

Информатика в техническом университете

Информатика
в техническом
университете

Информационные технологии в радиотехнических системах

Информационные технологии в радиотехнических системах

ISBN 5-7038-2263-7



9 785703 822630

Издательство МГТУ имени Н.Э.Баумана

Серия основана в 2000 году

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

д-р техн. наук *И.Б. Федоров* — главный редактор
д-р техн. наук *И.П. Норенков* — зам. главного редактора
д-р техн. наук *Ю.М. Смирнов* — зам. главного редактора
д-р техн. наук *В.В. Девятков*
д-р техн. наук *В.В. Емельянов*
канд. техн. наук *И.П. Иванов*
д-р техн. наук *В.А. Матвеев*
канд. техн. наук *Н.В. Медведев*
д-р техн. наук *В.В. Сюзев*
д-р техн. наук *Б.Г. Трусов*
д-р техн. наук *В.М. Черненький*
д-р техн. наук *В.А. Шахнов*

Информационные технологии в радиотехнических системах

Под редакцией И.Б. Федорова

*Допущено Министерством образования Российской Федерации
в качестве учебного пособия для студентов высших учебных
заведений, обучающихся по специальностям «Радиотехника»
и «Радиоэлектронные системы» направления подготовки
дипломированных специалистов «Радиотехника»*

УДК 621.396(075.8)

ББК 32

И741

Рецензенты:

кафедра «Радиотехнические системы» Спб ГЭТУ
(зав. кафедрой д-р техн. наук проф. В.М. Кутузов);
д-р техн. наук проф. Ю.Г. Сосулин

Авторы:

В.А. Васин, И.Б. Власов, Ю.М. Егоров, В.В. Калмыков, А.А. Кузнецов,
А.И. Николаев, В.Б. Пудловский, В.А. Родзивилов, Ю.Н. Себекин,
А.И. Сенин, Г.П. Слукин, И.Б. Федоров

Информационные технологии в радиотехнических системах:

И741 Учебное пособие / В.А. Васин, И.Б. Власов, Ю.М. Егоров и др.; Под
ред. И.Б. Федорова. — М.: Изд-во МГТУ им. Н.Э. Баумана, 2003. —
672 с.: ил. — (Сер. Информатика в техническом университете.)

ISBN 5-7038-2263-7

Изложены основы статистической теории радиосистем. Рассмотрены задачи обнаружения, различения, разрешения сигналов на фоне помех, а также измерения параметров сигналов, в том числе изменяющихся за время наблюдения. Сделана попытка обобщить методический подход к анализу и синтезу основных информационных технологий для различных радиотехнических систем: радиолокационных, спутниковых, радионавигационных и систем передачи информации.

Содержание учебного пособия соответствует курсам лекций, которые авторы читают в МГТУ им. Н.Э. Баумана.

Для студентов высших технических учебных заведений, обучающихся по направлению подготовки дипломированных специалистов «Радиотехника». Может быть полезным аспирантам, работникам научно-исследовательских организаций и промышленности, занимающихся вопросами разработки радиотехнических систем различного назначения.

УДК 621.396(075.8)

ББК 32

© Коллектив авторов, 2003

© МГТУ им. Н.Э. Баумана, 2003

ISBN 5-7038-2263-7

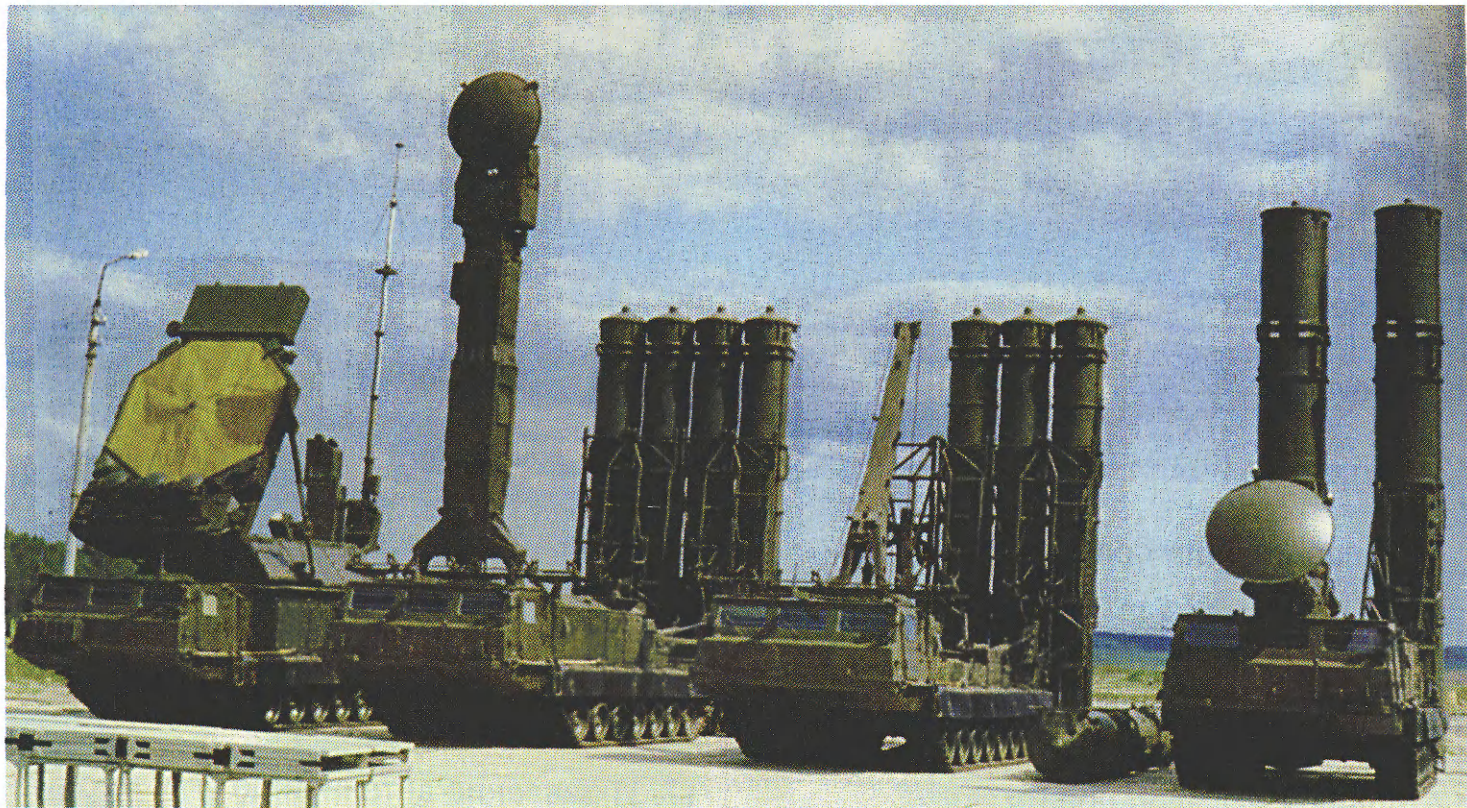


Рис.7. Зенитная ракетная система С-300В

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ.....	12
СПИСОК ОСНОВНЫХ СОКРАЩЕНИЙ.....	15
СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ.....	17
1. ОБЩИЕ СВЕДЕНИЯ О РАДИОТЕХНИЧЕСКИХ СИСТЕМАХ.....	19
1.1. Роль радиотехнических систем в современном обществе	19
1.2. Классификация радиотехнических систем	20
1.3. Тактико-технические характеристики радиотехнических систем	25
1.4. Энергетические соотношения в радиотехнических системах	28
1.5. Тенденции развития радиотехнических систем	32
Контрольные вопросы	36
2. СИГНАЛЫ И ПОМЕХИ В РАДИОТЕХНИЧЕСКИХ СИСТЕМАХ.....	37
2.1. Информация, сообщение, сигналы.....	37
2.2. Математические модели сигналов и помех.....	39
2.3. Векторное представление сигналов	43
2.4. Дискретизация непрерывных сигналов	47
2.5. Преобразование непрерывных сигналов в цифровую форму.....	56
2.6. Сложные сигналы	59
2.6.1. Линейные рекуррентные последовательности максимальной длины	61
2.6.2. Линейные рекуррентные последовательности не максимальной длины	69
2.6.3. Сегменты М-последовательностей.....	72
2.7. Системы сигналов.....	73
2.7.1. Ортогональные сигналы.....	75
2.7.2. Биортогональные сигналы	81
2.7.3. Симплексные сигналы	81
2.8. Моделирование сигналов и помех	83
2.8.1. Моделирование случайных величин	83
2.8.2. Моделирование случайных векторов	89
2.8.3. Моделирование гауссовских случайных процессов	94
2.8.4. Моделирование негауссовских случайных процессов	105
Контрольные вопросы	109

3. ОСНОВЫ ТЕОРИИ ОБНАРУЖЕНИЯ И РАЗЛИЧЕНИЯ СИГНАЛОВ	111
3.1. Общие сведения	111
3.2. Основные положения теории статистических решений. Оптимальные критерии	114
3.3. Обнаружение сигналов	120
3.3.1. Оптимальные алгоритмы обнаружения сигналов	120
3.3.2. Обнаружение детерминированного сигнала на фоне белого шума ..	134
3.3.3. Обнаружение сигнала со случайной начальной фазой на фоне белого шума	143
3.3.4. Обнаружение сигнала со случайными амплитудой и начальной фазой на фоне белого шума	148
3.3.5. Обнаружение пачки сигналов	151
3.4. Различение сигналов	160
3.4.1. Оптимальные алгоритмы различения сигналов	160
3.4.2. Различение двух детерминированных сигналов на фоне белого шума	160
3.4.3. Различение m детерминированных сигналов на фоне белого шума ..	168
3.4.4. Различение двух сигналов со случайной начальной фазой на фоне белого шума	173
3.4.5. Различение m сигналов со случайной начальной фазой на фоне белого шума	179
3.5. Оптимальный прием сигналов на фоне небелого шума	182
3.5.1. Обнаружение сигнала	182
3.5.2. Различение сигналов	190
3.6. Обнаружение сигналов в условиях априорной неопределенности	192
Контрольные вопросы	198
4. РАЗРЕШЕНИЕ СИГНАЛОВ	200
4.1. Понятие о разрешении сигналов	200
4.2. Функция рассогласования (неопределенности) в теории разрешения	203
4.2.1. Общие сведения о функции рассогласования	203
4.2.2. Время-частотные функции рассогласования	205
4.3. Функции рассогласования когерентных сигналов	209
4.3.1. Сигналы без внутримпульсной модуляции	209
4.3.2. Частотно-модулированные сигналы	214
4.4. Сигналы, обеспечивающие высокие разрешающие способности по времени запаздывания и частоте	220
4.4.1. Особенности выбора время-частотных функций рассогласования ...	220
4.4.2. Фазоманипулированные сигналы	221
4.5. Угло-поляризационные функции рассогласования	225
4.6. Согласованное и оптимальное разрешения сигналов	227
Контрольные вопросы	230
5. ОСНОВЫ ТЕОРИИ ИЗМЕРЕНИЯ ПАРАМЕТРОВ СИГНАЛОВ	231
5.1. Общие сведения	231

5.1.1. Измерение случайных параметров	233
5.1.2. Измерение неслучайных параметров	236
5.1.3. Особенности многократного измерения случайных параметров	237
5.2. Измерение параметров радиолокационных сигналов	238
5.2.1. Неследящие измерители дальности.....	240
5.2.2. Неследящие измерители скорости.....	241
5.2.3. Совместное измерение запаздывания и доплеровского сдвига частоты.....	242
5.3. Следящие измерители дальности и доплеровского сдвига частоты	245
5.4. Измерение угловых координат.....	252
5.4.1. Многоканальные измерители угловых координат.....	253
5.4.2. Одноканальные измерители угловых координат	255
5.4.3. Дискриминаторные методы измерения угловых координат.....	257
5.5. Точность измерения параметров	261
Контрольные вопросы	267
6. ОСНОВЫ ВТОРИЧНОЙ ОБРАБОТКИ РАДИОЛОКАЦИОННОЙ ИНФОРМАЦИИ.....	269
6.1. Общие сведения	269
6.1.1. Основные понятия и история вопроса	269
6.1.2. Основные операции вторичной обработки.....	274
6.2. Модели целевой и помеховой обстановки	280
6.2.1. Модели движения целей.....	280
6.2.2. Модели отсчетов	287
6.3. Оценка траекторных параметров по фиксированной выборке. Экстраполяция траекторных параметров	292
6.4. Рекуррентная оценка траекторных параметров	301
6.4.1. Основные соотношения калмановской фильтрации.....	301
6.4.2. Стационарный режим калмановского фильтра.....	306
6.4.3. Состоятельность калмановского фильтра.....	307
6.4.4. $(\alpha-\beta)$ -фильтры.....	311
6.4.5. Расширенный фильтр Калмана	313
6.5. Селекция отсчетов	316
6.5.1. Селекция отсчетов методом стробирования.....	320
6.5.2. Алгоритм селекции отсчетов по минимальному отклонению от центра строба	323
6.5.3. Алгоритмы сопоставления и привязки отсчетов к траекториям в многоцелевой ситуации.....	325
6.6. Обнаружение траекторий.....	327
6.6.1. Общие положения	327
6.6.2. Вероятность ложного обнаружения траектории.....	330
6.6.3. Вероятность правильного обнаружения траектории	334
6.7. Завязка траекторий	336
Контрольные вопросы	339

7. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В РАДИОЛОКАЦИОННЫХ СИСТЕМАХ	341
7.1. Принципы построения радиолокационных систем	341
7.1.1. Задачи и условия функционирования радиолокационных систем	341
7.1.2. Принципы получения радиолокационной информации и построения радиолокационных систем	344
7.2. Классификация радиолокационных систем	349
7.3. Тактико-технические характеристики радиолокационных систем	351
7.3.1. Тактические характеристики РЛС	351
7.3.2. Технические характеристики РЛС	353
7.4. Обобщенная структурная схема радиолокационных систем	357
7.4.1. Основные системы для получения радиолокационной информации	357
7.4.2. Дополнительные системы	361
7.5. Формирование отраженного радиолокационного сигнала	363
7.5.1. Вторичное излучение электромагнитных волн. Эффективная площадь рассеяния целей	363
7.5.2. Эффективная площадь рассеяния при различных соотношениях размеров цели и длины волны	365
7.5.3. Характер вторичного излучения и эффективная площадь рассеяния радиолокационных целей	371
7.5.4. Формирование отраженного радиолокационного сигнала для различных моделей целей	375
7.6. Особенности обработки радиолокационных сигналов на фоне стационарного аддитивного гауссовского белого шума	382
7.6.1. Общие сведения о современных методах обработки	382
7.6.2. Обработка одиночных и пачечных импульсных сигналов без внутримпульсной модуляции	384
7.6.3. Обработка частотно-модулированных радиоимпульсов	387
7.6.4. Обработка фазоманипулированных радиоимпульсов	394
7.7. Основные виды помех активной радиолокации	397
7.7.1. Естественные и взаимные маскирующие активные помехи и принципы защиты от них	398
7.7.2. Искусственные маскирующие активные помехи, особенности их воздействия и способы создания	400
7.7.3. Уравнение радиолокации, дальность действия и зоны видимости РЛС при воздействии маскирующих стационарных активных помех	402
7.7.4. Пассивные маскирующие помехи и способы их создания	405
7.8. Методы защиты от маскирующих активных помех	407
7.8.1. Основные направления защиты РЛС от маскирующих активных помех	408
7.8.2. Методы некогерентной и когерентной компенсации помех	410
7.8.3. Практические схемы автокомпенсаторов	413
7.9. Методы защиты от пассивных маскирующих помех	422

7.9.1. Основные различия сигналов целей и пассивных маскирующих помех	422
7.9.2. Оптимальное обнаружение сигнала на фоне пассивной помехи в виде стационарного небелого шума	426
7.9.3. Череспериодное вычитание — способ создания гребенчатых фильтров подавления	431
7.10. Методы стабилизации уровня ложных тревог	436
7.11. Основные характеристики некоторых современных РЛС	440
7.11.1. Радиолокационные станции управления воздушным движением ..	440
7.11.2. Радиолокационные станции обнаружения, наведения и целеуказания на средних и больших высотах	441
7.11.3. Радиолокационные станции обнаружения маловысотных целей	442
7.11.4. Радиолокационные станции наведения зенитных управляемых ракет	443
7.11.5. Радиолокационные станции и комплексы разведки на поле боя	444
7.11.6. Радиолокационные станции подповерхностного зондирования	445
7.11.7. Радиолокационные станции противоракетной обороны	445
7.11.8. Корабельные РЛС	446
7.11.9. Авиационные (самолетные) РЛС	446
Контрольные вопросы	447
8. СПУТНИКОВЫЕ РАДИОНАВИГАЦИОННЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ НА ИХ ОСНОВЕ	449
8.1. Введение	449
8.2. Основные принципы построения и функционирования СРНС	455
8.2.1. Обобщенная структурная схема СРНС	455
8.2.2. Подсистема НКА	457
8.2.3. Контрольно-измерительный комплекс	461
8.3. Структура сигналов и сообщений СРНС	464
8.3.1. Структура навигационных сигналов и навигационных сообщений ГЛОНАСС	464
8.3.2. Структура навигационных сигналов и навигационных сообщений GPS	469
8.4. Методы измерения навигационных параметров	470
8.4.1. Основные понятия	470
8.4.2. Методы определения координат по сигналам НКА	471
8.5. Навигационная аппаратура потребителя	477
8.5.1. Обобщенная структурная схема АП	477
8.5.2. Структура информационного обмена между элементами АП	479
8.5.3. Принципы и устройства первичной обработки навигационной информации	481
8.5.4. Поиск сигналов по задержке и частоте	482
8.5.5. Фильтрация радионавигационных параметров и дешифрация навигационной информации	484
8.5.6. Вторичная обработка навигационной информации	486

8.6. Точность навигационно-временных определений в СРНС	490
8.6.1. Погрешности, вносимые на НКА или КИК СРНС.....	490
8.6.2. Погрешности, вносимые на трассе распространения сигнала от НКА до АП	491
8.6.3. Погрешности аппаратуры потребителей	493
8.6.4. Геометрический фактор.....	494
8.7. Дифференциальная коррекция и относительные измерения в СРНС.....	496
8.7.1. Принцип дифференциальной коррекции	497
8.7.2. Параметры корректирующей информации.....	498
8.7.3. Прямой и инверсный ДМ	499
8.7.4. Методы относительных измерений.....	500
8.8. Угломерная навигационная аппаратура	503
8.8.1. Параметры угловой ориентации объектов.....	504
8.8.2. Принцип определения угловой ориентации объектов по сигналам СРНС.....	505
8.8.3. Радиоинтерферометрический метод измерения угловых координат.....	506
8.8.4. Методы разрешения фазовой неоднозначности при интерферометрических измерениях в спутниковой навигации	509
8.8.5. Факторы, ограничивающие точность угломерной АП СРНС.....	510
8.8.6. Особенности реализации угломерной АП	512
8.9. Информационные технологии на основе СРНС	516
8.9.1. Функциональные дополнения СРНС	517
8.9.2. Транспортные информационно-управляющие системы, использующие сигналы СРНС.....	521
8.9.3. Вторичные эталоны времени и частоты на основе АП СРНС	524
8.9.4. Использование СРНС в геодезии и для мониторинга деформации земной поверхности.....	525
8.9.5. Комплексируемые системы навигации	526
Контрольные вопросы	527
9. РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ..	529
9.1. Структурная схема и основные характеристики цифровых радиотехнических систем передачи информации	529
9.2. Каналы связи	533
9.3. Модели каналов связи	537
9.3.1. Физическая модель непрерывного канала связи	537
9.3.2. Математическая модель непрерывного канала связи	546
9.3.3. Математические модели дискретных каналов связи	548
9.4. Передача и прием дискретных сообщений.....	551
9.4.1. Модуляция и демодуляция. Модемы	551
9.4.2. Передача и прием дискретных сообщений в каналах с замираниями.....	557
9.4.3. Передача и прием дискретных сообщений в каналах с небелым шумом	564

9.5. Помехоустойчивое кодирование и декодирование	571
9.5.1. Принципы построения кодеров	571
9.5.2. Линейные блочные коды	576
9.5.3. Непрерывные (сверточные) коды	589
9.5.4. Сигнально-кодовые конструкции. Прием сигналов в целом	598
9.6. Многоканальные и многоадресные системы	601
9.6.1. Принципы многостанционного доступа	601
9.6.2. Системы с временным разделением	604
9.6.3. Системы с частотным разделением	607
9.6.4. Асинхронные адресные системы	608
9.7. Синхронизация в системах передачи дискретной информации	616
9.7.1. Принципы построения и основные характеристики систем синхронизации	616
9.7.2. Фазовая синхронизация модемов	621
9.7.3. Тактовая синхронизация модемов	628
9.7.4. Цикловая и кадровая синхронизация	631
9.7.5. Синхронизация модемов с широкополосными сигналами	637
9.8. Перспективные системы передачи информации	641
9.8.1. Системы наземной подвижной связи	641
9.8.2. Системы подвижной спутниковой связи	652
Контрольные вопросы	659
ЗАКЛЮЧЕНИЕ	661
СПИСОК ЛИТЕРАТУРЫ	663

ПРЕДИСЛОВИЕ

Двадцатое столетие было веком зарождения и бурного развития радиотехнических систем (РТС), без которых невозможно представить повседневную деятельность и развитие человеческого общества. Радиотехнические системы широко используют практически во всех сферах государственного управления, в промышленности, на транспорте и в связи, в сельском хозяйстве, в сфере образования, науки, культуры и других областях.

Изучение современных информационных технологий в РТС является важным этапом вузовской подготовки радиоинженера. Однако в последнее время такая подготовка встречает определенные трудности, основными из которых являются следующие.

Количество радиотехнических систем, различных по виду и назначению, непрерывно растет. Например, для передачи информации используют системы тропосферной, радиорелейной, спутниковой и сотовой связи, системы радиовещания и телевидения, радиотелеметрические системы, системы передачи команд и др. Системы извлечения информации включают: радиолокационные и навигационные, дистанционного зондирования окружающей среды, разведки ископаемых и состояния поверхности Земли, радиотехнической разведки и др. Непосредственное изучение по отдельности каждой системы не всегда оправдано.

Информационные технологии, применяемые в РТС, интенсивно развиваются, особенно в течение последних 10—15 лет. Широкое применение цифровых методов формирования и обработки сигналов, использование интегральной и функциональной электроники, гибридных интегральных схем, твердотельных СВЧ-устройств позволяют существенно расширить возможности РТС, увеличить объем перерабатываемой и используемой информации и многообразие решаемых задач. В последнее время разработаны и используются адаптивные РТС, которые могут приспосабливаться к внешней помеховой и целевой обстановке, к условиям распространения радиоволн и т. д. Все это не нашло должного отражения в учебниках и учебных пособиях для технических вузов.

Работ, посвященных применению информационных технологий в современных и перспективных РТС, достаточно много. Однако большинство из них опубликовано в журнальных и ведомственных изданиях. В силу ог-

раниченного тиража они мало доступны студенту вуза или начинающему инженеру.

Создание единого учебника или учебного пособия, в котором были бы охвачены все стороны РТС, затруднительно. Однако основные информационные технологии для большого класса систем являются общими. Настоящее учебное пособие посвящено изложению этих информационных технологий. Оно предназначено в какой-то мере восполнить перечисленные пробелы.

В основу книги положены материалы лекций, которые авторы читают в течение многих лет студентам МГТУ им. Н.Э. Баумана; ранее опубликованные учебные пособия и оригинальные работы авторов; результаты научно-исследовательских работ в области РТС различного назначения, выполненных с участием авторов; материалы научно-технических семинаров секции «Информационные технологии в радиолокации» Научного совета «Новые информационные технологии» Отделения информатики, вычислительной техники и автоматизации РАН, проводимых в МГТУ им. Н.Э. Баумана под руководством проф. И.Б. Федорова; работы отечественных и зарубежных специалистов, приведенные в списке литературы.

При написании глав 4 и 7 с любезного разрешения Я.Д. Ширмана использованы материалы справочника «Радиоэлектронные системы: основы построения и теория» / Я.Д. Ширман, Ю.П. Лосев, Н.Н. Минервин и др.; Под ред. Я.Д. Ширмана. – М.: ЗАО «Маквис», 1998.

Пособие соответствует программам учебных дисциплин: «Статистическая теория радиотехнических систем» (направление подготовки дипломированных специалистов 654200 «Радиотехника», специальность 20.07.00 «Радиотехника»), «Радиолокационные системы», «Радионавигационные системы», «Радиосистемы передачи информации» (направление подготовки дипломированных специалистов 654200 «Радиотехника», специальность 20.16.00 «Радиоэлектронные системы»). Оно может быть использовано при изучении курсов специализации по системам радиопреуправления, вторичной и третичной обработки радиолокационных сигналов, средствам радиоэлектронной защиты и др.

Учебное пособие состоит из девяти глав. Глава 1 является вводной. В ней обсуждается роль радиотехнических систем в современном обществе и тенденции развития радиотехнических систем, приводится классификация РТС, рассматриваются тактико-технические характеристики, энергетические соотношения в таких системах.

В главах 2—6 изложены вопросы статистической теории радиотехнических систем: основы теории статистических решений, обнаружение, различение и разрешение сигналов, а также измерение параметров сигналов, в том числе изменяющихся за время наблюдения.

Последние три главы (7—9) посвящены современным информационным технологиям, которые применяются в трех важнейших видах РТС: ра-

диолокационных системах, спутниковых радионавигационных системах и системах передачи информации.

Главы 1—3 написаны А.И. Сениным, А.А. Кузнецовым и И.Б. Федоровым, глава 4 — А.И. Николаевым, глава 5 — В.А. Родзивиловым, А.И. Николаевым и Ю.М. Егоровым, глава 6 — И.Б. Федоровым и Г.П. Слукиным, глава 7 — А.И. Николаевым, глава 8 — И.Б. Власовым и В.Б. Пудловским, глава 9 — В.А. Васиным, В.В. Калмыковым, Ю.Н. Себекиным, А.И. Сениным и И.Б. Федоровым. Общее редактирование выполнено И.Б. Федоровым.

Авторы выражают глубокую признательность рецензентам проф. Ю.Г. Сосулину и коллективу кафедры «Радиотехнические системы» Спб ГЭТУ (зав. кафедрой проф. В.М. Кутузов) за полезные замечания и многочисленные советы, которые были учтены авторами.

Будем признательны читателям за все замечания по содержанию книги, которые следует направлять по адресу: 105005, Москва, 2-я Бауманская ул., 5. МГТУ им. Н.Э. Баумана. Издательство МГТУ им. Н.Э. Баумана.

Авторы

СПИСОК ОСНОВНЫХ СОКРАЩЕНИЙ

ААС	— асинхронные адресные системы
АМ	— амплитудная модуляция (манипуляция)
АП	— аппаратура потребителя
АПВ	— апостериорная плотность вероятности
АС	— абонентская станция
АЦП	— аналого-цифровой преобразователь
БС	— базовая станция
ВКФ	— взаимная корреляционная функция
ВО	— вторичная обработка
ГЦСК	— геоцентрическая система координат
ДК	— дифференциальная коррекция
ДМ	— дифференциальный метод
ДН	— диаграмма направленности
ЗС	— земная станция
ИКМ	— импульсно-кодовая модуляция
КИК	— контрольно-измерительный комплекс
КФ	— корреляционная функция
ЛРП	— линейная рекуррентная последовательность
ЛЧМ	— линейная частотная модуляция
МДВР	— многостанционный доступ с временным разделением
МДКР	— многостанционный доступ с кодовым разделением
МДТР	— многостанционный доступ с пространственным разделением
МДЧР	— многостанционный доступ с частотным разделением
МСД	— многостанционный доступ
НВО	— навигационно-временные определения
НИ	— навигационная информация
НКА	— навигационный космический аппарат
НКС	— непрерывный канал связи
НП	— навигационный параметр
НС	— навигационное сообщение
ОП	— отношение правдоподобия
ОФ	— обесляющий фильтр
ОФМ	— относительная фазовая модуляция (манипуляция)
ПРЧ	— перестройка рабочей частоты

ПСП	— псевдослучайная последовательность
РЛИ	— радиолокационная информация
РЛС	— радиолокационная система (станция)
РНП	— радионавигационный параметр
РПН	— радиолокатор подсвета и наведения
РТР	— ретранслятор
РТС	— радиотехническая система
РЭС	— радиоэлектронная система
СВЧ	— сверхвысокочастотный (диапазон)
СДЦ	— селекция движущихся целей
СКО	— среднеквадратическое отклонение
СПИ	— система передачи информации
СРНС	— спутниковая радионавигационная система
ССПС	— сотовая система подвижной связи
СФ	— согласованный фильтр
ТЦСК	— топоцентрическая система координат
ФАП	— фазовая автоподстройка
ФАПЧ	— фазовая автоподстройка частоты
ФАР	— фазированная антенная решетка
ФМ	— фазовая модуляция (манипуляция)
ЦАП	— цифро-аналоговое преобразование
ЦС	— центральная станция
ЧАП	— частотная автоподстройка
ЧМ	— частотная модуляция (манипуляция)
ЧПВ	— череспериодное вычитание
ШВ	— шкала времени
ЭИ	— эфемеридная информация
ЭПР	— эффективная площадь рассеяния
GPS	— (Global Positioning System) — глобальная система местопределения
GSM	— (Global System for Mobile Communication) — глобальная система мобильной связи

СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ

B	— база сигнала
c	— скорость света
D	— вероятность правильного обнаружения
D_x	— дисперсия случайной величины x
d	— кодовое расстояние
E	— энергия сигнала
F_c	— ширина спектра сигнала
F_D	— доплеровская частота
$F_{л\tau}$	— вероятность ложной тревоги
$F(\cdot)$	— функция распределения
f	— частота
Δf	— полоса пропускания
G	— коэффициент усиления антенны
$G(f), G(\omega)$	— спектральная плотность мощности процесса
H_0	— гипотеза
H_1	— альтернативная гипотеза
$h^2 = E / N_0$	— отношение энергии сигнала E к односторонней спектральной плотности мощности белого шума N_0
$h(t)$	— импульсная характеристика
h	— высота
$K(j\omega)$	— комплексная частотная характеристика
$l(u)$	— отношение функций правдоподобия
$M\{\cdot\}$	— операция вычисления математического ожидания
$N(f)$	— спектральная плотность мощности помехи
$n(t)$	— помеха
p	— вероятность
$P_{\text{ош}}$	— средняя вероятность ошибки
P_c	— мощность сигнала
$P_{\text{пер}}$	— мощность передатчика
$P_{\text{ш}}$	— мощность шума
q^2	— отношение сигнал—шум
r	— расстояние
R	— скорость передачи информации

$R_n(\tau)$	— корреляционная функция случайного процесса $n(t)$
$r_n(\tau)$	— нормированная корреляционная функция случайного процесса $n(t)$
r_y	— коэффициент взаимной корреляции
$s(t)$	— полезный сигнал
T_n	— период повторения импульсов
T_c	— длительность сигнала
t	— время
$u(t)$	— входной сигнал
v_r	— радиальная скорость
$w(\cdot)$	— плотность распределения вероятностей
$x(t)$	— сигнал
β	— азимут цели
ε	— угол места цели
λ	— длина волны
$\delta(\cdot)$	— дельта-функция
σ_c	— эффективная площадь рассеяния цели
σ	— среднеквадратическое отклонение
σ^2	— дисперсия гауссовского случайного процесса
τ	— временная задержка
ω	— круговая частота
$P(\cdot)$	— функция потерь
P	— плотность потока мощности

1. ОБЩИЕ СВЕДЕНИЯ О РАДИОТЕХНИЧЕСКИХ СИСТЕМАХ

Показана роль радиотехнических систем в современном обществе, дана их классификация по назначению, характеру сообщений, используемым диапазонам частот и модулируемому параметру сигналов. Приведены тактико-технические характеристики радиотехнических систем и энергетические соотношения в таких системах. Рассмотрены тенденции развития радиотехнических систем.

1.1. Роль радиотехнических систем в современном обществе

По мере развития человеческого общества возникают все бóльшие требования к быстрому обмену информацией, ее извлечению, обработке и накоплению. Рост объема производства сопровождается увеличением номенклатуры изделий, углублением специализаций и т. п. При этом потребность в обмене информацией растет пропорционально квадрату коэффициента расширения производства. Необходимость обмена информацией в хозяйственной сфере возрастает ежегодно примерно на 10...15 %. Управление хозяйственной деятельностью в масштабах предприятий, объединений и отраслей, а также государства в целом также невозможно без обмена информацией.

С развитием транспорта, особенно воздушного и морского, возникает необходимость более интенсивного обмена информацией о погоде, заявках, получении сведений о местоположении и движении самолетов и кораблей.

Повышение благосостояния и культурного уровня человечества привело к тому, что потребность в обмене информацией между людьми возрастает быстрее, чем увеличивается численность населения. Решение задачи обеспечения информационных потребностей общества возлагается на системы и аппаратуру передачи, извлечения, обработки и накопления информации. Грандиозные успехи в развитии общества были бы немыслимы без этих технических средств, к которым, в первую очередь, относятся различные радиотехнические системы (РТС).

Радиосвязь снабжает информацией все области хозяйственной деятельности человека и личные потребности людей. Радиовещание и телевидение обеспечивают доставку последних известий, культурный досуг, образование и многое другое вне зависимости от расстояний. Радиолокация и радионавигация используются почти во всех видах транспорта. Без РТС не было бы современной авиации и морского флота.

Радиотехнические системы применяются в различных научных исследованиях, медицине, метрологии, геологии, физике. В настоящее время ни одна экспериментальная наука не обходится без сложных и уникальных радиоэлектронных установок при проведении физических и медико-биологических исследований. Широкое распространение получили РТС исследования космического пространства.

Трудно назвать область человеческой деятельности, которая не была бы связана с радиоэлектроникой. Без радиоэлектроники прогресс общества был бы невозможен.

1.2. Классификация радиотехнических систем

Классификация радиотехнических систем по назначению. Классифицировать РТС можно по различным признакам. Основное назначение РТС состоит в предоставлении информации потребителю, поэтому в качестве основного признака при их классификации обычно используют содержание информации или назначение системы.

Исходя из этого можно указать следующие основные типы систем: передачи информации (СПИ), извлечения информации, радиоуправления, разрушения информации и комбинированные системы [1—4].

Системы передачи информации. Характерной особенностью таких систем является наличие отправителя и получателя информации. На стороне отправителя сообщения преобразуются в радиосигналы, которые затем передаются по линии связи к получателю информации, где из принятых сигналов выделяется сообщение.

Структурная схема простейшей СПИ (рис. 1.1) состоит из источников сообщений (ИС), выдающих первичные сигналы; формирователя группового сигнала (ФГС), объединяющего несколько первичных сигналов; модулятора (М), предназначенного для формирования радиосигнала; радиопередающего устройства (РПУ); линии связи (ЛС), по которой радиосигнал передается к получателю сообщений; радиоприемного устройства (РПрУ), осуществляющего усиление и преобразование принятого сигнала; демодулятора (ДМ), выделяющего групповой сигнал; устройства разделения сигналов (УРС), предназначенного для разделения сигнала

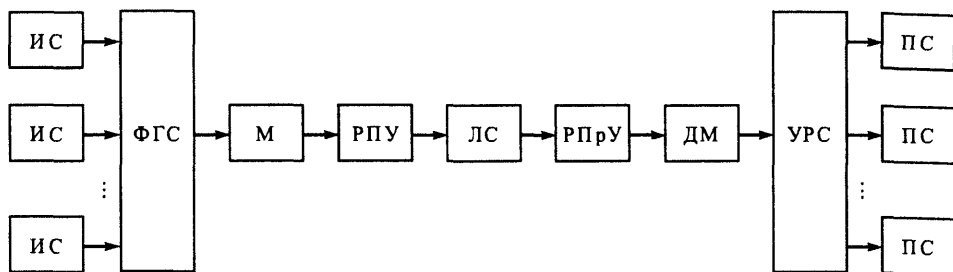


Рис. 1.1. Упрощенная структурная схема радиосистемы передачи информации

лов и выделения сообщений, которые затем поступают к получателям сообщений (ПС).

В линии связи на радиосигналы действуют различного рода помехи: внутрисистемные, обусловленные излучением передатчиков, работающих совместно с данной СПИ в составе сети связи; помехи от других РТС, работающих на близких частотах; атмосферные, космические, организованные помехи, а также помехи, возникающие в аппаратуре.

По назначению различают следующие РТС передачи информации: системы радиосвязи, системы радиосвязи между подвижными объектами, системы радиорелейной связи, спутниковые системы связи, системы радиовещания и телевидения, радиотелеметрические системы, системы передачи команд и др.

Радиотехнические системы извлечения информации. Характерной особенностью систем рассматриваемого класса является то, что полезная информация отображается в радиосигнале либо в процессе его распространения и отражения радиоволн, либо при не зависимом от рассматриваемой системы формировании и излучении радиоволн (естественных излучениях объектов, излучениях радиосредств противника и т. п.).

К системам извлечения информации относятся:

- радиолокационные системы (РЛС), за исключением РЛС с активным ответом;
- радионавигационные системы (РНС);
- системы дистанционного зондирования окружающей среды;
- системы разведки радиотехнических средств противника и др.

На рис. 1.2 представлена упрощенная структурная схема радиолокационной системы, состоящей из радиопередающего устройства (РПУ); антенны (A_1), излучающей радиоволны, направленные на цель (\mathcal{C}); антенны (A_2), принимающей отраженный от цели сигнал; радиоприемного устройства (РПрУ), которое обрабатывает принятый сигнал; измерительно-индикаторного устройства (ИИУ), осуществляющего выделение информации о параметрах движения цели.

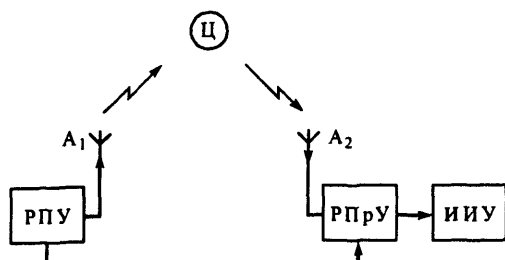


Рис. 1.2. Упрощенная структурная схема радиолокационной системы

ных объектов (ракет, искусственных спутников Земли (ИСЗ), космических аппаратов, самолетов, кораблей и т. п.). Характерной их особенностью является органическая связь радиотехнической части систем с управляемым объектом и зависимость выделяемой информации от выходных эффектов системы.

На рис. 1.3 изображена структурная схема системы самонаведения ракеты на цель (Ц). Она состоит из радиотехнического звена (РЗ), предназначенного для выявления соотношения между пространственным положением и движением цели и ракеты, а также для выдачи команд управления на автопилот (АП), который управляет рулями; динамического звена (ДЗ), отображающего реакцию ракеты на управляющее воздействие; кинематического звена (КЗ), определяющего связь положения и движения ракеты в пространстве с изменениями ее положения и движения относительно цели. Информация о цели I_c и ракете I_p обрабатывается в РЗ. На систему действует помеха $n(t)$.

Системы разрушения информации. Системы этого класса предназначены для противодействия радиотехническим средствам противника. Они создают помехи нормальной работе подавляемой системы излучением мешающего сигнала.

Комбинированные системы. К этим системам относятся, например, РЛС с активным ответом. Они выполняют функции извлечения и передачи информации.

Рассмотренная классификация РТС не является строгой. Реальные системы могут сочетать функции систем различных классов. Так, в систему радиоперехвата входят системы извлечения информации (радиолокаци-

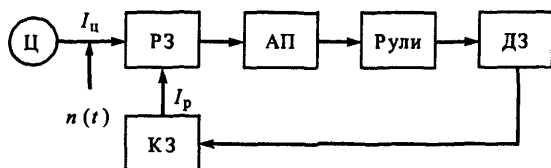


Рис. 1.3. Упрощенная структурная схема системы радиоперехвата

онные и радионавигационные) и передачи информации (радиотелеметрии, передачи команд и др.).

Классификация радиотехнических систем по характеру сообщений. В различных устройствах РТС (преобразователе сообщения в электрический сигнал, передатчике, модуляторе, демодуляторе) на стадиях передачи, извлечения, обработки и накопления информации используют различные виды сигналов.

В зависимости от характера сообщений и применяемых сигналов различают непрерывные, импульсные и цифровые РТС.

В непрерывных системах на основных этапах преобразования сообщения имеют непрерывный характер и отображаются в непрерывные изменения одного или нескольких параметров радиосигнала. К таким системам относят системы радиовещания и телевидения, некоторые типы навигационных систем и ряд других.

В импульсных РТС информация содержится в изменениях параметров импульсных радиосигналов. Типичными представителями таких систем являются импульсные радиолокационные системы, системы передачи информации с импульсной модуляцией и др.

В цифровых РТС сообщения отображаются в кодовые комбинации. Число различных символов, из которых состоят кодовые комбинации, называется *основанием кода*. Различные символы кодовой комбинации передаются соответствующими радиосигналами.

Классификация радиотехнических систем по используемым частотам. Радиотехнические системы могут работать в диапазоне от 3 кГц до 300 ГГц. Несущая частота имеет большое значение для свойств и возможностей РТС. Она существенно влияет на распространение, отражение и рассеяние радиоволн. Поэтому весь диапазон частот разделен на участки, каждый из которых имеет свои особенности (табл. 1.1).

Мириаметровые волны проникают в глубь почвы и воды, огибают Землю, отражаются от ионосферы днем и ночью, огибают, не отражаясь, обычные объекты.

Километровые волны поглощаются в Земле и частично огибают ее, отражаются от ионосферы ночью, огибают, не отражаясь, обычные объекты.

Гектометровые волны поглощаются в Земле, интенсивно отражаются от ионосферы ночью, огибают, не отражаясь, обычные объекты.

Декаметровые волны сильно поглощаются в Земле, избирательно отражаются от ионосферы, слабо отражаются от обычных объектов.

Метровые волны очень сильно поглощаются в Земле, не отражаются от ионосферы, распространяются в пределах прямой видимости, интенсивно отражаются от обычных объектов.

Таблица 1.1

Диапазон радиочастот	Длина волны	Название диапазона радиочастот	Название диапазона радиоволн
3...30 кГц	10...100 км	Очень низкие частоты (ОНЧ)	Мириаметровые волны
30...300 кГц	1...10 км	Низкие частоты (НЧ)	Километровые волны
0,3...3 МГц	100...1000 м	Средние частоты (СЧ)	Гектометровые волны
3...30 МГц	10...100 м	Высокие частоты (ВЧ)	Декаметровые волны
30...300 МГц	1...10 м	Очень высокие частоты (ОВЧ)	Метровые волны
300...3000 МГц	0,1...1 м	Ультравысокие частоты (УВЧ)	Дециметровые волны
3...30 ГГц	1...10 см	Сверхвысокие частоты (СВЧ)	Сантиметровые волны
30...300 ГГц	0,1...10 см	Крайне высокие частоты (КВЧ)	Миллиметровые волны

Дециметровые волны распространяются только в пределах прямой видимости, интенсивно отражаются от обычных объектов. Легко достигается направленность излучения и приема.

Сантиметровые волны распространяются только в пределах прямой видимости, интенсивно отражаются от объектов. Легко достигается высокая направленность излучения и приема.

Миллиметровые волны сильно поглощаются в атмосфере. Легко достигается очень высокая направленность излучения и приема.

Наиболее часто в РТС применяются диапазоны ОВЧ, УВЧ и СВЧ. Радиоволны этих диапазонов частот интенсивно отражаются от объектов, антенны компактны и обеспечивают высокую направленность излучения и приема.

Следует отметить, что использование того или иного диапазона радиочастот для систем различных назначений, а также ширина спектра частот, отводимого системе, регламентированы Международной комиссией распределения радиочастот (МКРР). Эти ограничения влияют на выбор вида радиосигнала и построение РТС и, в конечном счете, на ее характеристики.

Классификация радиотехнических систем по модулируемому параметру радиосигнала. Радиотехнические системы извлечения информации в зависимости от информационного параметра подразделяют на амплитудные, фазовые и частотные. К первым относятся, например, системы определения направления прихода радиоволн с помощью направленных

антенн; ко вторым — фазовые радионавигационные системы; к третьим — доплеровские системы измерения радиальной скорости.

В радиотехнических системах передачи информации сигналы формируются путем изменения тех или иных параметров переносчика информации по закону передаваемых сообщений. Процесс изменения параметров переносчика информации принято называть *модуляцией*, если передаваемые сообщения непрерывные, и *манипуляцией*, если передаваемые сообщения цифровые. В случае, когда переносчиком является гармоническое колебание, модулирующими параметрами могут быть его амплитуда, частота и фаза. Различают непрерывные РТС передачи информации с амплитудной (АМ), частотной (ЧМ) и фазовой (ФМ) модуляциями.

В импульсной РТС передачи информации модулируемыми параметрами могут являться амплитуда импульса, его длительность, частота следования и фаза (положение относительно точки отсчета), число импульсов, а также комбинация импульсов и пауз, определяющих код. Соответственно, различают РТС передачи информации с амплитудно-импульсной (АИМ), широтно-импульсной (ШИМ), частотно-импульсной (ЧИМ), фазоимпульсной (ФИМ), импульсно-кодовой (ИКМ) модуляциями. Возможны и другие виды систем.

В цифровых РТС передачи информации применяются относительно фазовая (ОФМ), частотная (ЧМ), амплитудная (АМ) манипуляции и другие более сложные виды.

Приведенная классификация позволяет выявить особенности РТС и учесть их при проектировании.

1.3. Тактико-технические характеристики радиотехнических систем

Характеристики РТС можно разделить на *тактические*, определяющие назначение и практические возможности системы (зона действия, разрешающая способность, точность, помехоустойчивость, пропускная способность, электромагнитная совместимость и др.), и *технические*, определяющие основные устройства систем (значение и стабильность частоты несущей, вид и параметры модуляции, используемых колебаний, диаграммы направленности антенн, мощность передатчика, чувствительность приемника и др.) [1—4].

Технические параметры характеризуют средства, необходимые для обеспечения заданных тактических параметров. Отклонение любого технического параметра от номинального значения оказывает влияние на те или иные тактические параметры, что может вызвать их выход за пределы установленных допусков (отказ системы).

Рассмотрим основные характеристики РТС.

Зона действия РТС — область пространства, в пределах которой возможны передача, извлечение или разрушение информации, либо область, в пределах которой возможно управление объектом. В сферической системе координат эта область ограничивается минимальной и максимальной дальностями и предельными значениями углов азимута и места. Иногда область действия приходится рассматривать в многомерном фазовом пространстве, координатами которого являются дальность, углы азимута и места, а также скорость и ускорение. Подобная ситуация встречается, в частности, в РТС управления, нормально функционирующих лишь при условии, что скорость и ускорение цели не превышают некоторых заданных значений.

Другим примером могут служить доплеровские РЛС, в которых обнаружение движущихся целей на фоне малоподвижных и неподвижных отражателей осуществляется благодаря различию их скоростей движения. Поэтому четвертой координатой, определяющей зону действия РЛС, является радиальная скорость цели.

Для РТС передачи информации зона действия ограничена только дальностью достоверного приема, которая в настоящее время измеряется сотнями миллионов километров (системы связи с космическими аппаратами).

Для РЛС зона действия ограничивается максимальной дальностью обнаружения с заданной вероятностью, минимальной дальностью, определяемой так называемой «мертвой зоной», и предельными углами азимута и места, определяемыми диаграммой направленности антенны и границами сканирования.

Разрешающая способность — свойство РТС разделять и независимо воспринимать информацию, заключенную в радиосигналах, мало отличающихся друг от друга значениями одного или нескольких параметров, например сдвигами по частоте, задержке и направлению прихода радиоволн.

Точность получаемой информации — степень искажения информации (величина ошибки) при определенных характеристиках сообщений, помеховой обстановки, дальностях, условиях эксплуатации.

Если извлекается информация о некоторой изменяющейся величине, принимающей конечное число значений, то качество системы целесообразно характеризовать вероятностью ошибки. Это имеет место, например, в РТС передачи дискретных сообщений.

В РТС передачи аналоговых сообщений и во многих РТС извлечения информации ошибка имеет непрерывный характер и качество системы целесообразно определять величиной ошибки.

Различают систематические ошибки и случайные. Систематические ошибки обуславливаются известными и закономерными факторами. Поэтому их можно оценить расчетным путем или экспериментально, а затем учесть при проведении измерений. Случайные ошибки возникают из-за дей-

ствия помех: шумов приемника, помех в среде распространения радиоволн, случайного изменения отражающих свойств объектов и т. п. Важными характеристиками случайных ошибок являются плотность распределения вероятности, математическое ожидание, дисперсия, корреляционная функция или спектральная плотность мощности.

Пропускная способность системы — максимальное количество информации, которое может быть передано или извлечено системой за единицу времени. Термин «пропускная способность» обычно применяют к РТС передачи информации и радиолокационным системам.

Быстродействие системы — длительность переходного процесса при подаче на вход системы единичного скачка. Термин «быстродействие системы» применяют к РТС управления. При этом под быстродействием системы понимают способность системы отслеживать быстрые изменения параметров входной величины.

Помехоустойчивость — способность РТС сохранять показатели качества (дальность, точность и т. п.) при воздействии помех. Она зависит от способов кодирования, модуляции, метода приема и т. п.

Для РТС передачи дискретных сообщений помехоустойчивость можно характеризовать вероятностью ошибки при заданном отношении средних мощностей сигнала и помехи в полосе частот, занимаемой сигналами, или требуемым отношением средних мощностей сигнала и помехи на входе приемника системы, при котором обеспечивается заданная вероятность ошибки.

Помехоустойчивость РТС передачи непрерывных сообщений или РТС извлечения непрерывных сообщений удобно оценивать средним квадратом ошибки или отношением средних мощностей сигнала и помехи на входе приемника системы, при котором обеспечивается заданный средний квадрат ошибки.

При сравнительной оценке РТС часто используют обобщенный выигрыш, определяемый по формуле

$$\chi = \frac{(P_c / P_{\text{ш}})_{\text{вых}}}{(P_c / P_{\text{ш}})_{\text{вх}}} \frac{F_{\text{сооб}}}{F_c},$$

где $(P_c / P_{\text{ш}})_{\text{вых}}$, $(P_c / P_{\text{ш}})_{\text{вх}}$ — отношения сигнал—шум на выходе и входе приемника соответственно; $F_{\text{сооб}}$, F_c — ширина соответственно спектра сообщения и сигнала, используемого для передачи сообщения.

Электромагнитная совместимость — способность РТС функционировать совместно с другими РТС.

Надежность аппаратуры — способность РТС выполнять заданные функции, сохраняя свои эксплуатационные показатели в заданных пределах в течение требуемого промежутка времени.

Скрытность действия — способность системы противостоять мерам радиотехнической разведки, направленным на обнаружение сигнала (энергетическая скрытность), определение его структуры (структурная скрытность) и раскрытие смысла передаваемой с помощью сигналов информации (информационная скрытность). Скрытность возрастает, в частности, при увеличении направленности, уменьшении мощности излучения, уменьшении длительности непрерывного функционирования, усложнении структуры сигналов, изменении параметров сигналов.

Масса, габариты, потребляемая мощность, удобства размещения и развертывания — характеристики, особенно важные при размещении РТС на подвижных объектах.

Перспективность — способность РТС в течение длительного времени удовлетворять потребностям общества.

Многие из перечисленных характеристик РТС являются и показателями их качества. В частности, к ним относятся такие, как дальность действия, точность, помехоустойчивость, пропускная способность и др. При рассмотрении характеристик и показателей качества РТС и возможностей их улучшения необходимо учитывать объективные природные ограничения, к которым относятся:

- ограниченный диапазон частот;
- наличие помех;
- особенности распространения радиоволн;
- ограниченный объем размещения аппаратуры РТС (самолеты, ракеты, ИСЗ и др.).

Необходимо также учитывать экономические факторы (улучшение показателей качества всегда связано с увеличением стоимости аппаратуры), а также психофизиологические возможности человека-оператора, обслуживающего ту или иную РТС, его способность к восприятию, обработке и накоплению информации.

1.4. Энергетические соотношения в радиотехнических системах

При разработке РТС возникает необходимость проведения энергетических расчетов, позволяющих гарантировать требуемые показатели качества. Для решения этой задачи учитывают дальность действия, диапазон частот, габариты антенны, предельную чувствительность приемников, мощ-

ность передатчиков, а также помехи, которые всегда присутствуют в радиолинии, собственные тепловые шумы приемника и фидерного тракта и внешние шумы (шумы радиоизлучения Солнца, Земли, космические шумы, индустриальные помехи и т. п.), принятые антенной.

Для достижения требуемых показателей качества РТС необходимо обеспечить определенное отношение $P_{с.вых} / P_{ш.вых}$, где $P_{с.вых}$ и $P_{ш.вых}$ — мощности сигнала и шума на выходе приемника. Соответственно, отношение сигнал—шум на входе приемника должно быть не меньше порогового уровня k_p , называемого *коэффициентом различимости*:

$$P_{с.вх} / P_{ш.вх} \geq k_p. \quad (1.1)$$

Из выражения (1.1) при известной мощности шума на входе приемника можно определить минимальную мощность сигнала на входе, при которой обеспечиваются требуемые показатели качества:

$$P_{с.вх}^{\min} = k_p P_{ш.вх}.$$

Величина $P_{с.вх}^{\min}$ называется *чувствительностью приемника*.

Мощность шума на входе приемника можно определить выражением

$$P_{ш.вх} = N_{ш} \Delta f_{эф} = k T_{\Sigma} \Delta f_{эф},$$

где $N_{ш}$ — спектральная плотность мощности шума; $\Delta f_{эф}$ — эффективная полоса пропускания приемника; $k = 1,38 \cdot 10^{-23}$ Дж/град — постоянная Больцмана; T_{Σ} — суммарная температура шума, К.

Суммарная температура шума определяется формулой

$$T_{\Sigma} = T_{пр} + T_a + T_{\phi},$$

где $T_{пр}$ — эффективная шумовая температура приемника, T_a — шумовая температура внешних шумов и антенны, T_{ϕ} — шумовая температура антенно-фидерного тракта.

Мощность шумов приемника

$$P_{ш.пр} = k T_{пр} \Delta f_{эф}$$

называется *предельной чувствительностью приемника*.

Величина $T_{пр}$ определяется входными каскадами приемника и изменяется в широком диапазоне от 1200...1900 К для кристаллического смесителя до 10...40 К для молекулярного усилителя с охлаждением жидким гелием.

Величина T_a лежит в пределах 40...120 К и зависит не только от внешних шумов, но и от тепловых шумов диэлектрического покрытия антенны.

Величина T_{ϕ} определяется конструкцией антенно-фидерного тракта и зависит от его КПД.

Энергетический расчет РТС зависит от типа радиолинии. Рассмотрим сначала случай, характерный для систем радиовещания, телевидения, радиосвязи, пассивной локации, когда передатчик и приемник разнесены в пространстве на расстояние r . Передача информации ведется из пункта, где находится передатчик, в пункт, где находится приемник.

Пусть G_1 и G_2 — коэффициенты усиления передающей и приемной антенн, η_1 и η_2 — КПД антенно-фидерных устройств передатчика и приемника, $P_{\text{пер}}$ — мощность передатчика. Тогда плотность потока мощности в месте расположения приемника имеет вид

$$\Pi = \frac{P_{\text{пер}} \eta_1 G_1}{4\pi r^2},$$

а мощность сигнала на входе приемника определяется формулой

$$P_{\text{с.вх}} = \Pi S_2 \eta_2 = P_{\text{пер}} \eta_1 \eta_2 G_1 G_2 \left(\frac{\lambda}{4\pi r} \right)^2 = \frac{P_{\text{пер}}}{L_{\text{св}}} \eta_1 \eta_2 G_1 G_2,$$

где $S_2 = G_2 \frac{\lambda^2}{4\pi}$ — эффективная площадь приемной антенны; $L_{\text{св}} = \left(\frac{\lambda}{4\pi r} \right)^{-2}$ — коэффициент, учитывающий потери электромагнитной энергии при распространении в свободном пространстве.

Уравнение

$$P_{\text{с.вх}} = \frac{P_{\text{пер}}}{L_{\text{св}}} \eta_1 \eta_2 G_1 G_2 \quad (1.2)$$

называется *уравнением радиосвязи в свободном пространстве*. Из него следует, что мощность сигнала, поступающего на вход приемника, обратно пропорциональна квадрату расстояния между передатчиком и приемником.

Учитывая поглощение сигнала в атмосфере и земной поверхности, уравнение радиосвязи (1.2) можно записать в виде

$$P_{\text{с.вх}} = \frac{P_{\text{пер}} \eta_1 \eta_2 G_1 G_2}{L_{\text{св}} L_p}, \quad (1.3)$$

где L_p — коэффициент, характеризующий потери энергии при распространении радиоволн в окружающей среде.

Рассмотрим другой тип радиолинии, в которой излученный радиосигнал отражается от объекта и поступает в приемное устройство. В радиосвязи такими объектами являются ионосфера, тропосфера, ионизированные следы метеоров, в активной радиолокации — корабли, самолеты, ракеты и т. д. Радиолинии такого типа применяются и в полуактивной радиолокации. Далее приводится методика энергетического расчета импульсных радиолокационных систем, в которых излучение и прием сигналов ведется одной антенной с коэффициентом усиления G_1 . Пусть $P_{\text{пер}}$ — импульсная мощность передатчика, $\tau_{\text{и}}$ — длительность импульса.

Плотность потока мощности Π в точке нахождения объекта определяется следующим равенством:

$$\Pi = \frac{P_{\text{пер}} \eta_1 G_1}{4\pi r^2}.$$

Мощность отраженного от цели сигнала имеет вид

$$P_{\text{отр}} = \Pi \sigma_{\text{ц}} = \frac{P_{\text{пер}} \eta_1 G_1 \sigma_{\text{ц}}}{4\pi r^2},$$

где $\sigma_{\text{ц}}$ — эффективная площадь рассеяния цели.

Плотность потока мощности в точке приема при предположении, что отраженная электромагнитная энергия рассеивается равномерно во всех направлениях, будем определять по формуле

$$\Pi_1 = \frac{P_{\text{отр}}}{4\pi r^2} = \frac{P_{\text{пер}} \eta_1 G_1 \sigma_{\text{ц}}}{(4\pi r^2)^2}.$$

Мощность сигнала на входе приемника имеет вид

$$P_{\text{с.вх}} = \Pi_1 S_2 = \frac{P_{\text{пер}} \eta_1^2 G_1^2 \sigma_{\text{ц}} S_2}{(4\pi r^2)^2} = \frac{P_{\text{пер}} \eta_1^2 G_1^2 \lambda^2 \sigma_{\text{ц}}}{(4\pi)^3 r^4}.$$

Уравнение

$$P_{\text{с.вх}} = \frac{P_{\text{пер}} \eta_1^2 G_1^2 \lambda^2 \sigma_{\text{ц}}}{(4\pi)^3 r^4} \quad (1.4)$$

называется *уравнением радиолокации в свободном пространстве*. Из него следует, что мощность сигнала, поступающего на вход приемника, обратно пропорциональна четвертой степени расстояния.

С учетом потерь при распространении радиоволн в окружающей среде уравнение радиолокации можно записать в виде

$$P_{\text{с.вх}} = \frac{P_{\text{пер}} \eta_1^2 G_1^2 \lambda^2 \sigma_{\text{ц}}}{(4\pi)^3 r^4 L_p}. \quad (1.5)$$

Приведенные соотношения (1.3) и (1.5) связывают дальность действия соответствующих РТС с их основными техническими характеристиками.

1.5. Тенденции развития радиотехнических систем

Научно-технический прогресс в области РТС проявляется в обновлении технической структуры РТС, в замене устаревших технических средств новыми. Вновь создаваемые РТС должны обладать лучшими показателями качества, более широкими функциональными возможностями и в большей степени удовлетворять требованиям получателя информации.

Основой развития РТС являются как достижения фундаментальных наук, открывающие новые физические принципы функционирования устройств и систем, так и успехи современной электроники.

В развитии РТС выделяют следующие принципиальные направления:

- интеллектуализация РТС на основе вычислительных средств;
- освоение в создаваемой радиоэлектронной технике широкого диапазона радиоволн: от миллиметрового до сверхдлинных;
- переход в современной аппаратуре от отдельных электронных элементов узкого назначения (транзисторов, логических ячеек, ячеек памяти) к функциональным сложным интегральным микросхемам;
- повышение роли устройств обработки информации в РТС;
- расширение областей применения РТС.

Развитие РТС в значительной мере определяется достижениями в области электроники. Для современной электроники характерным является все возрастающая степень интеграции, достигающая в настоящее время нескольких миллионов транзисторов на кристалл. Большой интерес вызывают разработки монолитных интегральных схем сантиметрового и миллиметрового диапазонов на базе биполярных и полевых транзисторов с гетеропереходами, в частности приемопередающих модулей систем с активными фазированными антенными решетками. Продолжает развиваться функциональная электроника: появились акустоэлектронные процессоры, приборы с зарядовой связью и устройства на их основе.

Усложнение функций, связанных с передачей, накоплением и обработкой информации, решается, главным образом, за счет устройств цифро-

вой техники. Цифровая техника используется в устройствах обработки сигналов, системах формирования луча и управления его сканированием в устройствах с фазированными антенными решетками, в системах связи, радиовещания и телевидения.

Наряду с микропроцессорной техникой быстро развиваются цифровые процессоры сигналов (ЦПС) — приборы, где цифровая техника наиболее тесно взаимодействует с аналоговой. Современные ЦПС характеризуются производительностью в несколько десятков миллионов операций в секунду.

Таким образом, РТС идут в своем развитии по пути повышения степени функциональной интеграции, что достигается увеличением в системе числа ячеек, выполняющих логические функции или функции хранения информации.

Повышение степени интеграции позволяет повысить надежность и быстродействие системы, снизить стоимость, перейти на высокоскоростные методы передачи и обработки информации, создать интегрированные многофункциональные комплексы с высоким уровнем искусственного интеллекта, адаптивные к помеховой обстановке.

За последнее десятилетие особенно существенные изменения претерпели РТС передачи информации. Еще совсем недавно в России практически не было междугородних цифровых систем передачи информации. Очень слабо была развита подвижная радиотелефонная связь. Существующие сети, в основном, аналоговые. Плохо удовлетворялся спрос на услуги международной связи. В зачаточном состоянии находились телематические службы «Телетекс», «Телефакс», «Бюрофакс», «Видеотекс» и другие, играющие существенную роль в информатизации общества (телематические службы, согласно определению Международного союза электросвязи, — это службы электросвязи (кроме телефонной, телеграфной и служб передачи данных), которые организуют с целью обмена информацией через сети электросвязи). Отсутствовала общенациональная сеть передачи данных с коммутацией пакетов, сотовые сети подвижной связи и т. п. [5].

Изменения в политической, экономической, культурной и общественной жизни, расширение производственных связей, интеграция в мировое сообщество способствовали бурному развитию систем передачи информации.

Одним из перспективных направлений совершенствования СПИ является создание высокоскоростных сетей для передачи всех видов информации. К приоритетным направлениям относятся такие направления, как создание высокоскоростных цифровых сетей связи, сетей передачи данных с коммутацией пакетов, телематических служб, совершенствование спутниковых систем связи, развитие современных систем подвижной радиосвязи, обеспечивающих как речевой, так и документальный обмен.

Эффективной движущей силой радикальных изменений облика телекоммуникаций являются успехи в развитии подвижной радиосвязи. Хорошо известны возможности такой связи. Системы подвижной связи имеют исключительное значение для больших регионов с низкой плотностью населения и большим числом малых городов и деревень, для труднодоступных районов.

При решении проблем информационного обмена как внутри страны, так и с зарубежными странами важна документальная электросвязь. В последнее время открыты широкие возможности по созданию современных телекоммуникационных технологий в России, резко возрос объем и расширилась номенклатура услуг, предоставляемых документальной электросвязью. Значительное развитие получила международная телеграфная сеть «Телекс». Существенную роль в обмене документальными сообщениями играет система электронной почты РЕЛКОМ, охватывающая практически все регионы России и предоставляющая возможность обмена сообщениями, в том числе с зарубежными корреспондентами. Значительно расширилась сеть факсимильной связи. Создана сеть пунктов пользования факсимильной связью, имеющая выход не только на все регионы страны, но и за рубеж.

К важнейшим средствам организации международной и междугородной телефонно-телеграфной связи, телевидения и радиовещания относится спутниковая связь. Особенно незаменимы спутниковые системы для больших территорий, для районов с малой плотностью населения, суровыми климатическими условиями. В настоящее время создана сеть телефонной связи через спутники с важнейшими районами и городами Дальнего Востока, Сибири, Крайнего Севера. Спутниковая связь позволила распространить многопрограммное телевизионное вещание из Москвы по всей территории России, создать региональное теле- и радиовещание, обеспечить ускоренную телефонизацию удаленных и труднодоступных регионов, организовать дополнительные линии связи и связь с подвижными объектами, ускорить развитие сетей передачи данных и международной спутниковой связи.

Одним из последних достижений в области развития проводной связи и радио являются службы передачи информации, получившие название телематические. Их появление стало результатом взаимопроникновения ЭВМ и новых средств связи. Развитие телематических служб идет в направлении создания новых, более совершенных услуг, что соответствует тенденциям максимального удовлетворения возрастающих потребностей пользователей сетей связи.

Дальнейшее совершенствование получили РЛС и РНС. Современные радиолокационные системы позволяют решать задачи, которые были недоступны единичным радиолокационным средствам. Они обладают высокой

разрешающей способностью по дальности, угловым координатам, радиальной скорости, что обеспечивается применением сверхширокополосных сигналов, когерентных сигналов большой длительности, антенн со сверхузкими диаграммами направленности.

В развитии РЛС наблюдаются тенденции к увеличению числа измеряемых координат (в частности, к созданию трехкоординатных РЛС), многорежимности зондирования и обзора пространства, повышению когерентности сигналов и эффективному ее использованию, автоматизации обработки сигналов и построению трасс целей, сокращению энергозатрат, сочетанию перспективных антенных решеток с более дешевыми антеннами при значительном уменьшении уровня боковых лепестков, существенному повышению надежности и ресурса РЛС, диагностике и ускоренному устранению неисправностей, сокращению количества обслуживающего персонала, ограничению функций персонала применением дистанционного контроля. В области РЛС обнаружения маловысотных целей предпочтение отдается антеннам, поднимаемым на вышки (мачты) высотой 10...20 м.

Большое внимание уделяется РЛС с активным ответом. К ним относятся прежде всего РЛС управления воздушным движением. Большинство РЛС военного назначения имеют также каналы опознавания государственной принадлежности. Для этих РЛС повышенное внимание уделяется защите от помех, а в ряде случаев и от самонаводящихся на излучение снарядов. В связи с широким применением малозаметных воздушных целей существенно повысилась роль РЛС метрового диапазона. Во всех диапазонах проводятся работы по распознаванию классов (и даже типов) воздушных целей.

В последние годы значительно возрос интерес к подповерхностной радиолокации [6]. Работа радиолокационной станции подповерхностного зондирования (георадара) основана на использовании классических принципов радиолокации. Передающая антенна излучает сверхкороткие электромагнитные импульсы, имеющие 1,0—1,5 периода квазигармонического сигнала. Такие импульсы относятся к классу широкополосных сигналов, у которых отношение ширины спектра к частоте его максимума близко к единице. Центральная частота сигнала и длительность определяются необходимой глубиной зондирования и разрешающей способностью георадара. Излучаемый импульс отражается от неоднородностей исследуемой среды и принимается приемной антенной. После обработки сигнала полученная информация отображается на индикаторе.

Подповерхностная радиолокация применяется в геофизике, инженерной геологии, транспортном, промышленном и гражданском строительстве, археологии, космических исследованиях планет Солнечной системы и их спутников, оборонной промышленности и т. д. Методы подповерхностной

радиолокации используются для построения геологических разрезов и профилей дна водоемов, определения положения уровня грунтовых вод и границ распространения полезных ископаемых, измерения толщины пресноводных и морских льдов, выявления местоположения инженерных сетей (металлических и пластиковых труб, кабелей и др.), оценки качества бетонных конструкций (мостов, дамб и плотин), обнаружения утечек из нефтепроводов и захоронения экологически опасных отходов, установления местонахождения археологических объектов и тайников, исследования структуры торфяных месторождений, песчаников, известняков, мерзлых почв. В оборонной промышленности георадары могут использоваться для обнаружения мест установки мин, расположения подземных тоннелей, коммуникаций, складов, а также для выявления подкопов к охраняемым объектам.

Контрольные вопросы

1. Дайте классификацию РТС: а) по назначению; б) по характеру сообщений; в) по используемым частотам; г) по модулируемому параметру радиосигнала.
2. Назовите основные характеристики РТС.
3. Укажите основные направления развития РТС.

2. СИГНАЛЫ И ПОМЕХИ В РАДИОТЕХНИЧЕСКИХ СИСТЕМАХ

Приведены основные сведения о сигналах и помехах в радиотехнических системах. Изложены вопросы дискретизации по времени и квантования по уровню непрерывных сигналов, а также способы их восстановления. Большое внимание уделено построению сложных сигналов и систем сигналов. Рассмотрены вопросы моделирования сигналов и помех.

2.1. Информация, сообщение, сигналы

В различных областях науки и техники, а также в повседневной жизни часто используют понятия «информация», «сообщение», «сигнал». В общем случае под *информацией* понимают совокупность сведений о каком-либо событии, объекте. Для хранения, обработки и преобразования информации применяют условные символы (буквы, математические знаки, рисунки, формы колебаний, слова и др.), позволяющие представить информацию в той или иной форме. Информация, выраженная в определенной форме, называется *сообщением*. Так, в случае телеграфной передачи информация представляется в виде букв и цифр. Соответственно, сообщением является совокупность этих знаков. В телефонных системах это — речь (непрерывное изменение звукового давления), в телевизионных системах — изображение (яркость его элементов). На практике часто информация представляется в двоичной форме, т. е. для ее представления используются только два условных символа, например 1 и 0. Соответственно, сообщением служит последовательность конечного числа двоичных символов.

Одни сообщения (речь, температура, давление) являются функциями времени, другие (например, текст телеграммы) не являются ими. Природа сообщений может быть как электрической, так и не электрической.

Для передачи сообщений используют физические процессы, способные распространяться в той или иной среде. Физический процесс, отображающий сообщение, называется *сигналом*. По своей природе сигналы могут быть электрическими, световыми, звуковыми и т. д. В радиотехнических системах используются электрические сигналы.

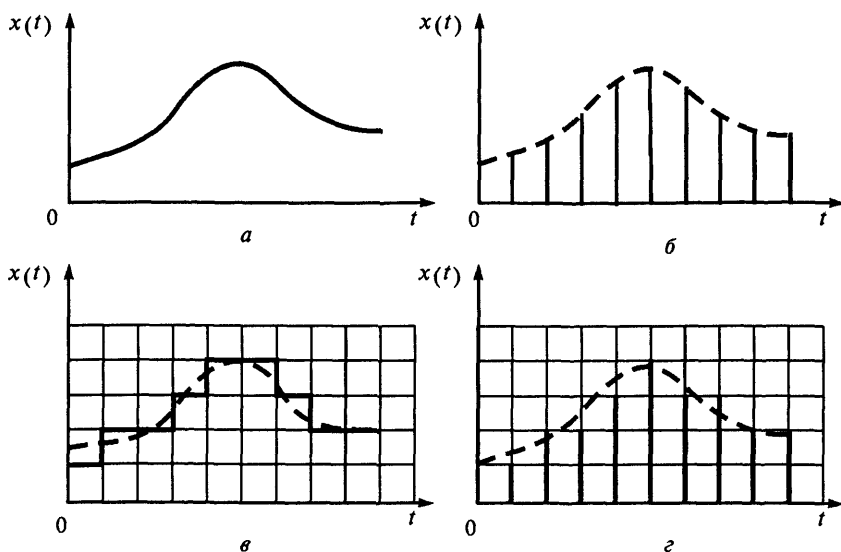


Рис. 2.1. Виды сигналов:

а — сигналы, непрерывные по величине и времени; *б* — сигналы, непрерывные по величине и дискретные по времени; *в* — сигналы, дискретные (квантованные) по величине и непрерывные по времени; *г* — сигналы, дискретные по величине и времени

Любой сигнал является функцией времени $x(t)$. В зависимости от области определения и области возможных значений этой функции различают следующие типы сигналов:

- непрерывные по величине и по времени (рис. 2.1, *а*);
- непрерывные по величине и дискретные по времени (рис. 2.1, *б*);
- дискретные (квантованные) по величине и непрерывные по времени (рис. 2.1, *в*);
- дискретные по величине и по времени (рис. 2.1, *г*).

Сигналы первого типа задаются на конечном или бесконечном временном интервале и могут принимать любые значения в некотором диапазоне. Примером таких сигналов являются сигналы на выходах микрофона, датчиков температуры, давления, положения и т. п. Поскольку такие сигналы являются электрическими моделями физических величин, их называют *аналоговыми*.

Сигналы второго типа, называемые *дискретными*, задаются в определенные дискретные моменты времени и могут принимать любые значения в некотором диапазоне. Их можно получить из непрерывных сигналов, сформулировав последовательность отсчетов в определенные моменты времени. Это преобразование называется *дискретизацией*. Шаг дискретизации T_d

(промежуток времени между двумя соседними отсчетами) может быть как постоянным, так и переменным. Обычно его выбирают исходя из допустимой погрешности при восстановлении непрерывного сигнала по конечному числу его дискретных отсчетов.

Сигналы третьего типа, называемые *квантованными по уровню*, задаются на некотором временном интервале и характеризуются тем, что принимают только вполне определенные дискретные значения. Их можно получить из непрерывных сигналов, применяя к ним операцию квантования по уровню. В результате этой операции непрерывный сигнал заменяется ступенчатой функцией. *Шаг квантования* Δx (расстояние между двумя соседними разрешенными уровнями) может быть как постоянным, так и переменным. Его обычно выбирают из условия обеспечения требуемой точности восстановления непрерывного сигнала из квантованного.

Сигналы четвертого типа задаются в определенные дискретные моменты времени и принимают определенные дискретные значения. Их можно получить из непрерывных сигналов, осуществляя операции дискретизации по времени и квантования по уровню. Такие сигналы можно легко представить в цифровой форме, т. е. в виде чисел с конечным числом разрядов, в связи с чем их обычно называют *цифровыми*.

Функции $x(t)$, описывающие сигналы, могут принимать как вещественные, так и комплексные значения. Соответственно, различают вещественные и комплексные модели сигналов.

Сигналы подразделяют на детерминированные и случайные. Детерминированные сигналы (колебания) — это сигналы, значения которых в любой момент времени известны, т. е. предсказуемы с вероятностью, равной единице. Случайные сигналы — это сигналы, значения которых в любой момент времени невозможно предсказать с вероятностью, равной единице. Все сигналы, несущие информацию, являются случайными, так как полностью известный (детерминированный) сигнал информации не содержит (он может быть создан в месте приема без канала связи).

Детерминированные сигналы применяют при изучении свойств линейных, нелинейных и параметрических цепей. Так, при анализе переходных процессов в линейных цепях часто используют единичный ступенчатый сигнал, синусоидальный сигнал, последовательности импульсов и др.

2.2. Математические модели сигналов и помех

При решении задач анализа и синтеза РТС широко используют математические модели сигналов и помех. Они позволяют отвлечься от физической природы сигналов и описывать только те свойства процессов,

которые являются существенными для решаемой задачи. В современной теории РТС общепринят вероятностный подход, при котором отдельные сообщения рассматриваются как реализации случайного процесса.

Математической моделью дискретных сигналов служит дискретная случайная последовательность $\{X_i\}$ — случайный процесс, областью определения и областью значений которого являются дискретные множества. В дальнейшем будем считать, что случайная величина X_i (элемент последовательности в момент t_i) принимает дискретные значения из множества $\alpha_1, \alpha_2, \dots, \alpha_m$.

Наиболее простой моделью является дискретная случайная последовательность с независимыми элементами (последовательность Бернулли). Для этой последовательности случайные величины X_i независимы и принимают значения из алфавита $\alpha_1, \alpha_2, \dots, \alpha_m$ с вероятностями $p(\alpha_r) = p_r$, $r = 1, 2, \dots, m$. Такая модель описывает сообщения *дискретного источника без памяти*.

Более общая модель — дискретная случайная последовательность с зависимыми элементами. Она описывает сообщения *дискретного источника с памятью*. Модель задается вероятностями

$$\begin{aligned} & p(\alpha_{r_{j+1}}^{(j+1)}, \alpha_{r_{j+2}}^{(j+2)}, \dots, \alpha_{r_{j+N}}^{(j+N)}) = \\ & = p(\alpha_{r_{j+1}}^{(j+1)}) p(\alpha_{r_{j+2}}^{(j+2)} | \alpha_{r_{j+1}}^{(j+1)}) \dots p(\alpha_{r_{j+N}}^{(j+N)} | \alpha_{r_{j+N-1}}^{(j+N-1)}, \dots, \alpha_{r_{j+1}}^{(j+1)}), \end{aligned} \quad (2.1)$$

определяемыми для всех последовательностей $\alpha_{r_{j+1}}^{(j+1)}, \alpha_{r_{j+2}}^{(j+2)}, \dots, \alpha_{r_{j+N}}^{(j+N)}$ длины N и для всех начальных моментов дискретного времени, где $p(\alpha_{r_{j+k}}^{(j+k)} | \alpha_{r_{j+k-1}}^{(j+k-1)}, \dots, \alpha_{r_{j+1}}^{(j+1)})$ — вероятность появления элемента $\alpha_{r_{j+k}}^{(j+k)}$ в момент времени t_{j+k} при условии, что предыдущими элементами были $\alpha_{r_{j+k-1}}^{(j+k-1)}, \dots, \alpha_{r_{j+1}}^{(j+1)}$.

Источник называется *стационарным*, если его статистическое описание (2.1) не зависит от начала отсчета времени t_j .

Математической моделью непрерывных сигналов является непрерывный случайный процесс $X(t)$. Наиболее полно описывается процесс n -мерной функцией распределения

$$F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P\{X(t_1) < x_1, X(t_2) < x_2, \dots, X(t_n) < x_n\} \quad (2.2)$$

или n -мерной плотностью распределения вероятности

$$w_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)}{\partial x_1 \partial x_2 \dots \partial x_n} \quad (2.3)$$

при $n \rightarrow \infty$.

Многомерные функции, заданные выражениями (2.2) и (2.3), определить сложно, а зачастую и невозможно. В то же время для решения многих практических задач, связанных с передачей сообщений, не требуется знания многомерных законов распределения. Поэтому в качестве моделей сообщений обычно используют случайные процессы, задаваемые одномерным и двумерным законами распределения, а во многих случаях — более простыми характеристиками — моментными функциями.

Реальные сообщения, как правило, являются нестационарными. Соответственно, их модели — нестационарные случайные процессы. Чаще всего нестационарные модели допускают квазистационарную трактовку: их можно считать практически стационарными на промежутках времени небольшой длительности. Переход к стационарной модели обусловлен тем, что решение задач с учетом нестационарности сообщений весьма затруднительно и требует сложного математического аппарата.

На практике в качестве стационарных моделей сообщений и помех часто используют *гауссовский* случайный процесс [7, 8]. Гауссовская модель достаточно хорошо описывает речевые и телевизионные сообщения, принимаемые радиолокационные сигналы, некоторые типы телеметрируемых процессов, а также шумы в каналах связи.

Среди моментных функций наибольшее применение получили:

— *математическое ожидание* случайного процесса

$$m_x(t) = \mathbf{M}\{X(t)\} = \int_{-\infty}^{\infty} x w(x; t) dx; \quad (2.4)$$

— *дисперсия* случайного процесса

$$D_x(t) = \mathbf{M}\{[X(t) - m_x(t)]^2\} = \int_{-\infty}^{\infty} [x - m_x(t)]^2 w(x; t) dx; \quad (2.5)$$

— *корреляционная функция* случайного процесса

$$\begin{aligned} R_x(t_1, t_2) &= \mathbf{M}\{[X(t_1) - m_x(t_1)][X(t_2) - m_x(t_2)]\} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1 - m_x(t_1)][x_2 - m_x(t_2)] w_2(x_1, x_2; t_1, t_2) dx_1 dx_2; \end{aligned} \quad (2.6)$$

— ковариационная функция случайного процесса

$$K_x(t_1, t_2) = \mathbf{M}\{X(t_1)X(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 w_2(x_1, x_2; t_1, t_2) dx_1 dx_2,$$

где $\mathbf{M}\{z\}$ — математическое ожидание величины z .

Для стационарных случайных процессов

$$m_x(t) = m_x = \text{const}, \quad D_x(t) = D_x = \text{const},$$

$$R_x(t_1, t_2) = R_x(t_2 - t_1) = R_x(\tau), \quad K_x(t_1, t_2) = K_x(t_1 - t_2) = K_x(\tau).$$

Иногда модель задается *спектральной плотностью мощности*, которая для стационарного центрированного процесса определяется следующим образом:

$$G_x(\omega) = \int_{-\infty}^{\infty} R_x(\tau) \exp(-j\omega\tau) d\tau. \quad (2.7)$$

В качестве моделей сообщений, сигналов и помех часто используют эргодический случайный процесс. Для него все характеристики, найденные статистическим усреднением (см., в частности, (2.4)—(2.6)), совпадают с характеристиками, найденными по его одной реализации $x(t)$ усреднением по времени. Так, для эргодического процесса имеем

$$m_x = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt,$$

$$D_x = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [x(t) - m_x]^2 dt, \quad (2.8)$$

$$R_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [x(t) - m_x][x(t + \tau) - m_x] dt,$$

$$K_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t + \tau) dt.$$

В ряде случаев достаточными для решения задач характеристиками непрерывных сигналов являются полоса частот Δf_c , средняя мощность P_c , пик-фактор K_n и динамический диапазон D_c .

Полоса частот сигнала определяется следующим образом:

$$F_c = F_B - F_H,$$

где $F_{\text{в}}$ и $F_{\text{н}}$ — верхняя и нижняя частоты спектра сигнала.

Средняя мощность сигнала находится усреднением мгновенной мощности $p(t) = x^2(t)$ за достаточно большой промежуток времени:

$$P_{\text{с}} = \frac{1}{T} \int_0^T p(t) dt.$$

Пик-фактор $K_{\text{п}}$ сигнала — это отношение его максимальной мгновенной мощности к средней:

$$K_{\text{п}} = P_{\text{max}} / P_{\text{с}}. \quad (2.9)$$

Часто пик-фактор выражается в децибелах:

$$K_{\text{п}} = 10 \lg(P_{\text{max}} / P_{\text{с}}). \quad (2.10)$$

Динамическим диапазоном называется отношение максимальной мгновенной мощности сообщения к минимальной мгновенной мощности, выраженное в децибелах:

$$D_{\text{с}} = 10 \lg(P_{\text{max}} / P_{\text{min}}). \quad (2.11)$$

Например, для телефонного речевого сообщения установлены верхняя $F_{\text{в}} = 3400$ Гц и нижняя $F_{\text{н}} = 300$ Гц частоты спектра, $F_{\text{с}} = F_{\text{в}} - F_{\text{н}} = 3100$ Гц, $K_{\text{п}} = 13 \dots 17$ дБ, $D_{\text{с}} = 35 \dots 40$ дБ [9].

При выборе модели необходимо учитывать содержание решаемой задачи, особенности математического аппарата, соответствующего данной модели, и ряд других факторов. Как правило, рациональной окажется наиболее простая модель, с помощью которой можно решить поставленную задачу с требуемой точностью.

2.3. Векторное представление сигналов

В современной теории радиотехнических систем для описания, анализа и преобразования сигналов широко используется геометрическое представление, при котором сигналы рассматриваются как элементы некоторого пространства, свойства сигналов — как свойства пространства, преобразование сигналов — как отображение одного пространства в другое.

Введем понятие пространства сигналов. Пусть S — множество сигналов, обладающих некоторым общим свойством. Элементы этого множества отличаются друг от друга теми или другими параметрами сигнала (амплитудой, длительностью, частотой и т. д.). В общем случае различия между

двумя любыми элементами можно охарактеризовать некоторым положительным числом, которое трактуется как количественная мера различия сигналов и называется *расстоянием*. Множество сигналов S с определенным подходящим образом расстоянием между элементами называется *пространством сигналов*.

Для определения расстояния между сигналами используют некоторый функционал d , отображающий каждую пару элементов s_i и s_j множества S на действительную ось. Обычно функционал выбирается таким, чтобы выполнялись требования, являющиеся формализацией свойств, интуитивно связываемых с понятием расстояния:

$$\begin{aligned} d(s_i, s_j) &\geq 0, & d(s_i, s_j) &= d(s_j, s_i), \\ d(s_i, s_k) &\leq d(s_i, s_j) + d(s_j, s_k), \\ d(s_i, s_j) &= 0 \text{ при } s_i = s_j. \end{aligned} \quad (2.12)$$

Функционал, удовлетворяющий условиям (2.12), называется *метрикой*. Множество S с метрикой d называется *метрическим пространством*.

Сигналы можно алгебраически суммировать. При этом результатом сложения является также сигнал. Сигналы можно усиливать или ослаблять. Все эти свойства выполняются, если в качестве пространства сигналов взять так называемое *линейное*, или *векторное*, пространство. Оно удовлетворяет следующим условиям:

1) для любых двух элементов пространства можно определить третий элемент, называемый *суммой* и входящий в данное пространство, такой, что

$$s_i + s_j = s_j + s_i, \quad s_i + (s_j + s_k) = (s_i + s_j) + s_k;$$

2) в пространстве сигналов имеется нулевой элемент 0 такой, что $s_i + 0 = s_i$ для любого s_i ;

3) для любого элемента s_i существует противоположный ему элемент $-s_i$, принадлежащий данному пространству, такой, что $s_i + (-s_i) = 0$;

4) любой элемент пространства можно умножить на любой элемент, принадлежащий скалярному множеству $\{\gamma_i\}$, на котором определены операции сложения и умножения с коммутативными и дистрибутивными свойствами и которое содержит в качестве элементов нуль и единицу, причем $\gamma_i s$ также является элементом пространства сигналов,

$$1 \cdot s = s, \quad (\gamma_i + \gamma_j)s = \gamma_i s + \gamma_j s, \quad \gamma_i(s_k + s_l) = \gamma_i s_k + \gamma_i s_l, \quad \gamma_i(\gamma_j s) = \gamma_i \gamma_j s.$$

Элементы линейного пространства обычно называются *векторами*. Практически все реальные сигналы можно рассматривать как векторы в не-

котором пространстве. Так, если сигналы представлены последовательностями N действительных чисел, то такие сигналы можно представить N -мерными векторами. Любой непрерывный сигнал также можно рассматривать в общем случае как бесконечномерный вектор.

В векторном пространстве между элементами определены простые алгебраические взаимосвязи. В частности, любой сигнал \mathbf{s}_j как вектор может быть представлен в виде комбинации независимых векторов \mathbf{e}_i , $i = 1, 2, \dots, N$, т. е.

$$\mathbf{s}_j = \sum_{i=1}^N c_i \mathbf{e}_i. \quad (2.13)$$

Представление (2.13) является единственным, если векторы \mathbf{e}_i , $i = 1, 2, \dots, N$, образуют *линейно независимую систему*.

Множество всех линейных комбинаций (2.13) образует N -мерное пространство. Множество линейно независимых векторов \mathbf{e}_i , $i = 1, 2, \dots, N$, называется *базисом* этого пространства. Упорядоченную последовательность скаляров c_i из формулы (2.13) обычно интерпретируют как *координаты* вектора \mathbf{s}_j в базисе \mathbf{e}_i , $i = 1, 2, \dots, N$. При этом базис интерпретируют как некоторую систему координат, в общем случае косоугольную.

Любой сигнал можно описать действительной или комплексной функцией, определенной на интервале T_c , который может быть и бесконечным. Множество таких функций образует также линейное пространство. Оно называется *функциональным*. В большинстве случаев функциональное пространство бесконечномерно.

Для количественной характеристики сигналов в линейном пространстве вводят *норму*, определяющую длину вектора \mathbf{s}_j , обычно обозначаемую символом $\|\mathbf{s}_j\|$ и удовлетворяющую следующим условиям:

$$\|\mathbf{s}_j\| \geq 0; \quad \|\mathbf{s}_j\| = 0, \text{ если } \mathbf{s}_j = 0; \quad \|\mathbf{s}_i + \mathbf{s}_j\| \leq \|\mathbf{s}_i\| + \|\mathbf{s}_j\|; \quad \|\gamma \mathbf{s}_j\| = |\gamma| \|\mathbf{s}_j\|,$$

где $|\gamma|$ — модуль скаляра γ .

Для N -мерного линейного пространства действительных или комплексных чисел $\{\mathbf{s}_j = (s_{j1}, s_{j2}, \dots, s_{jN})\}$, $j = 1, 2, \dots$, норма определяется следующим образом:

$$\|\mathbf{s}_j\| = \left[\sum_{k=1}^N |s_{jk}|^2 \right]^{1/2}, \quad (2.14)$$

а для функционального пространства

$$\|s_j\| = \left[\int_0^{T_c} |s_j(t)|^2 dt \right]^{1/2}. \quad (2.15)$$

При таком определении квадрат нормы представляет собой энергию сигнала.

В линейном нормированном пространстве в качестве метрики используется функционал

$$d(s_i, s_j) = \|s_i - s_j\|. \quad (2.16)$$

Для N -мерного линейного пространства действительных или комплексных чисел с учетом выражений (2.14) и (2.16) получаем

$$d(s_i, s_j) = \left[\sum_{k=1}^N |s_{ik} - s_{jk}|^2 \right]^{1/2},$$

а для функционального пространства

$$d(s_i, s_j) = \left[\int_0^{T_c} |s_i(t) - s_j(t)|^2 dt \right]^{1/2}. \quad (2.17)$$

Метрика (2.17) имеет определенный физический смысл: ее квадрат равен энергии разности двух сигналов, она полностью характеризует различие между сигналами (чем больше $d(s_i, s_j)$, тем больше это различие).

В линейном пространстве можно ввести понятие *скалярного произведения* двух элементов, которое часто используют при рассмотрении линейных способов обработки сигналов. Скалярное произведение определяют в случае функционального пространства формулой

$$(s_i, s_j) = \int_0^{T_c} s_i(t) s_j^*(t) dt \quad (2.18)$$

и в случае N -мерного линейного пространства — формулой

$$(s_i, s_j) = \sum_{k=1}^N s_{ik} s_{jk}^*, \quad (2.19)$$

где символом $*$ обозначена комплексно-сопряженная функция или величина.

В функциональном анализе доказывается, что в пространстве со скалярным произведением можно ввести норму, удовлетворяющую соотношению

$$\|s_i\| = (s_i, s_i)^{1/2}, \quad (2.20)$$

и метрику

$$d(s_i, s_j) = \|s_i - s_j\| = (s_i - s_j, s_i - s_j)^{1/2}. \quad (2.21)$$

Итак, пространство со скалярным произведением всегда является нормированным и метрическим. Такое пространство при конечном N называется *евклидовым* (обозначается R_N), а при бесконечном N — *гильбертовым* (обозначается L_2). Структура пространств R_N и L_2 определяет основные свойства сигналов и их взаимосвязи, а введение понятий пространства, нормы, метрики, базиса позволяет формализовать процессы, связанные с передачей, преобразованием и приемом сигналов.

Векторное представление применимо как для детерминированных функций, так и для случайных. Для последних скалярные произведения (2.18) и (2.19), норма (2.20) и расстояние (2.21) — случайные величины. Для случайных процессов также справедливо представление в виде (2.13). При этом коэффициенты c_i являются случайными величинами, а само разложение понимается в смысле среднеквадратической сходимости, т. е.

$$\lim_{N \rightarrow \infty} \mathbf{M} \left\{ \left[s_j - \sum_{i=1}^N c_i \mathbf{e}_i \right]^2 \right\} = 0. \quad (2.22)$$

В общем случае коэффициенты разложения c_i коррелированные. Решение многих задач существенно облегчается, если выбрать ортогональный базис, в котором эти коэффициенты оказываются некоррелированными. Разложение случайного процесса по такому базису называется *каноническим*. Для стационарных процессов каноническое разложение всегда возможно.

2.4. Дискретизация непрерывных сигналов

Дискретизация — это процесс представления непрерывного сигнала $x(t)$, заданного на интервале $(0, T_c)$, совокупностью координат c_1, c_2, \dots, c_N . В общем случае процессы представления и восстановления сигналов описываются выражениями

$$(c_1, c_2, \dots, c_N) = \mathbf{A}[x(t)], \quad (2.23)$$

$$\tilde{x}(t) = \mathbf{A}'[(c_1, c_2, \dots, c_N)], \quad (2.24)$$

где \mathbf{A} — оператор дискретного представления, \mathbf{A}' — оператор восстановления, $\tilde{x}(t)$ — восстановленный сигнал.

Операторы дискретного представления и восстановления бывают линейными и нелинейными. На практике обычно используют линейные операторы, как более простые в реализации.

При линейных процессах представления и восстановления сигналов выражения (2.23) и (2.24) можно записать в виде

$$c_i = \int_0^{T_c} x(t) \psi_i(t) dt, \quad i = 1, 2, \dots, N, \quad (2.25)$$

$$\tilde{x}(t) = \sum_{i=1}^N c_i \varphi_i(t), \quad (2.26)$$

где $\psi_i(t)$ и $\varphi_i(t)$ — *весовые* и *базисные (координатные)* функции соответственно.

В зависимости от системы используемых весовых функций $\psi_i(t)$, $i = 1, 2, \dots, N$, различают дискретное временное, дискретное обобщенное и дискретное разностное представления сигналов.

При *дискретном временном представлении* используется система весовых функций $\psi_i(t) = \delta(t - t_i)$, $i = 1, 2, \dots, N$, где $\delta(t - t_i)$ — дельта-функция. Координаты, как это видно из соотношения (2.25), определяются следующим образом: $c_i = x(t_i)$, $i = 1, 2, \dots, N$, т. е. совпадают с мгновенными значениями (отсчетами) непрерывной функции $x(t)$ в дискретные моменты t_i .

Представление называется *регулярным*, если шаг дискретизации $T_d = t_i - t_{i-1}$ является постоянным. В противном случае оно называется *нерегулярным (адаптивным)*.

Для представления сигналов регулярными отсчетами необходим выбор частоты дискретизации $F_d = 1/T_d$ и базисных функций $\varphi_i(t)$. Особенно важно найти минимальную частоту дискретизации, при которой еще имеется возможность восстановления непрерывного сигнала с заданной погрешностью. При решении этих задач следует учитывать свойства исходных сигналов, способы их восстановления и требуемую точность восстановления.

Для модели сигналов с ограниченным спектром решение указанных задач содержится в теореме Котельникова, которая утверждает, что любую непрерывную функцию со спектром, ограниченным полосой частот $(0 \dots F_{\max})$, можно однозначно определить последовательностью ее мгновенных значений, взятых через интервалы времени $T_d = 1/(2F_{\max})$.

Восстановление непрерывной функции осуществляется с помощью ряда:

$$\tilde{x}(t) = \sum_{i=-\infty}^{\infty} x(iT_d) \frac{\sin(2\pi F_{\max}(t - iT_d))}{2\pi F_{\max}(t - iT_d)}, \quad (2.27)$$

который называется *рядом Котельникова*. Базисными функциями в данном случае служат функции отсчетов

$$\varphi_i(t) = \frac{\sin(2\pi F_{\max}(t - iT_d))}{2\pi F_{\max}(t - iT_d)}, \quad i = \dots, -1, 0, 1, \dots$$

Они образуют ортогональную на бесконечном интервале $-\infty < t < \infty$ систему функций. Любую функцию $\varphi(t)$ можно получить на выходе идеального фильтра нижних частот, подав на его вход сигнал $\delta(t - iT_d)$.

В соответствии с выражением (2.27) восстановление непрерывного сигнала осуществляется подачей на вход идеального фильтра нижних частот с полосой пропускания $0 \dots F_{\max}$ последовательности δ -функций $\delta(t - iT_d)$, $i = \dots, -1, 0, 1, \dots$, умноженных на коэффициенты $x(iT_d)$. Однако ни сигнал в виде δ -функций, ни идеальный фильтр нижних частот физически не реализуемы. Поэтому на практике вместо δ -функций используют короткие импульсы, а вместо идеального фильтра нижних частот — фильтр нижних частот, что, естественно, приводит к погрешности восстановления сигнала.

Теорему Котельникова можно обобщить и на случайные сигналы [9]. При этом она формулируется следующим образом: для случайного процесса с односторонней спектральной плотностью мощности, удовлетворяющей условию $G_x(f) = 0$ при $f > F_{\max}$, ряд

$$\sum_{i=-\infty}^{\infty} X(iT_d) \frac{\sin(2\pi F_{\max}(t - iT_d))}{2\pi F_{\max}(t - iT_d)},$$

где $X(iT_d)$ — случайные величины, представляющие отсчеты случайного процесса, взятые через интервалы времени $T_d = 1/(2F_{\max})$, сходится в среднеквадратическом смысле (см. (2.22)) к процессу $X(t)$.

Теорема Котельникова справедлива и для идеализированных условий, среди которых следует отметить ограниченность спектра по частоте и бесконечное время наблюдения. Все реальные сигналы конечны по времени и имеют неограниченный по частоте спектр. Использование модели с ограниченным спектром и конечное время наблюдения приводит к погрешности при восстановлении непрерывного сигнала.

Тем не менее, теорема Котельникова имеет большое практическое значение. Спектр сигнала так или иначе ограничивается (например, при пе-

редаче непрерывного сигнала спектр $G(f)$ целесообразно ограничить частотой F_{\max} , при которой $G(f) < N(f)$, где $N(f)$ — спектральная плотность мощности шума на выходе канала). В этом случае по теореме Котельникова можно оценить частоту дискретизации, которую обычно определяют по приближенной формуле [10]

$$F_d \approx 2\lambda F_{\max},$$

где λ — некоторый коэффициент, выбираемый из интервала 1,25—2,5.

Ограничение спектра сигнала частотой F_{\max} посредством фильтрации приводит к погрешности восстановления, относительный средний квадрат которой определяется формулой

$$\delta_F^2 = \int_{F_{\max}}^{\infty} G(f) df \bigg/ \int_0^{\infty} G(f) df, \quad (2.28)$$

т. е. равен отношению мощности отфильтрованной части спектра к средней мощности исходного сигнала.

При отсутствии предварительной фильтрации сигнала в процессе его восстановления ошибка дискретизации возрастает. Пусть $S_x(j\omega)$ — спектральная плотность сигнала $x(t)$. Тогда спектральная плотность дискретизированного сигнала $x_d(t)$ [7] имеет вид

$$S_{xd}(j\omega) = \frac{1}{T_d} \sum_{n=-\infty}^{\infty} S_x \left[j \left(\omega - \frac{2\pi n}{T_d} \right) \right], \quad (2.29)$$

т. е. она равна с точностью до несущественного множителя $1/T_d$ сумме бесконечного числа «копий» спектра исходного сигнала (рис. 2.2). Эти копии располагаются на оси частот через равные промежутки $2\pi/T_d$.

При восстановлении сообщения идеальным фильтром нижних частот с полосой пропускания $-\frac{\pi}{T_d} \leq \omega \leq \frac{\pi}{T_d}$ возникает ошибка, относительный средний квадрат которой с учетом выражения (2.29) определяется соотношением

$$\delta_{\text{от.ф.}}^2 = \frac{\int_{|\omega| > \pi/T_d}^{\infty} |S_x(j\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |S_x(j\omega)|^2 d\omega} + \frac{\int_{-\pi/T_d}^{\pi/T_d} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \left| S_x \left[j \left(\omega - \frac{2\pi n}{T_d} \right) \right] \right|^2 d\omega}{\int_{-\infty}^{\infty} |S_x(j\omega)|^2 d\omega}. \quad (2.30)$$

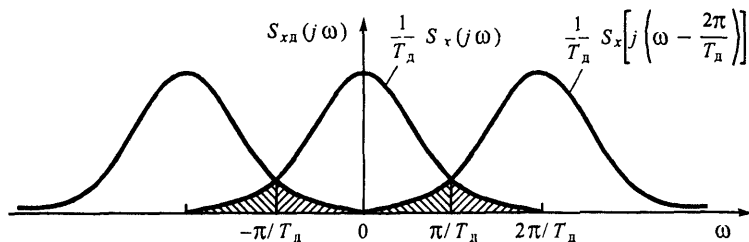


Рис. 2.2. Спектральная плотность дискретизованного сигнала

Первое слагаемое в выражении (2.30) характеризует ошибку, обусловленную тем, что составляющие сигнала $x(t)$ с частотами $|\omega| > \pi/T_d$ не попадают в полосу пропускания фильтра, и совпадает с (2.28). Второе слагаемое в выражении (2.30) характеризует ошибку, обусловленную попаданием в полосу пропускания фильтра составляющих боковых лепестков $S_x[j(\omega - 2\pi n/T_d)]$, $n = \pm 1, \pm 2, \dots$, спектра сигнала $x_d(t)$. Если ограничиться влиянием только двух боковых лепестков при $n = \pm 1$, то нетрудно видеть, что второе слагаемое также совпадает по значению с (2.28). Таким образом,

$$\delta_{от.ф}^2 = 2\delta_F^2, \quad (2.31)$$

и, следовательно, предварительная фильтрация сигнала с целью ограничения его спектра целесообразна.

Заметим, что обеспечить выполнение условия $G_x(f) = 0$ при $f > F_{\max}$ фильтрацией физически невозможно. Сигнал на выходе любого реализуемого фильтра будет содержать спектральные составляющие на частотах $f > F_{\max}$. Поэтому ошибка (2.28) является минимально возможной.

В общем случае восстановление (интерполяция) непрерывного сигнала $x(t)$ по его отсчетам выполняется в соответствии с формулой (2.26). При этом в качестве базисных функций обычно используют алгебраические полиномы. В частности, на практике применяют ступенчатую и линейную интерполяции. Для ступенчатой интерполяции (рис. 2.3, а) используют базисные функции $\varphi_1(t) = 1$, $\varphi_i(t) = 0$, $i = 2, \dots$, при этом $\tilde{x}(t) = x(t_i)$, $t_i \leq t \leq t_i + T_d$. Для линейной интерполяции (рис. 2.3, б) используют базисные функции $\varphi_1(t) = 1 - \tau/T_d$, $\varphi_2(t) = \tau/T_d$, $\varphi_i(t) = 0$, $i = 3, 4, \dots$, при этом $\tilde{x}(t) = x(t_i) + [x(t_{i+1}) - x(t_i)]\tau/T_d$, $t_i \leq t \leq t_i + T_d = t_{i+1}$, $\tau = t - t_i$.

Относительный средний квадрат погрешности восстановления сигнала зависит от нормированной корреляционной функции $r_x(\tau)$ исходного процесса $X(t)$, способа интерполяции и частоты дискретизации. В работе [11]

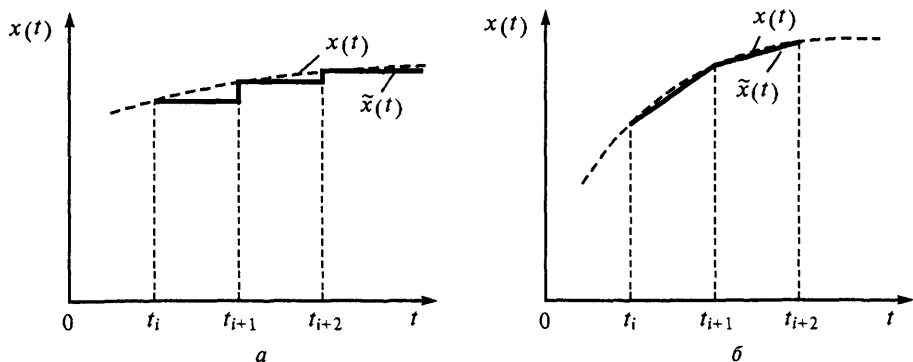


Рис. 2.3. Диаграммы, иллюстрирующие ступенчатую (а) и линейную (б) интерполяции

показано, что для любых стационарных процессов с нулевым математическим ожиданием при ступенчатой интерполяции погрешность восстановления определяется выражением

$$\delta_B^2 = 2 \left[1 - \frac{1}{T_d} \int_{-T_d/2}^{T_d/2} r_x(\tau) d\tau \right], \quad (2.32)$$

при линейной интерполяции — формулой

$$\delta_B^2 = \frac{5}{3} + \frac{1}{3} r_x(T_d) - \frac{4}{T_d} \int_0^{T_d} r_x(\tau) d\tau + \frac{4}{T_d^2} \int_0^{T_d} \tau r_x(\tau) d\tau. \quad (2.33)$$

Если задана погрешность интерполяции, формулы (2.32) и (2.33) используют для нахождения частоты дискретизации. Расчеты показывают, что частота F_d существенно превышает частоту дискретизации по Котельникову. Так, для сигнала с прямоугольной спектральной плотностью мощности, ограниченной частотой F_{\max} , отношение $F_d / (2F_{\max})$ равно $\pi / (6\delta)$ при ступенчатой интерполяции и $\pi / \sqrt[4]{600\delta^2}$ при линейной [11].

При обобщенном дискретном представлении дискретизация сигнала заключается в следующем: интервал T_c разбивается на подынтервалы $T_{a,j}$, $j = 1, 2, \dots$, на которых производится анализ сигнала. В результате обработки сигнала в соответствии с (2.25) в конце каждого j -го подынтервала находятся координаты c_1, c_2, \dots, c_{N_j} . Для регулярных методов представления значения $T_{a,j}$ и N_j не зависят от индекса j , т. е. $T_{a,j} = T_a$, $N_j = N$.

Решая рассматриваемую задачу, важно правильно выбрать длительность подынтервала анализа T_a . При этом необходимо иметь в виду, что с увеличением длительности этого подынтервала растет число координат N , необходимых для представления сигнала. Соответственно, усложняется аппаратура, увеличивается ее объем, масса и стоимость. В связи с этим значения T_a не должны быть слишком большими. На практике для непрерывного сигнала $X(t)$ часто вполне приемлема длительность интервала $T_a = (5 \dots 6) \tau_k$, где τ_k — интервал корреляции процесса $X(t)$ [11].

Другой важной задачей является выбор весовых и координатных функций. Вызывают интерес такие операторы A и A' , которые обеспечивают минимальную погрешность δ^2 при заданном числе координат N или минимальное число координат θ при заданной погрешности δ^2 .

Пусть $X(t)$ — случайный процесс с нулевым математическим ожиданием и непрерывной корреляционной функцией $R(t, t')$. Тогда можно показать [8], что при представлении процесса $X(t)$ рядом (2.26) математическое ожидание интегральной среднеквадратической ошибки

$$M \left\{ \int_0^{T_a} \left[X(t) - \sum_{i=1}^N c_i \varphi_i(t) \right]^2 dt \right\} \quad (2.34)$$

при любом фиксированном N будет минимальным, если весовые функции $\psi_i(t)$, $i = 1, 2, \dots, N$, совпадают с базисными функциями $\varphi_i(t)$, $i = 1, 2, \dots, N$, а базисные функции удовлетворяют однородному интегральному уравнению Фредгольма второго рода:

$$\sigma_i^2 \varphi_i(t) = \frac{1}{T_a} \int_0^{T_a} R_x(t, t') \varphi_i(t') dt', \quad (2.35)$$

где $\varphi_i(t)$ — собственные функции, σ_i^2 — собственные значения ядра $R_x(t, t')$ уравнения.

Собственные функции $\varphi_i(t)$, $i = 1, 2, \dots, N$, являются ортогональными и определяются уравнением с точностью до постоянного множителя, который можно выбрать таким, чтобы функции $\varphi_i(t)$, $i = 1, 2, \dots, N$, были ортонормированными. При этом координаты c_i в разложении (2.26) случайного процесса оказываются некоррелированными, а для гауссовского случайного

процесса — статистически независимыми. Кроме того, при $\mathbf{M}\{c_i\} = 0$ дисперсия D_{c_i} координаты c_i равна σ_i^2 .

Если базисные функции ортонормированные, то математическое ожидание интегральной среднеквадратической ошибки (2.34), найденное на интервале представления T_a , имеет вид

$$\begin{aligned}\bar{\varepsilon}_{T_a}^2 &= \frac{1}{T_a} \mathbf{M} \left\{ \int_0^{T_a} \left[X(t) - \sum_{i=1}^N c_i \varphi_i(t) \right]^2 dt \right\} = \\ &= \frac{1}{T_a} \int_0^{T_a} \mathbf{M} \{ X^2(t) \} dt - \frac{2}{T_a} \int_0^{T_a} \sum_{i=1}^N \mathbf{M} \{ c_i X(t) \} \varphi_i(t) dt + \\ &+ \frac{1}{T_a} \int_0^{T_a} \sum_{j=1}^N \sum_{i=1}^N \mathbf{M} \{ c_i c_j \} \varphi_i(t) \varphi_j(t) dt = D_x - \sum_{i=1}^N D_{c_i}. \quad (2.36)\end{aligned}$$

Выражение (2.36) позволяет находить число координат N , при котором обеспечивается заданная погрешность дискретного представления сигнала.

Разложение случайного процесса с непрерывной корреляционной функцией в ряд (2.26), в котором базисные функции являются собственными функциями уравнения (2.35), называется *разложением Карунена—Лозва*. Хотя это разложение обеспечивает минимальное число координат N при заданной погрешности дискретного представления случайного процесса $\bar{\varepsilon}^2$, его применение ограничено. Это обусловлено следующими причинами: корреляционная функция случайного процесса не всегда оказывается известной, процедура нахождения решения уравнения (2.35) в общем случае неизвестна, техническая реализация устройств разложения сигнала, за исключением случая, когда функции $\varphi_i(t)$ гармонические, сложная. Поэтому на практике в качестве базисных часто используют ортогональные функции, при которых погрешность представления близка к минимальной при использовании сравнительно простой аппаратуры. К ним относятся тригонометрические функции, полиномы Чебышева и Лежандра, функции Уолша и др. [7, 12].

При *дискретном разностном представлении* сигнала в качестве весовых функций $\psi_i(t)$ используют линейные комбинации дельта-функций:

$$\psi_i(t) = \sum_{k=0}^L (-1)^k C_L^k \delta(t - t_i + kT_d), \quad L = 1, 2, \dots, \quad i = 1, 2, \dots, \quad (2.37)$$

где C_L^k — число сочетаний из L по k [11]. Как следует из (2.25), координатами c_i являются конечные разности L -го порядка

$$\Delta^L x(t_i) = \sum_{k=0}^L (-1)^k C_L^k x(t_i - kT_d).$$

В частности, при $L = 1$

$$\psi_i(t) = \delta(t - t_i) - \delta(t - t_{i-1}), \quad c_i = \Delta x(t_i) = x(t_i) - x(t_{i-1}).$$

При адаптивной дискретизации непрерывных сигналов координаты c_i представляют мгновенные значения непрерывного сигнала в некоторых точках отсчета, не равноотстоящих друг от друга (рис. 2.4). На интервалах быстрого изменения значений сигнала отсчеты берутся чаще, а на интервалах медленного изменения — реже. Для представления сигнала стараются использовать как можно меньшее число отсчетов, но достаточное для его восстановления с заданной погрешностью. Отсчеты, позволяющие восстановить непрерывный сигнал на приемной стороне с заданной точностью, называются обычно *существенными*.

Известны различные способы адаптивной дискретизации, отличающиеся алгоритмом формирования существенных отсчетов и видом служебной информации [11]. Простейший алгоритм формирования существенных отсчетов заключается в следующем. Пусть последний существенный отсчет был в момент t_i . Для определения момента следующего отсчета сравнивают текущее значение функции $x(t)$ с $x(t_i)$. Момент t_{i+j} , при котором $|x(t_{i+j}) - x(t_i)| = \epsilon_{\max}$, соответствует очередной существенной выборке, где ϵ_{\max} характеризует максимальную погрешность представления.

При адаптивной дискретизации отсчеты передаются в случайные моменты времени. Поэтому для восстановления непрерывного сообщения по отсчетам приемная сторона должна знать, к каким тактовым моментам относятся принятые отсчеты. В связи с этим на приемную сторону приходится передавать дополнительную служебную информацию. Такой информацией могут быть значения тактовых моментов, соответствующих существенным выборкам. При сравнении различных способов представления сигнала это обстоятельство необходимо учитывать.

Адаптивные способы дискретизации широко применяют при отсутствии априорной информации о корреляционной функции или спектральной плотности мощности непрерывных сообщений.

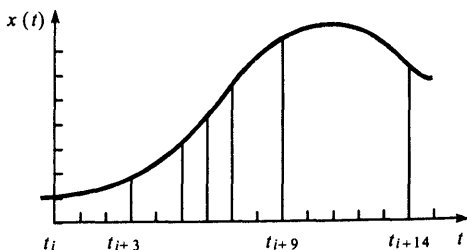


Рис. 2.4. Пример размещения существенных выборок при линейной интерполяции

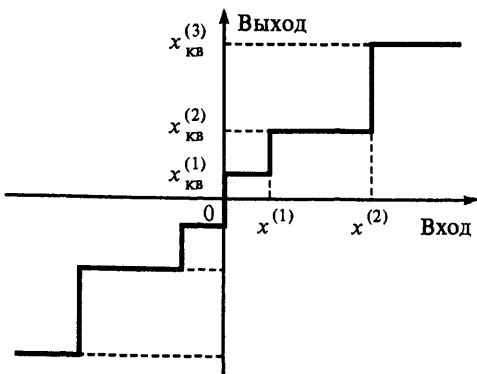
2.5. Преобразование непрерывных сигналов в цифровую форму

Непрерывные сигналы можно преобразовывать в цифровую форму (в последовательность символов некоторого алфавита, например двоичного) с помощью операций дискретизации по времени, квантования по уровню и кодирования. В общем случае квантованию могут подвергаться коэффициенты c_i , полученные в результате обобщенного или разностного дискретного преобразования.

Операция дискретизации по времени была описана в § 2.4. Операция квантования по уровню заключается в замене непрерывного множества значений, которые может принимать сигнал $x(t)$, дискретным множеством заранее определенных значений $x_{\text{кв}}^{(i)}$, $i = 1, 2, \dots, L_{\text{кв}}$, называемых *уровнями квантования*. Такое преобразование выполняет нелинейное устройство (квантователь) с характеристикой, изображенной на рис. 2.5, следующим образом. Диапазон возможных значений сигналов разбивается на $L_{\text{кв}}$ интервалов. При попадании отсчета сигнала в i -й интервал ему присваивается значение $x_{\text{кв}}^{(i)}$.

Различают *равномерное* и *неравномерное* квантования. При равномерном квантовании шаг Δx выбирают постоянным, а уровень $x_{\text{кв}}^{(i)}$ — соответствующим середине i -го интервала квантования. При неравномерном квантовании шаг Δx является переменным.

Замена непрерывного множества возможных значений сигналов дискретным множеством фиксированных значений приводит к погрешности, называемой *шумом квантования*. При равномерном квантовании дисперсия погрешности квантования определяется формулой [8]



$$D_{\text{кв}} = \sum_{i=1}^{L_{\text{кв}}} \int_{x_{\text{кв}}^{(i)} - \Delta x / 2}^{x_{\text{кв}}^{(i)} + \Delta x / 2} (x - x_{\text{кв}}^{(i)})^2 w(x) dx, \quad (2.38)$$

где $w(x)$ — плотность распределения вероятностей мгновенных значений сигнала $x(t)$.

В случае равномерного распределения значений сигнала из соотношения (2.38) находим

Рис. 2.5. Характеристика квантователя

$$D_{\text{кв}} = (\Delta x)^2 / 12. \quad (2.39)$$

Таким образом, для рассматриваемого случая дисперсия погрешности равномерного квантования зависит только от значения шага Δx или при заданном диапазоне изменения значений сигнала — от числа уровней квантования.

Заметим, что при большом количестве уровней квантования $D_{\text{кв}} \approx (\Delta x)^2 / 12$ при любом законе распределения мгновенных значений сигнала. Действительно, в этом случае плотность вероятности $w(x)$ в пределах любого i -го интервала можно считать постоянной и равной $w(x_{\text{кв}}^{(i)})$. Тогда

$$D_{\text{кв}} \approx \sum_{i=1}^{L_{\text{кв}}} \int_{x_{\text{кв}}^{(i)} - \Delta x / 2}^{x_{\text{кв}}^{(i)} + \Delta x / 2} (x - x_{\text{кв}}^{(i)})^2 w(x_{\text{кв}}^{(i)}) dx = \sum_{i=1}^{L_{\text{кв}}} w(x_{\text{кв}}^{(i)}) \frac{(\Delta x)^3}{12}.$$

Учитывая, что $\sum_{i=1}^{L_{\text{кв}}} w(x_{\text{кв}}^{(i)}) \Delta x \approx 1$, получаем $D_{\text{кв}} \approx (\Delta x)^2 / 12$.

Неравномерное квантование, несмотря на то, что оно сложнее в реализации, чем равномерное, довольно часто используют, например, при передаче речевых сигналов. Это объясняется следующими причинами. Одна из них заключается в том, что распределение мгновенных значений речевых сигналов отлично от равномерного; как правило, малые значения гораздо более вероятны, чем большие. Поэтому при равномерном квантовании вероятности попадания сигнала в разные интервалы квантования различны. Очевидно, что погрешность квантования можно уменьшить, если шаг квантования брать меньшим для более вероятных значений сообщения и большим для менее вероятных.

Другая причина заключается в том, что в телефонных системах средние значения речевых сигналов могут различаться на 30 дБ и более. Чтобы сохранить разборчивость речи «тихого» абонента, шаг квантования в области малых значений сигнала должен быть небольшим. В области больших значений сигнала шаг можно увеличить. Таким образом, вновь приходим к неравномерному квантованию.

Неравномерное квантование можно реализовать различными способами, например используя квантователь с соответствующей амплитудной характеристикой (непосредственное неравномерное квантование). При этом длины интервалов и уровни квантования (см. рис. 2.5) обычно выбирают из условия получения минимальной дисперсии погрешности $D_{\text{кв}}$. Такой же

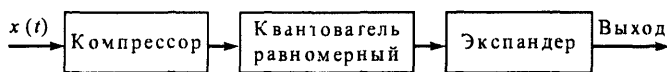


Рис. 2.6. Структурная схема компандирования

эффект можно получить сжатием (компрессированием) динамического диапазона сигнала, применением равномерного квантования и последующим расширением (экспандированием) после восстановления отсчетов на приемной стороне (рис. 2.6). Характеристики компрессора и экспандера должны быть взаимно обратными. Этот метод получил название *квантование с компандированием* сигнала. Характеристику компрессора выбирают из условия обеспечения минимума дисперсии погрешности квантования.

При неравномерном квантовании дисперсия погрешности квантования определяется выражением

$$D_{\text{кв}} = \sum_{i=1}^{L_{\text{кв}}} \int_{x^{(i)}}^{x^{(i+1)}} (x - x_{\text{кв}}^{(i)})^2 w(x) dx, \quad (2.40)$$

где $x^{(i)}$ и $x^{(i+1)}$ — нижняя и верхняя границы i -го интервала квантования, называемые обычно *порогами квантования*.

Пороги и уровни квантования выбирают из условия минимизации дисперсии (2.40). Для их нахождения продифференцируем выражение (2.40) по переменным $x_{\text{кв}}^{(i)}$ и $x^{(i)}$ и приравняем производные нулю. В результате получим

$$\int_{x^{(i)}}^{x^{(i+1)}} (x - x_{\text{кв}}^{(i)}) w(x) dx = 0, \quad (2.41)$$

$$[x^{(i)} - x_{\text{кв}}^{(i-1)}]^2 = [x^{(i)} - x_{\text{кв}}^{(i)}]^2. \quad (2.42)$$

Из соотношений (2.41) и (2.42) находим, что оптимальным значением

$x_{\text{кв}}^{(i)}$ является абсцисса центра тяжести криволинейной трапеции (рис. 2.7) с основанием $(x^{(i)}, x^{(i+1)})$, ограниченной сверху кривой $w(x)$, а порог квантования $x^{(i)}$ равен $(x_{\text{кв}}^{(i)} + x_{\text{кв}}^{(i-1)})/2$. В частности, из выражений (2.41) и (2.42) нетрудно

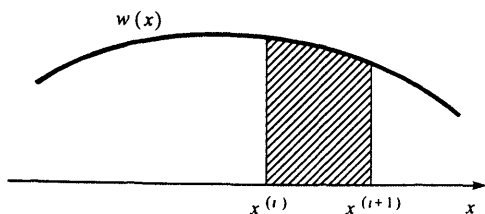


Рис. 2.7. Диаграмма, иллюстрирующая выбор уровня квантования

но видеть, что если распределение $x(t)$ равномерное, то квантование с постоянным шагом оптимальное.

Квантованные отсчеты можно передавать различными способами. На практике для этого чаще всего используют кодовые комбинации, каждая из которых соответствует определенному уровню квантования. При равномерном коде с основанием m длина кодовых комбинаций не может быть меньше k , где k выбирается из условия $L_{\text{кв}} \leq m^k$.

При выборе основания кода в первую очередь необходимо учитывать простоту, экономичность и удобство реализации цифрового представления непрерывных сообщений. На практике обычно применяют простые (безызыточные) двоичные коды, среди которых наибольшее применение нашли двоичный натуральный код, симметричный двоично-числовой код и код Грея [13].

Двоичный натуральный код — это код, комбинации которого представляют собой двоичные номера уровней квантования. Он прост в реализации и удобен при обработке на ЭВМ.

Симметричный двоично-числовой код используется для представления биполярных квантованных отсчетов. При этом высший разряд несет информацию о знаке отсчета, а остальные разряды — об абсолютном значении отсчета в натуральном двоичном коде.

Код Грея связан с двоичным натуральным кодом следующими соотношениями: $a_0^\Gamma = a_0 \oplus a_1$, $a_1^\Gamma = a_1 \oplus a_2$, ..., $a_{k-2}^\Gamma = a_{k-2} \oplus a_{k-1}$, $a_{k-1}^\Gamma = a_{k-1}$, где $a_{k-1}a_{k-2} \dots a_0$ — кодовая комбинация натурального кода, $a_{k-1}^\Gamma a_{k-2}^\Gamma \dots a_0^\Gamma$ — кодовая комбинация кода Грея, символ \oplus обозначает суммирование по модулю два. Этот код обладает следующими двумя особенностями, которые способствуют повышению быстродействия кодирующих устройств по сравнению с применением двоичного натурального кода: любые две кодовые комбинации, соответствующие соседним уровням квантования, отличаются друг от друга значениями только в одном разряде; смена значений элементов в каждом разряде при переходе от одной комбинации к другой происходит вдвое реже, чем в двоичном натуральном коде.

Кроме простых двоичных кодов используют помехоустойчивые коды, позволяющие обнаруживать и исправлять ошибки, возникающие при передаче дискретной информации.

2.6. Сложные сигналы

Под *сложными (шумоподобными, широкополосными) сигналами* понимают сигналы, для которых выполняется неравенство

$$B = F_c T_c \gg 1, \quad (2.43)$$

где B , F_c и T_c — база, ширина спектра и длительность сигнала соответственно. Для простых (узкополосных) сигналов $B \approx 1$.

Сложные сигналы можно образовать модуляцией гармонического колебания $A_0 \cos(\omega_0 t + \varphi(t))$ по амплитуде (АМ), частоте (ЧМ) и фазе (ФМ) специальной функцией $a(t)$, расширяющей спектр. При этом сигналы представляются в виде

$$s(t) = \begin{cases} a(t) A_0 \cos(\omega_0 t + \varphi(t)) & \text{при АМ,} \\ A_0 \cos\left(\omega_0 t + \Delta\omega \int a(t) dt + \varphi(t)\right) & \text{при ЧМ,} \\ A_0 \cos(\omega_0 t + \Delta\varphi a(t) + \varphi(t)) & \text{при ФМ,} \end{cases} \quad (2.44)$$

где $\Delta\omega$ и $\Delta\varphi$ — девиация частоты и фазы соответственно.

Расширяющие функции $a(t)$ должны быть детерминированными. Это требование связано с необходимостью получения идентичных реализаций функции $a(t)$ в передатчике и приемнике. Они должны обладать хорошими корреляционными и взаимно корреляционными свойствами. Именно эти свойства придают РТС, использующим сложные сигналы, такие характеристики, как высокую помехозащищенность, высокую точность измерений, подавление замираний в каналах с многолучевостью, обеспечение многостанционного доступа с одновременной работой многих РТС в одном и том же диапазоне частот, обеспечение электромагнитной совместимости с узкополосными РТС. Количество расширяющих функций должно быть как можно большим. Это обеспечивает структурную скрытность, имитостойкость, многостанционный доступ.

