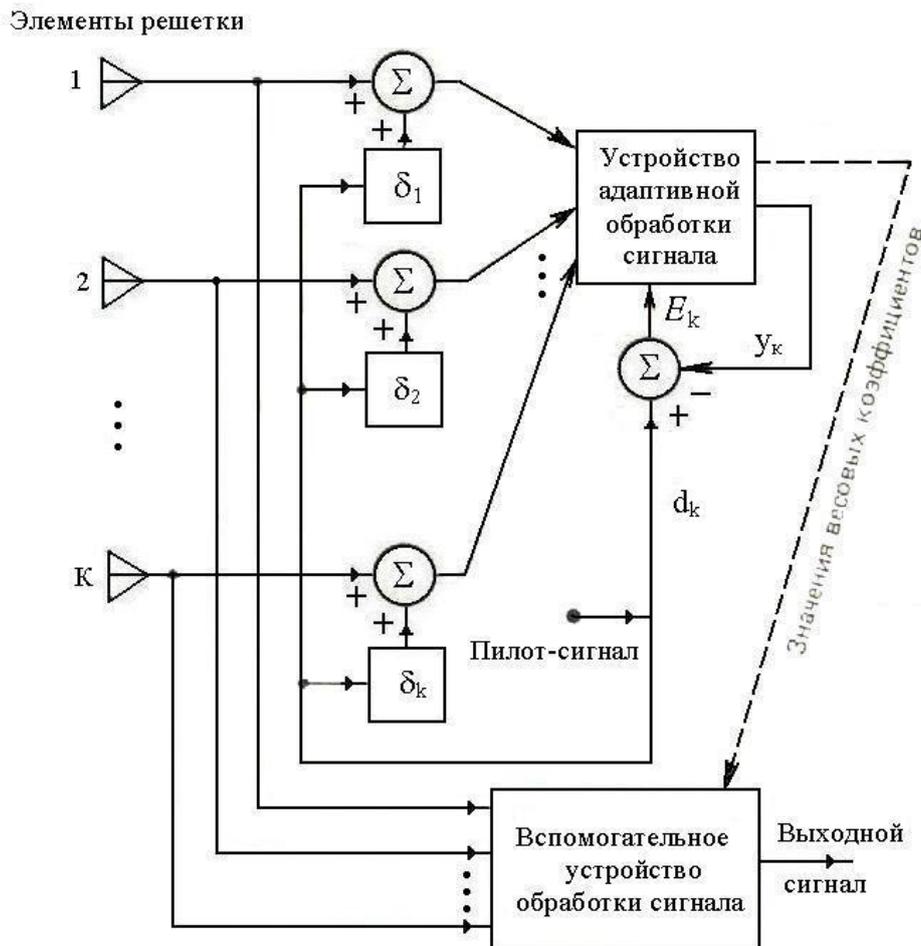


Рис. 16.18. Схема адаптации по пилот-сигналу

Вспомогательное устройство взаимодействует с адаптивным устройством обработки таким образом, что его весовые коэффициенты являются точными копиями соответствующих весовых коэффициентов адаптирующейся системы, т.е. для этого устройства нет необходимости принимать пилот-сигнал.

В приведенной на рис. 16.18 системе с одним режимом пилот-сигнал подключен постоянно. Процесс адаптации приводит к тому, что адаптивное устройство обработки с минимальной СКО воспроизводит пилот-сигнал и в то же время наилучшим в среднеквадратическом смысле образом режектирует все принятые элементами антенны сигналы, не коррелированные с пилот-сигналом. Таким образом, в результате адаптивного процесса ДН в полосе частот пилот-сигнала имеет нужную чувствительность в основном луче по направлению приема и провалы в направлениях помех для их полосы частот. При этом, как правило, чем мощнее помехи по сравнению с пилот-сигналом, тем выше уровень режекции.

Анализ приведенной на рис. 16.18 системы с одним режимом можно провести следующим образом. Оптимальные весовые коэффициенты адаптивного устройства обработки можно получить в виде произведения $\mathbf{R}^{-1}\mathbf{P}$.