Для образования сложных сигналов можно использовать как непрерывную (аналоговую), так и дискретную модуляции. Примером широкополосных сигналов, полученных непрерывной модуляцией, являются сигналы с линейной частотной модуляцией (ЛЧМ-сигналы). Они нашли широкое применение в радиолокационных системах (см. гл. 4).

В системах передачи информации, как уже упоминалось, необходимо иметь большой ансамбль сложных сигналов, которые можно сформировать дискретной модуляцией гармонического колебания по частоте (дискретные частотно-модулированные (ДЧМ) сигналы) и по фазе (фазоманипулированные (ФМ) сигналы). При этом расширяющие функции $a(t)$ представляют собой кодовые последовательности, обычно двоичные.

В современных РТС в качестве кодовых последовательностей применяются линейные рекуррентные последовательности максимального периода (М-последовательности), последовательности Баркера, Голда, Касами и ряд других [12].

2.6.1. Линейные рекуррентные последовательности максимальной длины

Линейной рекуррентной последовательностью (ЛРП) называется последовательность символов $\{a_i\} = a_0 a_1 a_2 \dots$, удовлетворяющая рекуррентному правилу

$$c_0 a_i = c + c_1 a_{i-1} + c_2 a_{i-2} + \dots + c_n a_{i-n}, \quad (2.45)$$

где значения как символов последовательности $\{a_i\}$, так и коэффициентов c и c_i принадлежат некоторому алфавиту $(0, 1, \dots, L-1)$, а операции сложения и умножения производятся по модулю L , причем L предполагается простым числом.

Соотношение (2.45) называется *правилом кодирования*, число n — *памятью последовательности*, а число L — ее *основанием*. В дальнейшем будем рассматривать только случай $L = 2$.

Без потери общности в выражении (2.45) коэффициент c можно положить равным нулю. Тогда рекуррентное правило запишется в виде

$$c_0 a_i = c_1 a_{i-1} \oplus c_2 a_{i-2} \oplus \dots \oplus c_n a_{i-n}. \quad (2.46)$$

Из соотношения (2.46) следует, что для построения ЛРП необходимо в каждый тактовый момент запоминать n последних символов $a_{i-1}, a_{i-2}, \dots, a_{i-n}$ последовательности $\{a_i\}$ и суммировать их по модулю два с весами c_1, c_2, \dots, c_n . Эти операции осуществляют сдвигающим регистром с обратной связью (рис. 2.8).

Любую ЛРП $\{a_i\}$ можно задать производящей функцией $G(x)$ [12], под которой понимается формальный степенной ряд

$$G(x) = \sum_{i=0}^{\infty} a_i x^i,$$

где a_i — i -й символ последовательности $\{a_i\}$, а суммирование производится по модулю два.

Пусть $a_i = \sum_{j=1}^n c_j a_{i-j}$. Тогда

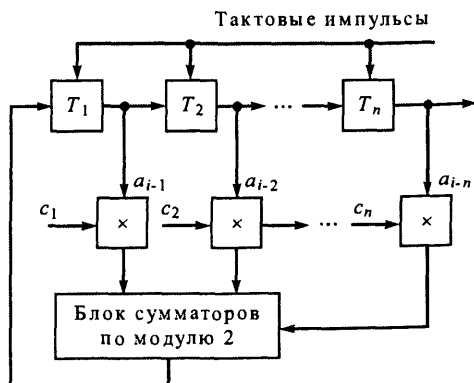


Рис. 2.8. Схема сдвигающего регистра с обратной связью

$$\begin{aligned}
G(x) &= \sum_{i=0}^{\infty} a_i x^i = \sum_{i=0}^{\infty} \sum_{j=1}^n c_j a_{i-j} x^i = \sum_{j=1}^n c_j x^j \sum_{i=0}^{\infty} a_{i-j} x^{i-j} = \\
&= \sum_{j=1}^n c_j x^j \left(a_{-j} x^{-j} \oplus \dots \oplus a_{-1} x^{-1} \oplus \sum_{i=0}^{\infty} a_i x^i \right) = \\
&= \sum_{j=1}^n c_j x^j (a_{-j} x^{-j} \oplus \dots \oplus a_{-1} x^{-1} \oplus G(x)), \tag{2.47}
\end{aligned}$$

где комбинация символов $a_{-j} a_{-j+1} \dots a_{-1}$ характеризует начальное состояние регистра сдвига, вырабатывающего последовательность $\{a_i\}$.

Из выражения (2.47) с учетом того, что для рассматриваемого случая все операции производятся по модулю два, можно получить

$$G(x) = \frac{\sum_{j=1}^n c_j x^j (a_{-j} x^{-j} \oplus \dots \oplus a_{-1} x^{-1})}{1 \oplus \sum_{j=1}^n c_j x^j} = \frac{q(x)}{f(x)},$$

где $q(x)$ и $f(x)$ — многочлены степени $r < n$ и n соответственно. Многочлен $f(x)$ полностью определяется рекуррентным правилом и называется *характеристическим многочленом*.

Выбирая многочлены $q(x)$ и $f(x)$ и производя их деление, можно получить различные рекуррентные последовательности. Пусть, например, $f(x) = 1 \oplus x \oplus x^5$. Тогда при $q(x) = 1$ получается последовательность (111110101001100010000)... с периодом 21; при $q(x) = 1 \oplus x \oplus x^3 \oplus x^4$ — последовательность (1001011)... с периодом 7; при $q(x) = 1 \oplus x^2 \oplus x^3$ — последовательность (110)... с периодом 3 и, наконец, при $q(x) = 0$ — последовательность (0)... с периодом 1.

Таким образом, каждому характеристическому многочлену $f(x)$ степени n соответствует некоторое множество из k последовательностей длины N_i , причем

$$\sum_{i=1}^k N_i = 2^n.$$

Пусть $\{a_i\}$ — ЛРП, соответствующая производящей функции $G(x) = q(x)/f(x)$, где $q(x)$ и $f(x)$ — взаимно простые многочлены. Тогда можно

показать, что периодом последовательности $\{a_i\}$ является наименьшее положительное целое число N , при котором $f(x)$ делит двучлен $1 \oplus x^N$.

Введем понятие *неприводимого* многочлена, под которым будем понимать многочлен степени n , не имеющий делителей, степень которых больше нуля, но меньше n . Рассмотрим порождающую функцию $G(x) = q(x)/f(x)$, где $f(x)$ является неприводимым многочленом степени n . Тогда можно утверждать, что период последовательности, соответствующей порождающей функции $G(x)$, не зависит от выбора многочлена $q(x)$. Выбирая различные многочлены $q(x)$, можно найти

$$k = (2^n - 1)/N \quad (2.48)$$

различных последовательностей с периодом N , соответствующих многочлену $f(x)$.

Если в выражении (2.48) $k = 1$, то период последовательности $\{a_i\}$ равен $2^n - 1$. Такие последовательности называются *линейными рекуррентными последовательностями максимальной длины (периода)*, или *М-последовательностями*. Очевидно, что характеристический многочлен М-последовательности должен делить двучлен $1 \oplus x^{2^n-1}$.

Неприводимость характеристического многочлена является необходимым, но не достаточным условием получения М-последовательности. Действительно, существуют неприводимые полиномы, которым соответствуют последовательности не максимальной длины. Например, многочлен $f(x) = x^6 \oplus x^3 \oplus 1$ является неприводимым. Однако максимальный период соответствующих ему ненулевых последовательностей равен 9.

Найдем достаточное условие получения М-последовательности. Известно, что любой неприводимый многочлен $f(x)$ степени n является делителем многочлена $1 \oplus x^{2^n-1}$. Тогда, учитывая, что $1 \oplus x^{2^n-1}$ делится также на любой многочлен $1 \oplus x^s$, если только s является делителем числа $2^n - 1$, можно утверждать, что период ЛРП с неприводимым характеристическим многочленом $f(x)$ степени n должен совпадать с одним из делителей числа $2^n - 1$. Очевидно, что для получения М-последовательности характеристический многочлен $f(x)$ степени n не должен делить никакой двучлен $1 \oplus x^s$ при $s < 2^n - 1$. Неприводимые многочлены $f(x)$ степени n , которые делят двучлен $1 \oplus x^{2^n-1}$ и не делят никакой двучлен $1 \oplus x^s$, $s < 2^n - 1$, называются *примитивными*. Таким образом, необходимым и достаточным условием существования М-последовательности является примитивность характеристического многочлена.

Каждому примитивному многочлену соответствует вполне определенная М-последовательность. Соответственно, число различных М-последовательностей памяти n равно числу примитивных многочленов степени n , которое определяется как $Q = \phi(2^n - 1)/n$, где $\phi(N)$ — функция Эйлера из теории чисел, равная количеству целых чисел, включая единицу, меньших числа N и взаимно простых с ним.

Приведем значения Q для некоторых n :

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Q	1	1	2	2	6	6	18	16	48	60	176	144	630	756	1800

В работе [14] даны все примитивные многочлены степени $n \leq 16$, а также примитивные многочлены с минимальным числом ненулевых коэффициентов степени $17 \leq n \leq 34$. В табл. 2.1 приведены примитивные многочлены по одному для $n = 3, 4, \dots, 10$.

Таблица 2.1

n	Многочлен	n	Многочлен
3	$x^3 \oplus x^2 \oplus 1$	7	$x^7 \oplus x^6 \oplus 1$
4	$x^4 \oplus x^3 \oplus 1$	8	$x^8 \oplus x^4 \oplus x^3 \oplus x^2 \oplus 1$
5	$x^5 \oplus x^3 \oplus 1$	9	$x^9 \oplus x^5 \oplus 1$
6	$x^6 \oplus x^5 \oplus 1$	10	$x^{10} \oplus x^7 \oplus 1$

Укажем основные структурные свойства М-последовательностей [12, 15—17].

1. Период М-последовательности равен $2^n - 1$, где n — степень ее характеристического многочлена.

2. М-последовательность имеет наибольший период среди ЛРП с равными степенями их характеристических многочленов.

3. В М-последовательности с периодом $2^n - 1$ содержатся все n -значные двоичные комбинации, кроме комбинаций из одних нулей, причем каждая n -значная комбинация встречается один раз.

4. В М-последовательности число единичных символов равно 2^{n-1} , а число нулей равно $2^{n-1} - 1$, т. е. на единицу меньше.

5. Если М-последовательность $\{a_i\}$ сложить по модулю два с последовательностью $\{a_{i+k}\}$, образованной из $\{a_i\}$ циклическим сдвигом на k символов, то получим исходную М-последовательность, но с некоторым другим циклическим сдвигом.

6. Пусть $\{a_i\}$ — М-последовательность памяти n , а d — любое положительное число. Тогда последовательность $\{a_{di}\}$, полученная из $\{a_i\}$ выборкой di -х членов, $i = 0, 1, \dots$, имеет период, равный $(2^n - 1)/(2^n - 1, d)$, где запись $(2^n - 1, d)$ означает наибольший общий делитель чисел $2^n - 1$ и d .

7. Если $\{a_i\}$ — М-последовательность памяти n , а d — любое из чисел $1, 2, 4, \dots, 2^n - 1$, то $\{a_{di}\}$, $i = 0, 1, \dots$, является исходной М-последовательностью с точностью до циклического сдвига.

8. Пусть $\{a_i\}$ — М-последовательность памяти n . Тогда существует ее циклический сдвиг $\{b_i\} = \{a_{i+m}\}$ такой, что $\{b_{2i}\} = \{b_i\}$, где m — некоторое целое число. М-последовательности, инвариантные относительно операции, заключающейся в выборке $2i$ -х членов последовательности, $i = 0, 1, 2, \dots, n - 1$, называются *характеристическими М-последовательностями*. Очевидно, что для характеристических М-последовательностей равенство $\{b_{2i}\} = \{b_i\}$ остается справедливым для всех $d = 2^k$, где $k = 1, 2, \dots, n - 1$.

9. Если $\{a_i\}$ — характеристическая М-последовательность, то ее производящая функция имеет вид

$$G(x) = \begin{cases} \left((xf(x))' \oplus f(x) \right) / f(x) & \text{для четных степеней } f(x), \\ (xf(x))' / f(x) & \text{для нечетных степеней } f(x). \end{cases}$$

В качестве примера в табл. 2.2 представлены производящие функции двух характеристических М-последовательностей, а также сами последовательности для $n = 5$.

Таблица 2.2

$G(x)$	М-последовательности
$(1 \oplus x^2) / (1 \oplus x^2 \oplus x^5)$	1000010101110110001111100110100
$1 / (1 \oplus x^3 \oplus x^5)$	1001011001111100011011101010000

10. Пусть $\{a_i\}$ — М-последовательность с характеристическим многочленом $f(x)$. Тогда последовательности $\{a_{di}\}$ являются М-последовательностями, если $(2^n - 1, d) = 1$. Это свойство позволяет по одной М-последовательности построить все М-последовательности того же периода.

11. Пусть $\{a_i\}$ — М-последовательность с характеристическим многочленом $f(x)$ степени n и $d_2 = d_1 2^j \pmod{(2^n - 1)}$, причем d_1 и d_2 — числа, вза-

именно простые с $2^n - 1$. Тогда последовательности $\{a_{d_i}\}$ и $\{a_{d_2 i}\}$ будут совпадать с точностью до циклического сдвига.

12. Пусть $\{a_i\}$ — М-последовательность памяти n , причем n — четное число. Тогда последовательность $\{a_{di}\}$, $i = 0, 1, \dots$, где $d = 2^{n/2} + 1$, является М-последовательностью памяти $n/2$.

Рассмотрим корреляционные свойства М-последовательностей. Найдем сначала периодическую корреляционную функцию (ПКФ)

$$R_a^n(k) = \sum_{i=0}^{N-1} a_i b_{i+k}, \quad k = 0, 1, \dots, \quad (2.49)$$

предварительно отметив, что далее все результаты приводятся для последовательностей, которые получаются из последовательностей, состоящих из символов 1 и 0, заменой 1 на -1 , а 0 на 1.

Тогда с учетом соотношения (2.49) и свойств 4, 5 М-последовательностей получаем

$$R_a^n(k) = \begin{cases} N, & k = 0 \pmod{N}, \\ -1, & k \neq 0 \pmod{N}, \end{cases}$$

где N — период М-последовательности.

Таким образом, ПКФ М-последовательности оказалась двухуровневой со значением боковых лепестков, равным -1 .

На рис. 2.9, б приведена ПКФ М-сигнала (рис. 2.9, а), представляющего собой последовательность видеоимпульсов, длительность которых равна τ_0 , а

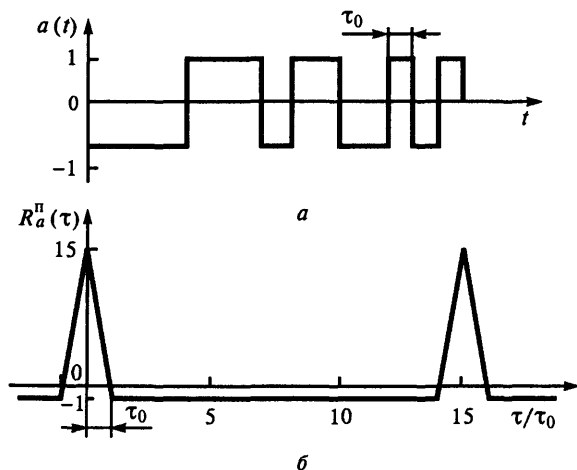


Рис. 2.9. М-сигнал (а) и его ПКФ (б)

полярность определяется М-последовательностью 111100010011010 (символу 1 соответствует видеоимпульс отрицательной полярности, а символу 0 — видеоимпульс положительной полярности). На практике используются сигналы, получаемые фазовой манипуляцией высокочастотного колебания по закону М-последовательности. У таких сигналов огибающие корреляционных функций совпадают с корреляционными функциями соответствующих М-сигналов.

Периодическая взаимокорреляционная функция (ПВКФ) М-последовательностей $\{a_i\}$ и $\{b_i\}$ с периодами N_1 и N_2 соответственно имеет вид

$$R_{ab}^n(k) = \sum_{i=0}^{\text{НОК}(N_1, N_2)-1} a_i b_{i+k},$$

т. е. является периодической функцией с периодом, равным наименьшему общему кратному чисел N_1 и N_2 ($\text{НОК}(N_1, N_2)$).

Периодическая взаимокоррелирующая функция достаточно просто определяется только для М-последовательностей с взаимно простыми периодами. Можно показать, что для них $R_{ab}^n(k) = 1$ при всех k . В других случаях ПВКФ не имеет общих закономерностей в своем поведении. На практике часто достаточно знать лишь максимальный уровень взаимной корреляции, а не детальное поведение ПВКФ. Укажем три подмножества М-последовательностей с определенным уровнем взаимной корреляции [16].

1. Пусть $\{a_i\}$ — некоторая характеристическая М-последовательность памяти n , где n — нечетное число. Образует из нее М-последовательность $\{b_i\} = \{a_{di}\}$; здесь

$$d = 2^i + 1, \quad (d, n) = 1. \quad (2.50)$$

Тогда ПВКФ последовательностей $\{a_i\}$ и $\{b_i\}$ определяется как

$$R_{ab}^n(k) = \begin{cases} -1, & \text{если } a_k = 1, \\ -(2^{(n+1)/2} + 1) \text{ или } 2^{(n+1)/2} - 1, & \text{если } a_k = -1. \end{cases} \quad (2.51)$$

2. Пусть $\{a_i\}$ — характеристическая М-последовательность. Образует М-последовательность $\{b_i\} = \{a_{di}\}$; здесь

$$d = \begin{cases} 2^{(n+1)/2} + 1, & \text{если } n - \text{нечетное число;} \\ 2^{(n+2)/2} + 1, & \text{если } n - \text{четное число, } n \neq 0 \pmod{4}. \end{cases} \quad (2.52)$$

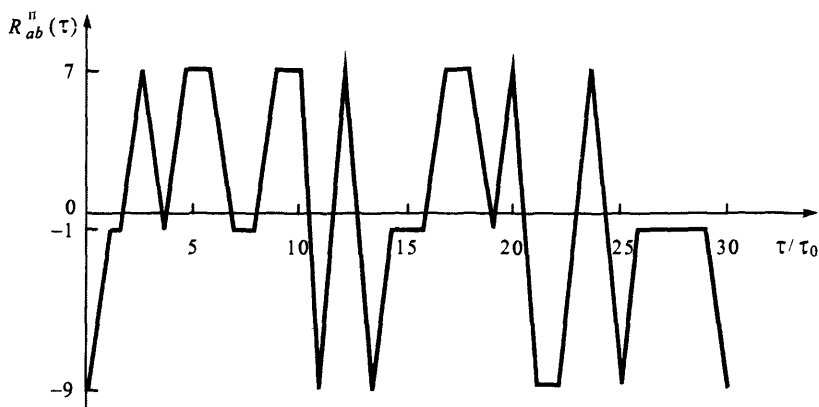


Рис. 2.10. ПВКФ сигналов, соответствующих М-последовательностям

Тогда ПВКФ удовлетворяет неравенству

$$|R_{ab}^n(k)| \leq \begin{cases} 2^{(n+1)/2} + 1, & \text{если } n - \text{нечетное число,} \\ 2^{(n+2)/2} + 1, & \text{если } n - \text{четное число.} \end{cases}$$

3. Пусть $\{a_i\}$ — характеристическая М-последовательность памяти n , причем $n \equiv 0 \pmod{4N}$. образуем последовательность $\{b_i\} = \{a_{di}\}$, где $d = 2^{(n+2)/2} - 1$. Тогда ПВКФ последовательностей $\{a_i\}$ и $\{b_i\}$ принимают только следующие четыре значения: $2^{(n+2)/2} - 1$, $2^{n/2} - 1$, -1 , $-(2^{n/2} + 1)$. Таким образом, $|R_{ab}^n(k)| \leq 2^{(n+2)/2} - 1$.

На рис. 2.10 изображена ПВКФ сигналов, соответствующих М-последовательностям

$$\{a_i\} = -11111-11-11-1-1-11-1-1111-1-1-1-1-111-1-11-111,$$

$$\{b_i\} = -111-111-1-11111-11-1-11-11-1111-1-1-11-1-1-1-1.$$

Сформулированные правила позволяют выбирать пары М-последовательностей с гарантированным уровнем взаимной корреляции, что имеет большое практическое значение. Так, для $n = 13$ существует 630 М-последовательностей. Среди них имеются пары, для которых значения ПВКФ достигают 703. Однако согласно формуле (2.51) имеются также пары с максимальным значением ПВКФ, равным 129.

В общем случае максимальный уровень боковых лепестков ПВКФ М-последовательностей одинакового периода лежит в пределах $(1,5...6)\sqrt{N}$ [17].

Непериодическая корреляционная функция (НКФ) М-последовательности имеет вид

$$R_a^H(k) = \sum_{i=1}^{N-k} a_i a_{i+k}.$$

Одной из ее характеристик является также максимальный уровень боковых лепестков. Для различных М-последовательностей он оказывается различным. Последовательности, для которых наибольший уровень боковых лепестков НКФ оказывается наименьшим, называются *минимаксными*. Для таких сигналов максимальное значение $R_a^H(k)$, $k \neq 0$, несколько меньше \sqrt{N} и стремится к этой величине с ростом N .

Непериодическая корреляционная функция обладает следующими двумя свойствами:

$$R_a^H(k) + R_a^H(N-k) = -1, \quad \sum_{k=1}^{N-1} R_a^H(k) = -\frac{N-1}{2}.$$

Непериодическая взаимокорреляционная функция (НВКФ) М-последовательностей не подчиняется каким-либо общим закономерностям. Согласно [17] максимальный уровень лепестков НВКФ лежит в пределах $(1,4...5,1)\sqrt{N}$.

2.6.2. Линейные рекуррентные последовательности не максимальной длины

Среди линейных последовательностей не максимальной длины наибольшее значение имеют последовательности Голда и Касами [16]. На их основе могут быть построены большие системы сигналов с малыми значениями периодических взаимокорреляционных функций.

Последовательности Голда образуются следующим образом. Пусть $f_a(x)$ и $f_b(x)$ — пара примитивных многочленов степени n , порождающих соответственно М-последовательности $\{a_i\}$ и $\{b_i\} = \{a_{di}\}$, где число d удовлетворяет условию (2.50) или (2.52). Тогда многочлен $f(x) = f_a(x)f_b(x)$ будет порождать $2^n + 1$ различных последовательностей с периодом $2^n - 1$, ПВКФ любой пары $\{c_i\}$ и $\{d_i\}$ которых принимают те же значения, что и ПВКФ М-последовательностей $\{a_i\}$ и $\{b_i\}$. Следовательно, они имеют трехуровневую ПВКФ, удовлетворяющую условию

$$|R_{ca}^n(k)| \leq \begin{cases} 2^{(n+1)/2} + 1, & \text{если } n - \text{нечетное число;} \\ 2^{(n+2)/2} + 1, & \text{если } n - \text{четное число.} \end{cases} \quad (2.53)$$

Образованные последовательности и называются *последовательностями Голда*.

Последовательности Голда можно получить почленным суммированием по модулю два циклических сдвигов М-последовательностей $\{a_i\}$ и $\{b_i\}$, порождаемых исходными многочленами $f_a(x)$ и $f_b(x)$. Действительно, производящую функцию $G(x) = q(x)/(f_a(x)f_b(x))$ любой последовательности Голда $\{c_i\}$ можно представить в виде

$$G(x) = \frac{q_1(x)}{f_a(x)} \oplus \frac{q_2(x)}{f_b(x)}, \quad (2.54)$$

где $q_1(x)$ и $q_2(x)$ — многочлены, степень которых меньше n . Из выражения (2.54) непосредственно следует, что $c_i = a_i \oplus b_i$. В множество последовательностей Голда входят и сами последовательности $\{a_i\}$ и $\{b_i\}$.

На рис. 2.11 изображена ПВКФ сигналов, соответствующих последовательностям Голда

$$\{c_i\} = 1000000000100101001001111010101,$$

$$\{d_i\} = 1111111111000110001110101100110,$$

формируемым сдвигающим регистром с функцией обратной связи

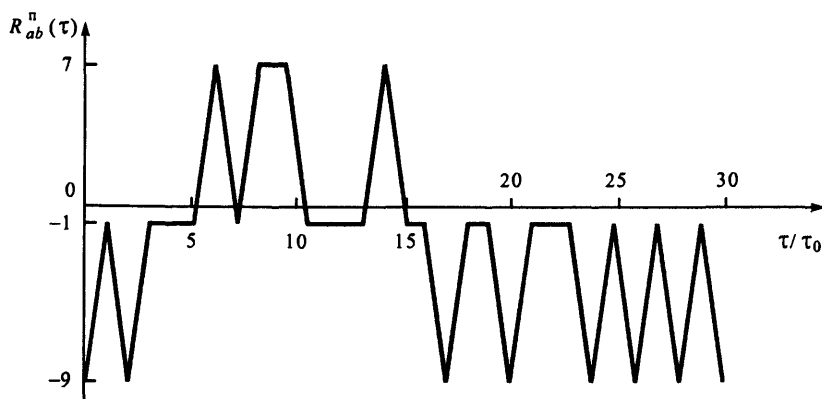


Рис. 2.11. ПВКФ сигналов, соответствующих последовательностям Голда

$$f(x) = f_a(x)f_b(x) = (1 \oplus x^2 \oplus x^5)(1 \oplus x^2 \oplus x^3 \oplus x^4 \oplus x^5) = \\ = x^{10} \oplus x^9 \oplus x^8 \oplus x^6 \oplus x^5 \oplus x^3 \oplus 1.$$

В данном случае получается 33 сигнала с максимальным значением модуля ПВКФ, равным 9.

При $n \equiv 0 \pmod{4}$ можно образовать множество последовательностей типа Голда. Пусть $f_a(x)$ и $f_b(x)$ — пара примитивных многочленов степени $n \equiv 0 \pmod{4}$, порождающих соответственно М-последовательности $\{a_i\}$ и $\{b_i\} = \{a_{di}\}$, где $d = 2^{(n+2)/2} + 1$. Тогда многочлен $f(x) = f_a(x)f_b(x)$ порождает 2^n различных последовательностей, которые и называются *последовательностями типа Голда*. Любая пара $\{c_i\}$ и $\{d_i\}$ этих последовательностей характеризуются тем, что их ПВКФ принимает только следующие значения: -1 , $-(2^{n/2} + 1)$, $(2^{n/2} - 1)$, $-(2^{(n+2)/2} + 1)$. Таким образом, для них $|R_{cd}^n| \leq 2^{(n+2)/2} + 1$.

Последовательности Касами образуются следующим образом. Пусть $\{a_i\}$ — характеристическая М-последовательность периода $N = 2^n - 1$, порождаемая многочленом $f_a(x)$ степени n , причем n — четное число. Образует последовательность $\{b_i\} = \{a_{di}\}$, где $d = 2^{n/2} + 1$. Полученная таким образом последовательность является М-последовательностью с периодом $2^{n/2} - 1$. Пусть многочлен $f_b(x)$ степени $n/2$ есть характеристический многочлен М-последовательности $\{b_i\}$. Тогда многочлен $f(x) = f_a(x)f_b(x)$ степени $3n/2$ порождает множество K_s последовательностей с периодом $2^n - 1$, каждая из которых может быть получена почленным суммированием по модулю два циклических сдвигов последовательностей $\{a_i\}$ и $\{b_i\}$. В множество K_s входит также последовательность $\{a_i\}$. Образованное множество последовательностей называется *малым множеством последовательностей Касами*.

Последовательности Касами характеризуются тем, что их ПВКФ принимает только значения -1 , $-(2^{n/2} + 1)$ и $2^{n/2} - 1$. Таким образом, для них

$$|R_{cd}^n| \leq 2^{n/2} + 1. \quad (2.55)$$

Из выражений (2.53) и (2.55) находим, что для последовательностей Касами при небольших n значения ПВКФ существенно меньше, чем для последовательностей Голда.

Определим класс последовательностей, носящий название *большого множества последовательностей Касами*. Пусть n — четное число, $f_a(x)$ — примитивный многочлен степени n , порождающий М-последовательность $\{a_i\}$, и $f_b(x)$ — примитивный многочлен степени n , порождающий М-последовательность $\{b_i\} = \{a_{di}\}$, где $d = 2^{(n+2)/2} + 1$, и $f_c(x)$ — примитивный многочлен степени $n/2$, порождающий М-последовательность $\{c_i\} = \{a_{qi}\}$, где $q = 2^{n/2} + 1$. Тогда многочлен $f(x) = f_a(x)f_b(x)f_c(x)$ порождает множество K_L последовательностей с периодом $2^n - 1$, каждая из которых может быть получена почленным суммированием по модулю два циклических сдвигов М-последовательностей $\{a_i\}$, $\{b_i\}$ и $\{c_i\}$. В множество K_L входят также все последовательности Голда при $n \not\equiv 0 \pmod{4}$. Образованное множество последовательностей и называется *большим множеством последовательностей Касами*. При $n \not\equiv 0 \pmod{4}$ множество K_L состоит из $2^{n/2}(2^n + 1)$ последовательностей, а при $n \equiv 0 \pmod{4}$ — из $2^{n/2}(2^n + 1) - 1$ последовательностей.

Последовательности, входящие во множество K_L , характеризуются тем, что их ПВКФ принимает только следующие значения: -1 , $-(2^{(n+2)/2} + 1)$, $2^{(n+2)/2} - 1$, $-(2^{n/2} + 1)$, $2^{n/2} - 1$. Таким образом, для них $|R_{lm}^n(k)| \leq 2^{(n+2)/2} + 1$.

В качестве примера укажем, что многочлен

$$x^{20} \oplus x^{19} \oplus x^{14} \oplus x^{12} \oplus x^9 \oplus x^8 \oplus x^7 \oplus x^2 \oplus 1$$

порождает 4111 последовательностей с периодом $N = 255$, ПВКФ любой пары которых принимает значения -1 ; -33 ; 31 ; -17 ; 15 .

2.6.3. Сегменты М-последовательностей

Большой класс сигналов с хорошими корреляционными свойствами можно получить, разбивая М-последовательности большого периода $2^n - 1$ на отрезки или сегменты длиной $n \ll N_{\text{сер}} \ll N$. Исследования показывают, что взаимокорреляционные функции (ВКФ) сегментов М-последовательностей подобны ВКФ самих М-последовательностей [17]. В частности, максимальный уровень лепестков НВКФ лежит в интервале $(1, 4 \dots 4, 3) \sqrt{N_{\text{сер}}}$, а максимальный уровень лепестков ПВКФ — в интервале $(1, 6 \dots 5, 0) \sqrt{N_{\text{сер}}}$.

Корреляционные свойства сегментов M -последовательностей значительно хуже, чем у M -последовательностей. Так, максимальное значение боковых лепестков НКФ равно $(1,45...4,1)\sqrt{N_{\text{сег}}}$, а максимальное значение боковых лепестков ПКФ равно $(1,6...4,3)\sqrt{N_{\text{сег}}}$. При фиксированном N с увеличением длины отрезка максимальный уровень лепестков ВКФ уменьшается, а при фиксированной длине отрезка увеличение N приводит к его росту.

В работе [12] приведены распределения значений ПВКФ для некоторых классов сегментов. Можно показать, что среднее значение (первый момент) и дисперсия (второй момент) этих распределений зависят только от длины сегмента $N_{\text{сег}}$ и периода N M -последовательности, из которой они образованы, и не зависят от характеристического многочлена используемой M -последовательности. Третий момент, характеризующий асимметрию распределения значений ПВКФ сегментов, зависит от характеристического многочлена M -последовательности. С точки зрения получения множества сегментов с хорошими взаимокорреляционными свойствами, необходимо использовать M -последовательности, для которых распределение значений ПВКФ сегментов симметричное.

Шумоподобные сигналы, описанные в настоящем параграфе, находят применение в радиолокации (гл. 4, 7), радионавигации (гл. 8), радиосистемах передачи информации (гл. 9).

2.7. Системы сигналов

При построении систем передачи информации, таких как многоканальные системы с кодовым уплотнением, m -ичные системы (системы, в которых для передачи сообщений используют m различных сигналов), асинхронно-адресные системы, необходимо располагать множеством сигналов. Совокупность сигналов, объединенных единым правилом построения, называется *системой сигналов*. Существует бесконечно большое число систем, отличающихся друг от друга индивидуальными и совместными свойствами сигналов. От выбора того или иного множества сигналов зависят такие характеристики системы передачи, как помехоустойчивость, полоса занимаемых частот, сложность и ряд других. Система сигналов, обеспечивающая максимальную помехоустойчивость при заданных априорных условиях передачи информации, называется *оптимальной*.

При действии в канале помехи типа белого гауссовского шума помехоустойчивость системы зависит от расстояния между сигналами

$$d(s_i, s_j) = \left[\int_0^{T_c} (s_i(t) - s_j(t))^2 dt \right]^{1/2}, \quad i, j = 1, 2, \dots, m, \quad (2.56)$$

причем чем больше минимальное из этих расстояний, тем выше помехоустойчивость системы.

Если сигналы имеют одинаковую энергию E , то выражение (2.56) можно упростить:

$$d(s_i, s_j) = (2E(1 - r_{ij}))^{1/2}, \quad (2.57)$$

где $r_{ij} = \frac{1}{E} \int_0^{T_c} s_i(t) s_j(t) dt$ — коэффициент взаимной корреляции между сигналами $s_i(t)$ и $s_j(t)$. Из формулы (2.57) следует, что для достижения большего расстояния коэффициент взаимной корреляции должен быть как можно меньше.

Для обеспечения одной и той же вероятности правильного приема при передаче любого сообщения необходимо, чтобы все коэффициенты r_{ij} были одинаковыми, т. е. $r_{ij} = r_0$ для всех i и j , $i \neq j$. Значение r_0 удовлетворяет неравенству $r_0 \geq -1/(m-1)$, которое вытекает из следующего соотношения:

$$\int_0^{T_c} \left(\sum_{i=1}^m s_i(t) \right)^2 dt = \int_0^{T_c} \sum_{i=1}^m \sum_{j=1}^m s_i(t) s_j(t) dt = mE + r_0 E(m^2 - m) \geq 0.$$

Очевидно, что для оптимальной системы сигналов

$$r_0 = -1/(m-1). \quad (2.58)$$

Сигналы $s_i(t)$, $i = 1, 2, \dots, m$, удовлетворяющие условию

$$\frac{1}{E} \int_0^{T_c} s_i(t) s_j(t) dt = \begin{cases} 1, & j = i, \\ -1/(m-1), & j \neq i, \end{cases}$$

называются *симплексными*, поскольку в $(m-1)$ -мерном пространстве они образуют правильный симплекс с числом вершин m . Симплексные сигналы являются эквидистантными, т. е. для всех пар сигналов $s_i(t)$ и $s_j(t)$ расстояние $d(s_i, s_j)$ одинаково.

На практике часто применяют ортогональные сигналы, для которых

$$r_{ij} = \frac{1}{E} \int_0^{T_c} s_i(t) s_j(t) dt = \begin{cases} 1, & j = i, \\ 0, & j \neq i, \end{cases} \quad i, j = 1, 2, \dots, m.$$

При больших значениях m ортогональные сигналы по помехоустойчивости близки к симплексным. Это следует из того, что значение r_0 , определяемое соотношением (2.58), при больших m стремится к нулю. Ортогональные сигналы с равной энергией также являются эквидистантными.

Другой системой, близкой при $m \gg 1$ к симплексной, является биортогональная система сигналов $s_i(t)$, $i = 1, 2, \dots, m$ (m — четное число), которая характеризуется тем, что для каждого сигнала $s_i(t)$ существует противоположный сигнал $-s_i(t)$, а остальные ортогональны сигналу $s_i(t)$.

Различают непрерывные и дискретные системы сигналов. В непрерывных системах сигналы могут быть построены на основе ортогональных полиномов Лежандра, Эрмита, Лаггера, функций Бесселя, Матье и т. п. [12]. Дискретные сигналы строятся с использованием матриц Адамара, а также различных кодовых последовательностей, таких как линейные рекуррентные последовательности максимальной длины (М-последовательностей), последовательностей Лежандра, Холла, Якоби, Голда, Касами и др. [12]. Поскольку непрерывные системы сигналов практически не нашли применения в РТС, далее будут рассматриваться только дискретные системы.

2.7.1. Ортогональные сигналы

В общем случае ортогональные сигналы можно сформировать следующим образом. Пусть $\varphi_j(t)$, $j = 1, 2, \dots, N$, — некоторая полная ортонормированная система функций. Тогда любой сигнал $s_i(t)$, $i = 1, 2, \dots, m$, с полосой частот F_c можно представить в виде

$$s_i(t) = \sum_{j=1}^N a_{ij} \varphi_j(t),$$

где $N = 2F_c T_c$ — число отсчетов на интервале T_c по теореме Котельникова,

$$a_{ij} = \int_0^{T_c} s_i(t) \varphi_j(t) dt, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, N,$$

— коэффициенты разложения.

Геометрически сигнал $s_i(t)$ можно представить вектором в N -мерном

пространстве с координатами $(a_{i1}, a_{i2}, \dots, a_{iN})$. Сигналы $s_i(t)$, $i = 1, 2, \dots, m$, будут ортогональны, если для любого i -го сигнала выполняется соотношение

$$a_{ij} = \begin{cases} \sqrt{E}, & j = i, \\ 0, & j \neq i. \end{cases}$$

Рассмотрим в качестве примера базисные функции

$$\varphi_j(t) = \begin{cases} \sqrt{2/T_c} \sin \omega_j t & \text{при } 0 \leq t \leq T_c, \\ 0 & \text{при других } t, \end{cases}$$

где частоты ω_j , $j = 1, 2, \dots, m$, выбираются из условия ортогональности функций $\varphi_j(t)$. Тогда сигналы

$$s_i(t) = \sqrt{\frac{2E}{T_c}} \sin \omega_i t, \quad 0 \leq t \leq T_c, \quad i = 1, 2, \dots, m,$$

образуют ортогональную систему.

Существует бесконечно большое число ортогональных систем функций, на основе которых могут быть сформированы ортогональные коды. При этом сами сигналы получают фазовой манипуляцией несущего колебания по закону кодовых комбинаций.

В общем случае построение ортогональных кодов связано с матрицами Адамара, являющимися квадратными ортогональными матрицами с элементами ± 1 . Поэтому строки (или столбцы) матрицы Адамара можно использовать для формирования комбинаций ортогонального кода (символ -1 заменяется символом 0).

Укажем два положения, касающихся существования и построения матриц Адамара.

1. Матрицы Адамара имеют порядок либо $N = 2$, либо $N = 4k$, $k = 1, 2, \dots$.

2. Матрица $A_{N_1 \times N_2}$ порядка $N_1 \times N_2$, полученная из матрицы Адамара A_{N_2} подстановкой матрицы Адамара A_{N_1} вместо элемента $+1$ и матрицы $-A_{N_1}$ вместо элемента -1 , есть также матрица Адамара.

Таким образом можно легко строить матрицы Адамара более высоких порядков.

Рассмотрим в качестве примера матрицы Адамара

$$\mathbf{A}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{A}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

Используя указанный способ, нетрудно получить матрицу Адамара порядка $N = 8$:

$$\mathbf{A}_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}.$$

Если первая строка и первый столбец матрицы Адамара состоят из единиц, то говорят, что матрица записана в нормальной форме.

Ортогональные коды можно построить на основе системы функций Уолша [12], которые достаточно просто генерируются.

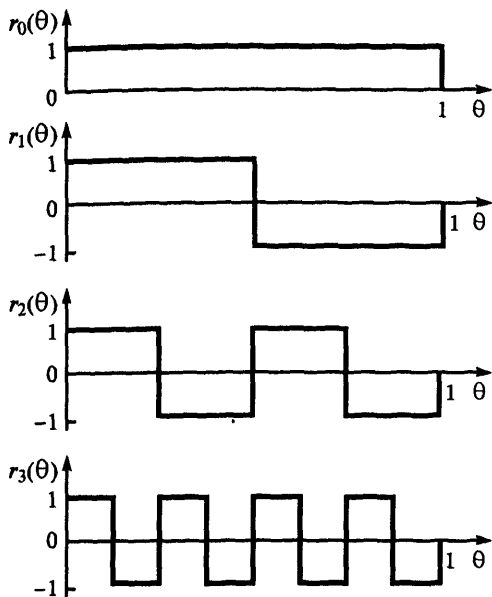
Система функций Уолша впервые была описана математиком Уолшем (J. Walsh) в 1923 г. В настоящее время существует ряд определений, позволяющих строить различные модификации этой системы, отличающиеся интервалом определения и порядком следования функций. Приведем сначала определение системы, практически совпадающей с системой, введенной Уолшем, в которой упорядочение функций производится по числу пересечений ими нулевого уровня. Система обычно обозначается как $\{\text{wal}_i(\theta)\}$, $i = 0, 1, 2, \dots$, где $0 \leq \theta = t/T < 1$, T — период функций. Далее будем рассматривать конечные системы, состоящие из $N = 2^n$ функций.

Введем предварительно функции Радемахера [12]

$$r_0(\theta) \equiv 1, \quad r_i(\theta) = \text{sign}[\sin(2^i \pi \theta)], \quad i = 1, 2, \dots, \quad (2.59)$$

где

$$\text{sign}[x] = \begin{cases} 1, & x > 0, \\ -1, & x < 0. \end{cases}$$

Рис. 2.12. Графики функций $r_i(\theta)$

Из выражения (2.59) следует, что эти функции являются дискретными и принимают только два значения: $+1$ на подынтервалах $(k/2^i, (k+1)/2^i)$, $k = 2j$, $j = 0, 1, \dots$, и -1 на остальных подынтервалах. На рис. 2.12 представлены первые четыре функции $r_i(\theta)$.

Система функций Радемахера является ортогональной на интервале $[0, 1]$, но неполной, так как на том же интервале существуют другие функции, ортогональные им.

Система функций Уолша $\{\text{wal}_i(\theta)\}$, $i = 0, 1, \dots, 2^{n-1}$, является расширением системы функций Радемахера до полной системы и определяется как

$$\varphi_0(\theta) = \text{wal}_0(\theta) \equiv 1, \quad \varphi_i(\theta) = \text{wal}_i(\theta) = \prod_{j=1}^n [r_j(\theta)]^{i_j^r}, \quad (2.60)$$

где i_j^r — значение j -го разряда в записи числа i в коде Грея.

Получение первых восьми функций Уолша в соответствии с выражением (2.60) наглядно показано в табл. 2.3, а на рис. 2.13 приведены их графики.

Функции Уолша являются дискретными (принимают значения ± 1), периодическими с периодом, равным 1. Они удовлетворяют условиям ортогональности, нормировки и мультипликативности:

$$\int_0^1 \text{wal}_i(\theta) \text{wal}_j(\theta) d\theta = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j; \end{cases} \quad \text{wal}_i(\theta) \text{wal}_j(\theta) = \text{wal}_{i \oplus j}(\theta),$$

где $i \oplus j$ — условная запись числа, двоичное представление которого получается поразрядным сложением по модулю два двоичных представлений чисел i и j . Следующий пример поясняет нахождение $k = i \oplus j$. Пусть $i = 7$, $j = 10$. Запишем i и j в двоичной системе исчисления и сложим их поразрядно по модулю два:

$$\begin{array}{r} 0111 \\ \oplus 1010 \\ \hline 1101 \end{array}$$

Полученное двоичное число является двоичной записью числа k . Таким образом, $k = i \oplus j = 13$.

Таблица 2.3

i	Двоичное представление числа i	Представление числа в коде Грея	$wal_i(\theta)$
0	000	000	$wal_0(\theta) = 1$
1	001	001	$wal_1(\theta) = r_1(\theta)$
2	010	011	$wal_2(\theta) = r_1(\theta)r_2(\theta)$
3	011	010	$wal_3(\theta) = r_2(\theta)$
4	100	110	$wal_4(\theta) = r_2(\theta)r_3(\theta)$
5	101	111	$wal_5(\theta) = r_1(\theta)r_2(\theta)r_3(\theta)$
6	110	101	$wal_6(\theta) = r_1(\theta)r_3(\theta)$
7	111	100	$wal_7(\theta) = r_3(\theta)$

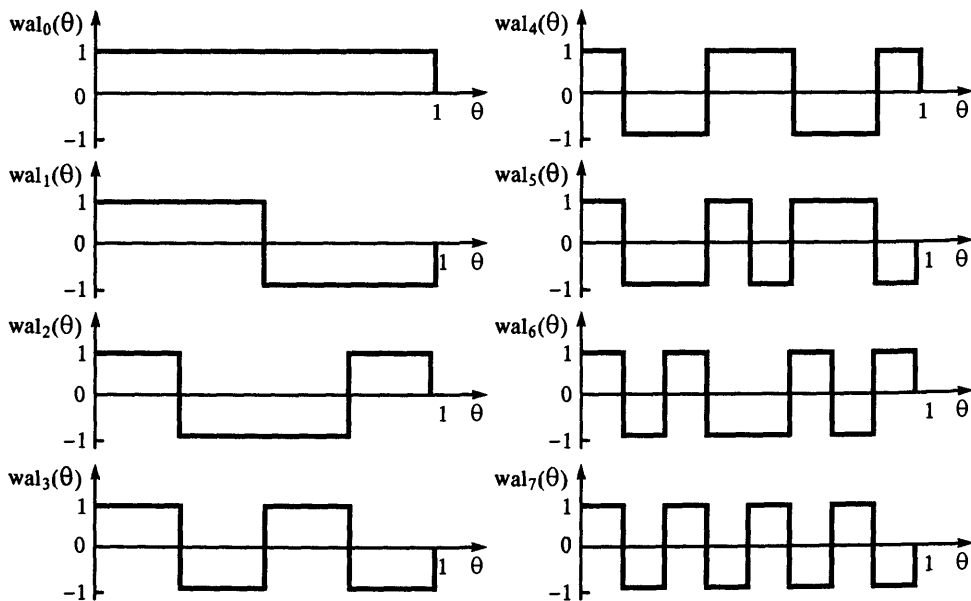


Рис. 2.13. Графики функций Уолша

Другая система функций Уолша, упорядоченная по числу пересечений нулевого уровня, — это система $\{\text{wal}_0(\theta), \text{cal}_i(\theta), \text{sal}_i(\theta)\}$, $i=1, 2, \dots$. Здесь буквосочетание al связано с фамилией Walsh, а первые буквы указывают на аналогию, в смысле четности и нечетности этих функций, с функциями $\cos \theta$ и $\sin \theta$.

Функции $\text{cal}_i(\theta)$, $i=1, 2, \dots$, являются четными, а $\text{sal}_i(\theta)$, $i=1, 2, \dots$, — нечетными. Параметр i равен половине числа пересечений нулевого уровня соответствующими функциями на интервале единичной длины.

Функции системы $\{\text{wal}_i(\theta)\}$ связаны с функциями $\text{cal}_i(\theta)$ и $\text{sal}_i(\theta)$ следующими соотношениями: $\text{cal}_i(\theta) = \text{wal}_{2i}(\theta)$, $\text{sal}_i(\theta) = \text{wal}_{2i-1}(\theta)$, $i=1, 2, \dots$.

На рис. 2.14 в соответствии с соотношениями (2.60) представлена структурная схема простого устройства для генерирования первых 16-ти функций Уолша. На выходах триггеров T_1, \dots, T_4 формируются функции $\text{wal}_1(\theta) = \text{sal}_1(\theta)$, $\text{wal}_3(\theta) = \text{sal}_2(\theta)$, $\text{wal}_7(\theta) = \text{sal}_4(\theta)$, $\text{wal}_{15}(\theta) = \text{sal}_8(\theta)$. Перемножители образуют из них систему функций Уолша от $\text{wal}_2(\theta)$ до $\text{wal}_{14}(\theta)$. Функция $\text{wal}_0(\theta)$ тождественно равна единице.

Функции Уолша не обладают хорошими корреляционными свойствами. Многие из них имеют большие боковые лепестки как КФ, так и ВКФ. По этой причине они применяются, в основном, в синхронных многоканальных системах.

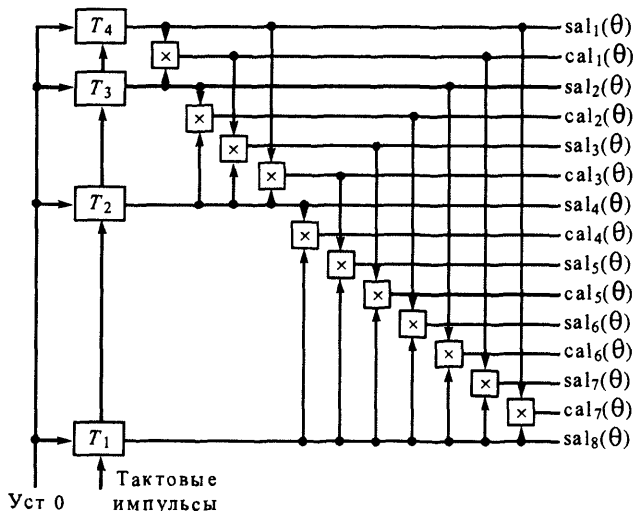


Рис. 2.14. Структурная схема устройства генерирования функций Уолша

2.7.2. Биортогональные сигналы

Система из m биортогональных сигналов формируется из $m/2$ ортогональных сигналов добавлением к каждому из них противоположного сигнала. Простейшей такой системой является система из четырех сигналов с одинаковой энергией.

Если в качестве базисных функций $\varphi_j(t)$ использовать функции

$$\varphi_1(t) = \sqrt{2/T_c} \cos \omega_0 t, \quad \varphi_2(t) = \sqrt{2/T_c} \sin \omega_0 t,$$

то при $m=4$ биортогональные сигналы будут отличаться только фазой и совпадут с сигналами, полученными фазовой манипуляцией.

2.7.3. Симплексные сигналы

В общем случае симплексные сигналы получаются из ортогональных сигналов следующим образом [10]. Пусть $\{a_i\}$, $i=1, 2, \dots, m$, — ортогональные сигналы. Добавив к каждому сигналу $s_i(t)$, $i=1, 2, \dots, m$, один и тот же сигнал $s(t)$, получим новую систему сигналов $s'_i(t) = s_i(t) + s(t)$. Заметим, что обе системы сигналов обеспечивают одинаковую помехоустойчивость. Суммарная энергия новых сигналов определяется следующей формулой:

$$E_\Sigma = \sum_{i=1}^m \int_0^{T_c} (s_i(t) + s(t))^2 dt. \quad (2.61)$$

Для симплексных сигналов энергия E_Σ должна быть минимальной. Минимизируя выражение (2.61) по $s(t)$, можно показать, что симплексные сигналы, получаемые на основе ортогональных $s_i(t)$, $i=1, 2, \dots, m$, имеют вид [10]

$$s'_i(t) = s_i(t) - \frac{1}{m} \sum_{i=1}^m s_i(t). \quad (2.62)$$

Поскольку с учетом выражения (2.62) $\sum_{i=1}^m s'_i(t) = 0$, то каждый из сигналов $s'_i(t)$ можно представить в виде линейной комбинации остальных. Отсюда следует, что симплексные сигналы $s'_i(t)$, $i=1, 2, \dots, m$, можно представить в виде векторов из $(m-1)$ -мерного евклидова пространства.

При одной базисной функции $\varphi(t)$ можно сформировать два симплексных сигнала, которые совпадут с противоположными сигналами.

При двух базисных функциях можно сформировать три симплексных сигнала. Концы их векторов лежат в вершинах равностороннего треугольника.

Симплексные сигналы могут быть получены на основе симплексных кодов. Рассмотрим равномерный код с основанием 2 и длиной m . Пусть его кодовые комбинации представляют последовательности из m символов, принимающих значения -1 и 1 . Введем понятие коэффициента взаимной корреляции между любой парой кодовых комбинаций $\{b_i\}$ и $\{b_j\}$:

$$r_{ij} = \frac{1}{m} \sum_{k=0}^{m-1} b_{ik} b_{jk}.$$

Потребуем, чтобы коэффициенты взаимной корреляции r_{ij} были бы равны r_0 , $i, j = 1, \dots, m$, $i \neq j$. Можно показать, что

$$r_0 \geq \begin{cases} -1/(m-1), & \text{если } m \text{ — четное число;} \\ -1/m, & \text{если } m \text{ — нечетное число.} \end{cases}$$

Код, все пары кодовых комбинаций которого имеют коэффициенты взаимной корреляции

$$r_0 = \begin{cases} -1/(m-1), & \text{если } m \text{ — четное число;} \\ -1/m, & \text{если } m \text{ — нечетное число,} \end{cases}$$

называется *симплексным*. Соответствующее название носят сигналы, полученные на его основе, например, путем фазовой манипуляции несущей.

Симплексные коды можно построить на основе матриц Адамара. Нетрудно показать, что если существует матрица Адамара порядка $N = 4k$, то можно построить симплексные коды для $m = 4k$, $4k - 1$, $2k$ и $2k - 1$, где k — целое положительное число. Действительно, пусть A_N — матрица Адамара порядка $N = 4k$, записанная в нормальной форме. Тогда, вычеркнув первый столбец, получаем матрицу, строки которой образуют симплексный код для $m = 4k$. Если, кроме того, вычеркнуть одну из строк, то получим симплексный код для $m = 4k - 1$.

Коды для $m = 2k$ и $m = 2k - 1$ получаются следующим образом. Рассмотрим j -й столбец ($j \neq 1$) и вычеркнем в матрице A_N строки, на которых элементы j -го столбца равны 1 (или -1). Вычеркнем также первый и j -й столбцы. Тогда окажется, что оставшиеся строки образуют симплексный код для $m = 2k$. Если вычеркнуть еще одну строку, то получим симплексный код для $m = 2k - 1$.

Особый интерес представляют симплексные коды, комбинации которых являются циклическими перестановками одной из них. Выясним, каким условиям должны удовлетворять такие кодовые последовательности.

Введем периодическую корреляционную функцию (ПКФ) кодовой последовательности b_0, b_1, \dots, b_{m-1} :

$$R^n(j) = \sum_{i=0}^{m-1} b_i b_{i+j}, \quad j = 0, 1, \dots, m-1.$$

Последовательности, ПКФ которых удовлетворяет условию $R^n(j) = a$, $j = 1, \dots, m-1$, называются *последовательностями с двухуровневой ПКФ*. Можно показать, что $|a| \geq 1$.

Для построения циклических симплексных кодов используют только последовательности с двухуровневой ПКФ, для которых $a = -1$. К таким последовательностям относятся последовательности максимальной длины (М-последовательности), последовательности Лежандра, Холла и Якоби [12].

2.8. Моделирование сигналов и помех

2.8.1. Моделирование случайных величин

Для формирования на ЭВМ случайных величин с различными законами распределения используют равномерно распределенные в интервале $(0, 1)$ случайные числа, которые получают на ЭВМ по стандартным программам, входящим в математическое обеспечение.

Плотность вероятности случайной величины X , равномерно распределенной на интервале (a, b) , имеет следующий вид:

$$w(x) = 1/(b - a).$$

Для рассматриваемого случая $a = 0$, $b = 1$, $w(x) = 1$. При этом распределение имеет математическое ожидание $m_x = 0,5$ и дисперсию $D_x = 1/12$.

Существуют различные методы преобразования случайных чисел с равномерным распределением в случайные числа с заданным законом распределения. Рассмотрим основные из них [18, 19].

Метод нелинейного преобразования, обратного функции распределения, использует соотношение

$$y = F^{-1}(x), \quad (2.63)$$

где y — значения случайной величины Y с заданной плотностью вероятности $w(y)$; x — значения случайной величины X с равномерным законом распределения на интервале $(0, 1)$; $F^{-1}(x)$ — функция, обратная функции распределения $F(y)$.

Пример. Определить алгоритм моделирования случайной величины Y с рэлеевским законом распределения:

$$w(y) = \frac{y}{\sigma^2} \exp\left(-\frac{y^2}{2\sigma^2}\right).$$

Функция распределения случайной величины Y имеет вид

$$F(y) = \int_0^y w(z) dz = \int_0^y \frac{z}{\sigma^2} \exp\left(-\frac{z^2}{2\sigma^2}\right) dz = 1 - \exp\left(-\frac{y^2}{2\sigma^2}\right).$$

С учетом выражения (2.63) получаем

$$x = F(y) = 1 - \exp\left(-\frac{y^2}{2\sigma^2}\right). \quad (2.64)$$

Отсюда следует алгоритм формирования заданной случайной величины

$$y = \sigma \sqrt{-2 \ln(1-x)}. \quad (2.65)$$

Выражение (2.65) можно представить и в другом виде:

$$y = \sigma \sqrt{-2 \ln(x)}.$$

Переход от $\ln(1-x)$ к $\ln(x)$ в последнем выражении основан на том, что случайные величины $1-X$ и X имеют одинаковые законы распределения.

В табл. 2.4 приведены алгоритмы моделирования случайных величин с наиболее распространенными законами распределения, полученные данным методом [18, 19].

Метод на основе преобразования нормально распределенных случайных чисел использует известные закономерности нелинейного преобразования случайных величин Z с нормальным распределением в случайные величины Y с другими распределениями. Алгоритмы моделирования случайных величин, полученные данным методом, представлены в табл. 2.4. Здесь же приводятся алгоритмы моделирования исходных нормальных случайных величин с использованием датчика случайных чисел с равномерным законом распределения на интервале $(0, 1)$ [19].

Метод Неймана применим для моделирования случайных величин Y с усеченными законами распределения (значения случайных величин принадлежат ограниченному интервалу (a, b)), а также случайных величин, законы распределения которых можно аппроксимировать усеченными законами. Процедура моделирования заключается в следующем.

Таблица 2.4

Распределение	Плотность распределения $w(y)$	Алгоритмы формирования y	
		x — числа, равномерно распределенные на интервале $(0,1)$	z — числа, нормально распределенные с параметрами $(0,1)$
Равномерное	$1/(b-a), a \leq y \leq b$	$a + (b-a)x$	—
Нормальное	$\frac{1}{\sqrt{2\pi}\sigma} \times \exp\left(\frac{-(y-m)^2}{2\sigma^2}\right)$	$\sigma\sqrt{2 \ln x_1} \sin 2\pi x_2 + m$ $\left(\frac{\sum_{i=1}^n x_i - n/2}{\sqrt{n/12}}\right) + m$	$\sigma z + m$
Рэлея	$\frac{y}{\sigma^2} \exp\left(\frac{-y^2}{2\sigma^2}\right), y \geq 0$	$\sigma\sqrt{-2 \ln x}$	$\sigma\sqrt{z_1^2 + z_2^2}$
Райса	$\frac{y}{\sigma^2} \exp\left(-\frac{y^2 + a^2}{2\sigma^2}\right) \times I_0\left(\frac{ay}{\sigma^2}\right), y \geq 0$	$(a^2 - 2\sigma^2 \ln x_1 - 2a\sigma \times \sqrt{-2 \ln x_1} \cos(2\pi x_2))^{1/2}$	$\sqrt{(\sigma z_1 + a)^2 + z_2^2}$
Логарифмически нормальное	$\frac{1}{y\sigma\sqrt{2\pi}} \times \exp\left(-\frac{(\ln y - m)^2}{2\sigma^2}\right), y \geq 0$	—	$\exp(\sigma z + m)$
Показательное	$\lambda \exp(-\lambda y), y \geq 0$	$-\frac{1}{\lambda} \ln x$	$\frac{1}{2\lambda} (z_1^2 + z_2^2)$
Арсинусное	$\frac{1}{\pi a \sqrt{1 - (y/a)^2}}, y \leq a$	$a \sin(2\pi x)$	—
Хи-квадрат	$\frac{1}{2^{n/2} \Gamma(n/2)} y^{(n/2)-1} e^{-y/2}, y \geq 0$	—	$\sum_{i=1}^n z_i^2$

Примечание. $I_0(\cdot)$ — модифицированная функция Бесселя нулевого порядка, $\Gamma(\cdot)$ — гамма-функция.

Из датчика равномерно распределенных на интервале $(0, 1)$ случайных чисел независимо выбирают пары чисел $x_1^{(i)}, x_2^{(i)}, i = 1, 2, \dots$, из которых формируются преобразованные пары

$$z_1^{(i)} = a + (b - a)x_1^{(i)}, \quad z_2^{(i)} = w_{\max}x_2^{(i)}, \quad (2.66)$$

где (a, b) — интервал возможных значений случайной величины Y с заданной плотностью вероятностью $w(y)$, w_{\max} — максимальное значение функции $w(y)$. В качестве реализации случайной величины Y выбирают число $z_1^{(i)}$ из тех пар $z_1^{(i)}, z_2^{(i)}$, для которых выполняется неравенство

$$z_2^{(i)} \leq w(z_1^{(i)}). \quad (2.67)$$

Пары чисел, не удовлетворяющие неравенству (2.67), отбрасывают.

Покажем, что сформированная таким образом случайная величина Y будет иметь заданное распределение $w(y)$. Из формулы (2.66) следует, что случайные числа $z_1^{(i)}$ и $z_2^{(i)}$ распределены равномерно на интервалах (a, b) и $(0, w_{\max})$ соответственно. Эти пары случайных чисел можно рассматривать как координаты случайных точек, лежащих внутри прямоугольника $aa'b'b$ (рис. 2.15). Пары $z_1^{(i)}, z_2^{(i)}$, удовлетворяющие условию (2.67), можно рассматривать как координаты случайных точек, лежащих внутри той части прямоугольника $aa'b'b$, которая расположена под кривой $w(y)$. Вероятность того, что случайная точка, находящаяся под кривой $w(y)$, попадет в элементарную полосу с основанием $(y, y + \Delta y)$, определяется произведением $w(y)\Delta y$. Поскольку попадание случайной точки в рассматриваемую элементарную полосу возможно только тогда, когда значения случайной величины Y попадают в интервал $(y, y + \Delta y)$, то можно утверждать, что вероятность выполнения условия $y \leq Y \leq y + \Delta y$

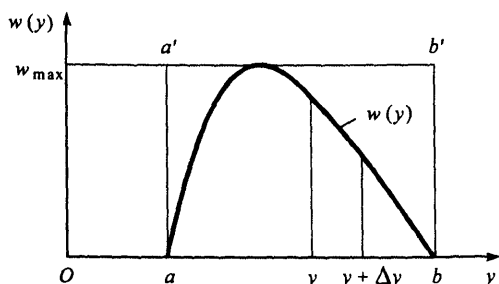


Рис. 2.15. Диаграмма, поясняющая метод Неймана

равна $w(y)\Delta y$. Отсюда следует, что сформированная случайная величина Y действительно имеет заданное распределение $w(y)$.

Рассмотрим метод кусочной аппроксимации. Пусть требуется получить случайную величину Y с плотностью вероятности $w(y)$. Предположим, что область возможных значений величины Y

ограничена интервалом (a, b) (неограниченное распределение можно приближенно заменить ограниченным). Разобьем интервал (a, b) на n достаточно малых интервалов (a_m, a_{m+1}) , $m = 0, 1, 2, \dots, n-1$, $a_0 = a$, $a_n = b$, так, чтобы распределение заданной случайной величины в пределах этих интервалов можно было достаточно точно аппроксимировать каким-нибудь простым распределением, например равномерным, трапецидальным и т. д. Далее будем использовать равномерное распределение (рис. 2.16, а). Пусть P_m — вероятность попадания случайной величины в интервал (a_m, a_{m+1}) , при-

чем $\sum_{m=0}^{n-1} P_m = 1$.

Процедура моделирования случайной величины с принятой аппроксимацией закона распределения заключается в следующем.

1. Случайным образом с вероятностью P_m выбирается интервал (a_m, a_{m+1}) . Выбор интервала производится с помощью датчика случайных равномерно распределенных чисел X на интервале $(0, 1)$. Для этого интервал $(0, 1)$ разбивается на n интервалов (x_m, x_{m+1}) , $m = 0, 1, \dots, n-1$, $x_0 = 0$, $x_n = 1$, длиной $x_{m+1} - x_m = P_m$ каждый (рис. 2.16, б). Из датчика случайных равномерно распределенных на интервале $(0, 1)$ чисел выбирается некоторая реализация x' . Последовательно сравнивая x' с x_m , определяем тот интервал (x_m, x_{m+1}) , в который попадает x' .

В основу этого процесса положен очевидный факт: вероятность попадания равномерно распределенной на интервале $(0, 1)$ случайной величины X в некоторый интервал (x_m, x_{m+1}) равна длине этого интервала.

2. Формируется реализация Δy_m случайной величины ΔY_m , равномерно распределенной на интервале $(0, a_{m+1} - a_m)$, с помощью другой реализации x'' равномерно распределенной случайной величины X на интервале $(0, 1)$:

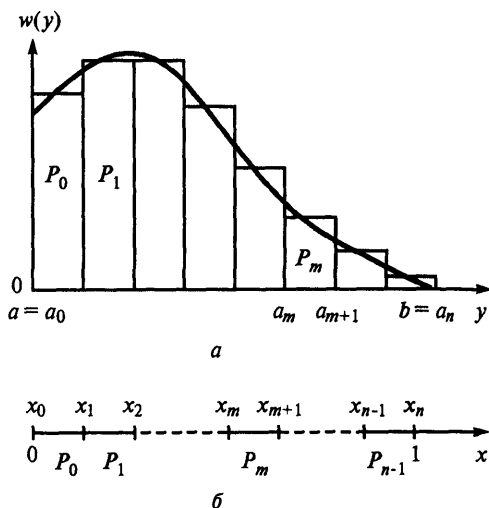


Рис. 2.16. Диаграммы, поясняющие метод кусочной аппроксимации:

а — аппроксимация плотности вероятности $w(y)$; б — разбиение интервала $(0, 1)$ на подынтервалы

$$\Delta y_m = (a_{m+1} - a_m) x^n.$$

3. Искомая реализация y получается по формуле

$$y = a_m + \Delta y_m = a_m + (a_{m+1} - a_m) x^n.$$

Метод кусочной аппроксимации широко используется для формирования дискретных случайных величин и случайных событий. Рассмотрим формирование последовательности чисел, подчиняющихся биномиальному закону распределения вероятностей. К биномиальному закону приводит задача повторения опытов. Если вероятность некоторого события A равна p , то вероятность того, что событие A при n опытах появится k раз, определяется следующей формулой:

$$P_k = C_n^k p^k (1-p)^{n-k},$$

где $C_n^k = \frac{n!}{(n-k)!k!}$, $k = 0, 1, \dots, n$.

Для наглядности изложения рассмотрим биномиальный закон P_k для $p = 0,5$ и $n = 4$ (рис. 2.17, а).

Процедура моделирования числовой последовательности с биномиальным законом распределения вероятностей состоит в следующем. Интервал $(0, 1)$ разбивается на подынтервалы, длина которых равна P_k (рис. 2.17, б). Генерируются равномерно распределенные на интервале $(0, 1)$ случайные числа x_i . При попадании числа x_i в k -й интервал, т. е. при выполнении неравенства

$$\sum_{j=0}^{k-1} P_j < x_i \leq \sum_{j=0}^k P_j,$$

случайному числу x_i приписывается значение k .

Дадим сравнительную характеристику методов моделирования случайных величин.

Для достижения высокой точности воспроизведения законов распределения случайных ве-

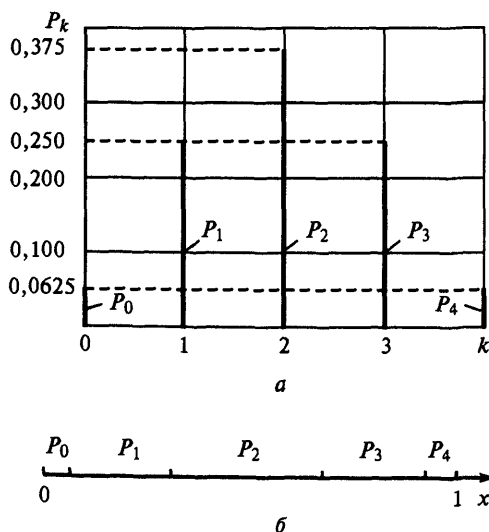


Рис. 2.17. Диаграммы, поясняющие алгоритм моделирования случайной величины с биномиальным законом распределения:

а — биномиальный закон ($p = 0,5$, $n = 4$); б — разбиение интервала $(0, 1)$ на подынтервалы

личин целесообразно использовать алгоритмы, не обладающие методической погрешностью. К ним относятся алгоритмы, полученные методом нелинейного преобразования, обратной функции распределения, и ряд алгоритмов на основе преобразования нормально распределенных случайных чисел. Погрешностью этих алгоритмов моделирования можно пренебречь, поскольку она определяется лишь погрешностью выполнения на ЭВМ необходимых нелинейных преобразований и отклонением закона распределения исходных случайных чисел от равномерного (нормального) закона распределения.

Другим преимуществом указанных алгоритмов является простота подготовительной работы, так как преобразования равномерного закона (или нормального) в требуемый описываются аналитическими зависимостями. Такого вида алгоритмы позволяют легко изменять форму распределения в процессе моделирования случайных величин, закон распределения которых зависит от параметров.

Основным недостатком этих алгоритмов является сравнительно низкое быстродействие, поскольку осуществление на ЭВМ нелинейных преобразований часто требует довольно большого количества элементарных операций.

В задачах, не предъявляющих высокие требования к точности распределения формируемых случайных величин, для сокращения количества элементарных операций рекомендуется использовать более экономные приближенные методы, например метод кусочной аппроксимации.

2.8.2. Моделирование случайных векторов

Существуют два основных метода моделирования случайных векторов с заданной плотностью распределения вероятностей.

1. *Метод условных распределений.* Алгоритм основан на рекуррентном вычислении условных плотностей вероятностей для координат случайного вектора.

Рассмотрим сначала двумерный случай, когда вектор имеет две координаты Y_1 и Y_2 . Одномерную плотность вероятности величины Y_1 можно представить через двумерную плотность вероятности величин Y_1 и Y_2 следующим образом:

$$w(y_1) = \int_{-\infty}^{\infty} w(y_1, y_2) dy_2. \quad (2.68)$$

Используя описанные выше способы моделирования случайных величин с заданным законом распределения, сформируем реализацию $y_1^{(i)}$ случайной величины Y_1 с плотностью вероятности (2.68). Затем найдем условное распределение случайной величины Y_2 :

$$w(y_2 | y_1^{(i)}) = w(y_1^{(i)}, y_2) / w(y_1^{(i)}). \quad (2.69)$$

Произведем выборку $y_2^{(i)}$ случайной величины Y_2 с плотностью вероятности $w(y_2 | y_1^{(i)})$. Полученная таким образом последовательность пар чисел $y_1^{(i)}, y_2^{(i)}$, $i = 1, 2, \dots$, будет иметь совместную плотность вероятности $w(y_1, y_2)$.

Пример. Составить алгоритм моделирования двумерного случайного вектора с плотностью вероятности вида

$$w(y_1, y_2) = cy_1 \exp(-y_1 y_2), \quad 0 \leq y_1 \leq 2, \quad y_2 \geq 0. \quad (2.70)$$

По выражению (2.68) находим одномерную плотность вероятности величины Y_1 :

$$w(y_1) = \int_0^{\infty} w(y_1, y_2) dy_2 = \int_0^{\infty} cy_1 \exp(-y_1 y_2) dy_2 = c, \quad 0 \leq y_1 \leq 2.$$

Находим значение c из условия нормировки

$$\int_0^2 w(y_1) dy_1 = \int_0^2 c dy_1 = 2c = 1,$$

откуда $c = 0,5$ и $w(y_1) = 0,5$, $0 \leq y_1 \leq 2$.

Из табл. 2.4 получаем следующий алгоритм моделирования Y_1 :

$$y_1^{(i)} = 2(x_1^{(i)}).$$

По выражению (2.69) находим условное распределение случайной величины Y_2 :

$$w(y_2 | y_1^{(i)}) = w(y_1^{(i)}, y_2) / w(y_1^{(i)}) = y_1^{(i)} \exp(-y_1^{(i)} y_2),$$

откуда видно, что при полученном значении $y_1^{(i)}$ случайная величина Y_2 имеет показательный закон распределения. Алгоритм моделирования такой случайной величины приведен в табл. 2.4:

$$y_2^{(i)} = -(1/y_1^{(i)}) \ln(x_2^{(i)}).$$

Таким образом, окончательно алгоритм моделирования случайного вектора с координатами Y_1, Y_2 и плотностью распределения, определяемой выражением (2.70), будет иметь следующий вид:

$$y_1^{(i)} = 2(x_1^{(i)}), \quad y_2^{(i)} = -\left(\frac{1}{2x_1^{(i)}}\right) \ln(x_2^{(i)}),$$

где $x_1^{(i)}, x_2^{(i)}$ — i -я пара выборок, полученная из датчика случайных чисел с равномерным законом распределения на интервале $(0,1)$.

Аналогично, если задана совместная плотность вероятности $w(y_1, y_2, y_3)$ трехмерного вектора, то выборка трех чисел осуществляется в соответствии с плотностями вероятностей

$$w(y_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(y_1, y_2, y_3) dy_2 dy_3,$$

$$w(y_2 | y_1^{(i)}) = \int_{-\infty}^{\infty} w(y_1^{(i)}, y_2, y_3) dy_3 / w(y_1^{(i)}),$$

$$w(y_3 | y_1^{(i)}, y_2^{(i)}) = w(y_1^{(i)}, y_2^{(i)}, y_3) / w(y_1^{(i)}) w(y_2^{(i)} | y_1^{(i)}).$$

Описанный прием позволяет моделировать, в принципе, многомерные случайные векторы с произвольно заданной плотностью вероятности. Однако использование этого способа связано с весьма громоздкими вычислениями, за исключением тех сравнительно редких случаев, когда интегралы в выражениях типа (2.68) имеют конечные пределы интегрирования. В противном случае приходится прибегать к приближенным вычислениям. При больших значениях N эти вычисления, как правило, оказываются также громоздкими и совершенно непригодны для практического использования.

2. Метод Неймана, обобщенный на многомерный случай. Идея метода такая же, как и в одномерном случае (см. п. 2.8.1), с той лишь разницей, что здесь имитируются случайные точки, равномерно распределенные не на плоскости под кривой $w(y)$, а в $(N+1)$ -мерном объеме под N -мерной поверхностью $w(y_1, y_2, \dots, y_N)$.

Пусть $w(y_1, y_2, \dots, y_N)$ — N -мерная плотность вероятности случайного вектора \mathbf{Y} . Случайные координаты Y_k , $k=1, 2, \dots, N$, имеют область определения (a_k, b_k) .

По аналогии с одномерным случаем для формирования реализации вектора \mathbf{Y} с помощью ЭВМ находят $N+1$ случайных чисел $Z_1, Z_2, \dots, Z_N, Z_{N+1}$, равномерно распределенных на интервалах (a_1, b_1) , (a_2, b_2) , ..., (a_N, b_N) , $(0, w_{\max})$ соответственно. Их реализации моделируются следующим образом:

$$\begin{aligned}
z_1^{(i)} &= a_1 + (b_1 - a_1)x_1^{(i)}, & x_1^{(i)} &\in (a_1, b_1); \\
z_2^{(i)} &= a_2 + (b_2 - a_2)x_2^{(i)}, & x_2^{(i)} &\in (a_2, b_2); \\
&\dots\dots\dots & \dots\dots\dots \\
z_N^{(i)} &= a_N + (b_N - a_N)x_N^{(i)}, & x_N^{(i)} &\in (a_N, b_N); \\
z_{N+1}^{(i)} &= w_{\max}x_{N+1}^{(i)}, & x_{N+1}^{(i)} &\in (0, w_{\max}),
\end{aligned}$$

где w_{\max} — максимальное значение функции $w(y_1, y_2, \dots, y_N)$, $x_1^{(i)}, x_2^{(i)}, \dots, x_{N+1}^{(i)}$ — реализации случайных независимых чисел X_1, X_2, \dots, X_{N+1} с равномерным законом распределения на интервале $(0, 1)$.

В качестве реализации случайного вектора Y , распределенного по закону $w(y_1, y_2, \dots, y_N)$, принимают реализации случайного вектора Z с координатами $z_1^{(i)}, z_2^{(i)}, \dots, z_N^{(i)}$, удовлетворяющими условию

$$z_{N+1}^{(i)} \leq w(z_1^{(i)}, z_2^{(i)}, \dots, z_N^{(i)}).$$

Реализации случайных чисел $z_1^{(i)}, z_2^{(i)}, \dots, z_N^{(i)}$, не удовлетворяющие этому условию, отбрасывают.

Моделирование случайных векторов с заданной матрицей корреляционных моментов может осуществляться *методом линейного преобразования*. Основная идея метода состоит в том, чтобы, получив N независимых случайных величин X_1, X_2, \dots, X_N с нулевыми математическими ожиданиями и одинаковыми дисперсиями, равными 1, подвергнуть их такому линейному преобразованию A , после которого полученные величины Y_1, Y_2, \dots, Y_N имели бы наперед заданную корреляционную матрицу

$$R = [R_{nm}] = [M\{Y_n \times Y_m\}],$$

где M — символ математического ожидания, $m = 1, 2, \dots, N$, $n = 1, 2, \dots, N$.

Известно, что произвольное линейное преобразование N -мерного вектора X сводится к умножению его на некоторую матрицу A N -го порядка:

$$Y = A \times X, \quad (2.71)$$

где Y, X — матрицы-столбцы с элементами $y_1, y_2, \dots, y_N, x_1, x_2, \dots, x_N$ соответственно; $A = [a_{nm}]$ — квадратная матрица преобразования.

Выберем матрицу преобразования A треугольной. Тогда в развернутой форме выражение (2.71) можно записать следующим образом:

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_N \end{bmatrix}. \quad (2.72)$$

Из формулы (2.72) следует, что

[illegible]

Учитывая, что случайные величины X_1, X_2, \dots, X_N независимы, а их дисперсии равны 1, элементы матрицы A определяются выражениями

$$\begin{aligned} R_{11} &= \mathbf{M}\{Y_1^2\} = a_{11}^2, \\ R_{12} &= \mathbf{M}\{Y_1 \times Y_2\} = a_{11} \times a_{21}, \\ R_{22} &= \mathbf{M}\{Y_2^2\} = a_{21}^2 + a_{22}^2, \dots \end{aligned} \quad (2.73)$$

Из соотношений (2.73) можно определить элементы a_{nm} :

$$\begin{aligned} a_{11} &= \sqrt{R_{11}}, & a_{21} &= R_{12} / \sqrt{R_{11}}, & a_{22} &= \sqrt{R_{22} - R_{12}^2 / R_{11}}, \\ a_{31} &= R_{13} / \sqrt{R_{11}}, & a_{32} &= \frac{R_{23} - R_{12}R_{13} / R_{11}}{\sqrt{R_{22} - R_{12}^2 / R_{11}}}, \\ a_{33} &= \sqrt{R_{33} - \frac{R_{13}^2}{R_{11}} - \frac{(R_{23} - R_{12}R_{13} / R_{11})^2}{R_{22} - R_{12}^2 / R_{11}}}, \dots \end{aligned}$$

Таким образом можно последовательно определить все элементы матрицы A . Тогда алгоритм моделирования случайных векторов с заданной корреляционной матрицей сведется к умножению реализаций вектора с независимыми случайными координатами на матрицу A . Составляющие вектора Y будут иметь нулевые средние значения. Вектор Z с ненулевым средним значением получается суммированием векторов Y и C , где C — вектор средних значений координат вектора Z .

Отметим, что рассмотренный процесс моделирования позволяет получить лишь необходимые корреляционные связи между координатами случайного вектора. Законы распределения координат не учитываются. Они

могут быть произвольными, например равномерными. Требуется только, чтобы координаты X_1, X_2, X_3 были независимы и их дисперсии равны 1.

Если законы распределения координат исходного вектора принять нормальными, то искомый вектор также будет нормальным (нормальный закон, как известно, инвариантен по отношению к линейному преобразованию).

2.8.3. Моделирование гауссовских случайных процессов

Стационарный гауссовский случайный процесс $\xi(t)$ однозначно определяется математическим ожиданием m и корреляционной функцией $R(\tau)$. В задачах моделирования можно предполагать, что математическое ожидание равно нулю, так как случайный процесс с произвольным m может быть получен следующим преобразованием:

$$\xi(t) = m + \xi'(t),$$

где $\xi'(t)$ — гауссовский случайный процесс с нулевым математическим ожиданием.

Обычно задача моделирования формулируется следующим образом. По известным характеристикам процесса (математическому ожиданию, корреляционной функции или спектральной плотности мощности) требуется определить вычислительный алгоритм, позволяющий получать на ЭВМ дискретные* реализации $\xi(n\Delta t) = \xi[n]$ случайного процесса $\xi(t)$, где Δt — интервал дискретизации, $n = 1, 2, \dots$

Известны два основных метода моделирования стационарных гауссовских случайных процессов: метод скользящего суммирования и метод рекуррентных алгоритмов [18, 19]. В основу этих методов положено линейное преобразование стационарной последовательности $x[n]$ независимых гауссовских случайных чисел с параметрами $m_x = 0$ и $D_x = 1$ (нормированный дискретный белый шум) в дискретные реализации $\xi[n]$ случайного процесса с заданными корреляционно-спектральными характеристиками. Плотность распределения вероятностей $w(x)$ и корреляционная функция $R_x[k]$ исходного нормированного дискретного белого шума $x[n]$ имеют следующий вид:

$$w(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right),$$

* Здесь и далее значение целочисленного аргумента, стоящего в квадратных скобках, означает, что речь идет о дискретной функции.

$$R_x[k] = \mathbf{M}\{x[n]x[n+k]\} = \begin{cases} 1, & k=0, \\ 0, & k \neq 0. \end{cases} \quad (2.74)$$

Возможные методы формирования на ЭВМ нормированного дискретного белого шума $x[n]$ рассмотрены в п. 2.8.1.

Метод скользящего суммирования. Дискретные значения моделируемого процесса $\xi[n]$ формируются в виде скользящей суммы значений $x[n]$ с весовыми коэффициентами c_k :

$$\xi[n] = c_1 x[n-1] + c_2 x[n-2] + \dots + c_N x[n-N] = \sum_{k=1}^N c_k x[n-k]. \quad (2.75)$$

Вид корреляционной функции случайного процесса, моделируемого с помощью алгоритма (2.75), определяется количеством и значениями весовых коэффициентов c_k , которые находятся на этапе предварительной подготовки к моделированию.

Алгоритм (2.75) описывает поведение некоторого дискретного линейного фильтра, который из нормированного дискретного белого шума формирует на выходе дискретный случайный процесс с заданными корреляционно-спектральными характеристиками. Передаточная функция $K(z)$ этого фильтра определяется как отношение z -преобразования выходного сигнала $\xi[n]$ к z -преобразованию входного сигнала $x[n]$:

$$K(z) = \frac{\mathbf{Z}\{\xi[n]\}}{\mathbf{Z}\{x[n]\}}, \quad (2.76)$$

где $\mathbf{Z}\{\xi[n]\} = \sum_{n=0}^{\infty} \xi[n]z^{-n}$, $\mathbf{Z}\{x[n]\} = \sum_{n=0}^{\infty} x[n]z^{-n}$, $z = \exp(p\Delta t)$, $p = \sigma + j\omega$ — комплексное число, действительная часть которого выбирается из условия сходимости рядов $\mathbf{Z}\{\xi[n]\}$ и $\mathbf{Z}\{x[n]\}$, Δt — интервал дискретизации.

Используя основные свойства z -преобразования: свойство линейности

$$\mathbf{Z}\{Ax_1[n] + Bx_2[n]\} = A\mathbf{Z}\{x_1[n]\} + B\mathbf{Z}\{x_2[n]\}$$

и свойство сдвига

$$\mathbf{Z}\{x[n-k]\} = z^{-k}\mathbf{Z}\{x[n]\},$$

где z^{-k} — оператор сдвига на k интервалов дискретизации, находим z -преобразование выходного сигнала $\xi[n]$ в следующем виде:

$$\mathbf{Z}\{\xi[n]\} = \mathbf{Z}\left\{\sum_{k=1}^N c_k x[n-k]\right\} = \mathbf{Z}\{x[n]\} \sum_{k=1}^N c_k z^{-k}.$$

Тогда в силу формулы (2.76) имеем

$$K(z) = \sum_{k=1}^N c_k z^{-k}. \quad (2.77)$$

Для моделирования случайных процессов методом скользящего суммирования требуется решить задачу нахождения весовых коэффициентов c_k по заданной корреляционной функции или спектральной плотности мощности.

Существуют различные способы определения c_k . Рассмотрим некоторые из них.

Способ нахождения весовых коэффициентов на основе решения нелинейной алгебраической системы уравнений. Используя выражение (2.75), найдем дискретные значения моделируемого процесса:

$$\begin{aligned} \xi[n] &= c_1 x[n-1] + c_2 x[n-2] + \dots + c_N x[n-N], \\ \xi[n+1] &= c_1 x[n] + c_2 x[n-1] + \dots + c_N x[n+1-N], \\ \xi[n+2] &= c_1 x[n+1] + c_2 x[n] + \dots + c_N x[n+2-N], \dots \end{aligned} \quad (2.78)$$

Зависимость (коррелированность) между дискретными значениями $\xi[n]$ и $\xi[n+k]$ случайного процесса обеспечивается в силу того, что в их формировании участвуют $N - k$ общих дискретных значений исходной последовательности $x[n]$. При $k \geq N$ значения $\xi[n]$ и $\xi[n+k]$ становятся некоррелированными.

Дискретные значения $R(k\Delta t) = R[k]$ корреляционной функции $R(\tau)$ в точках $\tau = k\Delta t$, где $k = 0, 1, 2, \dots, N$, сформированного случайного процесса $\xi[n]$ определяются следующим образом:

$$R[k] = \mathbf{M}\{\xi[n]\xi[n+k]\}. \quad (2.79)$$

Подставляя значения $\xi[n]$ и $\xi[n+k]$ из соотношения (2.78) в (2.79) и учитывая (2.74), получаем

$$\begin{aligned} R[0] &= c_1^2 + c_2^2 + \dots + c_N^2 = D_\xi, \\ R[1] &= c_1 c_2 + c_2 c_3 + \dots + c_{N-1} c_N, \\ &\dots \dots \dots \\ R[n-1] &= c_1 c_N, \\ R[N] &= 0. \end{aligned} \quad (2.80)$$

Первое уравнение системы (2.80) записано с учетом того, что $R(0) = D_\xi$, где D_ξ — дисперсия моделируемого процесса $\xi(t)$.

Решая систему уравнений (2.80) при заданных значениях $R[k]$, можно получить весовые коэффициенты c_k .

Пример. Определить алгоритм моделирования гауссовского случайного процесса с корреляционной функцией треугольной формы

$$R(\tau) = \begin{cases} \sigma^2 (1 - |\tau|/\tau_0), & |\tau| < \tau_0, \\ 0, & |\tau| \geq \tau_0. \end{cases}$$

Полагаем, что $N = \tau_0 / \Delta\tau$ — целое число. Пусть $N = 4$. Для этого случая система уравнений (2.80) принимает вид

$$\begin{aligned} R[0] &= c_1^2 + c_2^2 + c_3^2 + c_4^2 = \sigma^2, \\ R[1] &= c_1 c_2 + c_2 c_3 + c_3 c_4 = 0,75\sigma^2, \\ R[2] &= c_1 c_3 + c_2 c_4 = 0,5\sigma^2, \\ R[3] &= c_1 c_4 = 0,25\sigma^2, \\ R[4] &= 0. \end{aligned} \tag{2.81}$$

Нетрудно убедиться в том, что для системы уравнений (2.81) все весовые коэффициенты c_k оказываются одинаковыми, т. е. $c_1 = c_2 = \dots = c_N = c$.

Тогда из первого уравнения системы (2.81) определяем

$$c = \sigma / \sqrt{N}. \tag{2.82}$$

Подставляя равенство (2.82) в (2.75), получаем следующий алгоритм формирования случайного процесса:

$$\xi[n] = \left(\sigma / \sqrt{N} \right) \sum_{k=1}^N x[n-k].$$

При моделировании случайного процесса $\xi(t)$ интервал дискретизации по времени $\Delta t = \Delta\tau$ следует выбирать из условия $\Delta\tau \ll \tau_k$, где τ_k — интервал корреляции процесса $\xi(t)$.

Нахождение весовых коэффициентов c_k решением системы уравнений (2.80) в общем случае требует значительного объема вычислительной работы. Кроме того, при изменении интервала дискретизации необходимо повторно решать систему уравнений (2.80).

Способ нахождения весовых коэффициентов на основе разложения спектральной плотности мощности в ряд Фурье. При данном способе весовые коэффициенты вычисляются следующим образом [18, 19]:

$$c_k = \frac{1}{\omega_c} \int_0^{\omega_c} \sqrt{\frac{\omega_c}{\pi}} G(\omega) \cos(k\pi\omega/\omega_c) d\omega, \quad (2.83)$$

где $G(\omega)$ — спектральная плотность мощности моделируемого процесса, ω_c — граничная частота спектра процесса. При этом алгоритм моделирования гауссовского случайного процесса принимает вид

$$\xi[n] = \sum_{k=-m}^m c_k x[n-k]. \quad (2.84)$$

Параметр m , ограничивающий число весовых коэффициентов c_k , можно выбрать из условия

$$\left| 1 - \frac{1}{D_\xi} \sum_{k=-m}^m c_k^2 \right| < \varepsilon, \quad (2.85)$$

где D_ξ — дисперсия моделируемого случайного процесса, ε — погрешность моделирования.

Условие (2.85) следует из того, что сумма квадратов весовых коэффициентов c_k должна быть равна дисперсии моделируемого случайного процесса (см. (2.80)).

Пример. Определить алгоритм моделирования стационарного нормального случайного процесса, корреляционная функция которого имеет вид

$$R(\tau) = \frac{\sigma^2}{1 + \omega_0^2 \tau^2}.$$

Находим спектральную плотность мощности случайного процесса:

$$G(\omega) = \int_{-\infty}^{\infty} R(\tau) \exp(-j\omega\tau) d\tau = \frac{\pi\sigma^2}{\omega_0} \exp\left(-\left|\frac{\omega}{\omega_0}\right|\right).$$

Подставляя $G(\omega)$ в формулу (2.84), получаем

$$c_k = \frac{\sigma}{\sqrt{\omega_c \omega_0}} \int_0^{\omega_c} \exp\left(-\frac{\omega}{2\omega_0}\right) \cos\left(\frac{k\pi\omega}{\omega_c}\right) d\omega. \quad (2.86)$$

Интеграл в выражении (2.86) является табличным:

$$\int \exp(ax) \cos(nx) dx = \frac{\exp(ax)}{a^2 + n^2} [a \cos(nx) + n \sin(nx)].$$

После несложных преобразований находим

$$c_k = 2\sigma \sqrt{\frac{\gamma}{\pi}} \left(\frac{1 - (-1)^{|k|} \exp(-\pi/2\gamma)}{1 + 4\gamma^2 k^2} \right),$$

где $\gamma = \omega_0 \Delta t$, $\Delta t = \pi/\omega_c$.

При достаточно малом значении γ ($\gamma \ll 1$) имеем

$$c_k \approx 2\sigma \sqrt{\frac{\gamma}{\pi}} \left(\frac{1}{1 + 4\gamma^2 k^2} \right).$$

Алгоритм моделирования будет иметь вид (2.84).

Способ нахождения весовых коэффициентов на основе разложения спектральной плотности мощности процесса на множители. Этот метод применяется, когда спектральная плотность мощности $G(\omega)$ моделируемого стационарного случайного процесса является рациональной функцией, т. е. представляет собой отношение

$$G(\omega) = G_1(\omega)/G_2(\omega),$$

где $G_1(\omega)$ и $G_2(\omega)$ — полиномы степени f' и $m' > f'$ соответственно.

Случайные процессы с рациональной спектральной плотностью наблюдаются, как известно, на выходе линейной системы с постоянными параметрами при воздействии на входе белого шума. Действительно, комплексная частотная характеристика $K(j\omega)$ такой системы является дробно-рациональной функцией:

$$K(j\omega) = K_1(j\omega)/K_2(j\omega),$$

где $K_1(j\omega)$ и $K_2(j\omega)$ — полиномы по переменной $j\omega$ степени f и $m > f$ соответственно. При этом на выходе система при воздействии белого шума с единичной спектральной плотностью мощности ($G_0 = 1$) будет появляться случайный процесс со спектральной плотностью мощности

$$G(\omega) = |K(j\omega)|^2 = K(j\omega)K(-j\omega) = \frac{K_1(j\omega)K_1(-j\omega)}{K_2(j\omega)K_2(-j\omega)}. \quad (2.87)$$

Таким образом, линейную систему с соответствующей характеристикой $K(j\omega)$ можно использовать в качестве формирующего фильтра для получения случайного процесса $\xi(t)$ с заданной спектральной плотностью $G(\omega)$.

Определим комплексную частотную характеристику $K(j\omega)$. Это можно сделать путем разложения $G(\omega)$ на множители вида (2.87). На основании теоремы разложения неотрицательных дробно-рациональных функций на множители комплексную частотную характеристику $K(j\omega)$ можно представить в следующем виде:

$$K(j\omega) = \sqrt{C} \frac{\prod_{k=1}^f (j\omega - j\omega_{1k})}{\prod_{k=1}^m (j\omega - j\omega_{2k})}, \quad (2.88)$$

где ω_{1k} и ω_{2k} — корни соответственно числителя $G_1(\omega)$ и знаменателя $G_2(\omega)$ спектральной плотности $G(\omega)$ с положительной мнимой частью.

Множитель C определяется из условия

$$|K(j\omega)|^2 = G(\omega). \quad (2.89)$$

На основании (2.88) передаточная функция фильтра имеет вид

$$K(p) = \sqrt{C} \frac{\prod_{k=1}^f (p - p_{1k})}{\prod_{k=1}^m (p - p_{2k})},$$

где $p = j\omega$, $p_{1k} = j\omega_{1k}$, $p_{2k} = j\omega_{2k}$.

По найденной передаточной функции $K(p)$ можно найти импульсную характеристику формирующего фильтра $h(t)$, которая согласно известной теореме разложения Хевисайда имеет вид

$$h(t) = \sum_{v=1}^s \sum_{\mu=1}^{r_v} H_{v\mu} t^{r_v-\mu} \exp(p_v t), \quad (2.90)$$

где p_v , $v=1, 2, \dots, s$, — полюсы передаточной функции $K(p)$ с кратностями r_1, r_2, \dots, r_s соответственно, s — число различных корней ($r_1 + r_2 + \dots + r_s = m$),

$$H_{\nu\mu} = \frac{1}{(\mu-1)!(r_\nu-\mu)!} \frac{d^{\mu-1}}{dp^{\mu-1}} \left[(p-p_\nu)^{r_\nu} K(p) \right]_{p=p_\nu}. \quad (2.91)$$

Случайный процесс $\xi(t)$ на выходе формирующего фильтра с импульсной характеристикой $h(t)$ определяется интегралом свертки:

$$\xi(t) = \int_0^\infty h(\tau) x'(t-\tau) d\tau, \quad (2.92)$$

где $x'(t)$ — белый шум со спектральной плотностью мощности $G_0 = 1$.

Осуществляя дискретизацию процесса (2.92), получаем

$$\xi[n] = \Delta t \sqrt{D_{x'}} \sum_{k=0}^{\infty} h[k] x[n-k],$$

где $D_{x'}$ — дисперсия процесса $x'(t)$, $x[n]$ — нормированный дискретный белый шум.

Учитывая, что $\Delta t = \pi/\omega_c$, $D_{x'} = \omega_c/\pi$, и ограничивая количество членов суммы, окончательно можно записать

$$\xi[n] = \sum_{k=0}^{N-1} c_k x[n-k], \quad (2.93)$$

где

$$c_k = \sqrt{\Delta t} h[k]. \quad (2.94)$$

Пример. Определить алгоритм для моделирования случайного процесса с корреляционной функцией вида $R(\tau) = \exp(-\omega_0 |\tau|)$.

Находим спектральную плотность мощности моделируемого случайного процесса:

$$G(\omega) = \frac{2\omega_0}{\omega_0^2 + \omega^2}$$

и корень знаменателя $G(\omega)$ с положительной мнимой частью:

$$\omega_{21} = j\omega_0.$$

Комплексная частотная характеристика $K(j\omega)$ формирующего фильтра согласно выражению (2.88) имеет вид

$$K(j\omega) = \sqrt{C} \frac{1}{j\omega + \omega_0},$$

а передаточная функция

$$K(p) = \sqrt{C} \frac{1}{p + \omega_0}.$$

Из условия (2.89) получаем $\sqrt{C} = \sqrt{2\omega_0}$.

По формулам (2.90) и (2.91) при значениях $s = r_1 = 1$, $p_1 = -\omega_0$ находим импульсную характеристику формирующего фильтра:

$$h(t) = \sqrt{2\omega_0} \exp(-\omega_0 t).$$

Для дискретных значений $t = k\Delta t$ получаем

$$h[k] = \sqrt{2\omega_0} \exp(-\omega_0 k\Delta t).$$

По формуле (2.94) находим

$$c_k = \sqrt{\Delta t} \sqrt{2\omega_0} \exp(-\omega_0 k\Delta t).$$

Окончательно алгоритм моделирования (2.93) примет вид

$$\xi[n] = \sqrt{2\gamma} \sum_{k=0}^{N-1} \exp(-\gamma k) x[n-k], \quad \gamma = \omega_0 \Delta t. \quad (2.95)$$

Если при моделировании гауссовского случайного процесса $\xi[n]$ известно, что он является результатом воздействия белого шума на линейную систему с известной импульсной характеристикой $h(t)$, то данную линейную систему целесообразно использовать как формирующий фильтр. В этом случае весовые коэффициенты c_k алгоритма моделирования (2.93) будут определяться через дискретные значения импульсной характеристики $h[k]$ по формуле (2.94).

Метод рекуррентных алгоритмов. При этом методе значения $\xi[n]$ моделируемого случайного процесса $\xi(t)$ формируются на основании следующего алгоритма:

$$\begin{aligned} \xi[n] = & a_0 x[n] + a_1 x[n-1] + \dots + a_s x[n-s] - \\ & - b_1 \xi[n-1] - b_2 \xi[n-2] - \dots - b_m \xi[n-m] = \sum_{k=0}^s a_k x[n-k] - \sum_{k=1}^m b_k \xi[n-k], \end{aligned} \quad (2.96)$$

где $x[n]$ — нормированный дискретный белый шум. Параметры a_k и b_k определяют вид корреляционной функции случайного процесса, моделируемого с помощью алгоритма (2.96).

Уравнение (2.96) описывает поведение дискретного линейного фильтра, имеющего передаточную функцию

$$K(z) = \frac{\sum_{k=0}^s a_k z^{-k}}{1 + \sum_{k=1}^m b_k z^{-k}}. \quad (2.97)$$

Рекуррентные алгоритмы применимы только для моделирования случайных процессов с рациональным спектром.

Как и при использовании алгоритма скользящего суммирования, подготовительная работа сводится к нахождению параметров a_k и b_k . Существуют различные способы определения этих параметров. Рассмотрим один из них — способ факторизации.

Последовательность действий при осуществлении этого способа следующая:

1) находится спектральная плотность $F(z)$ моделируемого процесса $\xi[n]$ по корреляционной функции $R[k]$:

$$F(z) = \sum_{k=-\infty}^{\infty} R[k] z^{-k}; \quad (2.98)$$

2) осуществляется факторизация спектральной плотности $F(z)$:

$$F(z) = \frac{A(z)A(z^{-1})}{B(z)B(z^{-1})} = |K(z)|^2;$$

3) преобразуется передаточная функция $K(z)$ к виду (2.97) для нахождения параметров рекуррентного алгоритма (2.96).

Пример. Найти рекуррентный алгоритм моделирования гауссовского стационарного процесса $\xi(t)$, имеющего экспоненциальную корреляционную функцию

$$R(\tau) = \exp(-\omega_0 |\tau|).$$

Формирование такого процесса методом скользящего суммирования рассмотрено в предыдущем примере. Его спектральная плотность мощности имеет дробно-рациональный вид:

$$G(\omega) = \frac{2\omega_0}{\omega_0^2 + \omega^2}.$$

По заданной непрерывной корреляционной функции $R(\tau)$ находим ее дискретную модель:

$$R[k] = \exp(-\gamma|k|),$$

где $\gamma = \omega_0 \Delta t$, Δt — интервал дискретизации, равный π/ω_c .

Спектральную плотность мощности $F(z)$ моделируемого процесса $\xi(t)$ находим по формуле (2.98). Для вычисления двустороннего z -преобразования используем известное соотношение

$$F(z) = F^+(z) + F^+(z^{-1}) - R[0], \quad (2.99)$$

где $F^+(z) = \sum_{k=0}^{\infty} R[k] z^{-k}$ — одностороннее z -преобразование дискретной функции.

По таблицам [20] находим $F^+(z)$:

$$F^+(z) = \sum_{k=0}^{\infty} \exp(-\gamma k) z^{-k} = \frac{z}{z - \exp(-\gamma)}.$$

Подставляя это выражение в (2.99), после ряда преобразований получаем

$$F(z) = \frac{A(z)A(z^{-1})}{B(z)B(z^{-1})} = \frac{(\sqrt{1-\rho^2})^2}{(1-z\rho)(1-z^{-1}\rho)} = |K(z)|^2,$$

где $\rho = \exp(-\gamma)$.

Тогда дискретную передаточную функцию $K(z)$ можно представить в виде

$$K(z) = \frac{\sqrt{1-\rho^2}}{1-z^{-1}\rho}. \quad (2.100)$$

Из выражения (2.100) получаем следующие значения параметров рекуррентного алгоритма:

$$a_0 = \sqrt{1-\rho^2}, \quad b_1 = -\rho.$$

Окончательно рекуррентный алгоритм (2.96) принимает вид

$$\xi[n] = \sqrt{1-\rho^2} x[n] + \rho \xi[n-1], \quad (2.101)$$

где $\rho = \exp(-\gamma)$, $\gamma = \omega_0 \Delta t$.

В табл. 2.5 приведены алгоритмы моделирования стационарных гауссовских случайных процессов с типовыми корреляционными функциями.

Таблица 2.5

Алгоритмы моделирования гауссовских случайных процессов $\xi[n]$

Корреляционная функция $R(\tau)$	Спектральная плотность мощности $G(\omega)$	Алгоритм моделирования	Параметры алгоритма
$\sigma^2(1 - \omega_0 \tau)$ при $ \tau \leq 1/\omega_0$ 0 при $ \tau \geq 1/\omega_0$	$\frac{\sigma^2 \sin^2(\omega/2\omega_0)}{\omega_0 (\omega/2\omega_0)^2}$	$c_0 \sum_{k=0}^{N-1} x[n-k]$	$c_0 = \sigma/\sqrt{N}$ $N = [1/\gamma] + 1$ $[\delta]$ — целая часть числа δ , $\gamma = \omega_0 \Delta t$
$\frac{\sigma^2}{1 + \omega_0^2 \tau^2}$	$\frac{\pi \sigma^2}{\omega_0} \exp\left(-\left \frac{\omega}{\omega_0}\right \right)$	$\sum_{k=-p}^p c_k x[n-k]$	$c_k = 2\sigma \sqrt{\frac{\gamma}{\pi}} \frac{1}{1 + 4\gamma^2 k^2}$ $\gamma \leq 1/2$, $\gamma = \omega_0 \Delta t$
$\sigma^2 \exp(-\omega_0 \tau)^2$	$\frac{\sqrt{\pi} \sigma^2}{\omega_0} \exp\left(-\frac{\omega}{2\omega_0}\right)^2$	$\sum_{k=-p}^p c_k x[n-k]$	$c_k = \frac{\sigma \sqrt{2\gamma}}{\sqrt[4]{\pi}} \exp(-2\gamma^2 k^2)$ $\gamma \leq 1/2$, $\gamma = \omega_0 \Delta t$
$\sigma^2 \exp(- \omega_0 \tau)$	$\frac{2\sigma^2 \omega_0}{\omega^2 + \omega_0^2}$	$\sum_{k=0}^{N-1} c_k x[n-k]$	$c_k = \sqrt{2\gamma} \exp(-\gamma k)$ $\gamma = \omega_0 \Delta t$
$\sigma^2 \exp(- \omega_0 \tau)$	$\frac{2\sigma^2 \omega_0}{\omega^2 + \omega_0^2}$	$a_0 x[n] + b_1 [n-1]$	$a_0 = \sqrt{1 - \exp(-2\gamma)}$ $b_1 = \exp(-\gamma)$, $\gamma = \omega_0 \Delta t$

Примечание. σ^2 — дисперсия процесса, Δt — интервал дискретизации

2.8.4. Моделирование негауссовских случайных процессов

Рассмотрим метод моделирования случайных процессов, которые не являются гауссовскими, но порождены ими в нелинейных системах. Найдем алгоритм моделирования случайного негауссовского процесса, имеющего корреляционную функцию $R(\tau)$ и одномерную плотность распределения вероятностей $w(\xi)$.

Последовательность действий при решении этой задачи в общем виде следующая:

1) находят такое нелинейное безынерционное преобразование $y = f(x)$, которое преобразует гауссовский процесс $\xi_0(t)$ в процесс $\xi(t)$ с заданным законом распределения $w(\xi)$;

2) определяют по найденной функции $y = f(x)$ зависимость корреляционной функции $R(\tau)$ полученного процесса $\xi(t)$ от корреляционной функции $R_0(\tau)$ исходного гауссовского процесса $\xi_0(t)$:

$$R(\tau) = \varphi[R_0(\tau)];$$

3) получают корреляционную функцию исходного гауссовского процесса:

$$R_0(\tau) = \varphi^{-1}[R(\tau)],$$

где φ^{-1} — функция, обратная функции φ ;

4) находят алгоритм для моделирования гауссовского процесса $\xi_0(t)$, соответствующего требуемой корреляционной функции $R_0(\tau)$.

Моделирование рэлеевского случайного процесса. Одномерная функция плотности распределения вероятности такого процесса имеет вид

$$w(\xi) = \frac{\xi}{\sigma_0^2} \exp\left(-\frac{\xi^2}{2\sigma_0^2}\right), \quad \xi \geq 0,$$

где σ_0 — параметр распределения.

Известно [8, 19], что рэлеевский процесс $\xi(t)$ выражается через два независимых стационарных гауссовских случайных процесса $\xi_{10}(t)$ и $\xi_{20}(t)$ с параметрами $(0, 1)$:

$$\xi(t) = \sigma_0 \sqrt{\xi_{10}^2(t) + \xi_{20}^2(t)}. \quad (2.102)$$

При этом нормированная корреляционная функция рэлеевского процесса $r(\tau) = R(\tau)/R(0)$ связана с нормированной корреляционной функцией $r_0(\tau) = R_0(\tau)/R_0(0)$ процессов $\xi_{10}(t)$ и $\xi_{20}(t)$ следующей зависимостью [8]:

$$r(\tau) = \frac{\pi}{4(4-\pi)} \left[r_0^2(\tau) + \sum_{n=2}^{\infty} \frac{[(2n-3)!!]^2}{2^{2n-1}(n!)^2} r_0^{2n}(\tau) \right], \quad (2.103)$$

где $(2n-3)!!$ — произведение всех нечетных чисел натурального ряда от 1 до $2n-3$ включительно.

Из выражения (2.103) получим приближенную зависимость

$$r(\tau) \approx r_0^2(\tau),$$

на основании которой определим $r_0(\tau)$:

$$r_0(\tau) \approx \sqrt{r(\tau)}. \quad (2.104)$$

Таким образом, алгоритм моделирования рэлеевского случайного процесса с учетом формулы (2.102) имеет вид

$$\xi[n] = \sigma_0 \sqrt{\xi_{10}^2[n] + \xi_{20}^2[n]}, \quad (2.105)$$

где $\xi_{10}[n]$ и $\xi_{20}[n]$ — дискретные значения независимых гауссовских процессов с параметрами $(0, 1)$ и с нормированной корреляционной функцией $r_0(\tau)$, определяемой выражением (2.104).

Пример. Необходимо определить алгоритм моделирования рэлеевского случайного процесса, имеющего экспоненциальную корреляционную функцию

$$r(\tau) = \exp(-\omega_0 |\tau|). \quad (2.106)$$

Подставляя выражение (2.106) в (2.104), находим нормированную корреляционную функцию исходных гауссовских случайных процессов $\xi_{10}[n]$ и $\xi_{20}[n]$:

$$r_0(\tau) = \exp\left(-\frac{\omega_0 |\tau|}{2}\right). \quad (2.107)$$

Используя рекуррентный алгоритм (2.101) моделирования гауссовского случайного процесса с экспоненциальной корреляционной функцией вида (2.107), получаем

$$\begin{aligned} \xi_{10}[n] &= \sqrt{1-\rho^2} x_1[n] + \rho \xi_{10}[n-1], \\ \xi_{20}[n] &= \sqrt{1-\rho^2} x_2[n] + \rho \xi_{20}[n-1], \end{aligned} \quad (2.108)$$

где $\rho = \exp(-\omega_0 \Delta t / 2)$, $x_1[n]$ и $x_2[n]$ — независимые значения нормированного дискретного белого шума.

Подставляя $\xi_{10}[n]$ и $\xi_{20}[n]$ из (2.108) в формулу (2.105), получаем следующий алгоритм моделирования рэлеевского случайного процесса:

$$\xi[n] = \sigma_0 \sqrt{\left(\sqrt{1-\rho^2} x_1[n] + \rho \xi_1[n-1]\right)^2 + \left(\sqrt{1-\rho^2} x_2[n] + \rho \xi_2[n-1]\right)^2}.$$

Моделирование случайного процесса с показательным законом распределения. Одномерная функция плотности распределения вероятностей этого процесса имеет вид

$$w(\xi) = \frac{1}{2\sigma_0^2} \exp\left(-\frac{\xi}{2\sigma_0^2}\right), \quad \xi \geq 0,$$

где σ_0 — параметр распределения.

Показательный процесс можно представить как квадрат рэлеевского случайного процесса [8, 19] или, с учетом выражения (2.102), — как сумму квадратов двух одинаковых независимых стационарных гауссовских случайных процессов $\xi_{10}(t)$ и $\xi_{20}(t)$ с параметрами (0,1):

$$\xi(t) = \sigma_0^2 [\xi_{10}^2(t) + \xi_{20}^2(t)]. \quad (2.109)$$

Нормированная корреляционная функция $r(\tau)$ показательного случайного процесса выражается через нормированную корреляционную функцию $r_0(\tau)$ процессов $\xi_{10}(t)$ и $\xi_{20}(t)$ следующим образом [8]:

$$r(\tau) = r_0^2(\tau).$$

Из последнего выражения получаем

$$r_0(\tau) = \sqrt{r(\tau)}. \quad (2.110)$$

Таким образом, алгоритм моделирования показательного случайного процесса с учетом формулы (2.109) имеет вид

$$\xi[n] = \sigma_0^2 (\xi_{10}^2[n] + \xi_{20}^2[n]), \quad (2.111)$$

где $\xi_{10}[n]$ и $\xi_{20}[n]$ — дискретные значения независимых гауссовских процессов с параметрами (0, 1) и с нормированной корреляционной функцией $r_0(\tau)$, определяемой выражением (2.110).

Пример. Определить алгоритм моделирования показательного случайного процесса с экспоненциальной корреляционной функцией (2.106).

Подставляя $\xi_{10}[n]$ и $\xi_{20}[n]$ из соотношений (2.108) в (2.111), получаем следующий алгоритм моделирования показательного случайного процесса:

$$\xi[n] = \sigma_0^2 \left(\left(\sqrt{1-\rho^2} x_1[n] + \rho \xi_{10}[n-1] \right)^2 + \left(\sqrt{1-\rho^2} x_2[n] + \rho \xi_{20}[n-1] \right)^2 \right),$$

где $\rho = \exp(-\omega_0 \Delta t / 2)$, $x_1[n]$ и $x_2[n]$ — независимые значения нормированного дискретного белого шума.

Контрольные вопросы

1. Что такое информации, сообщение, сигнал?
2. Приведите примеры моделей дискретных сообщений.
3. Как задается модель непрерывных сообщений?
4. Дайте определения основных моментных функций случайного процесса: математического ожидания, дисперсии и корреляционной функции. Поясните их физический смысл.
5. Что такое спектральная плотность мощности случайного процесса?
6. Какие случайные процессы называются эргодическими?
7. Дайте определение пространства сигналов.
8. Что такое расстояние между сигналами, метрика, норма, скалярное произведение?
9. Какое пространство называется евклидовым? Какое пространство называется гильбертовым?
10. В чем особенность векторного представления случайного процесса?
11. Какое разложение случайного процесса называется каноническим?
12. Назовите основные способы дискретного представления непрерывных сообщений.
13. В чем суть дискретного временного представления непрерывных сообщений?
14. Сформулируйте теорему Котельникова для случайных процессов.
15. Как определяется погрешность дискретизации при наличии предварительной фильтрации? Является ли целесообразной предварительная фильтрация?
16. В чем суть адаптивной дискретизации сигналов? Когда она применяется?
17. Что такое дискретное обобщенное представление сигналов?
18. Поясните дискретное разностное представление сигналов.
19. Что такое равномерное квантование сигналов? Как определяется дисперсия погрешности при равномерном квантовании?
20. В чем суть неравномерного квантования? Как определяются пороги и уровни квантования? Как технически реализуется неравномерное квантование?
21. Когда применяется неравномерное квантование?
22. Какие коды используются для передачи уровней квантования?
23. Что такое линейная рекуррентная последовательность? М-последовательность?
24. Каким условиям должен удовлетворять характеристический многочлен М-последовательности?
25. Укажите основные структурные свойства М-последовательности.
26. Укажите основные корреляционные свойства М-последовательности.
27. Укажите свойства сегментов М-последовательности.
28. Как образуются последовательности Голда? Каковы их корреляционные свойства?

29. Как образуются последовательности Касами? Каковы их корреляционные свойства?
30. Дайте определение основных систем сигналов.
31. Что такое симплексный код? Приведите способы построения симплексных кодов.
32. Приведите способы построения ортогональных кодов.
33. Назовите основные методы моделирования случайных величин.
34. Поясните алгоритм моделирования дискретных случайных величин.
35. Сравните два основных метода моделирования случайных векторов с заданной плотностью распределения вероятностей.
36. Сравните два основных метода моделирования стационарных гауссовских случайных процессов.
37. Приведите способы нахождения весовых коэффициентов алгоритма скользящего суммирования для моделирования стационарных гауссовских случайных процессов.

3. ОСНОВЫ ТЕОРИИ ОБНАРУЖЕНИЯ И РАЗЛИЧЕНИЯ СИГНАЛОВ

Изложены основы теории статистических решений. Рассмотрены задачи обнаружения и различения сигналов. Приведены оптимальные алгоритмы обработки и структурные схемы обнаружителей и различителей сигналов, дана оценка помехоустойчивости устройств обработки. Рассмотрены модели сигналов и помех, наиболее часто встречающиеся при разработке радиотехнических систем различного назначения.

3.1. Общие сведения

В любой радиотехнической системе (РТС) сигнал при передаче искажается и сообщение на выходе приемника воспроизводится с некоторой ошибкой. Искажения сигнала обусловлены как случайными изменениями характеристик каналов, так и помехами, действующими в них. Частотные и временные характеристики канала определяют так называемые *линейные искажения*. Кроме того, канал может вносить *нелинейные искажения* при нелинейности тех или иных звеньев. Поскольку линейные и нелинейные искажения сигнала обусловлены известными характеристиками канала, то они могут быть устранены соответствующей их коррекцией. Искажения от помех носят случайный характер и поэтому не могут быть полностью устранены. В правильно спроектированной системе искажения сообщений обусловлены лишь воздействием помех.

Пусть приемник при отсутствии помех точно воспроизводит сообщение, содержащееся в полезном сигнале. При помехах сообщение будет воспроизводиться с некоторой погрешностью. Приемник, обеспечивающий минимальные искажения сообщения, называется *оптимальным*, или *идеальным*. Критерии оптимальности и количественные характеристики искажений зависят от назначения приемника. Минимальный уровень искажений при выбранном критерии характеризует потенциальную помехоустойчи-

вость. При заданных условиях приема помехоустойчивость реального приемника не может превзойти потенциальную.

Решение проблем, связанных с синтезом оптимальных устройств обработки сигналов, — основное направление теории оптимальных методов приема. При некоторых ограничениях, налагаемых на характеристики сигналов и помех, эта теория позволяет найти оптимальный алгоритм работы радиоприемного устройства и, соответственно, его структуру, оценить показатели синтезируемого устройства, определить наилучшие виды передаваемых сигналов, а также выяснить степень технического совершенства реальных приемников и возможные пути повышения их помехоустойчивости.

Основную задачу теории оптимального приема сигналов можно сформулировать следующим образом: предполагая заранее известными некоторые характеристики передаваемого сигнала, канала и помех, а также их функциональное взаимодействие, необходимо синтезировать оптимальное приемное устройство, которое бы наилучшим образом (в смысле выбранного критерия) воспроизводило сообщение, содержащееся в сигнале на входе приемника, или принимало решение с наименьшими ошибками [21].

При синтезе оптимальных приемных устройств важными являются: выбор математически обоснованного критерия оптимальности в соответствии с физическим смыслом и целевым содержанием решаемой задачи и четкая математическая формулировка задачи, включающая все априорные сведения и позволяющая решать ее в соответствии с выбранным критерием. В результате синтеза оптимальных приемных устройств должны быть найдены оптимальный алгоритм работы приемного устройства и, соответственно, структурная или функциональная схема приемника, определены параметры схемы. Отметим, что синтез радиотехнических устройств не исключает необходимости их анализа, целью которого является вычисление показателей качества работы приемника, выяснение степени чувствительности полученных алгоритмов и количественных характеристик к отклонениям от принятых априорных данных и т. п.

В зависимости от целевого назначения РТС функционируют при различных условиях и к ним предъявляются разные требования. Исходя из этих требований, для типовых систем можно сформулировать ряд частных задач, рассматриваемых в теории оптимальных методов радиоприема.

Пусть колебание на входе приемника имеет вид

$$u(t) = F(s(t, \lambda), n(t)), \quad (3.1)$$

где $F(\cdot)$ — некоторый оператор; $s(t, \lambda)$ — полезный сигнал; $\lambda = (\lambda_1, \dots, \lambda_m)$ — вектор параметров, от которых зависит сигнал; $n(t)$ — помеха. Априори считают известными: способ взаимодействия сигнала и помехи (оператор $F(\cdot)$),

функцию, описывающую полезный сигнал, а также все необходимые для решения задачи вероятностные характеристики векторной случайной величины λ и помехи $n(t)$.

В принятой модели полезного сигнала некоторые параметры могут быть заранее известны. Они составляют априорные сведения о полезном сигнале. Другие параметры сигнала неизвестны. Некоторые из них являются носителями информации и называются *информационными*, а остальные — *неинформационными*.

Далее будем рассматривать часто встречающийся на практике частный вид оператора $F(\cdot)$ из формулы (3.1), когда

$$u(t) = s(t, \lambda) + n(t). \quad (3.2)$$

Для колебания, представленного соотношением (3.2), можно сформулировать следующие основные задачи теории оптимальных методов приема сигналов, две из которых (задачи обнаружения и различения) будут рассмотрены в настоящей главе, а остальные — в последующих.

Обнаружение сигнала. Запишем выражение (3.2) в виде

$$u(t) = \theta s(t, \lambda) + n(t),$$

где θ — случайная величина, принимающая два значения: $\theta = 0$ (полезный сигнал отсутствует) и $\theta = 1$ (полезный сигнал присутствует). Задача обнаружения заключается в том, чтобы определить, присутствует или отсутствует сигнал в принятом колебании $u(t)$, или, что то же самое, в оценке значения параметра θ . Сформулированная задача является характерной для радиолокации. Она встречается также в двоичных системах передачи информации с пассивной паузой и др.

Различение сигналов. Пусть колебание $u(t)$ представляет собой сумму одного из m возможных полезных сигналов $s_i(t, \lambda)$, $i = 1, 2, \dots, m$, и помехи $n(t)$. Зная вероятности $p(s_i)$, $i = 1, 2, \dots, m$, наличия сигналов $s_i(t, \lambda)$, $i = 1, 2, \dots, m$, на входе приемника, необходимо по принятой реализации $u(t)$ оптимальным образом решить, какой из полезных сигналов содержится в ней. Задача характерна для систем передачи информации.

Оценка параметров сигнала. Пусть один из параметров сигнала $s(t, \lambda)$, является неизвестным. Необходимо наилучшим образом (в смысле выбранного критерия) оценить значение этого параметра по принятой реализации $u(t)$. Задача характерна для радиолокации, радионавигации, телеметрии, радиотелеуправления.

Заметим, что часто приходится решать задачу совместной оценки нескольких параметров сигнала, например времени появления сигнала и частоты несущей.

Фильтрация сообщений. Пусть информационный параметр λ_i полезного сигнала $s(t, \lambda_i)$ зависит от времени и представляет собой случайный процесс $\lambda_i(t)$ с известными статистическими характеристиками. Необходимо наилучшим образом выделить реализацию сообщения $\lambda_i(t)$, содержащуюся в колебании $u(t)$. Эта задача встречается в радиосвязи, телевидении, радиолокации.

К задачам теории оптимальных методов приема относится также задача разрешения сигналов, которую можно сформулировать следующим образом: на входе приемного устройства присутствуют два или более сигналов $s(t, \lambda)$, $i = 1, 2, \dots$; необходимо по принятой реализации $u(t)$ оптимальным образом разрешать эти сигналы по параметрам λ . Под термином «разрешить» понимают либо раздельное обнаружение сигналов, либо раздельное обнаружение и определение параметров λ во всех сигналах $s_i(t)$. Задача встречается в радиолокации, связи и др.

Разделение оптимальных методов радиоприема на указанные типы задач в значительной степени условно. Между ними нельзя провести четких границ. Например, задачу обнаружения можно трактовать как частный случай задачи различения двух сигналов, когда один из них тождественно равен нулю. Задачу обнаружения можно рассматривать так же, как частный случай измерения амплитуды сигнала, принимающей два значения: 0 и A . Задача оценки параметра сигнала является частным случаем задачи фильтрации сообщений. Тем не менее, приведенная классификация задач в методическом плане целесообразна.

В заключение отметим, что решение основных задач теории оптимальных методов радиоприема базируется на хорошо разработанных методах математической статистики.

3.2. Основные положения теории статистических решений.

Оптимальные критерии

Как уже упоминалось в § 3.1, при синтезе оптимальных устройств обработки сигналов необходим выбор критерия качества работы устройства. Поскольку теория статистических решений охватывает все многообразие статистических оптимальных критериев, целесообразно при решении задач синтеза оптимальных радиоприемных устройств воспользоваться результатами этой теории. Ниже излагаются основные положения теории статистических решений, рассмотренных в работах [22—26].

Пусть принятый сигнал имеет вид (3.1). Не нарушая общности, положим, что наблюдение колебания $u(t)$ проводится в дискретные моменты времени t_1, \dots, t_m . Тогда с учетом формулы (3.1) можно записать

$$\mathbf{u} = \mathbf{F}(\mathbf{s}, \mathbf{n}), \quad (3.3)$$

где $\mathbf{u} = (u_1, \dots, u_m)$, $\mathbf{s} = (s_1, \dots, s_m)$, $\mathbf{n} = (n_1, \dots, n_m)$ — m -мерные векторы (m -мерные выборки) соответственно принятого колебания, полезного сигнала и помехи.

Совокупность всех возможных векторов \mathbf{u} образует пространство U выборок принимаемого колебания $u(t)$. Аналогично, векторы \mathbf{s} и \mathbf{n} образуют соответственно пространство S выборок полезного сигнала $s(t)$ и пространство N выборок помехи $n(t)$.

Статистические характеристики помехи предполагаются известными и задаются в виде распределения $w(\mathbf{n}) = w_m(n_1, \dots, n_m)$. Известным считается также способ взаимодействия сигнала с помехой. При этом можно найти статистическое описание вектора \mathbf{u} принятого колебания для фиксированного вектора \mathbf{s} — условное распределение $w(\mathbf{u}|\mathbf{s}) = w_m(u_1, \dots, u_m|\mathbf{s})$.

Функцию $w(\mathbf{u}|\mathbf{s})$ называют *функцией правдоподобия*. Такое название объясняется следующим. После получения выборки u_1, \dots, u_m функция $w(\mathbf{u}|\mathbf{s})$ зависит только от \mathbf{s} и характеризует степень соответствия вектора принятого колебания тому или иному вектору полезного сигнала, т. е. показывает, насколько один вектор \mathbf{s} при известной выборке u_1, \dots, u_m более правдоподобен, чем другие.

При известном распределении $w(\mathbf{s}) = w_m(s_1, \dots, s_m)$ полезных сигналов, зная функцию $w(\mathbf{u}|\mathbf{s})$, можно найти совместное распределение векторов \mathbf{u} принимаемых колебаний и векторов \mathbf{s} полезных сигналов:

$$w(\mathbf{u}, \mathbf{s}) = w(\mathbf{s})w(\mathbf{u}|\mathbf{s}).$$

Задача заключается в том, чтобы на основе полученной выборки u_1, \dots, u_m и априорных данных о способе взаимодействия сигнала и помехи принять одно вполне определенное решение из набора $\gamma_0, \gamma_1, \dots, \gamma_k$ возможных образующих пространство решений Γ .

Выбор решения по выборке u_1, \dots, u_m проводится в соответствии с алгоритмом принятия решения $\Delta(\gamma|\mathbf{u})$. Функция $\Delta(\gamma|\mathbf{u})$ называется *решающей функцией*, или *решающим правилом*. Она представляет собой вероятности

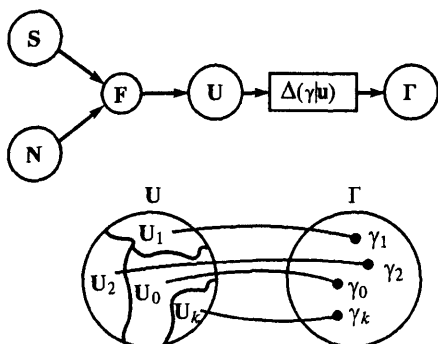


Рис. 3.1. Схема выбора решения

(или плотность вероятностей) принятия решения γ на основе принятых данных u . Механизм принятия решения можно представить следующим образом (рис. 3.1). Пространство принимаемых сигналов U разбивается на непересекающиеся области U_0, U_1, \dots, U_k и устанавливается соответствие между решениями и этими областями. Если каждой области ставится в соответствие только одно вполне определенное решение, то правило $\Delta(\gamma|u)$ называется

нерандомизированным (детерминированным).

Существуют правила, при которых для принятой выборки u_1, \dots, u_m допускается несколько решений с соответствующими вероятностями, например некоторой области U_i ставятся в соответствие решения $\gamma_{i1}, \dots, \gamma_{ik}$ с

вероятностями p_{i1}, \dots, p_{ik} соответственно, причем $\sum_{j=1}^k p_{ij} = 1$. Такие правила

называются *рандомизированными*.

Далее рассматриваются только нерандомизированные правила. Будем считать, что если выборка принятого колебания попадает в область U_i , то принимается решение γ_i . Соответственно, решающая функция имеет вид

$$\Delta(\gamma_i | u_j) = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j. \end{cases}$$

Очевидно, что для любого решающего правила при наличии помех всегда возможны ошибочные решения. Для количественной оценки ущерба (потерь), связанных с принятием решений, вводится так называемая *функция потерь* (штрафа, стоимости) $\Pi(s, \gamma)$. Ее конкретное значение $\Pi(s_i, \gamma_j)$ характеризует потери при принятии решения γ_j , в то время как правильным является решение γ_i . Функция потерь должна удовлетворять следующим свойствам:

$$\Pi(s_i, \gamma_j) > 0, \quad \Pi(s_i, \gamma_j) > \Pi(s_i, \gamma_i), \quad \Pi(s_i, \gamma_i) \leq 0.$$

Функция потерь выбирается заранее. Примеры ее задания будут приведены при решении конкретных задач.

Теперь можно сформулировать математически задачу выбора решения: на основе априорных данных о пространствах полезных сигналов S и помех N , распределениях вероятностей $w(s)$ и $w(n)$ на этих пространствах соответственно, способе взаимодействия сигнала и помехи и заданной функции потерь $\Pi(s, \gamma)$ необходимо по полученному сигналу u оптимальным образом найти решение γ — какой конкретно из полезных сигналов содержится на входе приемника.

Поскольку появление того или иного полезного сигнала s_i на входе приемника и принятие решения γ_j являются случайными событиями, то значение функции потерь $\Pi(s_i, \gamma_j)$ — случайная величина. Поэтому качество решения можно характеризовать математическим ожиданием функции потерь

$$R = M\{\Pi(s, \gamma)\} = \sum_j \sum_i w(s_i, \gamma_j) \Pi(s_i, \gamma_j), \quad (3.4)$$

где $w(s_i, \gamma_j)$ — совместная вероятность появления на входе приемника сигнала s_i и принятия решения γ_j . Величина R характеризует средние потери при принятии решения и называется *средним риском*. Чем меньше средний риск, тем лучше решение (наилучшим (оптимальным) решающим правилом будет такое, для которого значение среднего риска будет наименьшим). Правило, при котором минимизируется средний риск, называется *байесовским правилом*, или *байесовским критерием*. Часто его также называют *критерием минимума среднего риска*.

В случае, когда пространства сигналов S и решений Γ непрерывны, средний риск

$$R = \int_{\Gamma} \int_S \Pi(s, \gamma) w(s, \gamma) ds d\gamma, \quad (3.5)$$

где $w(s, \gamma)$ — совместная плотность вероятности сигнала s и решения γ .

Для нерандомизированных правил принятое колебание и решение связаны детерминированной зависимостью, а поэтому справедливо соотношение

$$w(s, \gamma) ds d\gamma = w(s, u) ds du. \quad (3.6)$$

С учетом формулы (3.6) выражение (3.5) записывается в виде

$$R = \int \int_S \Pi(s, \gamma_u) w(s, u) ds du, \quad (3.7)$$

где γ_u — решение, соответствующее принятому сигналу $u(t)$.

Заметим, что средний риск не является исчерпывающей характеристикой решающего правила. Более полной характеристикой качества обработки сигнала может служить, например, совокупность двух показателей: среднего риска R и дисперсии среднего риска D_R (в общем случае — совокупность моментов функции потерь $\Pi(s, \gamma)$).

Однако синтез решающего правила по совокупности показателей значительно сложнее, и на практике ограничиваются критерием минимума среднего риска.

Применение байесовского критерия требует большого объема априорной информации: необходимо знать функцию потерь и совместное распределение $w(s, u)$, или, что то же самое, распределения $w(s)$ и $w(u|s)$. Если априорное распределение сигналов $w(s)$ неизвестно, то применять критерий Байеса нельзя, так как при этом задача оптимизации, в смысле минимума среднего риска, оказывается неопределенной. В этом случае применяют минимаксный критерий.

Введем понятие условного среднего риска

$$r_{s_i} = \sum_j \Pi(s_i, \gamma_j) w(\gamma_j | s_i)$$

для дискретных пространств S и Γ и

$$r_s = \int_{\Gamma} \Pi(s, \gamma) w(\gamma | s) d\gamma = \int_U \Pi(s, \gamma_u) w(u | s) du$$

для непрерывных пространств S и Γ , где $w(\gamma_j | s_i)$ — вероятность принятия решения γ_j при условии, что на входе приемника присутствует полезный сигнал s_i , а $w(\gamma | s)$ и $w(u | s)$ — соответственно условные плотности вероятности случайных величин γ и u .

Согласно минимаксному критерию наилучшим считается решающее правило, которое обеспечивает наименьшее значение максимального условного риска r_s . Минимаксный критерий позволяет найти наилучшее решение для наихудшего случая. Он гарантирует, что минимально возможное значение условного риска не будет больше некоторого значения даже при самом неблагоприятном распределении $w(s)$.

На практике часто встречаются случаи, когда ошибочные решения одинаково нежелательны, например при передаче дискретных сообщений. При этом целесообразно функцию потерь задать следующим образом:

$$\Pi(s_i, \gamma_j) = \begin{cases} \text{const} & \text{при } j \neq i, \\ 0 & \text{при } j = i. \end{cases} \quad (3.8)$$

Обычно так же функция потерь задается и в тех случаях, когда ее обоснованный выбор затруднителен. Например, это имеет место в радиолокации. Действительно, при решении задачи обнаружения сигнала, как правило, трудно оценить потери, связанные с пропуском сигнала (пропуском цели) и ложным обнаружением (ложной тревогой).

Для функции потерь, определяемой выражением (3.8), средний риск определяется формулой

$$R = C \sum_i \sum_j w(s_i, \gamma_j), \quad i \neq j, \quad (3.9)$$

где C — любая постоянная.

Из выражения (3.9) следует, что величина R с точностью до постоянного множителя совпадает с полной вероятностью ошибочных решений. Поэтому байесовский критерий при задании функции потерь в виде выражения (3.8) имеет смысл назвать *критерием минимума полной вероятности ошибки*. Его часто называют *критерием идеального наблюдателя*, а также *критерием Котельникова—Зигерта*.

Другим критерием, который минимизирует полную вероятность ошибочных решений, является критерий максимума апостериорной вероятности

$$w(s|u) = \frac{w(s)w(u|s)}{w(u)}, \quad (3.10)$$

в соответствии с которым решение принимается в пользу сигнала s_i , если

$$w(s_i|u) > w(s_j|u), \quad j = 1, \dots, m, \quad j \neq i. \quad (3.11)$$

Действительно, все, что можно узнать о полезном сигнале s_i на основе принятого колебания u , заключено в апостериорной вероятности $w(s_i|u)$, которая представляет собой не что иное, как вероятность правильного приема сигналов. Поэтому если решение принимать в соответствии с выражением (3.11), то гарантируется, что полная вероятность ошибки будет минимальной.

При отсутствии априорных сведений не только о функции потерь, но и о распределении сигналов $w(s)$ считают, что $w(s) = \text{const}$, т. е. принимают распределение сигналов как равномерное. При этом апостериорная вероятность (3.10) с точностью до постоянного множителя совпадает с функцией правдоподобия $w(u|s)$. Для обеспечения минимального среднего риска решение принимается в пользу сигнала, для которого функция правдоподобия максимальна, т. е. в пользу сигнала s_i , если

$$w(u|s_i) > w(u|s_j), \quad j = 1, \dots, m, \quad j \neq i. \quad (3.12)$$

Критерий (3.12) целесообразно назвать *критерием максимального правдоподобия*. Он часто используется в математической статистике при оценке неизвестных величин.

В задачах обнаружения сигнала часто применяют критерий Неймана—Пирсона, который гарантирует, что вероятность ошибки типа «ложная тревога» не превысит заранее выбранную величину, а вероятность ошибки типа «пропуск сигнала» будет минимальной. Критерий Неймана—Пирсона не требует знания априорных вероятностей присутствия и отсутствия сигнала, а также функции потерь.

В рассмотренных выше решающих правилах выбор решений проводился на основе m -мерных векторов принимаемых сигналов, причем значение m оставалось постоянным. Существует правило выбора решений, при котором размер выборки сигнала n случаен и зависит от результатов наблюдения за сигналом на предыдущих $m - 1$ шагах (т. е. в моменты времени t_1, \dots, t_{m-1}). Это правило предложено Вальдом и называется *правилом последовательного анализа*. Оптимальность критерия Вальда заключается в том, что он при заданных вероятностях ошибочных решений гарантирует минимум среднего размера выборки сигнала, или, что то же самое, минимум среднего времени наблюдения, необходимого для принятия решения.

Рассмотренные в настоящем параграфе критерии будут использованы далее при синтезе оптимальных устройств обработки сигналов.

3.3. Обнаружение сигналов

3.3.1. Оптимальные алгоритмы обнаружения сигналов

Задача обнаружения (ее формулировка дана в § 3.1) является частным случаем общей задачи проверки статистических гипотез: необходимо на основе анализа принятого колебания $u(t)$ сделать выбор между гипотезой H_0 (утверждение, что сигнала на входе нет, т. е. параметр θ равен нулю) и альтернативной гипотезой H_1 (утверждение, что сигнал на входе присутствует, т. е. параметр θ равен единице). Поскольку в задаче обнаружения возможны только два решения: γ_0 (принимается гипотеза H_0) и γ_1 (принимается гипотеза H_1), она называется *двоичной (двухальтернативной)*.

Если распределение наблюдаемого колебания $u(t)$ зависит только от того, какое значение принял параметр θ , то гипотезы H_0 и H_1 называются *простыми*. Такой случай имеет место при обнаружении детерминированного сигнала. При этом

$$u(t) = \theta s(t) + n(t),$$

где $s(t)$ — полезный сигнал, все параметры которого известны, и приходится решать задачу проверки простой гипотезы H_0 при простой альтернативе H_1 .

Более типичным для задачи обнаружения является случай, когда распределение вероятностей наблюдаемого колебания $u(t)$ зависит не только от вида гипотезы, но и от неизвестного параметра λ (или вектора λ). При этом гипотезы H_0 и H_1 называются *сложными* и задача обнаружения состоит в проверке сложной гипотезы при сложной альтернативе.

Рассмотрим задачу обнаружения, когда гипотеза H_0 и альтернатива H_1 являются простыми. Задача заключается в нахождении правила $\Delta(\gamma|u)$, согласно которому любой наблюдаемой выборке u_1, \dots, u_m ставится в соответствие решение γ_0 или γ_1 . Процесс принятия решения в данном случае можно представить следующим образом. Пространство принимаемых сигналов U разбивается на две непересекающиеся области U_0 и U_1 . Если сигнал на входе приемника $u(t)$ попадает в область U_0 , то принимается решение γ_0 (гипотеза H_0), при попадании сигнала $u(t)$ в область U_1 принимается решение γ_1 (гипотеза H_1). Область U_0 называется *допустимой*, а область U_1 — *критической*.

Поскольку помеха $n(t)$ имеет случайный характер, принимаемое решение не всегда является достоверным. Поэтому при решении задачи обнаружения возможны следующие четыре случая.

1. Справедлива гипотеза H_0 . Принимается решение γ_0 . Этот случай называется *правильным необнаружением*. Вероятность такого события определяется формулой

$$P_n = \int_{U_0} w(u|H_0) du. \quad (3.13)$$

2. Справедлива гипотеза H_1 . Принимается решение γ_1 . Этот случай называется *правильным обнаружением*. Вероятность такого события

$$D = \int_{U_1} w(u|H_1) du. \quad (3.14)$$

3. Справедлива гипотеза H_0 . Принимается решение γ_1 , т. е. ошибочное решение. Этот случай называется *ложным обнаружением (ложной тревогой)*. Вероятность такого события

$$F_{лт} = \int_{U_1} w(u|H_0) du. \quad (3.15)$$

4. Справедлива гипотеза H_1 . Принимается решение γ_0 , т. е. ошибочное решение. Этот случай называется *пропуском сигнала*. Вероятность такого события

$$p_{\text{проп}} = \int_{U_0} w(u|H_1) du. \quad (3.16)$$

Таким образом, из четырех возможных случаев, имеющих место при обнаружении сигналов, в двух решения принимаются правильные, а в двух — ошибочные. Ошибку ложного обнаружения называют также *ошибкой первого рода*, ее вероятность $F_{\text{лт}}$ — *уровнем значимости правила выбора решения*. Ошибку, связанную с пропуском сигнала, называют *ошибкой второго рода*, а вероятность правильного обнаружения $D = 1 - p_{\text{проп}}$ — *мощностью правила выбора решения*.

Вероятности ошибочных решений $F_{\text{лт}}$ и $p_{\text{проп}}$, как следует из формул (3.15), (3.16), зависят от характера разбиения пространства принимаемых сигналов U на области U_0 и U_1 . Очевидно, что уменьшая область U_1 , можно уменьшить вероятность ложной тревоги $F_{\text{лт}}$, однако при этом возрастает вероятность ошибки пропуска сигнала $p_{\text{проп}}$. Уменьшение области U_0 приводит наряду с уменьшением $p_{\text{проп}}$ к увеличению $F_{\text{лт}}$. В связи с этим возникает задача, как наилучшим образом разбить пространство U на области U_0 и U_1 .

Рассмотрим сначала байесовское правило принятия решений.

Запишем колебание на входе обнаружителя:

$$u(t) = \theta s_1(t) + (1 - \theta) s_0(t) + n(t), \quad (3.17)$$

где $s_1(t) = s(t)$, $s_0(t) \equiv 0$.

Случайная величина θ принимает значения 1 и 0 с вероятностями p и $1 - p$ соответственно.

Функцию потерь зададим следующими выражениями:

$$\begin{aligned} \Pi(s_1, \gamma_0) &= \Pi_{10} > 0, & \Pi(s_0, \gamma_1) &= \Pi_{01} > 0, \\ \Pi(s_1, \gamma_1) &= \Pi_{11}, & \Pi(s_0, \gamma_0) &= \Pi_{00}, \end{aligned}$$

причем $\Pi_{10} > \Pi_{11}$ и $\Pi_{01} > \Pi_{00}$. Тогда средний риск, в соответствии с формулой (3.4), имеет вид

$$\begin{aligned} R &= \sum_{j=0}^1 \sum_{i=0}^1 \Pi(s_i, \gamma_j) w(s_i, \gamma_j) = \Pi_{00} w(s_0, \gamma_0) + \\ &+ \Pi_{01} w(s_0, \gamma_1) + \Pi_{10} w(s_1, \gamma_0) + \Pi_{11} w(s_1, \gamma_1) = \\ &= \Pi_{00} w(s_0) w(\gamma_0 | s_0) + \Pi_{01} w(s_0) w(\gamma_1 | s_0) + \\ &+ \Pi_{10} w(s_1) w(\gamma_0 | s_1) + \Pi_{11} w(s_1) w(\gamma_1 | s_1). \end{aligned}$$

Учитывая, что $w(s_1) = p$, $w(s_0) = 1 - p$ и

$$w(\gamma_0 | s_0) = p_n, \quad w(\gamma_1 | s_0) = F_{\text{лт}}, \quad w(\gamma_0 | s_1) = p_{\text{проп}}, \quad w(\gamma_1 | s_1) = D,$$

получим

$$R = \Pi_{00}(1-p)p_n + \Pi_{01}(1-p)F_{\text{лт}} + \Pi_{10}pp_{\text{проп}} + \Pi_{11}pD. \quad (3.18)$$

Подставляя в формулу (3.18) выражения для p_n , D , $F_{\text{лт}}$ и $p_{\text{проп}}$ из формул (3.13) — (3.16), находим

$$\begin{aligned} R = & \Pi_{00}(1-p) \int_{U_0} w(\mathbf{u} | H_0) d\mathbf{u} + \Pi_{01}(1-p) \int_{U_1} w(\mathbf{u} | H_0) d\mathbf{u} + \\ & + \Pi_{10}p \int_{U_0} w(\mathbf{u} | H_1) d\mathbf{u} + \Pi_{11}p \int_{U_1} w(\mathbf{u} | H_1) d\mathbf{u}. \end{aligned} \quad (3.19)$$

Учитывая, что

$$\begin{aligned} \int_{U_0} w(\mathbf{u} | H_1) d\mathbf{u} &= 1 - \int_{U_1} w(\mathbf{u} | H_1) d\mathbf{u}, \\ \int_{U_0} w(\mathbf{u} | H_0) d\mathbf{u} &= 1 - \int_{U_1} w(\mathbf{u} | H_0) d\mathbf{u}, \end{aligned}$$

согласно формуле (3.19) средний риск

$$\begin{aligned} R = & \Pi_{00}(1-p) + \Pi_{10}p - \int_{U_1} [p(\Pi_{10} - \Pi_{11})w(\mathbf{u} | H_1) - \\ & - (1-p)(\Pi_{01} - \Pi_{00})w(\mathbf{u} | H_0)] d\mathbf{u}. \end{aligned} \quad (3.20)$$

Первые два члена в соотношении (3.20) являются постоянными величинами, не зависящими от способа разбиения пространства U . Поэтому минимальное значение среднего риска достигается тогда, когда значение интеграла оказывается максимальным. Для этого необходимо область U_1 выбрать такой, чтобы в нее вошли все точки пространства U , для которых подынтегральное выражение неотрицательно, т. е.

$$p(\Pi_{10} - \Pi_{11})w(\mathbf{u} | H_1) - (1-p)(\Pi_{01} - \Pi_{00})w(\mathbf{u} | H_0) \geq 0. \quad (3.21)$$

Используя формулу (3.21), можно записать алгоритм работы оптимального по критерию Байеса обнаружителя сигнала:

$$\frac{w(\mathbf{u} | H_1)}{w(\mathbf{u} | H_0)} \underset{H_0}{\overset{H_1}{\geq}} \frac{(1-p)(\Pi_{01} - \Pi_{00})}{p(\Pi_{10} - \Pi_{11})}. \quad (3.22)$$

По алгоритму (3.22) необходимо вычислить отношение функций правдоподобия $l(\mathbf{u}) = w(\mathbf{u}|H_1)/w(\mathbf{u}|H_0)$ и сравнить его с пороговым значением (уровнем)

$$l_0 = \frac{(1-p)(\Pi_{01} - \Pi_{00})}{p(\Pi_{10} - \Pi_{11})}. \quad (3.23)$$

Если $l(\mathbf{u}) \geq l_0$, то принимается гипотеза H_1 о наличии сигнала, а если $l(\mathbf{u}) < l_0$, то принимается гипотеза H_0 об отсутствии сигнала.

Отношение $l(\mathbf{u}) = w(\mathbf{u}|H_1)/w(\mathbf{u}|H_0)$ называется *отношением правдоподобия*, или *коэффициентом правдоподобия*. Оно показывает, насколько гипотеза H_1 при принятом сигнале $u(t)$ правдоподобнее гипотезы H_0 . При известном распределении помехи $n(t)$ и известном способе взаимодействия полезного сигнала и помехи отношение $l(\mathbf{u})$ всегда может быть найдено.

Пороговый уровень l_0 зависит от априорной вероятности наличия сигнала p и значений функции потерь Π_{00} , Π_{01} , Π_{10} и Π_{11} . Он может быть вычислен заранее. Важно отметить, что априорные вероятности наличия и отсутствия сигнала и функция потерь не влияют на процедуру обработки сигнала. При изменении этих данных приходится лишь подстраивать пороговый уровень.

Задавая функцию потерь в виде соотношений

$$\Pi(s_1, \gamma_0) = \Pi(s_0, \gamma_1), \quad \Pi(s_1, \gamma_1) = \Pi(s_0, \gamma_0) = 0,$$

с учетом выражения (3.22) получаем алгоритм работы обнаружителя сигнала, построенного на основе критерия идеального наблюдателя:

$$l(\mathbf{u}) = \frac{w(\mathbf{u}|H_1)}{w(\mathbf{u}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{(1-p)}{p} = l_0. \quad (3.24)$$

В соответствии с выражением (3.24) необходимо вычислить отношение правдоподобия и сравнить его с пороговым уровнем

$$l_0 = \frac{(1-p)}{p}.$$

Из формулы (3.24), предполагая, что вероятности наличия и отсутствия сигнала равны, т. е. $1-p = p$, находим алгоритм работы обнаружителя сигнала, построенного на основе критерия максимального правдоподобия

$$l(\mathbf{u}) = \frac{w(\mathbf{u}|H_1)}{w(\mathbf{u}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} 1. \quad (3.25)$$

В радиолокации при решении задачи обнаружения часто применяется критерий Неймана—Пирсона. Алгоритм работы обнаружителя, построенного на основе этого критерия, можно записать в виде соотношения

$$l(\mathbf{u}) = \frac{w(\mathbf{u}|H_1)}{w(\mathbf{u}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} C, \quad (3.26)$$

где порог C находится из условия, что вероятность выполнения неравенства $l(\mathbf{u}) \geq C$ при гипотезе H_0 не превышает наперед заданной величины $F_{\text{лт}}$, т. е.

$$P\{l(\mathbf{u}) \geq C | H_0\} = F_{\text{лт}}.$$

Таким образом, для всех рассмотренных критериев: байесовского, идеального наблюдателя, максимального правдоподобия и Неймана—Пирсона, процедура принятия решения (рис. 3.2) сводится к вычислению отношения правдоподобия $l(\mathbf{u})$ (блок ОП) и сравнению его с пороговым значением l_0 (блок ПУ — пороговое устройство), равным соответственно $(1-p)(\Pi_{01}-\Pi_{00})/p(\Pi_{10}-\Pi_{11})$, $(1-p)/p$, 1 и C . Структура обнаружителя сигнала не зависит от выбранного критерия оптимальности.

Рассмотрим работу обнаружителя, построенного на основе критерия последовательного анализа (наблюдателя) или критерия Вальда. В ранее рассмотренных решающих правилах решение принимается после получения выборки сигнала u_1, \dots, u_m фиксированного объема m . Последовательная процедура обнаружения характеризуется тем, что попытка принять решение в пользу гипотезы H_0 или H_1 делается каждый раз по мере получения очередного элемента выборки. Процесс принятия решения состоит в следующем. При получении элемента выборки $u_1 = u(t_1)$ на основе его анализа либо принимается решение в пользу одной из гипотез H_0 или H_1 (решение γ_0 или γ_1), либо принимается решение γ_2 о продолжении наблюдения за сигналом. Если на первом шаге по элементу выборки u_1 принимается решение γ_0 или γ_1 , то процесс обнаружения на этом заканчивается. Если принято решение γ_2 , то продолжают наблюдение за сигналом. При получении второго элемента выборки $u_2 = u(t_2)$ на основе анализа двумерной выборки u_1, u_2 вновь проверяется возможность

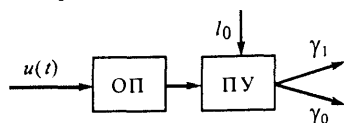


Рис. 3.2. Структурная схема обнаружителя

принятия решений в пользу одной из гипотез. Процесс обнаружения заканчивается на том шаге, на котором принимается решение γ_0 или γ_1 .

При последовательной процедуре обнаружения механизм принятия решения можно представить следующим образом. Пространство принимаемых сигналов U на каждом шаге наблюдения разбивается на три области: U_0 , U_1 , U_2 . Если анализируемая выборка u_1, u_2, \dots, u_i на i -м шаге наблюдения попадет в допустимую область U_0 , то принимается решение γ_0 об отсутствии сигнала. Если выборка попадает в критическую область U_1 , то выносится решение γ_1 о наличии сигнала. Наконец, если она попадет в область U_2 , то принимается решение γ_2 о продолжении наблюдения.

Эффективность последовательной процедуры обнаружения зависит от способа разбиения пространства U на области U_0 , U_1 и U_2 . Так же, как и при рассмотрении непоследовательных процедур, возникает вопрос: как наилучшим образом произвести разбиение пространства U ? Ответ на этот вопрос зависит от того, какой смысл вкладывается в понятие оптимальности процедуры.

На практике часто одной из важнейших характеристик решающего правила является средняя цена наблюдений, необходимых для принятия решения при обеспечении заданных условных вероятностей $F_{\text{лт}}$ и $p_{\text{проп}}$. Учитывая, что цена наблюдений пропорциональна длительности наблюдений, оптимальной при заданных вероятностях $F_{\text{лт}}$ и $p_{\text{проп}}$ будет процедура, которая обеспечивает минимальную среднюю продолжительность наблюдений, или, что то же самое, минимальный средний объем \bar{m} выборки принимаемого сигнала.

А. Вальдом совместно с Дж. Вольфовитцем в работе [27] было доказано, что для обеспечения минимального среднего размера выборки при условных вероятностях ложного обнаружения и пропуска сигнала, не превосходящих соответственно значений $F_{\text{лт}}$ и $p_{\text{проп}}$, необходимо на каждом i -м шаге, $i = 1, 2, \dots$, вычислять отношение правдоподобия

$$l_i(\mathbf{u}) = \frac{w(\mathbf{u}|H_1)}{w(\mathbf{u}|H_0)} = \frac{w_i(u_1, \dots, u_i|H_1)}{w_i(u_1, \dots, u_i|H_0)} \quad (3.27)$$

и сравнивать его с двумя порогами: верхним A и нижним B (рис. 3.3). Если отношение $l(\mathbf{u})$ превышает порог A (вектор \mathbf{u} попадает в область U_1), то принимается решение в пользу гипотезы H_1 . Если отношение $l(\mathbf{u})$ оказывается меньше порога B (вектор \mathbf{u} попадает в область U_0), то принимается решение в пользу гипотезы H_0 . Наконец, если отношение $l(\mathbf{u})$ удовлетворяет неравенству

$$B < l(\mathbf{u}) < A$$

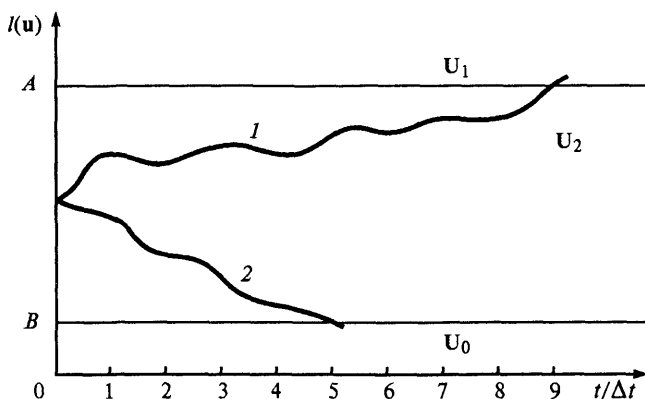


Рис. 3.3. Последовательная процедура принятия решения

(вектор \mathbf{u} попадает в область U_2), то принимается решение о продолжении наблюдения. Процесс обнаружения заканчивается при первом пересечении функции $l(\mathbf{u})$ с одним из порогов. На рис. 3.3 в приведенных примерах процесс обнаружения заканчивается принятием гипотезы H_1 (пример 1) и H_0 (пример 2) соответственно на девятом и пятом шагах.

Пороги A и B находятся по заданным вероятностям ошибок $F_{\text{лт}}$ и $p_{\text{проп}}$ на основе следующих рассуждений. Выпишем условие принятия решения в пользу гипотезы H_1 :

$$l(\mathbf{u}) = \frac{w(\mathbf{u}|H_1)}{w(\mathbf{u}|H_0)} \geq A,$$

или

$$w(\mathbf{u}|H_1) \geq Aw(\mathbf{u}|H_0). \quad (3.28)$$

Поскольку условие (3.28) справедливо для любой выборки, попадающей в область U_1 , обе части неравенства (3.28) можно проинтегрировать по этой области:

$$\int_{U_1} w(\mathbf{u}|H_1) d\mathbf{u} \geq A \int_{U_1} w(\mathbf{u}|H_0) d\mathbf{u}. \quad (3.29)$$

Левая часть неравенства (3.29) есть не что иное, как вероятность правильного обнаружения $D = 1 - p_{\text{проп}}$, а интеграл в правой части неравенства (3.29) равен $F_{\text{лт}}$. Следовательно, получаем

$$1 - p_{\text{проп}} \geq AF_{\text{лт}},$$

или

$$A \leq \frac{1 - p_{\text{проп}}}{F_{\text{лт}}} = \frac{D}{F_{\text{лт}}}. \quad (3.30)$$

Неравенство (3.30) является оценкой сверху порога A . Аналогично можно получить оценку снизу для порога B .

Выпишем условие принятия решения в пользу гипотезы H_0 :

$$l(\mathbf{u}) = \frac{w(\mathbf{u}|H_1)}{w(\mathbf{u}|H_0)} \leq B,$$

или

$$w(\mathbf{u}|H_1) \leq B w(\mathbf{u}|H_0). \quad (3.31)$$

Неравенство (3.31) справедливо для любой выборки \mathbf{u} , попадающей в область U_0 , поэтому обе части неравенства (3.31) можно проинтегрировать по области U_0 :

$$\int_{U_0} w(\mathbf{u}|H_1) d\mathbf{u} \leq B \int_{U_0} w(\mathbf{u}|H_0) d\mathbf{u}. \quad (3.32)$$

Левая часть неравенства (3.32) является вероятностью пропуска сигнала $p_{\text{проп}}$, а интеграл в правой части — вероятностью правильного необнаружения $p_{\text{н}} = 1 - F_{\text{лт}}$.

Поэтому

$$p_{\text{проп}} \leq B(1 - F_{\text{лт}}),$$

или

$$B \geq \frac{p_{\text{проп}}}{1 - F_{\text{лт}}} = \frac{1 - D}{1 - F_{\text{лт}}}. \quad (3.33)$$

Неравенство (3.33) является оценкой снизу порога B .

Оценки решающих порогов (3.30) и (3.33) определяются только заданными вероятностями $F_{\text{лт}}$ и $p_{\text{проп}}$ и не зависят от вида распределений $w(\mathbf{u}|H_1)$ и $w(\mathbf{u}|H_0)$. В случае близких гипотез, т. е. когда распределения $w(\mathbf{u}|H_1)$ и $w(\mathbf{u}|H_0)$ мало отличаются друг от друга, что имеет место при малых отношениях сигнал—шум, неравенства (3.30) и (3.33) переходят в приближенные равенства

$$A \approx \frac{1 - p_{\text{проп}}}{F_{\text{лт}}}, \quad B \approx \frac{p_{\text{проп}}}{1 - F_{\text{лт}}}. \quad (3.34)$$

При независимых элементах выборки вместо отношения правдоподобия удобно пользоваться его логарифмом. При этом выражение (3.27) можно переписать в виде

$$\ln l_i(\mathbf{u}) = Z_i = \sum_{j=1}^i z_j = \sum_{j=1}^i \ln \frac{w(u_j|H_1)}{w(u_j|H_0)}. \quad (3.35)$$

Верхний и нижний пороги берутся равными $a = \ln A$ и $b = \ln B$ соответственно. На практике условные вероятности ошибок не превышают 0,5 и пороги удовлетворяют неравенствам

$$a > 0, \quad b < 0.$$

Заметим, что при доказательстве оптимальности рассматриваемой процедуры в работе [27] предполагалось, что гипотезы H_0 и H_1 являются простыми, выборка u_1, \dots, u_i — независимой и однородной, т. е.

$$w_i(u_1, \dots, u_i | H_j) = \prod_{k=1}^i w(u_k | H_j), \quad w(u_k | H_j) = w(u | H_j), \quad k=1, \dots, i, \quad j=0, 1,$$

наблюдаемое распределение точно совпадает с ожидаемыми для гипотезы H_0 или H_1 , а перескоком решающей статистики $l_i(\mathbf{u})$ или $\ln l_i(\mathbf{u})$ через пороги можно пренебречь (гипотезы H_0 и H_1 близки). При нарушении одного или нескольких из указанных предположений поиск оптимальной последовательной процедуры усложняется.

Одной из важнейших характеристик последовательной процедуры является ее средняя длительность или математическое ожидание числа шагов при гипотезах H_0 и H_1 . Пусть решение принимается на шаге m наблюдения. Для независимой однородной выборки с учетом формулы (3.35) решающая статистика Z_m в момент принятия решения в пользу одной из гипотез представляет собой сумму случайного числа одинаково распределенных величин z_k , $k=1, 2, \dots$. Поэтому при гипотезах H_0 и H_1 математические ожидания числа шагов имеют вид

$$\bar{m}_0 = \mathbf{M}\{m|H_0\} = \mathbf{M}\{Z_m|H_0\} / \mathbf{M}\{z_k|H_0\}, \quad (3.36)$$

$$\bar{m}_1 = \mathbf{M}\{m|H_1\} = \mathbf{M}\{Z_m|H_1\} / \mathbf{M}\{z_k|H_1\}, \quad (3.37)$$

где

$$\mathbf{M}\{z_k|H_j\} = \bar{z}_j = \int_{-\infty}^{\infty} \ln \left(\frac{w(z_k|H_1)}{w(z_k|H_0)} \right) w(z_k|H_j) dz_k, \quad j=0, 1, \quad (3.38)$$

— математическое ожидание приращения решающей статистики за такт наблюдения при гипотезе H_j , причем $\bar{z}_0 < 0$, а $\bar{z}_1 > 0$.

Действительно, используя неравенство $\ln x \leq x - 1$ и учитывая, что $\ln x = x - 1$ только при $x = 1$, получим

$$\begin{aligned}\bar{z}_0 &= \int_{-\infty}^{\infty} \ln \left(\frac{w(z_k | H_1)}{w(z_k | H_0)} \right) w(z_k | H_0) dz_k \leq \int_{-\infty}^{\infty} \left(\frac{w(z_k | H_1)}{w(z_k | H_0)} - 1 \right) w(z_k | H_0) dz_k = \\ &= \int_{-\infty}^{\infty} [w(z_k | H_1) - w(z_k | H_0)] dz_k = 0.\end{aligned}$$

Поскольку $w(z_k | H_1) \neq w(z_k | H_0)$, находим, что $\bar{z}_0 < 0$.

Аналогично,

$$\begin{aligned}\bar{z}_1 &= \int_{-\infty}^{\infty} \ln \left(\frac{w(z_k | H_1)}{w(z_k | H_0)} \right) w(z_k | H_1) dz_k = \int_{-\infty}^{\infty} \ln \left(\frac{w(z_k | H_0)}{w(z_k | H_1)} \right)^{-1} w(z_k | H_1) dz_k = \\ &= \int_{-\infty}^{\infty} \ln 1 - \ln \left(\frac{w(z_k | H_0)}{w(z_k | H_1)} \right) w(z_k | H_1) dz_k.\end{aligned}$$

Учитывая неравенство

$$\int_{-\infty}^{\infty} \ln \left(\frac{w(z_k | H_0)}{w(z_k | H_1)} \right) w(z_k | H_1) dz_k < 0$$

(см. предыдущий вывод), находим, что $\bar{z}_1 > 0$.

Для близких гипотез решающие статистики $Z_m | H_0$ и $Z_m | H_1$ определяются выражениями

$$\begin{aligned}Z_m | H_0 &= \begin{cases} a & \text{с вероятностью } F_{\text{лт}}, \\ b & \text{с вероятностью } 1 - F_{\text{лт}}, \end{cases} \\ Z_m | H_1 &= \begin{cases} a & \text{с вероятностью } 1 - p_{\text{проп}}, \\ b & \text{с вероятностью } p_{\text{проп}}. \end{cases}\end{aligned}$$

Тогда

$$\mathbf{M}\{Z_m | H_0\} \approx F_{\text{лт}} a + (1 - F_{\text{лт}}) b, \quad (3.39)$$

$$\mathbf{M}\{Z_m | H_1\} \approx (1 - p_{\text{проп}}) a + p_{\text{проп}} b. \quad (3.40)$$

Подставив выражения (3.38)—(3.40) в формулы (3.36) и (3.37), получим

$$\bar{m}_0 = \mathbf{M}\{m|H_0\} = \frac{F_{\text{лт}}a + (1 - F_{\text{лт}})b}{\bar{z}_0}, \quad (3.41)$$

$$\bar{m}_1 = \mathbf{M}\{m|H_1\} = \frac{(1 - p_{\text{проп}})a + p_{\text{проп}}b}{\bar{z}_1}. \quad (3.42)$$

Заметим, что для рассматриваемого случая близких гипотез выполняется условие $\bar{z}_0 \approx -\bar{z}_1$. Тогда, учитывая выражения (3.41) и (3.42), находим

$$\bar{m}_1 / \bar{m}_0 \approx -a/b. \quad (3.43)$$

Для последовательной процедуры существенным является то, что среднее число шагов \bar{m} оказывается меньше фиксированного размера выборки m , необходимого для обнаружения сигнала по любому другому критерию при условных вероятностях ошибок, не превосходящих $F_{\text{лт}}$ и $p_{\text{проп}}$.

При $F_{\text{лт}} = p_{\text{проп}}$ выигрыш в среднем объеме выборки, обеспечиваемый процедурой Вальда по сравнению с эквивалентной по вероятностям ошибок процедурой Неймана—Пирсона, составляет примерно 2 раза (см. [27]), а при сильно различающихся значениях $F_{\text{лт}}$ и $p_{\text{проп}}$ ($F_{\text{лт}} \ll p_{\text{проп}}$), что характерно для радиолокации, выигрыш существенно возрастает.

Объясняется это следующим. При $F_{\text{лт}} < p_{\text{проп}}$ пороги a и b оказываются не симметричными относительно нулевого уровня и выполняется неравенство $a \gg |b|$. Поэтому, учитывая выражения (3.43), можно утверждать, что $\bar{m}_1 \gg \bar{m}_0$. Величина \bar{m}_1 приблизительно равна размеру выборки m , используемой в эквивалентной процедуре Неймана—Пирсона, и, следовательно, $\bar{m}_0 \ll \bar{m}_1$. Поскольку при радиолокационном наблюдении в большинстве элементов разрешения цели нет, средняя длительность процедуры Вальда будет существенно меньше длительности процедуры Неймана—Пирсона. Выигрыш в среднем времени наблюдения оценивается приближенно отношением $\ln F_{\text{лт}} / \ln p_{\text{проп}}$ и при типичных для радиолокации значениях $F_{\text{лт}} = 10^{-4} \dots 10^{-12}$ и $p_{\text{проп}} = 0,1 \dots 0,5$ составляет 5–20 раз.

Последовательный критерий отношения правдоподобия позволяет получить эффективные процедуры в смысле минимального значения среднего объема наблюдений (выборки) и в тех случаях, когда условия теоремы Вальда—Вольфовитца не выполняются, хотя строгого доказательства оптимальности этих процедур в настоящее время не существует.

Обратим внимание на то, что последовательная процедура Вальда дает выигрыш по сравнению с непоследовательными процедурами лишь в

среднем времени наблюдения. В отдельных случаях длительность процесса обнаружения по Вальду может оказаться недопустимо большой. Особенно это проявляется в тех случаях, когда задача обнаружения сигнала решается параллельно во многих каналах (например, в N радиолокационных каналах по дальности). Если решение в каждом из каналов принимается независимо, то общая длительность процедуры определяется тем каналом, где решение было принято последним. Очевидно, что с увеличением числа каналов N длительность процедуры увеличивается. При этом временные затраты последовательной процедуры растут быстрее, чем для эквивалентной по надежности (при одинаковых вероятностях $F_{\text{пт}}$ и D) процедуры Неймана—Пирсона [23]. Поэтому применение последовательной процедуры Вальда для многоканального обнаружения с независимым принятием решений при увеличении числа каналов становится все менее выгодным по сравнению с процедурой Неймана—Пирсона (предполагается, что каналы однородны).

Повысить эффективность радиолокационного последовательного обнаружения удалось, учитывая зависимость отношения сигнал—шум от дальности. Среднюю длительность многоканальной процедуры с независимыми решениями можно сократить, если перейти от постоянных (вальдовских) порогов к функциональным порогам, зависящим от номера шага, для которого область неопределенности по мере накопления данных сужается. Однако удобных для инженерной практики методов расчета зависимостей порогов $A(m)$ и $B(m)$, обеспечивающих минимизацию средней длительности процедуры при заданных вероятностях $F_{\text{пт}}$ и $p_{\text{проп}}$ в настоящее время не существует [23].

Вырожденным случаем такой процедуры является усеченная последовательная процедура [28]. При этом обнаружитель сначала работает как последовательный (рис. 3.4): вычисляет значения отношения правдоподобия $l_i(\mathbf{u})$, $i = 1, 2, \dots$ (блок ОП), которые через переключатель П поступают на

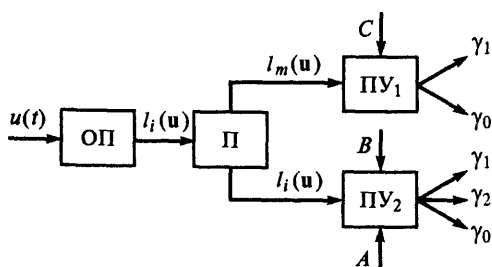


Рис. 3.4. Структурная схема последовательного обнаружителя с усечением времени наблюдения

двухпороговое устройство ПУ₂. Если окажется, что в течение заданного времени наблюдения (заданного числа m шагов) не будет принято решение в пользу гипотезы H_0 или H_1 , то выход блока формирования отношения правдоподобия переключается на однопороговое устройство ПУ₁. Производится сравнение величины $l_m(\mathbf{u})$ с одним порогом, выбираемым, например, по критерию

Неймана—Пирсона, и выносится решение γ_0 или γ_1 . На этом процедура обнаружения заканчивается.

Недостатком усеченной процедуры Вальда является то, что она часто не обеспечивает заданные вероятности ошибочных решений [29].

Дальнейшее повышение эффективности многоканального последовательного обнаружения достигается, если в ходе эксперимента целенаправленно изменяется не только решающее правило, но и условия наблюдения. Последние изменяются не на каждом шаге наблюдения, а по завершению некоторого этапа. Параметры первого этапа выбираются исходя из заданной вероятности пропуска сигнала, поэтому его средняя длительность может быть меньше величины, необходимой для поддержания расчетной вероятности ложной тревоги $F_{\text{лт}}$. Решение в пользу гипотезы H_0 на первом этапе означает окончание процедуры. Если гипотеза H_0 не подтверждается, то происходит переход к следующему этапу.

Примером алгоритма этого класса может служить двухэтапное радиолокационное обнаружение, при котором на первом этапе разрешающая способность по дальности уменьшается, например увеличением длительности зондирующего сигнала. На втором этапе, который проводится только при условии, что на первом этапе было принято решение об обнаружении сигнала, обеспечивается требуемая разрешающая способность и производится измерение дальности.

Многоэтапные процедуры относятся к классу последовательных, если даже на некоторых этапах правило принятия решения базируется на выборках фиксированного объема. При оптимальном выборе параметров двух-трехэтапные многоканальные последовательные процедуры приближаются по эффективности к квазиоптимальным невырожденным последовательным процедурам. Недостатком таких алгоритмов является сравнительная сложность их реализации и использования в нестационарных условиях наблюдения. С методами преодоления трудностей, возникающих при использовании многоэтапных процедур, можно познакомиться в работе [29].

На практике часто параметры сигнала (амплитуда, частота, фаза), а также параметры помехи (например, ее интенсивность) неизвестны. Поэтому наиболее типичной в задаче обнаружения являются ситуации, когда распределение наблюдаемого колебания $u(t)$ зависит не только от вида гипотезы, но и от неизвестного параметра λ (или вектора λ). При этом гипотезы H_1 и H_0 оказываются сложными и задача обнаружения состоит в проверке сложной гипотезы при сложной альтернативе.

Пусть $w(u|H_1, \lambda_1)$ и $w(u|H_0, \lambda_0)$ — условные плотности вероятностей наблюдаемого колебания $u(t)$ соответственно при гипотезах H_1 и H_0 , зависящие от неизвестных параметров λ_1 и λ_0 . При байесовском подходе к ре-

шению задачи обнаружения параметры λ_1 и λ_0 рассматриваются как случайные величины, распределения $w(\lambda_1)$ и $w(\lambda_0)$ которых известны. При этом можно найти усредненные плотности вероятностей:

$$\bar{w}(u|H_1) = \int_{\Lambda_1} w(u|H_1, \lambda_1) w(\lambda_1) d\lambda_1, \quad (3.44)$$

$$\bar{w}(u|H_0) = \int_{\Lambda_0} w(u|H_0, \lambda_0) w(\lambda_0) d\lambda_0, \quad (3.45)$$

где Λ_1 и Λ_0 — множества значений параметров λ_1 и λ_0 соответственно. В результате переходим к случаю проверки простой гипотезы при простой альтернативе, когда распределение вероятностей наблюдаемого колебания зависит только от типа гипотезы.

Таким образом, рассматриваемая задача обнаружения сводится к ранее рассмотренной. Оптимальный алгоритм обнаружения записывается следующим образом:

$$\bar{l}(u) = \frac{\bar{w}(u|H_1)}{\bar{w}(u|H_0)} = \frac{\int_{\Lambda_1} w(u|H_1, \lambda_1) w(\lambda_1) d\lambda_1}{\int_{\Lambda_0} w(u|H_0, \lambda_0) w(\lambda_0) d\lambda_0} \geq l_0, \quad (3.46)$$

где порог l_0 зависит от выбранного критерия. Неизвестные параметры λ_1 и λ_0 могут быть векторами, при этом интегралы в выражениях (3.44)—(3.46) будут многократными.

3.3.2. Обнаружение детерминированного сигнала на фоне белого шума

Пусть наблюдаемое колебание имеет вид

$$u(t) = \theta s(t) + n(t),$$

где θ — случайная величина, принимающая значения 1 и 0 с вероятностями p_1 и $p_0 = 1 - p_1$; $s(t)$ — полезный сигнал с известными параметрами; $n(t)$ — гауссовский белый шум с нулевым математическим ожиданием и корреляционной функцией

$$R_n(\tau) = \frac{N_0}{2} \delta(\tau).$$

Найдем алгоритм работы оптимального обнаружителя. Вначале рассмотрим дискретную обработку, когда наблюдения ведутся в дискретные

моменты времени t_1, t_2, \dots, t_n , а решение принимается на основе выборочных значений

$$u(t_i) = u_i = \theta s_i + n_i, \quad i = 1, 2, \dots, n.$$

Для принятия решения необходимо вычислить отношение правдоподобия

$$l(\mathbf{u}) = \frac{w(u_1, \dots, u_n | H_1)}{w(u_1, \dots, u_n | H_0)} = \frac{w(\mathbf{u} | H_1)}{w(\mathbf{u} | H_0)} \quad (3.47)$$

и сравнить его с порогом I_0 .

Найдем распределения $w(\mathbf{u} | H_1)$ и $w(\mathbf{u} | H_0)$.

Пусть справедлива гипотеза H_0 . Тогда наблюдаемое колебание представляет шум, т. е. $u(t) = n(t)$. Поскольку в качестве помехи рассматривается гауссовский белый шум, то распределение любого отсчета наблюдаемого колебания является гауссовским:

$$w(u_i | H_0) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{u_i^2}{2\sigma^2}\right).$$

Учитывая, что отсчеты белого шума статистически независимы, найдем распределение

$$w(u_1, \dots, u_n | H_0) \doteq \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n u_i^2\right). \quad (3.48)$$

При справедливости гипотезы H_1 имеем

$$u_i = s_i + n_i.$$

Используя формулу преобразования плотности вероятности и учитывая, что s_i — детерминированная величина, находим распределение

$$w(u_i | H_1) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(u_i - s_i)^2}{2\sigma^2}\right).$$

Поскольку отсчеты u_i , $i = 1, 2, \dots, n$, по-прежнему оказываются статистически независимыми, то распределение

$$w(u_1, \dots, u_n | H_1) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (u_i - s_i)^2\right). \quad (3.49)$$

Подставив выражения (3.48) и (3.49) в формулу (3.47), получим отношение правдоподобия

$$l(\mathbf{u}) = \exp \left(\frac{1}{\sigma^2} \sum_{i=1}^n u_i s_i - \frac{1}{2\sigma^2} \sum_{i=1}^n s_i^2 \right).$$

Соответственно, алгоритм работы обнаружителя примет вид

$$l(\mathbf{u}) = \exp \left(\frac{1}{\sigma^2} \sum_{i=1}^n u_i s_i - \frac{1}{2\sigma^2} \sum_{i=1}^n s_i^2 \right) \underset{H_0}{\overset{H_1}{\geq}} l_0. \quad (3.50)$$

Учитывая монотонную зависимость показательной функции от аргумента, алгоритм работы обнаружителя (3.50) можно записать в виде неравенств

$$\ln l(\mathbf{u}) = \frac{1}{\sigma^2} \sum_{i=1}^n u_i s_i - \frac{1}{2\sigma^2} \sum_{i=1}^n s_i^2 \underset{H_0}{\overset{H_1}{\geq}} \ln l_0$$

или

$$z' = \frac{1}{\sigma^2} \sum_{i=1}^n u_i s_i \underset{H_0}{\overset{H_1}{\geq}} \ln l_0 + \frac{1}{2\sigma^2} \sum_{i=1}^n s_i^2 = z'_0. \quad (3.51)$$

Из выражения (3.51) следует, что существенной операцией, которую необходимо выполнить при решении задачи обнаружения, является нахождение суммы произведений отсчетов принятого колебания и полезного сигнала.

Перейдем к рассмотрению непрерывной обработки сигнала. При переходе к непрерывному времени можно воспользоваться зависимостью

$$\sigma^2 = \frac{N_0/2}{\Delta t}, \quad \Delta t = t_i - t_{i-1}, \quad (3.52)$$

где $N_0/2$ — спектральная плотность мощности шума. Подставляя выражение (3.52) в формулу (3.51) и переходя к пределу при $\Delta t \rightarrow 0$, находим алгоритм работы обнаружителя непрерывного сигнала

$$z = \frac{2}{N_0} \int_0^{T_c} u(t)s(t)dt \underset{H_0}{\overset{H_1}{\geq}} \ln l_0 + \frac{E}{N_0} = z_0, \quad (3.53)$$

где T_c — длительность полезного сигнала*, а E — его энергия.

* В настоящей главе далее длительность сигнала T будем обозначать без индекса «с».

Заметим, что отношение правдоподобия для непрерывного сигнала имеет вид

$$l(u) = \exp \left[-\frac{E}{N_0} + \frac{2}{N_0} \int_0^T u(t)s(t)dt \right]. \quad (3.54)$$

Из формулы (3.54) следует, что при непрерывной обработке необходимо вычислить интеграл

$$z = \frac{2}{N_0} \int_0^T u(t)s(t)dt \quad (3.55)$$

и сравнить его с известным порогом. Интеграл (3.55) характеризует меру взаимной корреляции между наблюдаемым колебанием $u(t)$ и полезным сигналом $s(t)$ и называется *корреляционным*. Соответственно, *корреляционным* называется обнаружитель, построенный в соответствии с выражением (3.53) и состоящий (рис. 3.5) из перемножителя, интегратора и порогового устройства (ПУ). Перемножитель и интегратор образуют коррелятор.

Определим показатели качества обнаружителя: вероятность ложной тревоги $F_{\text{лт}}$ и вероятность правильного обнаружения $D = 1 - p_{\text{проп}}$. Для этого необходимо знать распределение решающей статистики z при отсутствии и наличии сигнала $s(t)$ на входе обнаружителя.

Пусть справедлива гипотеза H_0 . Тогда

$$z = \frac{2}{N_0} \int_0^T n(t)s(t)dt. \quad (3.56)$$

Из выражения (3.56) следует, что величина z является линейным преобразованием белого гауссовского шума, а следовательно, она имеет гауссовское распределение, которое полностью определяется математическим ожиданием $\mathbf{M}\{z|H_0\}$ и дисперсией σ_z^2 .

Первая характеристика находится тривиально:

$$\mathbf{M}\{z|H_0\} = \mathbf{M}\left\{ \frac{2}{N_0} \int_0^T n(t)s(t)dt \right\} = \frac{2}{N_0} \int_0^T \mathbf{M}\{n(t)\} s(t)dt = 0.$$

Вычисление дисперсии σ_z^2 несколько сложнее:

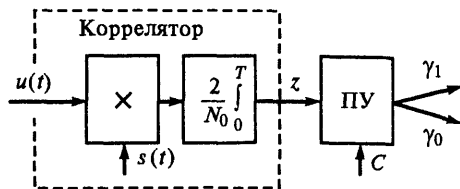


Рис. 3.5. Структурная схема корреляционного обнаружителя детерминированного сигнала

$$\begin{aligned}\sigma_z^2 &= \mathbf{M}\{z^2\} = \mathbf{M}\left\{\left[\frac{2}{N_0} \int_0^T n(t)s(t)dt\right]^2\right\} = \frac{4}{N_0^2} \mathbf{M}\left\{\int_0^T n(t_1)s(t_1)dt_1 \int_0^T n(t_2)s(t_2)dt_2\right\} = \\ &= \frac{4}{N_0^2} \int_0^T \int_0^T \mathbf{M}\{n(t_1)n(t_2)\} s(t_1)s(t_2)dt_1dt_2.\end{aligned}$$

Учитывая, что корреляционная функция белого шума выражается соотношением $\mathbf{M}\{n(t_1)n(t_2)\} = \frac{N_0}{2}\delta(t_2 - t_1)$, находим

$$\sigma_z^2 = \frac{2}{N_0} \int_0^T \int_0^T \delta(t_2 - t_1) s(t_1)s(t_2)dt_1dt_2.$$

Наконец, используя фильтрующее свойство δ -функции, получаем

$$\sigma_z^2 = \frac{2}{N_0} \int_0^T s^2(t_1)dt_1 = \frac{2E}{N_0}. \quad (3.57)$$

Таким образом,

$$w(z|H_0) = \frac{1}{\sqrt{2\pi \cdot 2E/N_0}} \exp\left(-\frac{z^2}{2 \cdot 2E/N_0}\right). \quad (3.58)$$

Пусть теперь справедлива гипотеза H_1 . При этом величина

$$z = \frac{2}{N_0} \int_0^T [s(t) + n(t)]s(t)dt.$$

по-прежнему является результатом линейного преобразования гауссовского шума и, следовательно, имеет гауссовское распределение.

Математическое ожидание и дисперсия величины z определяются формулами

$$\begin{aligned}\mathbf{M}\{z|H_1\} &= \mathbf{M}\left\{\frac{2}{N_0} \int_0^T [s(t) + n(t)]s(t)dt\right\} = \frac{2E}{N_0}, \\ \sigma_z^2 &= 2E/N_0.\end{aligned}$$

Теперь можно записать плотность вероятности величины z :

$$w(z|H_1) = \frac{1}{\sqrt{2\pi \cdot 2E/N_0}} \exp\left(-\frac{(z - 2E/N_0)^2}{2 \cdot 2E/N_0}\right). \quad (3.59)$$

Используя распределения (3.58) и (3.59), находим показатели качества обнаружения:

$$F_{\text{лт}} = P(z > z_0 | H_0) = \int_{z_0}^{\infty} w(z | H_0) dz =$$

$$= \int_{\ln I_0 + E/N_0}^{\infty} \frac{1}{\sqrt{2\pi \cdot 2E/N_0}} \exp\left(-\frac{z^2}{2 \cdot 2E/N_0}\right) dz = 1 - \Phi\left(\frac{\ln I_0 + E/N_0}{\sqrt{2E/N_0}}\right), \quad (3.60)$$

$$D = P(z > z_0 | H_1) = \int_{z_0}^{\infty} w(z | H_1) dz =$$

$$= \int_{\ln I_0 + E/N_0}^{\infty} \frac{1}{\sqrt{2\pi \cdot 2E/N_0}} \exp\left(-\frac{(z - 2E/N_0)^2}{2 \cdot 2E/N_0}\right) dz = 1 - \Phi\left(\frac{\ln I_0 - E/N_0}{\sqrt{2E/N_0}}\right), \quad (3.61)$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$ — интеграл вероятности.

Величины $F_{\text{лт}}$ и $p_{\text{проп}}$ численно равны площадям соответствующих участков под кривыми распределений $w(z | H_0)$ и $w(z | H_1)$ (рис. 3.6).

Используя выражения (3.60) и (3.61), можно для различных $F_{\text{лт}}$ (рис. 3.7, сплошные линии) получить зависимости $D = f(\sqrt{2E/N_0})$, которые называются *характеристиками обнаружения*. Пользуясь этими характеристиками, можно определить необходимое отношение $2E/N_0$, при котором для заданной вероятности $F_{\text{лт}}$ обеспечивается обнаружение сигнала с требуемой вероятностью правильного обнаружения D . Значение этого отношения характеризует так называемый *пороговый сигнал*.

Из полученных результатов вытекает следующий важный вывод: воз-

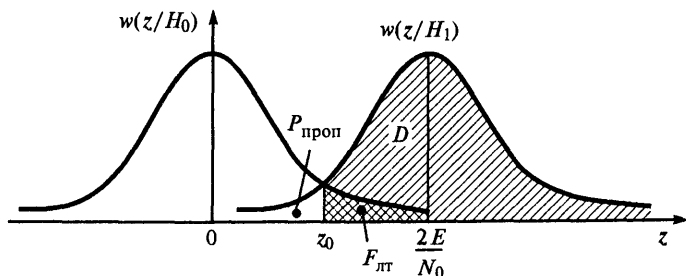


Рис. 3.6. Диаграмма расчета показателей качества обнаружения

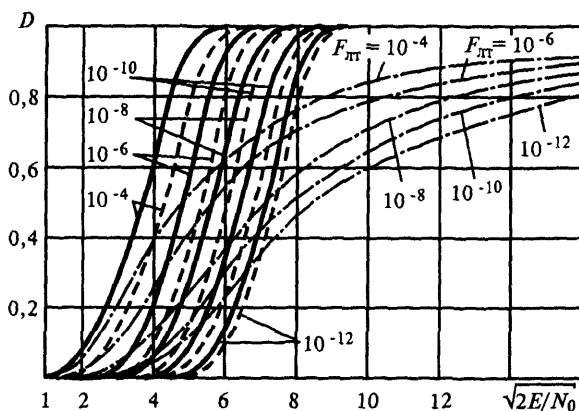


Рис. 3.7. Характеристики оптимального обнаружителя

возможность обнаружения сигнала с заданными вероятностями ошибок $F_{\text{пт}}$ и $p_{\text{проп}}$ при помехе типа белого гауссовского шума зависит только от отношения энергии сигнала E к спектральной плотности мощности шума $N_0/2$ и не зависит от формы сигнала.

Техническая реализация алгоритма обнаружения, представленного выражением (3.55), при использовании корреляционного приемника не является единственно возможной. Корреляционный интеграл (3.55) может быть сформирован с помощью линейного фильтра. Действительно, сигнал на выходе линейного фильтра с импульсной характеристикой $h(t)$ в момент $t = T$ имеет вид

$$u_{\text{вых}}(T) = \int_0^T u(t)h(T-t)dt.$$

Если импульсную характеристику фильтра выбрать такой, что

$$h(T-t) = \frac{2}{N_0} s(t),$$

или, что то же самое,

$$h(t) = \frac{2}{N_0} s(T-t), \quad (3.62)$$

то $u_{\text{вых}}(T)$ будет совпадать с корреляционным интегралом выражения (3.55).

Линейный фильтр, импульсная характеристика которого определяется выражением (3.62), или, в более общем случае, выражением

$$h(t) = as(t_0 - t),$$

где a — постоянная величина и $t_0 \geq T$, называется *согласованным*.

При фильтровой интерпретации интеграла (3.62) обнаружитель (рис. 3.8) состоит из согласованного фильтра (СФ) и порогового устройства (ПУ).

Рассмотрим основные характеристики согласованного фильтра. Импульсная характеристика была приведена ранее. Она дает временное описание согласованного фильтра. Найдем комплексную частотную характеристику:

$$K(j\omega) = \int_{-\infty}^{\infty} h(t) \exp(-j\omega t) dt = \int_{-\infty}^{\infty} as(t_0 - t) \exp(-j\omega t) dt.$$

Введя новую переменную $\tau = t_0 - t$, получим

$$K(j\omega) = \exp(-j\omega t_0) \int_{-\infty}^{\infty} as(\tau) \exp(j\omega \tau) d\tau = a \exp(-j\omega t_0) S^*(j\omega), \quad (3.63)$$

где $S^*(j\omega)$ — функция, комплексно-сопряженная спектральной плотности сигнала $s(t)$.

Из формулы (3.63) следует, что модуль комплексной частотной характеристики (амплитудно-частотная характеристика) согласованного фильтра имеет вид

$$|K(j\omega)| = a|S(j\omega)|, \quad (3.64)$$

а его фазовая характеристика определяется следующим образом:

$$\varphi(\omega) = -\varphi_c(\omega) - \omega t_0. \quad (3.65)$$

Равенство (3.64) означает, что амплитудно-частотная характеристика согласованного фильтра совпадает с точностью до постоянного множителя с амплитудным спектром полезного сигнала, а следовательно, согласованный фильтр избирательно пропускает частотные составляющие сигнала: чем больше интенсивность составляющей, тем с большим весом она передается.

Фазовая характеристика согласованного фильтра состоит из двух частей, одна из которых равна фазовому спектру сигнала $\varphi_c(\omega)$, взятому с обратным знаком, а вторая равна $-\omega t_0$. Благодаря такой характеристике фазовые сдвиги между частотными составляющими полезного сигнала компен-

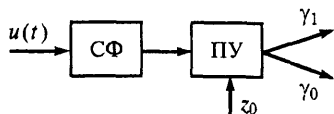


Рис. 3.8. Структурная схема фильтрового обнаружителя детерминированного сигнала

сируются так, что в момент $t = t_0$ все составляющие полезного сигнала оказываются в одной фазе и значение сигнала на выходе согласованного фильтра оказывается максимально возможным. Именно в этот момент принимается решение.

Определим отношение сигнал—шум на выходе согласованного фильтра. При наличии полезного сигнала на входе согласованного фильтра выходной сигнал в момент $t = t_0 = T$ имеет вид

$$\begin{aligned} u_{\text{вых}}(T) &= \int_0^T (s(t) + n(t))h(T-t)dt = \int_0^T (s(t) + n(t))as(t)dt = \\ &= a \int_0^T s^2(t)dt + a \int_0^T n(t)s(t)dt. \end{aligned} \quad (3.66)$$

Математическое ожидание и дисперсия выходного сигнала определяются формулами

$$\mathbf{M}\{u_{\text{вых}}(T)\} = a \int_0^T s^2(t)dt = aE, \quad (3.67)$$

$$\sigma_{\text{вых}}^2 = \mathbf{M}\{[u_{\text{вых}}(T) - aE]^2\} = \mathbf{M}\left\{\left[a \int_0^T n(t)s(t)dt\right]^2\right\} = a^2 E N_0/2. \quad (3.68)$$

С учетом выражений (3.67) и (3.68) отношение сигнал—шум по мощности в момент $t = T$ можно записать в виде

$$q^2 = \frac{[\mathbf{M}\{u_{\text{вых}}(T)\}]^2}{\sigma_{\text{вых}}^2} = \frac{2E}{N_0}. \quad (3.69)$$

Таким образом, отношение сигнал—шум на выходе согласованного фильтра в момент $t = T$ определяется только энергией сигнала и спектральной плотностью мощности шума и не зависит от формы сигнала.

Можно показать, что отношение сигнал—шум на выходе любого линейного фильтра при действии на него детерминированного сигнала и гауссовского белого шума удовлетворяет неравенству

$$q^2 \leq 2E/N_0. \quad (3.70)$$

Из формул (3.69) и (3.70) следует, что согласованный фильтр является оптимальным фильтром, обеспечивающим максимальное отношение сигнал—шум.

Заметим, что максимальное отношение сигнал—шум на выходе коррелятора (см. рис. 3.5) также равно $2E/N_0$.

Действительно,

$$q^2 = \frac{[\mathbf{M}\{z|H_1\}]^2}{\sigma_z^2} = \frac{(2E/N_0)^2}{2E/N_0} = \frac{2E}{N_0}. \quad (3.71)$$

Найдем вид сигнала на выходе согласованного фильтра при отсутствии помехи:

$$u_{\text{вых}}(t) = \int_0^t s(\tau)h(t-\tau)d\tau = \int_0^t s(\tau)as(t_0-t+\tau)d\tau. \quad (3.72)$$

Нетрудно заметить, что относительно момента времени $t=t_0$ выражение (3.72) с точностью до постоянного множителя совпадает с выражением для корреляционной функции полезного сигнала $s(t)$. Таким образом, сигнал на выходе согласованного фильтра при отсутствии шума совпадает по форме с корреляционной функцией сигнала, с которым фильтр согласован.

В заключение отметим, что рассмотренные схемы реализации оптимального обнаружителя (корреляционная и фильтровая) обеспечивают одинаковые показатели качества обнаружения. Независимо от вида приемника (корреляционный приемник или согласованный фильтр) информация о полезном сигнале содержится на приемной стороне и используется при приеме. В случае корреляционного приемника на приемной стороне генерируется копия сигнала $s(t)$, при согласованном фильтре информация о сигнале заложена в его комплексной частотной характеристике (или импульсной характеристике).

При использовании корреляционного приемника необходимо точно знать время прихода сигнала. Если оно неизвестно, то приходится применять набор корреляторов, каждый из которых соответствует определенному интервалу возможного времени прихода сигнала. При использовании согласованного фильтра нет необходимости в знании времени прихода сигнала. Действительно, фильтр, согласованный с некоторым сигналом $s(t)$, будет согласован и с сигналом $s(t-t')$, сдвинутым по времени на t' . Изменение времени прихода сигнала приводит к изменению времени достижения сигналом на выходе фильтра своего максимального значения.

3.3.3. Обнаружение сигнала со случайной начальной фазой на фоне белого шума

Пусть колебание на входе обнаружителя имеет вид

$$u(t) = \theta s(t, \varphi) + n(t) = \theta[s \cos(\omega_0 t + \varphi)] + n(t),$$

где $s(t, \varphi)$ — полезный сигнал, у которого все параметры, за исключением начальной фазы φ , известны.

Начальная фаза φ рассматривается как случайная величина с известным распределением. При отсутствии информации об априорном распределении начальной фазы вполне естественно считать это распределение равномерным:

$$w(\varphi) = \frac{1}{2\pi}, \quad 0 \leq \varphi \leq 2\pi. \quad (3.73)$$

Поскольку сигнал в рассмотренном случае может принимать различные значения в зависимости от φ , гипотеза H_1 является сложной. Поэтому алгоритм обнаружения согласно формуле (3.46) сводится к вычислению отношения правдоподобия

$$\bar{l}(u) = \frac{\int_0^{2\pi} w(u|H_1, \varphi) w(\varphi) d\varphi}{w(u|H_0)} \quad (3.74)$$

и сравнению его с порогом l_0 . При записи выражения (3.74) учитывалось, что распределение $w(u|H_0)$ не содержит неизвестных параметров.

Отношение $\bar{l}(u)$ можно найти следующим образом:

$$\bar{l}(u) = \int_0^{2\pi} l(u|\varphi) w(\varphi) d\varphi, \quad (3.75)$$

где $l(u|\varphi)$ — условное отношение правдоподобия (отношение правдоподобия при условии, что начальная фаза равна φ). Оно совпадает с отношением правдоподобия для детерминированного сигнала $s(t, \varphi)$, где φ — фиксированная величина.

С учетом выражения (3.54) имеем

$$\begin{aligned} l(u|\varphi) &= \exp \left(-\frac{E}{N_0} + \frac{2}{N_0} \int_0^T u(t) s(t, \varphi) dt \right) = \\ &= \exp \left(-\frac{E}{N_0} + \frac{2}{N_0} \int_0^T u(t) S \cos(\omega_0 t + \varphi) dt \right), \end{aligned} \quad (3.76)$$

где S — амплитуда.

Подставляя выражения (3.73) и (3.76) в формулу (3.75), получаем

$$\bar{l}(u) = \exp\left(-\frac{E}{N_0}\right) \int_0^{2\pi} \exp\left(\frac{2}{N_0} \int_0^T u(t) S \cos(\omega_0 t + \varphi) dt\right) \frac{1}{2\pi} d\varphi. \quad (3.77)$$

Преобразуем сначала корреляционный интеграл:

$$\begin{aligned} z &= \frac{2}{N_0} \int_0^T u(t) S \cos(\omega_0 t + \varphi) dt = \frac{2}{N_0} \int_0^T u(t) S (\cos(\omega_0 t) \cos \varphi - \sin(\omega_0 t) \sin \varphi) dt = \\ &= \frac{2}{N_0} \cos \varphi \int_0^T u(t) S \cos(\omega_0 t) dt - \\ &- \frac{2}{N_0} \sin \varphi \int_0^T u(t) S \sin(\omega_0 t) dt = \frac{2}{N_0} (z_1 \cos \varphi - z_2 \sin \varphi), \end{aligned}$$

$$\text{где } z_1 = \int_0^T u(t) S \cos(\omega_0 t) dt, \quad z_2 = \int_0^T u(t) S \sin(\omega_0 t) dt.$$

Введем обозначение $Z = \sqrt{z_1^2 + z_2^2}$; тогда

$$z = \frac{2}{N_0} Z \left(\frac{z_1}{Z} \cos \varphi - \frac{z_2}{Z} \sin \varphi \right).$$

Отношения z_1/Z и z_2/Z можно рассматривать соответственно как $\cos \theta$ и $\sin \theta$. Поэтому

$$z = \frac{2Z}{N_0} (\cos \theta \cos \varphi - \sin \theta \sin \varphi) = \frac{2Z}{N_0} \cos(\theta + \varphi). \quad (3.78)$$

Подставляя выражение (3.78) в формулу (3.77), находим

$$\bar{l}(u) = \exp\left(-\frac{E}{N_0}\right) \int_0^{2\pi} \frac{1}{2\pi} \exp\left(\frac{2Z}{N_0} \cos(\theta + \varphi)\right) d\varphi. \quad (3.79)$$

Интеграл в правой части выражения (3.79) является модифицированной функцией Бесселя нулевого порядка с аргументом $2Z/N_0$. Таким образом,

$$\bar{l}(u) = \exp\left(-\frac{E}{N_0}\right) I_0\left(\frac{2Z}{N_0}\right), \quad (3.80)$$

и алгоритм работы оптимального обнаружителя сигнала со случайной начальной фазой имеет вид

$$\bar{l}(u) = \exp\left(-\frac{E}{N_0}\right) I_0\left(\frac{2Z}{N_0}\right) \stackrel{H_1}{\geq} I_0$$

или

$$\ln I_0\left(\frac{2Z}{N_0}\right) \stackrel{H_1}{\geq} \ln I_0 + \frac{E}{N_0} = z_0. \quad (3.81)$$

Учитывая, что функция $I_0(Z)$ монотонна, алгоритм (3.81) можно заменить равносильным:

$$Z \stackrel{H_1}{\geq}_{H_0} C, \quad (3.82)$$

где порог C выбирается из обеспечения требуемой вероятности ложной тревоги $F_{\text{лт}}$.

Таким образом, существенной операцией, которую необходимо выполнить при решении задачи обнаружения сигнала со случайной начальной фазой, является вычисление величины

$$Z = \sqrt{z_1^2 + z_2^2} = \sqrt{\left[\int_0^T u(t)S \cos(\omega_0 t) dt\right]^2 + \left[\int_0^T u(t)S \sin(\omega_0 t) dt\right]^2}. \quad (3.83)$$

Оптимальный обнаружитель (рис. 3.9), построенный в соответствии с формулами (3.82) и (3.83), представляет собой схему с двумя квадратурными каналами, в каждом из которых осуществляются операции умножения, интегрирования и возведения в квадрат. Опорные напряжения в каналах сдвинуты по фазе относительно друг друга на 90° . Наличие двух квадратурных каналов позволяет получить результат обработки Z , не зависящий от истинного значения начальной фазы.

Учитывая, что величина Z является значением огибающей сигнала на

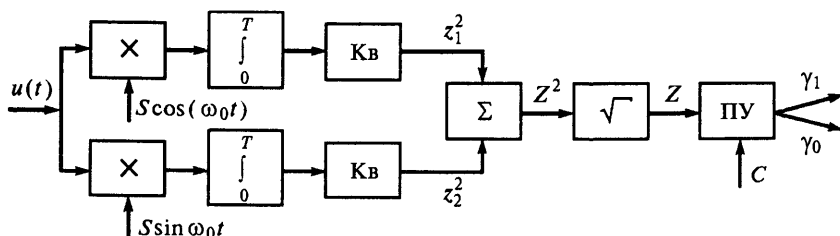


Рис. 3.9. Структурная схема корреляционного обнаружителя сигнала со случайной начальной фазой

выходе приемника, вычисляющего корреляционный интеграл $\int_0^T u(t)s(t, \varphi) d\varphi$,

возможна другая схема реализации оптимального обнаружителя (рис. 3.10). Она включает согласованный фильтр (СФ), детектор огибающей (ДО) и пороговое устройство (ПУ). Решение принимается в момент $t = T$.

Определим показатели качества обнаружения. Для этого надо знать законы распределения $w(Z|H_0)$ и $w(Z|H_1)$. Поскольку колебание на выходе согласованного фильтра представляет собой узкополосный гауссовский процесс, его огибающая будет распределена по закону Рэлея при гипотезе H_0 и по закону Райса при альтернативе H_1 . Соответственно, распределения величины Z при гипотезах H_0 и H_1 будут определяться формулами

$$w(Z|H_0) = \frac{Z}{\sigma^2} \exp\left(-\frac{Z^2}{2\sigma^2}\right), \quad (3.84)$$

$$w(Z|H_1) = \frac{Z}{\sigma^2} \exp\left(-\frac{Z^2 + E^2}{2\sigma^2}\right) I_0\left(\frac{ZE}{\sigma^2}\right), \quad (3.85)$$

где $\sigma^2 = EN_0/2$ — дисперсия шума на выходе согласованного фильтра.

На практике удобно пользоваться распределениями нормированной величины $v = Z/\sigma$. При этом

$$w(v|H_0) = v \exp\left(-\frac{v^2}{2}\right), \quad (3.86)$$

$$w(v|H_1) = v \exp\left(-\frac{v^2 + 2E/N_0}{2}\right) I_0\left(v\sqrt{\frac{2E}{N_0}}\right). \quad (3.87)$$

Используя распределения (3.86) и (3.87), найдем показатели качества обнаружения:

$$F_{\text{лт}} = \int_{C_n}^{\infty} w(v|H_0) dv = \int_{C_n}^{\infty} v \exp(-v^2/2) dv = \exp(-C_n^2/2), \quad (3.88)$$

$$D = \int_{C_n}^{\infty} w(v|H_1) dv = \int_{C_n}^{\infty} v \exp\left(-\frac{v^2 + 2E/N_0}{2}\right) I_0\left(v\sqrt{\frac{2E}{N_0}}\right) dv, \quad (3.89)$$

где $C_n = C/\sigma$ — нормированный порог.

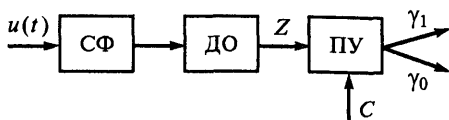


Рис. 3.10. Структурная схема фильтрового обнаружителя сигнала со случайной начальной фазой

Интеграл (3.89) нельзя выразить через элементарные функции. Поэтому вычисление вероятности правильного обнаружения производится численными методами.

Используя выражения (3.88) и (3.89), можно найти характеристики обнаружения сигнала (на рис. 3.7 они изображены штриховыми линиями). Сравнение их с характеристиками обнаружения детерминированного сигнала показывает, что незнание начальной фазы приводит к ухудшению показателей качества обнаружения, т. е. требует несколько большее пороговое отношение сигнал—шум.

3.3.4. Обнаружение сигнала со случайными амплитудой и начальной фазой на фоне белого шума

Пусть обнаруживаемый сигнал имеет вид

$$s(t, a, \varphi) = a \cos(\omega_0 t - \varphi), \quad (3.90)$$

где a и φ — случайные амплитуда и начальная фаза. На практике обычно величины a и φ статистически независимы и имеют соответственно рэлеевское и равномерное распределения:

$$w(a) = \frac{a}{\sigma_a^2} \exp\left(-\frac{a^2}{2\sigma_a^2}\right), \quad a \geq 0;$$

$$w(\varphi) = \frac{1}{2\pi}, \quad 0 \leq \varphi \leq 2\pi. \quad (3.91)$$

Для рассматриваемого случая отношение правдоподобия (3.46) можно представить выражением

$$\bar{I}(u) = \int_0^\infty \int_0^{2\pi} I(u|a, \varphi) w(a) w(\varphi) d\varphi da. \quad (3.92)$$

Условное отношение правдоподобия $I(u|a, \varphi)$ находим по формуле (3.54) при замене $s(t)$ на $s(t, a, \varphi)$:

$$I(u|a, \varphi) = \exp\left\{\frac{2}{N_0} \int_0^T u(t) s(t, a, \varphi) dt - \frac{1}{N_0} \int_0^T s^2(t, a, \varphi) dt\right\}.$$

Интегрируя правую часть выражения (3.92) по переменной φ , находим отношение правдоподобия

$$\bar{I}(u) = \int_0^\infty \exp\left(-\frac{a^2 E_1}{N_0}\right) I_0\left(\frac{2aZ_1}{N_0}\right) w(a) da, \quad (3.93)$$

где E_1 — энергия полезного сигнала при $a=1$, $Z_1 = Z/a$ — величина, определяемая формулой (3.83) при $S=1$ (значение огибающей сигнала на выходе оптимального приемника, настроенного на прием сигнала с амплитудой, равной 1).

Подставляя в формулу (3.93) выражение (3.91) и учитывая, что

$$\int_0^{\infty} x \exp(-\alpha x^2) I_0(\beta x) dx = \frac{1}{2\alpha} \exp\left(\frac{\beta^2}{4\alpha}\right), \quad (3.94)$$

окончательно получаем

$$\bar{l}(u) = \frac{N_0}{N_0 + \bar{E}} \exp\left[\frac{2\sigma_a^2 Z_1^2}{N_0(N_0 + \bar{E})}\right], \quad (3.95)$$

где

$$\bar{E} = \mathbf{M}\{a^2 E_1\} = E_1 \mathbf{M}\{a^2\} = E_1 \cdot 2\sigma_a^2 \quad (3.96)$$

— средняя энергия сигнала.

Поскольку отношение правдоподобия $\bar{l}(u)$ является монотонной функцией Z_1 , алгоритм работы оптимального обнаружителя можно представить в виде неравенств

$$Z_1 \geq C_1. \quad (3.97)$$

Из выражения (3.97) следует, что существенной операцией, которую необходимо выполнить при решении задачи обнаружения сигнала со случайными амплитудой и начальной фазой, является вычисление огибающей сигнала на выходе оптимального приемника, настроенного на прием сигнала с амплитудой, равной 1. Структурные схемы обнаружителя будут подобны схемам, изображенным на рис. 3.9 и 3.10.

Оценим показатели качества обнаружителя. При гипотезе H_0 распределение величины Z_1 будет рэлеевским:

$$w(Z_1|H_0) = \frac{Z_1}{\sigma_1^2} \exp\left(-\frac{Z_1^2}{2\sigma_1^2}\right),$$

где $\sigma_1^2 = E_1 N_0/2$. Вероятность ложной тревоги находится следующим образом:

$$F_{\text{лт}} = \int_{C_1}^{\infty} w(Z_1|H_0) dZ_1 = \int_{C_1}^{\infty} \frac{Z_1}{\sigma_1^2} \exp\left(-\frac{Z_1^2}{2\sigma_1^2}\right) dZ_1 = \exp\left(-\frac{C_1^2}{2\sigma_1^2}\right). \quad (3.98)$$

Для определения вероятности правильного обнаружения необходимо знать распределение

$$\bar{w}(Z_1|H_1) = \int_0^\infty w(Z_1|H_1, a) w(a) da, \quad (3.99)$$

где

$$\begin{aligned} \bar{w}(Z_1|H_1, a) &= \frac{Z_1}{\sigma_1^2} \exp\left(-\frac{Z_1^2 + a^2 E_1^2}{2\sigma_1^2}\right) I_0\left(\frac{Z_1 a E_1}{\sigma_1^2}\right) = \\ &= \frac{2Z_1}{N_0 E_1} \exp\left(-\frac{Z_1^2 + a^2 E_1^2}{N_0 E_1}\right) I_0\left(\frac{2Z_1 a}{N_0}\right). \end{aligned} \quad (3.100)$$

Подставляя выражения (3.91) и (3.100) в формулу (3.99) и используя значение интеграла (3.94), находим

$$\bar{w}(Z_1|H_1) = \frac{2Z_1}{N_0 E_1 + 2\sigma_a^2 E_1^2} \exp\left(-\frac{Z_1^2}{N_0 E_1 + 2\sigma_a^2 E_1^2}\right). \quad (3.101)$$

Вероятность правильного обнаружения определяется выражением

$$D = \int_{C_1}^\infty \bar{w}(Z_1|H_1) dZ_1 = \exp\left(-\frac{C_1^2}{N_0 E_1 (1 + 2\sigma_a^2 E_1 / N_0)}\right). \quad (3.102)$$

Используя выражения (3.98) и (3.102), можно найти характеристики обнаружения сигнала со случайными амплитудой и начальной фазой. Сравнение их с характеристиками обнаружения нефлуктуирующего радиосигнала необходимо проводить при равенстве средней энергии флуктуирующего сигнала и энергии нефлуктуирующего сигнала. Если принять амплитудный коэффициент нефлуктуирующего сигнала a равным 1, то, как видно из формулы (3.96), необходимо положить $\sigma_a^2 = 1/2$. При этом с учетом выражений (3.98) и (3.102) имеем

$$D = F_{\text{лт}}^{1/(1+E/N_0)}. \quad (3.103)$$

Зависимость (3.103) показана на рис. 3.7 штрих-пунктирными линиями. Из характеристик, изображенных на рис. 3.7, видно, что для обнаружения сигнала со случайными амплитудой и начальной фазой требуется значительно большее пороговое отношение сигнал—шум (при $D \geq 0,8$) по сравнению с обнаружением сигнала с известной амплитудой.

Таким образом, незнание амплитуды и начальной фазы сигнала приводит к существенному ухудшению показателей качества обнаружения. Для обеспечения заданных показателей необходимо увеличить энергию сигнала.

Заметим, что в ряде случаев в схемах рассмотренных обнаружителей вместо согласованных фильтров применяют более простые так называемые *квазиоптимальные фильтры*, у которых полоса пропускания выбирается из условия достижения как можно большего отношения сигнал—шум на выходе. При этом, конечно, помехоустойчивость обнаружителей уменьшается, так как квазиоптимальные фильтры лишь частично согласованы с полезным сигналом. Однако при использовании простых сигналов, например одиночных радиоимпульсов без внутриимпульсной частотной или фазовой модуляции, это уменьшение незначительно (энергетические потери не превышают 17 %).

3.3.5. Обнаружение пачки сигналов

Сигнал, отраженный от цели, представляет собой последовательность радиоимпульсов (пачку радиоимпульсов). Число N импульсов в пачке зависит от времени облучения $T_{\text{обл}}$ цели, периода повторения радиоимпульсов $T_{\text{п}}$ и определяется отношением

$$N = T_{\text{обл}} / T_{\text{п}}.$$

В режиме кругового обзора

$$T_{\text{обл}} = \Delta\theta_{0,5} / \Omega,$$

где $\Delta\theta_{0,5}$ — ширина диаграммы направленности по уровню 0,5, Ω — угловая скорость вращения луча антенны.

Огибающая пачки определяется формой диаграммы направленности антенны. В некоторых случаях удобно в качестве аппроксимации использовать пачку с прямоугольной огибающей, имеющую ту же энергию, но меньшее число импульсов.

Пачка радиоимпульсов называется *когерентной*, если начальные фазы радиоимпульсов связаны детерминированной зависимостью. Если начальные фазы радиоимпульсов случайны, то пачка радиоимпульсов называется *некогерентной*.

Когерентность пачки отраженных радиоимпульсов возможна при выполнении следующих условий:

- 1) излученные радиоимпульсы должны быть когерентны;
- 2) сдвиг фаз при отражении сигнала от объекта и при распространении его в среде должен быть одинаков для всех радиоимпульсов;
- 3) интервал корреляции флуктуаций должен быть больше длительности пачки:

$$\tau_{\text{к.фл}} > T_{\text{пачк}} = NT_{\text{п}}.$$

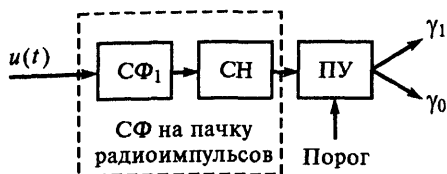


Рис. 3.11. Структурная схема оптимального обнаружителя когерентной пачки импульсов

Когерентность излучаемых сигналов можно обеспечить их формированием по схеме: высокостабильный генератор—модулятор—усилитель мощности.

При решении задачи обнаружения когерентной пачки радиоимпульсов рассматривают следующие основные модели:

- 1) пачка радиоимпульсов с полностью известными параметрами (детерминированная модель);
- 2) пачка радиоимпульсов со случайной начальной фазой;
- 3) пачка радиоимпульсов со случайными начальной фазой и амплитудой.

Синтез оптимальных обнаружителей для указанных выше моделей когерентной пачки радиоимпульсов практически ничем не отличается от синтеза обнаружителей одиночных сигналов, рассмотренного в п. 3.3.3. Структурные схемы обнаружителей когерентных пачек радиоимпульсов совпадают со структурными схемами обнаружителей, представленных на рис. 3.5 и 3.8. Однако в корреляционных схемах в качестве опорного сигнала необходимо использовать сигналы, описывающие пачки радиоимпульсов, а в фильтровых схемах оптимальный фильтр должен быть согласован с пачкой радиоимпульсов.

Согласованный фильтр для когерентной пачки радиоимпульсов можно представить (рис. 3.11) в виде каскадного соединения фильтра, согласованного с одиночным радиоимпульсом ($СФ_1$), и синхронного накопителя (СН). Синхронный накопитель (рис. 3.12) состоит из линии задержки с отводами, усилителей с коэффициентами усиления K_i , пропорциональными амплитудам импульсов пачки, и суммирующего устройства (Σ).

На рис. 3.13 представлена структурная схема оптимального обнаружителя когерентных пачек радиоимпульсов со случайной начальной фазой. Эта схема отличается от схемы, представленной на рис. 3.11, наличием детектора огибающей (ДО).

Заметим, что техническая реализация синхронного накопителя на радиочастоте представляет собой сложную задачу. Нестабильность параметров линии задержки и не-

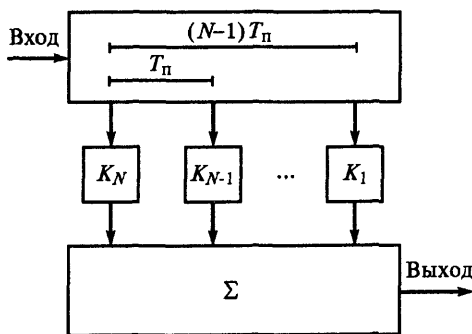


Рис. 3.12. Структурная схема согласованного фильтра для когерентной пачки радиоимпульсов

точность расположения отводов не позволяют в полной мере реализовать синхронное накопление сигнала, что приводит к снижению помехоустойчивости обнаружителя.

Для когерентной пачки радиоимпульсов со случайной начальной фазой можно использовать фильтрационно-корреляционную схему построения обнаружителя (рис. 3.14), включающую фильтр, согласованный с одиночным радиоимпульсом ($СФ_1$); два квадратурных канала, каждый из которых состоит из фазового детектора (ФД), синхронного накопителя (СН) и квадратора (Кв); сумматор (Σ) и пороговое устройство (ПУ) [28]. Частота ω_0 опорных колебаний $S \cos(\omega_0 t + \varphi_i)$ и $S \sin(\omega_0 t + \varphi_i)$ должна совпадать с частотой радиоимпульсов на выходе $СФ_1$, а их фаза φ_i должна меняться от одного периода следования к другому на величину $\omega_0 T_n$.

Рассмотренная схема обнаружителя обеспечивает такую же помехоустойчивость, что и схема, показанная на рис. 3.13, однако она проще в реализации, так как синхронное накопление осуществляется на видеочастоте.

Для обработки когерентной пачки радиоимпульсов со случайной начальной фазой можно применять также корреляционно-фильтровую схему (рис. 3.15) [30]. Пусть $s_1(t)$ — радиоимпульс длительности $T = (N-1)T_n + \tau_n$. Тогда когерентную пачку радиоимпульсов $s(t)$ можно представить формулой

$$s(t) = s_1(t) \times s_{\text{стр}}(t),$$

где $s_{\text{стр}}(t)$ — периодическая последовательность видеоимпульсов (стробов), длительность которых τ_n , а период повторения T_n .

С учетом изложенного выше корреляционный интеграл имеет вид

$$\int_0^T u(t)s(t)dt = \int_0^T u(t)s_{\text{стр}}(t)s_1(t)dt = \int_0^T u_{\text{стр}}(t)s_1(t)dt. \quad (3.104)$$

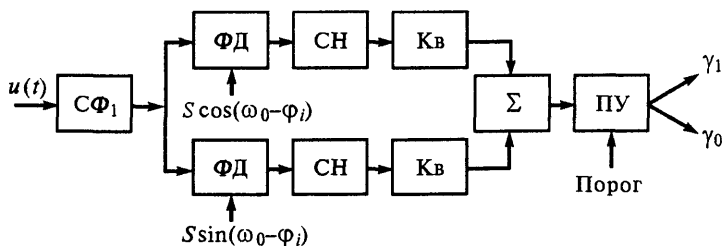


Рис. 3.14. Фильтрационно-корреляционная схема построения обнаружителя

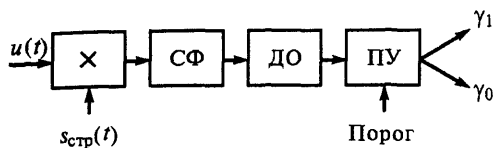


Рис. 3.15. Корреляционно-фильтровая схема построения обнаружителя

Из формулы (3.104) следует, что процедура обнаружения когерентной пачки радиоимпульсов со случайной начальной фазой сводится к перемножению принятого сигнала $u(t)$ и последовательности видеоимпульсов $s_{\text{стр}}(t)$ с дальней-

шей обработкой в фильтре, согласованном с сигналом $s_1(t)$.

В отличие от ранее рассмотренных схем (рис. 3.12—3.14) при корреляционно-фильтровой обработке отсутствуют сложные накопители на основе многоотводных линий задержек и квадратурные каналы. В то же время существенным недостатком корреляционно-фильтровой схемы является то, что она не обладает свойством инвариантности к запаздыванию сигнала. Поэтому при обработке когерентной пачки радиоимпульсов с неизвестным моментом прихода необходимо применять многоканальную схему. Фильтровая и фильтрационно-корреляционные схемы обработки инвариантны к запаздыванию сигнала и, соответственно, могут быть одноканальными по дальности.

Помехоустойчивость рассмотренных оптимальных обнаружителей когерентной пачки радиоимпульсов можно оценить по ранее полученным формулам (3.60), (3.61), (3.88), (3.89). При этом энергию E понимают как суммарную энергию всех импульсов, образующих пачку. В случае прямоугольной пачки (когда амплитуды всех импульсов одинаковы) отношение сигнал—шум на выходе синхронного накопителя (см. рис. 3.13, 3.14) будет в N раз больше, чем на выходе фильтра, согласованного с одиночным радиоимпульсом.

Рассмотрим теперь обнаружение некогерентной пачки радиоимпульсов, когда амплитуды радиоимпульсов известны, а их начальные фазы представляют собой независимые случайные величины с равномерным законом распределения. Для решения этой задачи необходимо вычислить отношение правдоподобия, которое с учетом статистической независимости начальных фаз радиоимпульсов можно записать в виде

$$l(u) = \prod_{i=1}^N \bar{l}_i(u),$$

где $\bar{l}_i(u)$ — отношение правдоподобия для i -го радиоимпульса, определяемое для рассматриваемого случая формулой (3.80), N — число импульсов в пачке.

Таким образом,

$$l(u) = \prod_{i=1}^N \exp\left(-\frac{E_i}{N_0}\right) I_0\left(\frac{2Z_i}{N_0}\right), \quad (3.105)$$

где E_i — энергия i -го радиоимпульса, Z_i — огибающая i -го радиоимпульса на выходе фильтра, согласованного с одиночным радиоимпульсом.

Прологарифмируем выражение (3.105) и перенесем постоянные величины в правую часть. Тогда алгоритм работы оптимального обнаружителя можно записать в виде

$$\sum_{i=1}^N \ln I_0\left(\frac{2Z_i}{N_0}\right) \underset{H_0}{\gtrless} C, \quad (3.106)$$

где C — величина порога, выбираемая из условия обеспечения заданной вероятности ложной тревоги.

В соответствии с алгоритмом (3.106) на рис. 3.16 представлена структурная схема оптимального обнаружителя некогерентной пачки радиоимпульсов с известными амплитудами и случайными фазами. Обнаружитель состоит из фильтра, согласованного с одиночным радиоимпульсом (СФ₁), детектора огибающей (ДО), синхронного накопителя (СН) и порогового устройства (ПУ).

На практике вместо согласованного фильтра СФ₁ часто используют квазиоптимальный фильтр, полоса пропускания которого выбирается из условия получения максимально возможного отношения сигнал—шум на выходе. В частности, при прямоугольной частотной характеристике фильтра ширина полосы пропускания определяется формулой Сифорова:

$$\Delta f_{\text{опт}} \approx 1,37/\tau_{\text{и}}.$$

Характеристика детектора огибающей в соответствии с (3.106) должна описываться функцией $f(x) = \ln I_0(x)$. Поскольку

$$\ln I_0(x) \approx \begin{cases} \frac{1}{4}x^2, & x \ll 1, \\ x, & x \gg 1, \end{cases}$$

то характеристика ДО при малых отношениях сигнал—шум является квадратичной, а при больших отношениях сигнал—шум — линейной. Соответственно, ре-

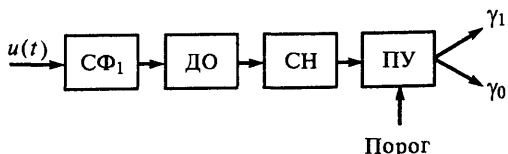


Рис. 3.16. Структурная схема оптимального обнаружителя некогерентной пачки импульсов

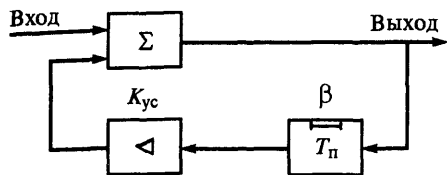


Рис. 3.17. Структурная схема рециркулятора

Однако из-за сложности технической реализации такого СН при обработке пачек большой длительности на практике для обеспечения синхронного накопления импульсов обычно используют рециркулятор (рис. 3.17) [28]. Он представляет собой схему с положительной обратной связью и состоит из линии задержки на время T_n , усилителя с коэффициентом усиления K_{yc} , обеспечивающего компенсацию затухания сигнала при его распространении по линии задержки, и суммирующего устройства. Для устойчивой работы схемы необходимо выполнение условия

$$K_{yc} \beta < 1,$$

где β — коэффициент затухания. Обычно $K_{yc} \beta \approx 0,8 \dots 0,95$.

Заметим, что при накоплении импульсов в рециркуляторе наблюдается эффект насыщения, проявляющийся в уменьшении вклада каждого последующего накапливаемого импульса в суммарный сигнал. Это приводит к снижению качества обнаружения. Подобный недостаток свойствен всем используемым на практике аналоговым накопителям (потенциалоскопам, магнитным барабанами, индикаторам на электронно-лучевых трубках, приборам с зарядовой связью и др.) и лишь цифровые накопители свободны от него.

Оценим помехоустойчивость рассмотренного обнаружителя. Алгоритм работы оптимального обнаружителя некогерентной пачки радиоимпульсов с известными амплитудами можно представить при больших отношениях сигнал—шум в виде

$$Z = \sum_{i=1}^N Z_i \underset{H_0}{\overset{H_1}{\gtrless}} C$$

и при малых — в виде

$$Z = \sum_{i=1}^N Z_i^2 \underset{H_0}{\overset{H_1}{\gtrless}} C',$$

где Z_i , $i=1, \dots, N$, — независимые случайные величины, плотности вероятности которых определяются формулами (3.84) и (3.85).

Зная законы распределения величин Z_i , $i = 1, \dots, N$, можно найти плотности вероятности решающих статистик Z для гипотез H_0 и H_1 . Интегрируя эти распределения в пределах от порогового значения до ∞ , находим условные вероятности ложной тревоги и правильного обнаружения.

При малых отношениях сигнал—шум распределение решающей статистики при гипотезе H_0 подчиняется χ^2 -закону:

$$w(Z|H_0) = \frac{Z^{N-1}}{2^N (N-1)!} \exp\left(-\frac{Z}{2}\right),$$

и вероятность ложной тревоги будет выражаться формулой

$$F_{\text{лт}} = \frac{1}{2^N (N-1)!} \int_C^\infty Z^{N-1} \exp\left(-\frac{Z}{2}\right) dZ. \quad (3.107)$$

Плотность вероятности решающей статистики Z при гипотезе H_1 определяется следующим выражением [31]:

$$w(Z|H_1) = \frac{1}{2} \left(\frac{Z}{Nq^2/2} \right)^{(N-1)/2} \exp\left(-\left(\frac{Nq^2}{2} + \frac{Z}{2}\right)\right) I_{N-1}\left(\sqrt{Nq^2 Z}\right).$$

Соответственно, вероятность правильного приема имеет вид

$$D = \frac{1}{2!} \int_C^\infty \left(\frac{Z}{Nq^2} \right)^{(N-1)/2} \exp\left(-\left(\frac{Nq^2}{2} + \frac{Z}{2}\right)\right) I_{N-1}\left(\sqrt{Nq^2 Z}\right) dZ, \quad (3.108)$$

где q^2 — отношение сигнал—шум на выходе фильтра, согласованного с одиночным радиоимпульсом; $I_{N-1}(x)$ — модифицированная функция Бесселя порядка $N-1$.

Вычисление вероятности ложной тревоги $F_{\text{лт}}$ (3.107) и вероятности правильного обнаружения D существенно упрощается при использовании неполной функции Торонто [31]:

$$T_a(m, N, r) = 2r^{N-m+1} \exp(-r^2) \int_0^a t^{m-N} \exp(-t^2) I_N(2rt) dt.$$

Если сделать замену переменной $t = \sqrt{Z/2}$, то выражения (3.107) и (3.108) можно представить в виде

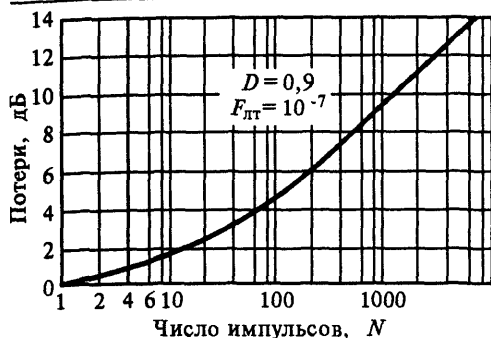


Рис. 3.18. Зависимость потерь некогерентного накопления по отношению к когерентному

неэффективно по сравнению с когерентным. В случае прямоугольной пачки при слабых сигналах и больших N отношение сигнал—шум на выходе некогерентного накопителя увеличивается пропорционально \sqrt{N} , а при когерентном накоплении — пропорционально N .

На рис. 3.18 представлена зависимость потерь некогерентного накопления по отношению к когерентному при $D = 0,9$, $F_{\text{лт}} = 10^{-7}$ от числа импульсов в пачке.

Используя решающую статистику $Z = \sum_{i=1}^N Z_i$, можно получить практи-

чески такую же помехоустойчивость, что и при статистике $Z = \sum_{i=1}^N Z_i^2$.

Рассмотрим обнаружение некогерентной пачки радиоимпульсов со случайными амплитудами. Различают следующие виды флуктуаций некогерентной пачки радиоимпульсов:

- независимые (быстрые); характеризуются тем, что амплитуды радиоимпульсов статистически независимы;
- дружные (медленные); характеризуются тем, что амплитуды радиоимпульсов полностью коррелированы;
- частично коррелированные; характеризуются тем, что интервал корреляции сравним с периодом повторения радиоимпульсов $T_{\text{п}}$ и длительностью пачки.

Пусть флуктуации амплитуд радиоимпульсов независимы и описываются законом Рэлея (3.91), начальные фазы распределены по равномерному закону. Тогда отношение функций правдоподобия для рассматриваемого случая с учетом (3.95) можно записать в виде

$$F_{\text{лт}} = 1 - T_a(2N - 1, N - 1, 0),$$

$$D = 1 - T_a\left(2N - 1, N - 1, \sqrt{Nq^2/2}\right),$$

где $a = \sqrt{C'/2}$.

Помехоустойчивость оптимального обнаружителя некогерентной пачки радиоимпульсов с известными амплитудами существенно ниже помехоустойчивости обнаружителя когерентной пачки. Это объясняется тем, что некогерентное накопление менее эффек-

$$l(u) = \prod_{i=1}^N \bar{l}_i(u) = \prod_{i=1}^N \frac{N_0}{N_0 + \bar{E}} \exp\left(\frac{2\sigma_a^2 Z_{li}^2}{N_0(N_0 + \bar{E})}\right), \quad (3.109)$$

где Z_{li} — огибающая i -го импульса на выходе оптимального приемника, настроенного на прием радиоимпульса единичной амплитуды.

Логарифмируя (3.109), алгоритм работы оптимального обнаружителя можно представить в виде

$$\sum_{i=1}^N Z_{li}^2 \underset{H_0}{\overset{H_1}{\geq}} C.$$

Структурная схема оптимального обнаружителя некогерентной пачки независимо флуктуирующих радиоимпульсов совпадает со схемой, показанной на рис 3.16. Характеристика детектора огибающей должна быть квадратичной.

При дружных флуктуациях амплитуд радиоимпульсов отношение правдоподобия можно записать в виде

$$l(u) = \int \prod_{i=1}^N \bar{l}_i(u|a) w(a) da,$$

где $\bar{l}_i(u|a) = \exp\left(-\frac{a^2 E_1}{N_0}\right) I_0\left(\frac{2aZ_{li}}{N_0}\right)$ — отношение правдоподобия для i -го радиоимпульса при условии, что амплитуда радиоимпульса равна a ; $w(a)$ — закон распределения амплитуд, обычно рэлеевский (3.91); начальные фазы распределены по равномерному закону.

Структурная схема обнаружителя оказывается такой же, как на рис. 3.16. Характеристика детектора огибающей должна быть линейной при больших отношениях сигнал—шум и квадратичной при малых.

При частично коррелированных флуктуациях синтез оптимального обнаружителя представляет более сложную задачу по сравнению с рассмотренными выше случаями.

Оценка помехоустойчивости оптимальных обнаружителей некогерентной пачки флуктуирующих радиоимпульсов достаточно сложна и осуществляется обычно численными методами на ЭВМ. Заметим, что при увеличении интервала корреляции флуктуаций характеристики обнаружения ухудшаются [28]. Поэтому на практике стараются обеспечить независимость флуктуаций радиоимпульсов пачки. Этого можно, в частности, добиться изменением частоты несущей от импульса к импульсу.

3.4. Различение сигналов

3.4.1. Оптимальные алгоритмы различения сигналов

Пусть колебание на входе приемника является суммой помехи и одного из сигналов $s_1(t), s_2(t), \dots, s_m(t)$. Задача состоит в том, чтобы по принятой реализации $u(t)$ решить, какой из сигналов передается. Такая задача характерна для систем связи.

Критерием оптимальности может служить один из ранее рассмотренных критериев. Поскольку в связанных системах ошибки в приеме различных сигналов, как правило, одинаково нежелательны, наиболее подходящим является критерий максимума апостериорной вероятности, реализуемый оптимальным приемником Котельникова.

При различении двух сигналов $s_0(t)$ и $s_1(t)$ в соответствии с критерием максимума апостериорной вероятности принимается решение в пользу сигнала $s_1(t)$, если отношение правдоподобия $l(u)$ удовлетворяет условию

$$l(u) = \frac{w(u|s_1)}{w(u|s_0)} \geq \frac{p_0}{p_1}, \quad (3.110)$$

где p_0 и p_1 — априорные вероятности появления сигналов $s_0(t)$ и $s_1(t)$.

При различении m равновероятных сигналов решение принимается в пользу $s_l(t)$, если

$$w(u|s_l) > w(u|s_i), \quad i = 1, 2, \dots, m, \quad i \neq l. \quad (3.111)$$

3.4.2. Различение двух детерминированных сигналов на фоне белого шума

Пусть сигнал на входе приемника имеет вид

$$u(t) = \theta s_1(t) + (1 - \theta) s_0(t) + n(t),$$

где θ — случайная величина, принимающая значения 0 и 1 с вероятностями p_0 и p_1 соответственно; $s_0(t)$ и $s_1(t)$ — полезные сигналы с известными параметрами; $n(t)$ — стационарный гауссовский белый шум с нулевым математическим ожиданием и корреляционной функцией $R_n(\tau) = \frac{N_0}{2} \delta(\tau)$.

Аналогично п. 3.3.2 можно показать, что отношение правдоподобия имеет вид

$$l(u) = \exp \left(-\frac{E_1}{N_0} + \frac{2}{N_0} \int_0^T u(t) s_1(t) dt \right) / \exp \left(-\frac{E_0}{N_0} + \frac{2}{N_0} \int_0^T u(t) s_0(t) dt \right),$$

где E_0 и E_1 — энергии сигналов $s_0(t)$ и $s_1(t)$, а логарифм отношения правдоподобия описывается формулой

$$\ln l(u) = -\frac{E_1 - E_0}{N_0} + \frac{2}{N_0} \int_0^T u(t)(s_1(t) - s_0(t)) dt.$$

Отсюда получаем, что решение принимается в пользу сигнала $s_1(t)$, если

$$z = \frac{2}{N_0} \int_0^T u(t)(s_1(t) - s_0(t)) dt \geq \ln \frac{p_0}{p_1} + \frac{E_1 - E_0}{N_0} = C_1. \quad (3.112)$$

Для симметричного канала, когда $p_0 = p_1 = 0,5$ и $E_0 = E_1 = E$, порог C_1 равен нулю и алгоритм различения принимает вид

$$z \underset{s_0}{\overset{s_1}{\geq}} 0. \quad (3.113)$$

Структурная схема оптимального когерентного приемника, соответствующая (3.112), представлена на рис. 3.19. Верхний и нижний корреляторы могут быть заменены согласованными фильтрами с импульсными характеристиками $h_1(t) = s_1(T-t)$ и $h_0(t) = s_0(T-t)$ соответственно (рис. 3.20).

Средняя вероятность ошибки записывается в виде

$$P_{\text{ош}} = p_0 P_{\text{ош}}(s_0) + p_1 P_{\text{ош}}(s_1),$$

где $P_{\text{ош}}(s_i)$ — вероятность ошибки при передаче сигнала $s_i(t)$, $i = 0, 1$.

При $p_0 = p_1 = 0,5$

$$P_{\text{ош}} = 0,5 (P_{\text{ош}}(s_0) + P_{\text{ош}}(s_1)). \quad (3.114)$$

Условные вероятности $P_{\text{ош}}(s_0)$ и $P_{\text{ош}}(s_1)$ определяются через распределения z при наличии соответственно сигналов $s_0(t)$ и $s_1(t)$ следующим образом:

$$\begin{aligned} P_{\text{ош}}(s_0) &= \int_{z < C_1} w(z|s_0) dz, \\ P_{\text{ош}}(s_1) &= \int_{z > C_1} w(z|s_1) dz. \end{aligned} \quad (3.115)$$

Легко видеть, что при сигнале $s_1(t)$ величина

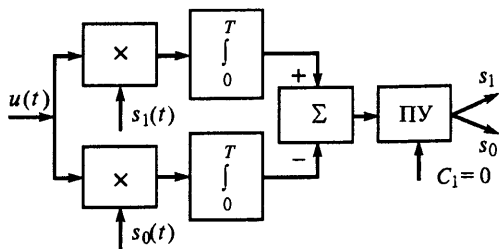


Рис. 3.19. Структурная схема корреляционного различителя двух детерминированных сигналов

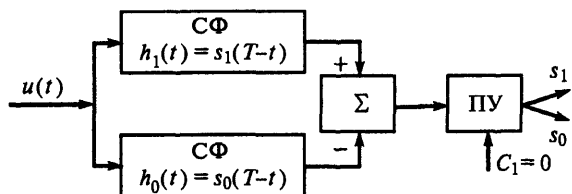


Рис. 3.20. Структурная схема фильтрового различителя двух детерминированных сигналов

$$z = z_1 = \frac{2}{N_0} \int_0^T (s_1(t) + n(t))(s_1(t) - s_0(t)) dt$$

распределена по нормальному закону с математическим ожиданием и дисперсией

$$\mathbf{M}\{z_1\} = \frac{2E}{N_0}(1 - r_s), \quad \sigma_{z_1}^2 = \frac{4E}{N_0}(1 - r_s),$$

где $r_s = \frac{1}{E} \int_0^T s_1(t)s_0(t)dt$ — коэффициент взаимной корреляции сигналов $s_1(t)$ и $s_0(t)$.

Аналогично, при сигнале $s_0(t)$ величина

$$z = z_0 = \frac{2}{N_0} \int_0^T (s_0(t) + n(t))(s_1(t) - s_0(t)) dt$$

распределена по нормальному закону с математическим ожиданием и дисперсией

$$\mathbf{M}\{z_0\} = -\frac{2E}{N_0}(1 - r_s), \quad \sigma_{z_0}^2 = \frac{4E}{N_0}(1 - r_s).$$

На рис. 3.21 представлены кривые распределений $w(z_1) = w(z|s_1)$ и $w(z_0) = w(z|s_0)$.

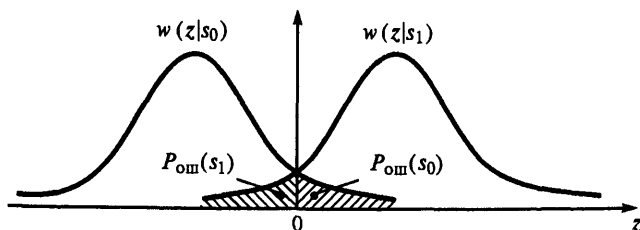


Рис. 3.21. Плотность вероятности решающей статистики при различении детерминированных сигналов

С учетом соотношений (3.113)—(3.115) и распределений $w(z_1)$ и $w(z_0)$ имеем

$$P_{\text{ош}} = 0,5 \left(\int_0^\infty w(z|s_0) dz + \int_{-\infty}^0 w(z|s_1) dz \right) =$$

$$= 1 - \Phi \left(\sqrt{\frac{E}{N_0}} (1 - r_s) \right) = 1 - \Phi(\sqrt{1 - r_s} h), \quad (3.116)$$

где $\Phi(z) = \frac{1}{2\pi} \int_{-\infty}^z e^{-t^2/2} dt$ — интеграл вероятности, $h^2 = \frac{E}{N_0}$.

Из формулы (3.116) видно, что средняя вероятность ошибки зависит не только от энергии сигнала и спектральной плотности мощности шума, но и от коэффициента взаимной корреляции между сигналами, т. е. от используемой системы сигналов. Интеграл вероятности $\Phi(z)$ является монотонно возрастающей функцией. Поэтому при одном и том же отношении E/N_0 помехоустойчивость системы оказывается тем выше, чем меньше коэффициент взаимной корреляции r_s .

Поскольку $-1 \leq r_s \leq 1$, то наибольшей помехоустойчивостью обладают сигналы с коэффициентом корреляции $r_s = -1$. Они имеют одинаковую форму, но противоположные знаки и называются *противоположными*. Для них

$$P_{\text{ош}} = 1 - \Phi(\sqrt{2E/N_0}) = 1 - \Phi(\sqrt{2}h). \quad (3.117)$$

Примером противоположных сигналов являются фазоманипулированные сигналы с манипуляцией фазы на π :

$$s_1(t) = S_0 \cos \omega_0 t, \quad s_0(t) = S_0 \cos(\omega_0 t + \pi), \quad 0 \leq t \leq T.$$

Меньшей помехоустойчивостью обладают ортогональные сигналы ($r_s = 0$). Для них

$$P_{\text{ош}} = 1 - \Phi(\sqrt{E/N_0}) = 1 - \Phi(h). \quad (3.118)$$

Сигналы при $r_s = 1$ являются одинаковыми, т. е. $s_1(t) = s_0(t)$, и их невозможно различить. Для них $P_{\text{ош}} = 0,5$.

Примером ортогональных сигналов являются фазоманипулированные сигналы с манипуляцией фазы на $\pi/2$:

$$s_1(t) = S_0 \cos \omega_0 t, \quad s_0(t) = S_0 \cos(\omega_0 t + \pi/2), \quad 0 \leq t \leq T.$$

Ортогональные сигналы можно получить, используя частотную манипуляцию. Действительно, в этом случае

$$s_1(t) = S_0 \cos(\omega_1 t - \varphi_1), \quad s_0(t) = S_0 \cos(\omega_0 t - \varphi_2).$$

При $\varphi_1 = \varphi_0 = \varphi$ коэффициент взаимной корреляции между этими сигналами имеет вид

$$r_s = \frac{\sin((\omega_1 - \omega_0)T)}{(\omega_1 - \omega_0)T} + \frac{\sin((\omega_1 + \omega_0)T - 2\varphi) + \sin 2\varphi}{(\omega_1 + \omega_0)T}.$$

При выполнении условия $(\omega_1 - \omega_0)T = 2\pi k$, $k = 1, 2, \dots$, коэффициент корреляции r_s равен нулю и сигналы оказываются ортогональными. На практике параметры ω_1 , ω_0 и T выбирают так, чтобы $(\omega_1 - \omega_0)T \gg 1$. При этом $r_s \approx 0$.

Заметим, что минимальное значение коэффициента взаимной корреляции r_s между частотно-манипулированными сигналами равно $-1/(1,5\pi)$. Оно достигается, когда $(\omega_1 - \omega_0)T = 1,5\pi$. При этом вероятность ошибок

$$P_{\text{ош}} \approx 1 - \Phi\left(\sqrt{\frac{1,21E}{N_0}}\right) = 1 - \Phi(\sqrt{1,21h}).$$

Оценим помехоустойчивость системы передачи, использующей амплитудно-манипулированные сигналы

$$s_1(t) = S_0 \cos(\omega_0 t + \varphi), \quad s_0(t) = 0, \quad 0 \leq t \leq T.$$

Алгоритм различения сигналов в рассматриваемом случае принимает вид

$$z = \frac{2}{N_0} \int_0^T u(t) s_1(t) dt \underset{s_0}{\overset{s_1}{\gtrless}} \frac{E}{N_0} + \ln \frac{p_0}{p_1} = C_1.$$

Плотности распределения вероятностей $w(z|s_1)$ и $w(z|s_0)$ описываются гауссовскими законами с параметрами $\mathbf{M}\{z\} = 2E/N_0$, $\sigma_z^2 = 2E/N_0$ и $\mathbf{M}\{z\} = 0$, $\sigma_z^2 = 2E/N_0$ соответственно.

При $p_1 = p_0 = 0,5$ средняя вероятность ошибки принимает вид

$$P_{\text{ош}} = 0,5 \left[\int_{-\infty}^{C_1} w(z|s_1) dz + \int_{C_1}^{\infty} w(z|s_0) dz \right].$$

Учитывая, что $C_1 = E/N_0$, и пренебрегая первым интегралом, находим

$$P_{\text{ош}} = 1 - \Phi(0,5\sqrt{2E/N_0}) = 1 - \Phi(h/\sqrt{2}). \quad (3.119)$$

На рис. 3.22 представлены зависимости вероятности ошибок от отношения $2E/N_0$ для фазо- (ФМ), частотно- (ЧМ) и амплитудно-манипулированных (АМ) сигналов (сплошные линии), вычисленные соответственно по формулам (3.117)—(3.119).

Таким образом, наибольшей потенциальной помехоустойчивостью обладают фазоманипулированные сигналы. Они обеспечивают энергетический выигрыш в два раза по сравнению с частотно-манипулированными сигналами и в четыре раза по сравнению с амплитудно-манипулированными сигналами. Частотно-манипулированные сигналы обеспечивают энергетический выигрыш в энергии сигнала по сравнению с амплитудно-манипулированными сигналами в два раза. Однако следует иметь в виду, что в отличие от фазовой и частотной манипуляции при амплитудной манипуляции передается только один сигнал. Поэтому если исходить из средне-энергетических затрат, то нетрудно видеть, что системы с АМ и ЧМ сигналами обладают одинаковой помехоустойчивостью.

Заметим, что величина $\sqrt{2E(1-r_s)}$ представляет расстояние между сигналами:

$$d = \left(\int_0^T [s_1(t) - s_0(t)]^2 dt \right)^{1/2}.$$

При этом формулу (3.116) можно записать в виде

$$P_{\text{ош}} = 1 - \Phi(d/\sqrt{2N_0}). \quad (3.120)$$

Из соотношения (3.120) следует, что при действии в канале гауссовского белого шума вероятность ошибки зависит только от расстояния между сигналами и спектральной плотности мощности шума. Этот вывод оказывается справедливым и для случая различения m сигналов ($m > 2$).

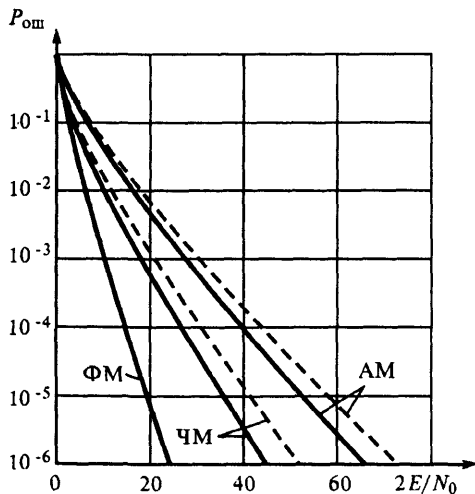


Рис. 3.22. Зависимость вероятности ошибки различения детерминированных сигналов при АМ, ЧМ и ФМ (сплошные линии) и различения сигналов со случайной начальной фазой при АМ и ЧМ (пунктирные линии)

При высоких требованиях к помехоустойчивости ($P_{\text{ош}} < 10^{-3}$) вероятность ошибки удобно определять по приближенной формуле:

$$P_{\text{ош}} \approx \frac{1}{\sqrt{2\pi} \sqrt{h^2(1-r_s)}} \exp\left(-\frac{h^2(1-r_s)}{2}\right), \quad (3.121)$$

которая получается при асимптотическом представлении интеграла вероятности:

$$\Phi(x) \approx 1 - \frac{\exp(-x^2/2)}{\sqrt{2\pi}x}. \quad (3.122)$$

Точность вычислений по формуле (3.121) не меньше 10 %, если $\sqrt{h^2(1-r_s)} \geq 3$.

Как указывалось ранее, ФМ сигналы обеспечивают наибольшую помехоустойчивость. Тем не менее, они практически не используются в системах передачи информации из-за трудностей реализации демодуляторов, связанных с созданием опорного колебания, имеющего неизменную начальную фазу.

В существующих системах передачи информации опорный сигнал формируется из принимаемого сигнала. В системах с фазовой манипуляцией задача затрудняется тем, что при равновероятных сигналах в их спектре отсутствует составляющая с частотой несущей и ее невозможно получить методом фильтрации. В этих случаях приходится применять способы формирования опорного колебания, основанные на снятии манипуляции принятого сигнала [23]. Однако всем им присущ одинаковый недостаток: при воздействии помех возможны скачки фазы опорного колебания на π , что приводит к инвертированию принимаемых символов (символ 1 регистрируется как 0, а символ 0 — как 1). Возникает так называемое явление «обратной работы», которое будет продолжаться до следующего скачка фазы.

Эффективным средством борьбы с явлением «обратной работы» является применение метода относительной фазовой модуляции (ОФМ), предложенного впервые Н. Т. Петровичем. Идея метода ОФМ состоит в том, что информация в сигнале определяется не абсолютным значением начальной фазы сигнала, как при обычной ФМ, а разностью $\Delta\phi$ начальных фаз двух соседних сигналов: $\Delta\phi = 0$, если передается символ 0, $\Delta\phi = 1$, если передается символ 1.

Формирователь ОФМ сигнала (рис. 3.23) состоит из относительного кодера (сумматор по модулю два (М2) и линия задержки на время T) и фазового манипулятора (ФМ). Работа кодера происходит в соответствии с правилом

$$b_k = a_k \oplus b_{k-1},$$

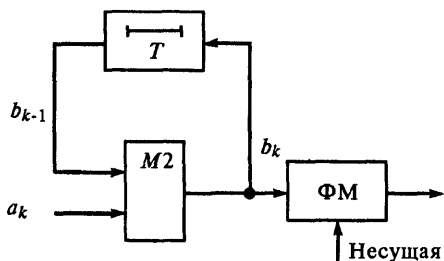


Рис. 3.23. Структурная схема формирователя ОФМ сигнала

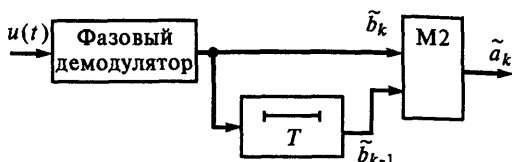


Рис. 3.24. Структурная схема оптимального демодулятора ОФМ сигнала

где $a_1, a_2, \dots, a_k, \dots$ — последовательность информационных символов;
 $b_1, b_2, \dots, b_k, \dots$ — последовательность символов на выходе кодера.