Матрица \mathbf{R} равна сумме матрицы полученных от антенны сигналов и матрицы множества пилот-сигналов, причем полагаем, что пилот-сигнал и сигналы, поступающие от антенны, являются некоррелированными. Вектор \mathbf{P} определяется только пилот-сигналом. Элементы этого вектора представляют собой значения автокорреляционной функции пилот-сигнала, при этом задержки определяются соответствующими задержками пилот-сигнала. Поскольку смещение главного лепестка зависит от выбора задержек пилот-сигнала, вектор \mathbf{P} непосредственно связан с углом этого смещения. Для определения характеристик функционирования адаптивных антенных систем проведено большое число экспериментов по моделированию на ЭВМ решеток с различной геометрической конфигурацией и различных видов сигналов и помех. Для простоты изложения приводимые в следующих подразделах примеры ограничиваются плоскими решетками, составленными из идеальных изотропных (ненаправленных) элементов. В каждом случае для адаптации используется метод наименьших квадратов. Во всех экспериментах начальные значения весовых коэффициентов приняты одинаковыми.

16.5. Моделирование для узкополосных сигналов

На рис. 16.19 показаны 12-элементная круговая решетка и устройство обработки сигнала, которые используются для анализа функционирования узкополосной системы, приведенной на рис.16.14. Здесь пилот-сигнал представляет собой синусоидальный сигнал единичной амплитуды с частотой ω_0 и мощностью $\sigma_p^2=0,5$ и предназначен для обучения решетки в направлении с $\theta=0^\circ$. Помеха состоит из синусоидального мешающего сигнала (с теми же частотой и мощностью, что и у пилот-сигнала) с направлением прихода $\theta=40^\circ$ и малой .случайной некоррелированной помехи в виде белого шума с мощностью σ_n^2 на каждом элементе решетки. При моделировании этой системы адаптация весовых коэффициентов осуществляется по алгоритму наименьших квадратов в двух режимах.