Оптимальный демодулятор (рис. 3.24) состоит из фазового демодулятора и относительного декодера (сумматора по модулю два (M2) и линии задержки на время T). Задача декодера — восстановить информационные символы. Это осуществляется в соответствии с правилом

$$\tilde{a}_k = \tilde{b}_k \oplus \tilde{b}_{k-1},$$

где \tilde{b}_k — k -й принятый символ. Нетрудно убедиться в том, что при каждом случайном скачке фазы опорного колебания в данном случае будет ошибочно принят только один символ, т. е. явления «обратной работы» не будет наблюдаться.

Помехоустойчивость демодулятора ОФМ сигналов легко определяется из следующих соображений. Очевидно, что ошибка в приеме информационного символа будет происходить в двух возможных случаях:

- символ \tilde{b}_k принят правильно, а символ \tilde{b}_{k-1} — ошибочно;
- символ \tilde{b}_k принят ошибочно, а символ \tilde{b}_{k-1} — правильно.

Вероятность каждого из этих событий равна $P_{\text{ош.ФМ}}(1 - P_{\text{ош.ФМ}})$, где $P_{\text{ош.ФМ}}$ — вероятность ошибочного приема символа при ФМ, определяемая выражением (3.117). Следовательно, вероятность ошибки приема символа при ОФМ имеет вид

$$P_{\text{ош.ОФМ}} = 2P_{\text{ош.ФМ}}(1 - P_{\text{ош.ФМ}}) \approx 2(1 - \Phi(\sqrt{2E/N_0})) = 2(1 - \Phi(\sqrt{2h})). \quad (3.123)$$

Таким образом, платой за устранение явления «обратной работы» при применении ОФМ является удвоение вероятности ошибки по сравнению с ФМ. Заметим, что энергетический проигрыш метода ОФМ методу ФМ не превосходит 1 дБ.

3.4.3. Различение m детерминированных сигналов на фоне белого шума

Пусть принятый сигнал имеет вид $u(t) = s_i(t) + n(t)$, $0 \leq t \leq T$, где $s_i(t)$, $i = 1, 2, \dots, m$ ($m > 2$), — возможные полезные сигналы на входе приемника; $n(t)$ — помеха типа белого гауссовского шума.

Предположим, что вероятность передачи любого сигнала равна $1/m$. Тогда решение, какой из сигналов $s_i(t)$, $i = 1, 2, \dots, m$, был передан, принимается на основе анализа $m - 1$ неравенств (3.111), которые для рассматриваемого случая можно переписать в виде

$$\frac{2}{N_0} \int_0^T u(t) s_l(t) dt - \frac{E_l}{N_0} > \frac{2}{N_0} \int_0^T u(t) s_i(t) dt - \frac{E_i}{N_0}, \quad i = 1, 2, \dots, m, \quad i \neq l;$$

или при $E_1 = E_2 = \dots = E_m = E$

$$\int_0^T u(t) s_l(t) dt > \int_0^T u(t) s_i(t) dt, \quad i = 1, 2, \dots, m, \quad i \neq l. \quad (3.124)$$

В соответствии с (3.124) оптимальный различитель m сигналов состоит из m корреляторов (рис. 3.25, а) или из m согласованных фильтров (рис. 3.25, б) и решающего устройства (РУ).

Оценим помехоустойчивость различителя. Очевидно, что ошибка при приеме сигнала возникает тогда, когда неравенства (3.124) не выполняются хотя бы для одного $i \neq l$. Пусть z_1, z_2, \dots, z_m — напряжения на выходах каналов различителя, а $w(z_1, z_2, \dots, z_m | s_l)$ — m -мерная плотность вероятности совокупности случайных величин z_1, z_2, \dots, z_m при условии, что на

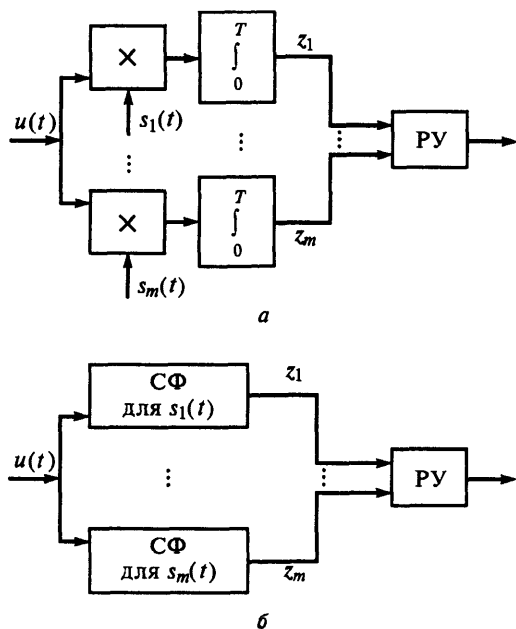


Рис. 3.25. Структурная схема:

а — корреляционный оптимальный различитель m детерминированных сигналов; б — фильтровый оптимальный различитель m детерминированных сигналов

входе приемника действует сигнал $s_l(t)$. Тогда с учетом алгоритма работы оптимального различителя вероятность правильного приема сигнала $s_l(t)$ определяется следующим образом:

$$P_{\text{пр}}(s_l) = \int_{-\infty}^{\infty} dz_l \int_{-\infty}^{z_l} \dots \int_{-\infty}^{z_l} w(z_1, z_2, \dots, z_m | s_l) dz_1 dz_2 \dots dz_{l-1} dz_{l+1} \dots dz_m. \quad (3.125)$$

Соответственно, вероятность ошибки имеет вид

$$P_{\text{ош}}(s_l) = 1 - P_{\text{пр}}(s_l). \quad (3.126)$$

Вероятность ошибки $P_{\text{ош}}$ зависит от ансамбля применяемых сигналов $s_i(t)$, $i = 1, 2, \dots, m$. Существует бесконечно большое число систем, отличающихся индивидуальными и совместными свойствами сигналов. Представляет интерес система сигналов, обеспечивающая максимальную помехоустойчивость при заданных априорных условиях передачи.

Определить помехоустойчивость m -ичных систем в общем случае трудно. Однако для равновероятных симплексных, ортогональных и биортогональных сигналов выражение (3.125) существенно упрощается и сводится к однократному интегралу, который можно оценить с помощью численных методов.

Рассмотрим сначала системы передачи с ортогональными сигналами. Пусть сигнал на входе приемника имеет вид

$$u(t) = s_l(t) + n(t), \quad 0 \leq t \leq T.$$

Тогда напряжение на выходе l -го канала $z_l = \int_0^T u(t)s_l(t)dt$ является гауссовской случайной величиной с математическим ожиданием $M\{z_l\} = E$ и дисперсией $\sigma_{z_l}^2 = EN_0/2$, а напряжения на выходах остальных каналов будут являться гауссовскими случайными величинами с нулевыми математическими ожиданиями и дисперсиями, равными $EN_0/2$.

Нетрудно показать, что в рассматриваемом случае величины z_1, z_2, \dots, z_m являются некоррелированными, а следовательно, с учетом их распределения и статистически независимыми. При этом m -мерную плотность вероятности можно записать в виде

$$w_m(z_1, z_2, \dots, z_m) = w(z_1)w(z_2)\dots w(z_m), \quad (3.127)$$

где

$$w(z_l) = \frac{1}{\sqrt{2\pi}\sqrt{EN_0/2}} \exp\left(-\frac{(z_l - E)^2}{EN_0}\right), \quad (3.128)$$

$$w(z_j) = \frac{1}{\sqrt{2\pi}\sqrt{EN_0/2}} \exp\left(-\frac{z_j^2}{EN_0}\right), \quad j = 1, 2, \dots, m, \quad j \neq l. \quad (3.129)$$

Подставляя соотношения (3.127)—(3.129) в (3.125), после преобразований получаем [21, 23]

$$\begin{aligned} P_{\text{пр}}(s_l) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{z^2}{2}\right) \Phi^{m-1}\left(z + \sqrt{\frac{2E}{N_0}}\right) dz = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(z - \sqrt{\frac{2E}{N_0}}\right)^2\right) \Phi^{m-1}(z) dz, \end{aligned}$$

где $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{t^2}{2}\right) dt$ — интеграл вероятности.

Нетрудно видеть, что вероятность правильного приема оказывается одинаковой для всех сигналов $s_i(t)$, $i = 1, 2, \dots, m$. Поэтому полная вероятность ошибки имеет вид

$$P_{\text{ош}} = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(z - \sqrt{\frac{2E}{N_0}}\right)^2\right) \Phi^{m-1}(z) dz. \quad (3.130)$$

Из (3.130) следует, что с увеличением количества сигналов m вероятность ошибки возрастает. Физически это объясняется увеличением вероятности превышения шумом на выходе какого-либо канала (в момент принятия решения) напряжения на выходе канала, принимающего полезный сигнал. Однако это не означает, что потенциальная помехоустойчивость m -ичных систем меньше, чем двоичных. При сравнении систем необходимо иметь в виду, что каждый равновероятный m -ичный сигнал несет в $\log_2 m$ раз большее количество информации, чем двоичный сигнал, или при той же скорости передачи информации имеет в $\log_2 m$ большую длительность.

На рис. 3.26 построены зависимости вероятности ошибки при когерентном приеме m ортогональных сигналов (сплошные линии) от отношения E_6/N_0 , где $E_6 = E/\log_2 m$ — энергия, затрачиваемая на 1 бит информации. Для сравнения представлена также зависимость вероятности ошибки

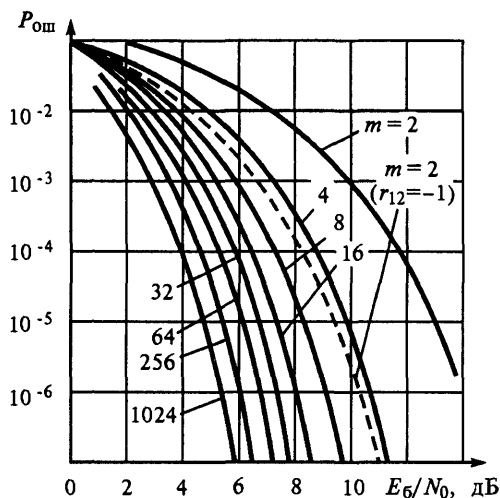


Рис. 3.26. Зависимость вероятности ошибки различения m детерминированных ортогональных сигналов (сплошные линии) и двух детерминированных противоположных сигналов (пунктирные линии)

от отношения E_b/N_0 для двух противоположных сигналов (штриховая линия).

Системы ортогональных сигналов с $m > 2$ позволяют обеспечить при одинаковой скорости передачи информации существенный выигрыш в энергетике по сравнению с двоичными сигналами [23]; например, при $m = 32$ и $P_{\text{ош}} = 10^{-5}$ он составляет почти два раза.

Расплатой за энергетический выигрыш является увеличение ширины полосы частот, занимаемой системой, и усложнение приемника, который для сигналов с одинаковыми энергиями содержит m корреляторов или согласованных фильтров (по числу сигналов) и решающее устройство.

Определим вероятность ошибки $P_{\text{ош}}$ в системе с симплексными сигналами $s_i(t)$, $i = 1, 2, \dots, m$; она связана простым соотношением с вероятностью ошибки для ортогональных сигналов. Действительно, пусть $s_i(t)$, $i = 1, 2, \dots, m$, — симплексные сигналы. Образует новый ансамбль сигналов длительностью $T(1 + |r_0|)$:

$$s'_i(t) = \begin{cases} s_i(t), & 0 \leq t \leq T, \\ \sqrt{E/T}, & T < t \leq T(1 + |r_0|), \end{cases} \quad (3.131)$$

где $r_0 \geq -1/(m-1)$.

Сигналы (3.131) являются ортогональными:

$$\int_0^{T(1+|r_0|)} s'_i(t)s'_j(t)dt = \int_0^T s_i(t)s_j(t)dt + \int_T^{T(1+|r_0|)} \frac{E}{T}dt = Er_0 + E|r_0| = 0, \quad i \neq j.$$

Энергия каждого сигнала $s'_i(t)$ равна $E(1 + |r_0|)$. Учитывая, что расстояния между сигналами обеих систем $\{s_i(t)\}$ и $\{s'_i(t)\}$ одинаковы, можно утверждать, что вероятность ошибки для исходного ансамбля сигналов

$s_i(t)$, $i = 1, 2, \dots, m$, равна вероятности ошибки для ортогонального ансамбля сигналов с энергией $E(1+|r_0|)$. Таким образом, вероятность ошибки для симплексных сигналов имеет вид

$$P_{\text{ош}} = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} \left(z - \sqrt{\frac{2E}{N_0} \frac{m}{m-1}} \right)^2 \right] \Phi^{m-1}(z) dz.$$

Зависимость вероятности ошибки от отношения E_6/N_0 для симплексных сигналов можно проследить по рис. 3.26, если на оси абсцисс вместо E_6/N_0 отложить величину $\frac{E_6}{N_0} \frac{m}{m-1}$.

Помехоустойчивость симплексных сигналов выше, чем ортогональных. Однако это различие уменьшается с увеличением m и при $m \gg 1$ помехоустойчивость обеих ансамблей оказывается практически одинаковой.

Оценим помехоустойчивость систем передачи с биортогональными сигналами. Оптимальный различитель биортогональных сигналов состоит из набора $m/2$ корреляторов, устройства нахождения максимального по абсолютной величине напряжения на их выходах и устройства определения знака этого напряжения.

При передаче любого сигнала $s_i(t)$ ошибка в приеме отсутствует, если выполняются неравенства

$$z_i > 0, \\ z_i > \max(|z_1|, |z_2|, \dots, |z_{i-1}|, |z_{i+1}|, \dots, |z_{m/2}|).$$

Вероятность правильного приема сигнала можно записать в виде

$$P_{\text{пр}}(s_i) = \int_0^{\infty} dz_i \int_{-z_i}^{z_i} \dots \int_{-z_i}^{z_i} w_m(z_1, \dots, z_{m/2} | s_i) dz_1 \dots dz_{i-1} dz_{i+1} \dots dz_{m/2}.$$

Случайные величины $z_1, \dots, z_{m/2}$ являются статистически независимыми и распределены так же, как и ортогональные сигналы. Это позволяет вывести следующее выражение для вероятности ошибки:

$$P_{\text{ош}} = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (2\Phi(z) - 1)^{\frac{m}{2}-1} \exp \left(-\frac{1}{2} \left(z - \sqrt{\frac{2E}{N_0}} \right)^2 \right) dz.$$

Помехоустойчивость биортогональных сигналов выше, чем ортогональных. Однако при $m \gg 1$ разница в их помехоустойчивости становится пренебрежимо мала.

Как уже указывалось выше, вычислить вероятность ошибки в m -ичной системе в общем случае нелегко. Поэтому на практике часто пользуются верхней границей для вероятности ошибки:

$$P_{\text{ош}}(s_i) \leq \sum_{j=1}^m P_{\text{ош}}(s_j | s_i), \quad (3.132)$$

где $P_{\text{ош}}(s_j | s_i)$ — вероятность ошибки при передаче сигнала $s_i(t)$ в двоичной системе, использующей сигналы $s_i(t)$ и $s_j(t)$. Оценка (3.132) справедлива для любой системы сигналов и любого канала.

Более простой, но менее точной является верхняя граница, определяемая следующим образом:

$$P_{\text{ош}} \leq (m-1) \max P_{\text{ош}}(s_j | s_i), \quad (3.133)$$

где $\max P_{\text{ош}}(s_j | s_i)$ — максимальная по всем парам i, j вероятность ошибки в двоичной системе, использующей сигналы $s_i(t)$ и $s_j(t)$.

3.4.4. Различение двух сигналов со случайной начальной фазой на фоне белого шума

Пусть сигнал на входе приемника имеет вид

$$u(t) = \theta s_1(t, \varphi_1) + (1 - \theta) s_0(t, \varphi_0) + n(t),$$

где θ — случайная величина, принимающая значения 1 и 0 с вероятностями p_1 и p_0 соответственно; φ_1 и φ_0 — начальные фазы, представляющие собой независимые случайные величины, распределенные равномерно на отрезке $[-\pi, \pi]$; $n(t)$ — помеха типа белого гауссовского шума со спектральной плотностью мощности $N_0/2$.

Отношение правдоподобия (ОП) здесь так же, как и в задаче обнаружения сигнала со случайной начальной фазой (см. п. 3.3.5), зависит от начальных фаз:

$$l(u | \varphi_1, \varphi_0) = \frac{\exp \left(-\frac{E_1}{N_0} + \frac{2}{N_0} \int_0^T u(t) s_1(t, \varphi_1) dt \right)}{\exp \left(-\frac{E_0}{N_0} + \frac{2}{N_0} \int_0^T u(t) s_0(t, \varphi_0) dt \right)}. \quad (3.134)$$

Усредняя числитель и знаменатель в выражении (3.134) по случайным параметрам φ_1 и φ_0 , получаем безусловное усредненное ОП:

$$\bar{I}(u) = \frac{\exp(-E_1/N_0)I_0(2Z_1/N_0)}{\exp(-E_0/N_0)I_0(2Z_0/N_0)},$$

где $I_0(\cdot)$ — модифицированная функция Бесселя нулевого порядка, E_0, E_1 — энергии сигналов. Величины $2Z_1/N_0$ и $2Z_0/N_0$ совпадают со значением огибающей на выходе согласованных с сигналами $s_1(t)$ и $s_0(t)$ фильтров.

В соответствии с критерием максимума апостериорной вероятности решение в пользу сигнала $s_1(t)$ принимается, когда

$$\bar{I}(u) = \frac{\exp((E_0 - E_1)/N_0)I_0(2Z_1/N_0)}{I_0(2Z_0/N_0)} \geq \frac{p_0}{p_1},$$

или

$$\ln I_0(2Z_1/N_0) - \ln I_0(2Z_0/N_0) \geq (E_1 - E_0)/N_0 + \ln(p_0/p_1) = C_1, \quad (3.135)$$

где p_0, p_1 — вероятности появления сигналов $s_0(t)$ и $s_1(t)$ соответственно.

Для симметричного канала $p_0 = p_1 = 0,5$, $E_0 = E_1 = E$, порог C_1 равен нулю, а алгоритм различения принимает вид

$$\ln I_0(2Z_1/N_0) \underset{s_0}{\overset{s_1}{\geq}} \ln I_0(2Z_0/N_0). \quad (3.136)$$

В силу монотонности функции $\ln I_0(\cdot)$ неравенство (3.136) эквивалентно неравенству

$$Z_1 \underset{s_0}{\overset{s_1}{\geq}} Z_0. \quad (3.137)$$

Оптимальный приемник, алгоритм работы которого описывается формулой (3.137) (рис. 3.27, а), состоит из двух каналов, вычисляющих по принятому колебанию $u(t)$ значения огибающих, сумматора и порогового устройства (ПУ). Каждый из каналов является оптимальным по отношению к соответствующему сигналу и реализуется по схеме, изображенной на рис. 3.10. Возможна реализация приемника на основе корреляторов (рис. 3.27, б), где $\hat{s}_i(t)$, $i=0,1$, — преобразование Гильберта от $s_i(t)$. При этом каждый канал представляет собой корреляционный приемник, схема которого показана на рис. 3.9.

Оценим помехоустойчивость различителя, предварительно отметив, что в данном случае для передачи информации нельзя использовать проти-

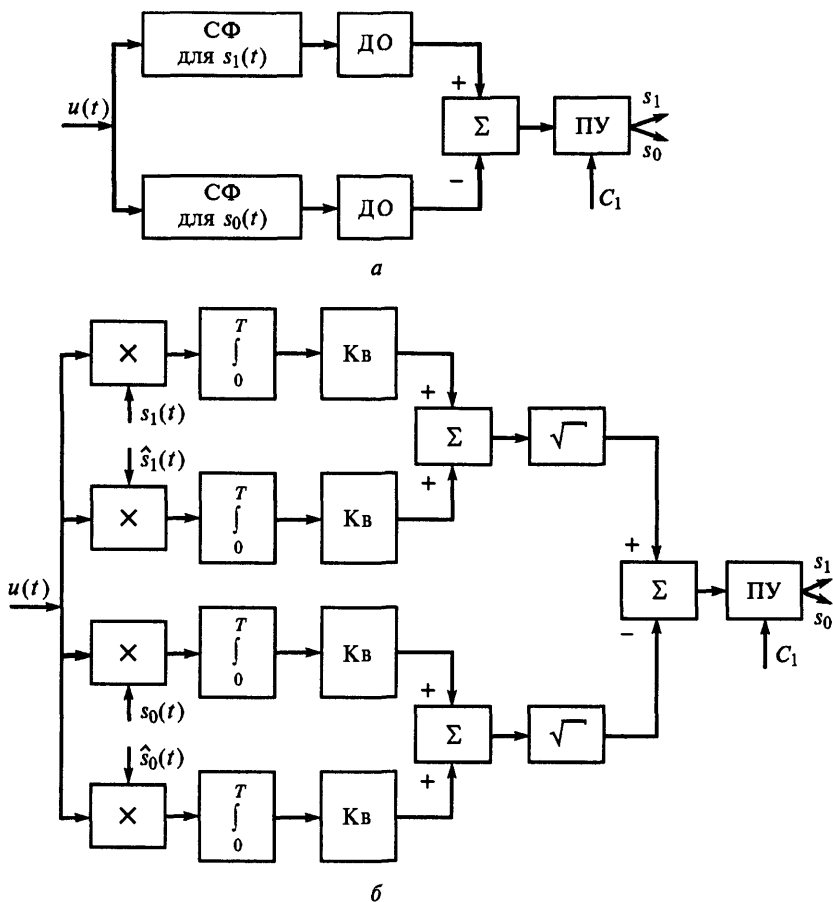


Рис. 3.27. Структурная схема:

а — фильтрового оптимального различителя двух сигналов со случайными начальными фазами; б — корреляционного оптимального различителя двух сигналов со случайными начальными фазами

воположные сигналы, отличающиеся сдвигом фаз на π , так как при случайной начальной фазе такие сигналы будут неразличимы. Обычно применяют ортогональные в усиленном смысле и амплитудно-манипулированные сигналы.

Рассмотрим сначала случай, когда используются ортогональные в усиленном смысле сигналы. Для таких сигналов справедливы соотношения

$$\int_0^T s_1(t)s_2(t)dt = \int_0^T s_1(t)\hat{s}_2(t)dt = 0, \quad (3.138)$$

где $\hat{s}_2(t)$ — преобразование Гильберта от $s_2(t)$. Примером таких сигналов являются ЧМ сигналы $s_0(t) = S_0 \cos(\omega_0 t + \varphi)$, $s_1(t) = S_0 \cos(\omega_1 t + \varphi)$, где φ — произвольная начальная фаза, а частоты ω_1 и ω_0 удовлетворяют соотношениям $\omega_1 = 2\pi k_1/T$, $\omega_0 = 2\pi k_0/T$; k_1 и k_0 — натуральные числа. Характерная особенность ортогональных в усиленном смысле сигналов состоит в следующем: если на вход согласованного фильтра, настроенного на сигнал $s_0(t)$, подать сигнал $s_1(t)$, то значение огибающей выходного напряжения в момент $t = T$ равно нулю.

Исследования показывают, что ортогональные в усиленном смысле сигналы с активной паузой обеспечивают в канале с неопределенной фазой и аддитивной гауссовской помехой минимальную вероятность ошибки, т. е. являются оптимальными для указанных условий.

Положим, что $p_1 = p_0$, $E_1 = E_0 = E$. Пусть для определенности передается сигнал $s_1(t)$. Тогда с учетом алгоритма (3.137) ошибка возникает, если выполняется неравенство $Z_0 > Z_1$ или

$$v_0 > v_1, \quad (3.139)$$

где $v_i = Z_i/\sigma$, $i = 0, 1$, — относительное значение огибающей.

Можно показать, что в рассматриваемом случае случайные величины Z_0 и Z_1 , а следовательно, v_0 и v_1 независимы [23]. Поэтому с учетом неравенства (3.139) вероятность ошибки при передаче $s_1(t)$ имеет вид

$$P_{\text{ош}}(s_1) = \int_0^\infty dv_1 \int_{v_1}^\infty w_2(v_1, v_0) dv_0 = \int_0^\infty w(v_1) \int_{v_1}^\infty w(v_0) dv_0 dv_1. \quad (3.140)$$

Учитывая, что огибающие v_0 и v_1 распределены по закону Рэлея (3.86) и Райса (3.87) соответственно, находим

$$\begin{aligned} P_{\text{ош}}(s_1) &= \int_0^\infty v_1 \exp\left(-\frac{v_1^2 + 2E/N_0}{2}\right) I_0\left(\sqrt{\frac{2E}{N_0}} v_1\right) \int_{v_1}^\infty v_0 \exp\left(-\frac{v_0^2}{2}\right) dv_0 dv_1 = \\ &= \int_0^\infty v_1 \exp\left(-\frac{v_1^2 + 2E/N_0}{2}\right) I_0\left(\sqrt{\frac{2E}{N_0}} v_1\right) \exp\left(-\frac{v_1^2}{2}\right) dv_1. \end{aligned}$$

Введем новую переменную $v = \sqrt{2} v_1$ и вынесем за знак интеграла множитель $\exp(-E/(2N_0))$. Тогда

$$P_{\text{ош}}(s_1) = \frac{1}{2} \exp\left(-\frac{E}{2N_0}\right) \int_0^{\infty} v \exp\left(-\frac{v^2 + E/N_0}{2}\right) I_0\left(\sqrt{\frac{E}{N_0}}v\right) dv. \quad (3.141)$$

Подынтегральное выражение в (3.141) представляет собой распределение Райса, а следовательно, интеграл равен 1. Таким образом,

$$P_{\text{ош}}(s_1) = 0,5 \exp(-E/(2N_0)).$$

Учитывая симметричность канала, вероятность ошибки при передаче сигнала $s_0(t)$

$$P_{\text{ош}}(s_0) = P_{\text{ош}}(s_1) = 0,5 \exp(-E/(2N_0)).$$

Соответственно, средняя вероятность ошибки

$$P_{\text{ош}} = 0,5 \exp(-E/(2N_0)). \quad (3.142)$$

На рис. 3.22 (штриховая линия — ЧМ) показана зависимость $P_{\text{ош}}$ от отношения E/N_0 , вычисленная по формуле (3.142). Анализ показывает, что некогерентный прием ортогональных сигналов дает небольшой энергетический проигрыш по сравнению с когерентным приемом. При малых вероятностях ошибки $P_{\text{ош}} = 10^{-4}$ он не превышает 1 дБ [23].

Рассмотрим случай, когда используются амплитудно-манипулированные сигналы. В данном случае

$$s_1(t) = S_0 \cos(\omega_0 t + \varphi), \quad s_0(t) = 0,$$

где начальная фаза φ является случайной величиной, распределенной равномерно на отрезке $[-\pi, \pi]$. По-прежнему полагаем, что $p_0 = p_1$.

Решение принимается на основе сравнения значения огибающей Z сигнала на выходе оптимального приемника (например, согласованного фильтра, настроенного на сигнал $s_1(t)$) с некоторым порогом $U_{\text{п}}$. При превышении порога принимается решение в пользу сигнала $s_1(t)$, в противном случае — в пользу $s_0(t)$.

Средняя вероятность ошибки имеет вид

$$P_{\text{ош}} = \frac{1}{2} (P_{\text{ош}}(s_1) + P_{\text{ош}}(s_0)) = \frac{1}{2} \left[\int_0^{z_0} w(v_1 | s_1) dv_1 + \int_{z_0}^{\infty} w(v_0 | s_0) dv_0 \right], \quad (3.143)$$

где v_1 и v_0 — относительные значения огибающих напряжений на выходе оптимального приемника в момент времени $t = T$ при передаче сигналов

$s_1(t)$ и $s_0(t)$ соответственно, $z_0 = U_{\text{п}}/\sigma$ — нормированный порог. Величина v_1 распределена по закону Райса (3.87), а v_0 — по закону Рэлея (3.86).

Подставляя распределения огибающих v_1 и v_0 в (3.143), получаем

$$P_{\text{ош}} = \frac{1}{2} \left[\int_0^{z_0} v_1 \exp\left(-\frac{v_1^2 + 2E/N_0}{2}\right) I_0\left(\sqrt{\frac{2E}{N_0}} v_1\right) dv_1 + \exp\left(-\frac{z_0^2}{2}\right) \right]. \quad (3.144)$$

Оптимальное значение порога z_0 находится из условия минимизации вероятности ошибки (3.144). Взяв производную $dP_{\text{ош}}/dz_0$ и приравняв ее нулю, имеем

$$z_0 \exp\left(-\frac{z_0^2 + 2E/N_0}{2}\right) I_0\left(\sqrt{\frac{2E}{N_0}} z_0\right) - z_0 \exp\left(-\frac{z_0^2}{2}\right) = 0,$$

или после упрощений

$$I_0\left(\sqrt{\frac{2E}{N_0}} z_0\right) = \exp\left(\frac{E}{N_0}\right). \quad (3.145)$$

Логарифмируя соотношение (3.145), получим $\ln I_0\left(\sqrt{\frac{2E}{N_0}} z_0\right) = \frac{E}{N_0}$.

Учитывая, что

$$\ln I_0(x) \approx \begin{cases} x, & x \gg 1, \\ x^2/4, & x \ll 1, \end{cases}$$

находим

$$z_{0\text{опт}} = \begin{cases} \sqrt{E/2N_0} & \text{при больших отношениях сигнал—шум,} \\ \sqrt{2} & \text{при малых отношениях сигнал—шум.} \end{cases} \quad (3.146)$$

Таким образом, с учетом (3.144) и (3.146) при больших отношениях сигнал—шум имеем

$$P_{\text{ош}} = \frac{1}{2} \left(\int_0^{\sqrt{E/2N_0}} v_1 \exp\left(-\frac{v_1^2 + 2E/N_0}{2}\right) I_0\left(\sqrt{\frac{2E}{N_0}} v_1\right) dv_1 + \exp\left(-\frac{E}{4N_0}\right) \right). \quad (3.147)$$

При $E/N_0 > 10$ первым слагаемым в (3.147) можно пренебречь. Тогда

$$P_{\text{ош}} \approx 0,5 \exp\left(-\frac{E}{4N_0}\right). \quad (3.148)$$

На рис. 3.22 (штриховая линия — АМ) показана зависимость вероятности $P_{\text{ош}}$ от отношения E/N_0 , рассчитанная по формуле (3.148). Сравнение с соответствующей кривой для когерентного случая позволяет сделать вывод, что при вероятности ошибки $10^{-3} \dots 10^{-6}$ некогерентный прием АМ сигналов проигрывает в энергетике на $1 \dots 0,5$ дБ.

При неоптимальном пороге вероятность ошибки может оказаться значительно больше $P_{\text{ош}}$, определяемой по формуле (3.148). Поэтому при изменении уровня принимаемого сигнала порог приходится подстраивать, что является существенным недостатком систем с пассивной паузой.

3.4.5. Различение m сигналов со случайной начальной фазой на фоне белого шума

Пусть на входе приемника сигнал имеет вид $u(t) = s_i(t, \varphi_i) + n(t)$, $i = 1, 2, \dots, m$, где начальная фаза φ_i представляет собой случайную величину, распределенную равномерно на отрезке $[-\pi, \pi]$; $n(t)$ — помеха типа белого гауссовского шума. Полезные сигналы $s_i(t)$, $i = 1, 2, \dots, m$, ортогональны в усиленном смысле, равновероятны и имеют одинаковую энергию.

Алгоритм принятия решения (3.111) для рассматриваемого случая можно записать в виде

$$Z_l > Z_i, \quad i = 1, 2, \dots, m, \quad i \neq l,$$

или

$$v_l > v_i, \quad i = 1, 2, \dots, m, \quad i \neq l, \quad (3.149)$$

где $v_i = Z_i / \sigma$.

Структурные схемы оптимальных различителей, выполненных на основе корреляторов и согласованных фильтров, представлены на рис. 3.28, а и 3.28, б соответственно.

Оценим помехоустойчивость оптимального различителя. Очевидно, что ошибка при приеме возникает тогда, когда неравенства (3.149) не выполняются хотя бы для одного индекса $i \neq l$. В общем виде вероятность правильного приема сигнала определяется формулой

$$P_{\text{пр}} = \int_0^\infty dv_l \int_0^{v_l} \dots \int_0^{v_l} w_m(v_1, v_2, \dots, v_m) dv_1 \dots dv_{l-1} dv_{l+1} \dots dv_m. \quad (3.150)$$

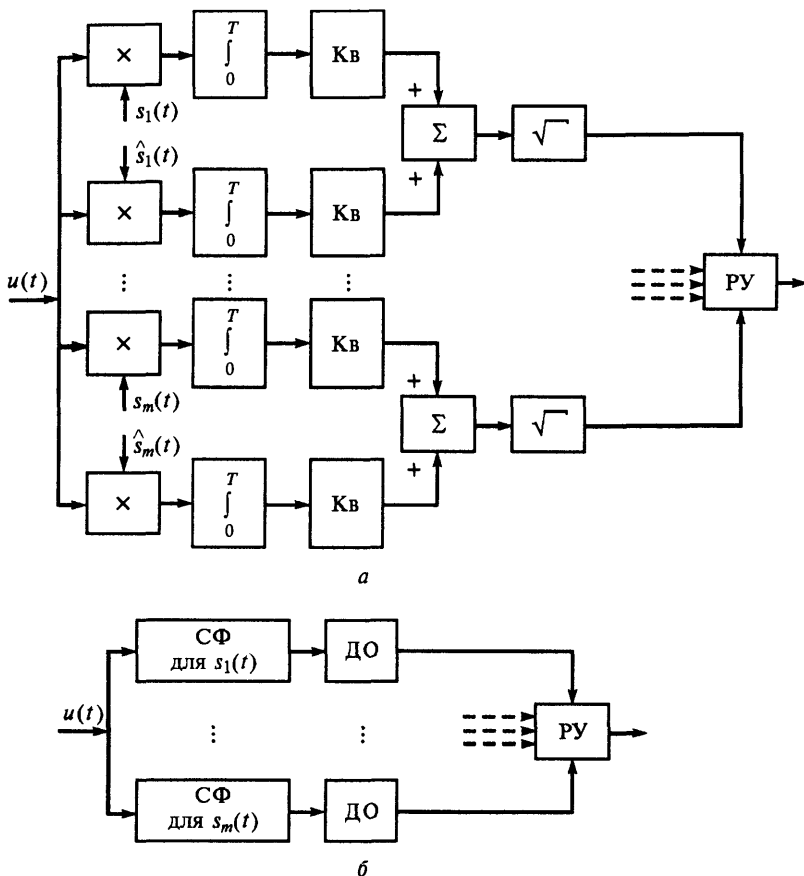


Рис. 3.28. Структурная схема:

a — корреляционного оптимального различителя m сигналов со случайными начальными фазами; b — фильтрового оптимального различителя m сигналов со случайными начальными фазами

При использовании ортогональных в усиленном смысле сигналов значения огибающих оказываются статистически независимыми и соотношение (3.150) можно переписать в виде

$$P_{\text{пр}} = \int_0^\infty dv_l \int_0^{v_l} \dots \int_0^{v_l} w(v_1) w(v_2) \dots w(v_m) dv_1 dv_2 \dots dv_{l-1} dv_{l+1} \dots dv_m.$$

Как и для двоичной системы, значение огибающей v_l распределено по закону Райса (3.87), а значения огибающих v_i , $i = 1, 2, \dots, m$, $i \neq l$, — по закону Рэлея (3.86). Используя (3.86), (3.87), находим

$$P_{\text{пр}} = \int_0^\infty v_l \exp\left(-\frac{v_l^2 + 2E/N_0}{2}\right) I_0\left(v_l \sqrt{\frac{2E}{N_0}}\right) \left[\int_0^{v_l} v_i \exp\left(-\frac{v_i^2}{2}\right) dv_i \right]^{m-1} dv_l =$$

$$= \exp\left(-\frac{E}{N_0}\right) \int_0^\infty v \exp\left(-\frac{v^2}{2}\right) I_0\left(\sqrt{\frac{2E}{N_0}} v\right) \left[1 - \exp\left(-\frac{v^2}{2}\right) \right]^{m-1} dv. \quad (3.151)$$

Учитывая, что

$$\left[1 - \exp\left(-\frac{v^2}{2}\right) \right]^{m-1} = \sum_{n=0}^{m-1} (-1)^n C_{m-1}^n \exp\left(-\frac{nv^2}{2}\right),$$

выражение (3.151) можно переписать в виде

$$P_{\text{пр}} = \exp\left(-\frac{E}{N_0}\right) \sum_{n=0}^{m-1} (-1)^n C_{m-1}^n \int_0^\infty v \exp\left(-\frac{(n+1)v^2}{2}\right) I_0\left(\sqrt{\frac{2E}{N_0}} v\right) dv.$$

После несложных вычислений получаем [23]

$$P_{\text{пр}} = \sum_{n=0}^{m-1} (-1)^n C_{m-1}^n \frac{1}{n+1} \exp\left(-\frac{nE}{(n+1)N_0}\right).$$

Вероятность ошибки имеет вид

$$P_{\text{ош}} = 1 - P_{\text{пр}} = \sum_{n=1}^{m-1} (-1)^{n+1} C_{m-1}^n \frac{1}{n+1} \exp\left(-\frac{nE}{(n+1)N_0}\right). \quad (3.152)$$

Так же, как и в п. 3.4.2, сравнение различных систем передачи информации при некогерентном приеме необходимо проводить на основе зависимостей $P_{\text{ош}} = f(E_6/N_0)$. Из рис. 3.29, на котором представлены эти зависимости, видно, что чем больше m , тем выше помехоустойчивость системы. Сравнение когерентного и некогерентного методов приема показывает, что при $m = 128$ различие в помехоустойчивости пренебрежимо мало.

На практике при $P_{\text{ош}} \ll 1$ часто пользуются верхней грани-

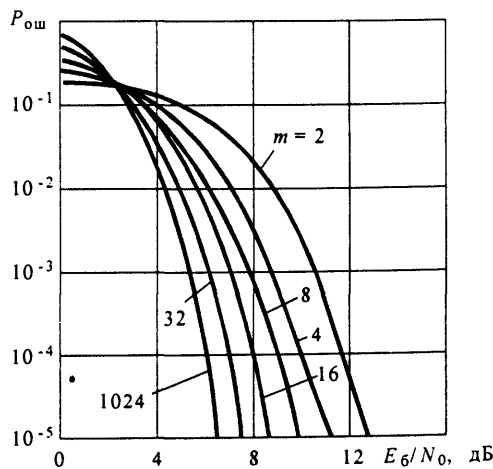


Рис. 3.29. Зависимость вероятности ошибки различения m сигналов со случайными начальными фазами

цей вероятности ошибки $P_{\text{ош}} < (m-1)\exp(-E/2N_0)/2$, совпадающей с первым членом суммы (3.152).

3.5. Оптимальный прием сигналов на фоне небелого шума

3.5.1. Обнаружение сигнала

Пусть колебание на входе приемника имеет вид

$$u(t) = \theta s(t) + n(t),$$

где θ — случайная величина, принимающая значения 1 и 0 с вероятностями p и $1-p$; $s(t)$ — полезный сигнал, параметры которого известны, $n(t)$ — гауссовский шум с корреляционной функцией $R_n(\tau)$.

Рассмотрим сначала дискретную обработку сигналов. При этом наблюдается выборка

$$\mathbf{u} = \theta \mathbf{s} + \mathbf{n},$$

где $\mathbf{u} = (u_1, u_2, \dots, u_m)$, $\mathbf{s} = (s_1, s_2, \dots, s_m)$ и $\mathbf{n} = (n_1, n_2, \dots, n_m)$ — векторы принятого, полезного и шумового сигналов соответственно, причем $u_i = u(t_i)$, $s_i = s(t_i)$ и $n_i = n(t_i)$, $i = 1, 2, \dots, m$; t_1, \dots, t_m — моменты наблюдения, выбираемые в соответствии с теоремой Котельникова.

Поскольку шум не является белым, то шумовые отсчеты будут коррелированы. Обозначим через \mathbf{R}_n корреляционную матрицу шума.

Оптимальный обнаружитель, по-прежнему, должен вычислить отношение функций правдоподобия $l(\mathbf{u}) = w(\mathbf{u}|\mathbf{H}_1)/w(\mathbf{u}|\mathbf{H}_0)$ и сравнить его с порогом C , значение которого определяется используемым критерием (см. § 3.3).

В рассматриваемом случае плотности вероятности $w(\mathbf{u}|\mathbf{H}_1)$ и $w(\mathbf{u}|\mathbf{H}_0)$ определяются n -мерными нормальными законами распределения:

$$w(\mathbf{u}|\mathbf{H}_1) = \frac{1}{(2\pi)^{n/2} |\mathbf{R}_n|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{u} - \mathbf{s})^T \mathbf{R}_n^{-1} (\mathbf{u} - \mathbf{s})\right), \quad (3.153)$$

$$w(\mathbf{u}|\mathbf{H}_0) = \frac{1}{(2\pi)^{n/2} |\mathbf{R}_n|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{u}^T \mathbf{R}_n^{-1} \mathbf{u}\right), \quad (3.154)$$

где $|\mathbf{R}_n|$ — определитель матрицы \mathbf{R}_n , \mathbf{R}_n^{-1} — матрица, обратная матрице \mathbf{R}_n , $\mathbf{u} - \mathbf{s}$ и \mathbf{u} — вектор-столбцы, элементами которых служат отсчеты $u(t_i) - s(t_i)$ и $u(t_i)$, $i = 1, 2, \dots, m$, соответственно.

С учетом соотношений (3.153) и (3.154) находим решающее правило:

$$l(\mathbf{u}) = \frac{\exp\left(-\frac{1}{2}(\mathbf{u}-\mathbf{s})^T \mathbf{R}_n^{-1}(\mathbf{u}-\mathbf{s})\right)}{\exp\left(-\frac{1}{2}\mathbf{u}^T \mathbf{R}_n^{-1}\mathbf{u}\right)} \underset{H_0}{\overset{H_1}{\geq}} C. \quad (3.155)$$

Заметим, что в случае белого шума $\mathbf{R}_n^{-1} = \sigma_n^{-2}\mathbf{I}$, где \mathbf{I} — единичная диагональная матрица, и выражение (3.155) переходит в (3.50).

Логарифмируя обе части неравенства (3.155), после преобразований получим

$$\ln l(\mathbf{u}) = -\frac{1}{2}(-\mathbf{s}^T \mathbf{R}_n^{-1}\mathbf{u} - \mathbf{u}^T \mathbf{R}_n^{-1}\mathbf{s} + \mathbf{s}^T \mathbf{R}_n^{-1}\mathbf{s}) \underset{H_0}{\overset{H_1}{\geq}} \ln C. \quad (3.156)$$

Матрица \mathbf{R}_n^{-1} является симметрической. Поэтому $\mathbf{s}^T \mathbf{R}_n^{-1}\mathbf{u} = \mathbf{u}^T \mathbf{R}_n^{-1}\mathbf{s}$ и выражение (3.156) приводится к виду

$$\mathbf{u}^T \mathbf{R}_n^{-1}\mathbf{s} \underset{H_0}{\overset{H_1}{\geq}} \ln C + \frac{\mathbf{s}^T \mathbf{R}_n^{-1}\mathbf{s}}{2} = C_1. \quad (3.157)$$

Поскольку параметры сигнала и статистические характеристики помехи известны, правая часть неравенства (3.157) может быть заранее вычислена. Поэтому существенной является лишь нахождение величины $\mathbf{q} = \mathbf{u}^T \mathbf{R}_n^{-1}\mathbf{s}$.

Техническая реализация алгоритма (3.157) многообразна. Приведем основные схемы построения оптимальных обнаружителей [29, 30].

Решающее правило (3.157) можно представить в одной из следующих форм:

$$\mathbf{u}^T \mathbf{s}' \underset{H_0}{\overset{H_1}{\geq}} C_1 \quad (3.158)$$

или

$$\mathbf{u}'^T \mathbf{s} \underset{H_0}{\overset{H_1}{\geq}} C_1, \quad (3.159)$$

где $\mathbf{s}' = \mathbf{R}_n^{-1}\mathbf{s}$, $\mathbf{u}' = \mathbf{R}_n^{-1}\mathbf{u}$.

В соответствии с неравенствами (3.158), (3.159) на рис. 3.30 представлены структурные схемы оптимального обнаружителя. Первая (рис. 3.30, а) отличается от схемы оптимального обнаружителя сигнала на фоне белого гауссовского шума тем, что вместо координат вектора \mathbf{s} используются координаты вектора \mathbf{s}' , хранящиеся в запоминающем устройстве

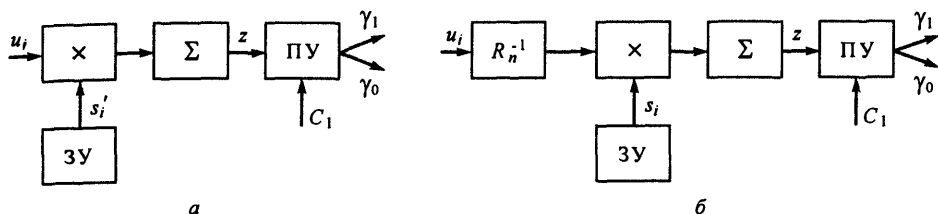


Рис. 3.30. Структурные схемы оптимального обнаружителя:

а — с преобразованием полезного сигнала; б — с матричным фильтром

(3У), а вторая (рис. 3.30, б) — наличием матричного фильтра с m входами и m выходами, осуществляющего преобразование входного вектора \mathbf{u} :

$$\mathbf{u}' = \mathbf{R}_n^{-1} \mathbf{u}, \text{ или, что то же самое, } u'_i = \sum_{j=1}^m R_{nij} u_j.$$

Возможна и другая реализация оптимального обнаружителя. Поскольку матрица \mathbf{R}_n является симметрической, то ее можно представить в виде $\mathbf{R}_n = \mathbf{L}\mathbf{L}^T$, где \mathbf{L} — нижняя треугольная матрица. При этом обратная матрица имеет вид $\mathbf{R}_n^{-1} = (\mathbf{L}^T)^{-1} \mathbf{L}^{-1} = \mathbf{V}^T \mathbf{V}$, где $\mathbf{V} = \mathbf{L}^{-1}$ — также является нижней треугольной матрицей, и решающее правило (3.157) можно представить в виде

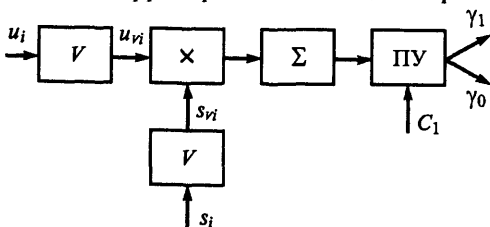
$$\mathbf{u}^T \mathbf{V}^T \mathbf{V} \mathbf{s} = \mathbf{u}_v^T \mathbf{s}_v \leq C_1, \quad (3.160)$$

где

$$\mathbf{u}_v = \mathbf{V} \mathbf{u}, \quad \mathbf{s}_v = \mathbf{V} \mathbf{s}. \quad (3.161)$$

Структурная схема оптимального обнаружителя в соответствии с (3.160) представлена на рис. 3.31. Она отличается от схемы оптимального обнаружителя сигнала на фоне белого гауссовского шума наличием двух матричных фильтров, осуществляющих преобразования (3.161).

Заметим, что на выходе матричного фильтра шумовые отсчеты оказываются некоррелированными с дисперсиями, равными единице. Действительно, корреляционная матрица шумовых отсчетов является единичной:



$$\begin{aligned} \mathbf{R}_{n_v} &= \mathbf{M}\{\mathbf{n}_v \mathbf{n}_v^T\} = \mathbf{V} \mathbf{M}\{\mathbf{n} \mathbf{n}^T\} \mathbf{V}^T = \\ &= \mathbf{V} \mathbf{R}_n \mathbf{V}^T = \mathbf{V} (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T = \mathbf{I}. \end{aligned}$$

Рис. 3.31. Структурная схема оптимального обнаружителя

Фильтры, осуществляющие декорреляцию шумовых отсчетов, получили название *обеляющих*.

Таким образом, в соответствии с алгоритмом (3.160) проводят сначала декорреляцию координат входного вектора, а затем его обрабатывают так же, как при приеме сигнала на фоне белого гауссовского шума.

Для обеспечения декорреляции координат входного вектора можно использовать ортогональные преобразования. Поскольку корреляционная матрица \mathbf{R}_n является симметрической, то всегда можно найти ортогональное преобразование \mathbf{Q} такое, что $\mathbf{Q}^T \mathbf{R}_n \mathbf{Q} = \mathbf{D}$, где \mathbf{Q} — ортогональная матрица, столбцами которой являются собственные векторы матрицы \mathbf{R}_n ; \mathbf{D} — диагональная матрица с элементами λ_i , $i = 1, 2, \dots, n$, на главной диагонали, равными собственным значениям матрицы \mathbf{R}_n .

Соответственно, обратная матрица \mathbf{R}_n^{-1} находится следующим образом:

$$\mathbf{R}_n^{-1} = (\mathbf{Q} \mathbf{D} \mathbf{Q}^T)^{-1} = \mathbf{Q} \mathbf{D}^{-1} \mathbf{Q}^T, \quad (3.162)$$

где \mathbf{D}^{-1} — диагональная матрица с элементами на главной диагонали, равными $1/\lambda_i$, $i = 1, 2, \dots, m$.

Таким образом, решающее правило с учетом соотношений (3.162) можно представить в виде

$$Z = \mathbf{u}^T \mathbf{Q} \mathbf{D}^{-1} \mathbf{Q}^T \mathbf{s} = \mathbf{u}_q^T \mathbf{D}^{-1} \mathbf{s}_q = \sum_{i=1}^m \frac{u_{qi} s_{qi}}{\lambda_i} \underset{H_0}{\overset{H_1}{\gtrless}} C_1, \quad (3.163)$$

где $\mathbf{u}_q = \mathbf{Q}^T \mathbf{u}$, $\mathbf{s}_q = \mathbf{Q}^T \mathbf{s}$.

В соответствии с алгоритмом (3.163) на рис. 3.32 представлена структурная схема оптимального обнаружителя. Преобразователь (матричный фильтр) \mathbf{Q} обеспечивает декорреляцию координат входного вектора \mathbf{u} . Все рассмотренные алгоритмы (3.158)—(3.160) и (3.163) обеспечивают одинаковую помехоустойчивость.

До сих пор предполагалось, что координаты векторов \mathbf{u} , \mathbf{s} и \mathbf{n} получаются в результате временной дискретизации непрерывных реализаций $u(t)$, $s(t)$ и $n(t)$ в соответствии с теоремой Котельникова. При этом в условиях действия белого шума шумовые отсчеты оказываются коррелированными, что усложняет обработку сигнала.

Однако разложение Котельникова не является единственным. Как известно [8, 30] (см. также § 2.4), случайный процесс, имеющий непрерывную корреляционную функцию, можно разложить по любой ортогональной системе функций в виде (2.26). Очевидно, что целесообразно для этого выбрать такую ортогональную систему, чтобы коэффициенты разложения (2.25) были некоррелированы, т. е. использовать разложение Карунена—Лозва.

При выборе системы ортонормированных функций, удовлетворяющих интегральному уравнению (2.35), коэффициенты разложения n_i , $i=1, 2, \dots, n$, гауссовского стационарного процесса $n(t)$ с нулевым математическим ожиданием будут представлять собой независимые гауссовские величины с нулевыми математическими ожиданиями и дисперсиями $\sigma_i^2 = \lambda_i$, $i=1, 2, \dots, n$, где λ_i — собственные числа ядра $R_n(t', t)$ интегрального уравнения (2.35).

Распределения $w(\mathbf{u}|H_1)$ и $w(\mathbf{u}|H_0)$ будут иметь вид

$$w(\mathbf{u}|H_1) = \frac{1}{(\sqrt{2\pi\lambda_i})^n} \exp\left(-\frac{1}{2} \sum_{i=1}^m \frac{(u_i - s_i)^2}{\lambda_i}\right), \quad (3.164)$$

$$w(\mathbf{u}|H_0) = \frac{1}{(\sqrt{2\pi\lambda_i})^n} \exp\left(-\frac{1}{2} \sum_{i=1}^m \frac{u_i^2}{\lambda_i}\right), \quad (3.165)$$

где u_i и s_i — коэффициенты разложения входного $u(t)$ и полезного $s(t)$ сигналов по используемой системе ортонормированных функций.

С учетом соотношений (3.164) и (3.165) находим логарифм отношения правдоподобия

$$\ln l(\mathbf{u}) = \sum_{i=1}^m \frac{u_i s_i}{\lambda_i} - \sum_{i=1}^m \frac{s_i^2}{2\lambda_i}.$$

Соответственно, алгоритм работы оптимального обнаружителя будет

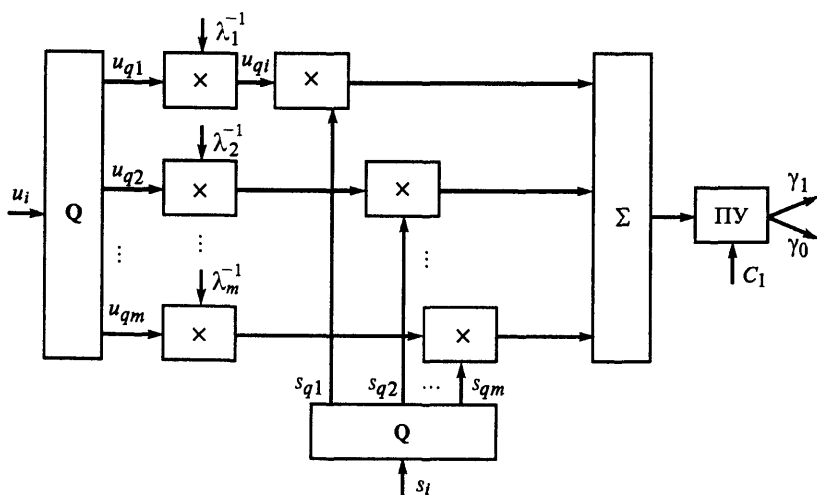


Рис. 3.32. Структурная схема оптимального обнаружителя

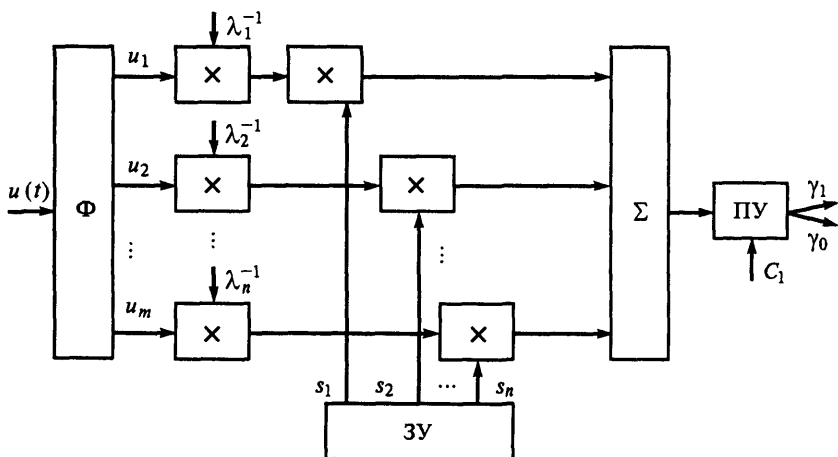


Рис. 3.33. Структурная схема оптимального обнаружителя

ИМЕТЬ ВИД

$$\sum_{i=1}^m \frac{u_i s_i}{\lambda_i} \underset{H_0}{\overset{H_1}{\gtrless}} \ln C + \sum_{i=1}^m \frac{s_i^2}{2\lambda_i} = C_1. \quad (3.166)$$

Структурная схема оптимального обнаружителя, построенного на основе (3.166), представлена на рис. 3.33, где блок Φ определяет коэффициенты разложения сигнала $u(t)$.

Заметим, что если $n(t)$ — стационарный белый шум, то уравнению (2.35) удовлетворяет любая ортонормированная система функций при $\lambda_i = N_0/2$. Для разложения сигналов удобными оказываются кусочно-постоянные функции, что обусловлено все более широким внедрением цифровых вычислительных устройств в технику обработки сигналов. Большинство работ по применению таких функций при синтезе устройств обработки сигналов связано с функциями Уолша и Хаара.

При синтезе обнаружителя на основе разложения Карунена—Лозва встречаются проблемы, о которых уже говорилось в § 2.4; процедура нахождения решения интегрального уравнения (2.35) в общем случае неизвестна, а техническая реализация обобщенной дискретизации сигналов более сложна, чем временная дискретизация. Реализация алгоритмов (3.158)—(3.160) и (3.163) достаточно трудоемка, поскольку требует большого быстродействия вычислительных устройств.

Рассмотрим непрерывную обработку сигналов. Оптимальный обнаружитель в данном случае должен вычислять отношение условных функционалов плотности вероятности $w[u(t)|H_1]$ и $w[u(t)|H_0]$ и сравнивать его

с порогом C , зависящим от используемого критерия. Таким образом, в общем виде решающее правило имеет вид

$$l[u(t)] = \frac{w[u(t)|H_1]}{w[u(t)|H_0]} \underset{H_0}{\overset{H_1}{\gtrless}} C. \quad (3.167)$$

Отношение функционалов (3.167) можно найти из отношения функций правдоподобия $l[\mathbf{u}]$, осуществляя предельный переход:

$$l[u(t)] = \lim_{\substack{n \rightarrow \infty \\ \Delta t \rightarrow 0}} l[\mathbf{u}].$$

Это позволяет использовать результаты, полученные при рассмотрении дискретной обработки сигналов, для нахождения решающего правила (3.167).

Как при дискретной обработке сигналов, так и при непрерывной решающее правило можно записать в различных формах [29, 30]. Соответственно, техническая реализация оптимального обнаружителя будет также различной. Одной из практических схем является схема обнаружителя на основе обеляющего фильтра (рис. 3.34). Алгоритм работы такого обнаружителя можно представить в виде

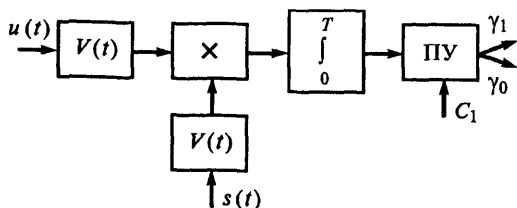
$$\int_0^T u_v(t) s_v(t) dt \underset{H_0}{\overset{H_1}{\gtrless}} \ln C + \frac{1}{2} \int_0^T s_v^2(t) dt, \quad (3.168)$$

где

$$u_v(t) = \int_0^T u(\tau) V(t-\tau) d\tau, \quad (3.169)$$

$$s_v(t) = \int_0^T s(\kappa) V(t-\kappa) d\kappa \quad (3.170)$$

и $V(t)$ — импульсная характеристика обеляющего фильтра.



Можно показать [30], что импульсная характеристика является решением интегрального уравнения

$$R_n^{-1}(t-\kappa) = \int_0^T V(\tau-t) V(\tau-\kappa) d\tau, \quad (3.171)$$

Рис. 3.34. Структурная схема обнаружителя на основе обеляющего фильтра

где функция $R_n^{-1}(t-\kappa)$, в свою очередь, находится из уравнения

$$\int_0^T R_n^{-1}(t-\kappa) R_n(\kappa-\nu) d\kappa = \delta(t-\nu).$$

Нетрудно убедиться в том, что шум на выходе фильтра с импульсной характеристикой, удовлетворяющей (3.171), является белым со спектральной плотностью мощности $N_0/2 = 1$. Очевидно, что комплексная частотная характеристика $K_{оф}(j\omega)$ обеляющего фильтра с учетом соотношения

$$F_{\text{вых}}(\omega) = |K(j\omega)|^2 F_{\text{вх}}(\omega),$$

связывающего спектральные плотности мощности случайных процессов на входе и выходе линейной системы, имеет вид

$$|K_{оф}(j\omega)| = 1/\sqrt{F_{\text{вх}}(\omega)}. \quad (3.172)$$

Учитывая, что решение интегрального уравнения (3.171) в общем случае является сложной и не всегда выполнимой задачей, для синтеза обеляющего фильтра обычно используется выражение (3.172).

Обнаружитель сигнала на фоне небелого гауссовского шума можно построить на основе оптимального фильтра, частотная характеристика которого определяется выражением

$$K(j\omega) = \frac{c S_c^*(j\omega) \exp(-j\omega t_0)}{F_{\text{вх}}(\omega)},$$

где $S_c^*(j\omega)$ — комплексно-сопряженный спектр сигнала $s(t)$, c — произвольная константа. Он представляет собой каскадное соединение обеляющего фильтра с комплексной частотной характеристикой $|K_{оф}(j\omega)| = 1/\sqrt{F_{\text{вх}}(\omega)}$ и фильтра, согласованного с сигналом $s_v(t)$. Структурная схема обнаружителя представлена на рис. 3.35.

Оценка помехоустойчивости обнаружителя сигналов на фоне небелого гауссовского шума проводится так же, как и для обнаружителя сигнала на фоне белого гауссов-

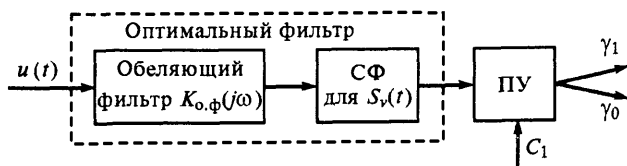


Рис. 3.35. Структурная схема обнаружителя на основе оптимального фильтра

ского шума и не вызывает трудностей. В частности, из рассмотрения структурной схемы обнаружителя (рис. 3.34) следует, что по помехоустойчивости он тождествен оптимальному обнаружителю сигнала $s_v(t)$, описываемого выражением (3.169), на фоне гауссовского белого шума $n_v(t)$ со спектральной плотностью мощности $N_0/2=1$. Качество обнаружителя полностью характеризуется отношением сигнал—шум на выходе коррелятора:

$$q^2 = \frac{2E_{s_v}}{N_0} = E_{s_v}, \quad (3.173)$$

где E_{s_v} — энергия сигнала $s_v(t)$, определяемая следующим образом:

$$\begin{aligned} E_{s_v} &= \int_0^T s_v^2(t) dt = \int_0^T dt \int_0^T s(\kappa) V(t-\kappa) d\kappa \int_0^T s(v) h(t-v) dv = \\ &= \int_0^T \int_0^T s(\kappa) R_n^{-1}(\kappa-v) s(v) d\kappa dv. \end{aligned} \quad (3.174)$$

Из (3.173) и (3.174) следует, что отношение сигнал—шум, а следовательно, помехоустойчивость обнаружителя сигнала на фоне небелого гауссовского шума, в отличие от случая действия белого гауссовского шума, зависит от формы сигнала. Фактически это обусловлено следующим. Коррелированная помеха имеет неравномерный спектр. Поэтому качество обнаружения будет зависеть от распределения энергии полезного сигнала по частоте. Желательно, чтобы основная доля энергии сигнала приходилась на те полосы частот, где составляющие помехи наименьшие.

Нетрудно оценить помехоустойчивость обнаружителя сигналов на фоне небелого гауссовского шума и при дискретной обработке. В частности, отношение сигнал—шум на выходе обнаружителя определяется формулой

$$q^2 = \mathbf{s}^T \mathbf{R}_n^{-1} \mathbf{s}. \quad (3.175)$$

3.5.2. Различение сигналов

Рассмотрим задачу различения двух детерминированных сигналов $s_1(t)$ и $s_0(t)$. Пусть сигнал на входе различителя имеет вид

$$u(t) = \theta s_1(t) + (1-\theta) s_0(t) + n(t), \quad 0 \leq t \leq T,$$

где θ — случайная величина, принимающая значения 1 и 0 с вероятностями p и $1-p$; $n(t)$ — гауссовский шум с корреляционной функцией $R_n(\tau)$.

Рассуждая точно так же, как и при рассмотрении задачи обнаружения, нетрудно получить алгоритм работы различителя при дискретной и непрерывной обработке сигналов.

В качестве примера приведем алгоритм работы оптимального различителя с обеляющим фильтром:

$$\int_0^T u_v(t)[s_{1v}(t) - s_{0v}(t)]dt \underset{s_0}{\overset{s_1}{\gtrless}} \ln C + \frac{1}{2} \int_0^T s_{1v}^2(t)dt - \frac{1}{2} \int_0^T s_{0v}^2(t)dt = C_1, \quad (3.176)$$

где

$$\begin{aligned} u_v(t) &= \int_0^T u(\tau)V(t-\tau)d\tau, \\ s_{1v}(t) &= \int_0^T s_1(\kappa)V(t-\kappa)d\kappa, \\ s_{0v}(t) &= \int_0^T s_0(\kappa)V(t-\kappa)d\kappa \end{aligned} \quad (3.177)$$

и $V(t)$ — импульсная характеристика обеляющего фильтра, удовлетворяющая уравнению (3.171).

На рис. 3.36 в соответствии с соотношениями (3.176), (3.177) представлена структурная схема различителя.

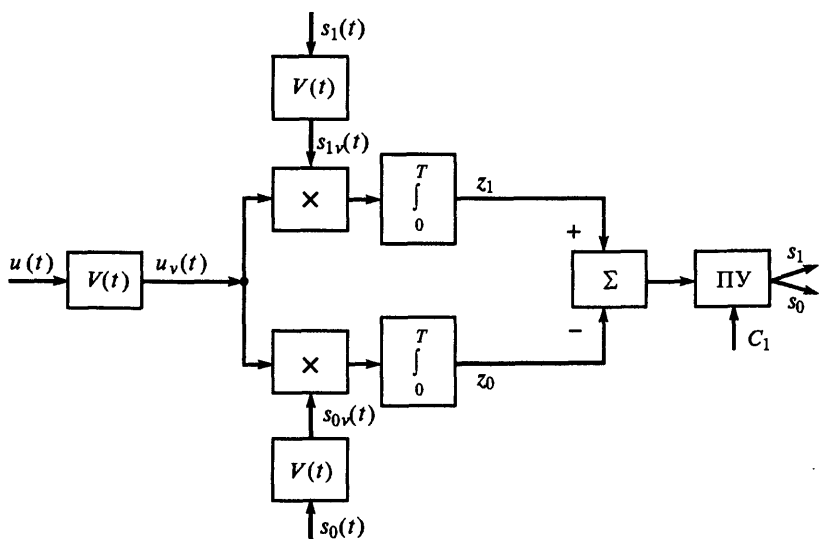


Рис. 3.36. Структурная схема различителя

Оценка помехоустойчивости оптимального различителя на фоне небольшого гауссовского шума проводится так же, как и в случае различителя сигналов на фоне белого гауссовского шума. В частности, различитель, изображенный на рис. 3.36, по помехоустойчивости тождествен оптимальному различителю сигналов $s_{1v}(t)$ и $s_{0v}(t)$, описываемых выражением (3.177), на фоне гауссовского белого шума $n_v(t)$ со спектральной плотностью мощности $N_0/2 = 1$.

Нахождение алгоритма работы оптимального различителя m сигналов не вызывает трудностей.

3.6. Обнаружение сигналов в условиях априорной неопределенности

В § 3.3 и 3.5 при решении задач обнаружения предполагалось, что распределения помех и параметров сигналов полностью известны, т. е. рассматривались задачи при полной статистической априорной информации. Однако в реальных условиях полное описание принимаемых сигналов практически невозможно. Современные радиотехнические системы различного назначения, как правило, работают в сложной помеховой обстановке, и при проектировании систем связи, радиолокации и других РТС приходится сталкиваться с задачами приема сигнала в шумах, когда их статистические характеристики неизвестны или подвержены изменениям. Поэтому в настоящее время прием сигнала общепринято трактовать как статистическую задачу с априорной неопределенностью.

В зависимости от полноты априорных сведений о помехах и параметрах сигналов различают следующие виды априорной неопределенности: параметрическую, непараметрическую и параметрико-непараметрическую [23—25, 28, 30].

Если вид распределений сигнала и помехи, а следовательно, распределений $w(u|H_1, \lambda_1)$ и $w(u|H_0, \lambda_0)$ наблюдаемого процесса при наличии и отсутствии сигнала известен, а некоторые параметры сигнала $\lambda_1 = (\lambda_{11}, \dots, \lambda_{1k})$ и помехи $\lambda_0 = (\lambda_{01}, \dots, \lambda_{0m})$, от которых зависят эти распределения, неизвестны, то априорная неопределенность называется *параметрической*. Неизвестными параметрами, общее число которых предполагается конечным, могут быть постоянная составляющая, мощность и другие параметры сигналов и помех. С таким видом априорной неопределенности встречаются, например, при решении задачи приема гауссовских случайных сигналов на фоне гауссовского шума неизвестной мощности.

Если неизвестен вид хотя бы одного из распределений $w(\mathbf{u} | H_1, \lambda_1)$, $w(\mathbf{u} | H_0, \lambda_0)$, априорную неопределенность называют *непараметрической*. При непараметрической неопределенности наблюдатель располагает небольшим объемом априорной информации. Например, может быть известно, что распределение $w(\mathbf{u} | H_1, \lambda_1)$ смещено относительно $w(\mathbf{u} | H_0, \lambda_0)$ в область больших значений координат вектора \mathbf{u} .

Как уже отмечалось в § 3.1, при решении задач приема сигналов на фоне помех необходимо использовать всю имеющуюся априорную информацию. С этой целью вводят *параметрико-непараметрические* модели априорной неопределенности, при которых для задания класса возможных распределений вероятностей наблюдаемого процесса используются известные и неизвестные распределения, параметрическое и непараметрическое описание [28]. Примером такой модели является класс ε -загрязненных распределений

$$W_1(w_0, \varepsilon) = \{w(\mathbf{u}): w(\mathbf{u}) = (1 - \varepsilon)w_0(\mathbf{u}) + \varepsilon w_1(\mathbf{u})\}, \quad (3.178)$$

где $w_0(\mathbf{u})$ — либо известная плотность распределения вероятностей, либо плотность вероятности, у которых функциональный вид известен, а параметры распределения неизвестны; $w_1(\mathbf{u})$ — неизвестная плотность распределения вероятностей; ε — известное число, причем $0 \leq \varepsilon \leq 1$. В качестве $w_0(\mathbf{u})$ обычно используют гауссовское распределение.

При $\varepsilon = 0$ и известных параметрах распределения $w_0(\mathbf{u})$ класс $W_1(w_0, \varepsilon)$, определяемый выражением (3.178), состоит из одной плотности вероятности $w_0(\mathbf{u})$. Это соответствует случаю полной априорной информации. При неизвестных параметрах распределения $w_0(\mathbf{u})$ и $\varepsilon = 0$ класс $W_1(w_0, \varepsilon)$ характеризует параметрическую априорную неопределенность, а при $\varepsilon = 1$ — непараметрическую неопределенность.

Известны и другие параметрико-непараметрические модели априорной неопределенности [28].

Для преодоления априорной неопределенности можно использовать:

- методы адаптации;
- методы непараметрической статистики;
- робастные методы [23—25, 28, 30].

В основе *методов адаптации* лежит процесс обучения, под которым понимается оценивание неизвестных функций распределений (при параметрической неопределенности) или неизвестных параметров распределений (при непараметрической неопределенности) по обучающей выборке. Полу-

ченные оценки используются вместо неизвестных характеристик наблюдаемых процессов.

Например, при решении задачи обнаружения сигнала в условиях параметрической неопределенности решающая статистика принимает вид

$$\frac{w(\mathbf{u}|H_1, \hat{\lambda}_1)}{w(\mathbf{u}|H_0, \hat{\lambda}_0)},$$

где $\hat{\lambda}_1$ и $\hat{\lambda}_0$ — оценки неизвестных параметров λ_1 и λ_0 , от которых зависят распределения наблюдаемого процесса при наличии и отсутствии сигнала.

В зависимости от того, известно или неизвестно распределение обучающей выборки, различают обучение с учителем и без учителя (самообучение), или, другими словами, обучение по классифицированной и неклассифицированной выборкам. Примером обучения с учителем является задача обнаружения сигнала на фоне шума с неизвестной мощностью, когда имеется возможность в некоторые моменты времени наблюдать только шум, а следовательно, имеется возможность получать оценку мощности шума.

Алгоритмы обработки сигналов, в которых используются полученные в результате обучения оценки функций распределения, их параметров или каких-либо других характеристик, называются *адаптивными*. Адаптивные алгоритмы сложнее неадаптивных, синтезированных при полностью известных распределениях, и уступают последним по эффективности. Однако с увеличением объема обучающей выборки, используемой при обучении, адаптивные алгоритмы сходятся к соответствующим оптимальным алгоритмам с полной априорной информацией.

Адаптивные процедуры находят широкое применение в локации, системах связи и управления и др. Примерами таких систем являются:

- автоматическая регулировка порога обнаружения, производимая по оценкам мощности шума;
- автоматическая регулировка скорости передачи информации в системах связи в зависимости от состояния канала;
- автоматическая перестройка частоты несущей в системах связи в зависимости от состояния канала и др.

Непараметрические методы применяются в условиях непараметрической априорной неопределенности. В теории обнаружения обнаружитель сигнала принято называть *непараметрическим*, если при гипотезе H_0 распределение вероятностей его решающей статистики не зависит от распределения шума [23]. Очевидно, что такой обнаружитель обеспечивает постоянную вероятность ложного обнаружения сигнала.

Непараметрический алгоритм уступает по эффективности алгоритму, полученному в условиях полной априорной определенности. Однако при изменении распределения шума непараметрические алгоритмы в общем случае более эффективны, чем классические алгоритмы, рассчитанные на определенный тип помехи.

При использовании методов непараметрической статистики выборочные данные преобразуются в статистики, вероятностные характеристики которых не чувствительны или слабо чувствительны к характеристикам сигналов и помех. В настоящее время не существует конструктивных методов построения наилучших непараметрических алгоритмов. Чаще всего их строят на основе знаковых и ранговых статистик [23, 24, 28, 30].

Пусть $\mathbf{u} = (u_1, u_2, \dots, u_n)$ — исходная последовательность наблюдаемых величин. *Знаковой статистикой* называется произвольная функция знакового вектора

$$\text{sign } \mathbf{u} = (\text{sign } u_1, \text{sign } u_2, \dots, \text{sign } u_n),$$

где

$$\text{sign } u_i = \begin{cases} 1, & u_i > 0; \\ 0, & u_i = 0; \\ -1, & u_i < 0. \end{cases}$$

Алгоритм, использующий только информацию о знаках элементов выборки, называется *знаковым*.

Знаковую статистику можно использовать, например, для обнаружения постоянного положительного (или отрицательного) сигнала на фоне помехи, имеющей симметричную плотность вероятности с нулевой медианой, при условии, что элементы выборки $\mathbf{u} = (u_1, u_2, \dots, u_n)$ независимы. Действительно, в этом случае при гипотезе H_0 количество положительных и отрицательных элементов в выборке равновероятно. Появление постоянного положительного (или отрицательного) сигнала увеличивает вероятность положительных (или отрицательных) элементов в выборке, что и используется для обнаружения сигнала.

Знаковый алгоритм обнаружения сигнала можно представить в виде

$$Z = \sum_{i=1}^n h(u_i) \underset{H_0}{\overset{H_1}{\gtrless}} C, \quad (3.179)$$

где

$$h(u_i) = \begin{cases} 1, & u_i > 0; \\ 0, & u_i \leq 0, \end{cases}$$

и C — порог, определяемый заданной вероятностью ложной тревоги $F_{\text{лт}}$.

Нетрудно заметить, что число единиц в сумме (3.179) эквивалентно числу положительных исходов в схеме испытаний Бернулли. Поэтому статистика Z будет распределена по биномиальному закону с параметром p , значение которого зависит от вида гипотезы: $p=1/2$ при гипотезе H_0 и $p>1/2$ при гипотезе H_1 и положительном сигнале.

Таким образом, распределения решающей статистики при гипотезах H_0 и H_1 имеют вид

$$P\{Z=k|H_1\} = C_n^k p^k (1-p)^{n-k}, \quad (3.180)$$

$$P\{Z=k|H_0\} = C_n^k \left(\frac{1}{2}\right)^n, \quad k=0, 1, \dots, n. \quad (3.181)$$

Соответственно, вероятности правильного обнаружения и ложной тревоги определяются выражениями

$$D = \sum_{k=C+1}^n C_n^k p^k (1-p)^{n-k}, \quad (3.182)$$

$$F_{\text{лт}} = \left(\frac{1}{2}\right)^n \sum_{k=C+1}^n C_n^k. \quad (3.183)$$

Как уже упоминалось ранее, обычно вероятность ложной тревоги $F_{\text{лт}}$ задана и по ней находится порог C . В рассматриваемом случае он находится как наибольшее целое число, удовлетворяющее неравенству

$$F_{\text{лт}} \leq \left(\frac{1}{2}\right)^n \sum_{k=C+1}^n C_n^k.$$

Из выражения (3.183) следует, что вероятность ложной тревоги не зависит от распределения помехи, а следовательно, рассмотренный знаковый обнаружитель является непараметрическим.

Знаковую статистику можно использовать и в тех случаях, когда медиана распределения шума неизвестна, а известно лишь то, что она меньше медианы распределения смеси полезного сигнала и шума. При этом используют двухвыборочную знаковую процедуру, основанную на подсчете знаков разностей пар наблюдений помеховой y_1, y_2, \dots, y_n и исследуемой u_1, u_2, \dots, u_n выборок. Решающую статистику формируют следующим образом:

$$Z = \sum_{i=1}^n h(u_i - y_i), \quad (3.184)$$

$$\text{где} \quad h(u_i - y_i) = \begin{cases} 1, & (u_i - y_i) > 0, \\ 0, & (u_i - y_i) \leq 0, \end{cases}$$

которую затем испытывают на порог C , определяемый по заданной вероятности $F_{\text{лт}}$.

Нетрудно видеть, что при гипотезе H_0 распределение решающей статистики (3.184) совпадает с распределением (3.181). Поэтому алгоритм обнаружения сигнала на основе статистики (3.184) будет непараметрическим.

Вероятность правильного обнаружения будет определяться выражением (3.182), в котором значение p вычисляется по формуле

$$p = P(u > y) = \int F_n(x) dF_{c+n}(x),$$

где $F_n(x)$ и $F_{c+n}(x)$ — интегральные функции распределения помехи и смеси сигнала и помехи соответственно.

Ранговая статистика определяется следующим образом [23]. Пусть $\mathbf{u} = (u_1, u_2, \dots, u_n)$ — исходная последовательность наблюдаемых величин. Перегруппируем элементы выборки \mathbf{u} , расположив их в порядке возрастания так, что $u^{(k)} \leq u^{(j)}$ при $k < j$. В результате получим упорядочную выборку или вариационный ряд

$$\mathbf{u}_y = (u^{(1)}, u^{(2)}, \dots, u^{(n)}).$$

Вектор \mathbf{u}_y называется *вектором порядковых статистик*, а его элементы — *порядковыми статистиками*. Порядковый номер выборки u_i в вариационном ряду \mathbf{u}_y называется *рангом* R_i этого элемента, вектор $\mathbf{R} = (R_1, R_2, \dots, R_n)$ — *ранговым вектором*. Произвольная функция от рангового вектора называется *ранговой статистикой*, а алгоритм на его основе — *ранговым*.

Если выборка \mathbf{u} однородная и независимая, то все ранговые векторы равновероятны. Поэтому при использовании рангового алгоритма обнаружения сигнала на фоне стационарной помехи вероятность ложного обнаружения не зависит от закона распределения помехи, т. е. ранговый обнаружитель оказывается непараметрическим. При появлении сигнала (при гипотезе H_1) выборка становится неоднородной, а распределение ранговых векторов перестает быть равномерным, что и позволяет обнаружить сигнал.

Ранговые статистики более информативны. Поэтому ранговые обнаружители эффективнее знаковых, т. е. обеспечивают большую вероятность правильного обнаружения сигнала при одной и той же вероятности ложной тревоги. Платой является большая сложность в реализации.

Возможны различные ранговые алгоритмы обнаружения. Наиболее простым в реализации и в то же время близким по эффективности к оптимальному ранговому, основанному на вычислении отношения правдоподобия вектора ранговой выборки $\mathbf{R} = (R_1, R_2, \dots, R_n)$

$$l(\mathbf{R}) = \frac{P(\mathbf{R}|H_1)}{P(\mathbf{R}|H_0)} = \prod_{i=1}^n \frac{P(R_i|H_1)}{P(R_i|H_0)},$$

является алгоритм обнаружения на основе суммы рангов [23, 24]:

$$Z = \sum_{i=1}^n R_i.$$

Робастные методы — это методы синтеза алгоритмов, эффективность которых близка к оптимальным для выбранных моделей и незначительно снижается при отклонении распределений наблюдений от исходных моделей. Таким образом, робастность предполагает малую чувствительность характеристик радиотехнических систем к изменению условий их функционирования.

Обычно под робастными понимают методы синтеза, занимающие промежуточное положение между параметрическими и непараметрическими [28]. Они используют меньший объем априорной информации по сравнению с параметрическими, но больший, чем непараметрические методы. Соответственно, робастные алгоритмы оказываются эффективнее непараметрических. Платой является сужение класса распределений, в котором сохраняется устойчивость алгоритмов.

Примеры построения робастных алгоритмов можно найти в [28, 30].

Контрольные вопросы

1. Сформулируете основные задачи радиоприема.
2. Поясните, чем отличаются регулярные решающие правила от рандомизированных.
3. Каким свойствам должна удовлетворять функция потерь?
4. Что такое условный риск? Средний риск? Апостериорный риск? Как они находятся?
5. В чем заключается оптимальная байесовская стратегия принятия решения?
6. Поясните суть минимаксного правила выбора решений.
7. Как находятся вероятности ошибок ложного обнаружения и пропуска сигнала?
8. Приведите алгоритмы работы оптимальных обнаружителей, построенных на основе критериев: Байеса, идеального наблюдателя, максимального правдоподобия, Неймана-Пирсона, минимаксного. В чем их сходство?
9. Поясните последовательную процедуру обнаружения сигнала.

10. Приведите алгоритм работы обнаружителя, построенного на основе критерия Вальда.
11. Как определяются верхний и нижний пороги в решающем устройстве последовательного обнаружителя?
12. Сравните по средней длительности наблюдения последовательное правило с правилом Неймана—Пирсона при равных вероятностях ложного обнаружения и пропуска сигнала.
13. Приведите алгоритм работы и структурную схему оптимального обнаружителя детерминированного сигнала.
14. Приведите методику расчета помехоустойчивости оптимального обнаружителя детерминированного сигнала.
15. Что такое рабочие характеристики? Характеристики обнаружения?
16. Приведите алгоритм работы и структурную схему оптимального обнаружителя сигнала со случайной начальной фазой.
17. Приведите методику расчета помехоустойчивости оптимального обнаружителя сигнала со случайной начальной фазой.
18. Приведите алгоритм работы оптимального обнаружителя сигнала со случайными амплитудой и начальной фазой.
19. Приведите методику расчета помехоустойчивости оптимального обнаружителя сигнала со случайными амплитудой и начальной фазой.
20. Что такое согласованный фильтр?
21. Как определяются комплексная частотная характеристика и импульсная характеристика согласованного фильтра?
22. Чему равно отношение сигнал—шум на выходе согласованного фильтра?
23. Поясните алгоритм работы оптимального обнаружителя когерентной пачки радиоимпульсов.
24. Приведите алгоритм работы оптимального обнаружителя некогерентной пачки радиоимпульсов.
25. Приведите алгоритм работы обнаружителя некогерентной пачки радиоимпульсов со случайными амплитудами.
26. Приведите алгоритм работы оптимального различителя двух детерминированных сигналов.
27. Приведите методику расчета помехоустойчивости оптимального различителя двух детерминированных сигналов.
28. Сравните по помехоустойчивости системы ФМ, ЧМ и АМ сигналов.
29. Поясните суть метода относительной фазовой модуляции.
30. Как оценивается помехоустойчивость демодулятора ОФМ сигналов?
31. Приведите алгоритм работы и методику расчета помехоустойчивости оптимального различителя m детерминированных сигналов.
32. Сравните по помехоустойчивости симплексные, ортогональные и биортогональные сигналы.
33. Приведите алгоритм работы оптимального различителя двух и более сигналов со случайной начальной фазой.
34. Приведите методику расчета помехоустойчивости оптимального различителя двух и более ортогональных в усиленном смысле сигналов со случайной начальной фазой.

4. РАЗРЕШЕНИЕ СИГНАЛОВ

Рассмотрены вопросы разрешения сигналов. Приведены сведения о пространственно-временной функции рассогласования, на основе которой рассмотрены время-частотные функции рассогласования основных видов радиолокационных сигналов: одиночного и пачки радиоимпульсов без внутримпульсной модуляции, линейно-частотно-модулированного и фазоманипулированных радиоимпульсов. Приведены примеры угловой и угло-поляризационной функций рассогласования применительно к линейной эквидистантной антенной решетке.

4.1. Понятие о разрешении сигналов

Разрешение сигналов состоит в выполнении задач обнаружения и измерения параметров произвольного сигнала в присутствии других сигналов. К разрешаемым параметрам α сигналов относят: временное положение (время запаздывания), частоту, длительность, направление прихода, поляризацию, производные частоты и угла прихода, их комбинации и др.

Разрешающей способностью принято называть способность приборов различать очень близкие в пространстве, по времени и по физическим свойствам объекты или процессы. Разрешающая способность — одна из важнейших характеристик радиолокационных и радионавигационных систем, радиотелескопов, систем радиотехнической разведки, аппаратуры разведки полезных ископаемых, медицинской диагностики и терапии, систем передачи информации и др. Повышение разрешающей способности — важное направление обеспечения помехозащищенности радиотехнических систем [31—36].

Количественную меру разрешающей способности целесообразно связать с возможностью разделения сигналов на выходе оптимального (согласованного) приемника. Очевидно, чем меньше протяженность выходного сигнала приемника по какому-либо параметру разрешения α , тем выше разрешающая способность радиотехнической системы. За количественную меру разрешающей способности обычно принимают величину $\Delta\alpha$, при кото-

рой огибающие выходных сигналов приемника пересекаются на уровне 0,5 от их максимального значения. Для сигналов, отличающихся только значением параметра α , величина $\Delta\alpha$ совпадает с шириной огибающей выходного сигнала на уровне 0,5.

В качестве примера рассмотрим разрешающую способность простейшего импульсного радиолокатора, измеряющего три координаты лоцируемых объектов: дальность r , азимут β и угол места ϵ цели. Разрешающую способность по координатам характеризуют элементарным объемом. Размеры такого элементарного объема: Δr по дальности, $\Delta\beta$ в азимутальной плоскости и $\Delta\epsilon$ в угломестной плоскости (рис. 4.1) — устанавливают так, что наличие цели в любом соседнем объеме практически не ухудшает показателей качества (эффективности) обнаружения и измерения координат цели, которая расположена в центре выделенного объема. Определенный таким образом элементарный объем называют *разрешаемым объемом* (при импульсном облучении цели — *импульсным объемом*).

Разрешающая способность по дальности характеризуется минимальным расстоянием Δr между двумя расположенными в створе с радиолокатором точечными целями, при котором одна цель не мешает обнаруживать вторую цель и измерять ее координаты. Чем меньше Δr , тем лучше разрешающая способность.

Пусть отраженные от целей прямоугольные радиоимпульсы без внутриимпульсной модуляции сдвинуты по времени на

$$\Delta t = t_{3_1} - t_{3_2} = \frac{2}{c}(r_2 - r_1) = \frac{2\Delta r}{c},$$

где $t_{3_{1,2}}$ — времена запаздывания сигналов, отраженных от первой и второй целей соответственно, $r_2 - r_1 = \Delta r$ — расстояние между этими целями, c — скорость света. На рис. 4.2 приведены огибающие выходных импульсов, соответствующие неискаженному приему прямоугольных радиоимпульсов (рис. 4.2, а) и оптимальному (согласованному) приему при отсутствии помех (рис. 4.2, б).

Величина минимального интервала Δt определяется возможностью раздельного наблюдения смежных импульсов. В рассматриваемом случае в качестве *условной меры разрешающей способности по времени* запаздыва-

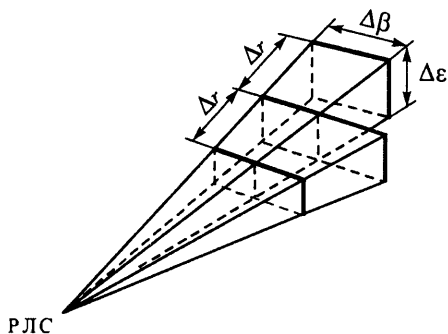


Рис. 4.1. Пояснение разрешающего объема и разрешающей способности по координатам

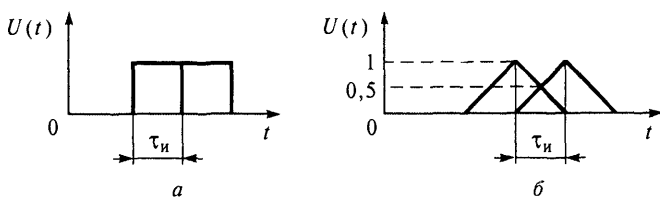


Рис. 4.2. Огибающие отраженных радиоимпульсов от двух близко расположенных целей при зондировании прямоугольными радиоимпульсами длительностью τ_n без внутриимпульсной модуляции:

а — неискаженный прием; *б* — оптимальный (согласованный) прием

ния принимают значение $\Delta t = \tau_n$, при котором максимуму огибающей сигнала (рис. 4.2, б), отраженного от одной цели, соответствует нулевое значение огибающей от другой цели. Соответственно, мерой разрешающей способности по дальности называют величину

$$\Delta r_{\min} = \frac{c\tau_n}{2}. \quad (4.1)$$

Разрешающая способность по угловым координатам обычно определяется шириной луча, которую принято отсчитывать по уровню половинной мощности диаграммы направленности антенны. Чем острее луч, тем выше разрешающая способность по угловым координатам и тем подробнее сведения о целях в секторе наблюдения.

Введенные таким образом количественные меры разрешающей способности характеризуют возможности разрешения объектов при согласованной обработке принимаемых сигналов, интенсивности которых являются величинами одного порядка, а отношение сигнал—шум достаточно велико. Реальная разрешающая способность при согласованной обработке хуже, чем приведенные выше количественные меры. Снижение разрешающей способности в общем случае зависит от искажения формы обрабатываемых сигналов в приемном тракте и индикаторе, от дискретности съема информации и т. д.

Максимально возможная (потенциальная) разрешающая способность достигается при оптимальной (в смысле статистического разрешения многих сигналов) обработке. Потенциальная разрешающая способность зависит не только от формы выходного сигнала, но и от отношения сигнал—шум. При увеличении отношения сигнал—шум потенциальная разрешающая способность возрастает.

Разрешающую способность $\Delta\alpha$ по дальности (и по другим параметрам) не следует путать с точностью измерения дальности (и, соответствен-

но, других параметров). Потенциальная точность измерения (при отсутствии других источников ошибок) определяется смещением по времени пика импульса из-за действия шумов. Поскольку это смещение меньше длительности импульса, ошибка измерения параметра сигнала меньше соответствующей меры разрешающей способности $\Delta\alpha_{\min}$. Отметим также, что рассмотренная мера разрешающей способности по дальности, определяемая формулой (4.1), справедлива только для радиотехнических систем, использующих простые радиоимпульсы (без внутриимпульсной модуляции).

4.2. Функция рассогласования (неопределенности) в теории разрешения

При согласованной обработке сигналов разрешающую способность радиотехнических (радиолокационных, радионавигационных и др.) систем характеризует так называемая *функция рассогласования*, которую часто называют также *функцией неопределенности* (неоднозначности) или *автокорреляционной функцией*.

4.2.1. Общие сведения о функции рассогласования

Радиолокационные и радионавигационные когерентные сигналы характеризуются *параметрами*, в общем случае векторными: временем запаздывания, соответствующим дальности; доплеровской частотой, соответствующей радиальной скорости цели; высшими производными по времени запаздывания; угловыми координатами и их производными; поляризационными параметрами; параметрами, связанными с особенностями распространения радиоволн в среде. Совокупный векторный параметр сигнала α_c часто *рассогласован* (отличается) по отношению к ожидаемому α , что проявляется при обнаружении и лежит в основе разрешения сигналов и измерения их параметров [31, 33—35].

Рассогласование параметров α_c и α оценивают по модулю сигнала $|Z(\alpha_c, \alpha)| = \psi(\alpha_c, \alpha)$ на выходе устройства оптимальной (согласованной) обработки при отсутствии помех. Устройство обработки полагают оптимизированным для ожидаемого сигнала $X(t, \alpha)$ на фоне некоррелированной стационарной помехи с одинаковой спектральной плотностью мощности в каналах приема. Определенную таким образом функцию $\psi(\alpha_c, \alpha)$ называют *функцией рассогласования*. Обычно используется *нормированная функция рассогласования*

$$\rho(\alpha_c, \alpha) = \psi(\alpha_c, \alpha) / \sqrt{\psi(\alpha_c, \alpha_c) \psi(\alpha, \alpha)}. \quad (4.2)$$

В знаменатель формулы (4.2) входят значения функций $\psi(\alpha_c, \alpha_c)$ и $\psi(\alpha, \alpha)$, в общем случае различающиеся между собой. Здесь

$$\psi(\alpha_c, \alpha) = \left| \int_{-\infty}^{\infty} \mathbf{X}^T(t, \alpha_c) \mathbf{X}^*(t, \alpha) dt \right|, \quad (4.3)$$

где T — знак транспонирования. Каждое значение ρ характеризует нормированное напряжение на выходе устройства оптимальной обработки при расогласовании параметров α_c и α . Для случая согласования $\alpha_c = \alpha$ имеем $\rho = 1$.

При выполнении условия $\mathbf{X}(t, \alpha) = X(t, \alpha_{вр}) \mathbf{X}(t, \alpha_{уп})$, где вектор $\alpha_{вр}$ включает временные параметры сигнала, а вектор $\alpha_{уп}$ — угловые (пространственные) и поляризационные, обработка разделяется на *временную (время-частотную)* и *угло-поляризационную (пространственно-поляризационную)*. Разделение обработки при плоском фронте волны имеет место, если запаздывание комплексной огибающей на совокупном раскрыве антенной системы много меньше $1/\Delta f_c$, где Δf_c — полоса частот сигнала. Нормированная функция расогласования сводится к произведению временной и угло-поляризационной нормированных функций расогласования:

$$\rho(\alpha_c, \alpha) = \rho_{вр}(\alpha_c, \alpha) \rho_{уп}(\alpha_c, \alpha). \quad (4.4)$$

Временная (время-частотная) функция расогласования учитывает расогласование только по временным (время-частотным) параметрам и определяется скалярной функцией времени $X(t, \alpha)$:

$$\rho_{вр}(\alpha_c, \alpha) = \frac{\left| \int_{-\infty}^{\infty} X(t, \alpha_c) X^*(t, \alpha) dt \right|}{\sqrt{\int_{-\infty}^{\infty} |X(t, \alpha_c)|^2 dt \int_{-\infty}^{\infty} |X(t, \alpha)|^2 dt}}. \quad (4.5)$$

Угло-поляризационная функция расогласования определяется не зависящим от времени вектором $\mathbf{X}(\alpha) = \|\mathbf{X}_i(\alpha)\|$, учитывающим зависимость ожидаемых колебаний от номера i канала приема, $i = 1, 2, \dots, M$. Согласно формулам (4.2) и (4.3) и правилу скалярного умножения векторов получаем

$$\rho_{\text{уп}}(\alpha_c, \alpha) = \frac{\left| \sum_{i=1}^M X_i(\alpha_c) X_i^*(\alpha) \right|}{\sqrt{\sum_{i=1}^M |X_i(\alpha_c)|^2 \sum_{i=1}^M |X_i(\alpha)|^2}}. \quad (4.6)$$

Разновидностью функций рассогласования (4.5) и (4.6) являются *комплексные функции рассогласования* $\dot{\rho}(\alpha_c, \alpha)$. Они описываются формулами (4.5) и (4.6), но без знаков модуля в числителях.

4.2.2. Время-частотные функции рассогласования

Пренебрегая деформацией комплексной огибающей $U(t)$ сигнала, обусловленной движением цели, выражение комплексной амплитуды ожидаемого сигнала можно представить в виде

$$X(t, \alpha) = U(t - t_3) e^{-j2\pi F_D t}. \quad (4.7)$$

Векторный параметр α ожидаемого сигнала заменен на два скалярных: время запаздывания t_3 и доплеровскую частоту F_D . Через два скаляра $t_{3c} = t_3 - \tau$ и $F_{Dc} = F_D - F$ выражают и векторный параметр α_c принимаемого сигнала, где τ и F — рассогласования по времени запаздывания и доплеровской частоте. Переходя в формуле (4.5) к новой переменной интегрирования $s = t - t_{3c}$, можно вынести за знак интеграла множитель $e^{j\varphi}$, $\varphi = -2\pi t_{3c}$, не зависящий от переменной интегрирования s . При замене модуля произведения произведением модулей учитываем, что $|e^{j\varphi}| = \sqrt{\cos^2 \varphi + \sin^2 \varphi} = 1$. В результате находим *нормированную время-частотную функцию рассогласования*

$$\rho(\tau, F) = \left| \int_{-\infty}^{\infty} U(s) U^*(s - \tau) e^{j2\pi F s} ds \right| / \int_{-\infty}^{\infty} |U(s)|^2 ds \quad (4.8)$$

как функцию τ, F разностей параметров ожидаемого и принимаемого сигналов. Иногда используют аналогичное (4.8) комплексное выражение $\dot{\rho}(\tau, F)$ без знака модуля в числителе.

От комплексных амплитуд в (4.8) можно перейти к их спектральным плотностям. Подставляя в (4.8) выражения

$$U(s) = \int_{-\infty}^{\infty} G(v) e^{j2\pi vs} dv,$$

$$U^*(s - \tau) = \int_{-\infty}^{\infty} G^*(v) e^{-j2\pi v(s - \tau)} dv$$

и учитывая, что $|e^{j2\pi F\tau}| = 1$, получаем

$$\rho(\tau, F) = \left| \int_{-\infty}^{\infty} G(v) G^*(v + F) e^{j2\pi v\tau} dv \right| / \left| \int_{-\infty}^{\infty} |G(v)|^2 dv \right|. \quad (4.9)$$

Время-частотная функция рассогласования $\rho(\tau, F)$ является по определению модулем нормированного напряжения на выходе устройства оптимальной (согласованной) обработки когерентного сигнала, когда на вход поступают колебания с параметрами, отличающимися на τ, F по отношению к ожидаемым. Нормирование состоит в обеспечении выполнения условия $\rho(0, 0) = 1$.

Функция рассогласования (неопределенности) $\rho(\tau, F)$, а также функция $\rho^2(\tau, F)$ описывают некоторые поверхности над плоскостью $\rho = 0$ и образуют пространственные фигуры, называемые *телами рассогласования* (неопределенности).

Сформулируем основные свойства время-частотных функций рассогласования $\rho(\tau, F)$.

1. Свойство *центральной симметрии*:

$$\rho(-\tau, -F) = \rho(\tau, F). \quad (4.10)$$

Чтобы убедиться в справедливости соотношения (4.10), достаточно заменить в числителе выражения (4.8) параметры τ, F на $-\tau, -F$ соответственно и провести замену переменной интегрирования $s = t - \tau$. Вынесем не зависящий от переменной интегрирования множитель $e^{j2\pi F\tau}$, имеющий единичный модуль, за знак интеграла, в результате, действительно, перейдем от функции $\rho(-\tau, -F)$ к функции $\rho(\tau, F)$.

2. Свойство *единичного объема тела*, ограниченного поверхностью $\rho^2(\tau, F)$ и плоскостью $\rho = 0$:

$$V_{\rho^2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho^2(\tau, F) d\tau dF = 1. \quad (4.11)$$

Это свойство называют иногда *принципом неопределенности в радиолокации*, имея в виду использование функции рассогласования в теории измерений. При заданном параметре обнаружения q степень убывания функции $\rho(\tau, F)$ в окрестности точки $\tau=0, F=0$ характеризует точность измерения параметров t_3, F_d : чем резче убывание, тем выше точность. Боковые выбросы функции $\rho(\tau, F)$ характеризуют возможную неоднозначность измерения. Наличие ошибок и неоднозначность измерения характеризуют обобщающим понятием *неопределенности*. Согласно (4.11) нельзя уменьшить объем тела неопределенности V_{ρ^2} , для любого сигнала $V_{\rho^2} = 1$.

Для обоснования справедливости равенства (4.11) подставим в него формулу (4.8) и воспользуемся соотношением

$$\left| \int_{-\infty}^{\infty} A(s) ds \right|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(s) A^*(\kappa) ds d\kappa.$$

Тогда получаем

$$\begin{aligned} V_{\rho^2} &= \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(s) U^*(s-\tau) e^{j\pi F s} ds \int_{-\infty}^{\infty} U(\kappa) U^*(\kappa-\tau) e^{-j\pi F \kappa} d\kappa d\tau dF}{\left[\int_{-\infty}^{\infty} |U(s)|^2 ds \right]^2} = \\ &= \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(s) U^*(s-\tau) U(\kappa) U^*(\kappa-\tau) \int_{-\infty}^{\infty} e^{j\pi F (s-\kappa)} dF d\kappa d\tau ds}{\left[\int_{-\infty}^{\infty} |U(s)|^2 ds \right]^2}. \end{aligned}$$

Здесь

$$\int_{-\infty}^{\infty} e^{j\pi F (s-\kappa)} dF = \delta(s-\kappa).$$

Используя фильтрующее свойство δ -функции, находим

$$V_{\rho^2} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\kappa) U^*(\kappa-\tau) U^*(\kappa) U(\kappa-\tau) d\tau d\kappa}{\left[\int_{-\infty}^{\infty} |U(s)|^2 ds \right]^2} = \frac{\int_{-\infty}^{\infty} |U(\kappa)|^2 \int_{-\infty}^{\infty} |U(\kappa-\tau)|^2 d\tau d\kappa}{\left[\int_{-\infty}^{\infty} |U(s)|^2 ds \right]^2}.$$

Внутренний интеграл по переменной τ равен энергии сигнала $U(t)$ и не зависит от сдвига κ . Тогда

$$V_{\rho^2} = \frac{\int_{-\infty}^{\infty} |U(\kappa)|^2 d\kappa \int_{-\infty}^{\infty} |U(t)|^2 dt}{\left[\int_{-\infty}^{\infty} |U(s)|^2 ds \right]^2} = 1.$$

3. Сечение тела $\rho(\tau, F)$ плоскостью $\tau = 0$ совпадает с нормированным амплитудно-частотным спектром квадрата модуля огибающей $|U(t)|^2$ сигнала.

Действительно, из формулы (4.8) при $\tau = 0$ получаем

$$\rho(0, F) = \frac{\left| \int_{-\infty}^{\infty} |U(s)|^2 e^{j2\pi F s} ds \right|}{\int_{-\infty}^{\infty} |U(s)|^2 ds}. \quad (4.12)$$

Выражение (4.12) характеризует значения нормированной огибающей напряжения на выходе согласованного фильтра при $\tau = 0$ и различных значениях F .

4. Сечение тела $\rho(\tau, F)$ плоскостью $F = 0$ совпадает с модулем нормированной корреляционной функции огибающей $U(t)$ сигнала.

Действительно, из формулы (4.8) при $F = 0$ получаем

$$\rho(\tau, 0) = \frac{\left| \int_{-\infty}^{\infty} U(s) U^*(s - \tau) ds \right|}{\int_{-\infty}^{\infty} |U(s)|^2 ds}. \quad (4.13)$$

Выражение (4.13) характеризует форму нормированной огибающей напряжения на выходе согласованного фильтра при воздействии сигнала с известной частотой ($F = 0$) при различных значениях τ . Если входной сигнал рассогласован по частоте на F_0 , то нормированная огибающая напряжения на выходе фильтра определяется сечением тела $\rho(\tau, F)$ плоскостью $F = F_0$.

Можно дать другую трактовку рассматриваемого сечения тела $\rho(\tau, F)$, используя формулу (4.9): сечение тела $\rho(\tau, F)$ плоскостью $F = 0$ совпадает с

нормированной огибающей преобразования Фурье квадрата модуля амплитудно-частотного спектра сигнала:

$$\rho(\tau, 0) \equiv \frac{\left| \int_{-\infty}^{\infty} |G(\nu)|^2 e^{j2\pi\nu\tau} d\nu \right|}{\int_{-\infty}^{\infty} |G(\nu)|^2 d\nu}. \quad (4.13')$$

4.3. Функции рассогласования когерентных сигналов

4.3.1. Сигналы без внутриимпульсной модуляции

Рассмотрим сигналы в виде одиночных радиоимпульсов $u(t) = \text{Re}[U(t)e^{j2\pi f_0 t}]$ с комплексными огибающими $U(t) = U_0(t)$ и составленные из них когерентные пачки с комплексной огибающей

$$U(t) = \sum_{i=1}^M U_0(t - t_i). \quad (4.14)$$

Типичными примерами *одиночных* радиоимпульсов являются прямоугольный (рис. 4.3, а) с огибающей

$$U_0(t) = \begin{cases} 1, & |t| \leq \tau_n/2, \\ 0, & |t| > \tau_n/2, \end{cases} \quad (4.15)$$

и колокольный (рис. 4.3, б) с огибающей

$$U_0(t) = \exp(-\pi(t/\tau_n)^2), \quad (4.16)$$

где τ_n — длительность импульса. Для колокольного радиоимпульса длительность τ_n определяется на уровне

$$\frac{U_0(\tau_n/2)}{U_0(0)} = \exp\left(-\frac{\pi}{4}\right) \approx 0,46.$$

Нормированная функция рассогласования прямоугольного радиоимпульса находится подстановкой равенства (4.15) в (4.8). Сомножители $U(s)$ и $U^*(s - \tau)$ числителя подынтегрального выражения в формуле (4.8) принимают значения, равные 1 при $-\tau_n/2 < s < \tau_n/2$ и $-\tau_n/2 < s - \tau < \tau_n/2$ соответственно и равные 0 вне этих интервалов. На рис. 4.3, в, г поясняется вза-

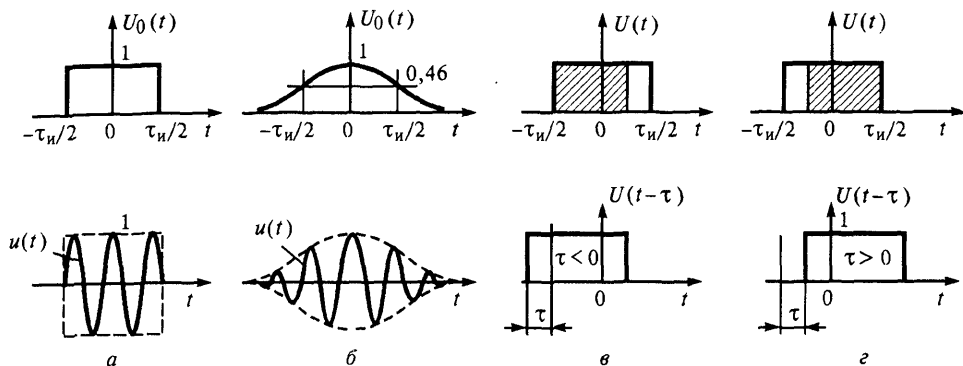


Рис. 4.3. Прямоугольный и колокольный радиоимпульсы, пояснение расчета функции рассогласования прямоугольного радиоимпульса

имный сдвиг графиков сомножителей при $\tau < 0$ и $\tau > 0$. Заштрихованы области ненулевых произведений, уточняющие пределы интегрирования $\tau + |\tau| - \tau_n/2 \leq s \leq \tau - |\tau| + \tau_n/2$ для $|\tau| \leq \tau_n$. Окончательно получим

$$\rho(\tau, F) = \begin{cases} |\sin(\pi F \tau_n (1 - |\tau|/\tau_n))| / (\pi F \tau_n), & |\tau| \leq \tau_n, \\ 0, & |\tau| > \tau_n. \end{cases} \quad (4.17)$$

Нормированная функция рассогласования колокольного радиоимпульса находится подстановкой выражения (4.16) в (4.8) и имеет вид

$$\rho(\tau, F) = \exp \left\{ -\frac{\pi}{2} \left[\left(\frac{\tau}{\tau_n} \right)^2 + (\tau_n F)^2 \right] \right\}. \quad (4.18)$$

Рассмотрим тела рассогласования и их сечения для прямоугольного и колокольного радиоимпульсов. На рис. 4.4, а, б, в приведены вертикальные $F = \text{const}$, $\tau = \text{const}$ и горизонтальные $\rho = \text{const}$ сечения, на рис. 4.4, г — аксонометрическое изображение тела рассогласования (неопределенности) $\rho(\tau, F)$ прямоугольного радиоимпульса. На рис. 4.4, д, е представлены аксонометрическое изображение и горизонтальные сечения тела рассогласования колокольного радиоимпульса. Зачерненное горизонтальное сечение на рис. 4.4, в, е соответствует уровню, близкому к 0,5, а заштрихованное — близкому к нулевому.

Рельеф тел $\rho(\tau, F)$ согласуется со сформулированными в п. 4.2.2 общими свойствами. С уменьшением длительности τ_n сечение $\rho(\tau, 0)$ сужается, а сечение $\rho(0, F)$ расширяется (рис. 4.4, б, в). Иначе, с повышением раз-

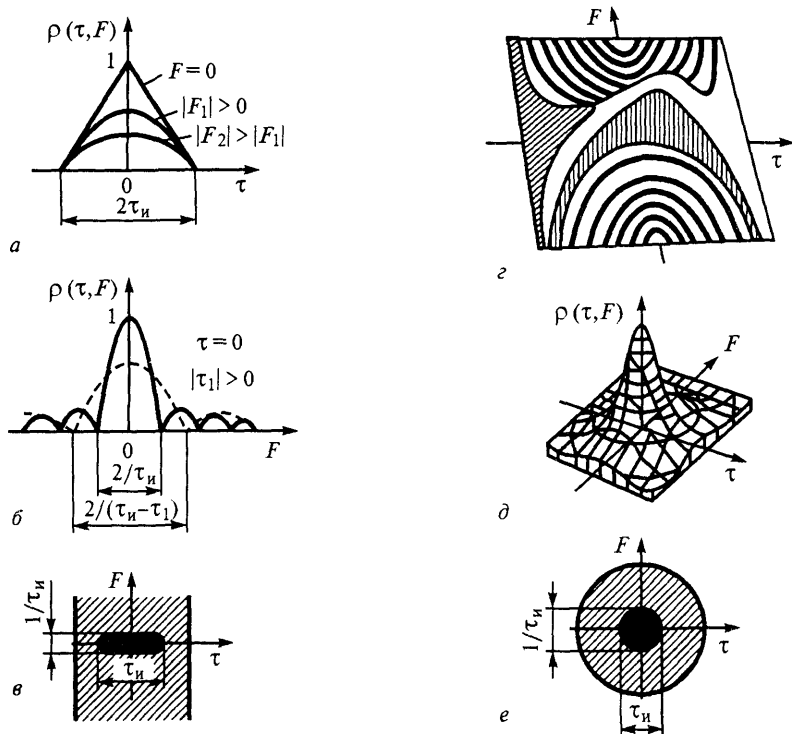


Рис. 4.4. Тела рассогласования прямоугольного и колокольного радиоимпульсов и их сечения

решающей способности по времени запаздывания (дальности) ухудшается разрешающая способность по частоте (радиальной скорости). Справедливо и обратное утверждение. Объем тела $\rho(\tau, F)$ при этом по свойству 2 (п. 4.2.2) остается неизменным.

Нормированная функция рассогласования пачки произвольно следующих радиоимпульсов находится подстановкой выражения (4.14) в (4.8) после замены переменной интегрирования $s - t_i = s'$ [31, 33—35]:

$$\rho(\tau, F) = \left| \frac{1}{M} \sum_{i=1}^M \sum_{k=1}^M \dot{\rho}_0(\tau - t_i + t_k, F) e^{j2\pi F t_i} \right|. \quad (4.19)$$

Здесь $\dot{\rho}_0(\tau, F)$ — комплексная функция рассогласования одиночного радиоимпульса произвольной формы.

Нормированная функция рассогласования пачки периодически следующих радиоимпульсов находится:

1) подстановкой в соотношение (4.19) $t_i = iT_{\pi} + \text{const}$ и $t_k = kT_{\pi} + \text{const}$, где T_{π} — период повторения импульсов;

2) введением разности $m = i - k$, где $-(M-1) \leq m \leq M-1$, $i, k = 1, 2, \dots, M$;

3) суммированием по k при фиксированных m геометрических прогрессий с членами $e^{j2\pi Ft_i}$, $i = k + m$, в интервале $1 \leq k \leq M - m$ при $m \geq 0$ и в интервале $1 - m \leq k \leq M$ при $m \leq 0$. Тогда

$$\rho(\tau, F) = \left| \sum_{m=-(M-1)}^{M-1} \frac{\sin(\pi(M-|m|)FT_{\pi})}{M \sin(\pi FT_{\pi})} \dot{\rho}_0(\tau - mT_{\pi}, F) e^{-j\pi m F T_{\pi}} \right|. \quad (4.20)$$

При любом значении τ для импульсов прямоугольной формы и $T_{\pi} \geq 2\tau_{\pi}$ не более чем одно слагаемое суммы (4.20) в силу (4.17) отлично от нуля.

Определяемое выражением (4.20) тело рассогласования прямоугольной пачки прямоугольных радиоимпульсов, его вертикальные $F = 0$, $\tau = 0$, горизонтальные $\rho = \text{const}$ сечения и аксонометрия представлены на рис. 4.5, а, б, в, г соответственно.

Сечение $F = 0$ (рис. 4.5, а) описывает огибающую реакцию (отклик) согласованного фильтра на пачечный когерентный сигнал при отсутствии расстройки по частоте. Каждый пик сечения имеет треугольную форму, со-

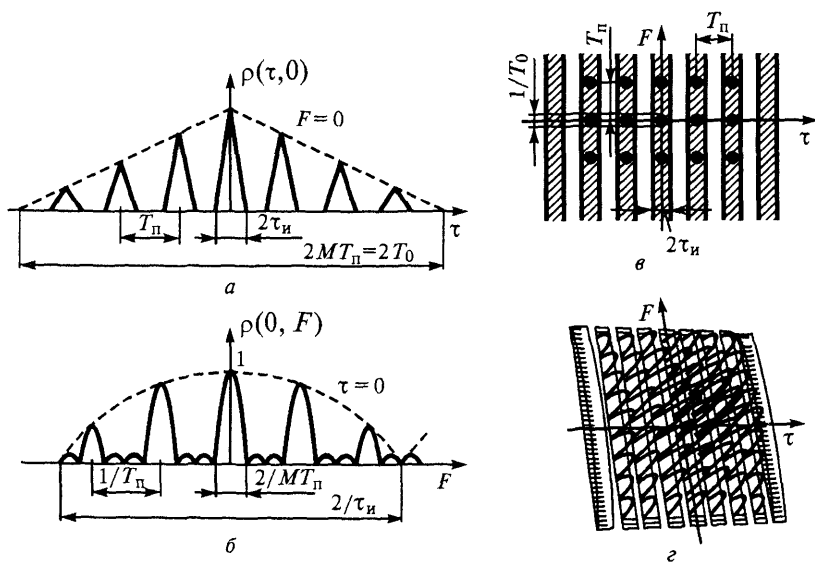


Рис. 4.5. Тела рассогласования пачечных сигналов и их сечения

ответствующую свертке огибающих прямоугольных радиоимпульсов длительности τ_n . Треугольную форму имеет и огибающая пиков. Она соответствует свертке прямоугольных огибающих импульсов пачки.

Сечение $\tau = 0$ (рис. 4.5, б) соответствует амплитудно-частотному спектру квадрата огибающей (а при постоянных значениях амплитуд — и самой огибающей) сигнала. Этот спектр состоит из пиков, взаимно сдвинутых на частоту повторения импульсов $1/T_n$. Огибающие отдельных пиков и совокупности огибающих пиков описываются при малых расстройках выражениями $|\sin(Mx)/\sin x|$ и $|(\sin x)/x|$ соответственно. Ширина центрального пика составляет $2/T_0$, где $T_0 = MT_n$ — длительность пачки. Ширина огибающей пиков равна $2/\tau_n$. На уровне 0,64 ширина пиков равна $1/T_0$ и $1/\tau_n$ соответственно. Большей протяженности T_0 пика во временной области соответствует меньшая протяженность $1/T_0$ пика в частотной; меньшей τ_n во временной области — большая $1/\tau_n$ в частотной.

Приближение формы огибающей пачки к колокольной (большая протяженность во временной области) приводит к колоколообразности отдельных пиков (малая протяженность в частотной области). Снижение при этом уровня боковых лепестков улучшает возможности разрешения по частоте.

Сравнение тел рассогласования пачечных и одиночных когерентных сигналов целесообразно проводить при одинаковой длительности радиоимпульсов. Разрешающая способность по дальности (по отношению к близко расположенным целям) для них одинакова. Пачечные когерентные сигналы обладают, однако, более высокой разрешающей способностью по частоте (скорости). В силу многолепесткового характера тел рассогласования пачечных сигналов (рис. 4.5, в, г) проявляется неоднозначность измерения времени запаздывания и частоты. Устранение неоднозначности по одному из параметров, достигаемое выбором периода T_n , ведет к появлению неоднозначности по другому. Если $T_n > 2\Delta r_{\max}/c$, где Δr_{\max} — диапазон наблюдаемых дальностей целей и других отражающих объектов, то можно говорить о сигналах с *однозначным измерением дальности*. Однако они обычно не обеспечивают однозначного измерения радиальных скоростей, в частности воздушных целей. Если же $1/T_n > |\Delta v_{r\max}/\lambda_0|$, где $-\Delta v_{r\max} \dots \Delta v_{r\max}$ — диапазон радиальных скоростей отражающих объектов, то можно говорить о системах с *однозначным измерением скорости*. Для этих систем необходима высокая частота следования импульсов (десятки-сотни килогерц), что обычно исключает однозначное измерение дальности.

Последовательности импульсов с высокой частотой следования $1/T_n$ имеют малую скважность. Их излучение называют *квазинепрерывным*. Для

определения истинной дальности при квазинепрерывном излучении используют:

- изменение периода посылок T_n ;
- улучшение селекции по угловым координатам;
- получение априорной информации о целях от других РЛС.

Возможность использования одной и той же антенны для передачи и приема — достоинство квазинепрерывного излучения по сравнению с непрерывным. Высокое качество селекции по скорости обеспечивается в сочетании с хорошей селекцией по дальности в пределах зон однозначности их измерения. Объем тела рассогласования V_p^2 распределяется по пикам неоднозначности без заметного увеличения остатков между пиками, в частности вдоль оси F . Хорошая селекция по скорости повышает возможности защиты от пассивных помех.

Недостатком квазинепрерывных сигналов является необходимость сложной и не всегда реализуемой процедуры устранения неоднозначности измерения дальности.

4.3.2. Частотно-модулированные сигналы

Частотно-модулированные (ЧМ) радиоимпульсы являются наиболее простыми разновидностями широкополосных когерентных сигналов [31, 33—35], для которых база $B = \Delta f_n \tau_n \gg 1$, что позволяет существенно повысить разрешающую способность по дальности. Для узкополосных сигналов повышение разрешающей способности по дальности ($\Delta f_n \tau_n \approx 1$) обеспечивается, как известно, уменьшением длительности τ_n зондирующих радиоимпульсов. Поскольку пиковые мощности импульсов ограничены, это часто ведет к уменьшению излучаемой энергии и к снижению дальности действия РЛС. Используя широкополосные (сложные) сигналы, можно увеличивать энергетику и дальность действия РЛС, не ухудшая, а даже улучшая разрешающую способность по дальности.

Линейно-частотно-модулированные (ЛЧМ) импульсы с прямоугольной и колокольной огибающими показаны на рис. 4.6, а, б. У этих сигналов мгновенная частота меняется по линейному закону (рис. 4.6, в):

$$f = f(t) = f_0 + (t/\tau_n) \Delta f_n, \quad (4.21)$$

а фаза — по квадратичному закону:

$$\psi(t) = 2\pi \int_{-\infty}^t f(s) ds = 2\pi f_0 t + b t^2 + \psi_0,$$

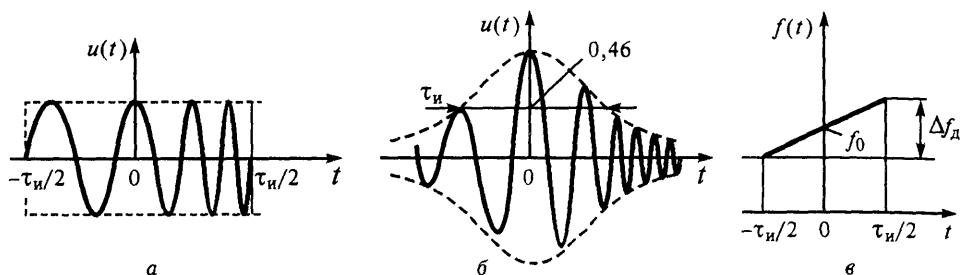


Рис. 4.6. Сигналы с линейной частотной модуляцией

где Δf_d — частотная девиация, τ_n — длительность импульса,

$$b = \pi \Delta f_d / \tau_n = \pi n / \tau_n^2 \quad (4.22)$$

— параметр фазовой модуляции сигнала, $n = \tau_n \Delta f_d$ — коэффициент широкополосности, ψ_0 — случайная начальная фаза (далее $\psi_0 = 0$). Комплексные амплитуды линейно-модулированных по частоте сигналов задаются выражениями $U(t) = U_a(t) e^{jbt^2}$, где множитель $U_a(t)$ характеризует амплитудную модуляцию. В частности, для прямоугольного ЛЧМ сигнала огибающая описывается следующим образом:

$$U(t) = \begin{cases} e^{jbt^2}, & |t| \leq \tau_n/2, \\ 0, & |t| > \tau_n/2, \end{cases} \quad (4.23)$$

для колокольного ЛЧМ сигнала

$$U(t) = e^{-\pi(t/\tau_n)^2} e^{jbt^2}. \quad (4.24)$$

Частотные спектры комплексных амплитуд ЛЧМ сигналов определяются выражением

$$G(f) = \int_{-\infty}^{\infty} U(t) e^{-2\pi f t} dt = |G(f)| e^{j\varphi(f)}, \quad (4.25)$$

где $|G(f)|$ — амплитудно-частотный, а $\varphi(f)$ — фазочастотный спектры. Амплитудно-частотный спектр колокольного ЛЧМ радиоимпульса описывается зависимостью (рис. 4.7, а)

$$|G(f)| = \frac{\tau_n}{\sqrt{1+n^2}} e^{-\pi f^2 / \Delta f_d^2} \quad (4.26)$$

и имеет на уровне 0,46 ширину спектра

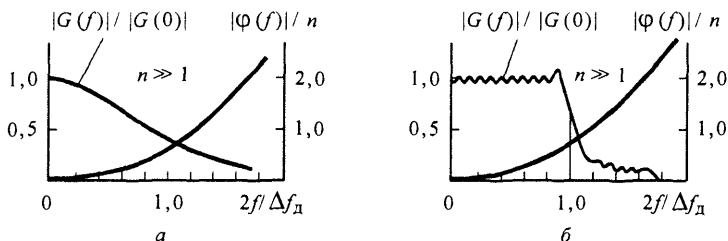


Рис. 4.7. Частотные спектры комплексных амплитуд колокольного и прямоугольного ЛЧМ сигналов

$$\Delta f_n = \frac{\sqrt{1+n^2}}{\tau_n}, \quad (4.27)$$

мало отличающуюся при больших n от частотной девиации. Фазочастотный спектр этого сигнала описывается уравнением параболы:

$$\varphi(f) = -\pi n \left(\frac{f}{\Delta f_d} \right)^2 + \frac{1}{2} \operatorname{arctg} n. \quad (4.28)$$

Амплитудно-частотный спектр прямоугольного ЛЧМ радиоимпульса (рис. 4.7, б) выражается через интегралы Френеля, а при $n \gg 1$ приближенно аппроксимируется прямоугольником, имеющим ширину Δf_d . Фазочастотный спектр в этих пределах приближенно описывается выражением (4.28). Независимо от характера огибающей чем больше частотная девиация ($n \gg 1$), тем шире спектр сигнала: $\Delta f_n \approx \Delta f_d$.

Нормированная функция рассогласования прямоугольного ЛЧМ радиоимпульса находится подстановкой выражения (4.23) в (4.8) и выбором пределов интегрирования согласно рис. 4.3, в, г. При $\tau \leq \tau_n$

$$\rho(\tau, F) = \frac{1}{\tau_n} \left| \int_{-(\tau-|\tau|+\tau_n)/2}^{(\tau-|\tau|+\tau_n)/2} e^{j2(bt+\pi F)s} ds \cdot e^{-jbt^2} \right|.$$

Интегрируя по s , учитывая единичное значение модуля сомножителя e^{-jbt^2} и используя соотношения (4.21), (4.22), можно получить

$$\rho(\tau, F) = \begin{cases} \left| \frac{\sin[\pi(n\tau/\tau_n + F\tau_n)(1-|\tau|/\tau_n)]}{\pi(n\tau/\tau_n + F\tau_n)} \right|, & |\tau| \leq \tau_n, \\ 0, & |\tau| > \tau_n. \end{cases} \quad (4.29)$$

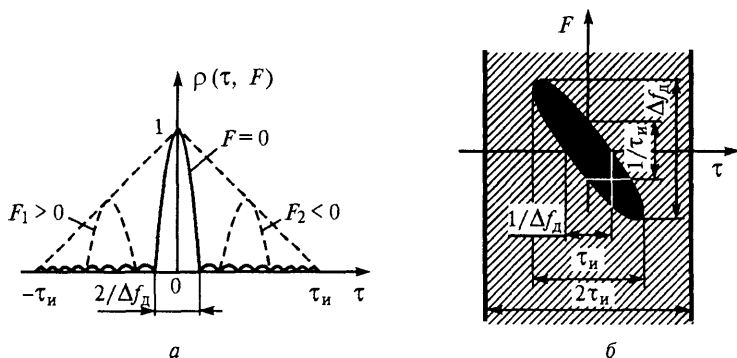


Рис. 4.8. Вертикальные (а) и горизонтальные (б) сечения функции рассогласования прямоугольного ЛЧМ радиоимпульса

При $n = 0$ выражение (4.29) преобразуется в (4.17).

Нормированная функция рассогласования колокольного ЛЧМ радиоимпульса находится подстановкой равенства (4.24) в (4.8):

$$\rho(\tau, F) = \exp \left[-\frac{\pi}{2} \left(\frac{1+n^2}{\tau_n^2} \tau^2 + 2n\tau F + \tau_n^2 F^2 \right) \right]. \quad (4.30)$$

При $n = 0$ формула (4.30) преобразуется в (4.18).

Рассмотрим тела рассогласования и их сечения для прямоугольного и колокольного ЛЧМ радиоимпульсов. На рис. 4.8, а, б приведены вертикальные $F = \text{const}$ и горизонтальные $\rho = \text{const}$ сечения для прямоугольного радиоимпульса, а на рис. 4.9, а, б — аксонометрические изображения тел рассогласования прямоугольного и колокольного радиоимпульсов соответ-

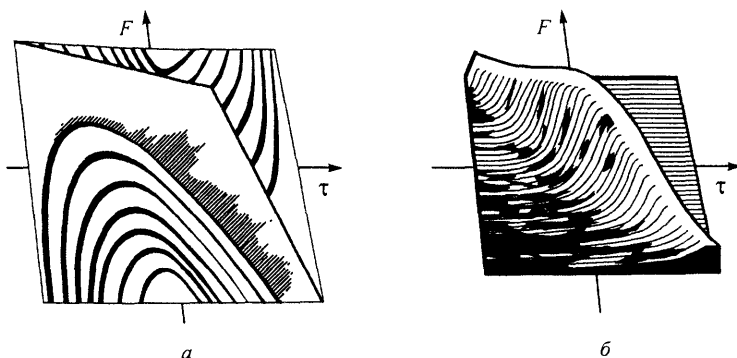


Рис. 4.9. Аксонометрические изображения тел рассогласования прямоугольного (а) и колокольного (б) радиоимпульсов

венно. В связи с частотной модуляцией сигналов тела рассогласования на рис. 4.9 повернуты и вытянуты относительно соответствующих тел рассогласования, изображенных на рис. 4.4 и 4.5, и поэтому обладают только центральной симметрией $\rho(-\tau, -F) = \rho(\tau, F)$, а симметрией относительно плоскостей $F = 0$, $\tau = 0$ не обладают. При $df/dt > 0$ имеет место поворот тела рассогласования против часовой, а при $df/dt < 0$ — по часовой стрелке. Кроме того, протяженность тела рассогласования в плоскости $F = 0$ при $n \gg 1$ оказывается существенно меньше длительности ЛЧМ сигнала. Наблюдается эффект сжатия сигнала.

Эффект сжатия радиоимпульсов непосредственно вытекает из формул (4.29) и (4.30), описывающих напряжения на выходе согласованного фильтра. Чем шире спектр, тем большее число независимых гармонических составляющих суммируется в фазе при $t = t_{\text{и}} + t_3$, тем уже пик радиоимпульса на выходе согласованного фильтра. Поскольку отклики согласованного фильтра характеризуются сечениями $F = \text{const}$ тел рассогласования, этот же вывод непосредственно следует из геометрической структуры рассматриваемых тел. Вследствие поворота тела рассогласования ЛЧМ сигнала вертикальное сечение его сужено по отношению к вертикальному сечению немодулированного сигнала, причем чем в большей степени сужено, тем больше коэффициент широкополосности $n = \tau_{\text{и}} \Delta f_{\text{д}}$. Форма сжатого радиоимпульса при $F = 0$ определяется в силу (4.13) исключительно амплитудно-частотным спектром входного сигнала. Фазочастотный спектр при отсутствии рассогласования (доплеровского, в частности) полностью компенсируется фазочастотной характеристикой фильтра и не влияет на форму выходного сигнала.

В силу принципа наложения (суперпозиции) воздействий, справедливого для линейных систем, сжатые радиоимпульсы могут не перекрываться при перекрытии ЧМ радиоимпульсов на входе фильтра. Этот эффект тем существеннее, чем больше коэффициент n . Разрешающая способность по времени запаздывания определяется, таким образом, не длительностью сигнала, а величиной, обратной ширине его амплитудно-частотного спектра. Последнее относится не только к согласованной фильтрации, но и к любой другой согласованной обработке, в частности к корреляционной обработке на основе многоканальных корреляционных устройств. Это непосредственно следует из свойства 4 (см. п. 4.2.2) тел рассогласования $\rho(\tau, F)$. Сечения $F = 0$ этих тел определяются преобразованиями Фурье квадрата амплитудно-частотного спектра сигнала.

Уровень боковых лепестков сжатого ЛЧМ радиоимпульса значительно ниже в случае скругленного амплитудно-частотного спектра (колоколь-

ный ЛЧМ сигнал), чем в случае амплитудно-частотного спектра, близкого к прямоугольному (прямоугольный ЛЧМ сигнал). С другой стороны, прямоугольная огибающая зондирующего сигнала облегчает оптимизацию работы мощных каскадов радиопередающего устройства. Поэтому если излучается сигнал с прямоугольной огибающей, то при обработке отраженного сигнала часто производят сглаживание его амплитудно-частотного спектра, видоизменяя соответствующим образом амплитудно-частотную характеристику тракта промежуточной частоты, хотя это ведет к некоторым энергетическим потерям и к растяжению сжатого импульса.

Рассмотрим неопределенность *дальность — скорость* ЛЧМ сигналов и возможные пути ее устранения. Рассогласование ЛЧМ радиоимпульсов по частоте F приводит (см. рис. 4.8, 4.9 и формулы (4.29), (4.30)) к временному смещению сжатых радиоимпульсов

$$\tau = -F/(df/dt) = \pm \tau_n F/\Delta f_d, \quad (4.31)$$

пропорциональному F . Знак минус соответствует положительному знаку производной df/dt , знак плюс — отрицательному.

При *однимпульсном* радиолокационном зондировании и неизвестном доплеровском рассогласовании $F = -2\Delta\nu_r/\lambda_0$ возможна *систематическая ошибка измерения дальности*, пропорциональная рассогласованию по радиальной скорости $\Delta\nu_r$ или самой радиальной скорости ν_r , если обработка ЛЧМ сигналов ведется в предположении, что $\nu_r = 0$. Однако при не очень больших длительностях зондирующих радиоимпульсов $\tau_n \ll 1/F_d$ величина скоростной ошибки в единицах времени значительно меньше длительности сжатого радиоимпульса и часто не имеет практического значения. Более того, даже при $\tau_n \gg 1/F_d$ любой отсчет дальности при отсутствии шума оказывается безошибочным, если его отнести к моменту наблюдения, смещенному на известную, не зависящую от $\Delta\nu_r$, величину

$$-\frac{c\tau/2}{\Delta\nu_r} = \frac{c}{\lambda_0} \frac{df}{dt} = \frac{f_0}{df/dt}. \quad (4.32)$$

Знак указанного смещения определяется знаком производной df/dt .

Если принимается последовательность ЛЧМ радиоимпульсов, то в ходе их вторичной обработки возможно достаточно точное определение текущей дальности и радиальной скорости, если время наблюдения превышает $f_0/|df/dt|$ [31, 33—35].

4.4. Сигналы, обеспечивающие высокие разрешающие способности по времени запаздывания и частоте

4.4.1. Особенности выбора время-частотных функций рассогласования

Важной задачей при разработке радиолокационных и других радиотехнических систем является выбор сигналов, функция рассогласования которых имеет низкий уровень боковых лепестков в плоскости $F = 0$ или для всей совокупности возможных значений τ , F . Обеспечение низкого уровня боковых лепестков только в плоскости $F = 0$ является достаточным в ситуациях, когда:

— максимальные значения доплеровских частот полученных сигналов и помех существенно меньше, чем величина разрешающей способности по доплеровской частоте;

— доплеровские частоты целей, находящихся в одном элементе разрешения по угловым координатам, одинаковы или достаточно близки.

Снижение уровня боковых лепестков функции рассогласования в плоскости $F = 0$ в этих случаях можно обеспечить при соответствующем увеличении уровня боковых лепестков вне рабочей области $F = 0$.

В общем случае, когда требуется уменьшение уровня боковых лепестков для совокупности значений τ , F во всей рабочей области, целесообразно выбирать сигналы, тела рассогласования которых приближаются к *игольчатому* (рис. 4.10). Сигналы с таким телом неопределенности при $\Delta f_n \tau_n \gg 1$ обеспечивают высокие разрешающие способности по времени запаздывания $1/\Delta f_n$ и частоте $1/\tau_n$. Часть объема V_{ρ^2} , относящаяся к пику, оценивается величиной $(1/\Delta f_n)(1/\tau_n)$. Остальная часть объема $1 - 1/(\Delta f_n \tau_n)$ распределяется

по большой площади $2\tau_n \cdot 2\Delta f_n$, чтобы уровень ρ_0^2 на этой площади был мал. Для идеализированного равномерного распределения ρ_0^2 имеем

$$\rho_0^2 \cdot 4\tau_n \Delta f_n \approx 1 - \frac{1}{\Delta f_n \tau_n},$$

$$\rho_0 \approx \frac{1}{2\sqrt{\Delta f_n \tau_n}}.$$

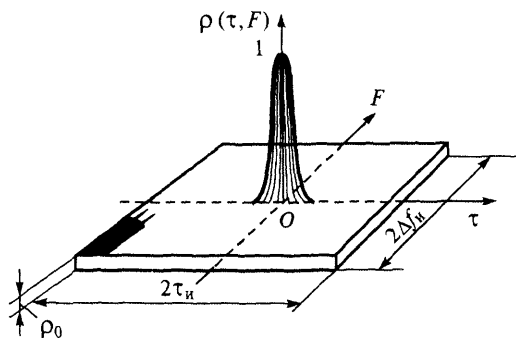


Рис. 4.10. Игольчатое тело рассогласования

Примеры реальных сигналов с телами рассогласования, близкими к идеальному игольчатому, приведены далее.

4.4.2. Фазоманипулированные сигналы

Фазовая (фазокодовая) манипуляция используется как средство расширения спектра импульсных и непрерывных сигналов для повышения разрешающей способности по дальности (времени запаздывания). Фазоманипулированный сигнал в общем случае — это совокупность сомкнутых радиоимпульсов $i = 1, 2, \dots, N$, имеющих одинаковые частоту колебаний f_0 и длительность τ_0 при ограниченном количестве возможных сдвигов фаз φ_q , $q = 0, 1, \dots, p - 1$, относительно опорного синусоидального колебания указанной частоты. Структура сигнала определяется кодом в виде цифровой последовательности q_j , $j = 1, 2, \dots, N$, элементы которой принадлежат к p -ичной системе счисления, j — номер позиции парциального радиоимпульса, q_i — номер начальной фазы на j -й позиции. Начальные фазы чаще всего равномерно распределены на интервале $(0, 2\pi)$, хотя для корректировки тел $\rho(\tau, F)$ используют иногда их неравномерное распределение. При равномерном распределении значения начальных фаз φ_q пропорциональны p -ичным цифрам q :

$$\varphi_q = 2\pi q/p, \quad q = 0, 1, \dots, p - 1. \quad (4.33)$$

Для $p = 2$ фазы принимают значения $0, \pi$; для $p = 3$ — значения $0, 2\pi/3, 4\pi/3$ и т. д.

В радиолокационных системах наиболее распространенными являются фазоманипулированные сигналы, составленные по двоичным ($p = 2$) кодам Баркера, М-кодам и т. д. (см. гл. 2). Иногда используют сигналы с изменением начальных фаз $0, \varphi$ ($\varphi \neq \pi$) и многофазовые ($p > 2$). Независимо от выбора кодовой последовательности значения q_j элементов кода используют в передатчике для формирования начальных фаз φ_q элементов зондирующего сигнала в соответствии с выбранной зависимостью φ_q и q [31, 33—35].

Импульсные сигналы с фазовой манипуляцией кодом Баркера — это сигналы, для которых уровень боковых лепестков (боковых пиков) тела $\rho(\tau, F)$ в сечении $F = 0$ при $(0, \pi)$ -манипуляции составляет $1/N$. Такие коды подобраны для ряда значений $N \leq 13$ (табл. 4.1). Для $N = 7$, например, баркеровский код описывается цифровой последовательностью 0001101, так что $\varphi_1 = \varphi_2 = \varphi_3 = \varphi_6 = 0$, а $\varphi_4 = \varphi_5 = \varphi_7 = \pi$. Уровень боковых лепестков в сечении $F = 0$ по мощности соответствует $1/N^2$, или $-20 \lg N$ в децибелах.

Таблица 4.1

Возможные законы фазовой манипуляции сигналов кодом Баркера

N	$q_j, j = 1, \dots, N$
3	0 0 1
4	0 0 1 0
5	0 0 0 1 0
7	0 0 0 1 1 0 1
11	0 0 0 1 1 1 0 1 1 0 1
13	0 0 0 0 0 1 1 0 0 1 0 1 0

При $N = 13$ этот уровень составляет -22 дБ. В сечениях $F \neq 0$ тела $\rho(\tau, F)$ содержат большие боковые пики, примерно равные $0,4 \dots 0,5$ для $N = 11$ или $N = 13$. Баркеровские коды используют поэтому лишь для радиоимпульсов не очень большой длительности, структура которых не искажается при отражении от движущихся целей.

Импульсные сигналы с фазовой манипуляцией по закону М-последовательности имеют тело рассогласования, близкое к игольчатому.

Как известно, для снижения уровня ρ вне пика сигнал должен быть протяженным и иметь ширину спектра $\Delta f_n \gg 1/\tau_n$. Рассогласования τ, F должны независимо разрушать имеющуюся корреляцию ожидаемых и принимаемых значений сигнала, чтобы равномернее распределить объем V_{ρ^2} по площади. Разрушенная из-за расстройки по τ корреляция не должна восстанавливаться где-либо при расстройках по F . Фазовая манипуляция импульсного сигнала М-последовательностью с большим числом элементов N позволяет обеспечить выполнение перечисленных требований.

Реальное тело $\rho(\tau, F)$ шумоподобного сигнала, включая область боковых лепестков, имеет в сечении $F = 0$ ограниченную протяженность $2\tau_n = 2N\tau_0$. В сечении $\tau = 0$ тело по-прежнему не ограничено, но протяженность его основной части $2\Delta f_n \approx 2/\tau_0$ определяется шириной спектра Δf_n парциального импульсного сигнала. Основной лепесток (область пика) тела рассогласования имеет в сечении $F = 0$ треугольную форму, его ширина по уровню $0,5$ равна τ_0 . В сечении $\tau = 0$ форма пика описывается выражением $|\sin x/x|$, его ширина по уровню $0,64$ составляет $1/\tau_n = 1/(N\tau_0)$. Распределение боковых пиков по плоскости τ, F оказывается, в целом, неравномерным.

Среди кодированных М-последовательностями сигналов можно выбрать сигналы с минимальным значением максимума боковых пиков (минимаксные сигналы). Максимумы пиков для этих сигналов имеют величину порядка $1/\sqrt{N}$, медленно снижаясь с увеличением N .

Использование шумоподобных фазоманипулированных сигналов с очень большим числом элементов N , как и других достаточно сложных сигналов, затрудняет разведку излучений РЛС, а значит, и наведение на них противолокационных ракет.

Непрерывные сигналы с фазовой манипуляцией по закону M -последовательности могут быть получены при фазовой манипуляции (или $(0, \pi)$, или $(0, \varphi)$) гармонического колебания двоичной рекуррентной M -последовательностью, что приводит к периодическому непрерывному сигналу с манипуляцией фазы по закону этой последовательности в каждом периоде. Обработка принимаемых колебаний сводится к *внутрипериодному* и *межпериодному* когерентному накоплению. Функцию рассогласования такого сигнала можно получить из формулы (4.20), заменив сомножитель $\dot{\rho}_0(\tau, F)$ комплексным выражением функции рассогласования для одиночного M -сигнала. Когда входящее в формулу (4.20) число периодов M стремится к бесконечности, тело $\rho(\tau, F)$ сводится в основной своей части к набору разнесенных по частоте на величину $1/T_n = 1/(N\tau_0)$ плоских элементов, протяженность которых $1/MT_n$ по оси F стремится к нулю, как это показано для $(0, \pi)$ -манипуляции на рис. 4.11. Подобная структура тела $\rho(\tau, F)$ согласуется со свойством 3 (см. п. 4.2.2) тел рассогласования, в соответствии с которым разрешающая способность по частоте (радиальной скорости) повышается с увеличением длительности когерентного сигнала. Обеспечивается однозначное разрешение по радиальной скорости в пределах частотных интервалов $1/(N\tau_0)$ при произвольном временном рассогласовании τ .

Двоичная $(0, \varphi)$ -манипуляция позволяет снизить уровень боковых лепестков тела рассогласования в сечении $F = 0$ до нуля (рис. 4.12). Это облегчает

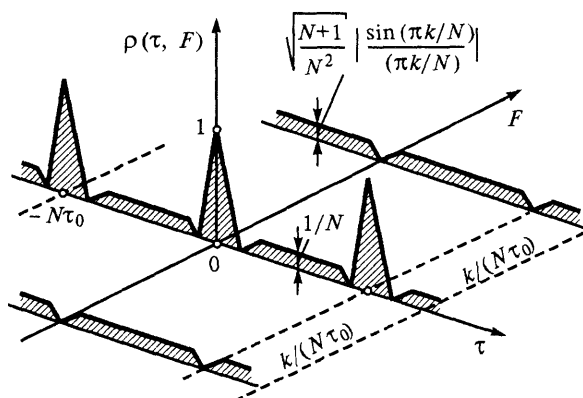
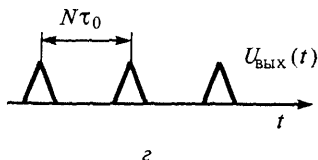
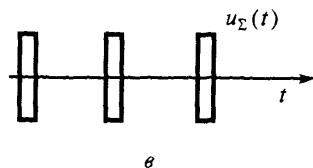


Рис. 4.11. Тело неопределенности непрерывного сигнала с фазовой манипуляцией по закону M -последовательности



Φ	0	0	Φ	Φ	Φ	0	0	Φ	0	0	Φ	Φ	Φ	0	0	Φ	0	0	Φ	$u(\tau)$
Φ	0	Φ	Φ	0	0	0	0	Φ	0	Φ	Φ	0	0	0	0	Φ	0	Φ	Φ	$u(t - \tau_0) e^{-j\varphi}$
Φ	0	Φ	0	0	Φ	Φ	Φ	Φ	0	Φ	0	Φ	Φ	Φ	0	Φ	0	Φ	Φ	$u(t - 2\tau_0)$
Φ	0	Φ	0	Φ	Φ	0	0	0	0	Φ	Φ	0	Φ	0	0	0	0	0	0	$u(t - 3\tau_0) e^{-j\varphi}$
0	0	0	Φ	0	Φ	Φ	Φ	0	0	0	Φ	0	Φ	0	Φ	0	Φ	0	Φ	$u(t - 4\tau_0) e^{-j\varphi}$
Φ	0	0	Φ	0	0	Φ	Φ	0	Φ	0	Φ	0	Φ	Φ	0	0	0	0	0	$u(t - 5\tau_0) e^{-j\varphi}$
0	0	Φ	Φ	Φ	0	Φ	0	0	Φ	0	Φ	Φ	Φ	0	0	0	Φ	Φ	Φ	$u(t - 6\tau_0)$

Рис. 4.12. К пояснению обработки сигналов с двоичной (0, ϕ)-манипуляцией

разрешение элементов групповой цели по дальности, перемещающихся с близкими радиальными скоростями, в процессе обнаружения или автоматического сопровождения по скорости. Значение φ в радианах выбирается из условия

$$\varphi = \pi - \arccos[(N-1)/(N+1)], \quad (4.34)$$

где N — период М-последовательности. Для значений N , равных 7, 15, 31, значения φ составят соответственно 139° , 151° , 160° .

Сформулированный результат справедлив как при корреляционной, так и при фильтровой обработке. Он поясняется ниже на примере обработки в согласованном фильтре (см. рис. 4.12, а) сигнала с чередованием фаз (0 ф ф 0 ф 0)..., соответствующим М-последовательности (0 1 1 1 0 1 0)... . Импульсная характеристика фильтра соответствует одному периоду М-последовательности. Внутрипериодная обработка в фильтре должна дополняться межпериодным накоплением. В момент достижения сигнала максимального значения все $N = 7$ подаваемых на сумматор задержанных и сдвинутых по фазе колебаний суммируются в фазе (рис. 4.12, б). В другие моменты времени из общего числа $N = 7$ парциальных импульсов $(N - 1)/2 = 3$ импульса имеют нулевые начальные фазы. Остальные парциальные импульсы, а именно $(N + 1)/2 = 4$, имеют ненулевые начальные фазы, в частности по $(N + 1)/4 = 2$ импульса имеют начальные фазы φ и $-\varphi$. Относительный уровень напряжения сигнала на выходе сумматора вне максимумов составляет

$$\frac{1}{N} \left| \frac{N-1}{2} e^{j0} + \frac{N+1}{4} (e^{j\varphi} + e^{-j\varphi}) \right| = \left| \frac{N-1}{2N} + \frac{N+1}{2N} \cos \varphi \right|,$$

т. е. равен $1/N$ при $\varphi = \pi$ и обращается в нуль (рис. 4.12, в) при выборе φ согласно формуле (4.34). Этот же уровень сохраняется и в результате согласованной фильтрации элементарных импульсов последовательности (рис. 4.12, з), приводящей к преобразованию прямоугольной огибающей в треугольную.

4.5. Угло-поляризационные функции рассогласования

Угло-поляризационная функция рассогласования когерентных сигналов определяется формулой (4.6) и зависит от пространственно-поляризационных характеристик антенной системы [35, 43, 47].

Рассмотрим М-элементную эквидистантную линейную антенную решетку с шагом d и равномерным амплитудным распределением коэффициентов передачи напряжений на сумматор. Подбором сдвигов фаз эта решетка согласуется с плоской гармонической волной, падающей под углом θ к нормали (рис. 4.13, а). Ожидаемое распределение комплексных амплитуд напряжений каналов приема источника, расположенного в дальней зоне, при отсутствии направленности одиночных вибраторов и движения цели относительно антенны определяется выражением

$$X_i(\theta) = e^{-j2\pi(i-i_0)d(\sin\theta)/\lambda},$$

$$i = 1, 2, \dots, i_0, \dots, M. \quad (4.35)$$

Если же волна (при согласованной поляризации) приходит под углом θ_c к нормали, то $X_i(\theta_c)$ определяются по формуле (4.35) после замены θ на θ_c .

Подставляя выражения для $X_i(\theta)$ и $X_i(\theta_c)$ в (4.6) и суммируя члены геометрической прогрессии

$$\exp(-j2\pi(i-i_0)d(\sin\theta_c - \sin\theta)/\lambda),$$

$$i = 1, 2, \dots, M,$$

находим нормированную угловую (пространственную) функцию рассогласования

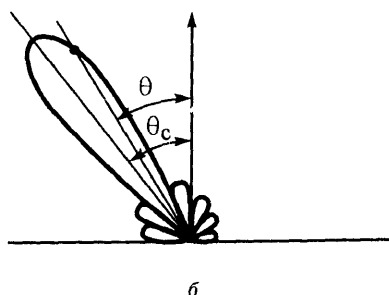
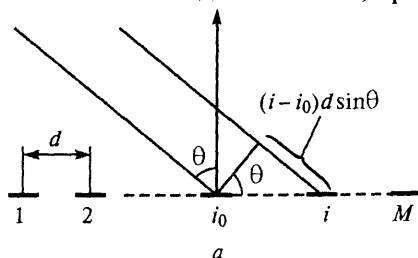


Рис. 4.13. Линейная эквидистантная решетка (а) и ее пространственная функция рассогласования (б)

$$\rho(\theta_c, \theta) = \left| \frac{\sin(\pi M d (\sin \theta_c - \sin \theta) / \lambda)}{M \sin(\pi d (\sin \theta_c - \sin \theta) / \lambda)} \right|. \quad (4.36)$$

Выражение (4.36) совпадает с выражением нормированной характеристики направленности рассматриваемой антенной решетки как функции угла прихода θ_c волны (рис. 4.13, б) при условии, что решетка согласована для угла прихода θ (т. е. значение ее характеристики направленности для этого угла максимально).

Фиксируя длину раскрыва решетки $L = Md$ и устремляя M к бесконечности, находим нормированную угловую функцию рассогласования непрерывной линейной антенны

$$\rho(\theta_c, \theta) = \left| \frac{\sin(\pi L (\sin \theta_c - \sin \theta) / \lambda)}{\pi L (\sin \theta_c - \sin \theta) / \lambda} \right|. \quad (4.37)$$

Определим *нормированную поляризационную функцию* рассогласования для когерентного поляризованного сигнала. Такой сигнал характеризуется двумя изменяющимися во времени ортогональными пространственными составляющими несущей частоты f_0 с амплитудами U_1 и U_2 и начальными фазами ψ_1 и ψ_2 [74]:

$$\mathbf{X}(t, \alpha) = \begin{bmatrix} U_1 \cos(2\pi f_0 t - \psi_1) \\ U_2 \cos(2\pi f_0 t - \psi_2) \end{bmatrix}. \quad (4.38)$$

Комплексную амплитуду такого сигнала можно записать в следующем виде:

$$\mathbf{X}(\alpha) = \begin{bmatrix} U_1 e^{-j\psi_1} \\ U_2 e^{-j\psi_2} \end{bmatrix} = U e^{-j\psi_1} \begin{bmatrix} \cos \gamma \\ e^{j\delta} \sin \gamma \end{bmatrix} = U e^{-j\psi_1} \cos \gamma \begin{bmatrix} 1 \\ e^{j\delta} \operatorname{tg} \gamma \end{bmatrix}. \quad (4.39)$$

В этом выражении U — результирующая комплексная амплитуда, $|U| = \sqrt{U_1^2 + U_2^2}$ — ее модуль, $\cos \gamma = U_1 / |U|$, $\sin \gamma = U_2 / |U|$, $\delta = \psi_1 - \psi_2$.

С точностью до множителя можно считать $\mathbf{X}(\alpha) = [X_1 \ X_2]^T$, где $X_1 = 1$, $X_2 = \alpha$, $\alpha = e^{j\delta} \operatorname{tg} \gamma$. Подставляя полученное выражение для комплексной амплитуды $\mathbf{X}(\alpha) = [1 \ \alpha]^T$ в формулу (4.6), получим нормированную функцию рассогласования по поляризации

$$\rho(\alpha_c, \alpha) = \frac{1 + \alpha_c \alpha^*}{\sqrt{(1 + |\alpha_c|^2)(1 + |\alpha|^2)}}. \quad (4.40)$$

Эта формула справедлива для эллиптически поляризованного сигнала.

При линейно-поляризованном колебании $\delta = 0$ или $\delta = \pi$ угол γ определяет плоскость поляризации. Тогда $\alpha_c = \text{tg } \gamma_c$, $\alpha = \text{tg } \gamma$, а функция рассогласования

$$\rho(\alpha_c, \alpha) = \rho(\gamma_c, \gamma) = \cos(\gamma_c - \gamma).$$

Найдем функцию рассогласования круго-поляризованного колебания с произвольным линейно-поляризованным колебанием. Для колебаний с круговой поляризацией имеем $\delta = \pm\pi/2$, $\gamma = \pm\pi/4$. Тогда, полагая $\alpha_c = \pm\gamma$, а $\alpha = \text{tg } \gamma$, из формулы (4.40) находим

$$\rho(\alpha_c, \alpha) = \rho(\gamma_c, \gamma) = 1/\sqrt{2}$$

независимо от γ . Из полученного выражения следует важный практический вывод: при приеме сигналов с неизвестной линейной поляризацией целесообразно использовать антенны с круговой поляризацией.

Приведем пример *угло-поляризационной функции рассогласования*. Пусть элементами эквидистантной линейной антенной решетки являются вертикальные и горизонтальные вибраторы, рассчитанные на отдельный прием соответствующих линейно-поляризованных колебаний. Функция рассогласования 2M-элементной решетки зависит в данном случае как от угловых, так и от поляризационных параметров. Если, как при выводе соотношения (4.35), не учитывать диаграммы направленности одиночного вибратора, то

$$\rho(\alpha_c, \alpha) = \rho(\theta_c, \theta) \rho(\alpha_c, \alpha),$$

где $\rho(\theta_c, \theta)$ определяется формулой (4.36), а $\rho(\alpha_c, \alpha)$ — формулой (4.40).

4.6. *Согласованное и оптимальное разрешения сигналов*

Различают *согласованное и оптимальное (рассогласованное)* разрешения сигнала [31—35, 74]. Согласованное разрешение, как уже отмечалось в § 4.1, реализуется при обработке, оптимизируемой для фона некоррелированной стационарной помехи с известными статистическими параметрами, а значит, без учета каких-либо других сигналов, кроме ожидаемого. Повышение качества разрешения при такой обработке обеспечивается лишь за счет выбора структуры сигнала (время-частотной, пространственно-поляризационной).

При воздействии интенсивных сигналов, являющихся по отношению к более слабым мешающими, подобная оптимизация оказывается в ряде случаев недостаточной. Для лучшего подавления мешающих сигналов приходится учитывать их особенности, а значит, переходить от согласованной обработки к оп-

тимальной. Оптимальная (рассогласованная) обработка может быть неадаптивной и адаптивной. *Неадаптивная обработка* ориентирована на заранее известную, типичную ситуацию. *Адаптивная обработка* заключается в приспособлении (адаптации) к не известной заранее, конкретной, ситуации.

Рассмотрим некоторые вопросы *согласованного разрешения*. Закономерности *оптимального (рассогласованного)* разрешения изложены в гл. 7, а также, например, в [32—36, 40, 74].

При синтезе устройств согласованного разрешения мешающие сигналы не учитываются и не влияют на прохождение полезного сигнала через линейный тракт обработки. Без дополнительного ослабления проходят через него и мешающие сигналы со случайными комплексными амплитудами

$$b_i X(t, \alpha_i) e^{j\beta_i}, \quad i = 1, 2, \dots, m,$$

где b_i — случайные амплитудные множители, β_i — случайные начальные фазы. Мешающие сигналы воздействуют после прохождения через линейный тракт как нестационарная помеха. Отношение их среднеквадратичных напряжений при настройке на полезный сигнал с учетом рассогласования (в том числе по времени наблюдения в случае согласованной фильтрации) к дисперсии стационарного шума при $\overline{b_i^2} = 1$ определяется выражениями $\sqrt{h_i} \rho(\alpha_c, \alpha_i)$, где $h_i^2 = q_i^2/2$, q_i^2 — параметры обнаружения мешающих сигналов. Для гауссовских стационарного и нестационарного шумов снижение значения параметра обнаружения q при воздействии мешающих сигналов можно оценить выражением

$$q^2/q_0^2 = 1 / \left[1 + \sum_{i=1}^m h_i^2 \rho^2(\alpha_c, \alpha_i) \right], \quad (4.41)$$

где m — число мешающих сигналов. Сигнал с параметром обнаружения q разрешается на фоне m мешающих сигналов, если этот параметр превосходит пороговое значение, выбранное для принятых показателей качества обнаружения D, F .

При разрешении двух сигналов ($m = 1$) с достаточно большими и равными энергиями ($q_0^2 = q_1^2 = 2h^2 \gg 1$) имеем $q^2 \approx 2/\rho^2(\alpha_1, \alpha_0)$. Принимая (с учетом некогерентного накопления) пороговое значение q^2 равным 8, приходим к условию взаимного согласованного разрешения сигналов с одинаковой средней энергией:

$$\rho = \rho(\alpha_1, \alpha_0) \leq 0,5. \quad (4.42)$$

Это приводит к упоминавшимся ранее мерам разрешающей способности, определяемым протяженностями основных лепестков функций и тел рассогласования когерентных сигналов на уровнях, близких к 0,5.

Таким образом, *меры разрешающей способности* характеризуют возможности согласованного разрешения по основным параметрам когерентных сигналов: времени запаздывания, частоте, угловым координатам.

Мера разрешающей способности *по времени запаздывания* (дальности) определяется величиной, обратной ширине спектра частот сигнала (импульса):

$$\Delta\tau = 1/\Delta f_n, \quad \Delta r = c/(2\Delta f_n).$$

Мера разрешающей способности *по частоте* (радиальной скорости) определяется величиной, обратной общей длительности сигнала (или интервалом когерентности принимаемого сигнала, если его протяженность меньше общей длительности сигнала):

$$\Delta F = 1/\tau_c, \quad \Delta v_r = \lambda/(2\tau_c).$$

Длительность τ_c сигналов выбираем следующим образом:

- для одиночного сигнала $\tau_c = \tau_n$;
- для когерентной пачки радиоимпульсов $\tau_c = MT_n$;
- для непрерывного сигнала $\tau_c = \tau_{\text{ког}}$.

Мера *угловой разрешающей способности* в радианах определяется обратной величиной от числа длин волн, содержащихся в проекции размера антенны на плоскость, перпендикулярную направлению прихода сигнала:

$$\Delta\theta = \lambda/(d \cos\theta).$$

В случае единой приемопередающей антенны возможно улучшение разрешающей способности примерно в $\sqrt{2}$ раз. Наряду с истинным раскрытием антенны при когерентной обработке может вводиться синтезированный раскрыт [28, 35].

Заметим, что рассматриваемые количественные меры разрешающей способности, определяемые шириной главного пика тела рассогласования (неопределенности), характеризуют разрешающую способность только при согласованной фильтрации сигналов, имеющих приблизительно равную интенсивность. Если же принимаемые сигналы существенно различаются по интенсивности, то слабый сигнал может быть замаскирован боковыми лепестками тела неопределенности сильного сигнала. Чтобы повысить разрешающую способность в данном случае, необходимо снижать уровень боковых лепестков. Для снижения уровня боковых лепестков по времени запаздывания и доплеровской частоте применяют корректирующие (не согласованные) фильтры, характеристики которых подбирают таким образом, чтобы выходной сигнал имел требуемые лепестки (см. гл. 7, а также [8, 28, 33, 34]). Для снижения уровня боковых лепестков по угловым координатам применяют специальные меры при конструировании антенных систем.

Контрольные вопросы

1. Что такое разрешающая способность?
2. Каким образом определяется разрешающая способность по дальности, по радиальной скорости, по угловым координатам?
3. Как определяется функция рассогласования?
4. При каких условиях возможно разделение пространственно-временной функции рассогласования на временную и угло-поляризационную (в частности, пространственную)?
5. Поясните формулу для время-частотной функции рассогласования.
6. Поясните принцип неопределенности в радиолокации.
7. Что характеризует вертикальное сечение времячастотной функции рассогласования плоскостью $\tau = 0$, плоскостью $F = 0$?
8. Каким параметром радиолокационного сигнала определяется разрешающая способность по времени запаздывания, по доплеровской частоте?
9. Изобразите сечения $\rho(\tau, 0)$, $\rho(0, F)$ и горизонтальное сечение на уровне $\rho(\tau, F) = 0,5$ тела неопределенности для прямоугольного радиоимпульса без внутриимпульсной модуляции.
10. Изобразите сечения $\rho(\tau, 0)$, $\rho(0, F)$ и горизонтальное сечение на уровне $\rho(\tau, F) = 0,5$ тела неопределенности когерентной пачки радиоимпульсов.
11. Изобразите сечения $\rho(\tau, 0)$, $\rho(0, F)$ и горизонтальное сечение на уровне $\rho(\tau, F) = 0,5$ тела неопределенности ЛЧМ сигнала с прямоугольной огибающей.
12. Поясните причину скоростной ошибки по дальности при использовании ЛЧМ сигнала. Приведите возможные способы устранения скоростной ошибки при измерении дальности.
13. Изобразите горизонтальное сечение на уровне $\rho(\tau, F) = 0,5$ тела неопределенности шумоподобного сигнала.
14. Каким образом можно приблизить тело неопределенности радиолокационного сигнала к идеализированному игольчатому?
15. Изобразите сечения $\rho(\tau, 0)$, $\rho(0, F)$ и горизонтальное сечение на уровне $\rho(\tau, F) = 0,5$ тела неопределенности радиоимпульса с фазовой манипуляцией кодом Баркера. Какой уровень боковых лепестков у таких сигналов по оси τ , на плоскости τ, F ?
16. Какой уровень боковых лепестков имеет функция рассогласования импульсного сигнала с фазовой манипуляцией по закону М-последовательности по оси τ ?
17. Поясните, каким образом удастся обеспечить нулевой уровень боковых лепестков по оси τ функции рассогласования непрерывного сигнала с фазовой манипуляцией по закону М-последовательности?
18. Изобразите угловую функцию рассогласования эквидистантной линейной решетки в дальней зоне.
19. Поясните, каким образом реализуется согласованное и рассогласованное разрешение радиолокационных сигналов?

5. ОСНОВЫ ТЕОРИИ ИЗМЕРЕНИЯ ПАРАМЕТРОВ СИГНАЛОВ

Приведены обобщенные показатели качества оценивания. Рассмотрены общие вопросы синтеза алгоритмов оценивания неслучайных и случайных параметров сигналов при различных функциях потерь. Изложены основные принципы построения следящих и неследящих измерителей дальности, скорости и угловых координат. Представлена методика определения потенциальной точности измерения параметров сигналов, получены оценки потенциальной точности измерения запаздывания, доплеровского сдвига частоты и угловых координат.

5.1. Общие сведения

Одной из основных задач теории оптимальных методов приема сигналов является измерение (оценивание) параметров принимаемых сигналов, характеризующих, например, пространственное положение объектов (дальность, угловые координаты), направление их перемещения (проекции векторов скорости и ускорения на оси используемой системы координат) и др. Эта задача формулируется в общем виде следующим образом. На интервале времени $(0, T)$ наблюдается процесс $\mathbf{u}(t)$, в общем случае векторный, представляющий аддитивную смесь сигнала $\mathbf{s}(t, \lambda)$, отраженного от цели, и шума $\mathbf{n}(t)$:

$$\mathbf{u}(t) = \mathbf{s}(t, \lambda) + \mathbf{n}(t). \quad (5.1)$$

Параметры λ сигнала (запаздывание, доплеровский сдвиг частоты и т. д.), характеризующие цель, неизвестны. В результате измерения по принятой реализации $\mathbf{u}(t)$ необходимо сформировать возможно более точные оценки неизвестных параметров (дальности, радиальной скорости, угловых координат и др.).

Процедура оценивания включает [37]:

— пространство оцениваемых параметров Λ (элементы этого пространства будем обозначать через λ);

- пространство наблюдений U (точку в этом пространстве будем обозначать вектором u);
- обусловленное влиянием шума и флуктуациями сигнала вероятностное отображение из пространства параметров в пространство наблюдений;
- собственно правило оценивания, являющееся отображением элементов пространства наблюдений в элементы пространства параметров $\hat{\lambda} = \lambda(u)$.

В зависимости от условий измерения оцениваемые параметры могут быть либо случайными величинами, описываемыми плотностями вероятности, либо неизвестными неслучайными величинами. Если априорные данные отсутствуют, измеряемые параметры целесообразно считать неизвестными неслучайными величинами. При наличии априорных данных, полученных, например, по результатам измерений в предыдущих циклах обзора, оцениваемые параметры целесообразно считать случайными величинами. Заметим, что в процессе радиолокационного наблюдения измеряемые параметры могут меняться. В этом случае мы приходим к задаче оценивания параметров случайного процесса, которая решается методами теории фильтрации.

Вследствие влияния шума и других факторов результат измерения параметров практически всегда отличается от истинного их значения на величину, называемую *ошибкой измерения*. Если приняты меры для исключения систематических ошибок, то ошибки измерения $\epsilon = \lambda - \hat{\lambda}$, где λ — оцениваемый вектор параметров, являются случайными, и их можно описывать статистическими характеристиками. В частности, показателями качества измерения одномерной случайной величины λ могут являться математическое ожидание, среднеквадратичная ошибка, вероятная (срединная) ошибка, максимальная ошибка, дисперсия, средний риск.

По значению математического ожидания все оценки можно классифицировать на:

- несмещенные, для которых математическое ожидание равно истинным значениям параметров, т. е. $M\{\hat{\lambda}(u)\} = \lambda$;
- имеющие известное смещение, т. е. $M\{\hat{\lambda}(u)\} = \lambda + b$, где b — известный вектор (систематическая ошибка b возникает, например, при задержке сигнала в элементах приемного тракта, когда расстояние до цели измеряется по времени запаздывания принимаемого сигнала относительно зондирующего);
- имеющие неизвестное смещение, зависящее от неизвестных параметров, т. е. $M\{\hat{\lambda}(u)\} = \lambda + b(\lambda)$.

Для произвольного закона распределения $w(\epsilon)$ случайных ошибок дисперсия ошибки измерения определяется из соотношения

$$\sigma_{\varepsilon}^2 = \int_{-\infty}^{+\infty} \varepsilon^2 w(\varepsilon) d\varepsilon.$$

Вероятная ошибка соответствует такому значению $\varepsilon_0 = \varepsilon_{\text{вер}}$, при котором

$$P(|\varepsilon| \leq \varepsilon_0) = P(|\varepsilon| \geq \varepsilon_0) = 0,5.$$

За максимальную ошибку обычно принимают такое значение $\varepsilon = \varepsilon_{\text{max}}$, вероятность превышения по модулю которого составляет 0,8 %.

В случае наиболее распространенного центрированного нормального закона распределения случайных ошибок среднеквадратичная ошибка полностью характеризует вероятную и максимальную ошибки:

$$\varepsilon_{\text{вер}} = \frac{2}{3} \sigma_{\varepsilon}, \quad \varepsilon_{\text{max}} = \frac{8}{3} \sigma_{\varepsilon}.$$

Обобщенным показателем качества оценивания являются усредненные по всем значениям параметров λ и возможным результатам наблюдения u потери R (средний риск):

$$R = \int \int_{U \Lambda} \Pi(\lambda, \hat{\lambda}(u)) w(\lambda, u) d\lambda du. \quad (5.2)$$

Здесь $w(\lambda, u)$ — совместная плотность вероятности оцениваемого параметра $\lambda \in \Lambda$ и наблюдаемой величины $u \in U$, $\Pi = \Pi(\lambda, \hat{\lambda}(u))$ — функция потерь. На практике наиболее широкое распространение получили функции потерь $\Pi = \Pi(\varepsilon)$, зависящие только от ошибки оценивания $\varepsilon = \hat{\lambda} - \lambda$. Графики некоторых типичных функций потерь для одномерного случая представлены на рис. 5.1. Из графика первой функции (см. рис. 5.1, а) видно, что потери пропорциональны квадрату ошибки: $\Pi(\varepsilon) \equiv \varepsilon^2$. Из графика второй функции (см. рис. 5.1, б), следует, что потери увеличиваются пропорционально абсолютной величине ошибки: $\Pi(\varepsilon) \equiv |\varepsilon|$. Функция потерь (см. рис. 5.1, в), принимает значения, равные 0 при $|\varepsilon| \leq \Delta/2$ и равные 1 при $|\varepsilon| > \Delta/2$. Конкретный вид функции потерь определяется типом решаемой задачи. При этом оптимальная процедура измерения должна обеспечивать минимум средних потерь.

5.1.1. Измерение случайных параметров

Задача оценивания случайных параметров сигналов состоит в нахождении процедуры (алгоритма) обработки результатов наблюдения $u \in U$,

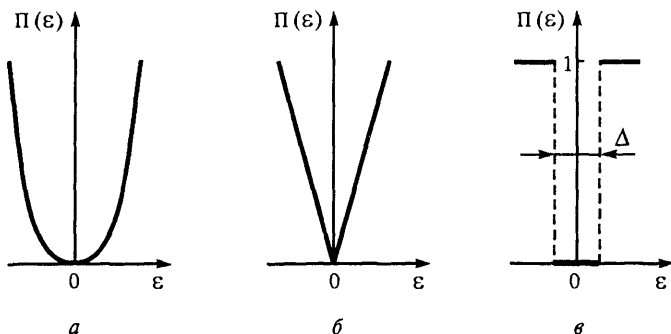


Рис. 5.1. Типичные функции потерь:

а — квадратичная; *б* — линейная по модулю; *в* — ступенчатая

обеспечивающей минимальное значение средних потерь R , определяемых выражением (5.2).

Для функции потерь, представленной на рис. 5.1, *а*, выражение для среднего риска имеет вид

$$R_{\text{кв}} = \int \int_{U \wedge \Lambda} (\lambda - \hat{\lambda}(\mathbf{u}))^2 w(\lambda, \mathbf{u}) d\mathbf{u} d\lambda. \quad (5.3)$$

Совместную плотность вероятности, входящую в (5.3), можно представить в виде

$$w(\lambda, \mathbf{u}) = w(\mathbf{u})w(\lambda|\mathbf{u}). \quad (5.4)$$

Здесь $w(\mathbf{u})$ — априорная плотность вероятности наблюдаемой величины, $w(\lambda|\mathbf{u})$ — апостериорная плотность вероятности оцениваемого параметра. Подставляя соотношение (5.4) в (5.3), будем иметь

$$R_{\text{кв}} = \int \int_{U \wedge \Lambda} (\lambda - \hat{\lambda}(\mathbf{u}))^2 w(\lambda|\mathbf{u})w(\mathbf{u}) d\lambda d\mathbf{u}. \quad (5.5)$$

Наименьшее значение потерь $R_{\text{кв}}$ можно получить, минимизируя внутренний интеграл в выражении (5.5). Поскольку сомножители подынтегрального выражения неотрицательны, продифференцировав его по λ и приравняв результат нулю, получим

$$\hat{\lambda}_{\text{кв}} = \int_{\Lambda} \lambda w(\lambda|\mathbf{u}) d\lambda. \quad (5.6)$$

Таким образом, при квадратичной функции потерь оптимальная оценка, минимизирующая средний риск, должна быть получена как среднее значение апостериорной плотности вероятности оцениваемого параметра.

Для функции, представленной на рис. 5.1, б (критерий минимума абсолютной величины ошибки), при измерении скалярного параметра средние потери определяются формулой

$$R_{\text{абс}} = \int_U \int_{-\infty}^{+\infty} |\lambda - \hat{\lambda}(\mathbf{u})| w(\lambda|\mathbf{u}) w(\mathbf{u}) d\lambda d\mathbf{u}. \quad (5.7)$$

Эти потери будут минимальны, когда внутренний интеграл минимален. Для нахождения оценки, минимизирующей $R_{\text{абс}}$, разобьем внутренний интеграл на два:

$$I(\mathbf{u}) = \int_{-\infty}^{\hat{\lambda}(\mathbf{u})} (\hat{\lambda}(\mathbf{u}) - \lambda) w(\lambda|\mathbf{u}) d\lambda + \int_{\hat{\lambda}(\mathbf{u})}^{\infty} (\lambda - \hat{\lambda}(\mathbf{u})) w(\lambda|\mathbf{u}) d\lambda. \quad (5.8)$$

Дифференцируя $I(\mathbf{u})$ по $\hat{\lambda}$ и приравнявая результат нулю, получим условие, которому должна удовлетворять оценка $\hat{\lambda}_{\text{абс}} = \lambda(\mathbf{u})$, минимизирующая средний риск $R_{\text{абс}}$:

$$\int_{-\infty}^{\hat{\lambda}_{\text{абс}}(\mathbf{u})} w(\lambda|\mathbf{u}) d\lambda = \int_{\hat{\lambda}_{\text{абс}}(\mathbf{u})}^{\infty} w(\lambda|\mathbf{u}) d\lambda. \quad (5.9)$$

Таким образом, оценка по критерию минимальной величины ошибки должна формироваться как абсцисса медианы апостериорной плотности вероятности.

Для ступенчатой функции потерь (рис. 5.1, в) имеем

$$R_{\text{ст}} = \int_{-\infty}^{\infty} w(\mathbf{u}) \left(1 - \int_{\hat{\lambda}(\mathbf{u}) - \Delta/2}^{\hat{\lambda}(\mathbf{u}) + \Delta/2} w(\lambda|\mathbf{u}) d\lambda \right) d\mathbf{u}. \quad (5.10)$$

Для минимизации потерь необходимо максимизировать внутренний интеграл, или, что то же самое, минимизировать вероятность того, что $|\varepsilon| > \Delta/2$. Очевидно, что для малых Δ наилучшей оценкой в этом случае будет то значение $\lambda \in \Lambda$, при котором апостериорная плотность $w(\lambda|\mathbf{u})$ принимает максимальное значение. Поскольку логарифмическая функция монотонна, то при условии, что измеряемые величины лежат внутри допустимой области изменения параметров Λ и функция $\ln w(\lambda|\mathbf{u})$ имеет непрерывные первые производные, необходимое условие максимума принимает вид

$$\left. \frac{d \ln w(\lambda|\mathbf{u})}{d\lambda} \right|_{\lambda=\hat{\lambda}(\mathbf{u})} = 0. \quad (5.11)$$

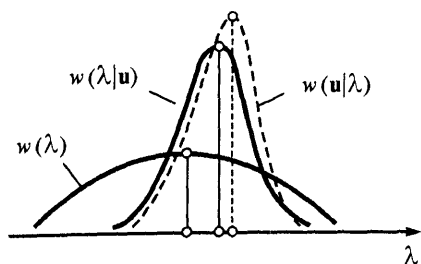


Рис. 5.2. Типичные зависимости априорной плотности вероятности $w(\lambda)$, апостериорной плотности вероятности $w(\lambda|u)$ и функции правдоподобия $w(u|\lambda)$ оцениваемого параметра

Выражение (5.11) называется *уравнением максимальной апостериорной вероятности*, оно часто используется для синтеза алгоритмов оценивания параметров сигналов в радиолокации. Применив формулу Байеса, получим

$$\ln w(\lambda|u) = \ln w(u|\lambda) + \ln w(\lambda) - \ln w(u). \quad (5.12)$$

Тогда уравнение для максимальной апостериорной вероятности можно записать в следующем виде:

$$\left. \frac{d \ln w(\lambda|u)}{d\lambda} \right|_{\lambda=\hat{\lambda}(u)} = \left. \frac{d \ln w(u|\lambda)}{d\lambda} \right|_{\lambda=\hat{\lambda}(u)} + \left. \frac{d \ln w(\lambda)}{d\lambda} \right|_{\lambda=\hat{\lambda}(u)} = 0. \quad (5.13)$$

Полученные соотношения для оценки скалярного параметра иллюстрируются графиками, представленными на рис. 5.2, где изображены зависимости априорной плотности вероятности $w(\lambda)$, апостериорной плотности вероятности $w(\lambda|u)$ и плотности вероятности $w(u|\lambda)$ (функции правдоподобия) оцениваемого параметра. Пологий характер кривой $w(\lambda)$ характеризует весьма ограниченную априорную информацию об оцениваемом параметре. Информация существенно уточняется после проведения измерения, что отражается на графике $w(u|\lambda)$ — пик становится существенно уже. Отметим, что при полном отсутствии априорных данных ($w(\lambda) = \text{const}$) график апостериорной плотности вероятности по форме совпадает с графиком функции правдоподобия.

5.1.2. Измерение неслучайных параметров

Синтез алгоритмов оценивания неслучайных параметров осуществляется обычно на основе метода максимума правдоподобия. Напомним, что плотность вероятности $w(u|\lambda)$, рассматриваемая как функция от оцениваемого параметра $\lambda \in \Lambda$, называется *функцией правдоподобия*. На практике удобно использовать логарифм этой функции. Оценка $\hat{\lambda}_{\text{мп}}(u)$ по максимуму правдоподобия — это такое значение $\lambda \in \Lambda$, при котором функция правдоподобия максимальна. Если точка максимума лежит внутри области из-

менения λ , а функция правдоподобия имеет непрерывную первую производную, то оценку по максимуму правдоподобия можно получить из уравнения правдоподобия

$$\left. \frac{d \ln w(\mathbf{u}|\lambda)}{d\lambda} \right|_{\lambda=\hat{\lambda}_{\text{mp}}(\mathbf{u})} = 0. \quad (5.14)$$

Сравнивая соотношения (5.14) и (5.13), видим, что оценка по максимуму правдоподобия и оценка по максимуму апостериорной плотности вероятности совпадают, если априорные данные отсутствуют. Отметим, что априорные данные об оцениваемых параметрах могут быть получены на основе предыдущих измерений.

5.1.3. Особенности многократного измерения случайных параметров

Единичные оценки, приведенные выше, формируются обычно за время радиолокационного контакта с целью. В процессе наблюдения за целью такие контакты и, следовательно, измерения могут повторяться и быть многократны. Таким образом, приходим к задаче синтеза алгоритмов обработки информации при многократных измерениях.

Предположим, что по результатам $n - 1$ радиолокационного контакта сформированы данные о параметрах цели, которые являются априорными при обработке n -го измерения. Далее будем рассматривать только случай регулярных измерений, для которых распределения оцениваемых параметров являются нормальными, а уравнение для логарифма апостериорной плотности вероятности имеет вид

$$\ln w(\lambda|\mathbf{u}) = -\frac{1}{2}(\lambda - \hat{\lambda}_0)^T \mathbf{D}_0^{-1}(\lambda - \hat{\lambda}_0) - \frac{1}{2}(\lambda - \hat{\lambda}_u)^T \mathbf{D}_u^{-1}(\lambda - \hat{\lambda}_u) + \text{const}, \quad (5.15)$$

где $\hat{\lambda}_0$ — оценка, а \mathbf{D}_0 — ковариационная матрица ошибок, сформированные по результатам предыдущих $n - 1$ измерений; $\hat{\lambda}_u$ — оценка неизвестных параметров, полученная по текущему измерению; \mathbf{D}_u — ковариационная матрица ошибок текущего измерения. При сделанных допущениях апостериорная плотность вероятности оцениваемых параметров является нормальной, и логарифм ее можно представить в виде

$$\ln w(\lambda|\mathbf{u}) = -\frac{1}{2}(\lambda - \hat{\lambda}_p)^T \mathbf{D}_p^{-1}(\lambda - \hat{\lambda}_p) + \text{const}. \quad (5.16)$$

Здесь λ_p , \mathbf{D}_p — математическое ожидание и ковариационная матрица результирующего апостериорного распределения соответственно. Из формул (5.15) и

(5.16) можно получить соотношения, связывающие параметры априорного $(\hat{\lambda}_0, \mathbf{D}_0)$, текущего $(\hat{\lambda}_u, \mathbf{D}_u)$ и апостериорного $(\hat{\lambda}_p, \mathbf{D}_p)$ распределений:

$$\mathbf{D}_p^{-1} = \mathbf{D}_0^{-1} + \mathbf{D}_u^{-1}, \quad (5.17)$$

$$\hat{\lambda}_p \mathbf{D}_p^{-1} = \mathbf{D}_0^{-1} \hat{\lambda}_0 + \mathbf{D}_u^{-1} \hat{\lambda}_u. \quad (5.18)$$

Откуда получаем

$$\hat{\lambda}_p = (\mathbf{D}_0^{-1} \hat{\lambda}_0 + \mathbf{D}_u^{-1} \hat{\lambda}_u) \mathbf{D}_p. \quad (5.19)$$

Поскольку весовые коэффициенты, входящие в соотношение (5.19), в соответствии с формулой (5.17) удовлетворяют условию

$$(\mathbf{D}_0^{-1} \mathbf{D}_p + \mathbf{D}_u^{-1} \mathbf{D}_p) = (\mathbf{D}_0^{-1} + \mathbf{D}_u^{-1}) \mathbf{D}_p = 1, \quad (5.20)$$

апостериорную оценку можно представить в виде суммы априорной оценки и невязки с коэффициентом \mathbf{B}_p :

$$\hat{\lambda}_p = \hat{\lambda}_0 + \mathbf{D}_u^{-1} \mathbf{D}_p (\hat{\lambda}_u - \hat{\lambda}_0) = \hat{\lambda}_0 + \mathbf{B}_p (\hat{\lambda}_u - \hat{\lambda}_0). \quad (5.21)$$

В частности, при оценке скалярного параметра результирующая дисперсия оценки и собственно оценка определяются соотношениями:

$$1/\sigma_p^2 = 1/\sigma_0^2 + 1/\sigma_u^2, \quad (5.22)$$

$$\hat{\lambda}_p = \hat{\lambda}_0 + \frac{\sigma_p^2}{\sigma_u^2} (\hat{\lambda}_u - \hat{\lambda}_0). \quad (5.23)$$

Выражение (5.23) показывает способ формирования апостериорной оценки: к априорной оценке $\hat{\lambda}_0$, сформированной по предыдущим измерениям, необходимо добавить невязку $\hat{\lambda}_u - \hat{\lambda}_0$ с коэффициентом, зависящим от результирующей дисперсии и дисперсии текущего измерения. Соотношение (5.22) поясняет механизм повышения точности при многократных измерениях. Следует отметить, что алгоритм формирования текущей отметки, описываемый выражением (5.23), по своей структуре близок к алгоритму фильтрации радиолокационных измерений, реализуемому в следящих измерителях [39].

5.2. Измерение параметров радиолокационных сигналов

Рассмотрим процедуру измерения параметров радиолокационного сигнала. Она заключается в том, что по принятой реализации необходимо сформировать функцию, пропорциональную апостериорной плотности ве-

роятности (5.12) или монотонной функции от нее в заданном диапазоне изменения параметров. Преобразуем выражение (5.12), введя отношение правдоподобия

$$l(\lambda) = \frac{w(\mathbf{u}|\lambda)}{w(\mathbf{u})}, \quad (5.24)$$

где $w(\mathbf{u}|\lambda)$ — плотность вероятности реализаций \mathbf{u} при наличии полезного сигнала с параметром λ , $w(\mathbf{u})$ — плотность вероятности помехи. Подставляя соотношение (5.24) в (5.12), получим

$$\ln w(\lambda|\mathbf{u}) = \ln l(\lambda) + \ln w(\lambda), \quad (5.25)$$

откуда следует, что устройство, предназначенное для измерения параметров радиолокационных сигналов, по принятой реализации должно формировать функцию, пропорциональную логарифму отношения правдоподобия, т. е. выполнять такие же операции, что и при обнаружении сигнала. Можно показать, что для сигнала со случайной начальной фазой логарифм отношения правдоподобия имеет вид [35]

$$\ln l(\lambda) \approx Z(\lambda) - q^2/2, \quad (5.26)$$

а для сигнала со случайными начальной фазой и амплитудой

$$\ln l(\lambda) = \frac{Z(\lambda)^2}{4(1 + q^2/2)} - \ln \left(1 + \frac{q^2}{2} \right). \quad (5.27)$$

Здесь $Z(\lambda)$ — модуль комплексного корреляционного интеграла

$$z(\lambda) = \frac{1}{N_0} \int_0^{T_c} u(t)s^*(t, \lambda)dt; \quad (5.28)$$

$q^2 = 2E/N_0$ — параметр обнаружения.

Таким образом радиолокационный измеритель должен вычислять либо $Z(\lambda)$, либо $Z(\lambda)^2$ в ожидаемом диапазоне изменения λ .

Отметим, что все измерители можно разделить на две большие группы: следящие и неследящие. Неследящие измерители обычно формируют сигнал, пропорциональный апостериорной плотности вероятности, во всем диапазоне возможных значений измеряемых параметров. Такой подход оправдан, если отсутствует априорная информация о параметрах сигнала. При наличии априорной информации можно ограничиться анализом области, в которую измеряемый параметр попадает с вероятностью, близкой к едини-

це, и корректировать положение этой области (следить за ней). На этом основана работа следящих измерителей. Далее рассмотрим основные структурные схемы измерителей, а также качество получаемых оценок.

5.2.1. Неследящие измерители дальности

Рассмотрим задачу измерения дальности r до цели при известных (фиксированных) остальных ее координатах. Оценку дальности \hat{r} определяют через оценку $\hat{\tau}_3$ времени запаздывания отраженного сигнала относительно зондирующего:

$$\hat{r} = \frac{c}{2} \hat{\tau}_3. \quad (5.29)$$

При измерении запаздывания необходимо по принятой реализации формировать статистику $Z(\tau)$ либо $Z(\tau)^2$ для всей области возможных значений запаздывания: $\tau_{\min} < \tau < \tau_{\max}$, где τ_{\min} , τ_{\max} — минимальное и максимальное время запаздывания сигнала цели. За оценку времени запаздывания следует принимать то значение τ , при котором статистика $Z(\tau)$ или $Z(\tau)^2$ принимает максимальное значение.

Вычисление корреляционного интеграла $z(\tau)$ в неследящих измерителях может производиться фильтровыми, корреляционными и корреляционно-фильтровыми устройствами.

Обобщенная схема неследящего фильтрового измерителя времени запаздывания приведена на рис. 5.3, а. Это устройство одноканальное, включает согласованный фильтр, детектор и решающее устройство, предназначенное для нахождения аргумента, соответствующего максимуму $Z(\tau)$. На рис. 5.3, б представлены также временные диаграммы напряжения на выходе

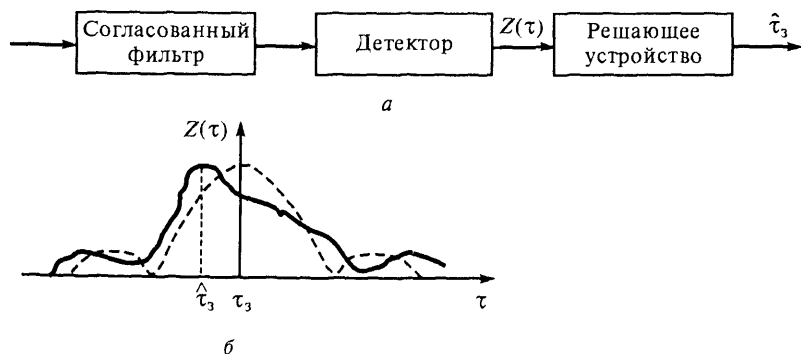


Рис. 5.3. Структурная схема неследящего фильтрового измерителя дальности (а) и временные диаграммы напряжения на выходе детектора (б)

детектора $Z(\tau)$ при отсутствии (штриховая линия) и наличии (сплошная линия) шума. Выделен момент времени $\hat{\tau}_3$, соответствующий максимуму выходного сигнала, фиксируемый с помощью решающего устройства. К смещению максимума, т. е. к ошибке измерения, приводит наличие шума.

Вычисление корреляционного интеграла корреляционными или корреляционно-фильтровыми устройствами влечет необходимость применения многоканальных измерителей. На рис. 5.4 представлена структурная схема корреляционного измерителя, который содержит набор каналов, настроенных на различные времена запаздывания с шагом Δt . По сигналу $u_{сч}$ решающее устройство осуществляет выбор канала с максимальным выходным напряжением. Информация о времени запаздывания принимаемого сигнала заложена в номере этого канала. Среднеквадратическая ошибка измерения запаздывания определяется, в основном, шумами приемного устройства и шагом Δt . Для реализации точности измерения, близкой к потенциальной, шаг Δt должен быть как минимум в 3—5 раз меньше разрешающей способности РЛС по дальности. При таком выборе шага отраженный от цели сигнал присутствует в нескольких соседних каналах. Это дает возможность дополнительно повысить точность измерений, для чего необходимо осуществлять параболическую аппроксимацию выходных данных в окрестности максимума, используя информацию об амплитуде сигналов в соседних каналах, а за оценку времени запаздывания принимать координату вершины аппроксимирующей функции.

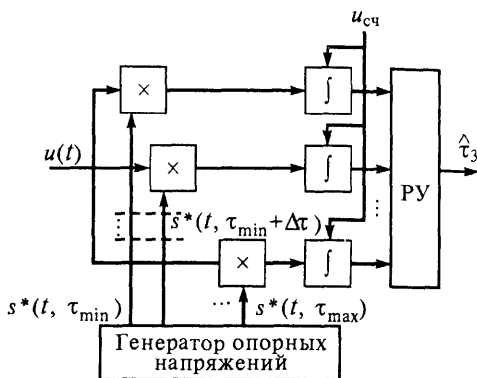


Рис.5.4. Структурная схема корреляционного измерителя дальности

5.2.2. Неследящие измерители скорости

Радиальную скорость цели обычно оценивают по результатам измерения доплеровского сдвига частоты F_d отраженного сигнала. Для получения оценки \hat{F}_d необходимо по принятой реализации вычислить модуль (либо квадрат модуля) комплексного корреляционного интеграла как функцию доплеровской частоты и найти аргумент (частоту), соответствующий максимуму этого модуля. В общем случае неследящий измеритель скорости состоит из k каналов, каждый из которых содержит согласованный фильтр

(СФ), настроенный на соответствующее значение доплеровского сдвига частоты, и амплитудный детектор. По номеру канала, на выходе которого амплитуда сигнала максимальна, оценивают доплеровскую частоту. При этом расстройка соседних частотных каналов должна быть как минимум в 3—5 раз меньше разрешающей способности измерителя по частоте. Для реализации точности измерения, близкой к потенциальной, по соседним отсчетам в окрестности максимума следует осуществлять параболическую аппроксимацию выходных данных, а за оценку доплеровского сдвига частоты принимать соответствующую координату вершины аппроксимирующей функции.

5.2.3. Совместное измерение запаздывания и доплеровского сдвига частоты

Устройство, предназначенное для совместного измерения запаздывания и доплеровского сдвига частоты, должно формировать функцию правдоподобия (или логарифм от нее) в интересующем нас диапазоне изменения оцениваемых параметров, а в качестве оценки принимать те значения, при которых эта функция имеет максимум. На рис. 5.5, *а* представлен пример обобщенной структурной схемы такого устройства. Оно состоит из k каналов, перекрывающих весь ожидаемый частотный диапазон доплеровского сдвига частоты сигнала цели. Каждый канал настроен на определенную частоту, обеспечивает согласованную обработку сигнала с фиксированным доплеровским сдвигом и включает согласованный фильтр и детектор огибающей. На выходе каждого канала (см. рис. 5.5, *б*) формируется функция правдоподобия по задержке для фиксированной частоты. Сигнал на выходе всего устройства воспроизводит функцию правдоподобия в координатах «запаздывание—доплеровский сдвиг частоты» в заданной области изменения параметров τ , F . Поскольку исходный процесс — сумма полезного сигнала и шума, выходной сигнал содержит три компонента. Первый компонент соответствует сигнальной составляющей. Его амплитуда зависит от мощности принимаемого сигнала, форма определяется функцией рассогласования, а координаты τ , F максимума соответствуют параметрам сигнала, отраженного от цели. Второй и третий компоненты обусловлены шумом и взаимодействием сигнала и шума при обработке. Они являются случайными, приводят к смещению максимума и, следовательно, к ошибкам измерения.

Если бы функция рассогласования сигнала имела один очень узкий центральный пик (см. рис. 4.10) и амплитуда сигнальной составляющей существенно превышала шумовую, то шум практически не влиял бы на положение максимума и, следовательно, не приводил бы к ошибкам измерения. Возможным приближением к такой функции может рассматриваться функция рассогласования сигнала в виде одиночного прямоугольного импульса

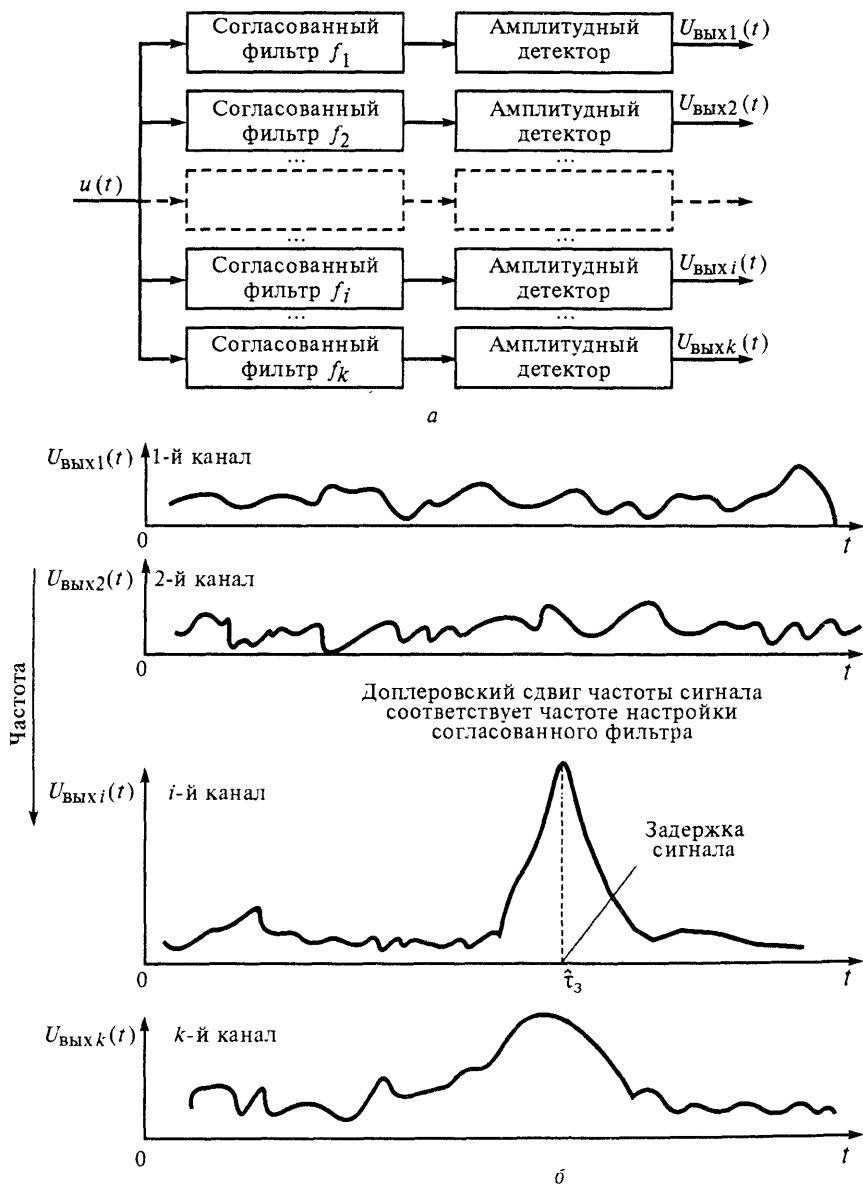


Рис. 5.5. Обобщенная структурная схема измерителя запаздывания и доплеровского сдвига частоты (а) и эпюры сигналов на выходах каналов (б)

длительностью τ_n . Она имеет единственный пик, ширина которого по оси τ прямо пропорциональна, а по оси F — обратно пропорциональна длитель-

ности импульса. Поскольку протяженность пика конечна, шум приводит к смещению максимума функции, т. е. к ошибкам измерения. При регулярных измерениях ошибки лежат в окрестности главного максимума, определяемой параметрами импульса. Уменьшение, например, длительности импульса приводит к уменьшению ширины пика функции рассогласования по оси t . При этом точность измерения времени запаздывания возрастает. Однако возрастают ошибки измерения частоты, поскольку ширина пика функции рассогласования по оси F увеличивается. Таким образом, варьируя единственным параметром — длительностью импульса, можно влиять на ошибки измерения дальности (скорости), однако уменьшение ошибок измерения по одной координате приводит к увеличению ошибок измерения по другой.

Если требуется одновременно уменьшать ошибки измерения и дальности, и скорости, необходимо переходить к сложным сигналам, например когерентным импульсным последовательностям. Анализ функции рассогласования сигнала данного типа (см. рис. 4.5) показывает, что ширину главного пика по оси частот, а следовательно, и ошибки измерения частоты, можно уменьшить, увеличив длительность последовательности. Укорачивая импульсы, можно уменьшить ширину главного пика по оси времени. Таким образом, увеличивая длительность последовательности и ширину спектра импульсов, можно уменьшать площадь области высокой корреляции функции рассогласования сигнала в окрестности максимума, что дает возможность одновременно повысить точность измерения и времени запаздывания, и доплеровского сдвига частоты. Однако платой за это является появление побочных пиков функции рассогласования, что может приводить к грубым ошибкам измерений.

Рассмотрим гауссовский импульс с внутриимпульсной линейной частотной модуляцией (ЛЧМ). Смещения максимума выходного сигнала, обусловленные шумами, при регулярных измерениях лежат в области высокой корреляции функции рассогласования. Анализ свойств функции рассогласования этого сигнала, проведенный в гл. 4, показывает, что увеличение τ_n и Δf_c приводит к уменьшению ширины сечения области высокой корреляции по осям t , F одновременно. Однако частотная модуляция изменяет область высокой корреляции функции рассогласования сигнала. При этом протяженность проекции области высокой корреляции на ось t остается фиксированной (см. рис. 4.8), поэтому точность измерения запаздывания не изменяется. Протяженность проекции области высокой корреляции на ось F увеличивается. Это при определенных условиях может привести к ухудшению точности измерения частоты.

Более подробно влияние параметров сигналов на точность измерения координат цели рассмотрим в § 5.5.

5.3. Следящие измерители дальности и доплеровского сдвига частоты

В процессе радиолокационного наблюдения координаты целей изменяются. В качестве примера на рис. 5.6 представлены полученные экспериментально результаты изменения доплеровского сдвига частоты сигналов, отраженных от групповой воздушной цели, состоящей из четырех летательных аппаратов, при длительном времени наблюдения [42]. По оси абсцисс отложено время (с), по оси ординат — значение доплеровского сдвига частоты отраженного сигнала (Гц). Видно, что доплеровский сдвиг частоты в процессе наблюдения весьма сильно меняется, что вызвано изменением ракурса, атмосферными флуктуациями, маневрированием и т. д. Таким образом, измеряемые параметры являются функциями времени. Радиолокационный измеритель должен формировать оценки с учетом данных, полученных ранее. Такая информация позволяет анализировать не весь диапазон возможного изменения параметров, а только часть его, вероятность нахождения полезного сигнала в котором близка к единице, и оценивать не измеряемый параметр, а отклонение его от прогнозируемого значения (невязку), уточняя последнее по результатам измерения. Такая процедура реализуется в следящих измерителях.

Обобщенная структурная схема следящего измерителя представлена на рис. 5.7. Измеритель включает в себя фильтр-экстраполятор, вырабаты-

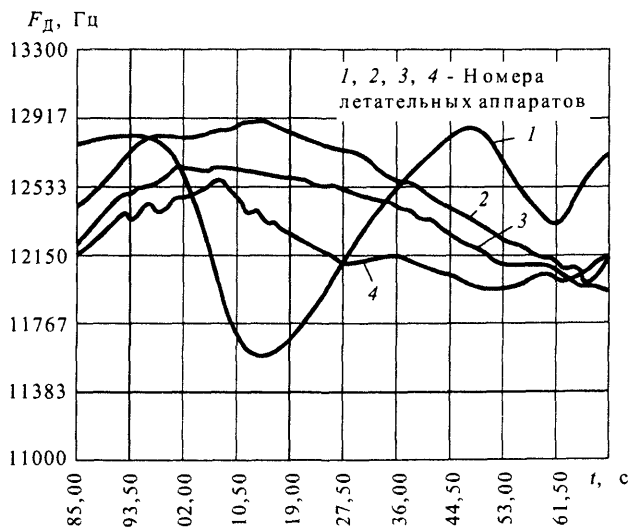


Рис. 5.6. Изменения доплеровского сдвига частоты сигналов, отраженных от воздушных целей

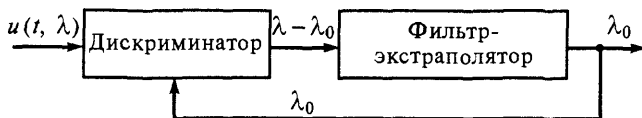


Рис. 5.7. Обобщенная структурная схема слеящего измерителя

вающий опорное напряжение, соответствующее прогнозируемому значению (априорной оценке λ_0) измеряемого параметра, и дискриминатор. В дискриминаторе формируется максимально правдоподобная оценка рассогласования между текущим и опорным значениями измеряемого параметра (оценка невязки). Для нахождения операций, выполняемых в дискриминаторе, разложим в ряд логарифм отношения правдоподобия в окрестности опорной точки λ_0 , ограничившись второй степенью [43]:

$$l(\lambda) \approx l(\lambda_0) + l'(\lambda_0)(\lambda - \lambda_0) + l''(\lambda_0)(\lambda - \lambda_0)^2 / 2. \quad (5.30)$$

Приравняв нулю первую производную функции $l(\lambda)$:

$$l'(\lambda) = l'(\lambda_0) + l''(\lambda_0)(\lambda - \lambda_0) = 0, \quad (5.31)$$

получим выражение для оптимальной оценки измеряемого параметра:

$$\hat{\lambda} = \lambda_0 - l'(\lambda_0) / l''(\lambda_0) = \lambda_0 + \Delta_\lambda, \quad (5.32)$$

где $\Delta_\lambda = -l'(\lambda_0) / l''(\lambda_0) = -Z'(\lambda_0) / Z''(\lambda_0)$ — оценка малого отклонения параметра, формируемая в дискриминаторе. Оценка (5.32) определяется из условия обращения в нуль первой производной логарифма отношения правдоподобия. Однако дискриминатор удобнее строить так, чтобы его выходная функция обращалась в нуль в опорной точке λ_0 , которая становится точкой настройки дискриминатора. Формально это можно обеспечить, поменяв местами $\hat{\lambda}$ и λ_0 в разности $\hat{\lambda} - \lambda_0$ и изменив знак первой производной. Тогда формируемую оптимальным дискриминатором оценку малого отклонения параметра можно представить формулой

$$\Delta_\lambda = Z'(\lambda) / Z''(\lambda). \quad (5.33)$$

Выражение (5.33) определяет структуру оптимального дискриминатора: в нем должны формироваться первая и вторая производные $Z(\lambda)$ в опорной точке, а затем вычисляться их отношение. При этом обеспечиваются: нечетная симметрия и линейность дискриминационной характеристики (которая является зависимостью выходного напряжения дискриминатора $U_{\text{дис}}$ от отклонения параметра) при малых рассогласованиях и независимость выходного напряжения дискриминатора от амплитуды принимаемого сигнала.

Пример дискриминационной характеристики представлен на рис. 5.8, *г*. Участок $\Delta\lambda$, в пределах которого дискриминационная характеристика линейна, называется *рабочим*. На том же рисунке приведены графики, иллюстрирующие изменения модуля корреляционного интеграла и его первой и второй производных от отклонения параметров. При большом отношении сигнал—шум зависимость модульного значения корреляционного интеграла от параметра рассогласования $\lambda - \lambda_0$ можно аппроксимировать квадратичной функцией (рис. 5.8, *а*), первой производной — линейной функцией (рис. 5.8, *б*), а второй производной — постоянным значением (рис. 5.8, *в*). Поскольку амплитуда сигнала входит как множитель в функции $Z'(\lambda)$ и $Z''(\lambda)$, отношение $Z'(\lambda)/Z''(\lambda)$ на рабочем участке от нее не зависит.

Учитывая, что корреляционный интеграл и его вторая производная меняются примерно одинаково, хотя и имеют противоположные знаки, можно упростить дискриминатор, заменив отношение $Z'(\lambda)/Z''(\lambda)$ на $-Z'(\lambda)/Z(\lambda)$. Модифицированный дискриминатор содержит два канала обработки входного сигнала: измерительный и опорный. В измерительном канале формируется нечетная функция рассогласования. Сигнал опорного канала используется для нормировки. Довольно часто вычисление корреляционного интеграла и нормировка (с помощью автоматической регулировки усиления) могут выполняться в общем тракте приемника. Также приближенно может вычисляться производная корреляционного интеграла. В основе приближенного вычисления лежит замена производной алгебраической разностью корреляционных интегралов, вычисленных для близких параметров рассогласования. Примеры таких дискриминаторов приведены ниже.

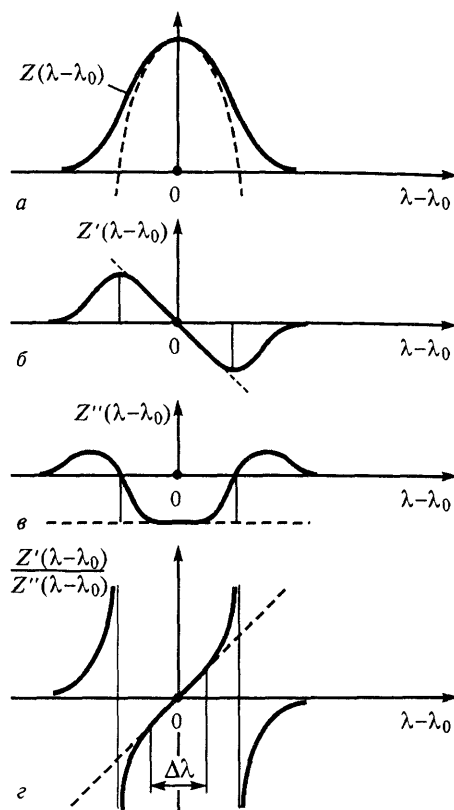


Рис. 5.8. Дискриминационная характеристика и ее формирование:

а — изменение модуля корреляционного интеграла $Z(\lambda - \lambda_0)$; *б* — первая производная $Z'(\lambda - \lambda_0)$; *в* — вторая производная $Z''(\lambda - \lambda_0)$; *г* — дисперсионная характеристика

Временные дискриминаторы. Рассмотрим примеры реализации временных дискриминаторов при обработке сигнала на фоне белого шума.

При реализации временного дискриминатора с линейным амплитудным детектором операцию дифференцирования $dZ(\tau)/d\tau$ можно заменить (с точностью до константы) вычислением конечной разности:

$$dZ(\tau)/d\tau|_{\tau=\tau_3} \approx Z(\tau_3 + \Delta\tau/2) - Z(\tau_3 - \Delta\tau/2). \quad (5.34)$$

Сигнал рассогласования $U = U(\Delta)$, пропорциональный производной корреляционного интеграла, можно получить на выходе устройства, схема которого представлена на рис. 5.9. Здесь генератор опорного напряжения формирует *стробирующий импульс*

$$U_{\text{стр}}(t) = U_1(t - (\hat{\tau}_3 + \Delta\tau/2)) - U_2(t - (\hat{\tau}_3 - \Delta\tau/2)), \quad (5.35)$$

где

$$U_1(t) = \begin{cases} 1 & \text{при } 0 < t < T_c, \\ 0 & \text{для остальных } t, \end{cases}$$

состоящий из двух полустробиов (рис. 5.10, б): $U_1 = U_1(\hat{\tau}_3 + \Delta\tau/2)$ и $U_2 = -U_1(\hat{\tau}_3 - \Delta\tau/2)$. Положение середины стробирующего импульса на оси времени соответствует точке априорной оценки времени запаздывания $\hat{\tau}_3$ и определяется напряжением, поступающим на управляющий вход генератора. На первый вход перемножителя поступает сигнал $U_{\text{ад}}$ с выхода приемника, а на второй — стробирующий импульс. При обработке прямоугольного импульса форма сигнала на выходе детектора показана рис. 5.10, а. В перемножителе формируются сигналы, пропорциональные площадям перекрытия импульса $U_{\text{ад}}$ со стробами, т.е. пропорциональные произведениям $U_{\text{ад}} U_1(\hat{\tau}_3 + \Delta\tau/2)$ и

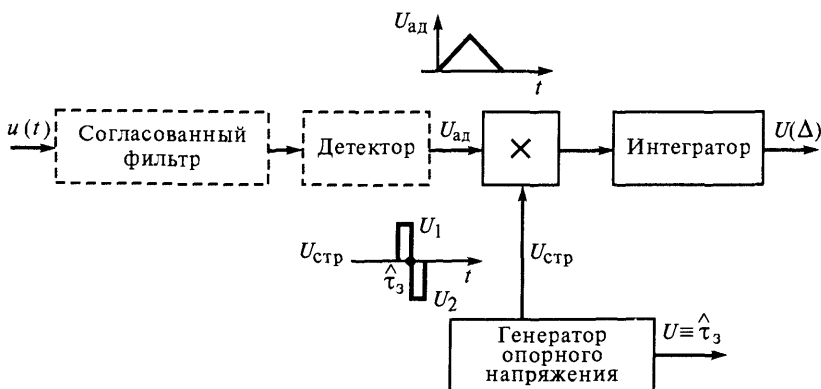


Рис. 5.9. Обобщенная структурная схема временного дискриминатора

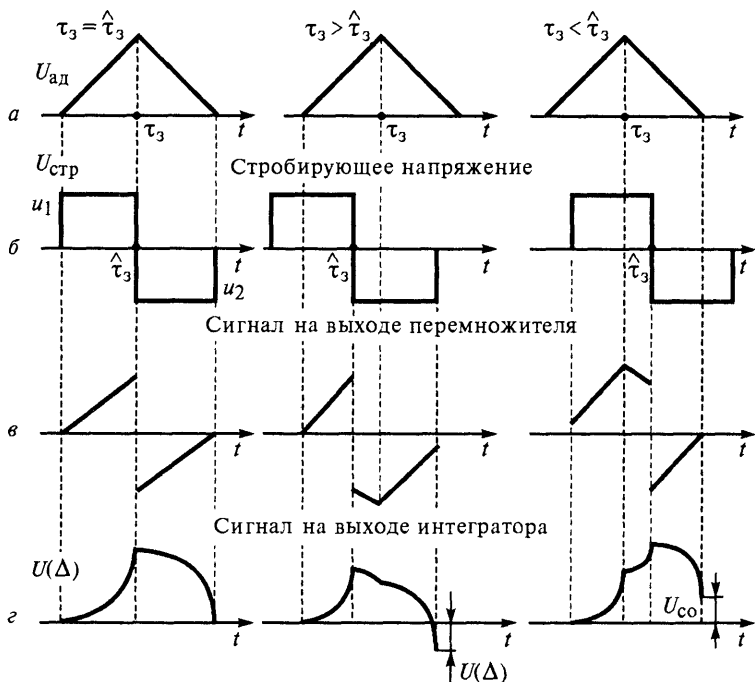


Рис. 5.10. Диаграммы, поясняющие работу временного дискриминатора:

а — сигнал на выходе детектора; б — стробирующее напряжение; в — сигнал на выходе перемножителя; г — сигнал на выходе интегратора

$U_{ад} U_1(\hat{\tau}_3 - \Delta\tau/2)$ (рис. 5.10, в). Выходное напряжение интегратора

$$U(\Delta) = \int_0^{\tau_c} U_{ад} U_{стр}(t) dt \quad (5.36)$$

пропорционально разности площадей перекрытия указанных импульсов. Оно характеризует величину и знак ошибки совмещения середины строба с полезным сигналом (рис. 5.10, г).

Во временных дискриминаторах с фазовым детектированием выходной сигнал пропорционален [39]

$$\Delta_\phi = \operatorname{Re} \left[\frac{dz(\tau)}{d\tau} z^*(\tau) \right]. \quad (5.37)$$

При практической реализации операция дифференцирования заменяется вычислением конечной разности (с точностью до константы):

$$\Delta_\phi \approx \operatorname{Re} \left[(z(\hat{\tau}_3 + \Delta\tau/2) - z(\hat{\tau}_3 - \Delta\tau/2)) z^*(\hat{\tau}_3) \right]. \quad (5.38)$$

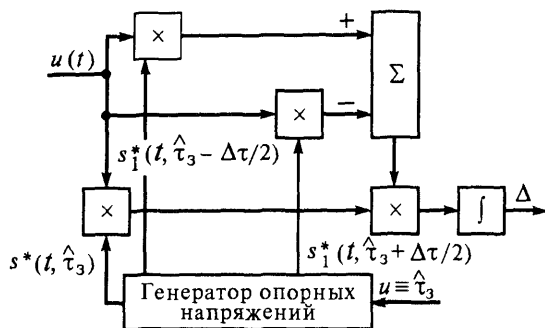


Рис. 5.11. Структурная схема временного дискриминатора с фазовым детектированием

амплитудами $s_1^*(t, \hat{t}_3)$, $s_1^*(t, \hat{t}_3 + \Delta\tau/2)$ (опережающее) и $s_1^*(t, \hat{t}_3 - \Delta\tau/2)$ (отстающее). При совпадении по времени принимаемого сигнала $s(t)$ во входной реализации $u(t) = s(t) + n(t)$, где $n(t)$ — шум, и опорного сигнала $s(t, \hat{t}_3)$ осуществляется полная фазовая демодуляция. После фильтрации это напряжение используется как опорное в выходном перемножителе.

Частотные дискриминаторы. Существует несколько разновидностей дискриминаторов, использующихся в системах слежения за частотой принимаемого сигнала.

В частотных дискриминаторах с амплитудным детектированием, схема которого приведена на рис. 5.12, вычисляется конечная разность вида

$$\Delta \approx Z(\hat{F}_D + \Delta F/2) - Z(\hat{F}_D - \Delta F/2). \quad (5.39)$$

Здесь $Z(\hat{F}_D + \Delta F/2)$, $Z(\hat{F}_D - \Delta F/2)$ — выходные сигналы каналов, расстроенных по частоте на ΔF . Управляемый гетеродин вырабатывает коле-

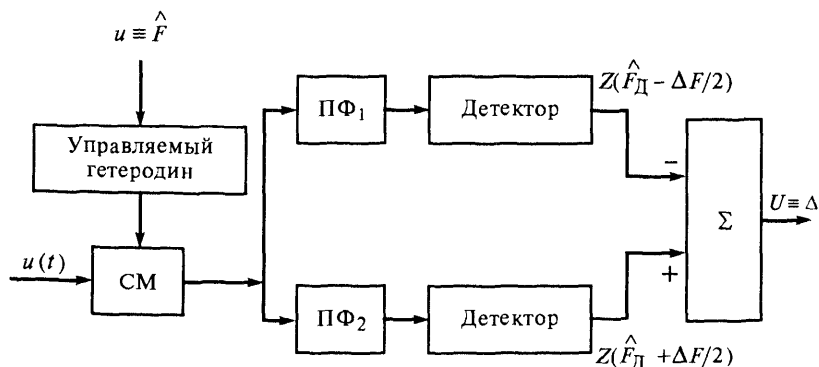


Рис. 5.12. Структурная схема частотного дискриминатора с амплитудным детектированием

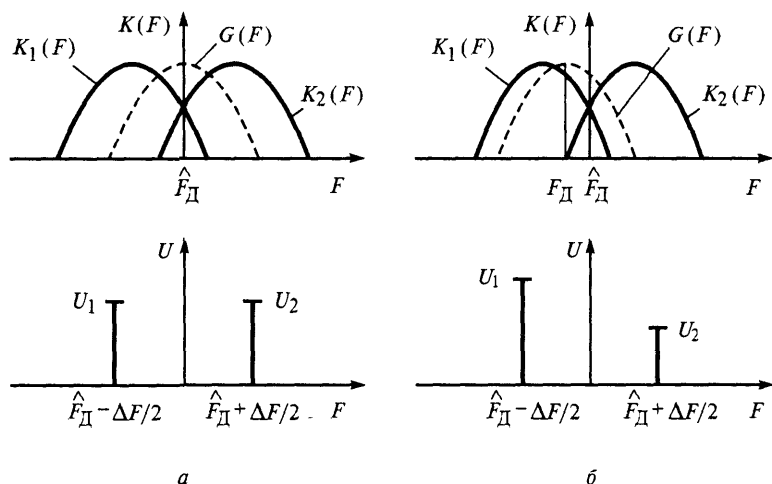


Рис. 5.13. Сигналы на выходе каналов частотного дискриминатора с амплитудным детектированием:

a — при $F_d = \hat{F}_d$; b — при $F_d < \hat{F}_d$

бание, частота которого определяется напряжением, пропорциональным ожидаемому доплеровскому сдвигу частоты сигнала \hat{F}_d . Принимаемое колебание преобразуется в смесителе на промежуточную частоту. Преобразованное (демодулированное) колебание поступает на два канала обработки, в которых полосовые фильтры ПФ₁ и ПФ₂ расстроены по частоте. В каналах осуществляется вычисление модулей $Z(\hat{F}_d + \Delta F/2)$ и $Z(\hat{F}_d - \Delta F/2)$. Разность напряжений на выходе детекторов (сигнал ошибки) будет пропорциональна разности априорной и текущей оценок доплеровской частоты. На рис. 5.13 представлены амплитудно-частотные характеристики $K_1(F)$ и $K_2(F)$ фильтров, а также спектр $G(F)$ сигнала на выходе смесителя. Напряжение на выходе детектора каждого канала определяется степенью перекрытия амплитудно-частотного спектра сигнала и амплитудно-частотной характеристики соответствующего фильтра. Если частота сигнала совпадает с \hat{F}_d , то амплитуды сигналов на выходах фильтров в среднем одинаковы (рис. 5.13, a). Если частота сигнала уменьшается, то амплитуда сигнала на выходе первого фильтра (с амплитудно-частотной характеристикой $K_1(F)$) возрастает (рис. 5.13, b), а на выходе второго фильтра (с амплитудно-частотной характеристикой $K_2(F)$) уменьшается. Если частота сигнала увеличивается, то картина меняется на обратную.

При построении частотных дискриминаторов с фазовым детектированием используют соотношение

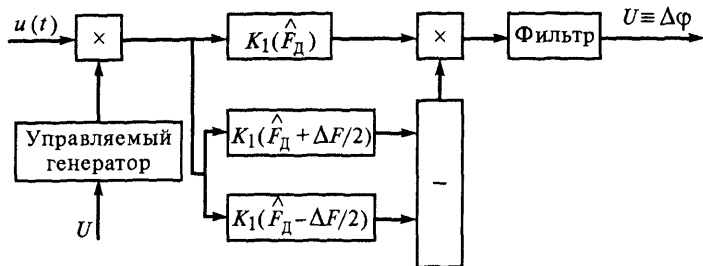


Рис. 5.14. Структурная схема частотного дискриминатора с фазовым детектированием

$$\Delta_{\Phi} = \operatorname{Re} \left[\frac{dz(F_D)}{dF_D} z^*(F_D) \right] \quad (5.40)$$

при $F_D = \hat{F}_D$, или его представление в виде

$$\Delta_{\Phi} \approx \operatorname{Re} \left[\left(z(\hat{F}_D + \Delta F/2) - z(\hat{F}_D - \Delta F/2) \right) z^*(\hat{F}_D) \right]. \quad (5.41)$$

На рис. 5.14 представлен вариант построения схемы частотного дискриминатора с фазовым детектированием, формирующего выходной сигнал в соответствии с формулой (5.41). На фазовый детектор поступают напряжения двух фильтровых каналов. Параметры первого канала, включающего фильтр $K_1(F)$, согласованы со спектром сигнала. На его выходе формируется величина, пропорциональная $z^*(\hat{F}_D)$. Сигнал на выходе второго канала, включающего фильтры, настроенные на частоты $\hat{F}_D + \Delta F/2$ и $\hat{F}_D - \Delta F/2$ и устройство вычитания, пропорционален разности $z(\hat{F}_D + \Delta F/2) - z(\hat{F}_D - \Delta F/2)$. После перемножения сигналов в фазовом детекторе и фильтрации формируется сигнал ошибки.

5.4. Измерение угловых координат

Измерение угловых координат основано на определении положения волнового фронта отраженной от цели электромагнитной волны относительно приемной апертуры. Процесс измерения угловых координат часто называют *пеленгацией*, а соответствующие измерители — *пеленгаторами*. Последние могут содержать один или несколько приемных каналов. Соответственно, методы пеленгации подразделяются на одноканальные и многоканальные.

5.4.1. Многоканальные измерители угловых координат

В общем случае сигнал $s(t, \lambda)$, отраженный от цели, принимается антенной решеткой, состоящей из N элементов. На рис. 5.15 обозначено: \mathbf{r}_0 — единичный вектор, характеризующий направление прихода электромагнитной волны; \mathbf{r}_i — вектор, определяющий положение i -го элемента антенной решетки; \mathbf{r} — единичный вектор, характеризующий положение основного лепестка диаграммы направленности антенны в пространстве.

Предположим также, что сигнал является узкополосным, т. е. интервал корреляции комплексной огибающей сигнала (для сигналов без внутримпульсной модуляции — длительность импульса) существенно превышает временной интервал между моментами прихода сигнала в наиболее разнесенные точки апертуры антенны. Это допущение позволяет разделить пространственно-временную обработку на пространственную и временную и далее рассматривать только пространственную обработку.

Сигнал на выходе элементов антенны запишем в виде N -мерного вектора $\mathbf{U}(t)$ с компонентами

$$u_i(t) = s\left(t + \frac{\mathbf{r}_i \mathbf{r}_0}{c}\right) + n_i(t), \quad 0 \leq i \leq N-1, \quad (5.42)$$

где $n_i(t)$ — компонента аддитивного шума, $\mathbf{r}_i \mathbf{r}_0$ — скалярное произведение векторов, $c = 3 \cdot 10^8$ м/с. В случае линейной эквидистантной антенной решетки $\mathbf{r}_i \mathbf{r}_0 = (i-1)d \sin \theta_0$, где θ_0 — угол между направлением распространения сигнала и нормалью к антенной решетке; d — расстояние между антенными элементами, i — порядковый номер антенного элемента. Введем вектор-столбец \mathbf{A} , называемый *опорным вектором* и характеризующий ориентацию основного лепестка диаграммы направленности антенны, элементы которого имеют вид

$$a_i = b_i \exp\left(j \frac{2\pi}{\lambda} \mathbf{r}_i \mathbf{r}\right), \quad (5.43)$$

где b_i — элемент вектора весовых коэффициентов.

Формирование результирующего сигнала происходит путем суммирования выходных сигналов элементов антенной решетки с весами b_i и задержками $\mathbf{r}_i \mathbf{r}/c$:

$$y = \mathbf{A}^* \mathbf{U},$$

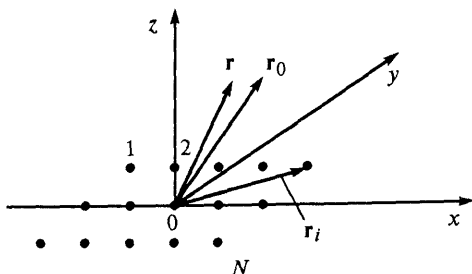


Рис. 5.15. Плоская антенная решетка

где $()^{*T}$ — операции сопряжения и транспонирования. Если задержка сигнала в каждом антенном элементе равна $r_1 r/c$, то основной лепесток диаграммы направленности антенны будет ориентирован по вектору r .

Фазированная антенная решетка позволяет сформировать несколько лучей, что дает возможность перейти к параллельным методам обзора пространства и измерять угловые координаты. При формировании m лучей алгоритм вычисления может быть записан в матричной форме:

$$Y = A^* U, \quad (5.44)$$

где Y — вектор выходных сигналов, соответствующий m лучам: $Y^T = (y_1, y_2, \dots, y_m)$, A — матрица преобразования размерности $N \times m$.

Один из возможных вариантов схемы многоканального измерителя угловой координаты представлен на рис. 5.16. Измеритель включает диаграммообразующую схему (ДОС), имеющую m выходов. Выходные сигналы ДОС образуются суммированием сигналов элементов антенной решетки в соответствии с соотношением (5.44), что обеспечивает формирование m диаграмм направленности. Выходные сигналы ДОС поступают на входы каналов временной обработки, каждый из которых включает согласованный фильтр (СФ) и детектор. Определяя максимум выходного напряжения, по номеру канала и запаздыванию сигнала формируют оценки угловой координаты и дальности соответственно.

Отметим, что соотношение (5.44) идентично записанному в матричной форме выражению для анализа спектральных компонент в наблюдаемой реализации U . В то же время нелинейная обработка сигналов позволяет получить спектральные оценки, обладающие повышенным разрешением (по сравнению с оценками, получаемыми на основе преобразования Фурье). Аналогично, нелинейная обработка сигналов, снимаемых с выходов элементов антенной решетки, дает возможность сформировать оценки угловых координат с повышенной разрешающей способностью [45].

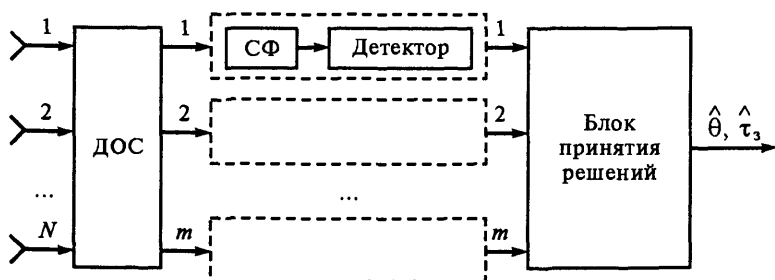


Рис. 5.16. Схема многоканального измерителя угловых координат

5.4.2. Одноканальные измерители угловых координат

Широкое практическое применение они нашли, в частности, в радиолокаторах с зеркальными антеннами. В них конструктивно жестко определено ожидаемое направление прихода отраженного сигнала. Как правило, оно совпадает с перпендикуляром к плоскости раскрыва антенны. Изменение положения антенны, осуществляемое, в основном, механически, позволяет оценить, насколько ожидаемое направление совпадает с направлением на цель, и измерить ее координаты.

Предположим, что РЛС измеряет угловые координаты цели в режиме обзора, причем в этом режиме луч антенны равномерно перемещается по угловым направлениям в зоне ответственности. Отраженный от цели сигнал (при импульсном методе локации) представляет пачку импульсов (рис. 5.17, а), огибающая $U(t)$ которой определяется формой диаграммы направленности антенны, а длительность — временем $T_{\text{обл}}$ облучения цели. Максимум огибающей будет наблюдаться в тот момент времени, когда антенна направлена на цель. Таким образом, полагая, что угловая скорость вращения антенны

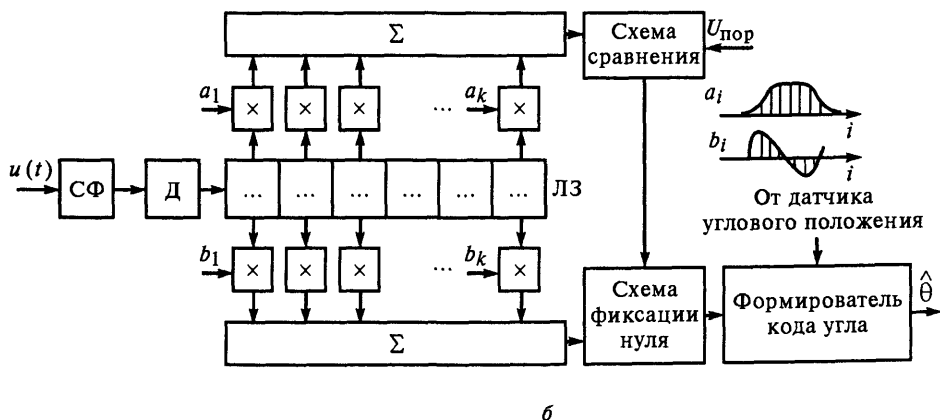
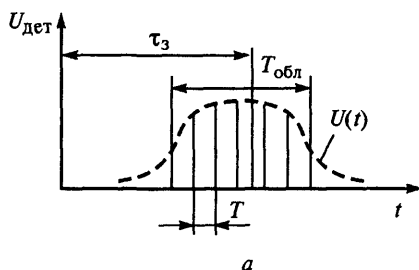


Рис. 5.17. Сигнал на выходе детектора (а) и структурная схема измерителя угловых координат (б)

постоянна, и фиксируя момент времени, когда огибающая пачки достигает максимума (в этой точке выполняется условие $dU(t)/dt = 0$), можно определить угловую координату цели. Упрощенная структурная схема такого измерителя угловой координаты представлена на рис. 5.17, б [43]. С выхода приемника на вход измерителя поступают импульсы, прошедшие согласованную обработку, которые запоминаются в регистре (линии задержки (ЛЗ)) на временном интервале, равном длительности пачки. Оценка производной обеспечивается суммированием отсчетов, снимаемых с выходов регистра, со специально подобранными весовыми коэффициентами b_i , пропорциональными производной огибающей. Момент времени, для которого производная огибающей равна нулю, определяется схемой фиксации. Чтобы исключить отсчет угловой координаты по шумовой выборке, угловая информация считывается только при наличии полезного сигнала. Решение о его наличии принимается в обнаружителе, где осуществляется накопление отсчетов с весовыми коэффициентами a_i , пропорциональными огибающей пачки, и сравнение накопленного сигнала с пороговым значением.

Ранее предполагалось, что сканирование пространства по угловым координатам осуществлялось путем механического изменения ориентации оси антенны. В РЛС с щелевыми антенными решетками можно организовать электронное сканирование путем изменения частоты (длины волны) несущего колебания, поскольку в соответствии с формулой (5.43) изменение частоты приводит к изменению положения основного лепестка диаграммы направленности. При изменении частоты по пилообразному закону с девиацией ΔF_m и периодом модуляции T_m , равном времени осмотра заданного углового сектора θ_c , положение основного лепестка диаграммы направленности антенны определяется текущей частотой $f(t)$ излучаемого сигнала. Таким образом, на интервале времени, равном периоду модуляции, несущую частоту сигнала можно записать следующим образом:

$$f(t) = f_n + kt, \quad 0 < t < T_m, \quad (5.45)$$

где f_n — значение частоты, соответствующее начальному положению луча антенны; k — коэффициент пропорциональности: $k = \Delta f_n / T_m$.

Отраженный от цели сигнал представляет собой радиоимпульс длительностью

$$\tau_n = T_m \left(\frac{\theta_a}{\theta_c} \right) \quad (5.46)$$

с девиацией частоты $\Delta F_n = \Delta F_m \frac{\theta_a}{\theta_c}$, средняя частота f_c которого зависит от угловой координаты θ_u :

$$f_c(\theta_{\text{ц}}) = f_n + \Delta F_m \frac{\theta_{\text{ц}}}{\theta_c}. \quad (5.47)$$

Здесь θ_a — ширина луча диаграммы направленности антенны.

Таким образом, в случае частотного сканирования измерение угловой координаты может быть сведено к измерению средней частоты сигнала. Схема такого измерителя включает набор фильтров, согласованных с импульсами длительности τ_n и девиацией частоты ΔF_n , отличающимися средними частотами. Номер фильтра с максимальной амплитудой выходного напряжения несет информацию об угловой координате цели.

5.4.3. Дискриминаторные методы измерения угловых координат

Простейшим дискриминаторным методом измерения углового рассогласования является коническое сканирование. Метод предполагает использование антенн с характеристикой направленности, имеющей круговую симметрию. При вращении облучателя, смещенного из фокуса антенны (рис. 5.18, а), обеспечивается вращение (сканирование) диаграммы направленности антенны по конической образующей с частотой $\Omega_{\text{ск}}$. Текущее угловое положение антенны определяется фазой управляющего сигнала, снимаемого с генератора опорного напряжения. Ось конической поверхности совпадает с равносигнальным направлением. Если цель находится на равносигнальном направлении, амплитуда принимаемого сигнала остается постоянной и не зависит от текущей ориентации диаграммы направленности (точка Ц₁, рис. 5.18, б). При небольших угловых отклонениях цели от равносигнального направления (точка Ц₂, рис. 5.18, а) огибающая принятого сигнала модулируется по гармоническому закону с частотой, равной частоте сканирования, причем глубина модуляции пропорциональна радиальному смещению цели в картинной плоскости относительно равносигнального направления, а фаза между модулирующим и опорным напряжениями характеризует направление смещения (рис. 5.18, а). Таким образом, в пеленгаторах, основанных на методе конического сканирования, при обработке должна быть выделена огибающая принимаемого сигнала, а сигналы ошибки по азимуту (α) и углу места (β) формируются после фильтрации произведений огибающей на опорные сигналы ($\sin(\Omega_{\text{ск}}t)$, $\cos(\Omega_{\text{ск}}t)$). К недостаткам РЛС с коническим сканированием, как и других одноканальных пеленгаторов, следует отнести ухудшение точности вследствие влияния амплитудных флуктуаций сигнала, отраженного от цели. Этот недостаток позволяют избежать моноимпульсные угловые дискриминаторы, которые получили в настоящее время широкое распространение.

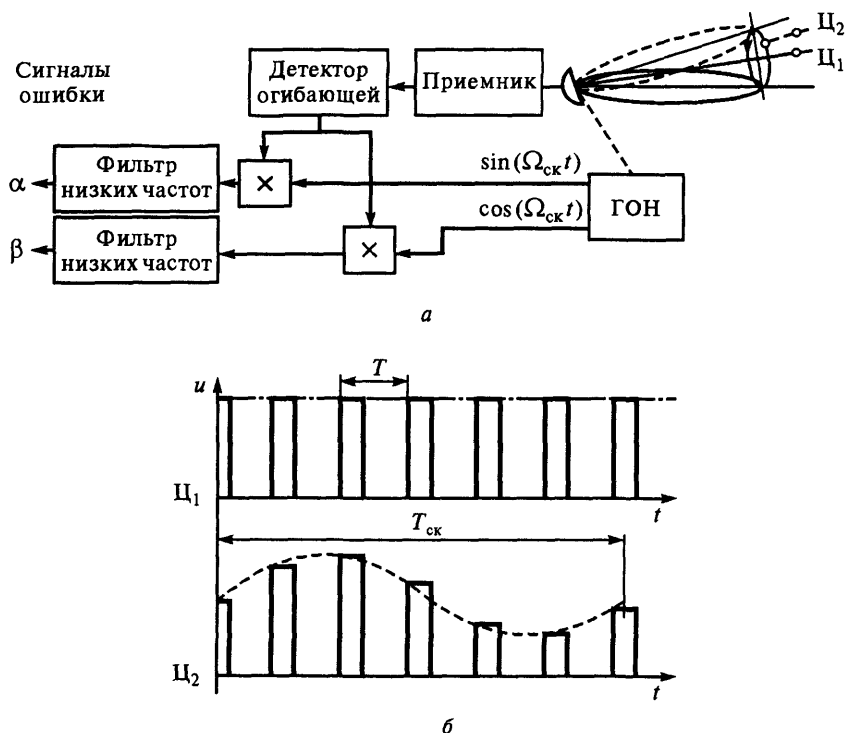


Рис. 5.18. Измерение угловых координат при коническом сканировании:

a — упрощенная структурная схема измерителя; b — сигналы на выходе детектора

Первоначально моноимпульсный метод был разработан для точного автоматического сопровождения целей. В настоящее время он применяется также и в обзорных радиолокационных системах. Моноимпульсный метод пеленгации основан на использовании отличий в амплитудно-фазовых характеристиках сигналов, принимаемых элементами антенны. Характер извлечения информации из принимаемых сигналов позволяет различать амплитудный и фазовый угловые моноимпульсные дискриминаторы.

Амплитудные дискриминаторы. Измерение угловых координат в амплитудных дискриминаторах основано на сравнении амплитуд колебаний, принимаемых парциальными лучами, формируемыми антенной системой. Парциальные диаграммы направленности имеют общий фазовый центр, а их максимумы смещены относительно равносигнального направления на угол $\Delta\theta/2$. Величина угла $\Delta\theta$ обычно выбирается таким образом, чтобы обеспечить максимальную точность измерения.

Простейшая схема моноимпульсного амплитудного дискриминатора, предназначенная для пеленгации цели в одной плоскости, содержит два идентичных приемных канала и устройство сравнения модульных значений корреляционных интегралов $Z(\theta_0 + \Delta\theta/2)$ и $Z(\theta_0 - \Delta\theta/2)$. В угловых дискриминаторах с амплитудным детектированием, как и в дискриминаторах, рассмотренных ранее, в качестве решающей статистики используют модуль корреляционного интеграла, а производная заменяется отношением

$$\frac{dZ}{d\theta} \approx \frac{Z(\theta_0 + \Delta\theta/2) - Z(\theta_0 - \Delta\theta/2)}{\Delta\theta}. \quad (5.48)$$

Здесь θ_0 — априорная оценка углового положения цели (равносигнальное направление).

Для исключения зависимости выходного сигнала пеленгатора от амплитуды полезного сигнала, изменяющегося, например, с изменением дальности до цели, можно вычислять не разность $Z(\theta_0 + \Delta\theta/2) - Z(\theta_0 - \Delta\theta/2)$, а отношение $Z(\theta_0 + \Delta\theta/2)/Z(\theta_0 - \Delta\theta/2)$, зависящее только от углового рассогласования между направлением на цель и равносигнальным направлением. Выходной сигнал дискриминатора можно также формировать в соответствии с выражением

$$\Delta_\theta = -\frac{Z'(\theta_0)}{Z''(\theta_0)} \approx -\frac{Z(\theta_0 + \Delta\theta/2) - Z(\theta_0 - \Delta\theta/2)}{Z(\theta_0)}. \quad (5.49)$$

Операция деления может выполняться с помощью блока автоматической регулировки усиления. Отметим, что ее можно заменить операцией вычитания логарифмов амплитуд сигналов, для чего в тракт обработки включаются логарифмические усилители (рис. 5.19). Моноимпульсные пеленгаторы — сложные многоканальные устройства, причем для их функ-

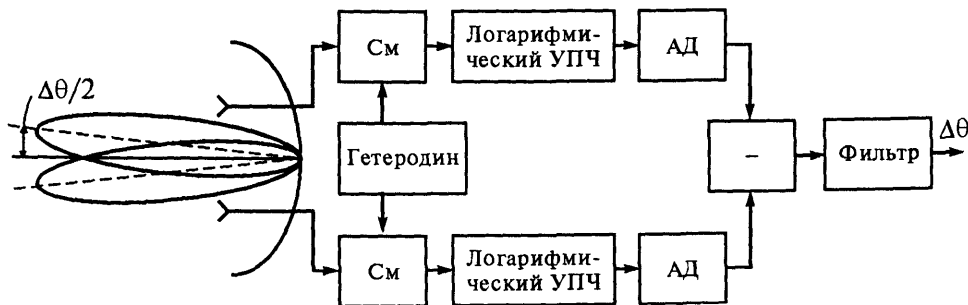


Рис. 5.19. Простейшая схема амплитудного моноимпульсного дискриминатора

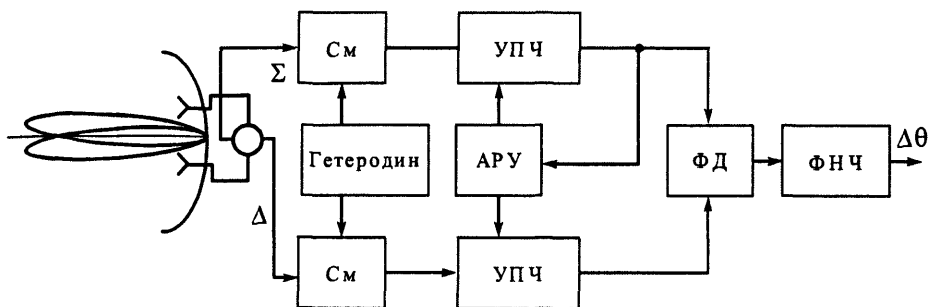


Рис. 5.20. Упрощенная структурная схема амплитудного суммарно-разностного углового дискриминатора

ционирования амплитудно-частотные характеристики различных каналов должны быть одинаковыми. Наличие аппаратных ошибок приводит к появлению систематических ошибок измерения. В частности, различия в коэффициентах усиления каналов пеленгатора (рис. 5.19) вызывают смещение равносигнального направления.

В целях ослабления влияния неидентичности приемных каналов на качество работы дискриминатора используют дискриминаторы с суммарно-разностной обработкой. Здесь в антенно-волноводном тракте формируются суммарный и разностный сигналы, а угловое рассогласование оценивается как

$$\Delta_{\theta} \approx - \frac{Z(\theta_0 + \Delta\theta/2) - Z(\theta_0 - \Delta\theta/2)}{Z(\theta_0 + \Delta\theta/2) + Z(\theta_0 - \Delta\theta/2)}. \quad (5.50)$$

Операция нормировки (деления сигналов разностного канала на сигнал суммарного канала) осуществляется с помощью системы автоматической регулировки усиления, работающей по суммарному сигналу. Пример структурной схемы такого дискриминатора представлен на рис. 5.20.

Фазовые угловые дискриминаторы. Фазовые методы измерения угловых координат основаны на сравнении фаз колебаний, принятых несколькими разнесенными в пространстве антеннами. Различия в амплитудных характеристиках направленности при этом не используются.

В случае приема сигнала двумя антеннами ($M = 2$), разнесенными на базу d (рис. 5.21), уравнения, описывающие измеритель с суммарно-разностной обработкой, принимают вид

$$\Delta_{\theta} = 2 \operatorname{Re}(-j z_c z_p^*) / z_c^2, \quad (5.51)$$

где z_c, z_p — сигналы на выходах фильтров (УПЧ) суммарного и разностного каналов. Множитель j учитывается введением в схему обработки (рис. 5.21)

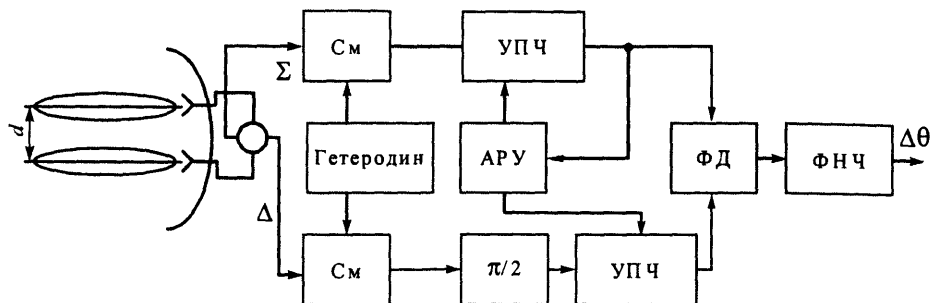


Рис. 5.21. Упрощенная структурная схема фазового суммарно-разностного углового дискриминатора

фазовращателя, обеспечивающего сдвиг фазы в одном из каналов на $\pi/2$. Вычисление реальной части произведения двух комплексно-сопряженных амплитуд осуществляется в фазовом детекторе. Операция деления обеспечивается введением автоматической регулировки усиления (АРУ) по суммарному каналу и использованием ее для регулировки усиления в канале разностного сигнала. Суммарно-разностная обработка находит широкое применение, поскольку позволяет существенно ослабить влияние фазовых неидентичностей каналов по сравнению со случаем, когда усиление и фильтрация сигналов в каналах осуществляется независимо.

5.5. Точность измерения параметров

В этом параграфе рассмотрим вопросы, связанные с потенциальной точностью измерения параметров сигналов. Будем предполагать, что измерения регулярны. Для регулярного измерения, характеризующегося тем, что отношение сигнал—шум велико (более 8...10), а функция $w(\mathbf{u}|\lambda)$ унимодальна и дважды дифференцируема, ошибки измерения можно считать распределенными по нормальному закону с нулевым математическим ожиданием. В этом случае при изменении одиночного параметра они полностью характеризуются дисперсией

$$\sigma^2(\hat{\lambda}(\mathbf{u}) - \lambda) = \mathbf{M}\{(\hat{\lambda}(\mathbf{u}) - \lambda)^2\},$$

а при оценке многомерных параметров — ковариационной матрицей ошибок измерения. Для границ дисперсии ошибок измерения можно использовать величины, полученные на основе неравенства Крамера—Рао:

$$\sigma^2(\hat{\lambda}(\mathbf{u}) - \lambda) \geq -\mathbf{M}\{[d^2 \ln w(\mathbf{u}|\lambda)/d\lambda^2]^{-1}\}.$$

Оценка называется *эффективной*, если указанное соотношение выполнено со знаком равенства.

Определение нижних границ дисперсии оценок удобно проводить в несколько этапов. При оценивании неслучайных параметров на первом этапе рассчитывается информационная матрица Фишера \mathbf{J} с элементами

$$j_{il} = -\mathbf{M} \left\{ \frac{d^2 \ln w(\mathbf{u}|\boldsymbol{\lambda})}{d\lambda_i d\lambda_l} \right\}. \quad (5.52)$$

Здесь $i, l \in 1, 2, \dots, m$, m — размерность вектора $\boldsymbol{\lambda}$ оцениваемых параметров.

На втором этапе определяются нижние границы ковариационной матрицы ошибок измерения \mathbf{D}_ε :

$$\mathbf{D}_\varepsilon = \mathbf{J}^{-1}. \quad (5.53)$$

При определении границ ошибок оценивания случайных параметров результирующая информационная матрица \mathbf{J}_R состоит из двух частей:

$$\mathbf{J}_R = \mathbf{J} + \mathbf{J}_a,$$

где элементы матрицы \mathbf{J} определяются выражением (5.52), а элементы матрицы \mathbf{J}_a , характеризующие априорную информацию, имеют вид

$$j_{a\,il} = -\mathbf{M} \left\{ \frac{d^2 \ln w(\boldsymbol{\lambda})}{d\lambda_i d\lambda_l} \right\}. \quad (5.54)$$

Нижние границы оценок ковариационной матрицы ошибок измерения \mathbf{D}_R случайных параметров определяются формулой

$$\mathbf{D}_R = \mathbf{J}_R^{-1}. \quad (5.55)$$

Диагональные элементы матриц \mathbf{J}^{-1} и \mathbf{J}_R^{-1} представляют оценки нижних границ дисперсий ошибок измерения соответствующих величин для неслучайных и случайных параметров сигнала соответственно, недиагональные элементы — оценки границ взаимно корреляционные функции ошибок.

Среднеквадратическая ошибка измерения времени запаздывания. Предположим, что измерение дальности (запаздывания) осуществляется устройством, структурная схема которого представлена на рис. 5.3. Приемник вычисляет модульное значение корреляционного интеграла, а в качестве оценки принимается то значение времени, для которого выходной сигнал максимален. При фиксированном отношении сигнал—шум среднее значение модуля корреляционного интеграла изменяется следующим образом:

$$Z(\tau, F) = q\rho(\tau, F),$$

где q — отношение сигнал—шум, $\rho(\tau, F)$ — функция рассогласования, τ, F — рассогласование по запаздыванию и частоте соответственно.

Учитывая, что в данном случае измеряется один параметр, и проводя дифференцирование по формуле (5.52), получим, что выражение для среднеквадратической ошибки измерения времени запаздывания при известной доплеровской частоте сигнала имеет вид

$$\sigma_\tau = \frac{1}{q\sqrt{|\rho''_{\tau\tau}(0, 0)|}}, \quad (5.56)$$

где $\rho''_{\tau\tau}(0, 0)$ — вторая производная $\rho(\tau, F)$ по τ при $\tau = 0, F = 0$, характеризующая остроту пика функции рассогласования. Ошибка тем меньше, чем больше отношение сигнал—шум и острее пик (больше вторая производная по τ) функции $\rho(\tau, 0)$.

Величина $\sqrt{|\rho''_{\tau\tau}(0, 0)|}$ имеет размерность частоты (Гц), и ей соответствует некоторая эффективная ширина спектра сигнала $\Delta f_{\text{эф}}$, связанная со спектральной плотностью комплексной амплитуды сигнала $G(F)$ соотношением

$$\Delta f_{\text{эф}} = \sqrt{|\rho''_{\tau\tau}(0, 0)|} = 2\pi \sqrt{\frac{\int_{-\infty}^{\infty} F^2 |G(F)|^2 dF}{\int_{-\infty}^{\infty} |G(F)|^2 dF}}. \quad (5.57)$$

В этом случае

$$\sigma_\tau = \frac{1}{q\Delta f_{\text{эф}}}. \quad (5.58)$$

Таким образом, среднеквадратическая ошибка измерения времени запаздывания тем меньше, чем больше эффективная ширина спектра сигнала $\Delta f_{\text{эф}}$ и выше отношение сигнал—шум.