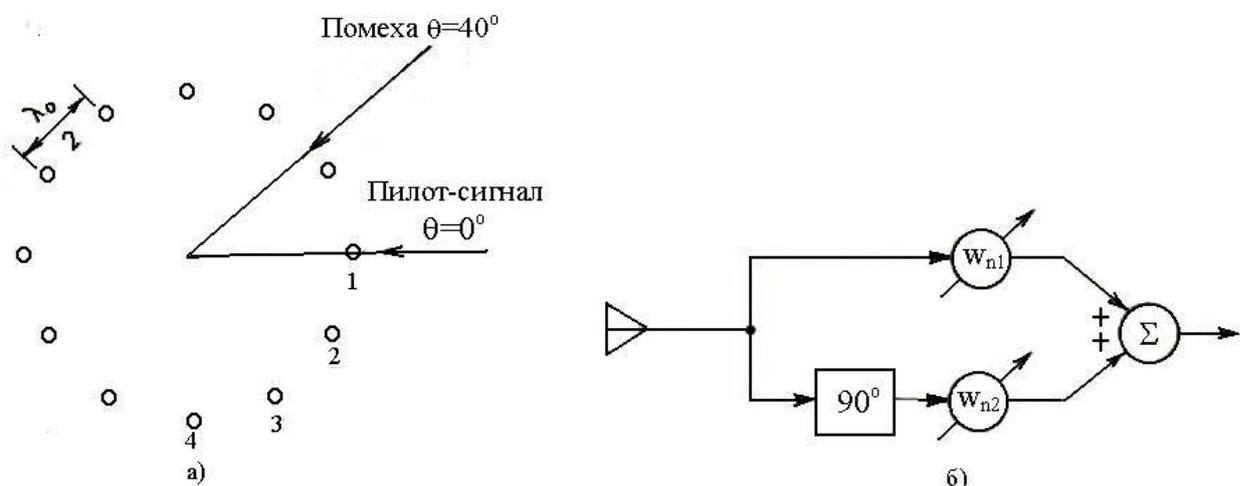


Рис. 16.19. Схема эксперимента для двух режимов с узкополосными

сигналами: а– расположение элементов; б– обработка сигнала n–го элемента

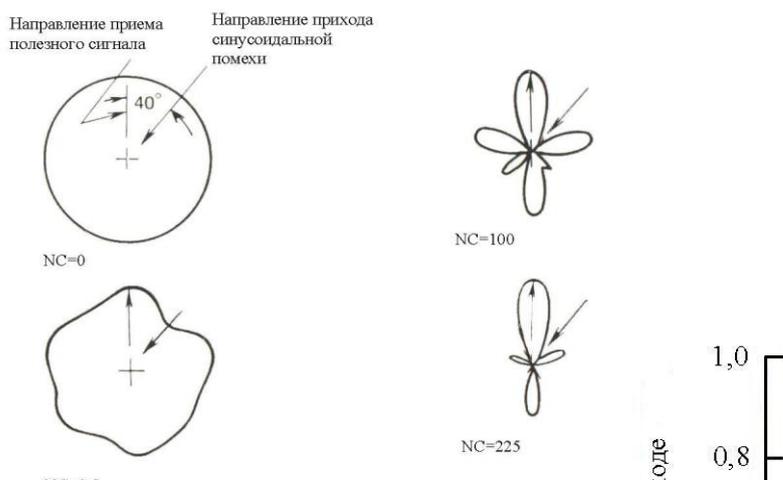
На рис. 16.20 показана последовательность ДН, изменяющихся в процессе обучения решетки. Эти диаграммы построены с помощью ЭВМ и представляют собой выраженную в децибелах чувствительность решетки на частоте ω_0 аналогично предыдущим ДН, рассмотренным в данной главе. Каждая ДН рассчитана для множества весовых коэффициентов, получаемых на различных стадиях процесса адаптации. Отметим, что начальная ДН из-за симметричности антенной решетки при равных начальных значениях весовых коэффициентов, по существу, является круговой. На каждой ДН указано число прошедших (на момент вычисления данной ДН) периодов NC частоты ω_0 . В этих экспериментах число циклов адаптации равно $20NC$. Отметим, что если $\omega_0/2T=1$ кГц, то $NC=1$ соответствует длительности 1 мс в реальном масштабе времени, если $\omega_0/2T=1$ МГц, то $NC=1$ соответствует 1 мкс и т. д.

По приведенной на рис. 16.20 последовательности ДН можно сделать некоторые выводы. В процессе адаптации чувствительность решетки в направлении приема по существу остается постоянной, а чувствительность в направлении источника синусоидальной помехи очень быстро падает; по мере развития процесса адаптации в ДН в направлении помехи формируется глубокий провал. После завершения переходного процесса чувствительность решетки в направлении помехи ниже чувствительности в заданном направлении приема на 27 дБ. Адаптивный алгоритм формирует провал при отсутствии априорных данных о частоте и о направлении прихода помехи.

Полная мощность помехи на выходе решетки равна сумме мощностей белого шума и направленной синусоидальной помехи. В процессе адаптации эта мощность обычно уменьшается до некоторого предельного уровня.

На рис. 16.21 приведена зависимость полной мощности помехи на выходе NC , представляющая собой обучающую кривую адаптивного формирователя лучей. При начальных значениях весовых коэффициентов уровень мощности составляет 0,65. После адаптации он уменьшается до 0,01. Для рассматриваемых помех отношение сигнал-шум на выходе после адаптации больше, чем для единичного изотропного приемного элемента, в 60 раз.

Для решетки на рис. 16.19 и адаптивного процесса с двумя режимами проведен эксперимент, в котором адаптивная решетка работала при наличии нескольких источников направленных помех.



В этом случае помеха представляет собой сумму помех одновременно действующих направленных синусоидальных помех с амплитудой 0,15 и мощностью 0,125 и некоррелированных составляющих белого шума с мощностью 0,5 на каждом из элементов антенны. Параметры направленных помех приведены в табл. 16.1.