Среднеквадратическая ошибка измерения частоты. В соответствии с формулой (5.52) среднеквадратическая ошибка измерения доплеровской частоты сигнала при известном времени запаздывания равна

$$\sigma_F = \frac{1}{q\sqrt{|\rho''_{FF}(0, 0)|}}, \quad (5.59)$$

где $\rho''_{FF}(0, 0)$ — вторая производная $\rho(\tau, F)$ по F при $\tau = 0, F = 0$, характеризующая остроту пика функции рассогласования по оси частот.

Величина $\sqrt{|\rho''_{FF}(0, 0)|}$ имеет размерность времени. Ей соответствует параметр $\tau_{\text{эф}}$, характеризующий эффективную длительность сигнала:

$$\tau_{\text{эф}} = 2\pi \sqrt{\frac{\int_{-\infty}^{\infty} t^2 |U(t)|^2 dt}{\int_{-\infty}^{\infty} |U(t)|^2 dt}}.$$

С учетом этих обозначений

$$\sigma_F = \frac{1}{q\tau_{\text{эф}}}. \quad (5.60)$$

Таким образом, среднеквадратическая ошибка измерения частоты тем меньше, чем больше отношение сигнал—шум и эффективная длительность сигнала.

Среднеквадратическая ошибка измерения угловой координаты. Аналогично выражениям (5.56), (5.59) среднеквадратическая ошибка измерения угловой координаты определяется формулой

$$\sigma_{\theta} = \frac{1}{q\sqrt{|\rho''_{\theta\theta}(\theta_0, \theta)|_{\theta=\theta_0}}}, \quad (5.61)$$

где $\rho''_{\theta\theta}(\theta_0, \theta)_{\theta=\theta_0}$ — вторая производная пространственной функции рассогласования по θ при $\theta = \theta_0$, которой можно дать определенное истолкование: корень квадратный из этой величины определяет эффективную длину апертуры антенны $L_{\text{эф}}$, нормированную по отношению к длине волны. Эффективная длина апертуры определяется следующим образом:

$$L_{\text{эф}} = \frac{\int_{-\infty}^{\infty} x^2 |A(x)|^2 dx}{\int_{-\infty}^{\infty} |A(x)|^2 dx}, \quad (5.62)$$

где $A(x)$ — функция распределения поля в раскрыве антенны. Отсюда выражение для среднеквадратической ошибки можно представить в виде

$$\sigma_{\theta} = \frac{\Delta\theta}{q}, \quad (5.63)$$

где $\Delta\theta = \lambda/(2L_{\text{эф}})$ — параметр, характеризующий разрешающую способность РЛС по угловой координате. Таким образом, среднеквадратическая ошибка измерения угловой координаты тем меньше, чем больше эффективная длина апертуры антенны и выше отношение сигнал—шум.

Потенциальная точность совместного измерения времени запаздывания и частоты колебаний. При одновременном измерении времени запаздывания и доплеровского сдвига частоты и регулярном измерении вектор

ошибок $\varepsilon = \begin{bmatrix} \tau_3 - \hat{\tau}_3 \\ F_{\text{Д}} - \hat{F}_{\text{Д}} \end{bmatrix}$ имеет нормальное распределение с нулевым математическим ожиданием и ковариационной матрицей \mathbf{D}_{ε} .

Информационная матрица является двумерной:

$$\mathbf{J} = \begin{bmatrix} -\mathbf{M} \left\{ \frac{\partial^2 \ln l(\tau, F)}{\partial \tau^2} \right\} & -\mathbf{M} \left\{ \frac{\partial^2 \ln l(\tau, F)}{\partial \tau \partial F} \right\} \\ -\mathbf{M} \left\{ \frac{\partial^2 \ln l(\tau, F)}{\partial \tau \partial F} \right\} & -\mathbf{M} \left\{ \frac{\partial^2 \ln l(\tau, F)}{\partial F^2} \right\} \end{bmatrix}, \quad (5.64)$$

а ковариационная матрица ошибок измерения определяется соотношением (5.55). Вычисляя входящие в это выражение производные функции $\ln l(\tau, F)$ и производя обращение матрицы, получим следующие выражения для среднеквадратических значений ошибок оценивания запаздывания и доплеровского сдвига частоты:

$$\sigma_{\tau} = \frac{1}{q \sqrt{|\rho''_{\tau\tau}(0, 0)|(1 - r^2)}}, \quad (5.65)$$

$$\sigma_F = \frac{1}{q \sqrt{|\rho''_{FF}(0, 0)|(1 - r^2)}}. \quad (5.66)$$

Здесь r — коэффициент корреляции измерений, определяемый формулой

$$r^2 = \frac{|\rho''_{\tau F}(0, 0)|^2}{(\rho''_{FF}(0, 0))(\rho''_{\tau\tau}(0, 0))},$$

где $\rho''_{\tau F}(0, 0)$ — вторая производная функции $\rho(\tau, F)$ по τ и F при $\tau = 0$, $F = 0$.

Множитель $\frac{1}{1-r^2}$ в выражениях (5.65), (5.66) отражает влияние корреляционной связи между измерениями дальности и доплеровского сдвига частоты. Если $r > 0$, то точность измерений уменьшается; если $r = 0$, то точность измерения максимальна, что имеет место только в том случае, когда функция рассогласования сигнала симметрична относительно плоскостей $\tau = 0$ и $F = 0$. Такую функцию рассогласования имеют, например, простые импульсные сигналы, сигналы в виде когерентных импульсных последовательностей и ФКМ сигналы. Для простого импульсного сигнала с гауссовой огибающей ошибки измерения времени запаздывания и частоты (скорости) независимы ($r = 0$), и их среднеквадратические значения определяются выражениями [34, 40]:

$$\sigma_{\tau} = \frac{\tau_n}{\sqrt{\pi q}}, \quad (5.67)$$

$$\sigma_F = \frac{2}{\sqrt{\pi q \tau_n}}. \quad (5.68)$$

Здесь τ_n — длительность импульса по уровню 0,46.

Из формул (5.67), (5.68) видно, что повысить точность измерения времени запаздывания можно, уменьшая длительность импульса, однако при этом ухудшается точность измерения частоты.

Рассмотрим точность измерений запаздывания и частоты для импульсного сигнала с гауссовой огибающей и линейной частотной модуляцией. В этом случае [37,40]

$$\sigma_{\tau} = \frac{\tau_n}{\sqrt{\pi q}}, \quad (5.69)$$

$$\sigma_F = \frac{B}{\sqrt{\pi q \tau_n}}, \quad (5.70)$$

где B — база сигнала. Из выражений (5.69), (5.70) видно, что при измерении двух параметров введение частотной модуляции не приводит к повышению точности измерения времени запаздывания, а уменьшает точность измерения частоты. Это связано с тем, что, как отмечалось ранее, площадь области неопределенности для немодулированного и модулированного сигналов одинаковы. Использование модуляции приводит к повороту области неопределенности. При этом проекция области высокой корреляции на ось частот и, следовательно, диапазон возможных ошибок измерения частоты возрастают. Таким образом, введение частотной модуляции не позволяет одно-

временно увеличить точность измерений запаздывания и доплеровского сдвига частоты. В то же время если рассматривать задачу измерения запаздывания при известном значении доплеровского сдвига частоты, то введение частотной модуляции позволяет уменьшить среднеквадратическую ошибку измерения запаздывания в B раз:

$$\sigma_{\tau} = \frac{\tau_{\text{и}}}{\sqrt{\pi q B}}. \quad (5.71)$$

Таким образом, если необходимо обеспечить точное измерение времени запаздывания и доплеровского сдвига частоты, то целесообразно использовать другие виды сигналов, например когерентные импульсные последовательности либо фазокодированные сигналы. Однако следует учитывать, что функция неопределенности этих сигналов содержит дополнительные выбросы, что увеличивает вероятность «аномальных» ошибок [37].

Контрольные вопросы

1. Перечислите основные параметры радиолокационных сигналов, которые оцениваются при измерении.
2. Перечислите показатели качества измерения одномерной случайной величины.
3. Каким образом определяются усредненные потери?
4. Какие функции потерь используются при решении задачи оптимального измерения параметров радиолокационного сигнала?
5. Каким образом определяется оптимальная оценка измеряемого параметра радиолокационного сигнала при квадратичной функции потерь?
6. Каким образом производится оценка параметров по максимуму апостериорной плотности вероятности, по максимуму правдоподобия?
7. Каковы особенности многократного измерения параметров радиолокационного сигнала?
8. Какие измерители параметров радиолокационных сигналов называют несledящими, какие следящими?
9. Каким образом радиолокационный приемник вычисляет функцию, пропорциональную апостериорной плотности вероятности (АПВ) или монотонную функцию от АПВ в заданном диапазоне изменения параметров?
10. Каким образом производится оптимальное несledящее измерение времени запаздывания (дальности)?
11. Каким образом производится оптимальное несledящее измерение доплеровской частоты (радиальной скорости)?
12. Приведите структурную схему измерения времени запаздывания и доплеровской частоты.
13. Приведите примеры практической реализации временных дискриминаторов.

14. Приведите пример практической реализации частотного дискриминатора.
15. Приведите примеры одноканальных измерителей угловых координат.
16. Изобразите простейшую схему моноимпульсного амплитудного дискриминатора с суммарно-разностной обработкой.
17. Изобразите схему моноимпульсного фазового дискриминатора с суммарно-разностной обработкой.
18. Приведите и проанализируйте формулу для среднеквадратической ошибки измерения времени запаздывания.
19. Приведите и проанализируйте формулу для среднеквадратической ошибки измерения частоты.
20. Приведите и проанализируйте формулу для среднеквадратической ошибки измерения угловой координаты.
21. От каких факторов зависит потенциальная точность совместного измерения времени запаздывания и доплеровской частоты радиолокационного сигнала?

6. ОСНОВЫ ВТОРИЧНОЙ ОБРАБОТКИ РАДИОЛОКАЦИОННОЙ ИНФОРМАЦИИ

Изложены задачи вторичной обработки (ВО) информации применительно к их реализации в радиолокационных системах. Приведена обобщенная структурная схема ВО. Рассмотрена реализация основных операций ВО и их особенностей. Основное внимание уделено базовым алгоритмам ВО и их физической сущности.

6.1. Общие сведения

6.1.1. Основные понятия и история вопроса

Вторичная обработка (ВО) радиолокационной информации, или, как еще иногда говорят, траекторная обработка, выполняется после первичной и применяется для решения задачи обнаружения траекторий целей в зоне ответственности РЛС и оценки их параметров. Траектория цели представляет собой след от перемещающегося с течением времени в некотором пространстве объекта наблюдения. Такое понимание траектории вытекает из естественных представлений, возникающих при наблюдении различных движущихся объектов.

Абстрагируясь от физической природы движущегося объекта, траекторию цели можно определить как линию — след от перемещения некоторой математической точки, соответствующей данному объекту. Заметим, что наиболее часто под этой точкой понимают центр масс соответствующего физического тела.

В каждый момент времени траектория цели может быть представлена некоторыми параметрами (фазовыми координатами): положением цели, ее скоростью, ускорением и т. п. Траектория цели считается известной, если для заданного интервала времени известна зависимость ее фазовых координат от времени.

Движение объекта можно рассматривать как некоторый (чаще всего, случайный) процесс, а траекторию объекта, характеризующую соответст-

вующими фазовыми координатами в течение некоторого интервала времени, — как представление этого процесса.

В течение такта первичной обработки радиолокационной информации по обнаруженным сигналам, порожденным либо движущимися объектами, либо шумом, в системе координат первичных наблюдений формируется совокупность случайных отсчетов, характеризующих обстановку в зоне контроля РЛС в момент зондирования. По результатам многократных первичных наблюдений в ходе вторичной обработки необходимо принимать решения об обнаружении целевых траекторий и оценивать их параметры. Появление шумовых отсчетов, исчезновение в некоторых тактах целевых отсчетов и погрешности измерений являются факторами, влияющими на качество проведения траекторной обработки.

Формально задача и первичной, и вторичной обработки радиолокационной информации одна и та же: оценка целевой ситуации* в контролируемой радиолокатором зоне. Необходимость ВО вызвана тем, что при малых временных интервалах наблюдения, которые характерны для принятия решения в ходе одного такта первичной обработки, надежность обнаружения и точность оценки координат цели, а часто и их состав** оказываются недостаточными для нужд потребителя радиолокационной информации. В ходе ВО при увеличении времени наблюдения, отводимого для принятия решения, появляется возможность повысить качество принимаемых решений.

Несмотря на внешнюю схожесть задач первичной и вторичной обработки, условия, в которых решаются эти задачи, и методы их решения существенно различаются.

При вторичной обработке учитывается, что цели могут быть перемещающимися, появляющимися и исчезающими в зоне контроля РЛС. При первичной обработке рассматривается только двухальтернативная ситуация: цель либо есть в элементе разрешения, либо ее там нет. Если в ходе первичной обработки радиолокационной информации параметры объекта наблюдения являются фактически неизменными, то в ходе вторичной они таковыми, естественно, быть не могут. Это связано с тем, что такт первичной обработки T_1 , определяемый временем обработки зондирующего сигнала, имеет, как правило, столь малые значения, что цель за время T_1 практически не изменяет своего положения в пространстве.

* Под целевой ситуацией понимаются цели, находящиеся в зоне контроля РЛС, их взаимные характеристики, вид целей, параметры их движения и т. п.

** Например, в ходе ВО, наблюдая за перемещениями цели, можно найти ее скорость, ускорение, даже если на этапе первичной обработки измеряется только положение объекта. Возможно определение и других принципиально ненаблюдаемых параметров цели.

Вторичная обработка выполняется в течение всего времени наблюдения за целью, которое, как правило, много больше T_I . Результаты ВО обычно относятся к текущему времени t и, вообще говоря, учитывают всю информацию о цели: с момента ее появления в зоне контроля до момента времени t . Иногда в рассмотрение вводят такт вторичной обработки T_{II} — время, в течение которого корректируется предыдущее решение об обнаружении траектории и оценивании ее фазовых координат — параметров траектории цели. Такт T_{II} может задаваться либо исходя из требований вышестоящей системы, либо из естественного цикла обновления информации, связанного с особенностями обзора контролируемого пространства (например, для РЛС кругового обзора такт T_{II} часто делают равным периоду вращения антенны).

Можно, таким образом, сказать, что на выходе первичной обработки получена совокупность отсчетов, каждый из которых является *случайной векторной величиной*, представляющей собой оценку параметров неизвестных случайных или неслучайных координат некоторой обнаруженной цели. На выходе вторичной обработки получаем совокупность траекторий, каждая из которых является *случайным векторным процессом*, представляющим собой динамическую оценку параметров неизвестных изменяющихся во времени случайных или неслучайных координат некоторой обнаруженной цели.

Естественно, что траекторная информация, как уже отмечалось, полнее отметочной, поскольку учитывает в текущей оценке предысторию наблюдения цели, включая все предыдущие отсчеты, и отображает взаимосвязь и изменение параметров цели во времени.

Вторичная обработка радиолокационной информации, основанная так же, как и первичная, на общей теории принятия решений и теории оценок, является специфической задачей многомерной фильтрации, дополненной необходимостью выбора (обнаружения) целевых траекторий. В ходе ВО решается задача отождествления отсчетов и траекторий, соответствующих одной и той же цели. В частном случае, при достоверной информации первичной обработки о получении отсчетов, относящихся исключительно к некоторой истинной цели, ВО сводится только к фильтрации траекторных параметров (это характерно для РЛС, обеспечивающих высокое значение отношения сигнал—шум при принятии решений в ходе первичной обработки).

Собственно фильтрация — это непрерывное воспроизведение некоторой переменной, являющейся параметром наблюдаемого случайного процесса. Можно выделить два основных подхода к решению задач фильтрации случайных процессов: на основе фильтра Винера и фильтра Калмана. В последнее время замечен повышенный интерес к развитию методов нелинейной фильтрации.

Теория винеровской фильтрации исторически была разработана первой. К одной из основополагающих работ в этой области следует отнести работу Н. Винера [50], для дискретных случайных процессов близкие вопросы еще ранее рассматривал А. Н. Колмогоров [51]. В основе теории винеровской фильтрации лежат следующие положения:

1) наблюдаемый случайный процесс есть аддитивная смесь оцениваемого и помехового случайных стационарных процессов с различными корреляционными функциями;

2) длительность наблюдения предполагается бесконечно большой (от $-\infty$ до текущего момента t);

3) критерием оптимальной фильтрации является минимум среднеквадратической ошибки оценки воспроизводимого параметра;

4) оптимальный фильтр находится в классе линейных фильтров.

При построении траекторий движения целей условия, для которых были получены выражения, описывающие винеровский фильтр, оказываются во многих случаях не адекватными реальной ситуации. Это связано, прежде всего, с тем, что, во-первых, оцениваемый случайный процесс (траекторию) невозможно представить в виде стационарного процесса с известной корреляционной функцией, во-вторых, появляющиеся и исчезающие объекты всегда наблюдаются в течение ограниченного времени.

Теория калмановской фильтрации при построении траекторий имеет большую свободу: непрерывный фильтр Калмана основан на представлении случайного процесса в виде стохастических дифференциальных уравнений, а цифровой — в виде соответствующей системы разностных уравнений, что соответствует широко распространенным общепринятым моделям движения объектов. Теория калмановской фильтрации была разработана позже винеровской и является существенным ее развитием. С использованием калмановского подхода можно решать нестационарные задачи при конечном интервале наблюдения. Основополагающими в этой области следует считать работы Р. Калмана [52, 53].

Вторичная обработка радиолокационной информации, как уже отмечалось, помимо фильтрации траекторных параметров, должна решать и задачи обнаружения целевых траекторий. Действительно, при наличии помеховых отсчетов по ним могут быть построены трассы несуществующих (ложных) целей. И это не единственная причина возможного появления ложных траекторий.

В многоцелевой ситуации даже при достоверном решении об обнаружении отсчета в некоторый момент времени на этапе первичной обработки в силу движения объекта отсчет от той же цели в следующий момент времени будет иметь другие параметры. Поэтому прежде чем производить фильтрацию траекторных параметров, необходимо убедиться в том, что вновь по-

ступивший отсчет относится именно к рассматриваемой траектории, а не к какой-либо другой. Вследствие действия помех, ограниченной точности измерений на этапе первичной обработки и конечной разрешающей способности по измеряемым координатам возможно перепутывание отсчетов и траекторий различных целей. Следовательно, при проведении ВО фильтрация (т. е. собственно оценивание траекторных параметров) должна дополняться операцией отождествления, обеспечивающей непоступление на ее вход «чужих» отсчетов.

При дальнейшем усложнении целевой и помеховой обстановки полезные (от целей) и помеховые отсчеты могут породить ложные трассы, среди которых принять достоверное решение о целевых траекториях возможно лишь при достаточно длительном наблюдении.

Естественно, что задачи обнаружения траектории и оценки ее параметров должны решаться совместно, поскольку в ходе выполнения одной из них используются результаты другой. Это учитывается во всех алгоритмах ВО, включая простейшие.

Началом разработки математического аппарата для оптимального решения всей совокупности задач ВО можно считать работы Р. Л. Стратоновича, который рассматривал проблемы общего описания случайных потоков и рекуррентные алгоритмы нахождения апостериорных характеристик марковских случайных процессов [54, 55].

В настоящее время существует большое количество публикаций по вопросам решения задач обнаружения траекторий и оценивания их параметров.

Детальный анализ отдельных операций и проблем особо сложных случаев траекторной обработки при сопровождении маневрирующих целей, построении и обнаружении траекторий для интенсивных потоков ложных сигналов и плотных потоков целей, а также для целей, движущихся с пересекающимися траекториями, можно найти в работах [56–67], в которых также приводится и обобщается большое число публикаций по многим другим вопросам ВО.

Следует выделить фундаментальную книгу И. А. Большакова [68], давшую толчок в развитии методов обнаружения—оценивания случайных потоков, работу Ю. С. Ачкасова [56], рассмотревшего вопросы обнаружения траекторий с использованием апостериорного анализа потоков, и, наконец, книгу под редакцией П. А. Бакута [57], в которой детально проанализированы многие вопросы эволюции апостериорных плотностей распределения вероятностей при наблюдении движущихся целей и методы их обнаружения. Некоторые результаты решения задач совместного обнаружения—оценивания траекторий на основе обобщения подходов И. А. Большакова и П. А. Бакута приводятся в статьях [58–60, 66].

В настоящей главе рассматриваются основные положения теории ВО радиолокационной информации. Решение задач, относящихся к сложной целевой и помеховой ситуации, лишь намечается. Основное внимание уделяется физической сущности рассматриваемых операций и базовым алгоритмам ВО.

6.1.2. Основные операции вторичной обработки

В ходе ВО радиолокационной информации производится обнаружение траекторий и оценивание их параметров по результатам наблюдений, выполненным на этапе первичной обработки.

Выходной информацией ВО являются рассчитанные на некоторый момент времени $t_{\text{пр}}$ (момент привязки измерений) оценки параметров траекторий $\hat{X}_j(t_{\text{пр}})$, $j=1, 2, \dots, N$ (j — номер траектории, N — число обнаруженных траекторий), показатели качества их обнаружения: условная вероятность правильного обнаружения траектории $D_{\text{тр}j}(t_{\text{пр}})$ и условная вероятность ложного обнаружения траектории $F_{\text{тр}j}(t_{\text{пр}})$, — а также точность оценки траекторных параметров, характеризуемая соответствующей ковариационной* матрицей $\hat{\Psi}_j(t_{\text{пр}})$ ошибок измерения траекторных параметров (фазовых координат траектории). Обычно траекторные параметры рассчитываются на момент последнего измерения $t_{\text{пр}} = t_k$.

Показатели качества ВО зависят от длительности наблюдения. Для определенности всегда необходимо указывать либо заданное время $t_{\text{об.тр}}$ обнаружения траектории, либо время $t_{\text{оц.тр}}$ оценивания траекторных параметров. Используются и другие показатели качества траекторной обработки [57, 67]. Векторный характер показателей качества объективно затрудняет сравнительный анализ различных алгоритмов траекторной обработки.

* Под ковариационной матрицей Ψ некоторого случайного вектора X (состоящего из случайных элементов) будем понимать среднее:

$$\Psi \equiv M\{(X - M\{X\})(X - M\{X\})^T\} = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} w(X)(X - M\{X\})(X - M\{X\})^T dX,$$

где $M\{\cdot\}$ — оператор усреднения, знак «Т» означает транспонирование матрицы, $w(X)$ — многомерная плотность распределения элементов случайного вектора X .

Входной информацией ВО являются параметры $Z_i(t_k)$ отсчетов, обнаруженных в некоторый момент времени t_k на этапе первичной обработки, где $i = 1, 2, \dots, m$ (i — номер отсчета, m — число обнаруженных отсчетов). Их достоверность характеризуется условными вероятностями правильного обнаружения отсчета $D_{от i}(t_k)$ и ложной тревоги $F_{от i}(t_k)$, а точность оценки — соответствующей ковариационной матрицей $R_i(t_k)$ ошибок оценивания параметров, измеряемых на этапе первичной обработки.

Будем говорить, что вектор $Z_i(t_k)$ определен в пространстве наблюдений, или измерений, а вектор $\hat{X}_j(t_{пр})$ — в пространстве оценок*.

Учитывая дискретность радиолокационных измерений во времени, операции обработки радиолокационной информации чаще рассматривают в дискретном времени. Временную привязку всех параметров в этом случае обозначают соответствующим индексом: $Z_i(t_k) \equiv Z_{ik}$, $X_i(t_k) \equiv X_{ik}$, $R_i(t_k) \equiv R_{ik}$ и т. д.

Схематически входные и выходные параметры ВО показаны на рис. 6.1.

Параметры отсчета Z (дополнительные индексы будем опускать в случаях, когда это не вызывает неоднозначного толкования) образуют вектор (матрицу-столбец размерности a), элементы которого — это первичные измерения. Например, если на этапе первичной обработки определены дальность до цели r , азимут цели β , угол места ε и радиальная скорость цели v_r , т. е. $a = 4$, то

$$Z = [r \ \beta \ \varepsilon \ v_r]^T. \quad (6.1)$$

Ковариационная матрица R ошибок оценивания первичных измерений в этом случае имеет вид

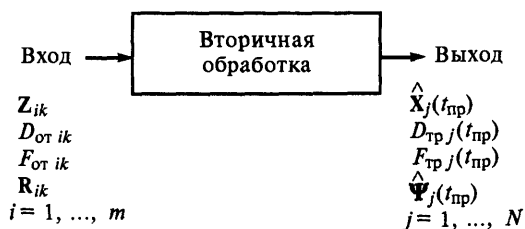


Рис. 6.1. Входные и выходные параметры алгоритма траекторной обработки

* Строго говоря, вектор $Z_i(t_k)$ является оценкой параметров отсчета (см. гл. 5), однако для ВО эта оценка является уже состоявшейся. После окончания первичной обработки полученный отсчет для ВО является входной информацией, т. е. наблюдением. Далее всегда будем считать параметры отсчета наблюдениями, или измерениями, а получаемые в ходе ВО параметры траектории — оценками.

$$\mathbf{R} = \begin{bmatrix} \sigma_r^2 & k_{r\beta}\sigma_r\sigma_\beta & k_{r\epsilon}\sigma_r\sigma_\epsilon & k_{rv_r}\sigma_r\sigma_{v_r} \\ k_{r\beta}\sigma_r\sigma_\beta & \sigma_\beta^2 & k_{\beta\epsilon}\sigma_\beta\sigma_\epsilon & k_{\beta v_r}\sigma_\beta\sigma_{v_r} \\ k_{r\epsilon}\sigma_r\sigma_\epsilon & k_{\beta\epsilon}\sigma_\beta\sigma_\epsilon & \sigma_\epsilon^2 & k_{\epsilon v_r}\sigma_\epsilon\sigma_{v_r} \\ k_{rv_r}\sigma_r\sigma_{v_r} & k_{\beta v_r}\sigma_\beta\sigma_{v_r} & k_{\epsilon v_r}\sigma_\epsilon\sigma_{v_r} & \sigma_{v_r}^2 \end{bmatrix}, \quad (6.2)$$

где σ_r^2 , σ_β^2 , σ_ϵ^2 , $\sigma_{v_r}^2$ — дисперсии ошибок измерения соответствующих параметров отсчета \mathbf{Z} ; $k_{\epsilon\beta}$, $k_{r\epsilon}$, $k_{r\beta}$, $k_{\beta\epsilon}$, $k_{\beta v_r}$, $k_{\epsilon v_r}$ — коэффициенты корреляции соответствующих параметров отсчета \mathbf{Z} .

Когда измерения отдельных параметров в РЛС при первичной обработке информации не коррелированы между собой, матрица \mathbf{R} становится диагональной.

Траекторные параметры \mathbf{X} зависят от выбранной системы координат, в которой определяются параметры траекторий, и от способа представления траекторий. Чаще всего для описания траекторий используют полиномы. Например, если движение цели в ходе траекторной обработки рассматривается в декартовой системе координат и траектория цели описывается полиномом первой степени по каждой координате, то

$$\mathbf{X} = [\lambda_{x0} \quad \lambda_{x1} \quad \lambda_{y0} \quad \lambda_{y1} \quad \lambda_{z0} \quad \lambda_{z1}]^T, \quad (6.3)$$

где λ_{x0} , λ_{y0} , λ_{z0} — соответствующие координаты положения объекта в момент привязки измерений t_{np} ; λ_{x1} , λ_{y1} , λ_{z1} — соответствующие скорости объекта в момент привязки измерений t_{np} .

В данном представлении траектория (как след от перемещения цели) будет описываться в декартовом пространстве следующими соотношениями:

$$\xi_x(t) = \lambda_{x0} + \lambda_{x1}(t - t_{np}),$$

$$\xi_y(t) = \lambda_{y0} + \lambda_{y1}(t - t_{np}),$$

$$\xi_z(t) = \lambda_{z0} + \lambda_{z1}(t - t_{np}).$$

При использовании полинома степени s уравнение траектории по некоторой координате, например по x , можно представить в виде степенного ряда:

$$\xi_x(t) = \sum_{l=0}^s \lambda_{x,l} \frac{(t - t_{np})^l}{l!}. \quad (6.4)$$

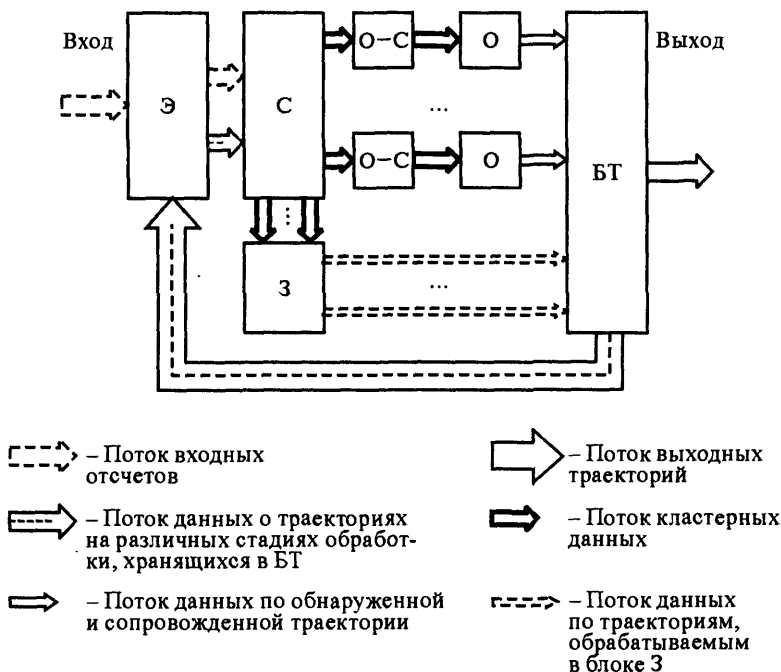


Рис. 6.2. Структурная схема алгоритма вторичной обработки

Производные приведенных соотношений определяют изменение скорости (траекторию скорости) цели в процессе ее движения:

$$\dot{\xi}_x(t) = \frac{\partial \xi_x(t)}{\partial t}, \quad \dot{\xi}_y(t) = \frac{\partial \xi_y(t)}{\partial t}, \quad \dot{\xi}_z(t) = \frac{\partial \xi_z(t)}{\partial t}.$$

Аналогичным образом могут быть получены траектории ускорений и т. д.

Структурная схема типичного алгоритма траекторной обработки приведена на рис. 6.2. Алгоритм включает в себя операции, которым соответствуют следующие блоки: блок экстраполяции (Э), блок селекции (С), блок обнаружения—сброса траекторий (О—С), блок оценки траекторных параметров (О), блок завязки траекторий (З), база траекторных данных (БТ). Разделение схемы на указанные блоки является в определенной степени условным. В принципе, в ходе траекторной обработки должно осуществляться совместное обнаружение—оценивание целевой обстановки в зоне ответственности РЛС. В различных модификациях алгоритмов некоторые блоки могут объединяться, некоторые — отсутствовать. В целом, приведенная структурная схема, объединяющая операции ВО в единый алгоритм, носит эвристический характер.

В ходе работы алгоритм ВО формирует траектории, находящиеся на различных стадиях обработки (информация о них хранится в базе траекторных данных). Среди них наиболее часто выделяют:

— *обнаруженные сопровождаемые траектории*, имеющие выходные параметры, сформированные в ходе ВО в полном объеме* и удовлетворяющие заданным вероятностным и точностным характеристикам; параметры этих траекторий выдаются потребителям;

— *завязанные траектории*, имеющие вектор выходных параметров полного объема, но показатели качества которых (вероятностные и точностные) удовлетворяют лишь некоторым промежуточным значениям, не достигая величин, заданных для обнаруженных сопровождаемых траекторий;

— *завязываемые траектории*, имеющие вектор выходных параметров, сформированный еще не в полном объеме, но показатели качества которых не достигают заданных значений даже для завязанных траекторий.

Возможно введение и ряда других стадий обработки траекторий со своими внутренними показателями качества.

Работа алгоритма ВО начинается с поступления на вход в момент времени t_k отсчетов с выхода первичной обработки. В блоке Э (рис. 6.2) параметры всех траекторий $\hat{X}_{j,k-1}$, $j = 1, 2, \dots, N$, которые были сформированы на предыдущих тактах обработки информации и хранятся в базе траекторных данных (блок БТ), пересчитываются (экстраполируются) на момент t_k и представляются в виде экстраполированных параметров X_{jk} , $j = 1, 2, \dots, N$. Затем в блоке С осуществляется их отождествление с поступившими отсчетами — выполняется операция селекции. В силу ошибок оценивания параметров отсчета на этапе первичной обработки и ошибок нахождения траекторных параметров невозможно абсолютно точно определить пару отсчет—траектория, относящуюся к некоторой одной цели. Обычно при экстраполяции траектории можно лишь указать область, попадание в которую отсчета не противоречит имеющимся ошибкам измерений параметров цели. Такие области являются зонами связи траекторий с отсчетами, они позволяют сформировать кластеры — совокупности соответствующих отсчетов и траекторий, возможно, относящихся к одной и той же цели.

В случае достаточно простой целевой и помеховой обстановки в некоторый момент времени t_k образуются кластеры следующих видов:

* Например, если отсчет включает только положение цели, то для построения траектории в виде полинома степени s принципиально необходимо соответствующее число тактов наблюдения.

1) траектория X_{ijk} и подтверждающий ее отсчет Z_{ik} (рис. 6.3, а);

2) траектория X_{ijk} , не имеющая подтверждающего отсчета (рис. 6.3, б);

3) отсчет Z_{ik} , не относящийся ни к одной траектории (рис. 6.3, в).

Предполагается, что каждый кластер соответствует некоторой цели, для которой строится траектория, находящаяся в той или иной стадии обработки. Алгоритмами ВО кластеры обрабатываются параллельно цепочками блоков О—С — О и блоком З.

Кластеры первого и второго вида с траекториями, обнаруженными и сопровождаемыми на предыдущих тактах, а также кластеры первого вида с завязанными траекториями направляются на дальнейшую обработку в свои блоки О—С и О. Кластеры второго вида с траекториями, находящимися в стадии завязки, и кластеры третьего вида направляются на обработку в блок З.

В более сложной помеховой и целевой ситуации возможно образование кластеров других видов:

4) одна траектория X_{ijk} , с высокой достоверностью подтверждаемая несколькими отсчетами $Z_{i_1k}, Z_{i_2k}, Z_{i_3k}$, один из которых, возможно, целевой (рис. 6.4, а);

5) несколько траекторий X_{j_1k}, X_{j_2k} и несколько отсчетов $Z_{i_1k}, Z_{i_2k}, Z_{i_3k}, Z_{i_4k}$, однозначное (достоверное) разделение которых на кластеры видов 1) — 3) или даже вида 4) не представляется возможным (рис. 6.4, б).

В блоке О—С подтверждается решение об обнаружении траектории или, напротив, принимается решение о ее сбросе. Это связано с тем, что, во-первых, цели в зоне контроля РЛС появляются и исчезают, во-вторых, из-за наличия помеховых отсчетов в алгоритме ВО помимо целевых траекторий обрабатываются и ложные, которые с течением времени (при правильной работе алгоритма) должны быть сброшены. В блоке О для каждого кластера уточняются оценки траекторных параметров. Достаточно часто очередность выполнения операций изменяется: блоки О и О—С меняются местами.

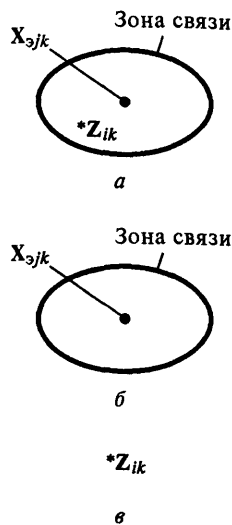


Рис. 6.3. Возможные варианты кластеров в простой обстановке:

а — кластер первого вида: одна траектория — один отсчет; б — кластер второго вида: одна траектория — ни одного отсчета, в — кластер третьего вида: «свободный» отсчет

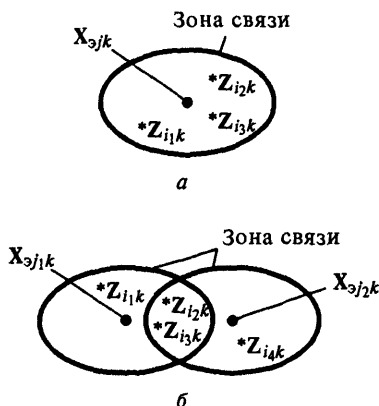


Рис. 6.4. Возможные варианты кластеров в сложной обстановке:

a — кластер четвертого вида: одна траектория — несколько отсчетов;
б — кластер пятого вида: несколько траекторий — несколько отсчетов

видов 1) и 2) (рис. 6.3, *a* и *б*), предполагается, что любой отсчет, поступивший с первичной обработки, соответствует обнаруженной цели. В этом случае под вторичной обработкой понимают, как уже было отмечено, прежде всего оценку тем или иным методом траекторных параметров.

При ухудшении помеховой ситуации, достаточно частом появлении кластеров четвертого вида (рис. 6.4, *a*) и плотных потоках наблюдаемых целей появление кластеров пятого вида (рис. 6.4, *б*), как минимум, усложняет операцию селекции.

К вопросам вторичной обработки радиолокационной информации в последнее время проявляется большой интерес. Это направление информационных технологий, особенно для сложной целевой и помеховой обстановки, активно развивается. В рамках настоящего учебного пособия ограничимся построением базовых алгоритмов ВО для достаточно типичной помеховой и сравнительно простой целевой обстановки.

6.2. Модели целевой и помеховой обстановки

6.2.1. Модели движения целей

Наблюдаемые радиолокационные цели: наземный транспорт, корабли, самолеты, космические аппараты и другие объекты — могут двигаться по самым разнообразным траекториям, имеющим, как правило,

случайный характер. Это вызвано тем, что движение любого реального объекта является чрезвычайно сложным, поскольку происходит под действием различных сил, зависящих от особенностей самого объекта, его конструкции, системы управления, свойств среды, в которой он движется, и других факторов. Сложность движения объекта затрудняет его изучение в полном объеме. Поэтому, как это обычно принято при исследовании, реальный процесс движения заменяется некоторой упрощенной моделью.

Математическая модель движения целей представляется в виде некоторых уравнений и ограничений, характеризующих представления о динамических свойствах объектов наблюдения и определяющих взаимосвязь их координат в различные моменты времени. С позиций теории систем совокупность используемых выражений, в общем случае векторных, описывает состояние некоторой *динамической системы*.

Математическая модель строится на основе всестороннего анализа поведения системы с использованием как априорных сведений, так и результатов статистических испытаний. Всегда существует противоречие между стремлением сделать модель системы наиболее полной, чтобы точнее описать реальный процесс, и необходимостью представления ее в достаточно простой форме, удобной для анализа и использования. Адекватный выбор модели системы, в данном случае — модели движения объекта, является одной из важнейших предпосылок, обеспечивающих построение алгоритма ВО, работающего с требуемым качеством.

Модели движения, используемые при рассмотрении различных объектов, разделяют на динамические и кинематические в соответствии с использованием или неиспользованием в них сведений об инерции движущегося объекта и силах, воздействующих на него. Кинематические уравнения обычно проще динамических, в них учитывается движение только центра масс без выделения причин движения, и именно они чаще всего используются для нужд вторичной обработки.

В соответствии с выбранной моделью могут рассматриваться системы, работающие в непрерывном или дискретном времени. Они могут быть описаны соответственно дифференциальными или разностными уравнениями. Эти уравнения называются *уравнениями состояния динамической системы*. Для задачи построения траекторий их также можно назвать *уравнениями модели процесса движения целей*, или просто *уравнениями движения цели*.

Собственно параметры траектории являются зависимой от времени векторной переменной — решением уравнения состояния динамической системы — и могут быть представлены вектором $X_j(t)$ состояния системы,

элементы которого — рассматриваемые траекторные параметры некоторой j -й траектории (некоторого j -го объекта)*.

Для системы непрерывного времени в общем случае уравнение состояния имеет вид

$$\frac{d\mathbf{X}}{dt} = \mathbf{q}(\mathbf{X}, \mathbf{E}, t),$$

где \mathbf{E} — вектор некоторых параметров, характеризующих объект, внешнюю среду, систему управления и др.; $\mathbf{q}(\cdot)$ — векторная функция аргументов \mathbf{X} , \mathbf{E} , t . Решая уравнение состояния при некоторых начальных условиях, можно получить вектор состояния динамической системы в любой момент времени t . Тем самым будет определено движение некоторого объекта, т. е. найдена зависимость его траекторных параметров (фазовых координат) от времени, и может быть построена траектория — след от движения объекта.

Для динамической системы дискретного времени в общем случае уравнение состояния имеет вид

$$\mathbf{X}_{k+1} = \mathbf{f}(\mathbf{X}_k, \mathbf{E}_k, k),$$

где k , $k+1$ — целочисленные индексы, обозначающие дискретные моменты времени, $\mathbf{f}(\cdot)$ — векторная функция, характеризующая переход системы из состояния \mathbf{X}_k в \mathbf{X}_{k+1} (изменение траекторных параметров во времени). Решение этого уравнения, как и в предыдущем случае, определит вектор состояния системы, но в дискретные моменты времени.

Динамические системы (в данном случае процессы, характеризующие движение объектов) могут быть детерминированными или стохастическими. Поскольку реальные системы и воздействие на них внешней среды не может адекватно описать ни одна детерминированная модель, в ходе построения траекторий для учета имеющихся неопределенностей и приближений чаще всего используется стохастическая модель системы. На практике, даже когда точно известна модель системы, но она описывается сложными математическими выражениями, имеет смысл упростить модель введением в нее некоторого эквивалентного случайного возмущения.

В настоящем учебном пособии рассматривается случай дискретного времени, как наиболее распространенный при проведении ВО.

* Чтобы не усложнять запись выражений, при рассмотрении только одной траектории (одного процесса) индекс, характеризующий номер траектории, здесь и далее в настоящей главе опускается.

В общем случае для стохастических динамических систем дискретного времени уравнение состояния имеет вид

$$\mathbf{X}_{k+1} = \mathbf{f}(\mathbf{X}_k, \mathbf{u}_k, \mathbf{v}_k, k), \quad (6.5)$$

где \mathbf{u}_k — p -мерный вектор детерминированных входных воздействий в момент k , \mathbf{v}_k — g -мерный вектор случайных возмущений в момент k .

Вектор состояния \mathbf{X} системы имеет размерность b , соответствующую размерности физического пространства, в котором рассматривается траектория, и числу фазовых координат в нем, которыми характеризуются траекторные параметры. Например, для представления вектора состояния в виде (6.3) размерность c физического пространства равна 3 (три координаты декартового пространства), число оцениваемых параметров по каждой физической координате равно 2 (дальность и скорость при аппроксимации траектории полиномом первой степени ($s = 1$)); в результате имеем $b = c(s + 1) = 6$. Размерности p и q векторов \mathbf{u}_k и \mathbf{v}_k определяются непосредственно соответствующими управляющими и возмущающими воздействиями.

Уравнение (6.5) в общем случае нелинейно, имеет стохастический характер, зависит как от детерминированных, так и от случайных параметров. Явная зависимость вектора состояния (см. формулу (6.5)) от k позволяет описывать нестационарные системы.

Наиболее часто используются линейные приближения, для которых уравнение состояния является векторным линейным разностным уравнением вида

$$\mathbf{X}_{k+1} = \mathbf{F}_{k+1}\mathbf{X}_k + \mathbf{G}_{k+1}\mathbf{u}_k + \mathbf{G}_{k+1}\mathbf{v}_k, \quad (6.6)$$

где $\mathbf{F}_{k+1}, \mathbf{G}_{k+1}, \mathbf{G}_{k+1}$ — известные матрицы размерности $b \times b$, $b \times p$, $b \times g$ соответственно.

Матрица \mathbf{F}_{k+1} получила название *матрицы экстраполяции*. Она устанавливает связь траекторных параметров в моменты k и $k + 1$ при отсутствии возмущающих воздействий. Матрица \mathbf{G}_{k+1} характеризует влияние детерминированных возмущений \mathbf{u}_k в момент k на траекторные параметры в момент $k + 1$. Матрица \mathbf{G}_{k+1} характеризует влияние случайных возмущений \mathbf{v}_k в момент k на траекторные параметры в момент $k + 1$.

Для заданных матриц $\mathbf{F}_{k+1}, \mathbf{G}_{k+1}, \mathbf{G}_{k+1}$ очередное состояние \mathbf{X}_{k+1} системы определяется текущим состоянием \mathbf{X}_k и входными воздействиями $\mathbf{u}_k, \mathbf{v}_k$.

При полиномиальной аппроксимации (см. (6.4)) траектории в одномерном пространстве ($c = 1$) переходная матрица \mathbf{F}_{k+1} имеет вид

$$\mathbf{F}_{k+1} = \begin{bmatrix} 1 & \tau_{k+1} & \tau_{k+1}^2/2! & \dots & \tau_{k+1}^s/s! \\ 0 & 1 & \tau_{k+1} & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 1 & \tau_{k+1} \\ \dots & \dots & \dots & \dots & 1 \end{bmatrix}, \quad (6.7)$$

где $\tau_{k+1} = t_{k+1} - t_k$.

Матрица \mathbf{F}_{k+1} имеет в данном случае размерность $b \times b = c(s+1) \times c(s+1)$ и определяется представлением траектории в виде полинома степени s .

В самом простом случае уравнение состояния является одновременно и линейным, и детерминированным:

$$\mathbf{X}_{k+1} = \mathbf{F}_{k+1} \mathbf{X}_k. \quad (6.8)$$

Естественно, область применения такой модели ограничена.

При построении траекторий обычно рассматривается линейная модель динамической системы, которая описывается линейным уравнением состояния

$$\mathbf{X}_{k+1} = \mathbf{F}_{k+1} \mathbf{X}_k + \mathbf{\Gamma}_{k+1} \mathbf{v}_k, \quad (6.9)$$

где \mathbf{v}_k — процесс типа белого гауссовского шума с нулевым средним и ковариационной (положительно определенной) матрицей

$$\mathbf{Q}_k = \mathbf{M}\{\mathbf{v}_k \mathbf{v}_k^T\}. \quad (6.10)$$

В уравнении (6.9) процесс \mathbf{v}_k учитывает возмущения траектории флуктуационного типа (обусловленные неоднородностью среды, неточностью системы управления и другими подобного рода факторами).

В радиолокационной практике принято цели разделять на неманеврирующие и маневрирующие. Цель называют *неманеврирующей*, если она движется по прямой с постоянной скоростью, т. е. по всем физическим координатам описывается полиномом не выше первой степени, и *маневрирующей* — во всех остальных случаях. Иногда допускают более расширенное толкование маневра. Например, для орбит спутников, которые принципиально имеют нелинейный вид, под маневром понимают лишь переход с одной орбиты на другую. Иногда любую детерминированную траекторию относят к случаю неманеврирующих целей. Если это не будет оговорено специально, мы будем понимать неманеврирующую цель как объект, движущийся прямолинейно и равномерно.

Заметим, что необходимо различать линейную модель системы, описываемую линейным уравнением состояния, и линейную траекторию движения цели. При линейной модели системы траектория (след от движения цели) может быть и линейной, и нелинейной. Строго говоря, рассмотренные выше линейные уравнения состояния только в случае (6.8) и только при использовании полинома первой степени описывают неманеврирующую цель, все другие случаи фактически относятся к маневрирующим целям.

В зависимости от особенностей наблюдаемых целей обычно выделяют преднамеренные и непреднамеренные маневры, маневры большой и малой интенсивности и т. п.

Тот или иной вид движения объекта можно рассмотреть на примере полета воздушной цели. Обычно ее траекторию делят на участки двух типов: практически прямолинейного движения и криволинейного. Используя линейные уравнения состояний, выбирая соответствующим образом их параметры, можно описать оба типа участков траектории. В первом случае движение цели описывают полиномом первой степени и представляют или моделью с отсутствием маневра (6.8), или моделью с непреднамеренным маневром малой интенсивности (6.9). Для описания движения во втором случае можно, например, использовать модель (6.8) при выборе полинома степени больше первой (детерминированная модель), можно использовать также и модель (6.9).

Обычно выделяют маневрирование по курсовой скорости и по направлению. Для летательных аппаратов маневрирование по курсовой скорости ограничено достижимым ускорением, которое редко превышает $(0,8 \dots 1,0)g_0$, где $g_0 = 9,81 \text{ м/с}^2$ — ускорение земного притяжения. Маневрирование по направлению ограничено допустимой перегрузкой $n_m = g_m / g_0 = 5 \dots 8$, где g_m — поперечное ускорение маневра. При движении по окружности со скоростью v_m минимальный радиус r_{\min} траектории связан с допустимой перегрузкой соотношением

$$r_{\min} = \frac{v_m^2}{g_0 \sqrt{n_m^2 - 1}}.$$

Типичные кинематические параметры различных целей приведены в табл. 6.1 [61].

Разработчики РЛС, как правило, стремятся упростить уравнения состояния при минимальном ухудшении точности модели. Чаще всего с этой целью, как уже отмечалось, в уравнения состояния вводится шум, который учитывает неполноту знаний об истинной модели движения и различные непредсказуемые явления.

Таблица 6.1

Кинематические параметры целей

Тип цели	Курсовая скорость, м/с	Параметры поперечного ускорения цели или скорости разворота
Корабль	0...20	2 град/с
Военные:		
самолет	50...1000	50...80 м/с ²
ракета	200...1200	до 100 м/с ²
вертолет	0...80	1,3...3 град/с
Гражданский самолет	50...300	1,3...3 град/с

Однако в ряде случаев, например при точных расчетах параметров орбит спутников, траекторий ракет, необходимо использовать подробное описание моделей движения в виде нелинейных функций, а в особо сложных случаях переходить от кинематических уравнений к динамическим.

Для траекторной обработки в некотором смысле традицией стало использование уравнений состояния (6.9), где по каждой физической координате траектория цели описывается полиномом не выше первой степени с возмущающим воздействием в виде случайного ускорения, представляемого белым гауссовским шумом с нулевым средним и некоторой дисперсией. Этой модели соответствует генеральное движение цели по прямой линии при непреднамеренном маневре шумового характера. Однако и при более сложном поведении цели данную модель используют достаточно часто, а возникающее несоответствие реальному движению компенсируют определенным увеличением дисперсии шума.

В уравнении состояния (6.9) при описании траектории в трехмерном декартовом пространстве вектор состояний имеет вид $\mathbf{X} = [\lambda_{x0} \lambda_{x1} \lambda_{y0} \lambda_{y1} \lambda_{z0} \lambda_{z1}]^T$, вектор возмущающих воздействий — $\mathbf{v} = [g_x \ g_y \ g_z]^T$, а матрицы имеют вид

$$\mathbf{F}_{k+1} = \begin{bmatrix} 1 & \tau_{k+1} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \tau_{k+1} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \tau_{k+1} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\Gamma_{k+1} = \begin{bmatrix} \tau_{k+1}^2/2 & 0 & 0 \\ \tau_k & 0 & 0 \\ 0 & \tau_{k+1}^2/2 & 0 \\ 0 & \tau_k & 0 \\ 0 & 0 & \tau_{k+1}^2/2 \\ 0 & 0 & \tau_{k+1} \end{bmatrix}. \quad (6.11)$$

Дальнейшее совершенствование модели траектории обычно вызывается необходимостью более точного учета маневров цели.

Для учета разнообразных вариантов движения цели в модели (6.9) можно увеличивать степень аппроксимирующего полинома, вводить ненулевую регулярную составляющую случайного ускорения, зашумление других фазовых координат, а также использовать негауссовские шумовые воздействия [61, 63].

6.2.2. Модели отсчетов

Для ВО входной информацией, как уже отмечалось в § 6.1, являются отсчеты $Z_{1k}, Z_{2k}, \dots, Z_{ik}, \dots, Z_{mk}$, поступающие в k -й момент времени с выхода первичной обработки. При формировании отсчетов определение координат соответствующих им целей производится на основании анализа некоторого сигнала (обычно, как показано в гл. 3, корреляционного интеграла). При наличии шумов (помех), во-первых, возможно необнаружение отсчета от какой-либо цели или обнаружение помеховых отсчетов, во-вторых, произведенные первичные измерения имеют некоторые случайные ошибки относительно истинных значений координат цели.

Поступающие с первичной обработки отсчеты являются для вторичной входными наблюдениями. При математическом описании отсчетов их можно рассматривать как некоторый поток случайных точек. Теория случайных потоков [68] предполагает, что каждая точка потока (отсчет) появляется с некоторой вероятностью в соответствующей области определения (зоне контроля РЛС) и имеет некоторые случайные параметры (измеренные координаты). Статистика отсчетов в существенной мере достаточно стабильна для большинства ситуаций, имеющих место на практике, поскольку входные эхо-сигналы подверглись целенаправленному воздействию процедур первичного обнаружения и измерения, ориентированных на извлечение информации оптимальным образом.

При предположении о том, что на выходе первичной обработки каждому объекту соответствует не более одного отсчета, а каждый полезный

отсчет порожден не более чем одним объектом, совокупность целевых отсчетов образует поток Бернулли [57]. Это соответствует типичному случаю наблюдения разрешенных целей: обнаружению отсчета от цели (см. гл. 3, 4) с вероятностью $D_{от} < 1$ и измерению его параметров тем или иным методом (см. гл. 5). Ошибки нахождения параметров отсчета, как правило, описываются нормальным законом распределения с дисперсиями, определяемыми зондирующим сигналом, отношением сигнал—шум и методом измерения. При оценке нескольких координат, например дальности, азимута, угла места, радиальной скорости, соответствующие ошибки могут быть как независимыми (чаще всего), так и зависимыми. Точностные характеристики каждого отсчета Z представляются ковариационной матрицей R ошибок оценивания первичных измерений, в которой диагональные элементы состоят из дисперсий ошибок измеряемых параметров, а недиагональные (ковариации) либо равны нулю, если соответствующие измерения независимы (коэффициент корреляции равен нулю), либо равны некоторым ненулевым значениям. Общий вид ковариационной матрицы Z приведен в выражении (6.2).

Поток ложных отсчетов в типичной ситуации является пуассоновским [57]. Интенсивность потока ложных отсчетов определяется вероятностью ложной тревоги $F_{от}$, длительностью такта T_1 первичной обработки и размером зоны контроля РЛС. Закон распределения параметров ложного отсчета можно считать равномерным в зоне контроля РЛС (или в некоторой части этой зоны). Могут быть и более сложные модели потоков полезных и ложных отсчетов [57, 68].

Параметры отсчетов, получаемые в ходе первичной обработки, представляются в станционной (радиолокационной) системе координат; параметры траекторий в соответствии с требованиями вышестоящей системы могут оцениваться в той же системе координат или в некоторой другой и, как правило, с большим числом фазовых координат (обычно за счет дополнительной оценки скорости, ускорения и т. д.).

Например, пусть параметры отсчета Z размерности $a = 4$ включают: дальность до цели r , азимут цели β , угол места ε , радиальную скорость цели v_r , т. е.

$$Z = [r \ \beta \ \varepsilon \ v_r]^T.$$

Вектор состояния X размерности b траектории, представленной в физической системе координат отсчета размерности $c = 3$, может быть получен для полинома первой степени ($b = c(s+1) = 3 \times 2 = 6$) в виде

* Чтобы не усложнять запись выражений, при рассмотрении только одного отсчета (или траектории) индексы, характеризующие номер отсчета (или траектории) и временную привязку, здесь и далее опускают.

$$\mathbf{X} = [\lambda_{r0} \ \lambda_{r1} \ \lambda_{\beta 0} \ \lambda_{\beta 1} \ \lambda_{\varepsilon 0} \ \lambda_{\varepsilon 1}]^T,$$

для полинома второй степени ($b = c(s+1) = 3 \times 3 = 9$) в виде

$$\mathbf{X} = [\lambda_{r0} \ \lambda_{r1} \ \lambda_{r2} \ \lambda_{\beta 0} \ \lambda_{\beta 1} \ \lambda_{\beta 2} \ \lambda_{\varepsilon 0} \ \lambda_{\varepsilon 1} \ \lambda_{\varepsilon 2}]^T,$$

где λ_{r0} , $\lambda_{\beta 0}$, $\lambda_{\varepsilon 0}$ — соответствующие координаты положения объекта в станционной (сферической в данном случае) системе координат в момент привязки измерений $t_{\text{пр}}$; λ_{r1} , $\lambda_{\beta 1}$, $\lambda_{\varepsilon 1}$ — соответствующие составляющие скорости объекта в станционной системе координат в момент привязки измерений $t_{\text{пр}}$; λ_{r2} , $\lambda_{\beta 2}$, $\lambda_{\varepsilon 2}$ — соответствующие составляющие ускорения объекта в станционной системе координат в момент привязки измерений $t_{\text{пр}}$.

Траектории, представленные не в физической системе координат отсчета, а, например, в трехмерной декартовой в виде полинома первой и второй степеней, имеют соответственно векторы состояний

$$\mathbf{X} = [\lambda_{x0} \ \lambda_{x1} \ \lambda_{y0} \ \lambda_{y1} \ \lambda_{z0} \ \lambda_{z1}]^T,$$

$$\mathbf{X} = [\lambda_{x0} \ \lambda_{x1} \ \lambda_{x2} \ \lambda_{y0} \ \lambda_{y1} \ \lambda_{y2} \ \lambda_{z0} \ \lambda_{z1} \ \lambda_{z2}]^T,$$

где λ_{x0} , λ_{y0} , λ_{z0} — соответствующие координаты положения объекта в декартовой системе координат в момент привязки измерений $t_{\text{пр}}$; λ_{x1} , λ_{y1} , λ_{z1} — соответствующие скорости объекта в декартовой системе координат в момент привязки измерений $t_{\text{пр}}$; λ_{x2} , λ_{y2} , λ_{z2} — соответствующие ускорения объекта в декартовой системе координат в момент привязки измерений $t_{\text{пр}}$.

Траекторные параметры некоторой цели для заданного момента времени, как правило, однозначно определяют соответствующие параметры отсчета. Например, если рассматривается равномерное прямолинейное движение цели на плоскости и вектор траекторных параметров имеет вид

$$\mathbf{X} = [\lambda_{x0} \ \lambda_{x1} \ \lambda_{y0} \ \lambda_{y1}]^T,$$

а отсчет $\mathbf{Z} = [r \ \beta]^T$ получен в полярной (станционной) системе с общим началом координат, то связь пространства параметров отсчета с пространством параметров траектории (при отсутствии шумов измерений) определяется векторным соотношением

$$\mathbf{h}(\mathbf{X}) = [r \ \beta]^T, \quad (6.12)$$

$$r = \sqrt{\lambda_{x0}^2 + \lambda_{y0}^2}, \quad \beta = \arctg\left(\frac{\lambda_{y0}}{\lambda_{x0}}\right).$$

Заметим, что вычисление вектора \mathbf{X} по единственному отсчету не всегда возможно, поскольку необходимые для этого измерения могут отсутствовать на этапе первичной обработки радиолокационной информации.

Функция связи $\mathbf{h}(\mathbf{X})$ пространства параметров отсчета с пространством параметров траектории может быть нелинейной (как в приведенном выше примере) и линейной. В простейшем случае линейная функция связи встречается при использовании одних и тех же физических координат для полиномиального представления фазовых координат отсчета и траектории. Например, пусть рассматривается равномерное прямолинейное движение цели в полярной системе координат с вектором траекторных параметров

$$\mathbf{X} = [\lambda_{r0} \ \lambda_{r1} \ \lambda_{\beta 0} \ \lambda_{\beta 1}]^T,$$

а отсчет получается в той же системе координат, но без скоростных компонент:

$$\mathbf{Z} = [r \ \beta]^T;$$

тогда связь пространства параметров отсчета с пространством параметров траектории или параметров отсчета с параметрами траектории (при отсутствии шумов измерений) определяется соотношением

$$\mathbf{h}(\mathbf{X}) = [r \ \beta]^T, \quad r = \lambda_{r0}, \quad \beta = \lambda_{\beta 0}. \quad (6.13)$$

Если истинные траекторные параметры j -й цели в k -й момент времени равны \mathbf{X}_{jk} , то параметры отсчета \mathbf{Z}_{ik} от той же цели в тот же момент времени, полученные на выходе первичной обработки информации, всегда будут искажены шумами измерения (см. гл. 5). В самом общем виде можно записать

$$\mathbf{Z}_{ik} = \mathbf{g}(\mathbf{X}_{jk}, \boldsymbol{\varepsilon}_{ik}, k), \quad (6.14)$$

где $\boldsymbol{\varepsilon}_{ik}$ — r -мерный вектор ошибок измерения параметров отсчета, $\mathbf{g}(\cdot)$ — векторная функция аргументов $\mathbf{X}_{jk}, \boldsymbol{\varepsilon}_{ik}, k$.

Соотношение (6.14) называется *уравнением измерений*. Пользуясь терминологией теории систем, можно сказать, что функция $\mathbf{g}(\cdot)$ отображает внутреннее состояние системы \mathbf{X}_{jk} на измеряемые параметры отсчета \mathbf{Z}_{ik} , а

также характеризует влияние случайных ошибок первичных измерений. Введение в явной форме в уравнение (6.14) зависимости от k так же, как и в (6.5), позволяет описывать нестационарные процессы в ходе первичных измерений. Функция $g(\cdot)$ может быть линейной и нелинейной.

В случае отсутствия ошибок измерений при нахождении отсчета функция $g(\cdot)$ совпадает с соответствующей функцией связи пространства параметров отсчета с пространством параметров траектории. Например, если рассматривается равномерное прямолинейное движение цели на плоскости, а отсчет получается в полярной системе координат, то функция $g(\cdot)$ описывается нелинейными соотношениями (6.12); если рассматривается равномерное прямолинейное движение цели в полярной системе координат и отсчет получается в той же системе координат, то функция $g(\cdot)$ описывается линейными соотношениями (6.13). Ошибки измерения параметров отсчета могут учитываться в уравнении измерений как нелинейно, так и линейно.

Особое место в теории систем и во вторичной обработке информации занимает случай, когда уравнение измерений является линейным. Тогда его можно представить в виде

$$\mathbf{Z}_{ik} = \mathbf{H}_k \mathbf{X}_{jk} + \varepsilon_{ik}, \quad (6.15)$$

где \mathbf{H}_k — известная матрица пересчета пространства состояния динамической системы (пространства траекторных параметров) в пространство отсчетов. Матрица \mathbf{H}_k получила название *матрицы связи пространства параметров отсчета с пространством траекторных параметров*. Для приведенного выше примера, когда функция $h(\mathbf{X})$ описывается соотношением (6.13), имеем

$$\mathbf{H}_k = [1 \ 0 \ 1 \ 0].$$

Обычно вектор ε_{ik} , характеризующий шум измерений параметров отсчета, является реализацией случайного процесса типа белого шума с нулевым средним и матрицей ковариации

$$\mathbf{M}\{\varepsilon_{ik}\varepsilon_{ik}^T\} = \mathbf{R}_{ik}.$$

Диагональные элементы матрицы \mathbf{R}_{ik} представляют дисперсии ошибок измерений параметров отсчета.

6.3. Оценка траекторных параметров по фиксированной выборке. Экстраполяция траекторных параметров

Оценка траекторных параметров движения цели в соответствии с общей структурной схемой ВО проводится в блоке О (см. рис. 6.2) по отсчетам, отобранным в ходе операции селекции и относящимся к одной цели. Не учитывая возможные ошибки селекции, положим, что параметры каждой j -й траектории должны оцениваться по набору относящихся к ней отсчетов, полученных с начала ее наблюдения t_1 до текущего момента времени t_k :

$$\Omega_j^{(k)} = [Z_j(t_1) \ Z_j(t_2) \ \dots \ Z_j(t_k)]^T. \quad (6.16)$$

Предполагается, что набор данных $\Omega_j^{(k)}$ по каждой траектории j формируется в процессе селекции кластером первого вида (см. рис. 6.3, а). В рамках базовых алгоритмов ВО оценивание траекторных параметров проводится по каждой траектории независимо друг от друга.

Для решения задачи оценивания траекторных параметров в настоящее время обычно используют два подхода: на основе фиксированной выборки измерений и на основе последовательных во времени измерений при рекуррентном уточнении параметров траектории [61, 63, 67].

В первом случае в текущий момент времени t_k по данным предыдущих n тактов первичной обработки сначала для некоторой j -й цели формируют фиксированную выборку наблюдений, являющуюся некоторой частью полного объема отсчетов $\Omega_j^{(k)}$, полученных по этой траектории:

$$\Omega_j^{(n/k)} = [Z_j(t_{k-(n-1)}) \ Z_j(t_{k-(n-2)}) \ \dots \ Z_j(t_k)]^T. \quad (6.17)$$

Затем по этой выборке проводится оценка $\hat{X}_j(t_{np})$ траекторных параметров на некоторый момент времени t_{np} (обычно предполагают, что $t_{np} = t_k$; см. п. 6.1.2).

Во втором случае для некоторой j -й цели оценку $\hat{X}_j(t_k)$ траекторных параметров вычисляют рекуррентно после получения в текущий момент времени t_k с первичной обработки соответствующего отсчета $Z_j(t_k)$ с учетом оценки $\hat{X}_j(t_{k-1})$ траекторных параметров на предыдущем такте.

При оценке по фиксированной выборке, как правило, используются очень простые, детерминированные и потому «грубые» модели (6.8) движе-

ния целей. Это вполне допустимо при малом временном интервале $\Delta t = (t_k - t_{k-(n-1)})$ формирования выборки (при выборе малого Δt любую кривую можно аппроксимировать с точностью, достаточной для практических задач, даже отрезком прямой линии). На практике оценку траекторных параметров по фиксированной выборке обычно применяют либо в режиме «скользящего» окна для заданного интервала Δt (или, что то же самое, по соответствующей величине объема выборки n), либо в ходе операции завязки, когда информация о параметрах движения цели еще крайне скудна и необходимо получить о траектории движения цели первые сведения.

Для нахождения оценки траекторных параметров по фиксированной выборке чаще всего используют метод максимального правдоподобия (см. гл. 5).

Опуская (для упрощения записи) номер j анализируемой траектории и считая $t_{k-(n-1)}$ первым моментом времени, $t_{k-(n-2)}$ — вторым, ..., t_k — n -м, входную выборку отсчетов можно записать в виде

$$\Omega^{(n)} = [Z_1 \ Z_2 \ \dots \ Z_n]^T.$$

Для линейного уравнения измерений (6.15) и стационарной матрицы связи $\mathbf{H} = \mathbf{H}_1, \dots, \mathbf{H}_n$ имеем

$$Z_1 = \mathbf{H}X_1 + \varepsilon_1,$$

$$Z_2 = \mathbf{H}X_2 + \varepsilon_2,$$

$$\dots\dots\dots$$

$$Z_n = \mathbf{H}X_n + \varepsilon_n.$$

где X_i , $i = 1, 2, \dots, n$, — траекторные параметры цели в i -й момент времени; ε_i , $i = 1, 2, \dots, n$, — ошибки измерения параметров цели (в пространстве наблюдений отсчета) на этапе первичной обработки в i -й момент времени.

Для полиномиальной детерминированной модели (6.8) движения цели, положив $t_{np} = t_n$, имеем

$$X_n = F_n X,$$

$$X_{n-1} = F_{n-1} X,$$

$$\dots\dots\dots$$

$$X_1 = F_1 X,$$

где $X = X_n$, $F_1 = F(\tau_1)$, $F_2 = F(\tau_2)$, ..., $F_n = F(\tau_n) = F(0) = \mathbf{I}$ (см. (6.7)), $\tau_1 = t_1 - t_n$, $\tau_2 = t_2 - t_n$, ..., $\tau_n = t_n - t_n = 0$ — времена экстраполяции, \mathbf{I} — единичная матрица. Тогда

$$\mathbf{Z}_1 = \mathbf{H}\mathbf{F}_1\mathbf{X} + \varepsilon_1,$$

$$\mathbf{Z}_2 = \mathbf{H}\mathbf{F}_2\mathbf{X} + \varepsilon_2,$$

$$\dots\dots\dots$$

$$\mathbf{Z}_n = \mathbf{H}\mathbf{F}_n\mathbf{X} + \varepsilon_n.$$

Полученную систему уравнений обычно представляют в векторном виде:

$$\mathbf{\Omega}^{(n)} = \mathbf{A}\mathbf{X} + \mathbf{E}^{(n)}, \quad (6.18)$$

где $\mathbf{A} = [\mathbf{H}\mathbf{F}_1^T \quad \mathbf{H}\mathbf{F}_2^T \quad \dots \quad \mathbf{H}\mathbf{F}_n^T]^T$, $\mathbf{E}^{(n)} = [\varepsilon_1 \quad \varepsilon_2 \quad \dots \quad \varepsilon_n]^T$.

Например, при полиномиальной аппроксимации траектории в пространстве с одной физической координатой $c = 1$ (см. (6.4)) матрица \mathbf{A} имеет вид

$$\mathbf{A} = \begin{bmatrix} 1 & \tau_1 & \tau_1^2/2! & \dots & \tau_1^s/s! \\ 1 & \tau_2 & \tau_2^2/2! & \dots & \tau_2^s/s! \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \tau_n & \tau_n^2/2! & \dots & \tau_n^s/s! \end{bmatrix}.$$

При гауссовской модели ошибок измерения параметров отсчета (см. п. 6.2.2) для совокупности ошибок $\mathbf{E}^{(n)}$ закон распределения будет также нормальным:

$$w[\mathbf{E}^{(n)}] = \frac{1}{(2\pi)^{an/2}(\det \mathfrak{R}^{(n)})^{1/2}} \exp \left\{ -\frac{1}{2} [\mathbf{E}^{(n)}]^T [\mathfrak{R}^{(n)}]^{-1} [\mathbf{E}^{(n)}] \right\}, \quad (6.19)$$

где a — размерность отсчета, $\mathfrak{R}^{(n)}$ — ковариационная матрица совокупности ошибок измерения параметров отсчетов $\mathbf{E}^{(n)}$.

Учитывая соотношения (6.18) и (6.19), можно получить функцию правдоподобия совокупности отсчетов $\mathbf{\Omega}^{(n)}$:

$$\Lambda(\mathbf{\Omega}^{(n)} | \mathbf{X}) = C \exp \left\{ -\frac{1}{2} [\mathbf{\Omega}^{(n)} - \mathbf{A}\mathbf{X}]^T [\mathfrak{R}^{(n)}]^{-1} [\mathbf{\Omega}^{(n)} - \mathbf{A}\mathbf{X}] \right\}, \quad (6.20)$$

где C — постоянный множитель.

Находя экстремум функции (6.20) после ее логарифмирования и дифференцирования по составляющим вектора оцениваемых траекторных параметров, приравнивая производную нулю при $\mathbf{X} = \hat{\mathbf{X}}$, получим векторное уравнение правдоподобия

$$\mathbf{A}^T [\mathfrak{R}^{(n)}]^{-1} [\mathbf{\Omega}^{(n)} - \mathbf{A}\hat{\mathbf{X}}] = 0. \quad (6.21)$$

Решая векторное уравнение (6.21), получим оценку траекторных параметров

$$\hat{\mathbf{X}} = \hat{\Psi} \mathbf{A}^T [\mathfrak{R}^{(n)}]^{-1} \mathbf{\Omega}^{(n)}, \quad (6.22)$$

где $\hat{\Psi}$ является ковариационной матрицей полученных оценок траекторных параметров:

$$\hat{\Psi} = \left(\mathbf{A}^T [\mathfrak{R}^{(n)}]^{-1} \mathbf{A} \right)^{-1}. \quad (6.23)$$

Если вектор \mathbf{X} состоит из b траекторных параметров, то ковариационная матрица $\hat{\Psi}$ имеет размерность $b \times b$:

$$\hat{\Psi} = \begin{bmatrix} \hat{\Psi}_{11} & \dots & \hat{\Psi}_{1b} \\ \dots & \dots & \dots \\ \hat{\Psi}_{b1} & \dots & \hat{\Psi}_{bb} \end{bmatrix},$$

каждый диагональный элемент этой матрицы является дисперсией соответствующего траекторного параметра, а недиагональный — ковариацией соответствующих траекторных параметров (см. п. 6.1.2).

Заметим, что полученные на основании критерия максимума функции правдоподобия соотношения (6.22) и (6.23) в случае нормальных ошибок измерений совпадают с результатами, которые следуют из критерия минимума средневзвешенных квадратов, а в предположении об априорном равномерном распределении оцениваемых параметров — и из критерия максимума апостериорной плотности вероятности.

Выражения (6.22) и (6.23) позволяют решать задачи, связанные с оценкой траекторных параметров по фиксированной выборке как для простейших случаев, так и для достаточно сложных [67].

Например, найдем оценку $\hat{\mathbf{X}} = [\hat{\lambda}_0 \ \hat{\lambda}_1]^T$ траекторных параметров линейной ($s = 1$) траектории $\mathbf{X} = [\lambda_0 \ \lambda_1]^T$ для некоторого одномерного ($c = 1$) физического пространства (индекс, обозначающий рассматриваемую координату физического пространства, для сокращения записи опущен) с числом фазовых координат $b = c(s + 1) = 2$ в условиях некоррелированных, равно- точных ошибок измерения параметров совокупности отсчетов $\mathbf{\Omega}^{(n)}$, состоящих из координатных членов z_i , $i = 1, 2, \dots, n$:

$$\Omega^{(n)} = [Z_1^T \ Z_2^T \ \dots \ Z_i^T \ \dots \ Z_n^T]^T = [z_1 \ z_2 \ \dots \ z_n]^T,$$

в той же физической системе координат, что и X (при этом $H = [1 \ 0]$), и получаемых через равные (равнодисcretные) промежутки времени $T = \text{const}$ (при этом $\tau_i = T(i - n)$).

В данном случае имеем

$$\mathcal{R}^{(n)} = \sigma_z^2 I,$$

где $\sigma_{z_1}^2 = \sigma_{z_2}^2 = \dots = \sigma_{z_n}^2 = \sigma_z^2$ — дисперсии ошибки единичного измерения параметров отсчетов в 1-й, 2-й, ..., n -й моменты времени (заметим, что в данном случае ковариационная матрица R_i каждого i -го отсчета есть скаляр $\sigma_{z_i}^2$).

Подставляя исходные данные в выражения (6.22), (6.23), получим

$$\hat{X} = \begin{bmatrix} \hat{\lambda}_0 \\ \hat{\lambda}_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n C_{0i} z_i \\ \frac{1}{T} \sum_{i=1}^n C_{1i} z_i \end{bmatrix}, \quad (6.24)$$

где

$$C_{0i} = \frac{2(3i - n - 1)}{n(n+1)}, \quad C_{1i} = \frac{6(2i - n - 1)}{n(n^2 - 1)} \quad \left(\sum_{i=1}^n C_{0i} = 1, \quad \sum_{i=1}^n C_{1i} = 0 \right); \quad (6.25)$$

получим также

$$\hat{\Psi} = \begin{bmatrix} \hat{\Psi}_{11} & \hat{\Psi}_{12} \\ \hat{\Psi}_{21} & \hat{\Psi}_{22} \end{bmatrix}, \quad (6.26)$$

где $\hat{\Psi}_{11} = \sigma_z^2 \frac{2(2n-1)}{n(n+1)}$, $\hat{\Psi}_{12} = \hat{\Psi}_{21} = \sigma_z^2 \frac{6}{n(n+1)T}$, $\hat{\Psi}_{22} = \sigma_z^2 \frac{12}{n(n^2-1)T^2}$. При этом $\hat{\Psi}_{11} \equiv \sigma_{\lambda_0}^2$ — дисперсия оценки ошибок траекторного параметра λ_0 (координатного члена) в момент привязки измерений $t_{\text{пр}} = t_n$, $\hat{\Psi}_{22} \equiv \sigma_{\lambda_1}^2$ — дисперсия оценки ошибок траекторного параметра λ_1 (скоростного члена) в момент привязки измерений $t_{\text{пр}} = t_n$, $\hat{\Psi}_{12} \equiv k_{\lambda_0 \lambda_1} \sigma_{\lambda_0} \sigma_{\lambda_1}$ — ковариация ошибок измерения параметров λ_0 и λ_1 , $k_{\lambda_0 \lambda_1}$ — коэффициент корреляции ошибок оценивания траекторных параметров λ_0 и λ_1 .

Зависимости нормированных безразмерных элементов ковариационной матрицы ошибок оценки параметров линейной траектории от объема выборки n показаны на рис. 6.5. Как видно из этого рисунка, для рассмотренной модели траектории движения цели и модели измерений при увеличении n уменьшаются ошибки оценки траекторных параметров. Для достижения заданной точности оценки траекторных параметров исходя из этих графиков можно найти необходимый объем выборки. При $n \rightarrow \infty$ ошибки стремятся к нулю.

Аналогичные результаты можно получить, используя соотношения (6.22) и (6.23) для любого набора параметров отсчета и траектории. Например, пусть параметры отсчета в i -й момент времени помимо измерения координаты z_{0i} включают и скорость ее изменения z_{1i} , т. е. $\mathbf{Z}_i = [z_{0i} \ z_{1i}]^T$, а точности измерения характеризуются ковариационной матрицей

$$\mathbf{R}_i = \mathbf{R} = \begin{bmatrix} \sigma_{z_0}^2 & k_{z_0 z_1} \sigma_{z_0} \sigma_{z_1} \\ k_{z_0 z_1} \sigma_{z_0} \sigma_{z_1} & \sigma_{z_1}^2 \end{bmatrix},$$

где σ_{z_0} , σ_{z_1} , $k_{z_0 z_1}$ — соответственно среднеквадратические отклонения (СКО) и коэффициент корреляции ошибок измерения координаты и скорости ее изменения.

Все измерения (отсчеты), которые после операции селекции сформированы для некоторой цели (индекс траектории для упрощения записи здесь также опускаем) в виде некоторой фиксированной выборки объемом n , запишем следующим образом:

$$\mathbf{\Omega}^{(n)} = [z_{01} \ z_{02} \ \dots \ z_{0n} \ z_{11} \ z_{12} \ \dots \ z_{1n}]^T.$$

Соответственно, ковариационная матрица $\mathbf{R}^{(n)}$ ошибок совокупности измерений при отсутствии корреляции в разные моменты времени будет иметь вид

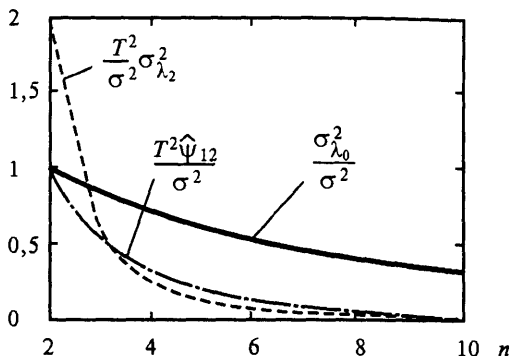


Рис. 6.5. Нормированные элементы ковариационной матрицы

$$\mathfrak{R}^{(n)} = \begin{bmatrix} \sigma_{z_0}^2 & 0 & \dots & 0 & \varphi_{z_0 z_1} & 0 & \dots & 0 \\ 0 & \sigma_{z_0}^2 & \dots & 0 & 0 & \varphi_{z_0 z_1} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_{z_0}^2 & 0 & 0 & \dots & \varphi_{z_0 z_1} \\ \varphi_{z_0 z_1} & 0 & \dots & 0 & \sigma_{z_1}^2 & 0 & \dots & 0 \\ 0 & \varphi_{z_0 z_1} & \dots & 0 & \dots & \sigma_{z_1}^2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \varphi_{z_0 z_1} & 0 & 0 & \dots & \sigma_{z_1}^2 \end{bmatrix},$$

где $\varphi_{z_0 z_1} = k_{z_0 z_1} \sigma_{z_0} \sigma_{z_1}$.

Если, например, вектор оцениваемых параметров является траекторией, описываемой полиномом второй степени, в той же физической системе координат, что и параметры отсчета, т. е. $\mathbf{X} = [\lambda_0 \lambda_1 \lambda_2]^T$, то для рассматриваемого примера матрица \mathbf{A} имеет вид

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ \tau_1 & \tau_2 & \dots & \tau_n & 1 & 1 & \dots & 1 \\ \tau_1^2/2 & \tau_2^2/2 & \dots & \tau_n^2/2 & \tau_1 & \tau_2 & \dots & \tau_n \end{bmatrix}^T.$$

Заметим, что матрицу \mathbf{A} для данного примера можно переписать в виде

$$\mathbf{A}^T = \nabla_{\mathbf{X}} \begin{bmatrix} \xi_x(\mathbf{X}, \tau_1) & \xi_x(\mathbf{X}, \tau_2) & \dots & \xi_x(\mathbf{X}, \tau_n) & \dot{\xi}_x(\mathbf{X}, \tau_1) & \dot{\xi}_x(\mathbf{X}, \tau_2) & \dots \\ & & & & \dot{\xi}_x(\mathbf{X}, \tau_n) \end{bmatrix}, \quad (6.27)$$

где $\nabla_{\mathbf{X}}$ — векторный оператор дифференцирования; $\xi_x(\mathbf{X}, \tau)$ — соотношение, описывающее траекторию движения цели; $\dot{\xi}_x(\mathbf{X}, \tau)$ — соотношение, описывающее траекторию скорости цели (см. п. 6.1.2).

Подставляя матрицы $\Omega^{(n)}$, $\mathfrak{R}^{(n)}$ и \mathbf{A} в выражения (6.22), (6.23), получим соотношения для оценки искомых траекторных параметров и их точностей. Для равнооточных равнодисcretных некоррелированных ошибок измерения параметров отсчетов в фиксированной выборке оценка $\hat{\mathbf{X}}$ и ковариационная матрица Ψ будут находиться из соотношений (6.24)—(6.26).

Ошибки оценки траекторных параметров при увеличении n также уменьшаются и при $n \rightarrow \infty$ стремятся к нулю. Полученный результат полностью объясним для выбранной детерминированной модели движения цели (6.8).

При равнооточных равнодисcretных измерениях оценка траекторных параметров, как видно из (6.24), производится нерекурсивным фильтром. Для получения таким фильтром текущей оценки на момент последнего измерения необходимо обеспечить его работу в режиме «скользящего окна», когда при поступлении нового измерения первое измерение отбрасывается, все остальные сдвигаются, новое измерение становится n -м.

В реальных условиях, рассматривая движение объекта на сравнительно небольшом временном интервале, методику нахождения оценки $\hat{\mathbf{X}}$ (6.22) и ковариационной матрицы $\hat{\Psi}$ (6.23) можно распространить и на изначально неполиномиальную траекторию, и на нелинейную связь пространства параметров отсчета с пространством параметров траектории (см. п. 6.2.2), произведя линеаризацию уравнения состояния и функции связи $\mathbf{h}(\mathbf{X})$. Тогда таким же образом, как в приведенных выше примерах, представляются параметры \mathbf{X} , $\Omega^{(n)}$, $\mathcal{R}^{(n)}$, а матрица \mathbf{A} находится в соответствии с выражением

$$\mathbf{A}^T = \nabla_{\mathbf{X}} \left[\mathbf{h}^T(\mathbf{X}(\tau_1)) \quad \mathbf{h}^T(\mathbf{X}(\tau_2)) \quad \dots \quad \mathbf{h}^T(\mathbf{X}(\tau_n)) \right],$$

после чего, подставляя матрицу \mathbf{A} в соотношения (6.22) и (6.23), получим искомую оценку траекторных параметров и соответствующую ковариационную матрицу.

Если априори не известна степень полинома, которым целесообразно описывать движение некоторого объекта наблюдения, то эту задачу можно решать следующим образом. Будем характеризовать качество полученных оценок траекторных параметров следующим квадратичным показателем:

$$J(n) = \left[\Omega^{(n)} - \mathbf{A}\hat{\mathbf{X}} \right]^T \left[\mathcal{R}^{(n)} \right]^{-1} \left[\Omega^{(n)} - \mathbf{A}\hat{\mathbf{X}} \right].$$

В случае гауссовского шума $J(n)$ имеет χ^2 -распределение с r степенями свободы ($r = na - b$, a — размерность параметров отсчета, b — размерность траекторных параметров).

Если выбранная степень полинома мала, то

$$J(n) > \chi_r^2(1 - \alpha),$$

где $\chi_r^2(1 - \alpha)$ — квантиль соответствующего χ_r^2 -распределения для вероятности ошибочного решения α , которая обычно выбирается в диапазоне от 0,01 до 0,1.

Если выбранная степень полинома слишком велика, то оценка будет статистически недостоверной. Этот факт можно установить из нарушения распределения параметров оценки компонент траекторных параметров по соответствующей физической координате, например x . Распределение каждого траекторного параметра $\hat{\lambda}_{xi}$, $i = 0, 1, \dots, s$, должно подчиняться нормальному закону $N[\lambda_{xi}, \hat{\psi}_{ii}]$, где $\hat{\psi}_{ii}$ — дисперсия оценки траекторного параметра λ_{xi} , а $N[m, \sigma^2]$ — оператор нормального закона распределения с математическим ожиданием m и дисперсией σ^2 . Если по результатам статистической проверки окажется, что величину траекторного параметра $\hat{\lambda}_{xi}$, $i > 0$, с высокой вероятностью можно считать равной нулю (верна гипотеза H_0), то лучше отказаться от оценки данной фазовой координаты.

Гипотезу H_1 (параметр λ_{xi} не равен нулю) можно принять, если

$$\frac{|\hat{\lambda}_{xi}|}{\sqrt{\hat{\psi}_{ii}}} > F\left(1 - \frac{\alpha}{2}\right),$$

где $F(1 - \alpha/2)$ — квантиль стандартного нормального распределения для вероятности ошибочного решения α .

Если гипотеза H_1 отвергается, то недостоверный траекторный параметр исключается из вектора состояний и оценки траекторных параметров находятся заново для вектора состояний меньшей размерности.

Часто необходимо предсказать (экстраполировать) параметры траектории на любой момент времени t_m , отличный от момента $t_{np} = t_k$ привязки траектории. Для полиномиальной модели траектории — уравнения состояния типа (6.8) — экстраполированную оценку траекторных параметров $\mathbf{X}_{\varepsilon m}$ можно найти с учетом линейного преобразования случайных величин аналогично (6.17):

$$\mathbf{X}_{\varepsilon m} = \mathbf{F}_m \hat{\mathbf{X}}_k, \quad (6.28)$$

где $\tau_m = t_m - t_k$ — время экстраполяции.

Можно показать, что экстраполированная ковариационная матрица в этом случае имеет вид

$$\mathbf{\Psi}_{\varepsilon m} = \mathbf{F}_m \hat{\mathbf{\Psi}}_k \mathbf{F}_m^T. \quad (6.29)$$

Оценка траекторных параметров, рассмотренная в данном параграфе, используется, как уже отмечалось, для сравнительно небольшого объема выборки, согласованного с особенностями движения объекта лоцирования и усло-

виями измерения. Чаще всего подобный подход реализуется на этапе завязки траектории (блок 3 на рис. 6.2) при небольших выборках, число элементов которых редко превышает 10.

6.4. Рекуррентная оценка траекторных параметров

6.4.1. Основные соотношения калмановской фильтрации

При рекуррентной (последовательной) оценке траекторных параметров некоторой j -й цели* их уточнение производится после поступления каждого нового k -го отсчета с выхода первичной обработки. Значения предыдущих отсчетов, как это предполагается при оценке по фиксированной выборке, не хранят, используют лишь данные о траекторных параметрах предыдущего шага. Такой подход применяется сейчас наиболее часто. Это удобно по ряду причин: при последовательной обработке точность оценки не ограничена фиксированным числом используемых данных, после поступления новых измерений с выхода первичной обработки не требуется повторения расчетов при использовании большого числа «старых» входных данных, как в режиме «скользящего окна», наконец, организация рекуррентного вычислительного процесса более удобна и естественна при наблюдении в течение неопределенного числа тактов за перемещающейся целью.

Алгоритм рекуррентной оценки в ходе своей работы уменьшает воздействие различных шумов на определяемые параметры, иначе говоря, осуществляет их фильтрацию, именно по этой причине он может рассматриваться в виде некоторого фильтра.

Рекуррентные соотношения для оценки траекторных параметров, являющиеся разновидностями формул калмановского фильтра, давно используют при построении траекторий различных объектов. Для дискретного времени уравнения калмановской фильтрации в классическом случае были получены при рассмотрении задачи оценивания состояния стохастической линейной динамической дискретной системы с уравнением состояния (6.6) по данным линейных измерений (6.15).

Априорные сведения о модели состояния системы (модели движения) и модели измерений предполагают следующие допущения [61, 62]:

— начальное состояние системы (параметры траектории в нулевой момент времени) является случайным вектором \mathbf{X}_0 с математическим ожиданием $\mathbf{M}\{\mathbf{X}_0\}$ и ковариационной матрицей Ψ_0 ;

* Индекс, показывающий принадлежность параметров траектории некоторой цели, в настоящем параграфе в дальнейшем опущен, чтобы не усложнять формулы.

— детерминированное входное воздействие (управление) u_k , если оно имеется, известно;

— случайное возмущение v_k траекторных параметров имеет характер гауссовского шума с нулевым средним и известной ковариационной матрицей Q_k ;

— ошибки измерений параметров отсчета (шум наблюдений ε_k) представляют собой процесс типа гауссовского шума с нулевым средним и известной ковариационной матрицей R_k ;

— случайные процессы v_k и ε_k взаимно не коррелированы;

— начальное состояние X_0 не коррелировано с возмущениями v_k, ε_k .

Изложенные условия образуют линейное гауссовское допущение.

Одним из наиболее распространенных подходов при выводе уравнений калмановского фильтра является использование критерия минимума среднего риска при квадратичной функции потерь. Показано (см., например, [62]), что оптимальный фильтр должен при оценивании параметров траектории в k -й момент времени вычислять условное среднее

$$\hat{X}_k = \int_{\Xi} w(X_k | \Omega^{(k)}) X_k dX_k, \quad (6.30)$$

где Ξ — область значений возможных состояний системы (область возможных траекторных параметров); $w(X_k | \Omega^{(k)})$ — многомерная условная плотность вероятности того, что траекторные параметры цели будут равны X_k , если к k -му моменту времени наблюдалась последовательность отсчетов $\Omega^{(k)}$, или, иначе говоря, апостериорная плотность вероятности того, что в k -й момент времени состояние динамической системы будет X_k , если с выхода первичной обработки получен отсчет $Z_i, i = 1, 2, \dots, k$, т. е.

$$\Omega^{(k)} = [Z_1 \quad Z_2 \quad \dots \quad Z_k]^T.$$

Используя введенное обозначение $M\{\cdot\}$ (см. п. 6.1.2) для нахождения условного среднего, можно записать

$$\hat{X}_k = M\{X_k | \Omega^{(k)}\}. \quad (6.31)$$

Ошибка оценки траекторных параметров имеет вид

$$\Delta X_k = X_k - \hat{X}_k. \quad (6.32)$$

Условная ковариационная матрица Ψ_k оценки траекторных параметров по определению вычисляется следующим образом:

$$\hat{\Psi}_k = \mathbf{M} \left\{ [\mathbf{X}_k - \hat{\mathbf{X}}_k][\mathbf{X}_k - \hat{\mathbf{X}}_k]^T \middle| \Omega^{(k)} \right\} = \mathbf{M} \{ \Delta \mathbf{X}_k \Delta \mathbf{X}_k^T \middle| \Omega^{(k)} \}. \quad (6.33)$$

Выражение (6.33) можно записать следующим образом:

$$\hat{\Psi}_k = \int_{\Xi} w(\mathbf{X}_k | \Omega^{(k)}) [\mathbf{X}_k - \mathbf{M}\{\mathbf{X}_k | \Omega^{(k)}\}] [\mathbf{X}_k - \mathbf{M}\{\mathbf{X}_k | \Omega^{(k)}\}]^T d\mathbf{X}_k. \quad (6.34)$$

Выражения (6.30) (или (6.31)) и (6.33) (или (6.34)) являются исходными для синтеза алгоритма оценивания траекторных параметров $\hat{\mathbf{X}}_k$ и соответствующей ковариационной матрицы $\hat{\Psi}_k$.

При сделанных выше предположениях были получены (подробный вывод формул калмановской фильтрации можно найти, например, в работах [55, 70]) рекуррентные соотношения для оценивания состояния $\hat{\mathbf{X}}_k$. Они позволяют вычислять оценку траекторных параметров $\hat{\mathbf{X}}_k$ на текущем k -м шаге с учетом текущих измерений параметров отсчета \mathbf{Z}_k по результатам оценивания траекторных параметров $\hat{\mathbf{X}}_{k-1}$ на предыдущем $(k-1)$ -м шаге:

$$\hat{\mathbf{X}}_k = \mathbf{X}_{\text{эк}} + \mathbf{W}_k \Delta \mathbf{Z}_k, \quad (6.35)$$

где $\mathbf{X}_{\text{эк}}$ — экстраполированная оценка состояния системы на k -й шаг; \mathbf{W}_k — коэффициент усиления фильтра; $\Delta \mathbf{Z}_k$ — невязка измерений.

Экстраполированная оценка состояния (экстраполированные параметры траектории) для линейной модели движения (см. п. 6.2.1) находится следующим образом:

$$\mathbf{X}_{\text{эк}} = \mathbf{F}_k \hat{\mathbf{X}}_{k-1} + \mathbf{G}_k \mathbf{u}_{k-1}. \quad (6.36)$$

Невязка измерений есть разница между полученными параметрами отсчета \mathbf{Z}_k на k -м шаге и экстраполированными параметрами отсчета $\mathbf{Z}_{\text{эк}}$:

$$\Delta \mathbf{Z}_k = \mathbf{Z}_k - \mathbf{Z}_{\text{эк}}. \quad (6.37)$$

Для рассматриваемой линейной модели измерений экстраполированные параметры отсчета связаны с экстраполированными параметрами траектории следующим образом:

$$\mathbf{Z}_{\text{эк}} = \mathbf{H}_k \mathbf{X}_{\text{эк}}. \quad (6.38)$$

Коэффициент усиления фильтра находится из соотношения

$$\mathbf{W}_k = \Psi_{\text{эк}} \mathbf{H}_k^T \mathbf{S}_{\text{эк}}^{-1}, \quad (6.39)$$

где $\Psi_{\varepsilon k}$ — экстраполированная на k -й шаг ковариационная матрица оценки траекторных параметров на k -м шаге:

$$\Psi_{\varepsilon k} = \mathbf{F}_k \hat{\Psi}_{k-1} \mathbf{F}_k^T + \mathbf{F}_k \mathbf{Q}_{k-1} \mathbf{F}_k^T; \quad (6.40)$$

$\mathbf{S}_{\varepsilon k}$ — экстраполированная ковариационная матрица параметров отсчета:

$$\mathbf{S}_{\varepsilon k} = \mathbf{H}_k \Psi_{\varepsilon k} \mathbf{H}_k^T + \mathbf{R}_k. \quad (6.41)$$

Выражения (6.35)—(6.41) определяют рекуррентный алгоритм вычисления оценки состояния системы — алгоритм оценки траекторных параметров.

Ковариационная матрица оценки траекторных параметров в k -й момент вычисляется следующим образом:

$$\hat{\Psi}_k = [\mathbf{I} - \mathbf{W}_k \mathbf{H}_k] \Psi_{\varepsilon k}, \quad (6.42)$$

где \mathbf{I} — единичная матрица.

Выражение (6.42) также может быть записано в тождественной форме:

$$\hat{\Psi}_k = \Psi_{\varepsilon k} - \mathbf{W}_k \mathbf{S}_{\varepsilon k} \mathbf{W}_k^T.$$

Ковариационную матрицу оценки траекторных параметров лучше вычислять, используя *форму Джозефа* [63], которая менее чувствительна к округлению ошибок и не приводит к отрицательным собственным значениям:

$$\hat{\Psi}_k = [\mathbf{I} - \mathbf{W}_k \mathbf{H}_k] \Psi_{\varepsilon k} [\mathbf{I} - \mathbf{W}_k \mathbf{H}_k]^T + \mathbf{W}_k \mathbf{R}_k \mathbf{W}_k^T. \quad (6.43)$$

Иногда уравнение (6.35) называется *уравнением обновленного состояния*, а уравнение (6.42) — *уравнением обновленной ковариации*. Эти уравнения определяют операцию калмановской фильтрации при построении траектории сопровождаемой цели.

В процессе фильтрации важную роль играет величина \mathbf{W}_k — коэффициент усиления фильтра. Из выражения (6.39) видно, что он пропорционален экстраполированной ковариационной матрице траекторных параметров $\Psi_{\varepsilon k}$ и обратно пропорционален экстраполированной ковариационной матрице параметров отсчета $\mathbf{S}_{\varepsilon k}$.

Таким образом, коэффициент усиления будет «большим», если прогноз (экстраполяция) параметров траектории является «грубым», а текущие измерения параметров отсчетов, поступившие с выхода первичной обработки, являются «точными», т. е. текущие измерения в этом случае будут учитываться в большей степени, чем оценки параметров траектории в предыдущий момент времени. Коэффициент усиления будет «малым», если прогноз параметров траектории является «точным», а текущие измерения параметров отсчета являются «грубыми».

Итак, последовательность операций при оценке траекторных параметров методами калмановской фильтрации (при условии, что все другие операции в соответствии со схемой, приведенной на рис. 6.2, выполнены) можно представить следующим образом.

До начала работы фильтра (априори), исходя из условий задачи, задаются:

- все компоненты уравнения (6.6) динамического состояния системы (модели движения объекта), т. е. матрицы F_k , G_k , Γ_k и ковариационная матрица Q_k ;

- все компоненты уравнения (6.12) измерений параметров отсчета — матрица H_k и ковариационная матрица R_k .

На предыдущем $(k-1)$ -м шаге необходимо вычислить или определить:

- оценку состояния \hat{X}_{k-1} системы (траекторные параметры) на $(k-1)$ -м шаге;

- ковариационную матрицу $\hat{\Psi}_{k-1}$ оценки состояния системы (ковариационную матрицу оценки траекторных параметров) на $(k-1)$ -м шаге.

На текущем k -м шаге вычисляют:

- экстраполированную оценку $X_{\varepsilon k}$ состояния (экстраполированные параметры траектории) по формуле (6.36);

- экстраполированную ковариационную матрицу $\Psi_{\varepsilon k}$ оценки состояния (ковариационную матрицу экстраполированных траекторных параметров) по формуле (6.40);

- экстраполированные (прогнозируемые) измерения $Z_{\varepsilon k}$ по формуле (6.38);

- экстраполированную ковариационную матрицу $S_{\varepsilon k}$ параметров отсчета по формуле (6.41);

- невязку измерений ΔZ_k по формуле (6.37);

- коэффициент усиления фильтра W_k по формуле (6.39);

- ковариационную матрицу $\hat{\Psi}_k$ состояния системы (ковариационную матрицу оценки траекторных параметров) на текущем k -м шаге по формуле (6.42);

- оценку состояния \hat{X}_k системы (оценку траекторных параметров) на текущем k -м шаге по формуле (6.35).

Для выполнения перечисленных операций необходимо определить параметры стартовой точки: \hat{X}_0 и $\hat{\Psi}_0$. Если значения выбранных параметров сильно отличаются от истинного, а выбранная ковариационная матрица этого не учитывает, то при малых значениях элементов ковариационной матрицы, а следовательно, при малых размерах строба сопровождения (см.

п. 6.1.2) может произойти срыв сопровождения. Если предположить, что стартовая точка выбрана с большой ошибкой, то, во-первых, время сходимости фильтра будет продолжительным, во-вторых, сильно увеличатся размеры строга сопровождения, а значит, возрастет вероятность появления в нем ложных отсчетов, что также может привести к срыву сопровождения.

Ошибка начальной оценки состояния должна быть *согласована* с начальной ковариационной матрицей. При выборе начальной ковариационной матрицы необходимо, чтобы ошибки по соответствующей координате по крайней мере в 2 раза превышали среднеквадратическое отклонение, в этом случае фильтр сходится достаточно быстро [63].

На практике выбор стартовой точки может быть сделан с использованием нескольких последовательных (обычно не более 2–4) измерений местоположения, как при построении траекторий по фиксированной выборке (см. § 6.3).

Например, если параметры отсчета представляют собой скаляр, измеряется только одна из координат z цели, а в качестве траекторных параметров оцениваются эта координата и скорость ее изменения, то из (6.22) и (6.23) по двум измерениям можно получить:

$$\hat{\mathbf{X}}_0 = \begin{bmatrix} \frac{z_{-1} + z_0}{2} \\ \frac{z_0 - z_{-1}}{T} \end{bmatrix}, \quad (6.44)$$

$$\hat{\Psi}_0 = \begin{bmatrix} \sigma^2 & \sigma^2 / T \\ \sigma^2 / T & 2\sigma^2 / T \end{bmatrix}, \quad (6.45)$$

где T — период измерений, σ^2 — дисперсия ошибок измерения дальности z .

6.4.2. Стационарный режим калмановского фильтра

Представляет интерес исследование калмановского фильтра при неограниченном увеличении интервала наблюдений. Важно знать, существуют ли предельные значения коэффициента усиления фильтра, ковариационных матриц ошибок измерения, экстраполяции, и если существуют, то при каких условиях они не зависят от начального состояния.

Предположим, что модели движения объекта и измерений параметров отсчетов инвариантны во времени, т. е. матрицы \mathbf{F} , \mathbf{G} , \mathbf{H} от времени не зависят. Предположим также, что возмущающее воздействие \mathbf{v} на парамет-

ры траектории и шум измерений ϵ параметров отсчетов являются стационарными в широком смысле, т. е. их ковариационные матрицы \mathbf{Q} и \mathbf{R} не зависят от времени. Даже в таких условиях сохраняется зависимость коэффициента усиления от времени, что обусловлено, в первую очередь, эволюцией ковариационной матрицы Ψ_{3k} .

Рекуррентное выражение для Ψ_{3k} при экстраполяции на один шаг имеет вид

$$\begin{aligned} \Psi_{3(k+1)} = & \mathbf{F}_{k+1} \left\{ \Psi_{3k} - \Psi_{3k} \mathbf{H}_k^T \left[\mathbf{H}_k \Psi_{3k} \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1} \times \right. \\ & \left. \times \mathbf{H}_k \Psi_{3k} \right\} \mathbf{F}_{k+1}^T + \mathbf{G}_{k+1} \mathbf{Q}_k \mathbf{G}_{k+1}^T. \end{aligned} \quad (6.46)$$

Уравнение (6.46) является *разностным матричным уравнением Рикатти* [63].

Ковариационная матрица экстраполяции при $k \rightarrow \infty$ принимает предельное значение Ψ_3 , являющееся решением алгебраического уравнения Рикатти:

$$\Psi_3 = \mathbf{F} \left[\Psi_3 - \Psi_3 \mathbf{H}^T \left\{ (\mathbf{H} \Psi_3 \mathbf{H}^T + \mathbf{R})^{-1} \Psi_3 \mathbf{H}^T \right\} \right] \mathbf{F}^T + \mathbf{G} \mathbf{Q} \mathbf{G}^T. \quad (6.47)$$

Выражение (6.47) справедливо, если при положительно определенной начальной ковариационной матрице траекторных параметров ($\hat{\Psi}_0 > 0$) модель движения цели и модель измерений предотвращают неопределенность состояния траекторных параметров и их оценок*.

Уравнение Рикатти описывает появление установившегося режима в фильтре Калмана. Установившийся, или стационарный, режим фильтра характеризует потенциальные возможности фильтрации траекторных параметров в ходе вторичной обработки.

6.4.3. Состоятельность калмановского фильтра

Калмановский фильтр, работающий в полном соответствии с условиями, для которых он был синтезирован (см. п. 6.4.1), позволяет получить несмещенную и эффективную оценку траекторных параметров [61]. Однако на практике при некоторых отступлениях от этих условий можно столкнуться с явлением *расходимости фильтра*. Под этим явлением понимают ситуацию, когда средняя квадратическая ошибка оценивания траекторных

* Более строго эти условия формулируют в теории систем с использованием понятий наблюдаемости, управляемости и контролируемости (см., например, [62, 69]).

параметров намного превышает величину самой оценки и значение средней квадратической ошибки со временем неограниченно возрастает. Расходимость фильтра может возникнуть в силу:

1) заметного несоответствия модели состояния динамической системы реальному процессу движения цели и (или) модели измерений — реальной помеховой остановке, а также неправильного выбора параметров стартовой точки;

2) ошибок в процессе численных вычислений (особенно при матричных вычислениях) из-за недостаточной точности представления чисел в ходе практической реализации алгоритма калмановской фильтрации на средствах вычислительной техники;

3) ошибок при выполнении операции селекции в условиях сложной помеховой и целевой обстановки.

Из анализа соотношений, определяющих работу калмановского фильтра, видно, что перечисленные факторы могут привести к возрастанию величины невязки измерений (6.37) до такого значения, при котором она становится несогласованной с коэффициентом усиления фильтра (6.39), в результате чего оценка (6.35) траекторных параметров не корректируется регулярно до нужных значений. Поскольку почти всегда любая модель содержит некоторые аппроксимации и погрешности, а численные расчеты выполняются с конечной точностью, вопрос о приемлемых текущих ошибках оценивания траекторных параметров, означающих отсутствие расходимости фильтра, является одним из важнейших при разработке алгоритмов ВО. Для характеристики с этих позиций работоспособности калмановского фильтра вводится понятие его состоятельности [63], несколько отличающееся от традиционного.

Состоятельной оценкой некоторого постоянного параметра называют оценку, которая сходится к истинному значению рассматриваемого параметра при увеличении числа наблюдений. Соответственно, алгоритм, производящий такое оценивание, можно назвать *состоятельным*. *Состоятельным калмановским фильтром* называют фильтр, в котором первый и второй моменты ошибок получаемых оценок траекторных параметров соответствуют теоретически предсказанным, т. е. средняя ошибка равна нулю (оценка является несмещенной), а второй момент соответствует ковариационной матрице ошибок, вычисляемой фильтром.

Для проверки состоятельности калмановского фильтра используется тот факт, что при линейном гауссовском допущении условная плотность распределения $w(X_k | \Omega^{(k)})$ траекторных параметров должна подчиняться нормальному закону распределения с параметрами \hat{X}_k и $\hat{\Psi}_k$, вычисляемыми фильтром в некоторый k -й момент времени:

$$w(\mathbf{X}_k | \Omega^{(k)}) = N[\mathbf{X}_k, \hat{\mathbf{X}}_k, \hat{\Psi}_k], \quad (6.48)$$

где $N[\mathbf{X}_k, \hat{\mathbf{X}}_k, \hat{\Psi}_k]$ — условная запись нормального закона распределения величины \mathbf{X}_k с математическим ожиданием $\hat{\mathbf{X}}_k$ и ковариационной матрицей $\hat{\Psi}_k$. Поэтому одним из способов проверки состоятельности фильтра является проверка справедливости выражения (6.48).

Статистические характеристики распределения можно задавать его моментами, причем для гауссовского распределения достаточно двух моментов. Тогда выражение (6.48) эквивалентно условиям

$$\mathbf{M}\{\mathbf{X}_k - \hat{\mathbf{X}}_k\} = \mathbf{M}\{\Delta\mathbf{X}_k\} = 0, \quad (6.49)$$

$$\mathbf{M}\{(\mathbf{X}_k - \hat{\mathbf{X}}_k)(\mathbf{X}_k - \hat{\mathbf{X}}_k)^\top\} = \hat{\Psi}_k, \quad (6.50)$$

которым должен удовлетворять фильтр.

Условие (6.49) является требованием несмещенности для оценок траекторных параметров (т. е. нулевой средней ошибки оценивания). Условие (6.50) — это требование согласованности моментов, т. е. того, что действительная ковариационная матрица (левая часть) равна ковариационной матрице, вычисленной фильтром (правая часть).

В качестве статистики, которую можно применить для проверки соотношений (6.49), (6.50), целесообразно использовать нормированный квадрат ошибки оценки траекторных параметров

$$\delta^2(\mathbf{X}) = \Delta\mathbf{X}_k^\top \hat{\Psi}_k^{-1} \Delta\mathbf{X}_k. \quad (6.51)$$

Если вычисленная фильтром ковариационная матрица $\hat{\Psi}_k$ соответствует оценкам траекторных параметров и действительной ковариационной матрице, то при гауссовском допущении эта статистика подчиняется χ^2 -распределению с b степенями свободы (b — размерность вектора траекторных параметров). При попадании этой статистики в заданный доверительный интервал анализируемый калмановский фильтр можно полагать состоятельным. Заметим, что статистика $\delta^2(\mathbf{X})$ чувствительна как к невыполнению условия (6.49), так и (6.50).

Тест на основе статистики $\delta^2(\mathbf{X})$ можно использовать при проверке состоятельности фильтра по результатам статистических испытаний. Пусть после моделирования с помощью метода Монте-Карло получена выборка M независимых случайных величин $\delta^2(\mathbf{X})_i$, $i = 1, 2, \dots, M$. Среднее значение δ^2 -статистики имеет вид

$$\bar{\delta}^2(\mathbf{X}) = \frac{1}{M} \sum_{i=1}^M \delta^2(\mathbf{X})_i. \quad (6.55)$$

Тогда величина $M\bar{\delta}^2(\mathbf{X})$ будет иметь χ^2 -распределение с $r = Mb$ степенями свободы.

Гипотеза о состоятельности фильтра принимается, если

$$M\bar{\delta}^2(\mathbf{X}) \in [c_1, c_2], \quad (6.56)$$

где границы двухстороннего доверительного интервала определяются квантилями соответствующего χ^2 -распределения для допустимой вероятности ошибки α :

$$c_1 = \chi_r^2\left(\frac{\alpha}{2}\right), \quad c_2 = \chi_r^2\left(1 - \frac{\alpha}{2}\right). \quad (6.57)$$

Например, для доверительной вероятности $P_{\text{дов}} = 0,95$, где $P_{\text{дов}} = 1 - \alpha$, $b = 2$, $M = 50$ имеем $c_1 = 74,2$ и $c_2 = 130$.

Условие (6.56) может не выполняться в силу различных причин, например из-за наличия смещения оценок (невыполнения условия (6.49)). Дополнительную проверку на наличие смещения можно провести, используя среднюю ошибку оценки траекторных параметров, нормированную к соответствующему значению среднеквадратической ошибки. При выполнении условий (6.49), (6.50) распределение нормированной средней ошибки должно подчиняться нормальному закону с нулевым математическим ожиданием и единичной дисперсией. Тестируя полученную после статистического моделирования последовательность ошибок оценок траекторных параметров, легко принять решение о наличии или отсутствии смещения.

Всеми указанными выше свойствами обладает невязка (см. п. 6.4.1)

$$\Delta \mathbf{Z}_k = \mathbf{Z}_k - \mathbf{H}_k \hat{\mathbf{X}}_k. \quad (6.58)$$

Квадрат нормированной невязки определяется выражением

$$\delta^2(\mathbf{Z}) = \Delta \mathbf{Z}_k^T \mathbf{S}_{\mathbf{Z}k}^{-1} \Delta \mathbf{Z}_k, \quad (6.59)$$

где $\mathbf{S}_{\mathbf{Z}k}$ — экстраполированная ковариационная матрица параметров отсчета.

Для гипотез, при которых фильтр является состоятельным, квадрат нормированной невязки имеет χ^2 -распределение с a степенями свободы, где a — размерность вектора параметров отсчета. После M независимых испы-

таний величина $M\bar{\delta}^2(X)$ тестируется аналогичным образом, исходя из χ^2 -распределения с Ma степенями свободы.

Заметим, что тестирование с использованием нормированного квадрата ошибки оценки траекторных параметров можно проводить только по результатам статистического моделирования, а тестирование с использованием нормированного квадрата невязки можно проводить и в реальном времени в процессе работы фильтра.

6.4.4. $(\alpha\text{-}\beta)$ -фильтры

В случае равнооточных ($R_k = R$) равнодискретных ($(t_k - t_{k-1}) = T$ для любых k) измерений параметров отсчета при неискаженном движении объекта ($Q = 0$) уравнения калмановской фильтрации существенно упрощаются.

При измерении в момент k на этапе первичной обработки одной обобщенной координаты z_k (z_k — скаляр) в случае, если оцениваются дальность $\lambda_{z0,k}$ и скорость $\lambda_{z1,k}$ цели, соответствующие оценки траекторных параметров после преобразования и упрощения соотношений (6.35)—(6.41) могут быть записаны в следующем виде:

$$\hat{X}_k = \begin{bmatrix} \hat{\lambda}_{z0,k} \\ \hat{\lambda}_{z1,k} \end{bmatrix} = \begin{bmatrix} \lambda_{z0,k} + \alpha_k (z_k - \lambda_{z0,k}) \\ \lambda_{z1,k} + \frac{\beta_k}{T} (z_k - \lambda_{z0,k}) \end{bmatrix}, \quad (6.60)$$

где

$$\alpha_k = \frac{2(2k-1)}{k(k+1)}, \quad \beta_k = \frac{6}{k(k+1)}. \quad (6.61)$$

Соотношения типа (6.60), (6.61) получили название $(\alpha\text{-}\beta)$ -фильтров в силу коэффициентов α_k и β_k (коэффициентов усиления), учитывающих влияние невязки измерений $\Delta z_k = z_k - \lambda_{z0,k}$ при коррекции экстраполированных параметров траектории $\lambda_{z0,k}$ и $\lambda_{z1,k}$. Для оценки скорости движения цели и ускорения применяют $(\alpha\text{-}\beta\text{-}\gamma)$ -фильтры. Их структура аналогична структуре $(\alpha\text{-}\beta)$ -фильтров.

Область применимости соотношений (6.60), (6.61) достаточно ограничена вследствие того, что коэффициенты α_k и β_k в данном случае с увеличением k стремятся к нулю. При фильтрации результаты последних измерений учитываются всё с меньшим весом и, наконец, фильтр перестает реагировать на изменения входного сигнала. Фильтр становится несостоятельным и расходится.

По указанным причинам фильтр, для которого оценки описываются соотношениями (6.60) и (6.61), используется без каких-либо дополнительных условий в радиолокационной практике только в ограниченных случаях, когда k не велико. Однако простота реализации $(\alpha-\beta)$ -фильтра и небольшие требования к вычислительным мощностям обусловили большой интерес к фильтрам с подобной структурой. Для решения практических задач был проведен значительный объем исследований $(\alpha-\beta)$ -фильтров для зашумленных кинематических моделей движения целей, прежде всего, для случаев (6.11), когда ускорение описывается белым шумом, а также когда ускорение описывается процессом Винера [55—61]. Естественно, коэффициенты α_k и β_k в этом случае рассчитываются по более сложным, чем формула (6.61), соотношениям.

Во многих случаях параметры $(\alpha-\beta)$ -фильтров могут быть получены в результате анализа их стационарных режимов. Как указывалось в п. 6.4.2, оценка и ковариационная матрица ошибок траекторных параметров для систем с постоянными коэффициентами в уравнениях, описывающих процесс движения целей и измерения параметров отсчета, будут сходиться при определенных условиях (выполняющихся для рассматриваемых моделей) к установившемуся значению. Это позволяет получить точные значения для ковариационной матрицы и коэффициента усиления фильтра и использовать их при нахождении параметров $(\alpha-\beta)$ -фильтров.

Например, пусть для модели траектории (6.11) и модели измерений (6.13) установившиеся компоненты ковариационной матрицы экстраполяции определяются соотношением (см. п. 6.4.2):

$$\lim_{k \rightarrow \infty} \Psi_{\Delta k} = [\Psi_{\Delta ij}], \quad i, j = 1, 2, \dots, b, \quad (6.62)$$

где b — размерность вектора параметров траектории. Тогда для оценивания дальности и скорости движения цели (как при выводе соотношений (6.60)) получим [61]

$$\begin{bmatrix} \alpha \\ \beta \\ \frac{\beta}{T} \end{bmatrix} = \begin{bmatrix} \frac{\Psi_{\Delta 11}}{\Psi_{\Delta 11} + \sigma_z^2} \\ \alpha \frac{\Psi_{\Delta 12}}{\Psi_{\Delta 11}} \end{bmatrix}, \quad (6.63)$$

где σ_z^2 — дисперсия ошибок измерения отсчета z .

Аналогично можно найти коэффициенты и для $(\alpha-\beta-\gamma)$ -фильтров, а также для других видов шумов измерений параметров отсчета.

Заметим, что $(\alpha-\beta)$ - и $(\alpha-\beta-\gamma)$ -фильтры являются простейшими из возможных фильтров. Они используют фиксированные или предварительно

вычисленные коэффициенты усиления. Естественно, они не являются оптимальными в течение переходного периода (в начале сопровождения), а также в случае, если шумы возмущения нестационарны. Поэтому без специальных мер коррекции рассмотренные фильтры мало пригодны для использования в автоматизированных системах обработки радиолокационной информации. Однако, как уже указывалось, они могут применяться для реализации алгоритмов сопровождения в реальном времени, где получаемые характеристики оценки траекторных параметров удовлетворяют потребностей информации.

6.4.5. Расширенный фильтр Калмана

В случае линейных моделей движения целей (6.6) и измерений параметров отсчетов (6.15) для возмущений параметров траектории гауссовского типа и гауссовских ошибок измерения параметров отсчета, как показано в п. 6.4.1, рекуррентный процесс вычисления оценок траекторных параметров реализует калмановский фильтр. Однако в практике радиолокации такая ситуация, строго говоря, является скорее исключением, чем правилом. Одна из причин этого — несовпадение пространства параметров отсчетов и траекторий. Действительно, измерения на этапе первичной обработки производятся обычно в радиолокационной системе координат, которая является полярной, а оценку траекторных параметров желательно выполнять в другой системе координат, чаще всего — в декартовой. В результате уже в простейшем случае, а тем более при наблюдении объектов со специальными видами маневра, связь параметров отсчетов с фильтруемыми параметрами траекторий становится нелинейной.

В этой ситуации можно применить синтез алгоритмов на основании теории нелинейной фильтрации. Однако эта теория в настоящее время окончательно не разработана, а известные нелинейные алгоритмы крайне сложны в реализации [61]. Поэтому в инженерной практике вместо оптимальных находят широкое применение субоптимальные алгоритмы с возможно более полным использованием подходов и структур, разработанных для линейных фильтров Калмана. К числу таких алгоритмов относится расширенный фильтр Калмана, который представляет субоптимальный нелинейный алгоритм.

Расширенный фильтр Калмана первого порядка основан на:

- линеаризации нелинейностей в уравнении модели процесса движения целей (6.5) и модели измерений параметров отсчетов (6.14);
- оценивании траекторных параметров так же, как и в линейном фильтре, на основе критерия минимума среднего риска при квадратичной функции потерь.

В расширенном фильтре Калмана второго порядка используется дополнительно второй член ряда в разложении соответствующих нелинейных функций в уравнениях (6.5) и (6.14).

Рассмотрим нелинейную модель движения объекта в виде

$$\mathbf{X}_k = \mathbf{f}_k(\mathbf{X}_{k-1}) + \gamma_k(\mathbf{X}_{k-1})\mathbf{v}_{k-1}, \quad (6.64)$$

где $\mathbf{f}_k(\cdot)$, $\gamma_k(\cdot)$ — матрицы нелинейных функций, зависящих от дискретного времени и дискретного вектора траекторных параметров соответствующих размерностей; модель измерений параметров отсчета рассмотрим в виде

$$\mathbf{Z}_k = \mathbf{h}_k(\mathbf{X}_k) + \varepsilon_k, \quad (6.65)$$

где $\mathbf{h}_k(\cdot)$ — матрица нелинейных, зависящих от дискретного времени функций связи пространства параметров отсчета \mathbf{Z}_k и траекторных параметров \mathbf{X}_k . Характеристики шумов возмущения траекторий \mathbf{v}_k и шумов измерения параметров отсчета ε_k предполагаются такими же, как в п. 6.4.1.

Из сравнения соотношений (6.64) и (6.65) с формулами (6.6) и (6.14) видно, что приведенные здесь модели не являются общими, однако опыт показывает, что с их помощью можно описать большинство задач, встречающихся в радиолокационной практике.

В расширенном фильтре Калмана начальное состояние $\hat{\mathbf{X}}_0$ и связанная с ним ковариационная матрица $\hat{\Psi}_0$ выбираются так же, как в п. 6.4.1.

Аналогично линейному случаю модели состояния систем и модели измерений считаем, что оценка в момент времени k есть приблизительно условное среднее:

$$\hat{\mathbf{X}}_k \approx \mathbf{M}\{\mathbf{X}_k | \Omega^{(k)}\}. \quad (6.66)$$

Эта оценка характеризуется соответствующей ковариационной матрицей $\hat{\Psi}_k$. Строго говоря, матрица $\hat{\Psi}_k$ не является истинной ковариацией оценки, а является матрицей, удобной для построения фильтра.

В рассмотренном фильтре Калмана нелинейные функции считаются достаточно гладкими, что позволяет разложить их в ряд Тейлора и аппроксимировать членами ряда невысоких порядков (чаще всего первого). Введем матрицы:

$$\mathbf{F}_k^T = \nabla_{\mathbf{X}} [\mathbf{f}_k(\mathbf{X})^T]_{\mathbf{X}=\mathbf{X}_0}, \quad (6.67)$$

$$\Gamma_k = [\gamma_k(\mathbf{X})^\top]_{\mathbf{X}=\mathbf{X}_0}, \quad (6.68)$$

$$\mathbf{H}_k^\top = \nabla_{\mathbf{X}} [\mathbf{h}_k(\mathbf{X})^\top]_{\mathbf{X}=\mathbf{X}_1}. \quad (6.69)$$

Заметим, что в выражениях (6.67)—(6.69) используются разложения функций \mathbf{F} , \mathbf{G} , \mathbf{H} в окрестности некоторой точки \mathbf{X} . Выбор точки разложения должен быть произведен таким образом, чтобы линейная аппроксимация соответствующих функций была достаточно корректной. Необходимо выбирать величину \mathbf{X} так, чтобы она была достаточно близкой к истинному значению \mathbf{X}_k . Обычно предполагается, что если аппроксимация допустима в окрестности истинного значения \mathbf{X}_k , то она допустима и в окрестности оценки этого вектора. Учитывая изложенное выше, в уравнении движения цели для матриц \mathbf{F}_k , \mathbf{G}_k целесообразно выбрать значение $\mathbf{X}_0 = \hat{\mathbf{X}}_{k-1}$, а в уравнении измерения параметров отсчета при вычислении матрицы \mathbf{H}_k — экстраполированное значение $\mathbf{X}_1 = \mathbf{X}_{\text{эк}}$.

С учетом сделанных допущений уравнения расширенного фильтра Калмана после поступления в k -й момент времени вновь полученных параметров отсчета \mathbf{Z}_k будут иметь вид

$$\mathbf{X}_{\text{эк}} = \mathbf{f}_k(\hat{\mathbf{X}}_{k-1}), \quad (6.70)$$

$$\Psi_{\text{эк}} = \mathbf{F}_k \hat{\Psi}_{k-1} \mathbf{F}_k^\top + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^\top, \quad (6.71)$$

$$\mathbf{Z}_{\text{эк}} = \mathbf{h}_k(\mathbf{X}_{\text{эк}}), \quad (6.72)$$

$$\mathbf{S}_{\text{эк}} = \mathbf{H}_k \Psi_{\text{эк}} \mathbf{H}_k^\top + \mathbf{R}_k, \quad (6.73)$$

$$\Delta \mathbf{Z}_k = \mathbf{Z}_k - \mathbf{Z}_{\text{эк}}, \quad (6.74)$$

$$\mathbf{W}_k = \Psi_{\text{эк}} \mathbf{H}_k^\top \mathbf{S}_k^{-1}, \quad (6.75)$$

$$\hat{\Psi}_k = [\mathbf{I} - \mathbf{W}_k \mathbf{H}_k] \Psi_{\text{эк}}, \quad (6.76)$$

$$\hat{\mathbf{X}}_k = \mathbf{X}_{\text{эк}} + \mathbf{W}_k \Delta \mathbf{Z}_k. \quad (6.77)$$

Расширенный калмановский фильтр является нелинейным, поскольку в его соотношения вводятся так или иначе сведения о предполагаемых параметрах траектории и отсчета, которые образуют цепь обратной связи.

Использование разложения в ряд при экстраполяции параметров траектории и экстраполяции параметров отсчета может привести к появлению непредусмотренных ошибок. В результате этого ошибки экстраполяции могут иметь смещенное среднее. При вычислении ковариационных матриц также возникают ошибки. Поэтому необходимо особое внимание к контролю возможной расходимости фильтра и проверке его на состоятельность.

Используемые тесты будут такими же, как и в случае обычного фильтра Калмана (см. п. 6.4.4).

Если фильтр оказывается несостоятельным, то для устранения этого можно применить ряд эвристических методов. Так, при экстраполяции целесообразно добавлять к ковариационной матрице шумов возмущения траектории некоторую положительно определенную матрицу Q'_k и использовать вместо Q_k матрицу

$$Q_{\Xi k} = Q_k + Q'_k. \quad (6.78)$$

Это увеличивает коэффициент усиления, что позволяет лучше учитывать данные последних измерений.

Устранить несостоятельность фильтра можно также с помощью умножения ковариационной матрицы экстраполированной ошибки на некоторую величину $\varphi > 1$ в каждом периоде выборки. Если указанные меры значительно ухудшают точности получаемых оценок траекторных параметров, то целесообразно использовать другие модели движения объекта и измерений, более адекватные реальной ситуации.

В некоторых случаях при описании движения целей нелинейными дифференциальными уравнениями экстраполяцию состояния можно получить численным интегрированием динамического уравнения. Возможны и другие методы устранения несостоятельности фильтра, более строго учитывающие возможные маневры цели [63].

Для оценки работоспособности расширенного фильтра Калмана и его характеристик необходимо использовать метод статистического моделирования. Только с его помощью при проверке состоятельности фильтра можно получить реальные границы использования расширенного фильтра Калмана.

6.5. Селекция отсчетов

В ходе операции селекции, как указывалось в § 6.1, осуществляется отбор совокупностей отсчет—траектория, относящихся к соответствующим целям. От выполнения этой операции зависит качество работы всего алгоритма ВО, и в ней, в свою очередь, учитываются результаты других операций ВО (а также качество выполнения первичной обработки радиолокационной информации).

Операция селекции, в основе которой лежит принцип отбора отсчета на k -м такте наблюдения, с наибольшей вероятностью относящемуся к прогнозируемому положению цели в этот же момент времени, оказывается осложнена рядом обстоятельств. Так, из-за ошибок измерения параметры по-

лученного отсчета никогда не совпадают с прогнозируемым положением цели, тоже определяемым с некоторыми ошибками. Поскольку целевой отсчет обнаруживается с вероятностью, меньшей единицы, в текущем такте он вообще может отсутствовать, а при наличии шумов может появиться ложный отсчет. В многоцелевой ситуации может также произойти перепутывание отсчетов от соседних целей.

При решении задачи селекции в процессе отбора отсчета, относящегося к траектории некоторой цели, в k -й момент времени так или иначе анализируется величина рассогласования экстраполированной оценки положения j -й цели (j -й траектории) с положением i -го отсчета, т. е. невязка измерений $\Delta Z_{jik} = \Delta Z_{ji}(t_k)$ (см. п. 6.4.1). В ходе операции селекции проверяется гипотеза о согласованности или несогласованности величины ΔZ_{jik} с характеристиками возможных ошибок.

Предположим, что измерения параметров отсчета Z_{ik} в k -й момент времени и ошибки экстраполяции — случайные процессы с ковариационными матрицами R_{ik} и Ψ_{jik} соответственно. Для определенности считаем, что рассматривается базовый случай модели движения объекта, когда возмущения v вызывают случайные ускорения, задаваемые гауссовским процессом с ковариационной матрицей Q_{jk} . При сделанных предположениях невязка ΔZ_{jik} является случайной величиной с нулевым математическим ожиданием и ковариационной матрицей

$$S_{jik} = H_k \Psi_{jik} H_k^T + R_{ik}, \quad (6.79)$$

где в соответствии с формулой (6.40)

$$\Psi_{jik} = F_k \hat{\Psi}_{j(k-1)} F_k^T + \Gamma_k Q_{j(k-1)} \Gamma_k^T. \quad (6.80)$$

Обычно селекция выполняется в несколько этапов. На первом этапе производится отбраковка отсчетов, которые заведомо не могут относиться к цели с траекторией X_j ; на последующих этапах осуществляется окончательная селекция, т. е. собственно отбор отсчета, с высокой вероятностью относящегося к траектории рассматриваемой цели.

На первом этапе, как обычно при проверке статистических гипотез, задается допустимая вероятность отбраковки правильного отсчета из-за слишком большого отклонения ΔZ_{jik} и в соответствии с теми или иными критериями согласия решается задача отбора или неотбора рассматриваемого отсчета Z_{ik} для продолжения рассматриваемой траектории X_j .

Операция селекции предполагает нахождение двух областей, одна из которых располагается в непосредственной близости от экстраполированного положения $Z_{эjk}$ цели, имеющей траекторию X_j , а другая далеко от него (попадание целевого отсчета в нее маловероятно). Первая область образует некоторый корреляционный строб (иначе говоря, строб сопровождения, или зону связи прогнозируемого положения цели с отсчетом от нее). Вероятность $P_{стр}$ нахождения отсчета в корреляционном стро́бе получила название *вероятности правильного стробирования*.

Выполнения операции стробирования оказывается достаточно для окончательного решения задачи селекции уже на первом этапе при выполнении всех следующих условий:

а) в корреляционном стро́бе не могут появиться отсчеты от нескольких целей;

б) вероятность $D_{от}$ правильного обнаружения отсчета от цели с рассматриваемой траекторией близка к 1;

в) вероятность $F_{ст}$ появления ложного отсчета в стро́бе близка к 0.

Ясно, что полностью обеспечить выполнение перечисленных условий в РЛС практически невозможно, но желательно к нему стремиться. Тогда операция селекции проводится наиболее просто. При этом вероятность правильного стробирования $P_{стр}$ практически совпадает с вероятностью правильной селекции P_c .

При явном невыполнении перечисленных условий возникает необходимость второго этапа селекции, в ходе которого конкретизируется отсчет, подтверждающий рассматриваемую траекторию. При правильном выборе параметров первого этапа характеристики двухэтапной процедуры селекции приближаются к характеристикам операции селекции на основе полного перебора всех гипотез отождествления отсчетов и траекторий [71]. В некоторых особо сложных случаях в явном виде операция селекции вообще не производится, однако при построении траекторий учитываются все возможные комбинации всех траекторий, обрабатываемых алгоритмом ВО, со всеми поступающими отсчетами (см., например, [57—60, 62—63]). Траектории, построенные по ложным или относящимся к другим траекториям отсчетам, сбрасываются с сопровождения, как только будет обнаружено какое-либо их несоответствие целевой траектории. Такие подходы требуют весьма больших вычислительных затрат.

Основные факторы, которые могут повлиять на выполнение операции селекции, изображены на рис. 6.6, где указано, какие параметры целевого потока, характеристики радиолокационной станции и условия ее работы необходимо учитывать при выборе метода селекции.

Часто изменения лишь одного фактора из перечисленных вызывают «цепную реакцию», неминуемо оказывающую влияние на операцию селекции и работу всего алгоритма ВО в целом. Например, ухудшение точности измерений, характеризуемых матрицей R_{ik} , приведет к такому изменению ковариационной матрицы S_{3ijk} , которое расширит границы корреляционного строга, что повлечет увеличение вероятности появления в нем ложных отсчетов. При ошибочной селекции калмановский фильтр перестает в некоторых тактах работы следить за целью, в результате ухудшается точность оценки траекторных параметров, возрастает вероятность срыва цели с сопровождения. Можно привести и другие аналогичные примеры.

В настоящее время детально разработаны методы селекции только для простейших случаев, несущественно отличающихся от идеального. Многие проблемы, возникающие при проведении ВО радиолокационной информации в случае наблюдения маневрирующих целей в условиях интенсивных помеховых сигналов и (или) в многоцелевой обстановке, еще не решены. В этом параграфе рассматриваются лишь основные, наиболее часто используемые в настоящее время методы селекции.

Условия операции селекции



Рис. 6.6. Условия операции селекции

6.5.1. Селекция отсчетов методом стробирования

Селекция методом стробирования применяется в качестве самостоятельного метода, когда условия селекции близки к идеальным (см. ранее). Допускается некоторое смягчение условия б): вероятность правильного обнаружения отсчета должна быть не ниже 0,8...0,9 (на 10 тактах наблюдения должно быть не более 1–2 пропусков полезного отсчета, пропуск подряд двух отсчетов маловероятен).

Допускается некоторое смягчение условия в): вероятность появления ложного отсчета в стробе должна быть много меньше 1, обычно $F_{\text{ст}} < 0,1$ (на 10 тактах наблюдения в стробе возможно появление не более одного ложного отсчета). Следует отметить, что вероятность $F_{\text{ст}}$ зависит от размеров корреляционного строба и элемента разрешения. В корреляционном стробе может быть большое число N_z элементов разрешения: десятки, сотни и даже тысячи.

Величина $F_{\text{ст}}$ связана с $F_{\text{от}}$ соотношением

$$F_{\text{ст}} = 1 - (1 - F_{\text{от}})^{N_z}. \quad (6.81)$$

Можно считать (если $N_z F_{\text{от}} < 1$ и отброшенные члены ряда пренебрежимо малы по сравнению с $N_z F_{\text{от}}$), что

$$F_{\text{ст}} \approx N_z F_{\text{от}}. \quad (6.82)$$

В случае гауссовских ошибок измерений параметров отсчета, гауссовских ошибок экстраполированных параметров измерений и гауссовских возмущений траектории движения цели, формируемый строб представляет собой эллипс, размеры которого определяются матрицей S_{zjk} и заданной вероятностью стробирования $P_{\text{ст}}$ — вероятностью того, что при имеющихся ошибках отсчет от цели будет находиться внутри этого строба. При этом ориентация строба зависит от взаимной ориентации главных осей эллипсоида ошибок измерений параметров отсчета, эллипсоида экстраполированных ошибок траекторных параметров, эллипсоида экстраполированных возмущений траектории движения цели и их соотношений.

При ориентации системы координат по главным полуосям эллипсоида плотность вероятности соответствующих случайных величин $\Delta\eta$, $\Delta\xi$, $\Delta\zeta$, характеризующихся среднеквадратическими отклонениями σ_η , σ_ξ , σ_ζ , имеет вид

$$w(\Delta\eta, \Delta\xi, \Delta\zeta) = \frac{1}{(2\pi)^{3/2} \sigma_\eta \sigma_\xi \sigma_\zeta} \exp\left(-\frac{r^2}{2}\right), \quad (6.83)$$

$$\text{где } r^2 = \frac{\Delta\eta^2}{\sigma_\eta^2} + \frac{\Delta\xi^2}{\sigma_\xi^2} + \frac{\Delta\zeta^2}{\sigma_\zeta^2}.$$

Вероятность попадания случайной точки в подобный эллипсоид определяется из выражения

$$P(r) = 2 \left[\Phi_0(r) - \frac{1}{\sqrt{2\pi}} r \exp\left(-\frac{r^2}{2}\right) \right], \quad (6.84)$$

где $\Phi_0(r) = \frac{1}{\sqrt{2\pi}} \int_0^r \exp\left(-\frac{t^2}{2}\right) dt$ — интеграл вероятности.

При r , равном 3,5, вероятность стробирования приблизительно равна 0,993.

Типичный пример селекции методом стробирования в штатном режиме на нескольких тактах работы алгоритма ВО для двухмерного случая показан на рис. 6.7. Для траектории с параметрами $\hat{X}_j(t_1)$ в момент времени t_1 экстраполируется положение цели $X_{yj}(t_2)$ на момент t_2 . Затем относительно экстраполированных параметров отсчета $Z_{yj}(t_2) = H_j X_{yj}(t_2)$ строится корреляционный строб. После выполнения в момент t_2 операций первичной обработки и получения отсчетов оказывается, что в корреляционный строб попал отсчет $Z_i(t_2)$. В соответствии, например, с формулами калмановской фильтрации находится оценка параметров траектории $\hat{X}_j(t_2)$ на момент t_2 . Та же процедура повторяется для всех последующих тактов работы.

При работе алгоритма возможны случаи, когда в строб сопровождения не попадут отсчеты от цели. Тогда траектория либо экстраполируется дальше, либо сбрасывается (в зависимости от алгоритма обнаружения—сброса).

Предположим, что в момент времени t_4 в стробе не оказалось подтверждающего отсчета. В этом случае (если траектория не сбрасывается с сопровождения) происходит экстраполяция уже на момент времени t_5 . При увеличении времени экстраполяции возрастает значение элементов экстраполированной ковариационной матрицы $S_{jji}(t_5)$ экстраполированного отсчета $Z_{yj}(t_5)$ и, следовательно, увеличиваются размеры корреляционного строба, как условно показано на рис. 6.7. При попадании в этот строб отсчета от цели $Z_i(t_5)$ уточняется оценка $\hat{X}_j(t_5)$, соответствующая ковариационная матрица и т. д.

В более сложных ситуациях возможно появление нескольких ложных отсчетов в стробе. В этом случае необходимо применять алгоритмы, пред-

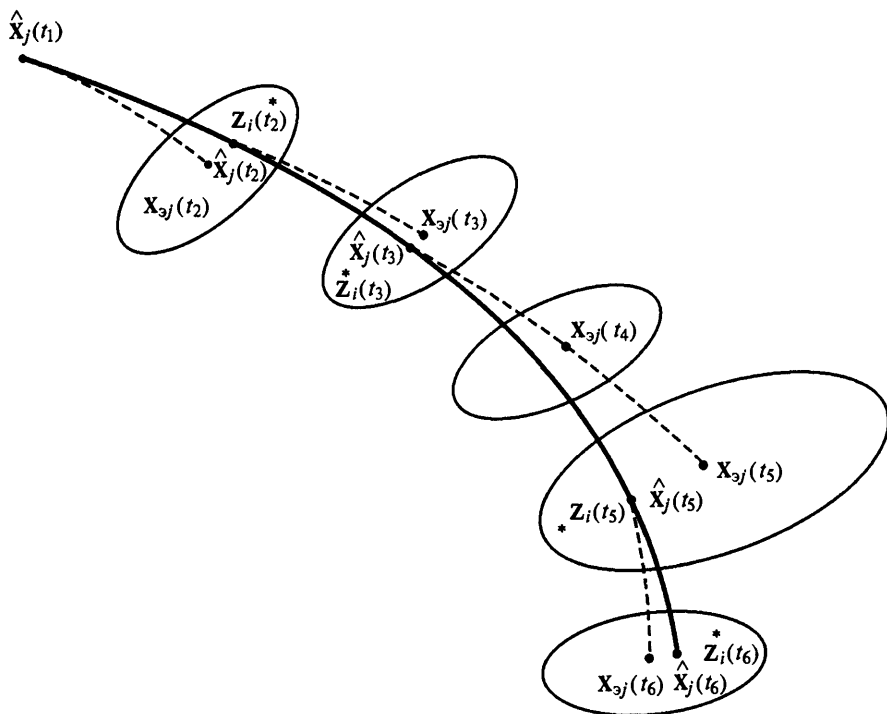


Рис. 6.7. Селекция методом стробирования

назначенные для второго этапа селекции. В таких ситуациях чаще всего для продолжения траектории отбирается отсчет, ближайший к центру строба (подробнее см. п. 6.5.2).

Метод стробирования может быть реализован различными способами. Учитывая гауссовское распределение величины невязки, в качестве меры рассогласования можно использовать соответствующую квадратичную форму

$$\delta^2(\mathbf{Z}_{ji}) = \Delta \mathbf{Z}_{ji}^T(t_k) \mathbf{S}_{эji}^{-1} \Delta \mathbf{Z}_{ji}(t_k). \quad (6.85)$$

Поскольку величина $\delta^2(\mathbf{Z}_{ji})$ при правильном отборе отсчета и траектории имеет χ^2 -распределение с числом степеней свободы, соответствующим количеству координат параметров отсчета, то, задавая вероятность $P_{\text{стр}}$, можно найти интервал (c_1, c_2) , при попадании в который значения квадратичной формы $\delta^2(\mathbf{Z}_{ji})$ принимается решение об отборе рассматриваемого отсчета для продолжения рассматриваемой траектории.

Иногда стробирование имеет смысл проводить геометрически. Однако построение эллипсоида и определение местоположения отсчетов относи-

тельно него нецелесообразно из-за трудоемкости вычислений. Поэтому для РЛС формируется строб более простой формы, обычно подобный элементу разрешения РЛС по измеряемым координатам. Например, это может быть строб, задаваемый размером по дальности $\Delta r_{\text{стр}}$ и по двум угловым координатам: по азимуту $\Delta \beta_{\text{стр}}$ и углу места $\Delta \epsilon_{\text{стр}}$. При этом величины $\Delta \beta_{\text{стр}}$, $\Delta \epsilon_{\text{стр}}$, $\Delta d_{\text{стр}}$ выбираются таким образом, чтобы в выделенном объеме помещался эллипсоид (6.83) любой ориентации с заданным r .

В результате форма стоба существенно упрощается, но объем эллипсоида увеличивается по сравнению с расчетным. Это уменьшает заданную ошибку стробирования, но увеличивает вероятность появления ложных отсчетов в стробе. Если это недопустимо, то объем «прямоугольного» стоба уменьшают для обеспечения заданной вероятности $P_{\text{стр}}$.

При наблюдении целей, которые могут осуществлять более сложные маневры, чем предусмотренные случайными ускорениями в виде гауссовского процесса с ковариационной матрицей Q , форма стоба может существенно отличаться от эллипсоида. Например, для военного самолета с учетом его возможностей ускорения форма стоба может иметь вид, приведенный на рис. 6.8 [61].

В настоящее время селекция методом стробирования, как уже отмечалось, в отдельности применяется редко, только в простой целевой и помеховой обстановке. Чаще этот метод используют в двухэтапных алгоритмах в качестве первого вспомогательного этапа. В этом случае с помощью стробирования отсекаются отсчеты, заведомо не подходящие для продолжения траектории. Это позволяет осуществлять более тонкое селектирование, не затрачивая вычислительные ресурсы на обработку заведомо не подходящих отсчетов.

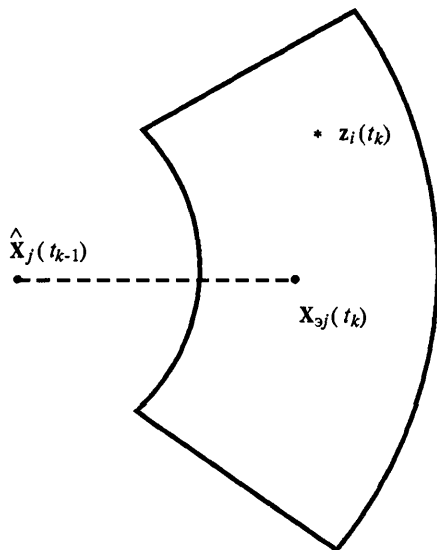


Рис. 6.8. Вариант стоба

6.5.2. Алгоритм селекции отсчетов

по минимальному отклонению от центра стоба

Алгоритм селекции отсчетов по минимальному отклонению от центра стоба обычно применяется в двухэтапной процедуре стробирования. Этот

алгоритм предназначен для работы в случаях, когда в стробе появляется несколько помеховых отсчетов.

Так же, как и в алгоритме стробирования, здесь считается, что рассогласование (невязка) ΔZ_{ji} для правильно выбранной пары траектория—отсчет есть случайная величина с нулевым математическим ожиданием и ковариационной матрицей (6.79). Предполагается, что ложные отсчеты распределены равномерно в зоне контроля со средней плотностью γ . Работа алгоритма основана на том, что в процессе наблюдений целевые отсчеты должны чаще оказываться ближе к центру строба траектории, чем помеховые.

В работе [67] при выполнении указанных условий получено выражение для вероятности $P_{м.о}$ правильного отбора целевых отсчетов по критерию минимального отклонения от центра строба. В случае двухмерного строба имеем

$$P_{м.о} = \frac{1}{1+2s} \left\{ 1 - \exp \left[- \left(\frac{1}{2} + s \right) r^2 \right] \right\}, \quad (6.86)$$

где $s = \pi \sigma_{\eta} \sigma_{\xi} \gamma$; σ_{η} и σ_{ξ} — главные полуоси эллипса рассогласования, определяемого ковариационной матрицей (6.80); r — коэффициент, определяющий размеры строба в зависимости от допустимой вероятности α отбраковки целевых отсчетов в методе стробирования (см. п. 6.5.1).

При $r \geq 3$ (что при стробировании почти всегда выполняется) приближенно имеем

$$P_{м.о} = \frac{1}{1+2\pi\gamma\sigma_{\eta}\sigma_{\xi}}. \quad (6.87)$$

Для трехмерного строба вероятность правильного отбора целевых отсчетов имеет вид

$$P_{м.о} = \frac{2}{\sqrt{2\pi}} \int_0^r t^2 \exp \left(-\frac{t^2}{2} \right) \exp(-Vt^2) dt, \quad (6.88)$$

где

$$V = \frac{4}{3} \pi \sigma_{\eta} \sigma_{\xi} \sigma_{\zeta} \gamma. \quad (6.89)$$

Для двухэтапного метода селекции «стробирование—отбор по минимальному расстоянию от центра строба» вероятность правильной селекции

$$P_c = P_{ст} P_{м.о}. \quad (6.90)$$

Рассмотренный двухэтапный метод селекции является в настоящее время наиболее распространенным методом селекции.

6.5.3. Алгоритмы сопоставления и привязки отсчетов к траекториям в многоцелевой ситуации

При высокой пространственной плотности траекторий, целевых, а также помеховых отсчетов могут образовываться кластеры вида, показанного на рис. 6.4, б. Это одна из самых трудных ситуаций для операции селекции и всей ВО в целом. Общепринятые методы, надежно решающие задачу селекции, для таких случаев еще не разработаны.

В качестве одного из подходов можно привести эвристический вариант селекции, использующий метод сопоставления и привязки совокупности всех пар траектория—отсчет. Работу алгоритма поясним на примере [61], показанном на рис. 6.9. В строб траектории X_1 попало два отсчета, в строб траектории X_2 попало три отсчета, в строб траектории X_3 попал один отсчет. Из-за пересечения стробов возникает неопределенность с отборами отсчетов.

Для решения возникшей задачи селекции сначала вычисляют все взаимные расстояния. Логично в качестве меры расстояния взять квадратичную форму (6.25). Результаты приведены в табл. 6.2.

Затем каждая траектория предварительно привязывается к ближайшему отсчету, после чего предварительные привязки проверяются в целях устранения использования отсчета более одного раза. Выполнение этой процедуры поясняется в табл. 6.3.

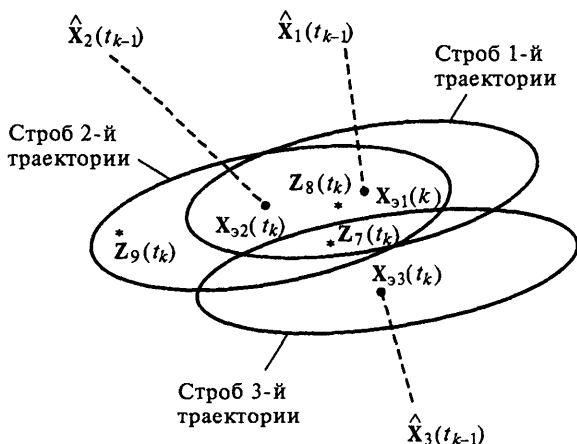


Рис. 6.9. Вариант многоцелевой ситуации

Таблица 6.2

Матрица расстояний: траектория—отсчеты

Номер отсчета	Номера траектории		
	1	2	3
7	4,2	5,4	6,3
8	1,2	3,1	∞
9	∞	7,2	∞

Таблица 6.3

Матрица близости

Номер траектории	Степень близости пар траектория—отсчет					
	I		II		III	
	Номер отсчета	Расстояние	Номер отсчета	Расстояние	Номер отсчета	Расстояние
1	8	1,2	7	4,2	—	—
2	8	3,1	7	5,4	9	7,2
3	7	6,3	7	6,3	—	—

Отсчет $Z_8(t_k)$, связанный с траекториями $X_{31}(t_k)$ и $X_{32}(t_k)$, отнесем к траектории $X_{31}(t_k)$, поскольку расстояние между отсчетом $Z_8(t_k)$ и этой траекторией меньше, чем расстояние между этим отсчетом и траекторией $X_{32}(t_k)$. Очевидно, в условиях, когда один отсчет может соответствовать только одной траектории, отсчет $Z_8(t_k)$ в дальнейшем отождествлении участвовать уже не может. Из оставшихся свободных пар после первого шага отбора расстояние минимально между отсчетом $Z_7(t_k)$ и второй траекторией. После исключения из отбора отсчета $Z_7(t_k)$ остается одно решение, траектория $X_{33}(t_k)$ не подтверждается ни одним из отсчетов. Решение о сбросе или несбросе этой траектории с сопровождения будет приниматься операцией обнаружение—сброс (см. § 6.6).

Наряду с приведенным выше простейшим эвристическим алгоритмом имеются и другие, при осуществлении которых, как правило, траектории разделяются на некоторые классы с использованием как траекторных, так и сигнальных признаков. В этом случае селекция становится возможной и в более сложной обстановке.

Эффективным является подход, при котором в сложной целевой и помеховой ситуации операцию селекции не проводят в качестве самостоятельной, поскольку принять правильное решение в ходе ее все равно не удастся.

Для этих условий в настоящее время интенсивно разрабатываются алгоритмы ВО, производящие совместно операции селекции—оценивания—обнаружения [56—66].

6.6. Обнаружение траекторий

6.6.1. Общие положения

В современных радиолокационных системах требуемые вероятностные и точностные характеристики обеспечиваются лишь после проведения этапа ВО. При этом в отличие от первичной обработки при проведении ВО приходится наблюдать появляющиеся, перемещающиеся и исчезающие объекты и принимать решение с учетом целевой ситуации во всей зоне ответственности, а не по отдельным элементам разрешения. Это решение не удастся свести только к задаче обнаружения цели; одновременно с обнаружением необходимо указывать траекторные и некоторые другие параметры целей, точность их оценки (см. § 6.1). Тем не менее, для анализа алгоритмов обнаружения и их сравнения желательно ограничиться небольшим числом характеристик, которые подобно вероятностям $D_{от}$ и $F_{от}$, используемым в первичной обработке, характеризовали бы качество обнаружения траекторий. Обычно в радиолокационной практике для оценки качества обнаружения траекторий используют три основных показателя: $f_{тр}$ — частоту ложно обнаруживаемых траекторий и $T_{тр}$ — время обнаружения траекторий с заданной вероятностью правильного обнаружения $D_{тр}$ с момента появления цели в зоне ответственности [57].

Величина $f_{тр}$ — это среднее число ложно обнаруженных целей (траекторий) за единицу времени в зоне ответственности. В зависимости от вида РЛС могут предъявляться различные требования к частоте ложных тревог: одна за час, одна за день, одна за год и т. п.

Вероятность правильного обнаружения траектории $D_{тр}$ является функцией времени. Совершенно очевидно, что существует тенденция увеличения $D_{тр}(t)$ с увеличением времени наблюдения t , но всегда есть определенные ограничения на временные ресурсы, отводимые для принятия решения о наличии цели. Поэтому наряду с вероятностью правильного обнаружения траектории обязательно указывается время $T_{тр}$, которое необходимо для достижения указанной вероятности $D_{тр}$. Отметим, что для появляющихся и исчезающих объектов помимо принятия решения об обнаружении траектории приходится принимать отдельные решения и о сбросе траекторий. Решение о сбросе траектории может быть правильным, если сбрасывается исчезнувший объект или начавшая завязываться ложная траектория, и мо-

жет быть ошибочным, если сбрасывается с сопровождения по тем или иным причинам (в том числе при срыве сопровождения в процессе фильтрации) траектория истинной цели. Вообще говоря, качество траекторной обработки характеризуется большой совокупностью точностных и вероятностных показателей, что указывает на необходимость более полного учета специфики задачи ВО информации как задачи совместного обнаружения—измерения.

В целом задача обнаружения—измерения траекторных объектов может быть отнесена к динамической задаче проверки сложных гипотез в условиях нестационарной входной помеховой и целевой обстановки с широким диапазоном изменения внутренних параметров и наличием ряда ограничений, также изменяющихся во времени. Общего решения этой задачи пока еще не существует.

Характеристики $f_{\text{тр}}$, $T_{\text{тр}}$, $D_{\text{тр}}(t)$ обнаружения траекторий зависят от большого числа факторов: качества информации, поступающей с выхода первичной обработки; используемых алгоритмов обнаружения; возможностей управления параметрами РЛС; динамики изменения целевой и помеховой обстановки.

Вероятность $D_{\text{тр}}(T_{\text{тр}})$ может быть использована как оценка качества обнаружения отдельной траектории. Частота ложных тревог $f_{\text{тр}}$ при обнаружении траекторий является интегральным показателем. Для алгоритма обнаружения траекторий удобнее использовать локальный показатель — вероятность ложного обнаружения траектории $F_{\text{тр}}$, определяемого как вероятность обнаружения траектории для некоторого ложного отсчета, поступившего с выхода первичной обработки.

Точную (функциональную) связь между $f_{\text{тр}}$ и $F_{\text{тр}}$ установить очень сложно, поскольку слишком большое число факторов внешней обстановки, зоны ответственности, параметров самых разнообразных алгоритмов и ограничений влияют (и не всегда однозначно) на эти величины.

Приблизительные соотношения между $f_{\text{тр}}$ и $F_{\text{тр}}$ можно получить следующим образом. Пусть зона ответственности РЛС характеризуется числом элементов разрешения, равным N_z . Тогда среднее число ложных отметок за такт T_1 первичной обработки равно $N_z F_{\text{от}}$, где $F_{\text{от}}$ — вероятность ложной отметки за такт первичной обработки (считаем здесь для определенности, что во всех элементах разрешения $F_{\text{от}} = F = \text{const}$). Очевидно, что во всей зоне ответственности частота ложных отметок равна $N_z F_{\text{от}} / T_1$.

Каждая ложная отметка после оценки ее параметров становится отсчетом, который может привести к возникновению ложной траектории. Если после появления первого ложного отсчета при последующих наблюдениях вероятность обнаружения ложной траектории будет равна $F'_{\text{тр}}$, то частоту ложных

тревог траекторий приближенно можно оценить следующим образом:

$$f_{\text{тр}} \approx \frac{N_3 F_{\text{от}}}{T_1} F'_{\text{тр}}.$$

Обозначив $F_{\text{тр}} = F_{\text{от}} F'_{\text{тр}}$ и понимая $F_{\text{тр}}$ как вероятность ложной тревоги траектории, можно записать

$$f_{\text{тр}} \approx \frac{N_3 F_{\text{тр}}}{T_1}. \quad (6.91)$$

Соотношение (6.91) определяет оценку сверху для частоты ложных тревог, поскольку некоторые ложные отсчеты на последующих тактах траекторной обработки не смогут участвовать в возникновении новых ложных траекторий, так как они уже будут отобраны для подтверждения ранее начавшихся. Более строгий подход к оценке $f_{\text{тр}}$ изложен в работах [57, 67].

В настоящем параграфе рассматривается базовый, наиболее распространенный в настоящее время подход к обнаружению траекторий на основе логических решающих правил типа « k из m » по аналогии с обнаружением пачки сигналов в ходе первичной обработки. Обнаруженными согласно этим правилам считаются те траектории, которые подтверждаются отсчетами, попавшими в строб селекции (см. § 6.5) k раз на m тактах наблюдения.

В ходе ВО при обнаружении траекторий одной из задач является фильтрация потока ложных отсчетов. При этом соответствующие вероятностные характеристики первичной обработки могут быть значительно улучшены. Например, если используется правило «3 из 3» и $F_{\text{от}} = 10^{-3}$, то, несколько упрощая ситуацию, можно считать, что в пределе вероятность ложной тревоги траектории может стать равной $F_{\text{тр}} = 10^{-9}$ (в действительности $F_{\text{тр}}$ будет несколько больше; см. п. 6.6.2).

В последнее время с помощью траекторной обработки все чаще стали решать задачу не только снижения вероятности ложной тревоги, но и повышения вероятности правильного обнаружения целей. Это связано с тем, что различные меры по снижению заметности целей привели к уменьшению вероятности $D_{\text{от}}$ на этапе первичной обработки. С увеличением времени наблюдения за целью естественным образом улучшаются энергетические ресурсы обработки, кроме того, параметры траектории служат дополнительным признаком для селекции помеховых сигналов. В этих условиях одним из путей решения задачи обнаружения малозаметных объектов может являться снижение порога обнаружения на этапе первичной обработки с последующим обнаружением на фоне помех истинных траекторий.

Обычно выделяются три этапа принятия решения о наличии целевой траектории или ее отсутствии.

На первом этапе решается вопрос о завязке траектории. Это делается в блоке завязки, обычно по правилу « m из m » (т. е. по обнаружению подряд m отсчетов в m тактах наблюдения), иногда по правилу « r из m » (т. е. по обнаружению r отсчетов в m тактах наблюдения). Завязка траектории в существенной мере определяет вероятность $F_{\text{тр}}$ ложной тревоги. При этом необходимо иметь в виду, что при слишком жестких условиях захвата (хороших фильтрующих свойствах обнаружителя) может существенно снизиться вероятность правильного обнаружения траектории. При невыполнении выбранного критерия алгоритм завязки осуществляет сброс траектории с завязки.

На втором этапе решается вопрос окончательного обнаружения траектории. Это делается в блоке обнаружения—сброса (см. рис. 6.2) с помощью работающего после завязки траектории правила « l из n ». Косвенно правило « l из n » включает в себя также критерий сброса траектории.

На третьем этапе работы блока обнаружения для более четкой фиксации сброса траектории вводится правило сброса, обычно это число пропущенных отсчетов подряд на нескольких тактах наблюдения s .

В общем виде алгоритм завязка—обнаружение—сброс записывается следующим образом: « r из m » + « l из n » — « s ».

Иногда в логическом критерии некоторые этапы опускаются, иногда добавляются промежуточные. Выбор того или иного варианта обнаружителя осуществляется после совместного анализа его вероятностных характеристик как в случае отсутствия полезных сигналов (при этом определяется вероятность $F_{\text{тр}}$), так и в случае их наличия (при этом рассчитывается вероятность правильного обнаружения траектории $D_{\text{тр}}(t)$).

6.6.2. Вероятность ложного обнаружения траектории

Структура простейшего алгоритма завязка—обнаружение—сброс «2 из m » + «1 из m » — « s » в виде направленного графа [67] приведена на рис. 6.10. Направленный граф отображает все возможные состояния логического обнаружителя и все возможные переходы одного состояния в другое. Переход логического обнаружителя из одного состояния в другое происходит в зависимости от подтверждения траектории в каждый момент времени отсчетом (появления логической единицы) или от ее неподтверждения. Вероятность подтверждения в общем случае зависит как от текущего времени t_k , так и от текущего состояния обнаружителя, т. е. $p_i = p_i(t_k)$, где i — номер состояния обнаружителя. Чтобы не усложнять рисунок, на нем зависимость появления логической единицы от времени t_k не отражена.

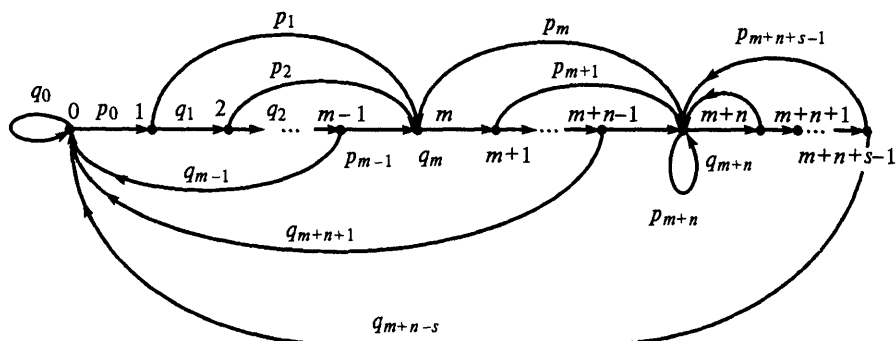


Рис. 6.10. Граф алгоритма «2 из m » + «1 из n » – « s » при обнаружении ложной траектории

Исходное состояние обнаружителя — нулевое (0). С приходом первого отсчета алгоритм переходит в состояние 1. Это событие характеризуется вероятностью p_0 . Очевидно, за величину p_0 можно принять вероятность ложного обнаружения отсчета $F_{от}$. На следующем такте вероятность появления ложного отсчета p_1 в стробе сопровождения зависит от вероятности обнаружения ложного отсчета в элементе разрешения $F_{от}$ и от размеров строба для первого такта, что характеризуется числом N_{31} элементов разрешения в нем. Можно считать, что

$$p_1 = 1 - (1 - F_{от})^{N_{31}} \approx N_{31} F_{от}.$$

С учетом логики работы алгоритма появление второго (ложного в данном случае) отсчета в корреляционном стробе (при выполнении правила «2 из m ») переводит его в состояние m — «завязка траектории».

Вероятности не появления отсчетов можно вычислить следующим образом: $q_0 = 1 - p_0$, $q_1 = 1 - p_1$ и т. д. Значение соответствующей вероятности p_i для i -го состояния рассчитывается аналогично p_1 . При переходе обнаружителя из одного состояния в другое вероятности p_i изменяются, что вызвано, прежде всего, размером строба сопровождения и, соответственно, числом N_{3k} элементов разрешения в нем.

В ходе траекторной обработки необходимо выбором всех параметров РЛС: размера элемента разрешения, величины корреляционного строба, вероятности $F_{от}$ — обеспечить выполнение неравенства $p_i \ll 1$.

При переходе обнаружителя в состояние m фиксируется факт завязки траектории. Затем начинается второй этап работы обнаружителя — этап окончательного обнаружения траектории. Таким же образом, как и раньше, в зависимости, прежде всего, от размера строба вычисляются вероятности p_k

и q_k и строятся соответствующие ветви графа. Траектория считается обнаруженной в данном примере при переходе графа в состояние $m + n$.

В дальнейшем траектория может оставаться в состоянии $m + n$ либо быть сброшена при пропуске подряд s отсчетов.

Вероятность обнаружения ложной траектории определяется вероятностью первого достижения состояния $m + n$.

Анализ графа производится с использованием теории марковских цепей. Обычно для этой цели ограничиваются простой цепью Маркова [67], которая задается матрицей переходных вероятностей $\Pi(t_k)$ и вектором-строкой начального состояния цепи $\mathbf{V}(t_0) = [b_0(t_0) \ b_1(t_0) \ b_2(t_0) \ \dots]$. Для упрощения записи зависимость от текущего времени t_k будем показывать номером соответствующего такта: $\Pi(t_k) \equiv \Pi(k)$,

$$\mathbf{V}(t_0) = [b_0(t_0) \ b_1(t_0) \ b_2(t_0) \ \dots] \equiv \mathbf{V}(0) = [b_0(0) \ b_1(0) \ b_2(0) \ \dots].$$

Для представленного на рис. 6.10 логического обнаружителя матрица переходных вероятностей имеет вид

$$\Pi(k) =$$

	0	1	...	m	$m+1$...	$m+n-1$	$m+n$	$m+n+1$...	$m+n+s-1$
0	q_0	p_0	...	0	0	...	0	0	0	...	0
1	0	q_1	...	p_1	0	...	0	0	0	...	0
2	0	0	...	p_2	0	...	0	0	0	...	0
...
$m-1$	q_{m-1}	0	...	p_{m-1}	0	...	0	0	0	...	0
m	0	0	...	0	q_m	...	0	p_m	0	...	0
$m+1$	0	0	...	0	0	...	0	p_{m+1}	0	...	0
...
$m+n-1$	q_{m+n-1}	0	...	0	0	...	0	p_{m+n-1}	0	...	0
$m+n$	0	0	...	0	0	...	0	p_{m+n}	0
$m+n+1$	0	0	...	0	0	...	0	0	0
...
$m+n+s-1$	$q_{m+n+s-1}$	0	...	0	0	...	0	$p_{m+n+s-1}$	0	...	0

(6.92)

Для расчета вероятностей состояния графа (обнаружителя) используется соотношение

$$\mathbf{V}(k) = \mathbf{V}(k-1)\Pi(k). \quad (6.93)$$

Задаваясь начальным состоянием обнаружителя $\mathbf{V}(0)$, можно рассчитать в любой k -й момент времени все состояния обнаружителя. Искомая вероятность ложного обнаружения траектории для рассматриваемого примера есть вероятность первого достижения состояния $m + n$, т. е.

$$F_{\text{тр}}(k) = b_m(k-1)p_m(k) + b_{m+1}(k-1)p_{m+1}(k) + \dots \\ \dots + b_{m+n-1}(k-1)p_{m+n-1}(k). \quad (6.94)$$

Произвести строгие вычисления по формулам (6.93) и (6.94) для нахождения $F_{\text{тр}}$ в силу нестационарности задачи достаточно сложно. Можно упростить задачу, сделав предположение о постоянстве матрицы перехода $\Pi(k) = \Pi$. В этом случае, естественно, с определенной погрешностью можно вычислить вероятность ложного обнаружения траектории по формулам (6.93), (6.94), воспользовавшись формулой финальных вероятностей

$$\mathbf{V} = \mathbf{V}\Pi, \quad (6.95)$$

приблизительно считая неизменными все параметры вектора состояний и матрицу переходных вероятностей (это справедливо для предельных состояний эргодической цепи Маркова).

В ряде случаев для несложных обнаружителей можно использовать и более простой вероятностный расчет. Например, для обнаружителя «2 из m » + «1 из n » для оценки вероятности ложной тревоги можно воспользоваться выражением

$$F_{\text{тр}} = F_{\text{от}} \left(1 - \prod_{i=1}^{m-1} q_i \right) \left(1 - \prod_{j=m}^{m+n-1} q_j \right), \quad (6.96)$$

где q_i — вероятность неполучения отсчета в первой серии наблюдений (в ходе завязки), q_j — вероятность неполучения отсчета во второй серии наблюдений (в ходе окончательного обнаружения).

Очень часто применяют обнаружители типа « r из m » — « s », в которых имеется только два этапа: завязка и сброс. Завязка при этом выполняется в блоке завязки, сброс — в блоке обнаружения—сброса. Некоторые вероятностные характеристики таких обнаружителей при фиксированных вероятностях появления логических единиц приведены на рис. 6.11.

Следует отметить, что сброс траектории с сопровождения при фиксированном числе пропусков отсчетов от цели не учитывает всю предысторию рассматриваемой траектории и всю дополнительную информацию, накопленную о цели в процессе сопровождения. В частности, можно корректировать критерий сброса траектории величиной требуемого корреляционного строга. Если он превышает определенные размеры, то сброс может производиться с меньшим числом пропусков, чем в случае строга с малыми размерами. Дополнительную информацию для уточнения критерия при принятии решения о сбросе может дать характер движения цели, наличие и отсутствие маневра и т. д.

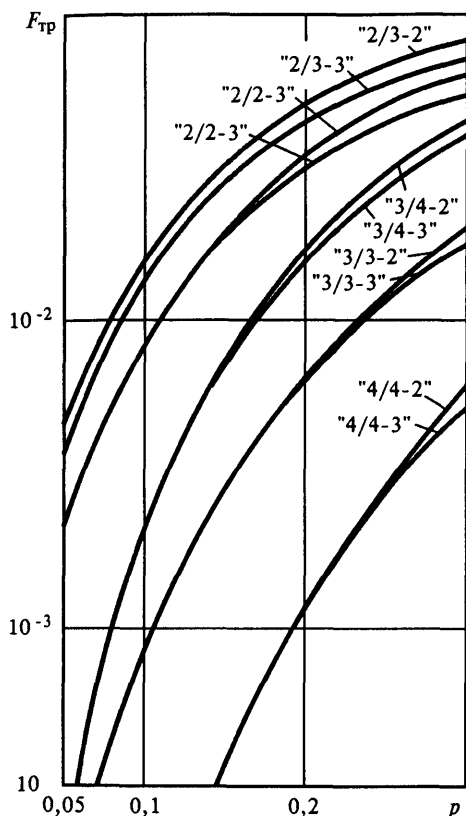


Рис. 6.11. Вероятности ложного обнаружения

при анализе вероятностей ложной тревоги, заменив в графе часть, описывающую критерий сброса, поглощающим экраном, поскольку вероятность правильного обнаружения траектории связана только с достижением соответствующего состояния обнаружения [67]. Такой граф для рассмотренного

Окончательно настройка алгоритма требует проведения моделирования и уточнения его параметров, а иногда и натурных экспериментов.

6.6.3. Вероятность правильного обнаружения траектории

При поступлении на вход обнаружителя отсчетов, полученных от некоторой цели, логика работы алгоритма остается той же, что и в случае ложных отсчетов. Траектория цели обнаруживается при выполнении условия, заложенного в используемое логическое правило. Метод расчета вероятности правильного обнаружения траектории близок к подходу, рассмотренному в п. 6.6.2 при нахождении вероятности ложного обнаружения траектории.

Для анализа вероятности правильного обнаружения $D_{гр}(k)$ можно воспользоваться графами, характеризующими работу обнаружителя

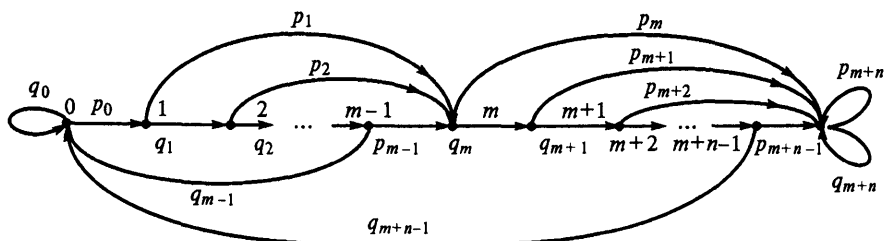


Рис. 6.12. Граф алгоритма «2 из m » + «1 из n » - « s » при обнаружении истинной траектории

в п. 6.6.2 примера обнаружителя «2 из m » + «1 из n » – « s » приведен на рис. 6.12. Для него вероятность обнаружения траектории равна вероятности состояния $m + n$.

Матрица переходных вероятностей $\Pi(k)$ в данном случае имеет вид

$$\Pi(k) =$$

	0	1	...	m	$m+1$...	$m+n-1$	$m+n$
0	q_0	p_0	...	0	0	...	0	0
1	0	q_1	...	p_1	0	...	0	0
2	0	0	...	p_2	0	...	0	0
...
$m-1$	q_{m-1}	0	...	p_{m-1}	0	...	0	0
m	0	0	...	0	q_m	...	0	p_m
$m+1$	0	0	...	0	0	...	0	p_{m+1}
...
$m+n-1$	q_{m+n-1}	0	...	0	0	...	0	p_{m+n-1}
$m+n$	0	0	...	0	0	...	0	1

(6.97)

Строго говоря, все вероятности p_i и q_i в (6.97), характеризующие вероятности появления отсчета в соответствующем корреляционном строке при переходе из одного состояния графа в другое, разные. Очевидно, что вероятность появления первого отсчета от цели при нахождении ее в зоне контроля равна $p_0 = D_{от}$, а последующие равны $p_i = D_{от} P_{ci}$, где P_{ci} — вероятность правильной селекции отсчета от цели при сопровождении на i -м такте после первого отсчета. Учитывая, что значение вероятности правильной селекции обеспечивается близкой к 1, то $p_i \approx D_{от}$. В процессе анализа обнаружителя при работе в простых помеховых и целевых условиях вероятностью появления ложного отсчета в корреляционном строке пренебрегают. По этой причине в первом приближении матрицу переходных вероятностей $\Pi(k)$ в процессе обнаружения некоторой целевой траектории можно считать неизменной.

В отличие от случая поступления ложных отсчетов вероятности состояния графа, характеризующиеся вектором-строкой $\mathbf{V}(k)$, при наблюдении цели принципиально зависят от времени (такта наблюдения). Поэтому при расчете вероятностей $\mathbf{V}(k)$ необходимо использовать рекуррентное соотношение (6.93).

Задавшись начальным состоянием обнаружителя $\mathbf{V}(0)$, можно определить нарастающую вероятность правильного обнаружения траектории $D_{тр} = D_{тр}(k)$. В качестве начального состояния чаще всего используют век-

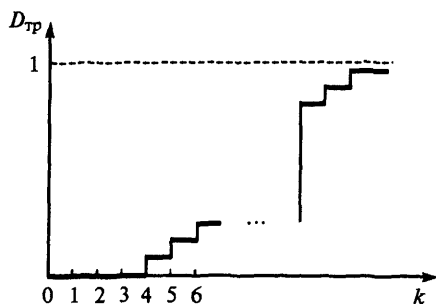


Рис. 6.13. Нарастающая вероятность правильного обнаружения траектории

траектории в зависимости от такта наблюдения цели $D_{\text{тр}} = D_{\text{тр}}(k)$ приведен на рис. 6.13.

Объединяя результаты вычисления вероятностей ложного и правильного обнаружения траекторий, можно получить все искомые вероятностные характеристики рассматриваемых логических алгоритмов.

6.7. Завязка траекторий

Завязка траекторий — это операция формирования новой траектории для ее дальнейшего анализа алгоритмом ВО.

Операция завязки является в определенном смысле вспомогательной, поскольку информация о существовании целевой траектории дополнительно проверяется в ходе дальнейшей обработки, в процессе которой также уточняются ее параметры. Однако от характеристик завязки во многом зависит построение других алгоритмов ВО и жесткость предъявляемых к ним требований.

Традиционно в условиях достаточно хорошего качества входной информации при выполнении операции завязки происходит обнаружение траектории и формирование предварительных данных о ее параметрах, при выполнении других операций — собственно сопровождение траектории и ее фильтрация [67]. В этом случае алгоритм завязки непосредственно определяет вероятностные характеристики всей ВО.

При усложнении целевой и помеховой обстановки удастся лишь на основе тесного взаимодействия всех операций ВО, реализующих не последовательное, а совместное обнаружение—оценивание траекторий, достичь требуемых как вероятностных, так и точностных характеристик. В этой ситуации алгоритм завязки отвечает за рациональное использование ресурсов РЛС, поддерживая компромисс между допустимым числом ложных траек-

тор-строку $\mathbf{B}(0) = [1 \ 0 \ 0 \ \dots]$, т. е. в данном случае предполагается, что в момент поступления первого отсчета от цели обнаружитель находился в 0-м состоянии. В некоторых случаях в качестве начального состояния выбираются финальные состояния, полученные при анализе обнаружителя для ложных отсчетов.

Типичный вид зависимостей вероятностей правильного обнаружения

торий, анализируемых алгоритмом ВО, скоростью завязки и ошибочным сбросом целевых отсчетов.

В ходе операции завязки анализируются отсчеты, принадлежность которых к уже сопровождаемым траекториям или местным предметам не установлена. Необходимо проверить любой такой «свободный» отсчет: не является ли он возможным началом траектории некоторой цели.

Наиболее часто решение о завязке (предварительном обнаружении) траектории принимается логическим обнаружителем после получения в течение нескольких тактов наблюдения отсчетов, не противоречащих возможным перемещениям цели в пределах корреляционных стробов. Поскольку в начале наблюдения априорная информация о траекторных параметрах, как правило, крайне скудна, размеры строба при завязке выбираются большими, главным образом, исходя из представлений о максимальной и минимальной скорости возможных целей.

Алгоритм завязки начинает строить траектории по каждому «свободному» отсчету, полученному после операции стробирования в некотором такте t_k наблюдения. С этой целью на следующем такте t_{k+1} для этого отсчета строится корреляционный строб. В зависимости от наличия или отсутствия отсчетов в этом стробе и в соответствии с критерием завязки и сброса (см. § 6.6) осуществляется завязывание траектории. По двум отсчетам уже может формироваться оценка скорости движения объекта, корреляционный строб на следующем такте с учетом этой информации может быть соответствующим образом уменьшен. Работа алгоритма завязки, включающая этапы: получение «свободного» отсчета, построение корреляционного строба, отбор отсчета, попавшего в строб, или констатация факта его отсутствия, проверка ситуации логическим критерием завязки и сброса, нахождение оценок завязываемой траектории, построение нового корреляционного строба и др., — продолжается несколько тактов. На этапе завязки выполняется практически полный набор всех операций ВО, но с использованием достаточно приближенных моделей целевой обстановки.

Как правило, оценка траекторных параметров на этапе завязки осуществляется с использованием соотношений для фиксированной выборки (см. § 6.3). Как только вектор оцениваемых траекторных параметров X_j некоторой j -й траектории становится полным и достигнуты вероятностные характеристики завязки, траектория по этим отсчетам считается завязанной и дальнейшая ее обработка происходит в основном тракте ВО. Полученные параметры траектории при этом обычно используются в качестве стартовой точки для алгоритма калмановской фильтрации. Если критерий завязки не выполняется, то происходит сброс завязываемой траектории.

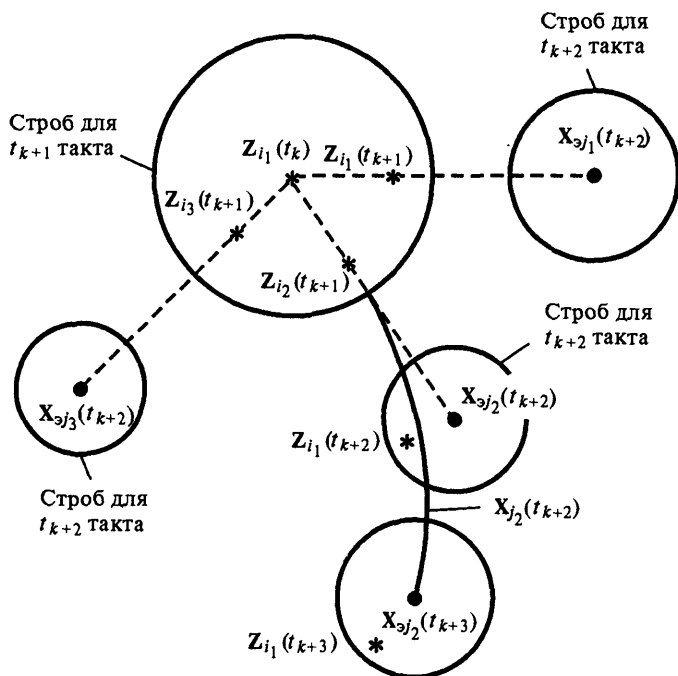


Рис. 6.14. Завязка траекторий: $X_{j_2}(k+2)$ — завязанная траектория

Для завязки траектории чаще всего используются логические критерии, имеющие хорошие фильтрующие (по помеховым отсчетам) свойства типа «2 из 2», «3 из 3» и т. п. Пример возможных вариантов нормального течения процедуры завязки для правила «3 из 3» показан на рис. 6.14.

Для некоторого отсчета $Z_{i_l}(t_k)$, полученного на k -м такте, строится строб для $(k+1)$ -го такта. В этот строб, как видно из рис. 6.14, на $(k+1)$ -м такте попало три отсчета $Z_{i_1}(t_{k+1})$, $Z_{i_2}(t_{k+1})$, $Z_{i_3}(t_{k+1})$. По ним завязываются и строятся три соответствующие траектории, для которых находятся экстраполированные параметры траекторий $X_{эli}$ и отсчетов $Z_{эli} = \mathbf{H}X_{эli}$ и формируются соответствующие стробы для $(k+2)$ -го такта. После получения информации с выхода первичной обработки на $(k+2)$ -м такте траектория X_{j_1} сбрасывается, поскольку в ее строб не попало ни одного отсчета, траектория X_{j_2} завязывается по попавшему в строб отсчету (выполняется правило «3 из 3»), траектория X_{j_3} сбрасывается, поскольку в ее строб не попало ни

одного отсчета. Траектория X_{j_2} на $(k + 2)$ -м такте считается завязанной и передается на обработку в основной тракт ВО.

Необходимо, чтобы критерий сброса также учитывал размеры строга экстраполяции. Если размеры корреляционного строга превышают размеры начального, то траектория сбрасывается даже без проверки на попадание в него отсчетов.

Операция завязки во многом носит эвристический характер, она учитывает по возможности всю доступную, обычно сравнительно скудную на этапе завязки, априорную информацию. По этой причине для получения нужных характеристик операции завязки требуется моделирование и дополнительное определение ее параметров в ходе испытаний.

Рассмотренный подход к реализации операции завязки применяется в случае разрешенного потока целей и сравнительно низкой интенсивности помеховых сигналов. При усложнении целевой и помеховой обстановки — увеличении плотности помеховых и целевых сигналов — завязку делают по всем отсчетам, попавшим в корреляционные стробы на протяжении нескольких тактов, строя всевозможные варианты траекторий, с тем чтобы ложные траектории сбросить в ходе дальнейшей обработки. Это приводит к необходимости обработки все возрастающего объема информации и разработки обоснованных критериев сброса ложных траекторий.

Контрольные вопросы

1. Что такое траектория движения цели?
2. Перечислите основные операции вторичной обработки радиолокационной информации.
3. Назовите отличия условий проведения вторичной обработки радиолокационной информации от первичной.
4. Дайте определение вектора траекторных параметров и вектора параметров отсчета.
5. Поясните уравнения состояния динамической системы.
6. Поясните уравнения измерений.
7. В чем разница между линейной траекторией и линейной моделью?
8. Перечислите ограничения при нахождении оценок траекторных параметров по фиксированной выборке.
9. Поясните методику проверки допустимости аппроксимации траектории полиномом выбранной степени.
10. Как определяется селекция методом стробирования?
11. Как определяется селекция методом «ближайшего соседа»?
12. Сформулируйте предположения о модели движения цели и модели измерений при проведении калмановской фильтрации.
13. От чего зависит коэффициент усиления фильтра Калмана?

14. Что такое стационарный режим работы фильтра Калмана?
15. Как проверяется состоятельность фильтра Калмана?
16. Когда допустимо применение $(\alpha-\beta)$ -фильтров?
17. Что такое расширенный фильтр Калмана?
18. Поясните методику расчета вероятностных характеристик логического обнаружителя при гипотезе H_0 .
19. Поясните методику расчета вероятностных характеристик логического обнаружителя при гипотезе H_1 .
20. Какие задачи решает операция завязки траекторий?

7. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В РАДИОЛОКАЦИОННЫХ СИСТЕМАХ

Рассмотрена обобщенная структурная схема РЛС, приведены основные технологии, которые используются или могут быть использованы в современных и перспективных РЛС. Изложены особенности обработки радиолокационных сигналов на фоне помех, в том числе методы адаптивной обработки радиолокационных сигналов на фоне активных и пассивных маскирующих помех. Описаны устройства адаптивной обработки сигналов, использующие корреляционные автокомпенсаторы. Приведены основные тактико-технические характеристики ряда современных РЛС.

7.1. Принципы построения радиолокационных систем

7.1.1. Задачи и условия функционирования радиолокационных систем

Радиолокация — отрасль радиотехники, обеспечивающая получение сведений об объектах путем приема и анализа радиоволн [31, 35, 43, 72—84]. Объекты, сведения о которых необходимо получить, называют *радиолокационными целями*. Различают следующие цели: аэродинамические (самолеты, крылатые ракеты, вертолеты и др.), наземные и надводные (автомашины, танки, корабли и др.), космические (космические аппараты, баллистические ракеты и др.), подземные и подводные (полости в грунте, различные объекты в земле и воде и др.), природного происхождения (облака, естественные ориентиры на местности, метеоры, планеты) и другие.

Совокупность сведений о наличии целей в отдельных областях пространства, об их координатах и других параметрах движения, о числе целей и их характеристиках называют *радиолокационной информацией*. Технические средства получения радиолокационной информации называют *радиолокационными средствами*, *радиолокационными станциями* (РЛС), или *радиолокаторами*. Для расширения информационных возможностей радиолокационных средств их объединяют в *радиолокационные системы*

(комплексы), включающие средства передачи данных и управления [31, 35, 43, 72—84].

Термин «радиолокация» составлен из латинских слов *locus* — *место* и *radio* — *излучение*, характеризующих важнейшую из решаемых задач и пути ее решения. В зарубежной литературе используется термин «радар» (*radar*), происходящий из словосочетания *radio detection and ranging* (от англ. *обнаружение и измерение дальности с помощью радиоволн*). В современных РЛС используются электромагнитные излучения декаметровых, метровых, дециметровых, сантиметровых и миллиметровых волн.

Основными информационными задачами радиолокации являются следующие:

- обнаружение целей;
- измерение координат целей и других параметров их движения;
- разрешение целей;
- классификация целей.

Эти задачи решаются на всех этапах обработки радиолокационной информации: первичной, вторичной и третичной.

Задача *обнаружения* состоит в принятии решения о наличии или отсутствии цели в каждом выделенном элементе пространства, входящем в зону ответственности (контроля) РЛС, с минимальными вероятностями ошибок при первичной обработке и во всей зоне ответственности РЛС при вторичной (третичной) обработке [31, 43, 73, 74, 83].

Задача *измерения* сводится к оцениванию координат и других параметров движения целей с минимально возможными погрешностями. Измеряют, в первую очередь, *дальность* до цели $r_{ц}$, ее *азимут* $\beta_{ц}$, *угол места* $\epsilon_{ц}$, а также *производные координат* (в частности, *радиальную скорость* $v_{ц} = dr_{ц}/dt$), *элементы траектории* (см. гл. 5 и 6). Могут измеряться параметры, не связанные непосредственно с координатами целей: элементы поляризационной матрицы рассеяния, радиальная протяженность целей и др.

Задача *разрешения* заключается в обнаружении и измерении параметров произвольной цели в присутствии других объектов (целей) (см. гл. 4).

Задача *классификации* (распознавания) состоит в установлении принадлежности цели к определенному классу и разделяется на решение двух основных задач. Первая состоит в определении государственной принадлежности «свой—чужой» с помощью запросно-ответных устройств *опознавания*, установленных на своих объектах, вторая — в распознавании цели, не отвечающей на запрос [31, 43, 74, 85].

Всю совокупность информационных задач радиолокации характеризуют часто термином *радиолокационное наблюдение*.

Информационные задачи решаются за ограниченное время. Для первичной обработки это время определяется временем однократного контакта РЛС с целью, для вторичной — временем нахождения цели в зоне ответственности РЛС. Обнаружение, измерение и разрешение часто представляют единый процесс *обнаружения—измерения—разрешения*, а визуальное наблюдение дополняется автоматизированным (без участия оператора) или полуавтоматизированным (с участием оператора). Скоротечность изменения радиолокационной обстановки требует высокого темпа выдачи данных.

На радиолокационные средства воздействуют помехи природного происхождения (естественные), от других радиоэлектронных средств (взаимные), а в ряде случаев — организованные (умышленные). Помехи природного происхождения в той или иной степени воздействуют на любое радиолокационное средство. Влияние взаимных помех в последнее время значительно возросло в связи с внедрением различных радиоэлектронных средств в народное хозяйство и военную технику; стало актуальным обеспечение электромагнитной совместимости радиоэлектронных средств. Основными способами противодействия средствам военной радиолокации со стороны противника являются создание организованных помех и прямое уничтожение РЛС, т. е. их радиоэлектронное подавление и огневое поражение.

Независимо от происхождения различают помехи в виде мешающих излучений — *активные*, мешающих отражений — *пассивные*, их комбинаций — *комбинированные*. Помехи могут *маскировать* полезные сигналы или *имитировать* цели, снижая эффективность радиолокационных средств и систем. В связи с этим к РЛС предъявляются требования по обеспечению *помехозащищенности*, т. е. по поддержанию качества информации в помеховых ситуациях на допустимом уровне [86—88]. Для повышения помехозащищенности используют различные меры защиты от помех, в том числе приспособление (адаптацию) к помеховой обстановке и использование помех для непосредственного получения информации о целях (постановщиках помех). Целесообразное объединение радиолокационных средств в системы — одна из мер повышения качества информации в сложных конфликтных ситуациях. Объединять необходимо достаточно информативные средства. Однако объединение и малоинформативных средств может повысить эффективность радиолокации [74, 76, 83, 89, 98].

Для радиолокационных средств военного назначения важным требованием является обеспечение их *живучести* в условиях применения против них ракет, самонаводящихся на излучение, и другого высокоточного оружия противника без ухудшения информационных возможностей [74, 90].

Таким образом, усложнение условий работы, необходимость повышения качества наблюдения и живучести радиолокационных средств и систем требуют развития информационных технологий для всестороннего исполь-

зования современных возможностей получения радиолокационной информации в пределах допустимых экономических затрат.

7.1.2. Принципы получения радиолокационной информации и построения радиолокационных систем

Носителями информации о целях являются принимаемые радиолокационные сигналы. Возможность приема радиолокационных сигналов обеспечивается в результате вторичного излучения, переизлучения или собственного излучения радиоволн. Различают, соответственно, активную радиолокацию с пассивным ответом, активную радиолокацию с активным ответом и пассивную радиолокацию. Возможна также пассивная радиолокация с активным ответом [43].

Активная радиолокация с пассивным ответом основана на использовании *эффекта вторичного излучения* (отражения) радиоволн (рис. 7.1, а). Активный ее характер состоит в облучении пространства мощными зондирующими колебаниями. Пассивным ответом на облучение является вторичное излучение радиоволн. На характер активной радиолокации, кроме особенностей вторичного излучения, существенно влияет также *характер размещения при-*

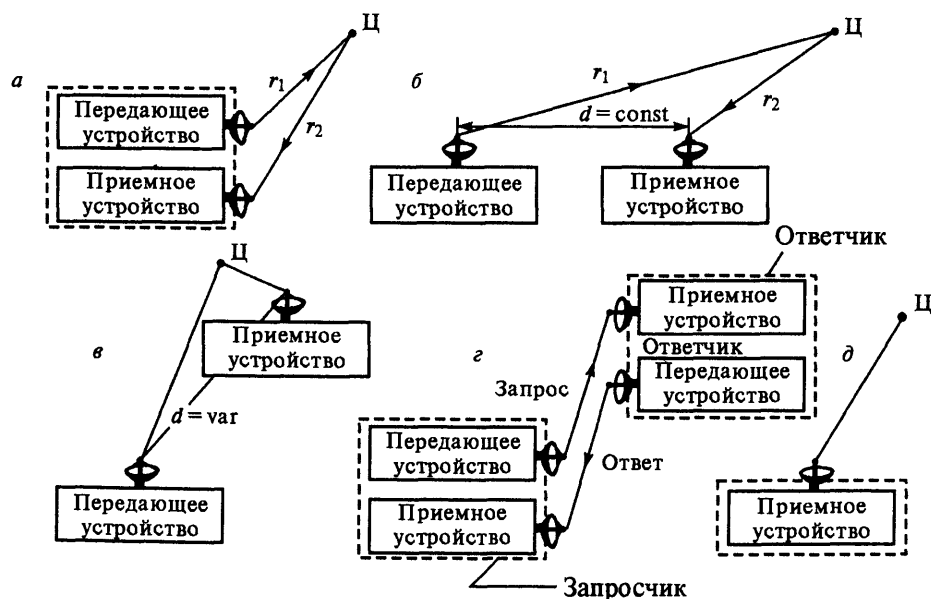


Рис. 7.1. Виды радиолокации:

а — активная радиолокация с пассивным ответом; б — разнесенная активная радиолокация с постоянной базой и в — с переменной базой; г — активная радиолокация с активным ответом; д — пассивная радиолокация

емной и передающей аппаратуры. Если приемная позиция совмещена с передающей, то активное радиолокационное средство называют *совмещенным*. Совмещенное средство часто содержит одну антенну, коммутируемую поочередно на передачу и прием сигналов. Возможен разнос приемной и передающей позиций на расстояние d , называемое *базой*. Базы бывают не только постоянными, $d = \text{const}$ (рис. 7.1, б), но и *переменными* (рис. 7.1, в). Приемный пункт — головка самонаведения (см. рис. 7.1, в) — располагается, в частности, на ракете, так что $d = \text{var}$. Наряду с однобазовыми (двухпозиционными) разнесенными активными радиолокационными средствами возможны многобазовые (многопозиционные). В связи с усложнением задач радиолокации интерес к разнесенным радиолокационным средствам в последнее время возрастает [73, 76, 89].

Радиолокация с активным ответом (вторичная радиолокация) — это активная радиолокация с активным ответом. Такой вид локации позволяет получать надежную информацию о своих объектах (например, о кораблях, самолетах и т. д.). Их облучают запросными (зондирующими) сигналами (рис. 7.1, г). На объектах устанавливают ответчики, т. е. приемопередатчики, переизлучающие (ретранслирующие) принятые сигналы. Несущие частоты и законы модуляции (коды) запросных и ответных сигналов могут изменяться в широких пределах. Это обеспечивает опознавание государственной принадлежности объектов («свой—чужой») и индивидуальное опознавание. Активный ответ широко применяется также для радионавигации самолетов, морских судов и т. д.

Пассивная радиолокация использует собственные излучения элементов цели и их ближайшей окрестности. Излучения создают нагретые участки поверхности, радиолокационные и радионавигационные средства, средства радиосвязи и радиоэлектронного подавления, ионизированные образования различного вида и др. В общем случае средство пассивной радиолокации может быть размещено на одной (рис. 7.1, д) или нескольких разнесенных позициях. На принципах пассивной радиолокации работают, в частности, средства радиотехнической разведки излучений.

Пассивные и активные радиолокационные средства можно использовать совместно. Тогда говорят об *активно-пассивных* радиолокационных средствах (системах или комплексах).

Важное значение для активных и активно-пассивных радиолокационных средств имеет характер зондирования пространства. Высокая направленность зондирующего излучения обеспечивает концентрацию его энергии, облегчая последующее выделение отраженных сигналов, поэтому зондирование различных участков пространства часто проводится одновременно, т. е. наряду с одновременным обзором участков пространства реализуется последовательный обзор. Поскольку колебания, излучаемые в каждом направлении, обычно модулированы во времени, законы модуляции для раз-

личных направлений не совпадают. Это позволяет говорить о *пространственно-временной модуляции зондирующих колебаний*. Она достигается временной модуляцией в передатчиках и перемещением диаграмм направленности передающих антенн. Возможные виды пространственно-временной модуляции обеспечивают последовательный обзор пространства по *жесткой* либо по *гибкой* программе, в зависимости от результатов текущих наблюдений. Для повышения оперативности обзора используют антенны с *электрическим* управлением положением луча: антенные решетки, антенны с частотным качанием (сканированием луча).

На характеристики принимаемых радиолокационных сигналов и их использование для измерения координат целей влияют *свойства среды*, в которой распространяются радиоволны. Простейшим является случай распространения радиоволн в свободном пространстве, которое считаем однородным, изотропным, недиспергирующим. Это значит, что скорость распространения радиоволн:

- одинакова для всех элементов этого пространства;
- не зависит от направления распространения и поляризации волны;
- не зависит от частоты колебаний ($c \approx 3 \cdot 10^8$ м/с).

Зондирующий и отраженный сигналы распространяются по прямолинейным траекториям без искажения своей формы. Времена запаздывания t_3 отраженных от точечных целей сигналов относительно зондирующих определяются для разнесенных (см. рис. 7.1, б) и совмещенных (см. рис. 7.1, а) радиолокаторов соответственно соотношениями

$$t_3 = (r_1 + r_2)/c, \quad t_3 = 2r/c.$$

Информация о дальности до цели заключена во временной структуре принимаемых колебаний. Для совмещенных радиолокаторов дальность однозначно определяется временем запаздывания:

$$r_u = ct_3/2.$$

При использовании разнесенных пунктов приема или одной многоэлементной антенны можно говорить о *пространственно-временной структуре* принимаемых колебаний. Набор временных запаздываний характеризует не только дальности, но и *угловые координаты целей*.

При малом разnose приемных элементов (в пределах антенной решетки), когда разностью запаздываний огибающих сигналов до приемных элементов можно пренебречь, угловая координата цели определяется распределением начальных фаз принимаемых колебаний. С этим же распределением связано формирование характеристик направленности антенн.

При вращении антенной системы (сканировании), можно сравнительно просто измерять угловые координаты — азимуты и углы места целей,

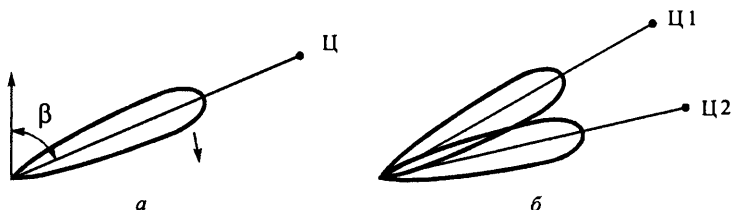


Рис. 7.2. Определение угловой координаты цели (а), угловое разрешение целей (б)

например по оценке временного положения максимума отраженного сигнала (рис. 7.2, а), обеспечивать их угловое разрешение (рис. 7.2, б). При одноканальном приеме информация о различных угловых направлениях поступает последовательно во времени, при многоканальном, когда характеристики (см. рис. 7.2, б) относятся к разным каналам приема, ее можно получать параллельно, практически одновременно.

Реализацию основных операций обнаружения целей, измерения их угловых координат и дальности поясним на примере структурной схемы (рис. 7.3) простейшего совмещенного активного импульсного радиолокатора с общей приемопередающей антенной и одним приемником. Важным элементом радиолокатора является *синхронизатор*, определяющий последовательность работы его основных элементов. Зондирование короткими радиоимпульсами обеспечивает неодновременность приема и излучения. Это позволяет использовать общую антенну, коммутируемую антенным переключателем на передачу и прием. После излучения зондирующего радиоимпульса антенна соединяется с приемником. Индикаторное устройство обеспечивает возможность обнаружения оператором вторичного излучения цели, измерения дальности до цели и ее угловых координат. Системы автоматики связывают индикаторное устройство с антенной, что позволяет

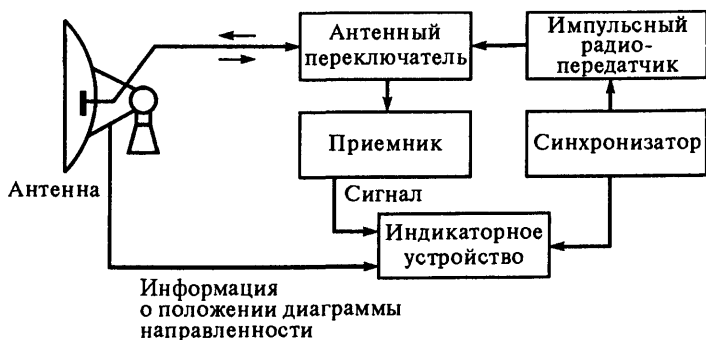


Рис. 7.3. Структурная схема простейшего совмещенного активного импульсного радиолокатора

получать информацию о положении диаграммы направленности антенны, а значит, об угловых координатах целей, а также управлять антенной (система управления на рис. 7.3 не показана).

В общем случае прием может быть многоканальным, длительность сигнала не должна быть обязательно малой. Приемная и передающая антенны могут быть пространственно разделены (даже в условиях совмещенной локации).

Возможности съема и обработки данных существенно расширяются при использовании средств вычислительной техники, обеспечивающих более полную автоматизацию радиолокационного наблюдения.

В случае радиолокации движущихся целей происходит изменение временных запаздываний отдельных элементов, а значит, всей структуры сигналов. Так, радиальное движение цели относительно совмещенного импульсного радиолокатора:

- 1) изменяет запаздывания последовательно принимаемых импульсов;
- 2) приводит к известному из физики изменению несущей частоты — эффекту Доплера.

Оба эффекта в отдельности можно использовать для измерения радиальных скоростей целей и их скоростного разрешения. Селекцию по скорости широко используют для защиты от пассивных помех.

При любом из методов радиолокации приходящие сигналы оказываются часто слабыми. Особенно это относится к активной радиолокации, когда ослабление обусловлено двукратным рассеянием энергии: на пути до цели и обратно. Принимают ряд мер для выделения слабых сигналов: увеличивают по возможности габариты передающей и приемной антенн, среднюю мощность зондирующих колебаний; применяют высокочувствительные (малошумящие) входные элементы радиоприемных устройств. Оптимизируют наряду с этим обработку принимаемых колебаний с учетом внешних помех и внутренних шумов приемника.

Оптимизация обработки означает наилучший (в статистическом смысле) учет взаимных различий сигналов и помех. Учет этот существен на всех этапах получения радиолокационной информации, в первую очередь, при обнаружении цели и измерении их координат и параметров движения. Радиоприемное устройство по существу становится специализированной ЭВМ, точно или приближенно выполняющей линейные и нелинейные операции оптимальной обработки принимаемых колебаний. Постепенно стираются грани между обработкой в цепях приемника, элементах автоматики и вычислительной техники. Существенна лишь совокупность выполняемых операций, подлежащих совместной оптимизации. При использовании многоэлементных антенн необходимые операции оптимальной обработки проводят уже над колебаниями, принятыми элементами антенн. Антенные операции оказываются начальными звеньями единой цепи обработки (аналого-

вой, цифровой, комбинированной). Оптимизируя обработку, учитывают также условия распространения радиоволн в средах, отличающихся от свободного пространства.

Параметры помех и сред распространения обычно заранее неизвестны. Важна поэтому адаптация (приспособление) пространственно-временной обработки сигналов к текущим условиям работы радиолокационного средства или системы. Наряду с этим повышается роль адаптации по отношению к пространственно-временной модуляции излучаемых сигналов.

7.2. Классификация радиолокационных систем

Классификацию радиолокационных систем так же, как и классификацию радиотехнических систем, можно проводить по различным признакам. В зависимости от используемых классификационных признаков радиолокационные системы подразделяются: по месту установки (на наземные, корабельные, авиационные, космического базирования); по назначению (на РЛС обнаружения целей, управления оружием, обеспечения полетов, метеорологические, навигационные, опознавания государственной принадлежности, многофункциональные); по рабочему диапазону длин волн (на станции дециметрового, метрового, дециметрового, сантиметрового, миллиметрового диапазонов длин волн, многодиапазонные); по виду излучения (на РЛС импульсного, непрерывного, квазинепрерывного, шумового и комбинированного излучения); по числу измеряемых координат (на двухкоординатные — обычно дальность и азимут, трехкоординатные — обычно дальность, азимут и угол места); по числу занимаемых позиций (на однопозиционные и многопозиционные).

Рассмотрим задачи, решаемые РЛС в зависимости от места их установки.

Наземные РЛС можно разделить на РЛС надгоризонтного обнаружения (НГО), загоризонтного обнаружения (ЗГО) и подповерхностной радиолокации [6, 35, 72, 77, 79].

Системы НГО работают в метровом, дециметровом, сантиметровом и миллиметровом диапазонах длин волн (0,03...300 ГГц), подразделяются на стационарные и подвижные (мобильные): самоходные, буксируемые, возимые, переносные. По решаемым задачам (назначению) они подразделяются на РЛС:

- управления воздушным движением (УВД);
- обнаружения, наведения и целеуказания;
- обнаружения маловысотных целей;
- наведения зенитных управляемых ракет;
- орудийной наводки;

— радиолокационной разведки на поле боя: наземной разведки, наземной артиллерийской разведки, обнаружения стреляющих минометов и стартующих ракет;

— высотомеры;

— предупреждения о ракетном нападении (ПРН);

— противоракетной обороны (ПРО);

— контроля космического пространства (ККП);

— полигонные;

— метеорологические и др. [35, 72].

РЛС загоризонтного обнаружения (или загоризонтные РЛС) основаны на использовании эффекта отражения радиоволн декаметрового диапазона (3...30 МГц) от ионосферы Земли и работают в режиме обратного рассеяния радиоволн, при котором сигнал принимается в месте излучения, или же в режиме прямого рассеяния, когда падающий и рассеянный потоки радиоволн распространяются в одну сторону [78]. Загоризонтные РЛС могут быть односкачковыми и многоскачковыми. Они предназначены для:

— наблюдения на больших площадях за состоянием поверхности морей и океанов, а также за движением кораблей и самолетов;

— обнаружения областей с отчетливо выраженной плазменной неоднородностью, создаваемой стартующими баллистическими ракетами и метеорными следами;

— ионосферных исследований.

Дальность действия загоризонтных РЛС достигает нескольких тысяч километров [72, 78].

Радиолокаторы подповерхностного зондирования предназначены для обнаружения полостей в грунте, различных объектов, сооружений из бетона, определения толщины льда, подводной радиолокации и др. Достижимая глубина проникновения может составлять до нескольких сотен метров [6, 72].

Корабельные РЛС предназначены для обнаружения и сопровождения воздушных и надводных целей, обзора надводной и береговой поверхности, целеуказания, наведения зенитных управляемых ракет и орудий, а также для кораблевождения и навигации (определения местонахождения кораблей, их скорости и проверки курса). Для обеспечения необходимого обзора антенные системы корабельных РЛС устанавливаются на мачтах, а для устранения влияния качки корабля стабилизируют или расширяют сектор обзора по углу места. На современных кораблях число РЛС может быть более 50 [72, 80].

Авиационные (самолетные) РЛС делятся на РЛС обзора воздушного пространства, РЛС землеобзора, multifunctional РЛС.

К РЛС обзора воздушного пространства относятся системы:

- перехвата и прицеливания;
 - дальнего радиолокационного обнаружения (дозора) и наведения (управления);
 - защиты своих самолетов;
 - обхода препятствий в воздухе (например, грозовых образований) и др.
- К радиолокационным системам *землеобзора* относятся:
- панорамные;
 - бокового обзора с антенной, расположенной вдоль фюзеляжа;
 - бокового обзора с синтезированной апертурой (РСА);
 - подповерхностной радиолокации [72, 81].

РЛС космического базирования применяются для дистанционного исследования (в том числе и картографирования) Земли и планет, обеспечения сближения, стыковки и посадки космических аппаратов. Обсуждаются также возможности создания и использования РЛС космического базирования для решения задач противовоздушной и противокосмической обороны [72, 82].

7.3. **Тактико-технические характеристики радиолокационных систем**

Радиолокационные системы, как отмечено в гл. 1, относятся к классу радиотехнических систем извлечения информации. Поэтому РЛС можно описать теми же характеристиками, которые применяются для описания радиотехнических систем в целом (см. § 1.3), но с учетом задач, решаемых радиолокационными системами.

Различают тактические и технические характеристики РЛС. *Тактические* характеристики показывают назначение и возможности радиолокационной системы, а *технические* — описывают основные параметры ее устройств. В совокупности они представляют тактико-технические характеристики (ТТХ) радиолокационных систем [35, 43, 77].

7.3.1. **Тактические характеристики РЛС**

К тактическим характеристикам РЛС относятся назначение, сектор или зона работы, время обзора этого сектора, качественные показатели (показатели эффективности) обнаружения объекта, число измеряемых координат и параметров движения объекта и точность этих измерений, вид выходных данных, разрешающая способность, пропускная способность, помехозащищенность, надежность, экономичность и др.

Назначение радиолокационных систем приведено в § 7.2.

Сектор или зона работы описываются пределами изменения координат, обнаруживаемых или обследуемых объектов, т. е. минимальными и максимальными значениями дальности, азимута и угла места (или высоты).

Время обзора заданного сектора — это характеристика, определяющая темп выдачи данных по результатам работы РЛС и являющаяся исключительно важной, особенно при работе по высокоскоростным объектам (самолет, баллистическая ракета, космический аппарат и др.). Время обзора заданного сектора радиолокационной станцией должно быть в несколько раз меньше времени, которое требуется лоцируемому объекту для пересечения этого сектора в любом направлении.

Качественные показатели (показатели эффективности) обнаружения объекта — это условная вероятность правильного обнаружения D и условная вероятность ложной тревоги $F_{\text{лт}}$. (Определение и вычисление этих показателей проводились в гл. 3.)

Число измеряемых координат и параметров движения объекта и точность этих измерений. В радиолокационных системах противовоздушной и особенно противоракетной и противокосмической обороны требуется измерение не только трех координат лоцируемых объектов, но и их первых, а иногда и вторых производных. Для наблюдения за наземными и надводными объектами достаточно измерения только двух координат: дальности и азимута. В некоторых случаях ограничиваются измерением и одной координаты. Точность измерения координат объекта и параметров его движения характеризуется корреляционной матрицей ошибок измерения, в частности среднеквадратическими ошибками, и другими параметрами. Вопросы точности измерения рассматривались в гл. 5.

Вид выходных данных — это световой, звуковой или иной сигнал об обнаружении объекта, отметка на экране электронно-лучевой трубки, показания стрелочного прибора о координате объекта или параметре его движения, цифровые коды указанных координат и параметров, передаваемые по каналу связи потребителей (ЭВМ командного пункта управления системой ПВО, ПРО, пункта управления воздушным движением и др.).

Разрешающая способность — это способность раздельного обнаружения и измерения координат или параметров движения близко расположенных объектов. Вопросы разрешения сигналов (и объектов) рассматривались в гл. 4.

Пропускная способность — это способность РЛС работать с большим числом объектов. Количественно характеризуется максимальным числом объектов, одновременно обслуживаемых радиолокационной системой. Например, импульсный радиолокатор может одновременно работать приблизительно с тысячей объектов. Его пропускная способность равна отноше-

нию разности максимальной и минимальной дальностей к разрешающей способности по дальности:

$$n \approx (r_{\max} - r_{\min}) / \Delta r.$$

В последнее время пропускную способность РЛС характеризуют числом трасс целей, выдаваемых за один обзор.

Помехозащищенность — это способность РЛС поддерживать на заданном уровне тактические характеристики (и, прежде всего, показатели качества обнаружения и измерения параметров) в условиях действия различного рода помех. Помехозащищенность РЛС определяется ее скрытностью, т. е. способностью затруднять противнику создание организованных радиопомех на входе приемника, и помехоустойчивостью, характеризующей способность нормального функционирования РЛС в условиях наличия на ее входе естественных, взаимных и организованных помех. В свою очередь, помехозащищенность является составной частью *живучести*, т. е. способности РЛС сохранять свою эффективность в условиях огневого и помехового воздействия.

Помехозащищенность (помехоустойчивость) характеризуют коэффициентом подавления помех, дальностями обнаружения цели в условиях воздействия помех с различной интенсивностью.

Надежность — это свойство РЛС сохранять на установленном уровне тактические характеристики при заданных условиях эксплуатации. Количественно надежность описывается вероятностью безотказной работы в течение установленного интервала времени, средним временем исправной работы, частотой отказов в работе, средним временем восстановления.

Экономичность характеризуется стоимостью затрат на разработку, производство и эксплуатацию (включая и ремонт).

Мобильность (маневренность) — это способность РЛС быстро перемещаться, развертываться и вести работу в высоком темпе при различных условиях обстановки. Характеризуется временем развертывания и свертывания РЛС, скоростью перемещения.

7.3.2. Технические характеристики РЛС

Основными техническими характеристиками РЛС являются следующие: режим работы и вид модуляции (манипуляции) зондирующих сигналов, диапазон частот (длин волн), энергетические характеристики (импульсная и средняя мощности передатчика, чувствительность приемника), характеристики антенн, способы обзора пространства, характеристики вычислительных средств, используемая элементная база, число и тип выходных уст-

роиств, весогабаритные характеристики, мощность первичных источников питания и др.

Режим работы и вид модуляции (манипуляции) зондирующих сигналов. Режим работы РЛС определяется видом (характером) излучения зондирующих сигналов и может быть непрерывным или импульсным. В свою очередь, РЛС с импульсным режимом излучения иногда разделяют на РЛС с большой скважностью $Q \approx 50 \dots 2000$ и на РЛС с малой скважностью $Q \approx 2 \dots 50$ излучения или РЛС с квазинепрерывным излучением, $Q = T_{\text{п}}/\tau_{\text{и}}$. Импульсные сигналы можно представить как непрерывный гармонический сигнал, промодулированный по амплитуде импульсной последовательностью. Кроме амплитудной, используют частотную (чаще линейную частотную) модуляцию, фазовую (фазокодовую) и частотную манипуляции, шумовую модуляцию и комбинированные виды модуляции и манипуляции зондирующих сигналов независимо от режимов их излучения.

Диапазоны длин волн (частот). В РЛС используют диапазоны декаметровых, метровых, дециметровых, сантиметровых и миллиметровых волн. Декаметровый диапазон длин волн используют в загоризонтной радиолокации. Метровый, дециметровый, сантиметровый и миллиметровый диапазоны длин волн используют в надгоризонтной локации. Это объясняется следующими причинами.

Радиолокация основана на использовании явления отражения радиоволн от объектов. Высокая интенсивность отражений от некоторого объекта обеспечивается, если длина волны λ много меньше габаритного размера $l_{\text{об}}$ этого объекта ($\lambda \ll l_{\text{об}}$) или соизмерима с его размерами, причем в этом случае желательно выполнение условия $\lambda \sim l_{\text{об}}/2$. Поскольку наиболее часто наблюдаемые объекты (космические, аэродинамические, наземные и др.) имеют габаритные размеры от одного до десяти метров, то используемая в РЛС длина волны должна быть меньше единиц метров.

Для получения большой дальности действия, высокой разрешающей способности по угловым координатам и точности их измерений необходимо использование остронаправленных антенн. Направленное излучение и прием радиоволн достигаются, как известно, только в том случае, когда используемая длина волны много меньше габаритных размеров антенны ($\lambda \ll l_{\text{а}}$). Поэтому с точки зрения конструирования остронаправленных и компактных антенн желательно применение малых длин волн.

Для получения высокой разрешающей способности по дальности $\Delta r = c/(2\Delta f_{\text{с}})$ необходимо применять сигналы с большой шириной спектра $\Delta f_{\text{с}}$. Величина $\Delta f_{\text{с}}$ обычно много меньше частоты несущей ($\Delta f_{\text{с}} \ll f_0$). Из этих двух выражений следует, что $\lambda = c/f_0 \ll 2\Delta r$, т. е. длина волны должна

быть много меньше удвоенного значения разрешающей способности по дальности.

Для достижения больших дальностей действия, высокой эффективности обнаружения сигналов и точности измерения их параметров уровень помех должен быть как можно меньше. В указанных диапазонах радиоволн это условие выполняется, так как уровень атмосферных и большинства промышленных помех достаточно мал, а квантовые эффекты еще не сказываются. Поэтому основным видом помех в РЛС является внутренний шум радиоприемных устройств. Для обработки информации, искаженной действием только внутренних шумов, обычно удается выполнить приемное устройство, близкое к оптимальному, и получить при этом практически потенциальные качественные показатели работы РЛС (малые вероятности ошибок обнаружения, высокую точность измерения параметров сигнала, большую дальность действия и т. д.).

Для повышения точности измерения малых радиальных скоростей объекта по величине доплеровского смещения частоты $F_D = 2v_r/\lambda$ необходимо выбирать длину волны как можно меньше.

При использовании в радиолокации диапазонов миллиметровых, сантиметровых, дециметровых и метровых радиоволн дальность действия РЛС ограничивается пределами прямой видимости, так как в указанных диапазонах радиоволны, как правило, распространяются без явлений рефракции и отражений от ионосферы, т. е. в первом приближении прямолинейно, при этом волны длиннее 3 см еще сравнительно слабо затухают в атмосфере и распространяются практически независимо от метеорологических условий (времени суток и года, наличия тумана, облаков, осадков и т. п.).

Наиболее часто в радиолокации используются сантиметровые и дециметровые волны с $\lambda = 3 \dots 30$ см. В то же время диапазон используемых длин волн постоянно расширяется.

Энергетические характеристики. Импульсная мощность зондирующих сигналов лежит в диапазоне единицы ватт — десятки мегаватт. Средняя мощность при импульсном режиме работы так же, как и мощность излучаемых сигналов при непрерывном режиме, может изменяться от долей ватта до десятков киловатт.

Чувствительность приемного устройства определяется предельной чувствительностью (см. формулу (1.1) в гл. 1). Обычно радиоприемные устройства РЛС имеют чувствительность от 10^{-12} до 10^{-22} Вт.

Характеристики антенны определяются диаграммой направленности антенны (ДНА). К основным характеристикам относятся ширина $\Delta\theta$ ДНА на уровне 0,707 (по напряжению) или 0,5 (по мощности) от максимального

значения и коэффициент направленного действия G (выигрыш по мощности в данном направлении вследствие направленности антенны), причем $G = 4\pi S_a / \lambda^2$, где S_a — эффективная площадь антенного устройства, пропорциональная геометрической площади S_r раскрыва антенны: $S_a = K S_r$. Коэффициент пропорциональности K , часто называемый *коэффициентом использования площади* (КИП), характеризует использование площади раскрыва антенны и зависит от распределения поля на этой площади. Если поле равномерно и синфазно по всей площади раскрыва антенны, то $K = 1$. Для большинства антенн РЛС $\Delta\theta \approx 0,1 \dots 10^\circ$, $G \approx 10^2 \dots 10^7$, а $K \approx 0,3 \dots 0,7$ [35, 73, 77].

Способ обзора пространства определяет последовательность просмотра различных разрешаемых объемов в заданном секторе обзора. Он выбирается в зависимости от назначения и структуры РЛС, ее тактических (размер сектора обзора, число определяемых координат, разрешающая способность по угловым координатам и т. д.) и технических (форма диаграммы направленности, вид ее сканирования и др.) характеристик. В однопозиционных РЛС используются одновременный (параллельный), последовательный и смешанный обзор. Последовательный обзор может быть круговым, секторным, спиральным и кадровым.

Основными *характеристиками вычислительных средств* являются быстродействие и объем памяти используемых процессоров. Быстродействие современных процессоров достигает сотен миллионов операций в секунду, их возможности постоянно растут.

Используемая элементная база СВЧ-трактов приемопередающих устройств. В зависимости от используемой СВЧ-технологии РЛС может быть полностью твердотельной или использующей СВЧ электровакуумную технологию.

Весогабаритные характеристики. Для мобильных РЛС обычно приводят значения массы и габариты кабин, для стационарных РЛС — размеры зданий, в которых размещается аппаратура.

Мощность первичных источников питания характеризуется максимальной мощностью, которая требуется для обеспечения работы РЛС.

Число и вид выходных устройств. Различают следующие виды съема выходных данных РЛС: автоматический (без участия человека-оператора), автоматизированный (с участием оператора), визуальный и ручной. Используют следующие виды индикаторов выходных данных: стрелочный (например, обычный вольтметр), цифровой (например, цифровой вольтметр), на электронно-лучевых трубках (дальности; кругового обзора, в котором применяется радиально-круговая развертка; секторного обзора; «дальность—угол места»; «дальность—высота»; «дальность—направление» и т. д.), дисплей цветного изображения.

Рассмотренные выше основные тактические и технические характеристики РЛС не являются изолированными друг от друга, а тесно связаны между собой.

7.4. Обобщенная структурная схема радиолокационных систем

Структура радиолокационных систем существенно зависит от их назначения, места установки, источника сигнала, методов измерения координат и т. д. Однако при всем многообразии радиолокационных систем можно выделить элементы, которые в той или иной степени присущи большинству современных, а также перспективных РЛС.

На рис. 7.4 приведена обобщенная структурная схема РЛС, содержащая системы (подсистемы), предназначенные для непосредственного получения радиолокационной информации (РЛИ), и дополнительные средства, которые могут включаться или не включаться (кроме системы питания) в состав РЛС в зависимости от ее назначения, места установки и т. д. Рассмотрим основные задачи, решаемые каждой системой, и возможные технологии, которые используются или предполагают использовать в современных и перспективных РЛС.

7.4.1. Основные системы для получения радиолокационной информации

Антенная система служит для излучения и приема радиоволн. В зависимости от назначения РЛС, места установки и используемого диапазона длин волн применяют следующие типы антенн:

- проволочные: спиральные и вибраторные (поперечного или осевого излучения);
- акустического типа: волноводные и рупорные;
- оптического типа: зеркальные (рефлекторные) и линзовые;
- поверхностных волн;
- с обработкой сигнала;
- с частотным качанием (сканированием) луча;
- фазированные антенные решетки (ФАР).

В современных РЛС все чаще используют активные ФАР с одно- и двумерным управлением, многодиапазонные приемопередающие ФАР, цифровые ФАР, адаптивные ФАР и др. [72, 73].

Приемопередающий модуль размещается как можно ближе к антенной системе и включает радиопередающие и радиоприемные устройства.

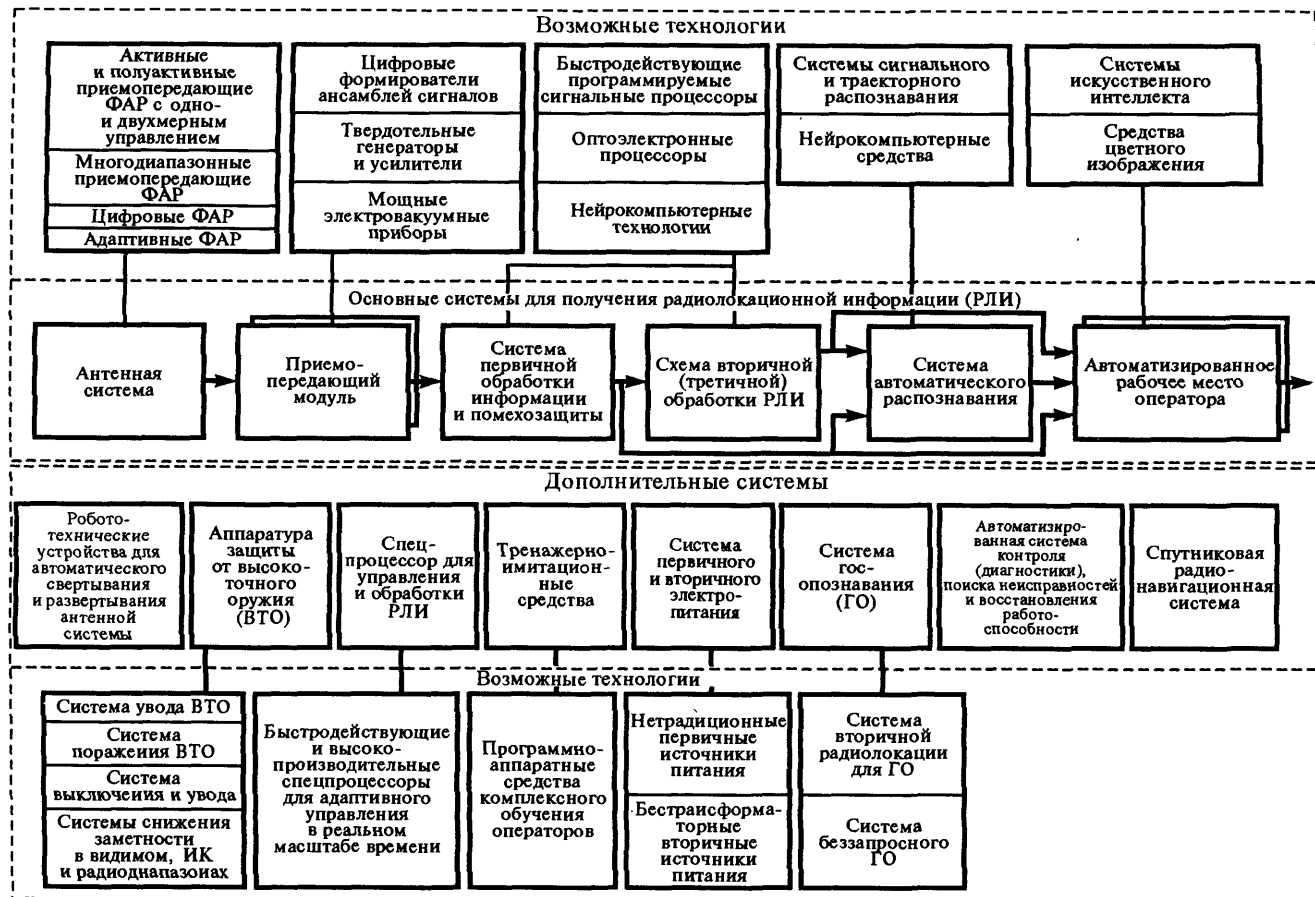


Рис. 7.4. Обобщенная структурная схема радиолокационной системы

Радиопередающее устройство (радиопередатчик) генерирует импульсы или непрерывные высокочастотные (сверхвысокочастотные) колебания с заданными параметрами амплитудной и угловой (частотной или фазовой) модуляции.

Основными задачами радиопередающего устройства являются генерирование колебаний, управление параметрами колебаний (модуляция) и усиление мощности. Соответственно, функционально необходимыми узлами являются: автогенератор (возбудитель), модулятор и усилитель мощности. В состав радиопередающего устройства входят также источники вторичного электропитания, умножители (делители) частоты, системы автоподстройки частоты и параметров модуляции сигналов и др. В зависимости от тактико-технических требований к РЛС радиопередающие устройства делят на однокаскадные и многокаскадные. В однокаскадных передатчиках генерирование колебаний обеспечивается автогенератором необходимой мощности, в многокаскадных — колебания высокостабильного маломощного автогенератора усиливаются до необходимого уровня мощности. По сравнению с однокаскадными многокаскадные радиопередающие устройства обеспечивают более высокие мощность и стабильность параметров зондирующих сигналов, возможность генерирования сложных сигналов.

Для генерирования и усиления колебаний используются электровакуумные и твердотельные электронные приборы. К электровакуумным приборам относятся электронные лампы (триоды и тетроды), магнетроны, клистроны, лампы бегущей волны (для усиления колебаний), лампы обратной волны, гибридные (твистронные) и другие приборы. К твердотельным относятся: транзисторные (на биполярных транзисторах, на полевых транзисторах), диодные (лавинно-пролетные диоды, диоды Ганна, варакторы) и другие полупроводниковые приборы. Все более широкое применение в РЛС находят цифровые формирователи сигналов (ансамблей сигналов) [72, 82, 84].

Радиоприемные устройства (радиоприемники) совместно с системой первичной обработки информации и помехозащиты обеспечивают усиление, преобразование и выделение (оптимальную обработку) принятых антенной сигналов на фоне внутренних шумов и внешних помех.

Первичная обработка радиолокационной информации включает операции обнаружения и измерения (оценки) параметров сигналов, полученных за период (цикл) обзора РЛС. Совокупность оценок параметров сигналов и их показателей качества образуют радиолокационную отметку от цели. Задачи обнаружения сигналов на фоне помех рассматривались в гл. 3 и будут дополнены в § 7.6 применительно к различным видам радиолокационных сигналов. Задачи измерения (оценки) параметров сигналов рассмотрены в гл. 5.

В зависимости от способа выделения принимаемых сигналов на фоне помех используют фильтровые, корреляционные, корреляционно-фильтро-

вые, взаимокорреляционные (в многопозиционных РЛС), автокорреляционные (в пассивных РЛС) радиоприемники.

Система вторичной (траекторной) обработки радиолокационной информации. Вторичная обработка проводится по совокупности радиолокационных отметок, полученных за несколько циклов обзора от одной совмещенной РЛС или нескольких приемных позиций многопозиционной РЛС, и обеспечивает формирование траекторной информации. Анализ алгоритмов вторичной (траекторной) обработки проведен в гл. 6.

Третичная обработка состоит в объединении и отождествлении информации, полученной отдельными РЛС, входящих в радиолокационную систему, или информации, полученной отдельными радиолокационными системами.

Следует отметить, что распределение операций первичной, вторичной и третичной обработки между подсистемами и средствами радиолокационной системы, а также области применения цифровых методов и устройств не являются окончательно установившимися. Вторичная и третичная обработки уже осуществляются с помощью электронных вычислительных машин. Первичная обработка проводится как цифровыми, так и аналоговыми средствами. В качестве аналоговых устройств применяют приборы на поверхностных акустических волнах, приборы с переносом заряда, оптоэлектронные процессоры и др. [35, 73—75]. При цифровой обработке используют быстродействующие программируемые сигнальные процессоры, нейрокомпьютерные технологии и др. [67, 82, 84].

Система автоматического распознавания (классификации, различения) решает задачи определения, к каким классам и типам относятся соответствующие наблюдаемые объекты. Для распознавания различных целей (аэродинамических, наземных, надводных) можно использовать, в частности, *траекторные* признаки, т. е. параметры траекторий целей, определяемые их тактико-техническими характеристиками, учитывающими соответствующие закономерности движения целей. Наряду с траекторными используют *сигнальные* признаки, учитывающие особенности сигналов, отраженных от целей различных классов. В общем случае для решения задачи распознавания требуется высокая производительность систем вычислительной техники ($10^8 \dots 10^{10}$ операций в секунду и более) [85]. Среди многочисленных предложенных алгоритмов распознавания большое внимание последнее время уделяется нейрокомпьютерным алгоритмам [35, 85].

Автоматизированное рабочее место (АРМ) оператора является выходным устройством РЛС. В современных и перспективных АРМ используются системы искусственного интеллекта, системы цветного отображения радиолокационной информации и др.

7.4.2. Дополнительные системы

Робототехнические устройства для автоматического свертывания и развертывания антенных систем используются, в частности, в РЛС космического базирования и мобильных РЛС военного назначения.

Аппаратура защиты от высокоточного оружия (ВТО) применяется в РЛС военного назначения. Высокоточное оружие, обеспечивающее огневое поражение РЛС, можно разделить на две группы:

- противорадиолокационные ракеты (ПРР), самонаводящиеся на радиоизлучение активных РЛС; такие ракеты применяют пассивные радиолокационные головки самонаведения (ГСН);

- управляемые ракеты класса «воздух—поверхность» и управляемые (корректируемые) авиабомбы, использующие активные, полуактивные радиолокационные (лазерные) и пассивные инфракрасные, телевизионные и другие ГСН; такие средства могут наводиться и на неизлучающий радиолокатор, используя, в частности, соответствующий контраст между РЛС и окружающей поверхностью.

Для защиты от ПРР используют дополнительные передатчики, обеспечивающие смещение точки прицеливания в сторону от РЛС; системы, обеспечивающие своевременное выключение (или ограничение времени выхода в эфир) РЛС при подлете ПРР и при этом дополнительный увод противорадиолокационной ракеты, повышение скрытности излучения РЛС. Для защиты от управляемых ракет добиваются снижения заметности РЛС в видимом, ИК и радиодиапазонах. Общими методами защиты РЛС от высокоточного оружия являются использование существующих или создание специальных сравнительно недорогих средств огнестрельного поражения ВТО (например, зенитных ракетно-пушечных комплексов малой и сверхмалой дальности) и непосредственное повышение защищенности (живучести) позиций РЛС за счет применения специального инженерного оборудования позиций, разнесения на местности отдельных систем РЛС, бронирования ее наиболее важных и уязвимых элементов и др. [35, 90].

Спецпроцессор для управления и обработки радиолокационной информации выполняется в виде распределенного комплекса вычислительных средств. Он предназначен для решения разнообразных задач. Приведем для примера задачи, решаемые вычислительными средствами в РЛС 39Н6Е «Каста-2Е2» (см. также п. 7.1.3):

- автоматическое формирование координатных отметок;
- автоматический захват и сопровождение целей, вычисление параметров движения, обнаружение маневра;
- автоматическое и полуавтоматическое управление режимами работы РЛС;

- выделение и автоматическое сопровождение наземных и малоскоростных целей;
- автоматическое управление каналом определения государственной принадлежности и привязка его информации к сопровождаемым целям;
- формирование и выдача информации на отображение (описание целей, линий государственной границы и коридоры движения воздушных средств, тексты помощи оператору, информация о состоянии и загрузке вычислительных средств и т. д.);
- формирование и выдача информации внешним потребителям [92].

Тренажно-имитационные средства используют различные программно-аппаратные средства для комплексного обучения и тренировок операторов РЛС как в составе боевого расчета РЛС, так и в составе, например, войскового соединения или войсковой части.

Система первичного и вторичного электропитания является неотъемлемой частью любой РЛС. К основным системам первичного питания относятся внешняя электрическая сеть или дизель-электростанция. Могут использоваться также и иные источники питания: химические, ядерные и др. В системах вторичного питания, кроме традиционных источников, использующих трансформаторы, применяются бестрансформаторные источники.

Система опознавания государственной принадлежности (система госопознавания) может быть запросной или беззапросной. Используется в радиолокаторах, работающих прежде всего по воздушным целям. Система госопознавания (иногда ее называют системой вторичной радиолокации) — это самостоятельный активный радиолокатор с активным ответом (запросная система) или пассивный радиолокатор с активным ответом (беззапросная система). Для эффективной работы беззапросных систем требуются высокостабильные генераторы как в РЛС, так и на цели, или использование системы единого времени [35, 43].

Автоматизированная система контроля (диагностики) поиска неисправностей и восстановления работоспособности. Современные РЛС оснащаются встроенными автоматизированными (автоматическими) системами функционального контроля (диагностики), использующими методы сигнатурного анализа и допускового контроля для сквозных и локальных проверок всей аппаратуры РЛС. Современные системы обеспечивают в большинстве случаев поиск неисправного элемента с достоверностью не менее 0,99 и дают оператору информацию о расположении отказавших элементов и месте хранения соответствующих запасных частей. Встроенные средства измерения основных параметров дают оператору информацию о состоянии РЛС без прекращения ее работы, а наличие автоматических систем регулировки и защиты приемного тракта и передатчика обеспечивают стабильность рабочих характеристик РЛС в течение срока службы без трудоемких работ по ее настройке.

Использование автоматизированных систем диагностики, поиска неисправностей и восстановления работоспособности обеспечивает длительную непрерывную работу РЛС без выключения на обслуживание и с сохранением всех основных параметров на 20 суток и более, а время восстановления работоспособности — не более чем 0,3...0,5 ч [92, 93].

Спутниковая навигационная система позволяет осуществлять быструю высокоточную топопривязку (см. также гл. 8). Использование такой аппаратуры повышает мобильность РЛС.

7.5. Формирование отраженного радиолокационного сигнала

7.5.1. Вторичное излучение электромагнитных волн.

Эффективная площадь рассеяния целей

Явление вторичного излучения, лежащее в основе активной радиолокации, свойственно волнам любой природы. Оно возникает всякий раз, когда волна встречает препятствие на пути своего распространения. Падающую на препятствие волну называют *первичной*, отраженную, или рассеянную, — *вторичной*. Препятствие в этом случае является пассивным вторичным излучателем.

Препятствием для радиоволн служит любая неоднородность электрических параметров среды (абсолютной диэлектрической проницаемости ϵ_a , абсолютной магнитной проницаемости μ_a , проводимости σ). В радиолокации интерес представляют как объекты с большой проводимостью ($\sigma \rightarrow \infty$), так и объекты с малой проводимостью (диэлектрики): гидрометеоры, неоднородности тропосферы и др. Под действием электрического поля волны на облучаемой поверхности, например проводящей, возникают колебания электрических зарядов. Наведенные при этом токи проводимости являются источником излучения вторичных электромагнитных волн. В диэлектрике таким же источником являются токи смещения.

Характер вторичного излучения зависит от многих факторов, основными из которых являются электрические свойства, геометрическая форма, движение и взаимное перемещение элементов отражающего объекта, соотношение размера объекта и длины облучающей его волны, соотношение размеров объекта и разрешаемого объема пространства (объект считается сосредоточенным, если он попадает в пределы одного разрешаемого объема, и объемно распределенным, если занимает несколько разрешаемых объемов), закон модуляции и поляризация облучающей электромагнитной волны [31, 43, 84, 94—96].

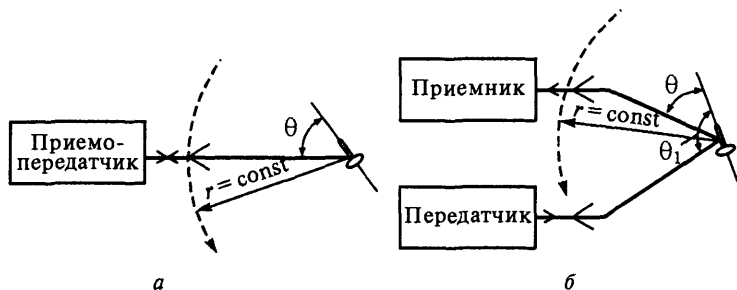


Рис. 7.5. Экспериментальное определение диаграмм:

a — обратного вторичного излучения; *б* — вторичного излучения

Наиболее важной энергетической характеристикой вторичного излучения сосредоточенного объекта (цели) в точке приема, не зависящей от интенсивности первичной волны, является его *эффективная площадь* $\sigma_{\text{ц}}$. На рис. 7.5, *a*, *б* поясняются условия возбуждения и приема: *a* — обратного вторичного излучения в совмещенном радиолокаторе; *б* — вторичного излучения в разнесенном локаторе. Пусть в окрестности цели создается плотность потока мощности первичной волны $\Pi_{\text{ц}}$, Дж/с · м² = Вт/м². На расстоянии r от цели в точке приема создается плотность потока мощности вторичной волны $\Pi_{\text{пр}}$. Реальный вторичный излучатель можно заменить эквивалентным, создающим равномерную плотность потока $\Pi_{\text{пр}}$ на сфере радиуса r с площадью поверхности, равной $4\pi r^2$, именно такую, которая создается в направлении приемника реальной целью. Отношение мощности $P = 4\pi r^2 \Pi_{\text{пр}}$, рассеиваемой эквивалентной целью без потерь, к плотности потока мощности у цели имеет размерность площади и служит искомой характеристикой

$$\sigma_{\text{ц}} = \frac{P}{\Pi_{\text{ц}}} = \frac{4\pi r^2 \Pi_{\text{пр}}}{\Pi_{\text{ц}}},$$

которую называют *эффективной площадью* (*эффективной поверхностью*) *вторичного излучения цели* (ЭПВИ, ЭПЦ, ЭП). Широкое распространение получило также название *эффективная площадь рассеяния* (ЭПР).

Таким образом, эффективная площадь рассеяния $\sigma_{\text{ц}}$ — это площадь поверхности такого эквивалентного вторичного отражателя, который равномерно рассеивает всю падающую на него энергию и создает в точке приема такую же плотность потока энергии, что и реальная цель.

Выражая $\Pi_{\text{пр}}$ и $\Pi_{\text{ц}}$ через квадраты амплитуд напряженностей электрического и магнитного полей, формулу для определения ЭПР цели представляют в виде

$$\sigma_{\text{ц}} = \frac{4\pi r^2 |E_{\text{пр}}|^2}{|E_{\text{ц}}|^2} = \frac{4\pi r^2 |H_{\text{пр}}|^2}{|H_{\text{ц}}|^2}. \quad (7.1)$$

Знаки модулей позволяют вводить в расчет комплексные амплитуды колебаний.

Зависимость ЭПР от направления на приемопередатчик в равноудаленных от цели точках называют *характеристикой (диаграммой) обратного вторичного излучения* $\sigma_{\text{ц}}(\theta)$. Для ее измерения приемопередатчик перемещают вокруг цели (см. рис. 7.5, а), либо цель поворачивают относительно направления на приемопередатчик.

Зависимость ЭПР от направления на приемник в равноудаленных от цели точках при фиксированном направлении облучения цели называют *характеристикой (диаграммой) вторичного излучения* $\sigma_{\text{ц}}(\theta, \theta_1)$ при $\theta_1 = \text{const}$. Для ее измерения приемник перемещают вокруг модели цели (см. рис. 7.5, б), сохраняя положение передатчика. Эффективная площадь вторичного излучения для фиксированной длины волны является функцией $\sigma_{\text{ц}} = \sigma_{\text{ц}}(\varepsilon, \beta)$ двух углов при однопозиционной локации и функцией $\sigma_{\text{ц}} = \sigma_{\text{ц}}(\varepsilon, \beta, \varepsilon_1, \beta_1)$ четырех углов при разнесенной локации.

7.5.2. Эффективная площадь рассеяния при различных соотношениях размеров цели и длины волны

Поля вторичного излучения, позволяющие определить ЭПР объектов, находят точными и приближенными расчетными методами, в том числе методом математического моделирования, наряду с которым широко используют физическое моделирование и натурный эксперимент. Полагают, что вторичные излучатели расположены в свободном пространстве с диэлектрической и магнитной проницаемостями соответственно $\varepsilon = 1$ и $\mu = 1$. На практике ЭПР может быть рассчитана различными методами. Точные решения обеспечиваются на основе использования уравнений Максвелла и решения соответствующих волновых уравнений при заданных граничных условиях. Решения таких частных модельных задач найдены для тел сравнительно простой формы: шара, полуплоскости, клина, плоских пластин, цилиндров, усеченных конусов, сфероидов, прямых и искривленных ребер. Для решения модельных задач также используют приближенные методы, к основным из которых относятся следующие методы: геометрической оптики, физической оптики, геометрической теории дифракции, физической теории дифракции; смешанные методы [35, 84, 85, 94—97].

Метод геометрической оптики основывается на предположении, что энергия рассеянного поля распространяется вдоль узких трубок лучей по законам оптики, когда длина падающей волны стремится к нулю.

При использовании *метода физической оптики* пренебрегают вкладом в рассеянное поле затененных участков объекта при условии, что длина волны много меньше его размеров.

Метод геометрической теории дифракции развивает метод геометрической оптики и основывается на предположении, что основной вклад во вторичное поле вносят токи вблизи геометрических разрывов объекта.

При использовании *метода физической теории дифракции* считается, что основной вклад в рассеянное поле вносят освещенные участки объекта и его острые кромки.

Задачи расчета полей вторичного излучения вполне естественно разбиваются на три группы, связанные с относительными размерами целей:

- цели малых по сравнению с длиной волны размеров ($l_{ц}/\lambda \ll 1$);
- цели, размеры которых соизмеримы с длиной волны ($l_{ц}/\lambda \approx 1$);
- цели, размеры которых больше длины волны ($l_{ц}/\lambda \gg 1$).

Зависимость ЭПР от размеров цели при фиксированной длине волны рассмотрим на примере объекта простой формы — проводящего шара радиуса ρ . На рис. 7.6, а представлена зависимость ЭПР шара $\sigma_{ш}$, нормированная к его видимой площади $\pi\rho^2$, от отношения радиуса шара к

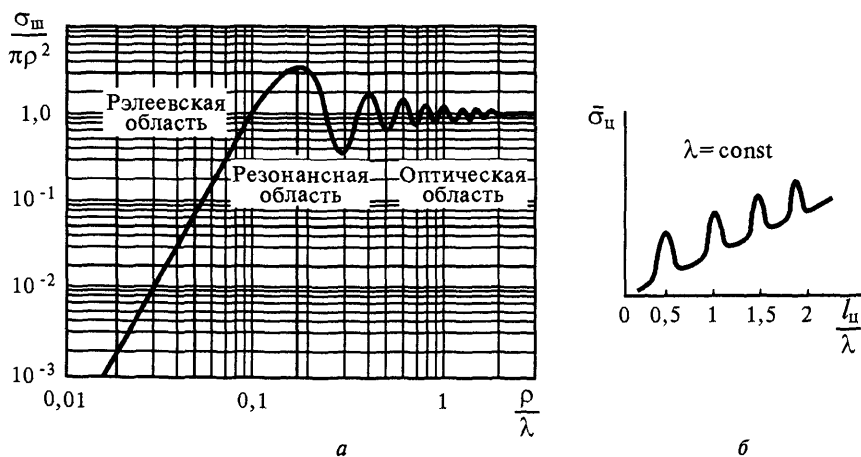


Рис. 7.6. Зависимость ЭПР шара, нормированной к площади его поперечного сечения $\pi\rho^2$, от величины отношения радиуса шара к длине волны ρ/λ (а); зависимость ЭПР вибратора от величины отношения длины вибратора к длине волны $l_{ц}/\lambda$ (б)

длине волны [35, 84]. На графике можно выделить три характерных области.

Первая область ($\rho/\lambda \leq 0,1$) называется *областью рэлеевского рассеяния* (низкочастотной, дифракционной). В рэлеевской области функция и аргумент связаны зависимостью четвертой степени:

$$\frac{\sigma_{\text{ш}}}{\rho^2} \equiv \left(\frac{\rho}{\lambda} \right)^4. \quad (7.2)$$

Такая зависимость характерна для любого объекта, наибольший размер которого значительно меньше длины волны.

Вторая область — *резонансная* ($\pi\rho/\lambda \sim 1$). Колебательный характер зависимости значений ЭПР от отношения ρ/λ обусловлен интерференцией зеркально отраженной от поверхности шара волны и волн рассеиваемой «невидимой» (обратной) частью шара. Максимум ЭПР наблюдается в случае, когда шар как бы эквивалентен полуволновому вибратору и вдоль его полуокружности укладывается полуволна тока ($\pi\rho/\lambda = 0,5$).

Третья область ($\rho/\lambda \gg 1$) называется *оптической (высокочастотной, зеркальной)* областью. Ее характерная особенность состоит в том, что значение ЭПР постепенно приближается к площади поперечного сечения (видимой площади) цели.

Рассмотрим особенности отражения для других объектов радиолокации.

Для объектов первой группы ($l_{\text{ц}}/\lambda \ll 1$) независимо от конфигурации и электрических свойств материала остается в силе зависимость (7.2):

$$\sigma_{\text{ц}} \equiv \frac{l_{\text{ц}}^6}{\lambda^4}, \quad \text{или} \quad \frac{\sigma_{\text{ц}}}{S_{\text{г}}} \equiv \left(\frac{l_{\text{ц}}}{\lambda} \right)^4, \quad (7.3)$$

где $S_{\text{г}} = l_{\text{ц}}^2$ — квадрат линейного размера объекта (видимая геометрическая площадь объекта).

Примерами объектов первой группы являются гидрометеоры, аэрозоли для РЛС дециметрового диапазона, птицы — для РЛС метрового диапазона и т. д.

Из формулы (7.3) следует, что мешающее влияние метеобразований наиболее существенно в коротковолновой части сантиметрового диапазона, а также в миллиметровом диапазоне. С другой стороны, РЛС этих диапазонов можно использовать для разведки физических условий на трассе распространения радиоволн. Как уже отмечалось выше (см. § 7.3), низкая эффективность вторичного излучения при $\lambda \gg l_{\text{ц}}$ ограничивает длины волн (частоты) колебаний, используемых в радиолокации.

Для объектов второй группы ($l_{\text{ц}}/\lambda \approx 1$) рассеяние является резонансным. Так, полуволновый вибратор $l_{\text{ц}} = \lambda/2$ (используемый для создания пассивных помех РЛС) в направлении, перпендикулярном его оси, имеет ЭПР $\sigma_{\text{max}} = 0,86\lambda^2$, существенно большую его «видимой» площади. При изменении λ или $l_{\text{ц}}$ условия резонанса нарушаются и частотная зависимость характеризуется резонансной кривой, которая тем острее, чем тоньше вибратор.

В пространстве диполи ориентируются произвольно относительно вектора облучающей волны. В предположении равновероятности ориентации усредненное значение ЭПР диполя определяется формулой

$$\sigma_{\text{ц.ср}} = \bar{\sigma}_{\text{ц}} = 0,17\lambda^2. \quad (7.4)$$

Зависимость ЭПР диполя от отношения $l_{\text{ц}}/\lambda$ (см. рис. 7.6, б) носит осциллирующий характер, присущий и другим проводящим телам (шар, эллипсоид) при $l_{\text{ц}} \approx \lambda$. К рассматриваемой группе объектов радиолокации можно отнести, например, боевые блоки баллистических ракет, крылатые ракеты, беспилотные летательные аппараты при условии, что они облучаются радиолокаторами метрового диапазона.

Для объектов третьей группы ($l_{\text{ц}} \gg \lambda$), в которую входит большинство объектов радиолокации, особенностью является интерференционный характер поля вторичного излучения, зависящий от геометрической формы и размеров объекта, длины волны, направления распространения зондирующей и отраженной электромагнитных волн.