Таблица 16.1

Направление помехи, град	Относительная частота помехи (относительно ω_0)	Уровень подавления помехи относительно приема, дБ
67	1,10	-26
134	0,95	-30
191	1,00	-28
236	0,90	-30
338	1,05	-38

На рис. 16.22 показаны процесс изменения ДН для частоты ω_0 от начала адаптации до конечного установившегося (оптимального) состояния, которое достигается через $NC=682$ периодам частоты ω_0 (а) и обучающая кривая для этого случая (б).

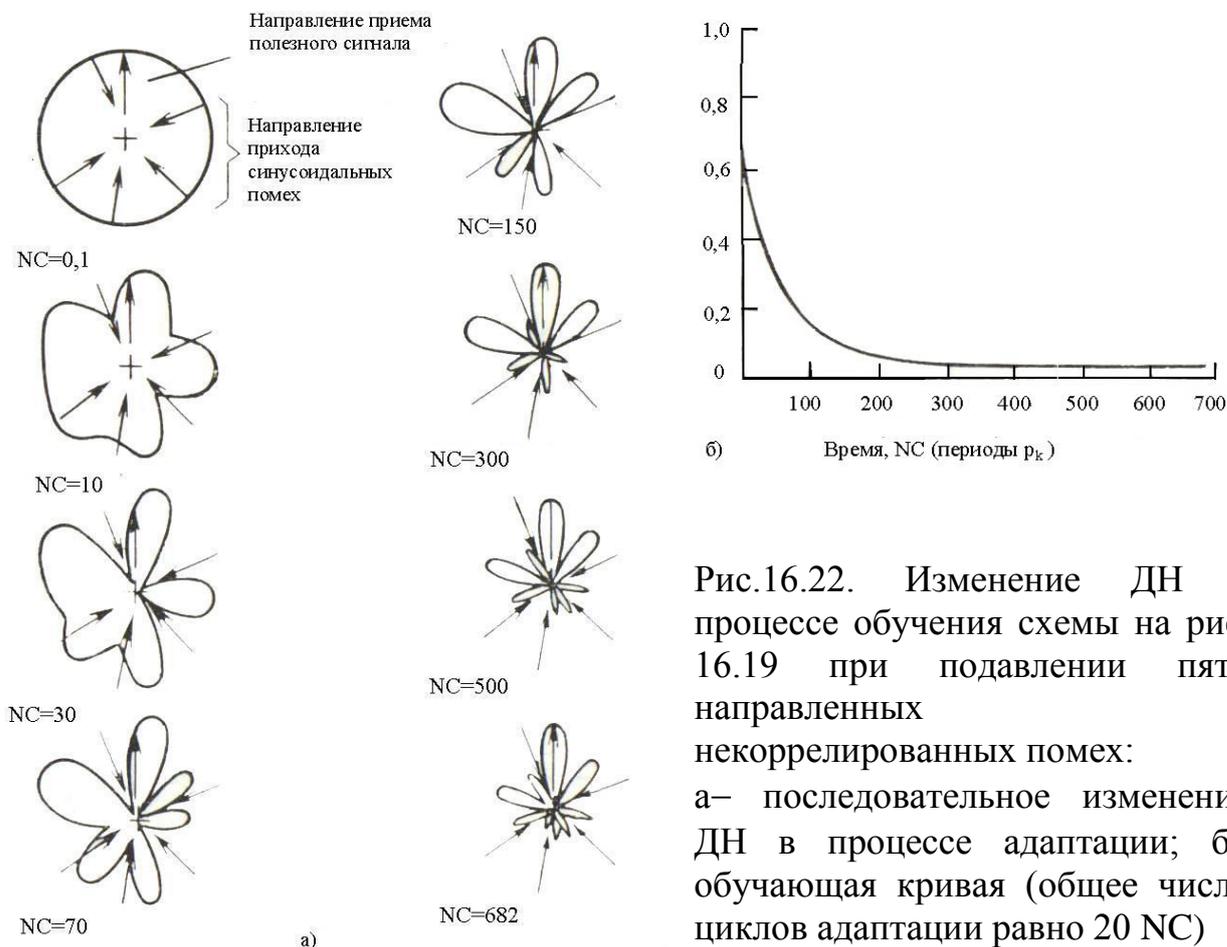


Рис.16.22. Изменение ДН в процессе обучения схемы на рис. 16.19 при подавлении пяти направленных и некоррелированных помех: а– последовательное изменение ДН в процессе адаптации; б– обучающая кривая (общее число циклов адаптации равно $20 NC$)

В табл. 16.1 приведены результирующие значения чувствительности решетки по пяти направлениям помех относительно чувствительности по направлению приема. Отношение сигнал-шум по сравнению с единичным изотропным излучателем в этом случае повышается в 15 раз. Из рис. 16.22,б следует, что постоянная времени обучающей кривой равна примерно 70 периодам. Поскольку на один период частоты ω_0 приходится 20 шагов адаптации, постоянная времени обучающей кривой $\tau_{ско}=1400$. Установившееся состояние адаптивного процесса наступает примерно в течение 400 периодов частоты ω_0 .

При $\omega_0/2\pi=1$ МГц реальное время установления равно 400 мкс. Относительное среднее значение СКО для этого адаптивного процесса:

$$M = \frac{\text{число весовых коэффициентов}}{4\tau_{ско}} = \frac{24}{4\tau_{ско}} = \frac{6}{1400} = 0,43\% \quad (16.46)$$

Такое небольшое полученное значение указывает на то, что адаптивный процесс является очень медленным и приводит к точному результату. Это следует также из характера приведенной на рис. 16.22,б обучающей кривой для данного эксперимента, которая является сглаженной и не имеет шумовой составляющей.

Если все синусоидальные помехи имеют одну и ту же частоту, то адаптивное устройство формирования лучей также осуществляет их подавление, но при этом в общем случае не формируются провалы в направлении их прихода. Адаптивный формирователь лучей принимает все эти помехи и устанавливает для них такие амплитудно-фазовые соотношения, что при их суммировании помеховая составляющая на выходе приближенно равна нулю. Если частоты различны, как это имело место в эксперименте, то адаптивное устройство формирования лучей для подавления синусоидальных помех обязательно формирует провалы.

Литература

1. Радиотехнические системы: Учеб. для вузов / Ю. П. Гришин, В. П. Ипатов, Ю. М. Казаринов и др.; Под ред. Ю. М. Казаринова. М.: Высшая школа, 1990. 496 с.
2. Теория передачи сигналов: Учеб. для вузов / А. Г. Зюко, Д. Д. Кловский, М. В. Назаров, Л. М. Финк. М.: Радио и связь, 1986. 304 с.
3. Радиотехнические системы передачи информации: Учеб. для вузов / В. А. Борисов, В. В. Калмыков, Я. М. Ковальчук и др.; Под ред. В. В. Калмыкова. М.: Радио и связь, 1990. 304 с.
4. Баскаков С. И. Радиотехнические цепи и сигналы: Учеб. для вузов. М.: Высшая школа, 1988. 448 с.
5. Тихонов В. И. Оптимальный прием сигналов. М.: Радио и связь, 1983. 320 с.
6. Аграновский К. Ю., Златогурский Д. Н., Киселев В. Г. Радиотехнические системы: Учебное пособие для вузов. М.: Высшая школа, 1979. 333 с.
7. Никольский В. В., Никольская Т. И. Электродинамика и распространение радиоволн: Учебное пособие для вузов. М.: Наука, 1989. 544 с.
8. Распространение оптического и миллиметрового излучения в атмосфере: Учебное пособие / Е. М. Калабанов, В. И. Юдин. Воронеж: ВГТУ, 1997. 57 с.
9. Теоретические основы радиолокации / А. А. Коростылев, А. Ф. Клюев, Ю. А. Мельник и др. Под ред. В. Е. Дулевича. М.: Сов. радио, 1978.
10. Кочержевский Г. Н., Ерохин Г. А., Козырев Н. Д. Антенно-фидерные устройства. М.: Радио и связь, 1989. 352 с.
11. Сазонов Д. М. Антенны и устройства СВЧ Учеб. для вузов. М.: Радио и связь, 1988. 432 с.
12. Черный Ф. Б. Распространение радиоволн. М.: Сов. радио, 1972. 463 с.
13. Максимов М. В., Горгонов Г. И. Радиоэлектронные системы самонаведения. М.: Радио и связь, 1982. 304 с.
14. Антенны и устройства СВЧ: Проектирование фазированных антенных решеток / Д. И. Воскресенский, В. Л. Гостюхин, Р. А. Грановская и др.; Под ред. Д. И. Воскресенского. М.: Радио и связь, 1981. 431 с.
15. Дымова А. И., Альбац М. Е., Бонч-Бруевич А. М. Радиотехнические системы. Учеб. для вузов. М.: Сов. радио, 1975. 440 с.
16. Твердотельные устройства СВЧ в технике связи / Л. Г. Гассанов, А. А. Липатов, В. В. Марков, Н. А. Могильченко. М.: Радио и связь, 1988. 288 с.
17. Вартанесян В. А. Радиоэлектронная разведка. М.: Воениздат, 1991. 254 с.
18. Малашин М. С., Каминский Р. П., Борисов Ю. Б. Основы проектирования лазерных локационных систем: Учебное пособие для вузов. М.: Высшая школа, 1983. 207 с.
19. Справочник по лазерной технике. Киев: Техника, 1978. 288 с.
20. Протопопов В. В., Устинов Н. Д. Инфракрасные лазерные локационные системы. М.: Воениздат, 1987. 175 с.
21. Б. Уидроу, С. Стирнз. Адаптивная обработка сигналов / Пер. с англ. под ред. В. В. Шахгильдяна. М.: Радио и связь. 1989. 440 с.

СОДЕРЖАНИЕ

Предисловие	3
I. Введение в теорию радиотехнических систем	4
§ 1.1. Определение и классификация радиотехнических систем	4
§ 1.2. Общая модель радиосистемы	4
§ 1.3. Эффективность радиотехнических систем, их тактические и технические характеристики	6
II. Введение в статистическую теорию радиотехнических систем	9
§ 2.1. Модели радиосигналов и помех в радиосистемах	9
§ 2.2. Прием сигналов, как статистическая задача проверки гипотез	11
§ 2.3. Отношение правдоподобия. Функция правдоподобия	15
§ 2.4. Статистические критерии оптимальности при решении задач приема и обработки сигналов	16
III. Обнаружение и различение сигналов	19
§ 3.1. Алгоритмы оптимального обнаружения детерминированных сигналов	19
§ 3.2. Структуры оптимальных приемников обнаружения детерминированных сигналов	21
§ 3.3. Оптимальное обнаружение сигналов со случайной начальной фазой	26
§ 3.4. Оптимальное обнаружение пачек радиоимпульсов	29

§ 3.5. Оптимальное различение детерминированных сигналов.	33
IV. Оптимальная оценка параметров сигналов	38
§ 4.1. Задача оптимальной оценки параметров сигналов	38
§ 4.2. Байесовская оценка неэнергетических параметров сигналов	40
§ 4.3. Устройства байесовской оценки неэнергетических параметров сигнала	44
§ 4.4. Оптимальная оценка начальной фазы и смещения частоты сигнала	47
V. Основы теории разрешения сигналов	50
§ 5.1. Разрешение сигналов.	50
§ 5.2. Двумерная автокорреляционная функция	50
§ 5.3. Функция неопределенности и ее свойства. Принцип неопределенности в радиолокации	53
§ 5.4. Широкополосные сигналы и их особенности	56
VI. Общие сведения о радиолокации	60
§ 6.1. Основные понятия и определения	60
§ 6.2. Классификация РЛС, их тактические и технические характеристики	61
§ 6.3. Основные сведения о радиотехнических методах измерения координат и их производных	62
§ 6.4. Устройство, принцип действия и основные характеристики РЛС кругового обзора	63

VII. Физические основы радиолокации	67
§ 7.1. Объекты радиолокационного наблюдения и формирования отраженных сигналов	67
§ 7.2. Эффективная площадь рассеяния объекта (ЭПР)	68
§ 7.3. ЭПР простейших объектов	69
VIII. Дальность действия радиотехнических систем	72
§ 8.1. Дальность действия радиолиний	72
§ 8.2. Влияние условий распространения радиоволн на дальность действия РТС	77
§ 8.3. Влияние отражения радиоволн от земной поверхности на дальность действия РЛС	78
IX. Поиск объектов в зоне действия РЛС	82
§ 9.1. Радиолокационный обзор пространства	82
§ 9.2. Методы последовательного обзора пространства	83
§ 9.3. Параллельный и управляемый обзор пространства	87
X. Измерение дальности, радиальной скорости и угловых координат объектов	89
§ 10.1. Методы и устройства измерения дальности объектов	89
§ 10.2. Методы и устройства измерения радиальной скорости объектов	92
§ 10.3. Совместное измерение дальности и радиальной скорости объектов частотным методом	93
§ 10.4. Методы и устройства радиопеленгования	95

<i>XI. Радиоэлектронная борьба и защита РТС от помех</i>	<i>101</i>
§ 11.1. <i>Радиоэлектронное подавление и защита РТС от активных помех</i>	<i>101</i>
§ 11.2. <i>Защита РТС от пассивных помех</i>	<i>103</i>
§ 11.3. <i>Системы радиоэлектронной разведки</i>	<i>105</i>
<i>XII. Основы построения радиотеплолокаторов и оптических локационных систем</i>	<i>107</i>
§ 12.1. <i>Физические основы радиотеплолокации</i>	<i>107</i>
§ 12.2. <i>Обнаружение радиотепловых сигналов. Принципы построения радиотеплолокаторов</i>	<i>108</i>
§ 12.3. <i>Физические основы оптической локации</i>	<i>111</i>
§ 12.4. <i>Дальность действия лазерных локационных систем</i>	<i>113</i>
<i>XIII. Адаптивные системы</i>	<i>116</i>
§ 13.1. <i>Определения и свойства</i>	<i>116</i>
§ 13.2. <i>Области применения адаптивных устройств и систем</i>	<i>117</i>
§ 13.3. <i>Общие свойства адаптивных устройств и систем</i>	<i>118</i>
§ 13.4. <i>Адаптация с обратной связью и без обратной связи</i>	<i>119</i>
§ 13.5. <i>Приложения алгоритмов адаптации с обратной связью</i>	<i>122</i>
§ 13.6. <i>Простой пример адаптивной системы</i>	<i>124</i>
<i>XIV. Адаптивный линейный сумматор</i>	<i>126</i>
§ 14.1. <i>Общее описание адаптивного линейного сумматора</i>	<i>126</i>
§ 14.2. <i>Векторы входного сигнала и весовых коэффициентов</i>	<i>127</i>
§ 14.3. <i>Полезный отклик и сигнал ошибки</i>	<i>129</i>

§ 14.4. Рабочая функция	130
§ 14.5. Минимальная среднеквадратическая ошибка и градиент	132
§ 14.6. Пример анализа рабочей функции	133
§ 14.7. Альтернативное представление градиента	136
§ 14.8. Декорреляция сигнала ошибки и элементов входного сигнала	137
XV. Поиск рабочей функции	138
§ 15.1. Методы поиска параметров рабочей функции	138
§ 15.2. Основные принципы методов градиентного поиска	138
§ 15.3. Простой алгоритм градиентного поиска	140
§ 15.4. Устойчивость и скорость сходимости	141
§ 15.5. Обучающая кривая	142
§ 15.6. Градиентный поиск методом Ньютона	143
§ 15.7. Метод Ньютона для многомерного пространства	146
§ 15.8. Градиентный поиск методом наискорейшего спуска	148
§ 15.9. Сравнение обучающих кривых	152
XVI. Введение в адаптивные антенные решетки и адаптивное формирование лучей	155
§ 16.1. Подавление боковых лепестков	156
§ 16.2. Формирование лучей по пилот-сигналу	170
§ 16.3. Пространственные схемы	175
§ 16.4. Адаптивные алгоритмы	178
§ 16.5. Моделирование для узкополосных сигналов	181

<i>Литература</i>	<i>186</i>
<i>Содержание</i>	<i>187</i>