Рассмотрим методику расчета ЭПР гладкого идеально проводящего выпуклого тела (рис. 7.7) в приближении геометрической оптики ($l_{\text{ц}} \gg \lambda$,

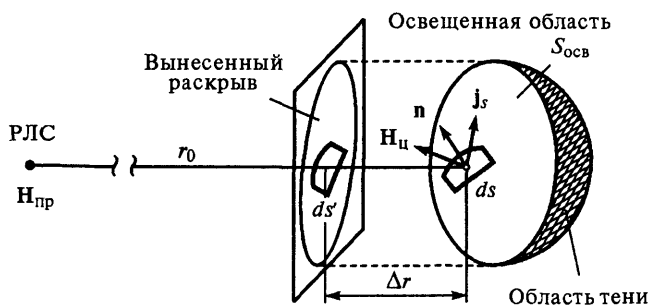


Рис. 7.7. Определение ЭПР гладкого идеально проводящего выпуклого тела в приближении геометрической оптики ($l_{\text{ц}} \gg \lambda$, радиусы кривизны $\rho_{1,2} \gg \lambda$)

радиусы кривизны $\rho_{1,2} \gg \lambda$). Будем предполагать, что вклад в отраженное поле дает только освещенная область цели с резкой границей (областью полутени, наличие которой обусловлено дифракцией, пренебрегаем). Зондирующая электромагнитная волна наведет на освещенной области поверхности токи с плотностью $\mathbf{j}_s = \mathbf{n} \times \mathbf{H}_u$, где \mathbf{n} — единичный вектор нормали. Эти токи создают поле вторичного излучения. Напряженность $\mathbf{H}_{пр}$ поля в точке приема рассчитываем в приближении физической оптики: ток на криволинейном элементе поверхности заменим током на касательном плоском элементе. Для упрощения воспользуемся методом вынесенного раскрыва: спроецируем все элементы на плоскость, перпендикулярную направлению облучения, и затем, пользуясь формулами для составляющих поля элементарного вибратора, найдем

$$\frac{H_{пр}}{H_u} = \frac{1}{\lambda r_0} \left| \int_{S_{осв}} \exp\left(-j \frac{2\pi}{\lambda} 2\Delta r\right) ds' \right|, \quad (7.5)$$

где r_0 — расстояние от РЛС до вынесенного раскрыва; Δr — расстояние от элемента ds освещенной поверхности до его проекции ds' .

Подставив отношение (7.5) в (7.1), получим

$$\sigma_u = \frac{4\pi}{\lambda^2} \left| \int_{S_{осв}} \exp\left(-j \frac{2\pi}{\lambda} 2\Delta r\right) ds' \right|^2. \quad (7.6)$$

Применим формулу (7.6) для расчета поля вторичного излучения идеально проводящих параболоида и прямоугольной пластины, расположенных так, как показано на рис. 7.8, 7.9 соответственно. После необходимых вычислений получим ЭПР параболоида [31, 35, 43, 73, 74]

$$\sigma_{\Pi} = \pi \rho_1 \rho_2, \quad (7.7)$$

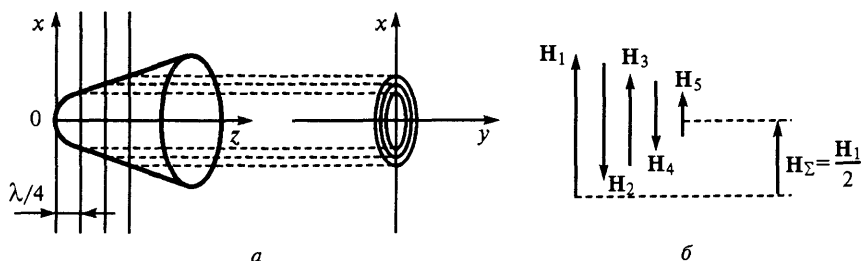


Рис. 7.8. Зоны Френеля при вторичном излучении параболоида (а) и соответствующие им векторы поля (б)

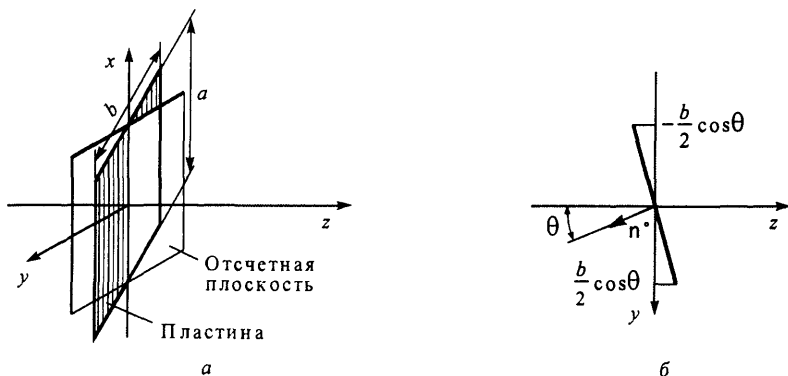


Рис. 7.9. Изображения пластины и отсчетной плоскости $z = 0$:

a — в аксонометрии; b — на плоскости y, z

где ρ_1, ρ_2 — главные радиусы кривизны в точке касания поверхности параболоида с фронтом падающей электромагнитной волны (ЭМВ). Поясним физический смысл формулы. Сечения параболоида плоскостями $z = n\lambda/4$, $n = 1, 2, 3, \dots$, разбивают его поверхность на зоны Френеля. Каждая зона по отношению к другой является противофазным источником обратного вторичного излучения. Амплитуды полей этих источников согласно (7.5) пропорциональны площадям проекций зон (рис. 7.8, a). При большом числе зон напряженность \mathbf{H}_Σ суммарного поля стремится к векторной величине, равной половине вектора \mathbf{H}_1 поля, создаваемого первой зоной Френеля (рис. 7.8, b). В итоге вторичное излучение определяется первой зоной Френеля, т. е. носит локальный характер.

Таким образом, отражающие свойства поверхности второго порядка определяются как бы «светящимся» элементом поверхности, который называют *блестящим элементом* (*блестящей точкой*).

Эффективная площадь рассеяния прямоугольной пластины (с размерами a, b), рассчитанная по формуле (7.6), определяется выражением

$$\sigma_{\text{пл}}(\theta) = \frac{4\pi}{\lambda^2} (ab \cos \theta)^2 \left| \frac{\sin\left(\frac{2\pi}{\lambda} b \sin \theta\right)}{\frac{2\pi}{\lambda} b \sin \theta} \right|. \quad (7.8)$$

При $\theta = 0$ величина ЭПР имеет максимальное значение

$$\sigma_{\text{пл max}} = \frac{4\pi}{\lambda^2} (ab)^2, \quad (7.9)$$

существенно большее ее геометрической площади, поскольку $(ab)/\lambda^2 \gg 1$. Формула (7.9) справедлива для пластин произвольной формы.

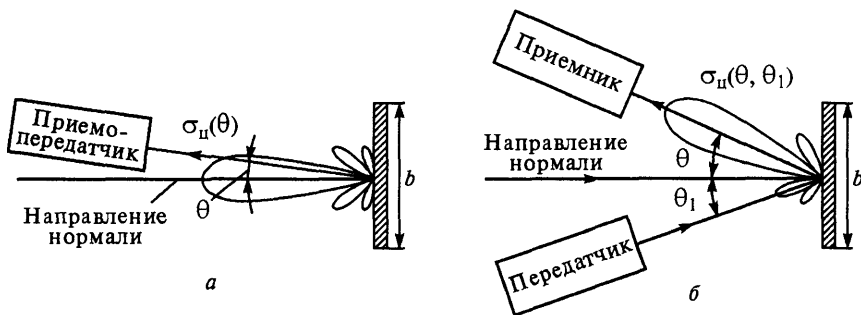


Рис. 7.10. Вид диаграммы обратного вторичного излучения $\sigma_{\text{ц}}(\theta)$ (а) и диаграммы вторичного излучения $\sigma_{\text{ц}}(\theta, \theta_1)$ при $\theta_1 = \text{const}$ (б) в полярных координатах для прямоугольной пластины

Вид характеристик обратного вторичного излучения и вторичного излучения, приведенных на рис. 7.10, а, б соответственно, свидетельствует об интерференционном характере поля вторичного излучения. Ширина главного лепестка диаграммы обратного вторичного излучения равна λ/b , боковых лепестков — $\lambda/(2b)$.

7.5.3. Характер вторичного излучения и эффективная площадь рассеяния радиолокационных целей

Радиолокационные цели обладают, как правило, достаточно сложной конфигурацией. Поверхность их характеризуется наличием блестящих точек в виде выпуклых участков поверхностей двойной кривизны и резонансных элементов, а также шероховатых диффузно рассеивающих участков. Поля, отраженные перечисленными элементами с различными амплитудами и фазами, интерферируют. Характеристики вторичного излучения и обратного вторичного излучения оказываются многолепестковыми. Количество и ширина лепестков зависят от отношения $l_{\text{ц}}/\lambda$. Так, средняя ширина лепестка характеристики обратного вторичного излучения определяется формулой

$$\theta_{\text{л}} = \lambda/(2l_{\text{ц}}),$$

где $l_{\text{ц}}$ — эквивалентный размер цели, не превышающий геометрического размера ее проекции на плоскость, нормальную направлению облучения.

Примерные значения средних ЭПР $\sigma_{\text{ц,ср}}$ для различных целей, имеющих размеры $l_{\text{ц}} \gg \lambda$, приведены в табл. 7.1 [35].

Таблица 7.1

**Примерные значения средних ЭПР различных целей,
имеющих размеры $l_{ц} \gg \lambda$**

Радиолокационная цель	$\sigma_{ц,ср}, М^2$
Корабль водоизмещением 10^4 т	$2 \cdot 10^4$
Средний корабль водоизмещением $3 \cdot 10^4 \dots 10^3$ т	$10^4 \dots 3 \cdot 10^3$
Малые корабли водоизмещением 200...60 т	250...50
Подводные лодки в надводном положении	140...35
Транспортный самолет, дальний бомбардировщик	100...20
Средний бомбардировщик	30...5
Бронетранспортер	30...5
Тягач тяжелый	20...15
Тяжелый истребитель	10...5
Танк	9...6
Автомобиль	8...6
Легкий истребитель	5...2
Человек ($\lambda = 3$ см)	1,2...0,5
Рубка подводной лодки	1
Крылатая ракета	$1 \dots 10^{-2}$
Стая птиц, попадающих в разрешаемый объем РЛС	$1 \dots 10^{-2}$
Головка баллистической ракеты	$1 \dots 10^{-3}$
Голубь ($\lambda = 3$ см)	$10^{-2} \dots 10^{-3}$
Пчела рабочая ($\lambda \approx 10 \dots 0,8$ см)	$6 \cdot 10^3 \dots 10^{-8}$

Случайный характер интерференции волн, порожденных вторичным излучением реальных целей, вызывает амплитудные и фазовые флуктуации принимаемого РЛС сигнала, которые называют *амплитудным* и *фазовым шумом* соответственно. Эти шумы обусловлены случайным характером (ориентацией) диаграммы обратного вторичного излучения. Пусть цель облучается гармонической волной. Тогда в процессе движения цели меняется уровень лепестков этой диаграммы в направлении на РЛС (рис. 7.11, а). Сигнал на входе приемника претерпевает случайные изменения как по амплитуде, так и по фазе. Таким образом, амплитудный и фазовый шумы цели приводят к амплитудной и фазовой модуляции отраженного сигнала (рис. 7.11, б).

При отражении от сложной цели возникает также *угловой шум*, обусловленный тем, что положение фазового центра системы блестящих точек зависит от соотношения амплитуд и фаз отраженных колебаний. Фронт отраженной волны оказывается отличным от сферического, направление нормали к нему может отклоняться не только от направления на какую-либо

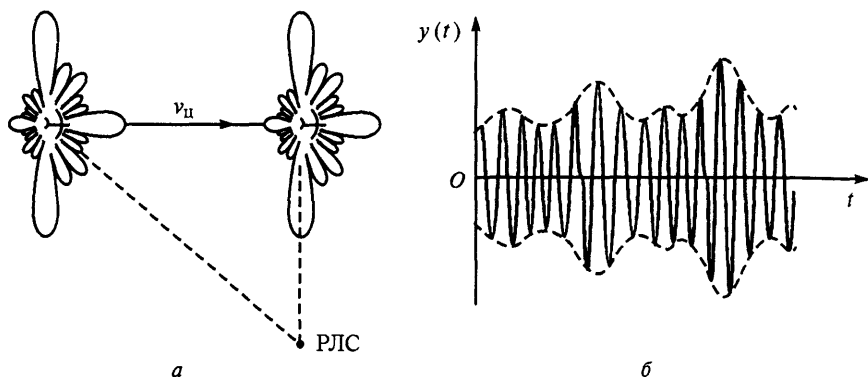


Рис. 7.11. Влияние движения цели (а) на амплитудную и фазовую модуляции отраженного гармонического сигнала (б)

блестящую точку, но и выходить за пределы углового размера цели. На рис. 7.12 представлена типичная картина углового шума для двухмоторного самолета. По оси ординат отложено линейное смещение кажущегося положения цели относительно истинного в плоскости, перпендикулярной направлению РЛС—цель. Для реальных самолетов среднеквадратическое отклонение $\sigma_{\delta L}$ кажущегося положения цели от истинного δL находится в интервале $0,15L \leq \sigma_{\delta L} \leq 0,3L$, где L — габаритный размер цели [84]. Угловой шум проявляется сильнее для целей с малым числом блестящих точек.

Сложной цели свойствен также *поляризационный* шум, обусловленный тем, что цель может преобразовывать поляризацию облучающей электромагнитной волны. В отраженной волне появляются составляющие на ортогональной поляризации. Такие цели являются поляризационно-избирательными.

В зависимости от ракурса цели вклад различных компонентов конструкции летательного аппарата неодинаков. На рис. 7.13 показана зависи-

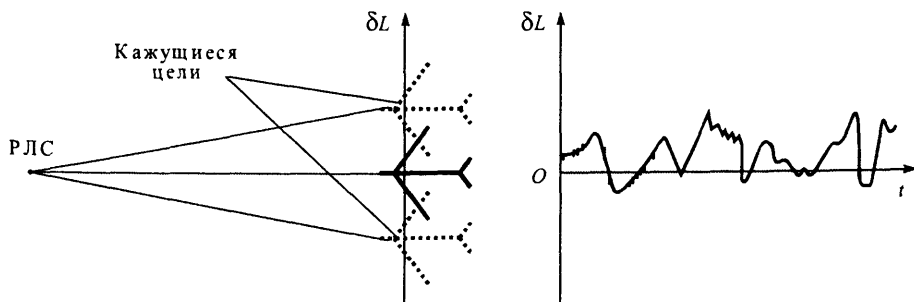


Рис. 7.12. Влияние углового шума на определение угловых координат движущейся цели

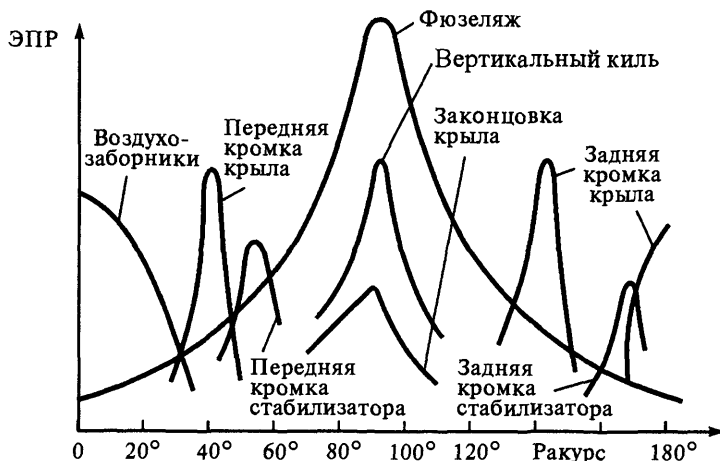


Рис. 7.13. Влияние различных элементов конструкции летательного аппарата на величину ЭПР в зависимости от ракурса

мость ЭПР основных элементов планера истребителя от его ракурса [84]. Под ракурсом цели понимают угол между продольной осью цели и направлением на РЛС.

Уменьшение ЭПР цели является одной из основных задач развития и модернизации средств воздушного нападения. Решение этой задачи ведется по многим направлениям, к основным из которых можно отнести следующие [35, 84, 85].

Совершенствование формы летательных аппаратов. Техническая реализация этого направления проявляется в устранении острых кромок, разрывов поверхностей (например, в самолетах F-111, ATF-1, F-16), в уменьшении площади вертикальных килей за счет перехода к двухкилевой схеме, отклонение килей внутрь к оси фюзеляжа (в самолетах SR-71, F-19). Эффективными методами уменьшения ЭПР являются применение аэродинамической схемы «летающее крыло» (в самолетах B-2, АТВ и т. п.); интеграция системы «планер—воздухозаборник—двигатель» и «планер—вооружение» путем размещения воздухозаборников и сопел на верхней поверхности самолета (как это реализовано в SR-71), применение конформных систем подвески вооружения (например, в самолете B-2B). Кроме того, придание фюзеляжу и передним кромкам аэродинамических поверхностей формы, уменьшающей зеркальное отражение в сторону облучающей РЛС.

Применение радиопрозрачных и радиопоглощающих материалов и устройств. Технически это достигается использованием композиционных материалов для создания элементов конструкции летательных аппаратов. Так, термопластики могут составлять до 60 % массы самолета ATF. Широко

применяются радиопоглощающие покрытия (ферромагнитные материалы и полимеры), для уменьшения отражения от входного канала двигателя используются аттенуаторы. В самолетах SR-11, U-2, TR-1, F-19 применяются токопроводящие покрытия для устранения разрывов между материалами с различными электрическими свойствами.

Снижение заметности бортовых антенных систем. Для реализации этого направления используют следующие пути: ориентируют антенны таким образом, чтобы исключить зеркальное отражение в сторону РЛС (так, в бомбардировщике B-1В антенна многофункциональной РЛС отклонена вниз); уменьшают число антенн; используют обтекатели с управляемыми характеристиками радиопрозрачности.

7.5.4. Формирование отраженного радиолокационного сигнала для различных моделей целей

Облучение цели и формирование отраженного сигнала – очень важный этап в работе радиолокационного тракта. Именно на этом этапе у сигнала появляются информативные параметры $\alpha = (\alpha_1, \alpha_2, \alpha_3, \dots)$, связанные с координатами и параметрами движения цели.

Рассмотрим формирование отраженного сигнала $x(t, \alpha)$ для постепенно усложняющейся модели цели (неподвижная точечная, движущаяся равномерно точечная, движущаяся многоточечная).

Сигнал $x(t, \alpha_1)$, отраженный от неподвижной точечной цели, находящейся на некотором расстоянии от РЛС, отличается от зондирующего сигнала $x_0(t)$ существенно меньшей амплитудой и наличием запаздывания:

$$x(t, \alpha_1) = Bx_0(t - t_3), \quad (7.10)$$

где $B \ll 1$ — некоторый коэффициент, а $t_3 = \alpha_1 = 2r/c$ — время запаздывания.

В случае движения точечной цели с постоянной радиальной скоростью к описанным по формуле (7.10) преобразованиям добавляется изменение масштаба времени, обусловленное эффектом Доплера, так что

$$x(t) = Bx_0(K_t(t - t_3)), \quad (7.11)$$

где

$$K_t = (1 \mp v_r/c)/(1 \pm v_r/c) \approx 1 \mp 2\frac{v_r}{\lambda} \text{ при } v_r \ll c. \quad (7.12)$$

Изменение масштаба времени проявляется только на достаточно больших промежутках времени. На малых отрезках времени заметно лишь

изменение несущей частоты. Несущая частота отраженного сигнала изменяется на величину доплеровской поправки частоты

$$F_D = \alpha_2 = 2f_0 \frac{v_r}{c} = \frac{2v_r}{\lambda}, \quad (7.13)$$

знак которой зависит от того, удаляется ($F_D < 0$) или приближается ($F_D > 0$) цель. Наличие доплеровской поправки частоты обуславливает также изменение фазы за период посылки импульсов.

Таким образом, при облучении изотропно отражающей движущейся точечной цели принимаемый сигнал (без учета влияния среды) является нефлуктуирующим и содержит временные (точнее, времячастотные) информативные параметры $\alpha = (\alpha_1, \alpha_2)$.

Рассмотрим особенности поля вторичного излучения цели, образованного двумя изотропно отражающими точками с ЭПР σ_1, σ_2 . Геометрия такой цели иллюстрируется на рис. 7.14, а, где θ — угол между нормалью к линии базы излучателей 1, 2 и направлением на РЛС с совмещенными передающей и приемной антеннами.

Пусть передатчик излучает гармоническое колебание частоты f_0 вида

$$x(t) = X(t) \cos(2\pi f_0 t - \psi_0) = \operatorname{Re}(X(t)e^{-j\psi_0}e^{j2\pi f_0 t}), \quad (7.14)$$

где $X(t)$ — огибающая колебания $x(t)$, ψ_0 — начальная фаза, значение которой не влияет на полученные ниже результаты.

Колебания на входе приемной антенны находят как результат интерференции полей E_1 и E_2 , обусловленных изотропно отражающими точками 1 и 2. Результирующее поле в точке приема является векторной суммой:

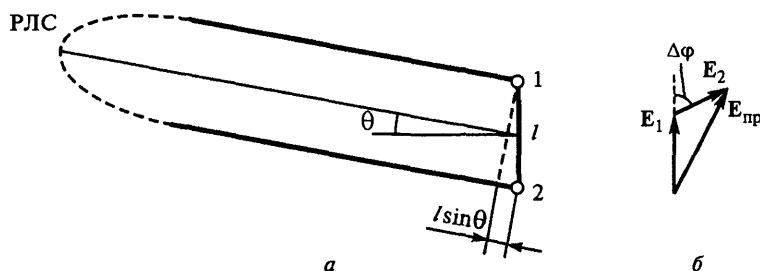


Рис. 7.14. Определение эффективной площади рассеяния группового вторичного излучателя:

а — геометрия системы двух изотропно отражающих точечных целей, *б* — определение значения результирующего вектора $E_{пр}$ в точке приема

$$E_{\text{пр}} = E_1 + E_2. \quad (7.15)$$

Разность фаз $\Delta\varphi$ полей E_1 и E_2 в точке приема зависит от угла θ и пропорциональна разности хода волн от РЛС к цели и обратно:

$$\Delta\varphi = 2kl \sin \theta, \quad k = 2\pi/\lambda = 2\pi f_0/c. \quad (7.16)$$

Как видно из рис. 7.14, б, по теореме косинусов имеем

$$E_{\text{пр}}^2 = E_1^2 + E_2^2 + 2E_1E_2 \cos \Delta\varphi. \quad (7.17)$$

В соответствии с введенным в п. 7.5.1 определением ЭПР получим

$$\sigma_{\text{ц}} = E_{\text{пр}}^2, \quad \sigma_1 = E_1^2, \quad \sigma_2 = E_2^2, \quad (7.18)$$

и эффективную площадь рассеяния двухточечной цели можно записать в следующем виде:

$$\sigma_{\text{ц}} = \sigma_1 + \sigma_2 + 2\sqrt{\sigma_1\sigma_2} \cos \Delta\varphi. \quad (7.19)$$

Как следует из соотношения (7.19), эффективная площадь рассеяния двухточечной цели зависит от ЭПР каждого вторичного излучателя и сдвига фаз возбуждаемых этими излучателями вторичных волн в точке приема. Если излучения суммируются в фазе $\Delta\varphi = 0$, то $\sigma_{\text{ц}}$ принимает максимальное значение:

$$\sigma_{\text{ц max}} = (\sqrt{\sigma_1} + \sqrt{\sigma_2})^2. \quad (7.20)$$

В случае противофазного излучения получаем минимальное значение ЭПР цели:

$$\sigma_{\text{ц min}} = (\sqrt{\sigma_1} - \sqrt{\sigma_2})^2. \quad (7.21)$$

На рис. 7.15, а, б показаны диаграммы обратного вторичного излучения группового излучателя $\sigma_{\text{ц}}(\theta)$ для $\sigma_1 = \sigma_2$ и различных l/λ . Видно, что

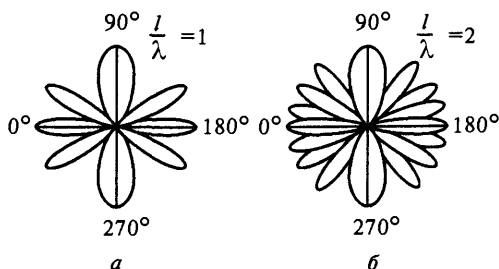


Рис. 7.15. Диаграммы обратного вторичного излучения двухточечной цели группового излучателя для $\sigma_1 = \sigma_2$ и $l/\lambda = 1$ (а), $l/\lambda = 2$ (б)

при увеличении отношения l/λ ширина лепестков диаграммы обратного вторичного излучения уменьшается, а число их растет. Глубина провалов в диаграмме тем больше, чем ближе σ_1 к σ_2 .

Для сложной многоточечной цели эффективная площадь рассеяния находится из выражения

$$\sigma_{\text{ц}} = \sum_{i=1}^n \sigma_i + 2 \sum_{i \neq j} \sqrt{\sigma_i \sigma_j} \cos(\varphi_i - \varphi_j). \quad (7.22)$$

Здесь n — число блестящих точек, φ_i и φ_j — фазы поля, отраженного от соответствующих точечных излучателей.

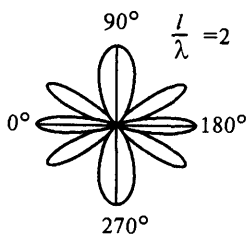
Диаграмма вторичного излучения, полученная при разнесении в пространстве приемной и передающей антенн, также носит многолепестковый характер [31, 35, 43, 73, 74]. Однако количество лепестков в рассматриваемом случае меньше, чем для диаграммы обратного вторичного излучения (рис. 7.16 и 7.15, б соответственно).

Учитывая, что при движении цели относительно РЛС число блестящих точек, амплитуды и фазы отраженных от них полей изменяются случайным образом, приходим к выводу, что результирующее поле ЭПР будет случайными функциями времени. Значит, необходимо статистическое описание ЭПР и отраженного сигнала.

Пусть, как и раньше, цель облучается гармоническим колебанием вида (7.14). В этом случае отраженный от реальной цели сигнал будет случайным образом промодулирован по амплитуде и фазе:

$$x(t) = B(t)X(t) \cos(2\pi(f_0 - F_{\text{д.ср}})t - \varphi(t) + \beta(t)). \quad (7.23)$$

Здесь $X(t)$, $\varphi(t)$ — законы изменения неслучайных компонент амплитуды и фазы отраженного от цели сигнала; $F_{\text{д.ср}}$ — средняя частота Доплера, обусловленная движением центра отражения с радиальной скоростью $v_{\text{р.ср}}$;



$B(t)$, $\beta(t)$ — функции изменения случайных компонент амплитуды и фазы отраженного сигнала, обусловленные амплитудными и фазовыми шумами цели.

Случайный процесс $B(t)$ характеризуется нормированной корреляционной функцией

$$r(\tau) = \frac{\mathbf{M}\{B(t)B(t-\tau)\}}{\mathbf{M}\{B(t)B(t)\}}. \quad (7.24)$$

Рис. 7.16. Диаграмма вторичного излучения группового излучателя для $\sigma_1 = \sigma_2$ и $l/\lambda = 2$

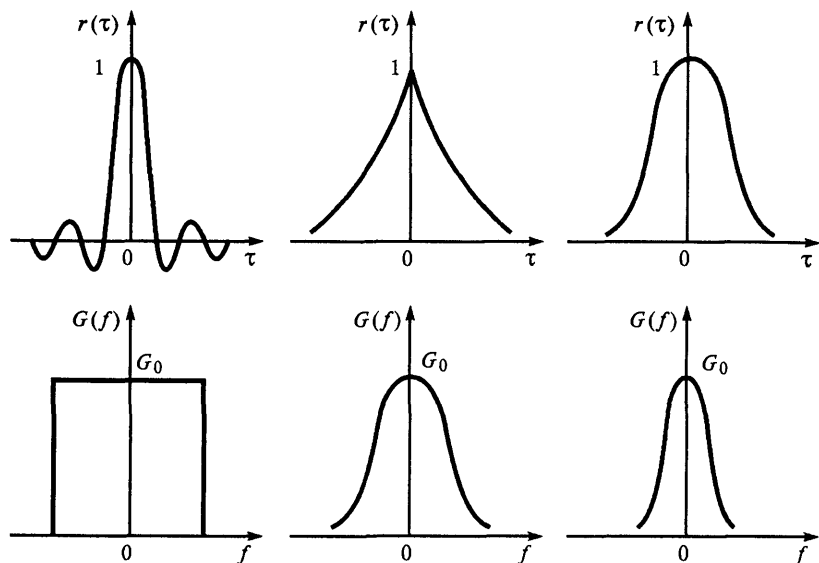


Рис. 7.17. Примеры автокорреляционных функций $r(\tau)$ и соответствующих им энергетических спектров $G(f)$

В частотной области процесс $B(t)$ описывается энергетическим спектром $G(f)$, взаимно связанным с корреляционной функцией $r(\tau)$ преобразованиями Фурье:

$$G(f) = \int_{-\infty}^{\infty} r(\tau) \exp(-j2\pi f\tau) d\tau, \quad (7.25, a)$$

$$r(\tau) = \int_{-\infty}^{\infty} G(f) \exp(j2\pi f\tau) df. \quad (7.25, б)$$

Функции $G(f)$ и $r(\tau)$ могут быть получены в результате статистической обработки случайных процессов $B(t)$, $\beta(t)$, получаемых в физическом эксперименте на реальных объектах или моделях, а также в результате математического моделирования. Возможные виды корреляционных функций $r(\tau)$ и связанных с ними энергетических спектров $G(f)$ представлены на рис. 7.17. Здесь показаны усредненные нормированные корреляционные функции сигналов, отраженных от цели, облучаемой гармонической волной.

Функция $r(\tau)$ убывает с ростом $|\tau|$ и можно указать интервал времени $\tau_{\text{фл}}$, связанный с шириной спектра флуктуаций, в пределах которого величина $r(\tau)$

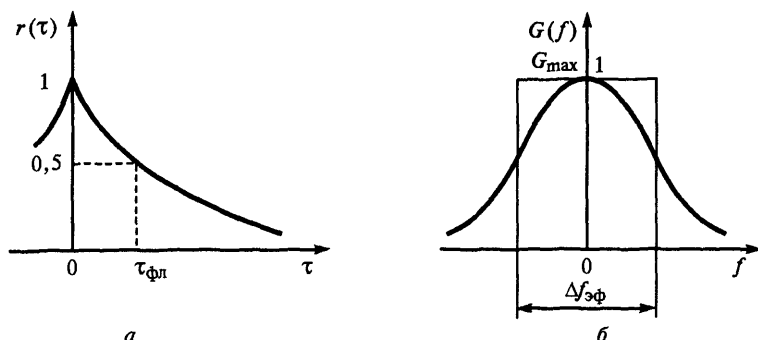


Рис. 7.18. Возможный способ определения интервала корреляции флуктуации (а) и эффективной ширины энергетического спектра (б)

уменьшается от 1 до заданного значения (например, до 0,5). Таким образом, интервал корреляции флуктуаций — это промежуток времени между сечениями случайного процесса, в пределах которого статистические (корреляционные) связи этого процесса достаточно сильные. Другими словами, возможен вероятностный прогноз любой реализации случайного процесса на время порядка $\tau_{\text{фл}}$, если известна информация об ее поведении в прошлом. Однако любая попытка осуществить прогнозирование на время, существенно превышающее интервал корреляции, окажется безрезультатной: мгновенные значения, столь далеко отстоящие по времени, практически некоррелированы. На рис. 7.18, а показан один из способов определения интервала корреляции флуктуаций.

На практике численные значения интервала корреляции флуктуаций, например для аэродинамических целей, составляют десятки–сотни миллисекунд [31, 42]. Интервал корреляции флуктуаций связан обратно пропорциональной зависимостью с эффективной шириной энергетического спектра флуктуаций [31]:

$$\tau_{\text{фл}} \approx \frac{1}{\Delta f_{\text{эф}}}. \quad (7.26)$$

Эффективную ширину энергетического спектра при этом определяют следующим образом (рис. 7.18, б):

$$\Delta f_{\text{эф}} = \frac{1}{G_{\text{max}}} \int_{-\infty}^{\infty} S(f) df. \quad (7.27)$$

Наличие флуктуаций отраженного от цели сигнала разрушает его закономерную фазовую структуру. Для того чтобы ее сохранить, необходимо

выбирать время наблюдения сигнала T_n существенно меньше интервала корреляции флуктуаций (рис. 7.19):

$$T_n \ll \tau_{\text{фл}}. \quad (7.28)$$

За это время флуктуации амплитуды и фазы будут пренебрежимо малы, т. е. амплитуда сигнала практически не изменится, а фазовая структура останется закономерной.

Конкретные значения как амплитуды сигнала на интервале наблюдения, так и его начальной фазы β могут быть произвольными. Физически это означает, что за время наблюдения T_n временное расположение блестящих точек относительно РЛС практически не изменяется, а отраженный от цели сигнал на интервале времени T_n является *когерентным* (если, конечно, обнаружимый цель сигнал является когерентным на интервале времени $\tau_{\text{ког}} > T_n$). Термин «когерентность», как известно, заимствован из оптики, где его используют для описания корреляции световых волн. Когерентность — это свойство двух процессов (полей), характеризующее их способность взаимно усиливать или ослаблять друг друга при сложении. Слово «когерентность» происходит от латинского слова *cohaerens* — *находящийся в связи*. Это понятие в равной мере применимо к колебаниям, к волнам любой физической природы любого диапазона частот.

В радиотехнике понятие когерентности трактуют несколько иначе. Под когерентностью понимают детерминированную (закономерную, неслучайную, жесткую) связь между фазами двух или нескольких сигналов.

Рассмотрим теперь понятие когерентности применительно к радиолокации. Обработку сигналов в радиолокации, как правило, можно условно разделить на пространственную и временную. В силу этого раздельно рассматривают понятия *временной когерентности* и *пространственной когерентности*.

Понятие *временной когерентности*, в отличие от данного выше определения, относят к одному сигналу (непрерывному, импульсному, пачке импульсов). Этот сигнал может быть как детерминированным, так и случайным. *Когерентным во времени* называют сигнал, обладающий на интервале наблюдения закономерной фазовой структурой. Примерами радиолокационных сигналов с временной когерентностью являются детерминированный сигнал, т. е. сигнал с полностью известными параметрами, и квазидетерминированный сигнал, т. е. сигнал, у которого случайными являются началь-

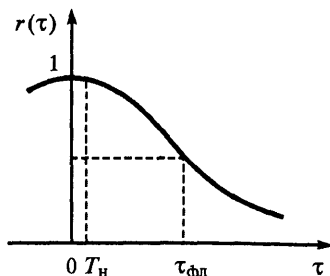


Рис. 7.19. Определение времени наблюдения сигнала, при котором сохраняется фазовая структура сигнала, отраженного от флуктуирующей цели

ная фаза и амплитуда. Когерентным можно считать также флуктуирующий случайный сигнал, наблюдаемый на ограниченном временном промежутке, не превышающем интервал корреляции флуктуаций.

Знание интервала корреляции флуктуаций необходимо в радиолокационной практике для выбора длительности τ_c импульсного сигнала или времени когерентного накопления $T_{\text{кн}}$ непрерывного сигнала. При выполнении условия

$$\tau_c \ll \tau_{\text{фл}} \quad (7.29)$$

или условия

$$T_{\text{кн}} \ll \tau_{\text{фл}} \quad (7.30)$$

наличие флуктуационной модуляции не разрушает когерентность сигнала, т. е. существенно не искажает структуру сигнала. Значит, на краях интервала τ_c (или $T_{\text{кн}}$) значения функции $B(t)$ оказываются коррелированными, что позволяет пренебречь ее временной зависимостью и считать $B(t) = B$ комплексной случайной величиной. Таким образом, флуктуации ЭПР при выполнении условий (7.29), (7.30) не разрушают когерентность сигнала, но обуславливают случайную амплитуду B и случайную начальную фазу β принимаемого сигнала. Отметим, что когерентное накопление на интервале $0 \dots T_{\text{кн}}$, $T_{\text{кн}} \ll \tau_{\text{фл}}$, можно проводить даже для непрерывного шумового зондирующего сигнала, если при излучении его структура запоминается для последующего учета в схеме обработки.

Понятие *пространственной когерентности* будет рассмотрено далее (см. § 7.7).

7.6. Особенности обработки радиолокационных сигналов на фоне стационарного аддитивного гауссовского белого шума

7.6.1. Общие сведения о современных методах обработки

Обработка принимаемых радиолокационных сигналов на фоне помех сводится, как показано в гл. 3, к вычислению достаточных статистик или практически реализуемых их приближений, т. е. к выполнению определенных математических операций над принимаемой смесью полезных сигналов и помех. Методы вычислений подразделяют на *аналоговые* и *цифровые*.

Аналоговые методы вычислений применяют к непрерывно изменяющимся величинам без их дискретизации и перевода в цифровую форму. Не-

обходимые вычислительные операции реализуют, используя аналогию законов изменения вычисляемых (или промежуточных) величин и величин произвольной физической природы, соответствующих входным данным. Обработку сигналов на основе аналоговых вычислений называют *аналоговой обработкой*.

Цифровые методы вычисления применяют к множествам чисел (двоичных, в частности) с ограниченной разрядностью. Обработку на основе цифровых вычислений называют *цифровой*. Цифровая обработка преобразованных по частоте принимаемых колебаний предусматривает их предварительную дискретизацию по времени (рис. 7.20, а, см. также гл. 2) и уровням мгновенных значений (рис. 7.20, б). Устройства, обеспечивающие преобразование непрерывно изменяющихся величин в цифровую форму (рис. 7.20, а—в) называют *аналого-цифровыми преобразователями* (АЦП) или *преобразователями напряжение—код*.

Промежуточное место между аналоговыми и цифровыми методами обработки занимают *дискретно-аналоговые*. Дискретизация мгновенных значений по времени сочетается в них с последующей аналоговой обработкой этих значений. Дискретно-аналоговые методы реализуют, используя полупроводниковые приборы с переносом заряда и зарядовой связью, способные запоминать и выдавать с задержками большое число аналоговых дискретизированных по времени величин [35, 74].

К устройствам обработки радиолокационных сигналов предъявляют все возрастающие и одновременно противоречивые требования, а именно: расширение динамического диапазона входных сигналов (для работы в сложных помеховых ситуациях); обеспечение обработки широкополосных сигналов (для повышения разрешающей способности по дальности); обеспечение обработки протяженных сигналов (для повышения разрешающей

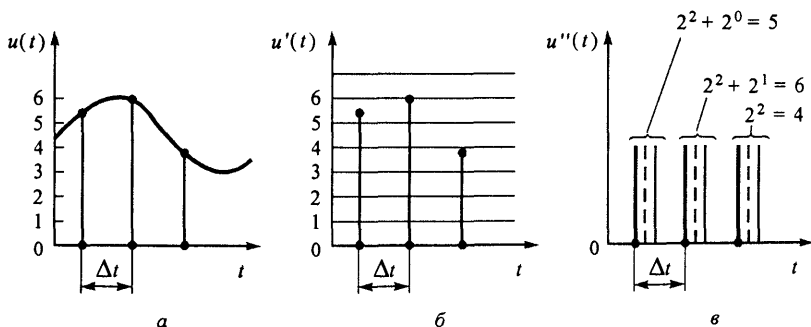


Рис. 7.20. Преобразование непрерывно изменяющихся величин в цифровую форму: предварительная дискретизация по времени (а) и уровням мгновенных значений (б), преобразование дискретных значений в цифровой код (в)

способности по радиальной скорости); повышение точности выполнения вычислительных операций; повышение надежности; стандартизация; микроминиатюризация; упрощение эксплуатации; снижение стоимости.

Наибольшее значение в настоящее время приобрели цифровые методы обработки (несмотря на отдельные ограничения по широкополосности и стоимости). Это не означает, что аналоговая обработка потеряла свое значение. Развиваются новые методы аналоговой обработки: *акустические, акустооптические, спиновые* и др. [35, 40, 74, 75, 84].

В настоящем параграфе выявляются основные операции, необходимые при оптимальной обработке радиолокационных сигналов на фоне аддитивного стационарного гауссовского белого (внутриприемного) шума. Их техническая реализация может быть как аналоговой, так и цифровой.

7.6.2. Обработка одиночных и пачечных импульсных сигналов без внутримпульсной модуляции

Одиночные радиоимпульсы. Рассмотренные в гл. 3 устройства оптимальной (согласованной) обработки одиночных радиоимпульсов со случайными параметрами предназначены для обнаружения сигналов с известной доплеровской частотой или с частотой, удовлетворяющей условию $F_D < 1/\tau_n$. Однако в радиолокации часто возникает задача обнаружения сигналов, отраженных от движущихся целей. Тогда принимаемый сигнал содержит не только неизвестное время запаздывания, но и неизвестную доплеровскую частоту. Устройство фильтровой обработки (близкое к оптимальному при $F_D < 1/\tau_n$) становится многоканальным по доплеровской частоте (рис. 7.21). Взаимные расстройки ΔF средних частот амплитудно-частотных характеристик согласованных фильтров (СФ) многоканального

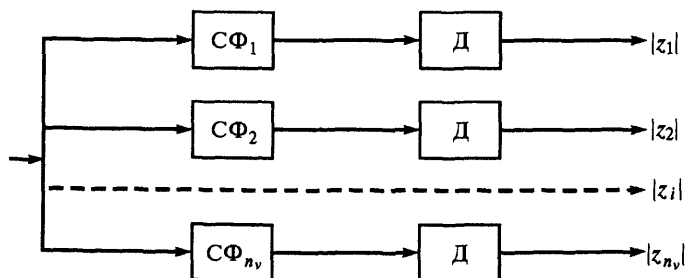


Рис. 7.21. Многоканальное устройство фильтровой обработки сигналов с неизвестной доплеровской частотой:

СФ — согласованный фильтр; Д — детектор

обнаружителя выбираются исходя из допустимых энергетических потерь $\nu = 10 \lg(1/\rho^2(0, \Delta F/2))$ дБ на краях полосы пропускания. Для прямоугольных радиоимпульсов при $\Delta F = 1/\tau_n$ энергетические потери равны 4 дБ. Необходимое число согласованных фильтров при этом определяется из формулы

$$n_\nu = 1 + 2F_{D\max} \tau_n, \quad (7.31)$$

где $F_{D\max}$ — максимально возможная доплеровская частота движущейся цели (в сантиметровом диапазоне длин волн для большинства аэродинамических целей $F_{D\max} \leq 100$ кГц).

Одиночные когерентные импульсные сигналы без внутриимпульсной модуляции встречаются в некогерентных импульсных РЛС, в доплеровских РЛС с непрерывным излучением и др. В некогерентных импульсных РЛС обычно используют короткие когерентные радиоимпульсы (длительностью $\tau_n \leq 10$ мкс), из которых в процессе обзора по угловым координатам образуется некогерентная пачка. В этом случае $\tau_n F_{D\max} < 1$, селекция по скорости нереализуема, а число каналов обработки равно 1 [31, 34].

В доплеровских РЛС с непрерывным излучением импульсные сигналы образуются в процессе обзора пространства по угловым координатам. Длительность подлежащего обработке когерентного радиоимпульса определяется временем облучения цели и может составлять единицы—десятки миллисекунд. При обработке сигналов с такими длительностями появляется возможность селекции движущихся целей по радиальной скорости, а устройство обработки становится многоканальным.

Когерентные пачки радиоимпульсов. Фильтр, согласованный с пачкой радиоимпульсов, как показано в гл. 3, включает согласованный фильтр для одиночного импульса и когерентный сумматор импульсов пачки, снимаемых с линии задержки с отводами (см. рис. 3.11). Подбирая расположение отводов, их число и коэффициенты передачи, можно получить согласованную импульсную характеристику фильтра в виде пачки радиоимпульсов с заданной огибающей. Это позволяет решить задачу согласованной обработки когерентной пачки импульсов с неизвестным временем запаздывания и известной или несущественной ($F_D < 1/(MT_n)$) доплеровской частотой. Если доплеровская частота существенна ($F_D > 1/(MT_n)$) и неизвестна, то необходим согласованный фильтр, многоканальный по доплеровской частоте (по радиальной скорости). На рис. 7.22 показан принцип построения многоканального фильтра с нониусными линиями задержки для пачки из нечетного числа $M = 7$ радиоимпульсов. Фильтр пригоден для работы в диапазоне ско-

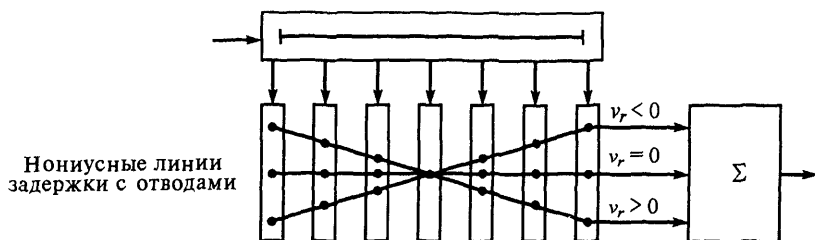


Рис. 7.22. Многоканальное устройство фильтровой обработки когерентной пачки радиопульсов с неизвестной доплеровской частотой

ростей при произвольно большой длительности пачки (ускорение цели отсутствует). Нониусные линии задержки учитывают растяжение ($v_r > 0$) или сжатие ($v_r < 0$) пачки при отражении от движущейся цели в $(c + v_r)/(c - v_r)$ раз. Для используемых на практике длительностях пачек импульсов нониусные линии задержки можно заменить фазовращателями [35, 74]. Необходимое число каналов обработки определяется из формулы

$$n_v = 1 + 2F_{D\max}MT_n. \quad (7.32)$$

Если $F_{D\max} \geq 1/T_n$, то

$$n_v = 1 + 2M. \quad (7.33)$$

На рис. 7.23 представлено устройство корреляционно-фильтровой обработки когерентной пачки радиопульсов с неизвестными временем запаздывания и доплеровской частотой, использующее видеоимпульсное стробирование [34, 35]. Это приводит к двойной многоканальности обработки: по дальности и радиальной скорости. Каждый канал дальности строится пачкой видеоимпульсов. Длительность каждого стробирующего импульса равна τ_n , период следования импульсов равен T_n , а длительности совокупности пачек стробирующих импульсов соответствуют ожидаемой длительности принимаемого сигнала, которая, в свою очередь, определяется временем облучения цели. Каждый канал дальности содержит набор узкополосных фильтров (Φ) с шириной полосы пропускания $\sim 1/(MT_n)$, перекрывающих диапазоны доплеровских частот. Выходные напряжения каналов детектируются амплитудными детекторами (D) и могут дополнительно подаваться на некогерентный накопитель (на рис. 7.23 не показан). Выходные напряжения частотных подканалов каждого дальностного канала подаются на пороговые устройства ($ПУ$) и сумматоры. Выходные напряжения сумматоров дальностных каналов повторно стробируются видеоимпульсами «дальности». Исходная длительность отраженных радиопульсов при этом

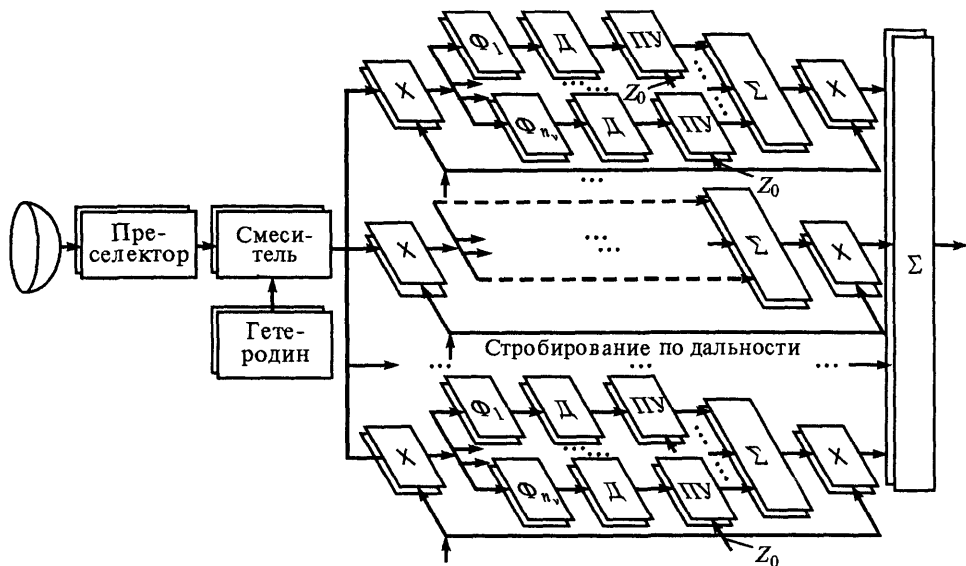


Рис. 7.23. Устройство корреляционно-фильтровой обработки когерентной пачки радиоимпульсов:

× — перемножитель; Φ — узкополосный фильтр; Д — амплитудный детектор; ПУ — пороговое устройство

восстанавливается. Простробированные импульсы всех каналов объединяют в выходном сумматоре. Обычно видеоимпульсные последовательности любых двух смежных каналов смещены относительно друг друга на элемент разрешения обрабатываемого сигнала по времени запаздывания ($\Delta\tau = \tau_n$), а средние частоты любых смежных узкополосных фильтров отличаются на величину разрешающей способности по доплеровской частоте $\Delta F \sim 1/(MT_n)$. В этом случае число каналов обработки по дальности определяется формулой $n_r = t_{3\max} / \tau_n$, если $t_{3\max} < T_n$. Если выполняется условие $t_{3\max} \geq T_n$, то $n_r = T_n / \tau_n = Q$, где Q — скважность пачки, а число каналов обработки по скорости, как и для фильтровой схемы, определяется формулами (7.32) и (7.33).

Довольно часто число каналов обработки может быть меньше этих значений. Тогда в устройстве корреляционно-фильтровой обработки предусматривается дополнительный поиск по дальности и скорости.

7.6.3. Обработка частотно-модулированных радиоимпульсов

Особенности обработки частотно-модулированных (ЧМ) сигналов рассмотрим на примере обработки линейно-частотно-модулированных радиоимпульсов (ЛЧМ), широко используемых в современных РЛС.

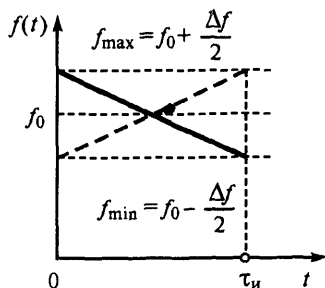


Рис. 7.24. Законы изменения мгновенной частоты во времени: ЛЧМ сигнала (штриховая линия), импульсной характеристики согласованного фильтра (сплошная линия)

Согласованная фильтровая обработка ЛЧМ радиоимпульсов проводится с использованием дисперсионных фильтров, т. е. линейных систем, групповое запаздывание в которых зависит от частоты [31, 35, 74]. Импульсная характеристика согласованного фильтра представляет собой ЛЧМ колебание, зеркальное сигнальному (см. гл. 3). Если мгновенная частота сигнала со временем нарастает (рис. 7.24, штриховая линия), то мгновенная частота колебаний в импульсной характеристике убывает (рис. 7.24, сплошная линия), и наоборот.

Дисперсионная характеристика согласованного фильтра — зависимость группового запаздывания $t_{гр}(f)$ в фильтре от частоты — определяется производной от фазочастотной характеристики фильтра. Нелинейность фазочастотной характеристики линейных сред, приводящую к зависимости группового запаздывания от частоты, в физике называют *явлением дисперсии*, что и определило термин «дисперсионная характеристика фильтра». Дисперсионная характеристика согласованного фильтра имеет вид

$$t_{гр}(f) = t_0 + (2\pi)^{-1} d\varphi(f)/df, \quad (7.34)$$

где $\varphi(f)$ — фазочастотная характеристика согласованного фильтра, t_0 — постоянная задержка в фильтре.

Для больших $n = \tau_n \Delta f_d$, когда $\arctg n \approx \pi/2$, дисперсионная характеристика (7.34) в соответствии с соотношением (4.28) для ЛЧМ сигналов приобретает вид [31, 35, 74]

$$t_{гр}(f) = t_0 - \tau_n (f - f_0) \Delta f. \quad (7.35)$$

Выражение (7.35) поясняет формирование импульсной характеристики согласованного с ЛЧМ сигналом фильтра под воздействием дельта-функции. Группы гармонических составляющих дельта-функции более низких мгновенных частот задерживаются в фильтре на большее время и формируются на выходе позже, чем группы составляющих более высоких частот.

Дисперсионные фильтры используют не только при обработке, но и при формировании сигналов, осуществляемом под воздействием на них коротких радио- или видеоимпульсов.

Поясним механизм сжатия ЛЧМ сигналов. Группы колебаний более низких мгновенных частот (штриховая линия на рис. 7.24) воздействуют в моменты $t_{\text{возд}}(f)$ раньше, но задерживаются на большее время $t_{\text{гр}}(f)$. Группы колебаний более высоких частот воздействуют позже, но задерживаются на меньшее время. Это обеспечивает совмещение групп радиоимпульсов с различными несущими частотами. При их синфазном наложении в момент времени t_0 формируется более короткий сжатый радиоимпульс. Условие наложения

$$t_{\text{возд}}(f) + t_{\text{гр}}(f) = t_0 = \text{const} \quad (7.36)$$

является необходимым условием синтеза фазочастотной характеристики фильтра в режиме согласования. Видоизмененные трактовки механизма сжатия рассматривались в гл. 3 и 4.

Аналоговые дисперсионные фильтры часто выполняются на основе акустических линий задержки с поверхностными или объемными волнами. Акустические линии позволяют получать приемлемые задержки при малых габаритах вследствие существенно меньшей скорости распространения механических колебаний по сравнению с электромагнитными. На основе использования поверхностных акустических (ультразвуковых) волн можно строить многоотводные линии задержки, обеспечивающие эффект искусственной дисперсии, а на основе использования объемных волн — линии задержки с естественной дисперсией [35].

Фильтры сжатия с искусственной дисперсией на основе акустических линий задержки с поверхностными волнами (рис. 7.25, а). Звукпровод такой линии выполняют из пьезоэлектрического монокристалла, например из кварца, ниобата лития и т. д. На поверхность кристалла наносят проводящие металлические входные и выходные электроды, например, фотоспособом. Под воздействием входного напряжения между электродами создается переменное электрическое поле. Вследствие пьезоэлектрического эффекта вдоль поверхности электрода возбуждается акустическая волна. Под воздействием этой волны в соединенных между собой электродах в силу обратного пьезоэлектрического эффекта наводятся ЭДС и создаются токи в выходной нагрузке. Требуемая зависимость группового запаздывания от частоты обеспечивается путем расстановки электродов с интервалами, изменяющимися по закону, близкому к закону арифметической прогрессии. При возбуждении линии дельта-импульсом через частотно-неселективный электрод это обеспечивает получение импульсной характеристики в виде набора коротких импульсов (в пределе дельта-импульсов), распределенных на длительности τ_n с интервалами при $f_0 \gg \Delta f$ по закону арифметической прогрессии. Последовательность таких видеоимпульсов с неизменным пе-

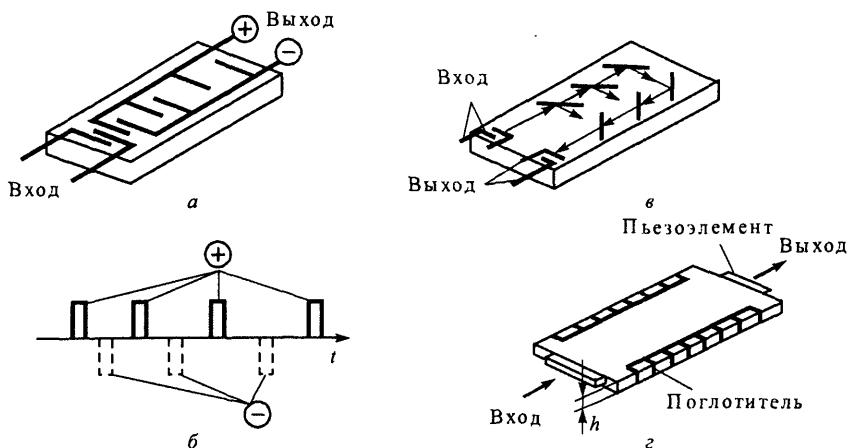


Рис. 7.25. Примеры аналоговых фильтров сжатия: с искусственной дисперсией на основе акустической линии задержки с поверхностными волнами (а), импульсная характеристика фильтра (б), дисперсионный фильтр на поверхностных волнах отражательного типа (в), фильтр сжатия с естественной дисперсией на основе объемных волн в акустических волноводах (г)

приодом $T_0 = 1/f_0$ свелась бы к наложению гармоник и постоянной составляющей. Последовательность видеоимпульсов с изменяющейся скважностью (рис. 7.25, б) позволяет получить совокупность ЛЧМ колебаний с девиациями мгновенной частоты $m(f_{\max} - f_{\min})$, где $m = 1, 2, \dots$. Неравномерность ЛЧМ сигналов по амплитуде устраняют путем вариаций длины электродов (аподизации): удлинения в местах разрежения и укорочения в местах сгущения отводов. За счет элементов связи или внешних цепей выделяют ЛЧМ колебание с фиксированным номером m , в частности $m = 1$. Балансное построение фильтра (см. рис. 7.25, а, б), для которого четные гармоники отсутствуют, облегчает выделение нужной нечетной гармоники.

Многоотводные линии описанных конструкций позволяют реализовать фильтры сжатия импульсов с длительностями линейного участка дисперсионной характеристики от единиц до сотен микросекунд, с полосами частот от единиц до сотен мегагерц, с произведениями $n = \tau_{\text{л}} \Delta f_{\text{д}}$ порядка десятков, сотен и тысяч [35, 74].

Для уменьшения уровня боковых лепестков сжатых радиоимпульсов проводится скругление амплитудно-частотного спектра (всесовое суммирование во времени). Для этого обеспечивают необходимое скругление амплитудно-частотной характеристики (АЧХ) фильтра, например, за счет разновидности аподизации: уменьшения длины электродов, формирующих край-

ние мгновенные частоты. В целях скругления АЧХ фильтра используют даже многоэлементные входные и выходные электроды как самостоятельные корректирующие фильтры (каскады) с требуемой частотной зависимостью.

Применение такой несогласованной фильтрации приводит, очевидно, к потерям в отношении сигнал—шум по сравнению со значением $2E/N_0$ — максимальным отношением сигнал—шум на выходе согласованного фильтра. Кроме того, на выходе фильтра расширяется главный пик ЛЧМ сигнала. Однако эти недостатки окупаются положительным фактором — снижением уровня боковых лепестков выходного сигнала по оси времени (дальности).

В качестве корректирующих могут быть использованы фильтры, АЧХ которых описываются функциями

$$|K(f)| = a + (1 - a) \cos\left(\frac{b\pi f}{\Delta f}\right), \quad a < 1,$$

где a и b имеют различные значения. Наименьший уровень боковых лепестков для этого класса функций при $b = 2$ обеспечивается при $a = 0,88$ (фильтр Хэмминга). Если при согласованной фильтрации ЛЧМ импульса уровень максимального бокового лепестка относительно главного составляет $-13,2$ дБ, то при использовании фильтра Хэмминга данный уровень равен $-42,8$ дБ. При этом главный максимум расширяется примерно в 1,5 раза, а потери в отношении сигнал—шум составляют 1,34 дБ.

Существуют и другие способы подавления боковых лепестков выходного сигнала. Они сводятся к специальному подбору закона частотной модуляции, отличного от линейного, или формы огибающей зондирующего импульса.

Дисперсионные фильтры на поверхностных волнах отражательного типа содержат прорезы, отражающие акустическую волну. Расстояния между прорезами выбираются по закону, близкому к закону арифметической прогрессии (рис. 7.25, в). Достоинством подобных фильтров считают менее жесткие допуски на точность выполнения.

Фильтры сжатия с естественной дисперсией на основе объемных волн в акустических волноводах (рис. 7.25, г) обеспечивают реализацию линейных участков дисперсии от десятых до десятков миллисекунд. Выполняются в виде металлических полосок (лент), толщина h которых соизмерима с длиной акустической волны в металле. Входной и выходной пьезоэлементы обеспечивают возбуждение и съём акустических волн, обычно продольных. Для снижения уровня паразитных отражений боковые торцы волновода покрывают поглотителем. Чтобы ослабить влияние нежелательных типов волн, способных распространяться в акустическом волноводе, используют частотную селекцию на его входе и выходе. Дисперсионная характеристика

$t_{\text{гр}}(f)$ на рабочем типе волны имеет практически линейный участок. Его средняя частота f_0 обратно пропорциональна толщине h , а временная протяженность пропорциональна длине волновода. Акустические фильтры волноводного типа в принципе проще многоотводных на поверхностных акустических волнах, но при реализуемых значениях h сравнительно узкополосны, их параметры поддерживаются менее жестко, чем параметры многоотводных фильтров.

Отметим также, что ЛЧМ радиоимпульсы и совокупность одновременно излучаемых ЛЧМ радиоимпульсов с одинаковыми параметрами модуляции, но взаимно смещенные по частоте на одну и ту же величину $F_0 \geq \Delta f_d$, могут обрабатываться с помощью рециркуляторов со смещающим гетеродином [99].

Линейно-частотно-модулированный сигнал обладает важным достоинством: он единственный из сложных (широкополосных) радиолокационных сигналов, обеспечивающий одноканальную обработку не только по дальности, но и по радиальной скорости [100], если $\Delta f_d \geq F_{D \max}$ (что обычно выполняется). Поэтому устройство обработки одиночных ЛЧМ радиоимпульсов обычно состоит из одного согласованного (дисперсионного) фильтра и амплитудного детектора (детектора огибающей).

Если $\Delta f_d < F_{D \max}$ (что возможно, например, при обнаружении высокоскоростных целей в миллиметровом диапазоне длин волн), то устройство обработки ЛЧМ радиоимпульсов соответствует схеме, показанной на рис. 7.21. В этом случае согласованные фильтры (СФ) являются дисперсионными, а число каналов по скорости определяется формулой

$$n_v = 1 + 2F_{D \max} / \Delta f_d. \quad (7.37)$$

Корреляционно-фильтровая обработка ЛЧМ сигналов имеет определенные особенности в зависимости от того, осуществляется полная или частичная фазочастотная демодуляция.

1. *Полная фазочастотная демодуляция.* Если закон модуляции гетеродина полностью повторяет закон модуляции ожидаемого сигнала, то последующая обработка сводится к накоплению демодулированного сигнала в узкополосном фильтре (контуре) с импульсной характеристикой в виде немодулированного по фазе (частоте) колебания. Узкополосный фильтр выполняет роль накопителя (интегратора) на промежуточной частоте.

Рассмотрим случай полной демодуляции непрерывного ЛЧМ сигнала с периодическим законом изменения частоты. Здесь предполагается, что сигнал с законом изменения частоты

$$\begin{aligned} f(t) &= f_{\min} + (t/T)\Delta f, & 0 \leq t \leq T, \\ f(t+T) &= f(t), \end{aligned} \quad (7.38)$$

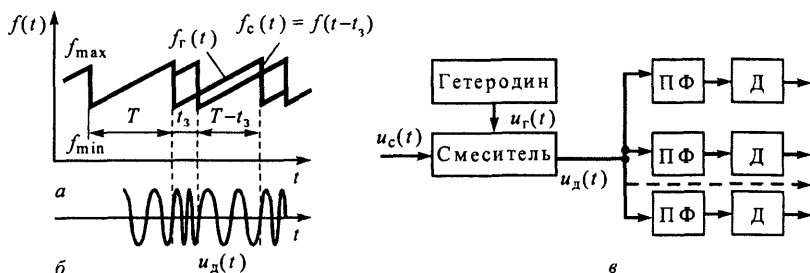


Рис. 7.26. Корреляционно-фильтровая обработка непрерывного ЛЧМ сигнала с периодическим законом изменения частоты: закон изменения мгновенной частоты отраженного $f_c(t) = f(t - t_3)$ и гетеродинного $f_r(t)$ колебаний (а), колебания разностной частоты $u_d(t)$ (б), многоканальный спектроанализатор (в)

излучается через частотно-независимую несканирующую антенну. От целей приходят непрерывные отраженные сигналы $u_c(t)$ с законами изменения мгновенной частоты $f_c(t) = f(t - t_3)$ (рис. 7.26, а, в).

Если $t_3 \ll T$, то в качестве гетеродинного $u_r(t)$ можно взять смещенное на промежуточную частоту зондирующее колебание. На выходе смесителя формируются колебания $u_d(t)$ разностной частоты (рис. 7.26, б), так называемой частоты дальности, принимающей на большей части периода T значение

$$f(t) - f(t - t_3) = \frac{\Delta f}{T} \frac{2r}{c} = f_r. \quad (7.39)$$

Для выделения колебаний $u_d(t)$ на схеме (рис. 7.26, в) предусмотрен спектроанализатор в виде набора полосовых фильтров (ПФ) с амплитудными детекторами (Д). Полосы пропускания фильтров выбираются обратными длительностям когерентно обрабатываемых сигналов. Когда условие $t_3 \ll T$ не соблюдается, спектроанализатор разбивается на группы фильтров. На предшествующие им смесители подаются сдвинутые по времени гетеродинные напряжения. Если дополнительно реализуется обзор пространства, например путем частотного сканирования, то принимаемые колебания оказываются импульсными, что учитывается при подборе полос пропускания фильтров спектроанализатора. Моменты прихода импульсов могут давать информацию об угловой координате.

2. Частичная фазочастотная демодуляция. Этот случай рассмотрим на примере обработки ЛЧМ сигнала с большой частотной девиацией, обеспечивающей разрешение целей (элементов цели) в некотором строке даль-

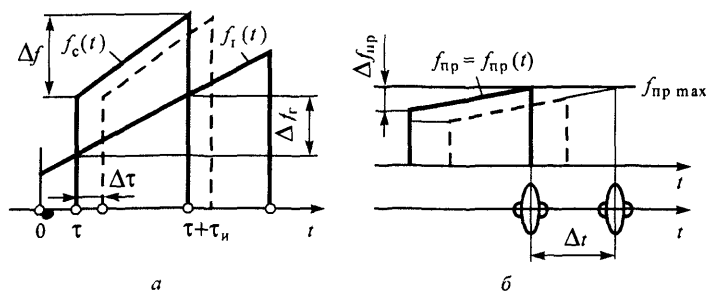


Рис. 7.27. Корреляционно-фильтровая обработка ЛЧМ сигнала с частичной фазочастотной модуляцией: *а* — зависимости изменения мгновенной частоты по времени обрабатываемых и гетеродинных колебаний, *б* — сигналы на выходе устройства

ности, выставляемом, например, по результатам предыдущей более узкополосной локации. В качестве гетеродина можно использовать генератор ЛЧМ колебаний, длительность которых превышает сумму длительностей сигнала и наблюдаемого строба, а производная частоты df_r/dt которых близка, но отличается от соответствующей производной колебаний сигнала. В результате гетеродинирования девиация Δf принимаемого сигнала существенно снижается до величины $\Delta f_{пр}$, отличной, однако, от нуля (рис. 7.27).

Несмотря на сужение полосы частот, разрешающая способность по дальности в выделенном стробе не ухудшается. В дополнение к имеющемуся взаимному временному смещению отраженные целями (элементами цели) радиоимпульсы получают при гетеродинировании дополнительные частотные смещения и, как следствие, дополнительные временные смещения в фильтре сжатия, зависящие от дальности. Расширение сжатых импульсов из-за сужения полосы частот канала обработки сопровождается общим растяжением масштаба времени и не ведет к ухудшению разрешающей способности по дальности. Уменьшение же частотной девиации упрощает фильтр сжатия и последующий тракт обработки в условиях малоцелевой локации [35, 74].

7.6.4. Обработка фазоманипулированных радиоимпульсов

Рассмотрим согласованную фильтрацию фазоманипулированного радиоимпульса длительностью τ_n (рис. 7.28). Внутримпульсная модуляция обеспечивается изменением (манипуляцией) на π начальной фазы отдельных элементов сигнала — парциальных радиоимпульсов длительности

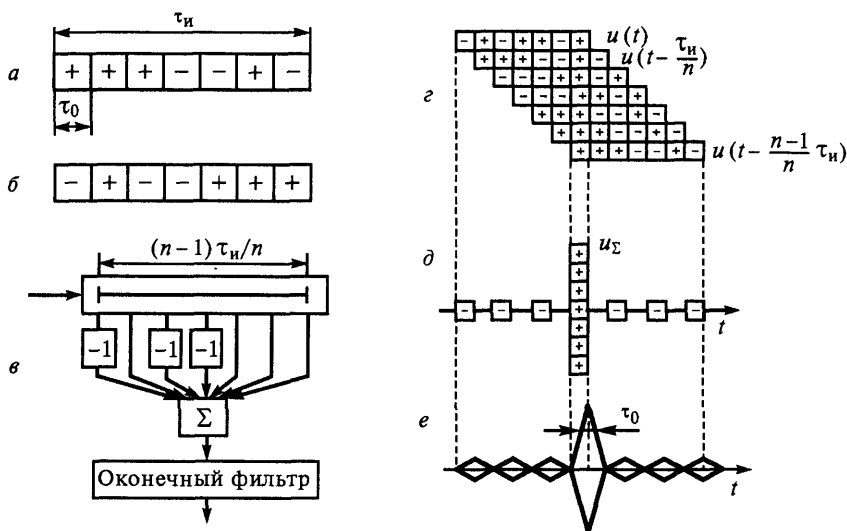


Рис. 7.28. Согласованная фильтрация фазоманипулированного радиоимпульса: условное изображение ФМ радиоимпульса (а) и импульсной характеристики согласованного фильтра (б), согласованный фильтр (в), условное изображение обрабатываемых колебаний на входе (г) и выходе сумматора (д), сигнал на выходе согласованного фильтра (е)

$\tau_0 = \tau_n / n$, где $n = 7$. Парциальные радиоимпульсы с неизменной и измененной на π начальной фазой условно показаны на рис. 7.28, а знаками + и – соответственно. Условное изображение импульсной характеристики согласованного фильтра, зеркальной обрабатываемому сигналу, показано на рис. 7.28, б аналогичным образом. Такой импульсной характеристикой обладает фильтр, изображенный на рис. 7.28, в. Основой его является линия задержки с отводами и общим сумматором. Часть отводов подключена к сумматору через инверсные каскады. Напряжение сумматора проходит через оконечный фильтр, согласованный с парциальными радиоимпульсами. При воздействии на фильтр (см. рис. 7.28, в) дельта-функции с выхода сумматора снимается последовательность манипулированных по знаку дельта-функций, каждая из которых возбуждает в оконечном фильтре парциальный радиоимпульс с установленной начальной фазой. Совокупность состыкованных таким образом радиоимпульсов образует импульсную характеристику фильтра (см. рис. 7.28, б). Зеркальная характеристика «перевернута» по отношению к сигналу.

На рис. 7.28, г показаны условные изображения незадержанного и задержанных в линии задержки фазоманипулированных радиоимпульсов (часть из которых прошла инверсные каскады) на входе сумматора, а на

рис. 7.28, δ — результат суммирования. Ясно виден эффект когерентного (с точностью до фазы) накопления элементов сигнала в пике. На рис. 7.28, e представлен результат согласованной фильтрации просуммированных парциальных радиоимпульсов в оконечном фильтре (высокочастотное заполнение на рис. 7.28, e не показано). В результате согласованной фильтрации происходит преобразование прямоугольных парциальных радиоимпульсов в ромбовидные с длительностью τ_0 на уровне 0,5 от максимального значения. Фазоманипулированный сигнал на выходе согласованного фильтра сжимается во времени в $n = \tau_n / \tau_0$ раз, что позволяет соответственно повысить разрешающую способность РЛС по дальности. Отметим, что согласованная обработка внутри каждого парциального радиоимпульса может быть заменена квазисогласованной (см. гл. 3) при приемлемых энергетических потерях.

Определенным недостатком согласованной обработки фазоманипулированного радиоимпульса (так же как и ЛЧМ радиоимпульса) является сравнительно высокий относительный уровень боковых лепестков, составляющий в рассматриваемом примере величину, равную $1/n$.

Уменьшение уровня боковых лепестков по оси времени для ФМ сигнала, как и для ЛЧМ сигнала, достигается применением специальных несогласованных фильтров. Общая структурная схема фильтра, сжимающего ФМ сигнал, а также подавляющего боковые лепестки, включает в себя линию задержки с максимальной задержкой τ_3 , большей длительности импульса τ_n . Как и в согласованном фильтре, отводы линии задержки расположены так, чтобы сигнал между ними задерживался на время, равное длительности дискрета τ_0 , однако число отводов k больше, чем у согласованного фильтра. На выходе отводов имеются фазовращатели ϕ_i и аттенюаторы a_i , $i = 1, 2, \dots, k$. После суммирования выходных сигналов аттенюаторов осуществляется фильтрация в полосе $1/\tau_0$. Применение такого фильтра приводит к проигрышу в отношении сигнал—шум по сравнению с согласованным фильтром. В отличие от ЛЧМ импульса подавление боковых лепестков в случае ФМ сигнала не сопровождается расширением главного максимума, вместо этого расширяется вся область боковых лепестков выходного сигнала.

Рассматриваемый фильтр при соответствующей настройке фазовращателей и аттенюаторов позволяет значительно уменьшить уровень боковых лепестков, причем степень их подавления возрастает с ростом числа отводов $k = \tau_3 / \tau_0 + 1$ линии задержки и, следовательно, с увеличением максимальной задержки τ_3 . Если, например, используется фильтр с 33 отводами для обработки ФМ сигнала, манипулированного 13-позиционным кодом

Баркера, то уровень боковых лепестков удастся уменьшить с 22 до 40 дБ при энергетических потерях менее 0,4 дБ [8].

При обработке ФМ сигналов, отраженных от движущихся целей ($F_D > 1/\tau_n$), фильтровое устройство становится многоканальным по доплеровской частоте (радиальной скорости). Возможный вариант построения согласованного фильтра для семизлементного ФМ радиопульса показан на рис. 7.29. Так же, как и для когерентной пачки, используются нониусные линии задержки, которые для наиболее распространенных параметров ФМ сигналов можно заменить на фазовращатели. Число каналов обработки по доплеровской частоте определяется формулой (7.20).

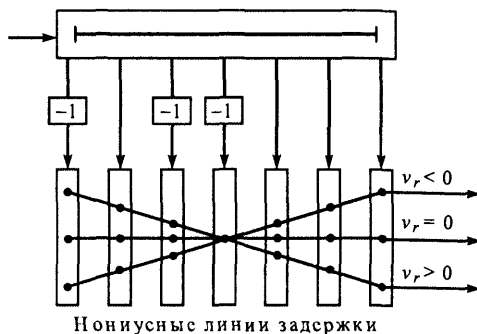


Рис. 7.29. Многоканальное устройство фильтровой обработки ФМ радиопульса с неизвестной доплеровской частотой

7.7. Основные виды помех активной радиолокации

Как и в любой радиотехнической системе, в радиолокации может существенно сказываться влияние различного рода помех [31, 35, 40, 43, 73—75, 86—88]. Роль помех в активной радиолокации может оказаться еще большей, чем в других РТС, поскольку обычно имеет место существенное ослабление сигнала на пути распространения до цели и обратно. Кроме того, в радиолокации важное значение имеют некоторые специфические виды помех, с которыми гораздо реже приходится считаться, например в радиосвязи. Такими помехами являются, в частности, *пассивные помехи*, вызываемые переотражениями от мешающих объектов. По своему происхождению помехи могут быть естественными, взаимными и искусственными.

Естественными являются помехи природного происхождения. Например, естественные пассивные помехи образуются в результате переотражений от холмов, гор, облаков и т. д. Естественные активные помехи создаются излучениями Солнца и других внеземных источников.

Взаимными называют активные помехи, вызываемые влиянием излучений различных радиоэлектронных средств друг на друга. Наряду с взаимными активными помехами иногда наблюдаются также взаимные пассивные помехи, когда в гористой местности помеха радиолокатору создается за счет переотражения колебаний, излучаемых другим радиолокатором.

Искусственные активные и пассивные помехи создаются для радиолокаторов военного назначения. Такие помехи широко применялись во время боевых действий в ходе Второй мировой войны, войн в Корее, во Вьетнаме, на Ближнем Востоке, в Югославии и других локальных конфликтах. Создание помех является одной из форм радиоэлектронной войны (борьбы), а радиоэлектронная война считается важной составной частью информационной войны.

По характеру воздействия на подавляемое средство помехи делятся на *маскирующие* и *имитирующие*. Маскирующие помехи создают фон, на котором трудно выделить сигнал, прикрываемый помехой; наряду с этим они обычно подавляют сигнал в нелинейных элементах приемника РЛС. Имитирующие помехи создают эффект ложных целей, затрудняя получение информации об истинных целях. Каждая из трех указанных выше разновидностей помех — естественная, взаимная и искусственная — может быть в свою очередь маскирующей и имитирующей. Ниже будут рассмотрены возможности создания маскирующих помех и принципы защиты от активных и пассивных маскирующих помех.

7.7.1. Естественные и взаимные маскирующие активные помехи и принципы защиты от них

Существуют два основных вида источников естественных маскирующих активных помех: дискретные и распределенные. К дискретным источникам помех относятся Солнце, Луна и радиозвезды. К распределенным источникам — галактические шумы, излучение атомарного водорода и шумы атмосферы. Из дискретных источников практически влияние на работу радиолокационных станций СВЧ диапазона могут оказывать Солнце и в меньшей степени Луна. Плотность потока мощности Солнца на длине волны 10 см оказывается порядка $(10^{-20} \dots 10^{-18}) \text{ Вт/м}^2 \cdot \text{Гц}$, где большее число соответствует повышенной солнечной активности. Эта плотность превышает плотность излучения абсолютно черного тела при температуре 6000 К в $10 \dots 1000$ раз. На длине волны 1 м плотность потока мощности будет $(10^{-23} \dots 10^{-17}) \text{ Вт/м}^2 \cdot \text{Гц}$ [31, 43]. Из распределенных источников преобладающим является собственное тепловое излучение атмосферы.

В последнее время очень важную роль начинают играть взаимные помехи. По мере стремительного увеличения числа используемых радиоэлектронных средств резко возрастает опасность их взаимных влияний. Чтобы устранить эти влияния, практикуется плановое распределение рабочих частот между различными радиоэлектронными средствами (радиолокационными, в частности) как на основе международных соглашений, так и на основе внутренних регламентаций в пределах каждой страны, каждой отрасли на-

родного хозяйства и военного дела. Тем не менее, при отсутствии должных мер защиты от взаимных помех наблюдается взаимное влияние радиоэлектронных средств даже с различными рабочими частотами. Последнее имеет место при наличии *внеполосных* и *побочных излучений* радиоэлектронных средств. Наряду с внеполосными и побочными излучениями причиной взаимных помех являются *побочные каналы приема* в супергетеродинных приемных устройствах. Известно, что при воздействии на смеситель входящих колебаний частоты f и колебаний гетеродина частоты f_r на выходе смесителя образуются колебания ряда комбинационных частот $|nf \pm mf_r|$. Если какая-либо из этих частот совпадает с промежуточной, на которую настроены последующие каскады приемника, она усиливается и образуется побочный канал приема. В реальных условиях, когда $f_r \gg f_{пр}$, побочные каналы приема образуются на частотах входных колебаний

$$f_{mn} = \frac{1}{n}(mf_r \pm f_{пр}).$$

Характеристики направленности приемных и передающих антенн для внеполосных излучений, побочных излучений и каналов приема обычно отличаются от характеристик направленности для основных каналов излучения и приема, в первую очередь, значительно большим уровнем боковых лепестков.

Во многих случаях может создаться достаточно сложная обстановка. Действительно, в одном и том же районе передатчики радиоэлектронных средств создают основные, внеполосные и побочные излучения, а приемные устройства этих средств наряду с основными имеют побочные каналы приема. Если основной или побочный канал приема случайно совпадает с основным или побочным каналом излучения и интенсивность излучаемого колебания достаточно велика, может иметь место взаимная помеха, в частности маскирующая. Так, например, частотно-модулированные и амплитудно-модулированные непрерывные колебания линий связи могут создать маскирующую помеху импульсным радиолокационным приемникам.

Совокупность мер, направленных на исключение взаимных помех, обеспечивает *электромагнитную совместимость*. Наряду с правильным распределением частот и другими организационными мероприятиями электромагнитная совместимость достигается за счет фильтрации побочных излучений в передающих устройствах, гетеродинных колебаний в приемных трактах, за счет правильного использования условий распространения, особенностей местности, выбора режимов работы радиолокационных средств [31, 35].

7.7.2. Искусственные маскирующие активные помехи, особенности их воздействия и способы создания

В качестве искусственных маскирующих активных помех может быть использовано излучение шумовых колебаний. Такие помехи называют *прямошумовыми*, они в наибольшей степени приближаются к нормальному шуму. Прямошумовые помехи могут создаваться двумя способами. В первом — шумовые колебания формируются генератором СВЧ непосредственно в рабочем диапазоне частот РЛС и после усиления по мощности излучаются в пространство. В качестве первичных источников шума на СВЧ используют газоразрядные лампы, специальные магнетроны, работающие в шумящем режиме. Второй способ создания прямошумовой помехи заключается в формировании шумовых колебаний на низкой частоте, переносе средней частоты этих колебаний в заданную область высоких (сверхвысоких) частот с помощью гетеродинирования, для последующего усиления по мощности и излучения в пространство. В качестве первичных источников шума на низкой частоте используют диоды прямого накала, тиратроны в магнитном поле, фотоэлектронные умножители [31, 86].

При достаточно большом динамическом диапазоне приемника шумовые колебания создают эффект, аналогичный резкому увеличению внутреннего шума, что затрудняет обнаружение и измерение параметров радиолокационного сигнала при больших дальностях до цели. Очень мощные искусственные активные помехи, как и взаимные, могут действовать, в принципе, и по побочным каналам приема.

Если динамический диапазон приемника недостаточен и имеет место амплитудное ограничение (например, в последних каскадах УПЧ, после сужения полосы пропускания), отношение сигнал—помеха после ограничителя еще более ухудшается. Это поясняется на рис. 7.30, где показано прохождение через ограничитель немодулированной импульсной помехи (огигающие помехи обозначены цифрами 1, 2 и 3) вместе с импульсом сигнала. Видно, что по мере увеличения интенсивности помехи может произойти полное подавление сигнала. Аналогичный эффект имеет место и в случае воздействия шумовой помехи при недостаточном динамическом диапазоне приемника. Поэтому воздействие маскирующей помехи при малом динамическом диапазоне приемника особенно опасно. Но даже и при очень большом динамическом диапазоне приемника воздействие помехи, эквивалентное увеличению внутреннего шума, может значительно ухудшить или полностью исключить радиолокационное обнаружение или сопровождение.

Для упрощения аппаратуры создания помех наряду с генерацией шума используют генерацию колебаний, модулированных шумом по амплитуде или частоте. Так, например, модулированное по частоте колебание

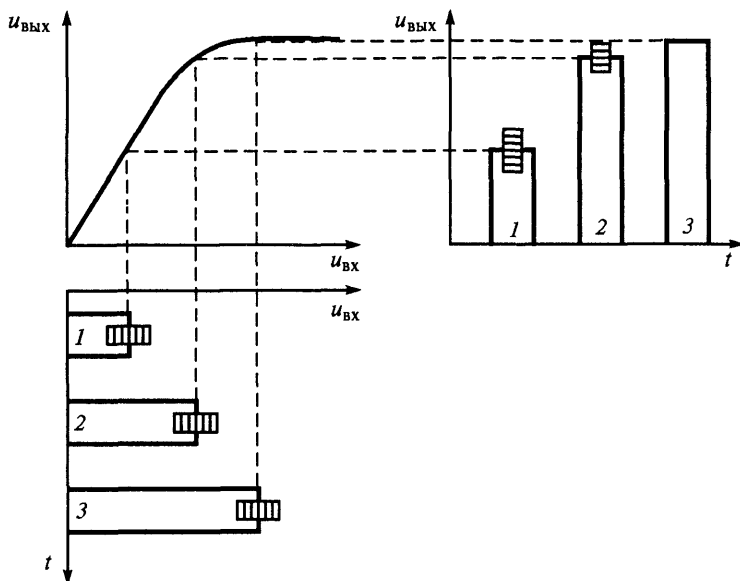


Рис. 7.30. Влияние слабой (1) и сильной (2, 3) помехи на прохождение сигнала

$$u(t) = U \cos \left(\omega_0 t + \int_0^t \Delta \omega(t) dt \right), \quad (7.40)$$

где $\Delta \omega(t)$ — случайная модулирующая функция, будет представлять собой частотно-модулированную шумом помеху, если изменение частоты $\Delta \omega(t)$ происходит в соответствии с некоторым шумовым колебанием.

К числу электронных приборов, позволяющих сравнительно просто осуществить частотную модуляцию, относятся лампы обратной волны. Лампы обратной волны допускают частотную девиацию от единиц до сотен мегагерц, что позволяет создавать помехи в сравнительно широком и узком спектрах частот, называемые соответственно *заградительными* и *прицельными*.

Модулированная шумом помеха не полностью эквивалентна внутреннему шуму приемника. У шума с равномерным спектром практически не коррелированы мгновенные значения напряжения через интервал порядка $1/\Delta f_{\text{ш}}$, где $\Delta f_{\text{ш}}$ — ширина энергетического спектра шума. Для модулированной шумом помехи некоррелированными будут значения, разделенные интервалом $1/\Delta F_{\text{мод}}$, где $\Delta F_{\text{мод}}$ — ширина спектра модулирующих колебаний. Тем не менее, при воздействии на приемник с полосой, меньшей $\Delta F_{\text{мод}}$ и де-

виации частоты, помеха, модулированная по частоте шумом, создает практически такой же эффект, как и шумовая помеха. В этом случае за время переходного процесса в узкополосном приемнике накладывается ряд независимых воздействий, соответствующих попаданию мгновенной частоты в полосу пропускания приемника. Статистика суммы налагающихся мгновенных значений в каждый момент времени приближается к гауссовой, закон распределения амплитуды становится рэлеевским, корреляционные связи оказываются такими же, как и для шума, прошедшего через узкополосную колебательную систему.

Как шумовые, так и модулируемые шумом передатчики помех могут дополнительно перестраиваться (скользить по частоте). Получаемая при этом помеха называется *скользящей*. Скользящая помеха является нестационарным случайным процессом. Однако даже и нескользящая помеха в реальных условиях воздействия также не полностью сводится к стационарному процессу. Например, характеристика направленности обзорного радиолокатора модулирует не только сигнал, но и помеху. Однако в ряде случаев нестационарность помехи не имеет решающего значения. Именно для таких случаев ниже будет приведен анализ условий обнаружения на фоне активных маскирующих помех.

7.7.3. Уравнение радиолокации, дальность действия и зоны видимости РЛС при воздействии маскирующих стационарных активных помех

При достаточном динамическом диапазоне приемника условие обнаружения цели в маскирующих стационарных активных помехах типа белого шума имеет вид

$$E_{\text{пр}} \geq v(N_0 + N_{\text{п.вх}}), \quad (7.41)$$

где $E_{\text{пр}}$ — энергия отраженного от цели сигнала, принимаемого на входе приемника РЛС; v — коэффициент различимости при заданных показателях эффективности обнаружения или измерения; $N_0 = kTK_{\text{ш}}$ — спектральная плотность внутреннего шума приемника, k — постоянная Больцмана, равная $1,38 \cdot 10^{-23}$ Вт · с/град, T — абсолютная температура (обычно $T = 290$ К), $K_{\text{ш}}$ — коэффициент шума приемника; $N_{\text{п.вх}}$ — спектральная плотность маскирующей помехи на входе приемника.

Если на вход приемника воздействуют колебания от нескольких источников (постановщиков) активных помех ($i = 1, 2, \dots, m$), то

$$N_{\text{п.вх}} = \sum_{i=1}^m \frac{P_{\text{п}i} G_{\text{п}i} A_i}{4\pi r_{\text{п}i}^2 \Delta f_{\text{п}i}} \gamma_i \alpha_i. \quad (7.42)$$

Здесь P_{ni} , Δf_{ni} — мощность шума и ширина его энергетического спектра; G_{ni} — значение коэффициента усиления антенны в направлении на РЛС для i -го источника помех; A_i — эффективная площадь приемной антенны в направлении на i -й источник помех; r_{ni} — расстояние от РЛС до i -го источника; γ_i — коэффициент, учитывающий различие поляризации помехи, приходящей от i -го источника помех, и поляризации, оптимальной для приемной антенны (принимает значения от 1 до 0); α_i — коэффициент, учитывающий возможное ухудшение качества помехи от i -го источника за счет использования модуляции шумом (для шумовой помехи $\alpha = 1$).

Если предельную дальность действия радиолокатора в помехах, как и ранее, обозначить r_{\max} , то, заменяя неравенство (7.41) равенством и учитывая формулу (7.42), приходим к уравнению радиолокации в присутствии помех

$$\frac{EG\sigma_{\text{ц}}A}{(4\pi r_{\max}^2)^2} = v \left(N_0 + \sum_{i=1}^m \frac{1}{4\pi} \frac{P_{ni}G_{ni}A'_i}{r_{ni}^2\Delta f_{ni}} \gamma_i \alpha_i \right). \quad (7.43)$$

Полученное уравнение иногда называют уравнением *противорадиолокации*. Используя уравнение противорадиолокации, различают следующие режимы прикрытия сигнала от цели помехой: самоприкрытие, внешнее и коллективное прикрытие.

Самоприкрытием называют случай, когда цель прикрывается передатчиком помех, размещенным на ее борту. *Внешним прикрытием* называют случай, когда нешумящая цель прикрывается постановщиками помех, *коллективным* — когда помехи ставятся и с борта цели, и с других постановщиков [31, 88].

В случае самоприкрытия $m = 1$, $r_{np} = r_{\max}$, $A'_1 = A$. Обозначая отношение максимальных дальностей действия при наличии и отсутствии помех $r_{\max}/r_{0\max} = y$, выражение (7.43) сводим к биквадратному уравнению

$$y^4 + y^2/a^2 = 1, \quad (7.44)$$

где a — отношение дальности действия в помехах r'_{\max} к дальности действия без помех при условии, что дальность действия РЛС на фоне маскирующих помех определяется в пренебрежении внутренним шумом приемника. Для последнего случая из уравнения (7.43) получаем

$$r'_{\max} = \sqrt{\frac{EG\sigma_{\text{ц}}\Delta f_{\text{п}}}{4\pi P_{\text{п}}G_{\text{п}}v\gamma\alpha}}. \quad (7.45)$$

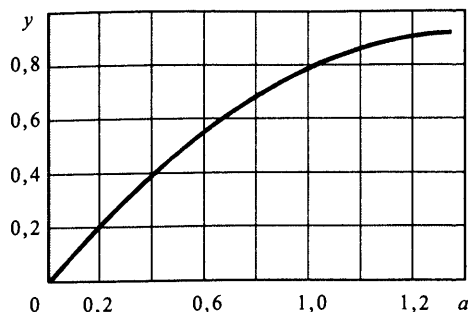


Рис. 7.31. К расчету дальности действия при самоприкрытии по точной формуле

большинстве случаев учет поправки не требуется, так как $a \ll 1$, $y = a$ и $r_{\max} = r'_{\max}$.

Для режима внешнего прикрытия характерно отличие значений $A'_i(\beta, \epsilon)$ от значения A . Наиболее заметно это в том случае, когда помеха действует по боковым лепесткам диаграммы направленности. Если антенна совершает обзор, например, по азимуту, то

$$A(\beta, \epsilon) = A_{\max} F^2(\beta - \beta_0, \epsilon),$$

где β_0 — направление оси диаграммы направленности. Пусть цель находится на азимуте β_0 , а источник помехи — на азимуте β_i ; тогда соответствующие значения эффективной площади антенны будут $A = A_{\max} F^2(0, \epsilon)$ и $A'_i = A_{\max} F^2(\beta_i - \beta_0, \epsilon)$. Подставляя эти выражения в (7.43), можно определить дальность действия РЛС для каждого азимута при фиксированном угле места цели, а значит, и установить *зону видимости* РЛС для данного угла ϵ .

Примерное изображение зоны видимости в случае двух постановщиков помех показан на рис. 7.32 (кривая 2). Наблюдается сокращение дальности действия по сравнению со случаем отсутствия помех (кривая 1), даже когда помеха действует по боковым лепесткам диаграммы направленности. Наибольшее сокращение дальности действия имеет место тогда, когда источник

Решение биквадратного уравнения (7.44) определяется выражением

$$y = \sqrt{\frac{1}{2} \left(\sqrt{\frac{1}{a^4} + 4} - \frac{1}{a^2} \right)}$$

и графиком, изображенным на рис. 7.31, который позволяет найти поправку к решению уравнения (7.44), полученному в пренебрежении внутренним шумом. Однако в

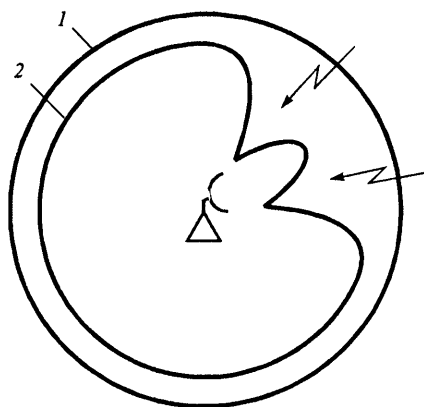


Рис. 7.32. Зоны видимости РЛС в отсутствие (1) и при воздействии помех с двух направлений (2) для фиксированного угла места ϵ

помех воздействует по главному лепестку диаграммы направленности антенны РЛС. В окрестностях направления главного лепестка диаграммы направленности антенны на источник помех могут создаваться *секторы эффективного подавления*. Степень уменьшения дальности в каждом случае так же, как и ширина сектора подавления, зависит от параметров радиолокатора и спектральной плотности мощности помехи на входе приемника РЛС.

Если аналогичным образом исследовать зависимость дальности

от угла ϵ , то можно убедиться в том, что не только уменьшается дальность действия, но и снижается потолок обнаружения и поднимается нижняя кромка зоны видимости РЛС.

Полученные выражения и выводы справедливы только для РЛС с необходимым динамическим диапазоном приемного тракта. Если это не соблюдается, то наряду с потерей возможности обнаруживать цель на больших дальностях, будет теряться возможность обнаружения цели и на малых дальностях, поскольку уровень помехи может превышать уровень ограничения в тракте приемника. Для случая, изображенного на рис. 7.30, при недостаточном динамическом диапазоне приемника на индикаторе будет наблюдаться картина, показанная на рис. 7.33. Для увеличения динамического диапазона приемных трактов используют системы автоматической регулировки усиления и усилители с линейно-логарифмическими характеристиками [35, 68].

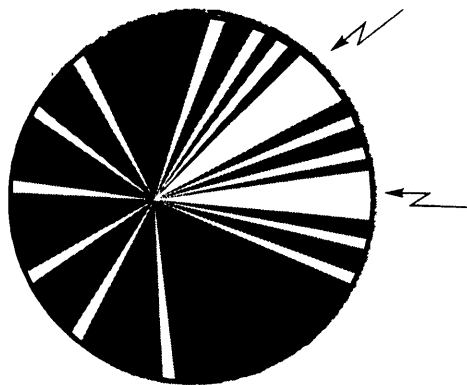


Рис. 7.33. Вид экрана индикатора кругового обзора при воздействии помех с двух направлений

7.7.4. Пассивные маскирующие помехи и способы их создания

Как уже указывалось выше, к естественным пассивным помехам относятся радиопомехи, создаваемые природными отражателями (местными предметами, водной поверхностью, гидрометеорами, северными сияниями и т. д.). Эти помехи могут существенно нарушать работу аэродромных радиолокаторов, обеспечивающих посадку самолетов, и радиолокаторов военного назначения, используемых для обнаружения целей, особенно на малых высотах [31, 35, 86—88].

Наибольшее распространение из искусственных маскирующих пассивных помех получили помехи, создаваемые дипольными противорадио-

локационными отражателями. Они представляют собой пассивные полуволновые вибраторы, изготовленные из металлизированных бумажных лент, фольги или металлизированного стеклянного или капронового волокна. Длина узкополосных резонансных вибраторов выбирается примерно равной половине длины волны подавляемой РЛС. Ширина лент в зависимости от их длины может быть в пределах от нескольких миллиметров до нескольких сантиметров, а диаметр волокна — от десятков до сотен микрон при толщине металлического покрытия порядка единиц микрон.

Обычно дипольные отражатели собираются в пачки таким образом, чтобы каждая пачка по своим отражающим свойствам имитировала реальную цель ($\sigma_{п.ср} = \sigma_{ц.ср}$). Число отражателей n в пачке зависит от диапазона длин волн, в котором работает подавляемая РЛС, и может быть определено по известной формуле:

$$n = \sigma_{п.ср} / (0,17\lambda^2). \quad (7.46)$$

Основной недостаток таких пачек — узкий диапазон перекрываемых частот (5...10% от резонансной). Полоса частот расширяется, если пачки комплектовать из вибраторов различной длины или увеличивать длину и поперечные размеры диполей. Пачки помещают между специальными лентами, которые наматываются на барабаны, расположенные в кассетах. Ими могут снаряжаться противорадиолокационные патроны. Возможно также создание пассивных помех с нарезкой дипольных отражателей на борту самолета в зависимости от разведанного диапазона частот подавляемой РЛС.

Для маскировки воздушных целей дипольные отражатели сбрасываются в окружающее пространство при помощи автоматов или бомб (в заднюю полусферу) или выстреливаются при помощи пушек и ракет (в переднюю и заднюю полусферы). При этом могут создаваться как сплошные полосы (облака) пассивных отражателей, так и разрывные.

Облако отражателей характеризуется своей плотностью. Плотность пассивных отражателей, определяемая количеством пачек на единицу пути, при полете на или от РЛС находится по формуле

$$m_1 = \frac{z_a}{v_{ц} t_{сбр}}, \quad (7.47)$$

где z_a — число автоматов сбрасывания отражателей; $v_{ц}$ — скорость поставщика помех; $t_{сбр}$ — время между сбросами пачек диполей.

Зная плотность диполей, можно вычислить количество пачек в одном импульсном объеме

$$m_{\Delta r} = m_1 \Delta r, \quad (7.48)$$

где Δr — разрешающая способность по дальности подавляемой РЛС.

Если пассивными помехами маскируется объем V пространства, в котором строй разнесенных по азимуту и углу места самолетов проходит расстояние R (рис. 7.34), то количество пачек, сбрасываемых в этом объеме, определяется выражением

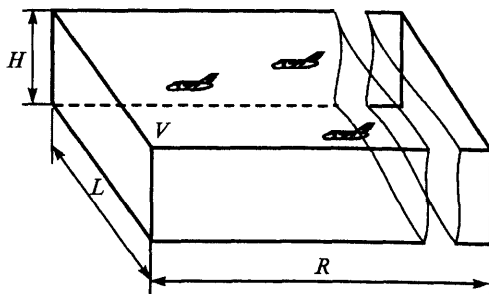


Рис. 7.34. Пояснение к расчету количества пачек в объеме V

$$m_{R,L,H} = m_{\Delta r} \frac{R}{\Delta r} \frac{L}{r\Delta\beta} \frac{H}{r\Delta\varepsilon}, \quad (7.49)$$

где R, L, H — размеры маскируемого объема V ; $\Delta r, r\Delta\beta, r\Delta\varepsilon$ — размеры импульсного объема РЛС на расстоянии r от станции.

Для маскировки целей необходимо, чтобы средняя эффективная площадь диполей, занимающих импульсный объем РЛС, превышала среднюю эффективную площадь целей, находящихся в этом объеме. Пренебрегая взаимной экранировкой дипольных отражателей, а также их деформацией и разрушением при сбрасывании, можно определить среднюю эффективную поверхность облака диполей, занимающих импульсный объем, по формуле

$$\sigma_{\text{п.ср}} \approx 0,17\lambda^2 n m_{\Delta r} = \sigma_{\text{п.ср}} m_{\Delta r}.$$

Условие подавления РЛС, не защищенной от пассивных помех, можно записать в виде $\sigma_{\text{п.ср}} > \sigma_{\text{ц.ср}}/\nu$, а для защищенной РЛС — в виде

$$\sigma_{\text{п.ср}} > \sigma_{\text{ц.ср}} K_{\text{пв}}/\nu, \quad (7.50)$$

где $K_{\text{пв}}$ — коэффициент подпомеховой видимости аппаратуры защиты.

7.8. Методы защиты от маскирующих активных помех

Методы защиты от маскирующих активных помех основаны на использовании различий статистических характеристик сигнала и помехи, которые для гауссовских статистик сводятся к различиям соответствующих корреляционных матриц. В свою очередь, корреляционные матрицы могут различаться дисперсиями и коэффициентами корреляции, обусловленными амплитудными, поляризационными, частотными и пространственными раз-

личиями сигналов и помех. Меры защиты могут быть достаточно эффективными только в том случае, если не происходит подавления сигнала за счет недостаточного динамического диапазона приемника. В этом случае возможно использование амплитудной, поляризационной, частотной, пространственной селекции [31, 35, 43, 74, 75, 86—88].

7.8.1. Основные направления защиты РЛС от маскирующих активных помех

Анализ уравнения противорадиолокации также показывает, что основные направления защиты РЛС от маскирующих активных помех связаны с использованием амплитудных, поляризационных, частотных и пространственных различий между принимаемыми сигналами и помехами. Действительно, как видно из равенства (7.43), для повышения дальности действия r_{\max} РЛС в помехах необходимо увеличивать числитель левой части или уменьшать правую часть этого равенства. Например, повышение энергии зондирующего сигнала позволяет увеличить дальность действия в помехах пропорционально $\sqrt[4]{E}$ в режиме внешнего прикрытия и \sqrt{E} в режиме самоприкрытия. Увеличение коэффициента усиления передающей антенны в направлении на цель позволяет увеличить дальность действия в помехах также пропорционально $\sqrt[4]{G}$ в режиме внешнего прикрытия и \sqrt{G} в режиме самоприкрытия. В отдельных случаях уменьшение поляризационного коэффициента γ может снизить воздействие помехи по сравнению с воздействием сигнала. В целом, дальность действия в режиме самоприкрытия оказывается обратно пропорциональной $\sqrt{\gamma}$ и $\sqrt{\nu}$. Уменьшение относительного уровня боковых лепестков диаграммы направленности A'/A (или даже образование провалов в главном лепестке в направлении на источники помех) позволяет увеличить дальность действия в режиме внешнего прикрытия пропорционально $\sqrt[4]{A'/A}$. Рассмотрим более подробно практическую реализацию перечисленных выше основных направлений защиты РЛС от активных маскирующих помех.

Увеличение энергии зондирующего сигнала может осуществляться путем повышения мощности, увеличения длительности сигнала и числа импульсов в пачке за счет соответствующего увеличения времени облучения и (или) частоты следования импульсов. Энергия зондирующего сигнала будет рационально использоваться при приеме только в случае приближения обработки принимаемых колебаний к оптимальной, иначе возрастает величина коэффициента ν в правой части равенства (7.43).

Увеличение коэффициента усиления антенны в направлении на цель может в то же время соответственно замедлить обзор пространства. В настоящее время используют и развивают методы управляемого обзора с последовательным анализом, когда время, в течение которого антенна направлена на цель, зависит от условий обнаружения и, в частности, от помеховой обстановки. Особенно широкие возможности для использования программного автоматически управляемого обзора открываются при применении передающих антенн с электронным управлением луча в виде фазированных решеток.

Известно, что приемная антенна обычно настроена на определенную поляризацию принимаемого сигнала: линейную, круговую либо (в общем случае) эллиптическую. Возможны антенны с регулируемой поляризацией. Если поляризация антенны соответствует поляризации помехи, эффект воздействия помехи будет наибольшим. Например, для вертикальной поляризации помехи воздействие будет наибольшим, если прием ведется на вертикальный вибратор; для круговой поляризации с вращением вектора поля по часовой стрелке эффект воздействия будет наибольшим, если антенна рассчитана на этот же вид поляризации. С учетом этого антенну можно тем или иным способом перестроить на ортогональную поляризацию, т. е. для приведенных примеров — на горизонтальную либо на круговую поляризацию, но с вращением против часовой стрелки. Для эллиптически поляризованной волны ортогональным является также эллиптически поляризованное колебание, но со сдвинутым на 90° положением эллипса поляризации. Во всех указанных случаях можно достичь существенного ослабления помехи. Будет ли при этом происходить ослабление полезного сигнала, зависит от поляризации колебаний сигнала. Если поляризация колебаний полезного сигнала точно совпадает с поляризацией колебаний помехи, одновременно с помехой и в той же мере будет ослаблен сигнал. Поскольку поляризация сигналов, отраженных от реальных целей, случайная и в общем случае не совпадает с поляризацией помехи, имеются принципиальные возможности ослабить помеху в большей степени, чем сигнал.

Для повышения помехозащищенности целесообразно *уменьшать коэффициент различимости ν* . Уменьшение коэффициента ν достигается за счет приближения приема к оптимальному. Если помехой является стационарный белый шум, то уменьшение ν достигается оптимизацией приема для таких помех (см. гл. 3). При фильтровом приеме, в частности, используют оптимальную частотную характеристику, т. е. осуществляют оптимальную частотную селекцию. Частотная селекция тем более эффективна, чем шире спектр помехи по сравнению со спектром сигнала, так как спектральная плотность мощности *заградительной* по частоте помехи при заданной мощности передатчика помех снижается обратно пропорционально ширине по-

лосы частот помехи. *Прицельные* помехи (с меньшей полосой частот), как правило, более эффективны, но их труднее реализовать. Создание прицельных помех затрудняется при повышении скрытности РЛС, например при использовании быстрой перестройки частоты радиолокатора, при многочастотном или широкополосном зондирующем сигнале и т. д. Если полоса частот помехи заметно уже ширины спектра принимаемого сигнала, то результирующий шум нельзя считать белым. В этом случае оптимальной является частотная характеристика с подавлением в полосе частот помехи (см. § 3.13); иначе говоря, целесообразно использование различного рода настраиваемых *режекторных* фильтров для колебаний помехи, приводящее к существенному уменьшению коэффициента различимости ν .

Снижение уровня боковых лепестков диаграммы направленности может существенно ослабить влияние помех и представляет собой самостоятельную задачу, особенно важную в случае внешнего прикрытия. Из теории антенн известно, что снижение уровня боковых лепестков может быть достигнуто увеличением размеров антенны, рациональным распределением поля в раскрыве, повышением точности изготовления, снижением влияния переотражений от близлежащих объектов. Повышение избирательности антенны позволяет улучшить пространственную селекцию принимаемых колебаний.

7.8.2. Методы некогерентной и когерентной компенсации помех

Для улучшения *пространственной* селекции сигнала на фоне помех, приходящих с отдельных направлений, кроме мер, перечисленных выше, могут быть также использованы методы *некогерентной* и *когерентной компенсации помеховых колебаний* [31, 35, 40, 43, 73, 74, 103]. Для этого наряду с основной могут быть задействованы дополнительные антенны (в фазированной антенной решетке — ее отдельные элементы). Возможности компенсации помех были сформулированы академиком Н.Д. Папалекси еще несколько десятилетий тому назад [31].

Если сигналы, принимаемые дополнительной антенной и боковыми лепестками основной антенны, компенсируются после детектора, следует говорить о *некогерентной компенсации*. Если такая компенсация производится на высокой (промежуточной) частоте, ее можно называть *когерентной*. Для обеспечения как когерентной, так и некогерентной компенсации помеховые колебания должны быть *пространственно когерентны* (коррелированы).

На рис. 7.35 схематически показана система, включающая основную и две дополнительные антенны. Каждой антенне соответствует свой канал приема. Колебания, прошедшие через соответствующие каналы приема, по-

даются на сумматор. При этом по крайней мере в двух дополнительных каналах по амплитуде и фазе регулируются комплексные коэффициенты передачи K_1 и K_2 . Если комплексные характеристики направленности каналов имеют вид $F_0(\theta)$, $F_1(\theta)$, $F_2(\theta)$, то суммарную комплексную характеристику направленности можно представить в виде

$$F_{\Sigma}(\theta) = F_0(\theta) + K_1 F_1(\theta) + K_2 F_2(\theta). \quad (7.51)$$

Тогда, подбирая коэффициенты K_1 и K_2 , можно достичь образования провалов в результирующей характеристике направленности для угловых координат θ_1 и θ_2 источников помех. Необходимые значения K_1 и K_2 определяются из системы уравнений

$$\begin{aligned} F_0(\theta_1) + K_1 F_1(\theta_1) + K_2 F_2(\theta_1) &= 0, \\ F_0(\theta_2) + K_1 F_1(\theta_2) + K_2 F_2(\theta_2) &= 0. \end{aligned} \quad (7.52)$$

Провалы в характеристике направленности, образуемые благодаря когерентной компенсации помех, создают дополнительный резерв пространственной селекции помех, воздействующих как по главному, так и по боковым лепесткам характеристики направленности.

Подбор коэффициентов в многоканальных схемах, подобных схеме, приведенной на рис. 7.35, можно осуществить, используя *принцип корреляционной обратной связи* [31, 103]. На рис. 7.36, а показана схема с двумя входами, на которые поступают напряжения одной и той же частоты с комплексными амплитудами $U_0(t)$ и $U_1(t)$ (например, от основной и дополнительной антенн, рис. 7.36, б). На сумматоре образуется напряжение

$$U_{\Sigma}(t) = U_0(t) - KU_1(t). \quad (7.53)$$

Имеется цепь обратной связи с выхода сумматора на управляемый элемент — умножитель в цепи подачи первого напряжения. В эту цепь включено состоящее из перемножителя и интегратора устройство вычисления оценки корреляционного момента $\overline{U_{\Sigma} U_1^*}$. Последний с точностью до постоянной γ используется в качестве управляющего множителя K , подаваемого на управляемый элемент. Используя уравнения $K = \gamma \overline{U_{\Sigma} U_1^*}$ и (7.53), можно найти

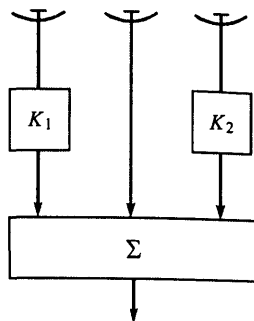


Рис. 7.35. Система с двумя дополнительными приемными каналами для образования провалов в результирующей диаграмме направленности

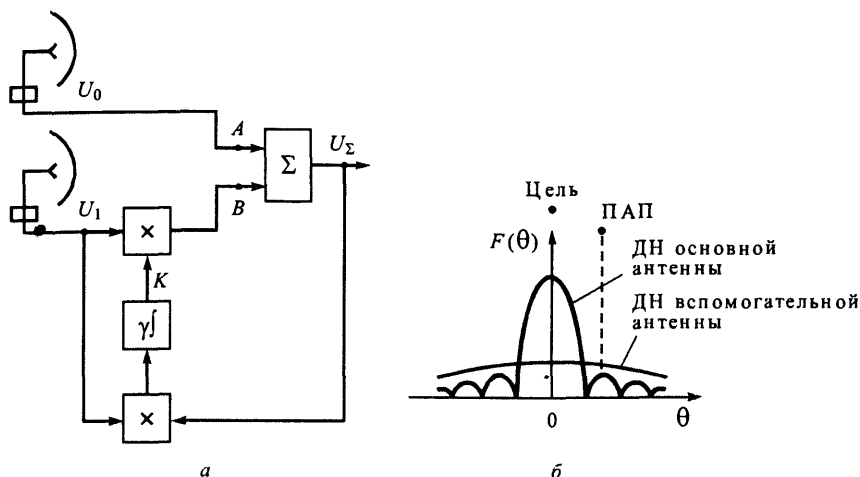


Рис. 7.36. Автоматическая компенсация помех в устройстве с корреляционной обратной связью:

а — одноканальный автокомпенсатор, *б* — диаграмма направленности антенн: ПАП — постановщик активной помехи, ДН — диаграмма направленности

$$K = \frac{\overline{\gamma U_0 U_1^*}}{1 + \gamma |U_1|^2}, \quad (7.54)$$

$$U_\Sigma = U_0 - \frac{\overline{\gamma U_0 U_1^*}}{1 + \gamma |U_1|^2} U_1. \quad (7.55)$$

Легко видеть, что при $\gamma \rightarrow \infty$ и достаточной корреляции U_0 и U_1 (например, при $U_1 = CU_0$, где $C = \text{const}$) происходит полная компенсация пространственно коррелированных помех, т. е. U_Σ обращается в нуль. Умножение комплексных амплитуд, как известно, может быть осуществлено, например, путем преобразования частоты, усреднение — за счет интегрирования в узкополосном фильтре. Те же операции могут быть произведены путем аналоговой квадратурной обработки или цифровыми методами.

Рассмотренный *одноканальный автокомпенсатор* (АК) имеет один регулируемый компенсационный (вспомогательный, дополнительный) канал (отсюда название *одноканальный*) и один нерегулируемый основной канал.

В [34, 74] показано, что при отсутствии полезного сигнала в реализации U весовой коэффициент K можно оценить следующим образом:

$$T \frac{dK}{dt} + K = -\gamma U_{\Sigma} U_1^* / 2, \quad (7.56)$$

где T — постоянная времени интегратора, а

$$U_{\Sigma} = U_0 + KU_1. \quad (7.57)$$

В установившемся режиме справедливо равенство $\mathbf{M}\{dK/dt\} \approx 0$. Определяя математическое ожидание от обеих частей равенства (7.56), находим установившееся значение весового коэффициента

$$K_{\text{уст}} = -\gamma \mathbf{M}\{U_{\Sigma} U_1^* / 2\}. \quad (7.58)$$

При отсутствии корреляции помех (при наличии только внутренних шумов) $K_{\text{уст}} = 0$, т. е. компенсационный канал отключается.

Отметим, что структурная схема автокомпенсатора, приведенная на рис. 7.36, а, носит обобщенный характер, поскольку в ней не конкретизируется техническая реализация элементов перемножения и вычисления комплексных функций. Рассмотрим поэтому практические схемные решения автокомпенсаторов, получивших широкое распространение в радиолокационной технике.

7.8.3. Практические схемы автокомпенсаторов

Квадратурный автокомпенсатор [31, 40, 74, 103]. В таком автокомпенсаторе формирование весового (управляемого) напряжения осуществляется на видеочастоте. В этой связи представим комплексную амплитуду весового коэффициента K через реальную k_c и мнимую k_s части (рис. 7.37):

$$K = k_c - jk_s.$$

Подставляя последнее выражение в (7.57), получим

$$U_{\Sigma} = U_0 + k_c U_1 + k_s (-jU_1)$$

или

$$U_{\Sigma} = U_0 + k_c U_1 + k_s U_{1\perp}. \quad (7.59)$$

Здесь $U_{1\perp} = -jU_1$ — квадратурная составляющая напряжения U_1 , получаемая на выходе фазовращателя на 90° (так как $e^{-j\pi/2} = -j$).

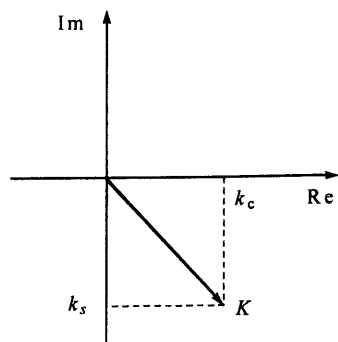


Рис. 7.37. Представление весового коэффициента

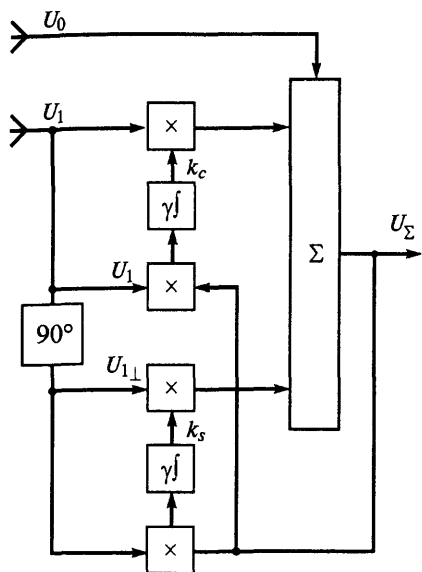


Рис. 7.38. Квадратурный автокомпенсатор

имеет два регулируемых квадратурных канала (подканала) с вещественными передаточными коэффициентами k_c и k_s и называется *квадратурным*. В каждом подканале имеется балансный усилитель с управляемым и инвертирующим коэффициентом передачи (перемножитель) и коррелятор (сочетание перемножителя и интегратора), выходное напряжение которого регулирует коэффициент усиления. С помощью фазовращателя достигается фазовый сдвиг на 90° между напряжениями в квадратурных подканалах.

Принцип компенсации помех можно пояснить с помощью векторных диаграмм, показанных на рис. 7.39, а, б (модуль коэффициента корреляции помехи полагается равным 1). Вектором U_0 обозначено напряжение помехи в основном канале, а U_1 — в компенсационном канале. Фазовый сдвиг между напряжениями основного и компенсационного каналов — это угол φ_n между векторами U_0 и U_1 . Этот сдвиг может образоваться за счет разности фазовых центров основной и компенсационных антенн, различия электрических длин основного и вспомогательного каналов и т. п. Напряжение компенсационного канала представляется двумя составляющими U_1 и $U_{1\perp}$ (см. рис. 7.39, а). На выходах регулируемых усилителей формируются компенсирующие напряжения $k_c U_1$ и $k_s U_{1\perp}$ (см. рис. 7.39, б). Результирующий вектор суммарного напряжения компенсационных подканалов $k_c U_1 + k_s U_{1\perp}$ в установившемся режиме и при полной корреляции помехи оказывается рав-

Мгновенное значение суммарного выходного напряжения определяется соотношением

$$u_\Sigma(t) = u_c(t) + k_c u_1(t) + k_s u_{1\perp}(t). \quad (7.60)$$

В свою очередь, дифференциальные уравнения для квадратурных составляющих k_c и k_s определяются в виде

$$\begin{aligned} T \frac{dk_c(t)}{dt} + k_c(t) &= -\gamma u_1(t) u_\Sigma(t), \\ T \frac{dk_s(t)}{dt} + k_s(t) &= -\gamma u_{1\perp}(t) u_\Sigma(t). \end{aligned} \quad (7.61)$$

Структурная схема одноканального автокомпенсатора (АК), реализующего уравнения (7.59)—(7.61), представлена на рис. 7.38. Данный автокомпенсатор

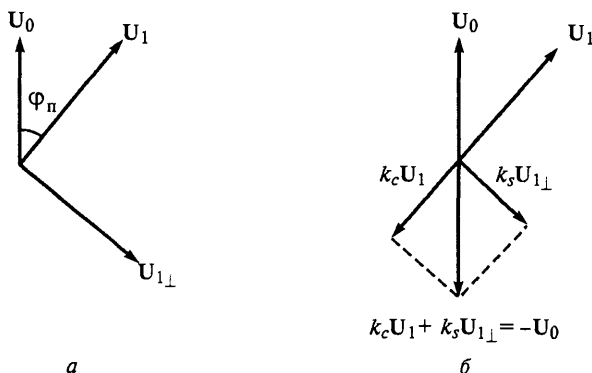


Рис. 7.39. Пояснение принципа компенсации помех в одноканальном квадратурном автокомпенсаторе:

a — напряжение помех на входах основного U_0 и квадратурных U_1 , $U_{1\perp}$ каналов, *б* — напряжение помех на выходах основного U_0 и квадратурных $k_c U_1$, $k_s U_{1\perp}$ каналов

ным (по длине) и противофазным поступающему на вход сумматора вектору напряжения основного канала. Выходное напряжение АК U_Σ полностью декоррелируется при этом с напряжением компенсирующего канала U_1 , т. е.

$$\mathbf{M}\{U_\Sigma U_1^*\} \Big|_{K=K_{\text{опт}}} = 0. \quad (7.62)$$

Декорреляция выходного и компенсирующего напряжений, обеспечиваемая за счет корреляционной обратной связи, оказывается признаком минимизации дисперсии помехи. Значение комплексного коэффициента передачи K , при котором достигается минимум дисперсии помехи σ_Σ на выходе АК в установившемся режиме, находится из условия $\partial \sigma_\Sigma^2 / \partial K = 0$ ($\partial \sigma_\Sigma^2 / \partial k_c$, $\partial \sigma_\Sigma^2 / \partial k_s$) при $K = K_{\text{опт}}$ и определяется выражением

$$K_{\text{опт}} = -r \frac{\sigma_0}{\sigma_1}, \quad (7.63)$$

где $r = \frac{\mathbf{M}\{U_0 U_1^* / 2\}}{\sigma_0 \sigma_1}$ — коэффициент корреляции комплексных амплитуд компенсируемого и компенсирующего напряжений (на входах основного и компенсационных каналов); σ_0 , σ_1 — среднеквадратические значения помехи на входах основного и компенсационного каналов.

В свою очередь (для рассматриваемого примера), значение коэффициента передачи K в установившемся режиме, как отмечалось выше, равно

$$K_{уст} = -\gamma \mathbf{M}\{U_{\Sigma} U_1^* / 2\}. \quad (7.64)$$

Подставляя значение $U_{\Sigma} = U_0 + KU_1$ в (7.55), при сильной обратной связи ($\gamma \gg 1/\sigma_1^2$) после преобразований получим

$$K_{уст} = K_{опт} = -r \frac{\sigma_0}{\sigma_1} = -(r_c - jr_s) \frac{\sigma_0}{\sigma_1}. \quad (7.65)$$

Тогда значения k_c и k_s в установившемся режиме можно найти соответственно из формул

$$k_c = -r_c \frac{\sigma_0}{\sigma_1}, \quad k_s = -r_s \frac{\sigma_0}{\sigma_1}. \quad (7.66)$$

При точном соблюдении квадратурности обработки принимаемых колебаний шумы компенсационных каналов некоррелированы и подканалы настраиваются независимо.

Результирующую дисперсию помехи на выходе автокомпенсатора σ_{Σ}^2 определим следующим образом:

$$\begin{aligned} \sigma_{\Sigma}^2 &= \frac{1}{2} \mathbf{M}\{|U_{\Sigma}|^2\} = \frac{1}{2} \mathbf{M}\{|U_0 + KU_1|^2\} = \frac{1}{2} \mathbf{M}\{|U_0|^2\} + 2KM\{U_0^* U_1 / 2\} + \\ &+ |K|^2 \frac{1}{2} \mathbf{M}\{|U_1|^2\} = \sigma_0^2 - 2|r|^2 \sigma_0^2 + |r|^2 \sigma_0^2 = \sigma_0^2 (1 - |r|^2). \end{aligned} \quad (7.67)$$

Показателем качества работы автокомпенсатора является коэффициент подавления помехи

$$K_n = \frac{\sigma_0^2}{\sigma_{\Sigma}^2} = \frac{1}{1 - |r|^2}. \quad (7.68)$$

Для эффективной компенсации помех значение модуля коэффициента корреляции $|r|^2$ должно быть близко к единице. Поэтому необходимо обеспечивать высокую идентичность амплитудно- и фазочастотных характеристик приемных трактов и антенных элементов. Решение данной задачи в ряде случаев облегчается при использовании в этих трактах цифровых устройств.

В аналоговом квадратурном компенсаторе при выработке управляющего напряжения операция усреднения по времени (интегрирования)

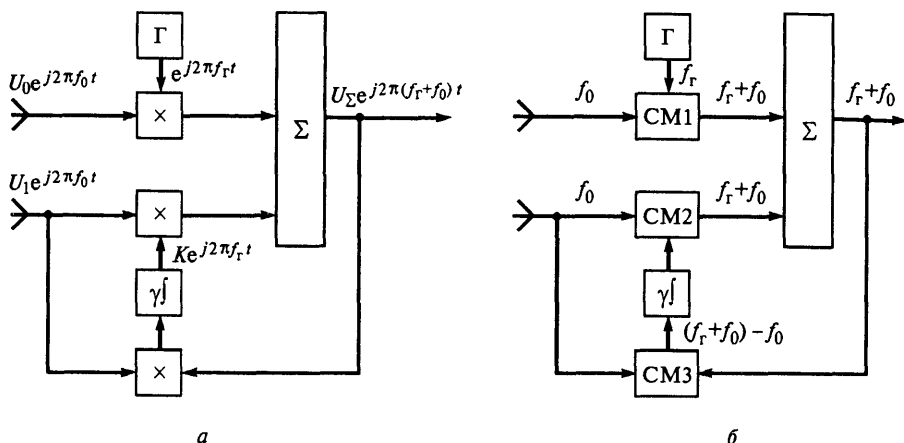


Рис. 7.40. Одноканальный гетеродинный автокомпенсатор:

а — пояснение принципа работы автокомпенсатора, б — структурная схема

осуществляется на постоянном токе. На практике используется также гетеродинный автокомпенсатор, где данная операция осуществляется на переменном токе путем интегрирования (накопления) в узкополосном контуре.

Гетеродинный автокомпенсатор [34, 35, 40, 74]. Структурная схема гетеродинного автокомпенсатора представлена на рис. 7.40, а, б. Алгоритм работы гетеродинного автокомпенсатора можно получить, умножая равенство (7.57) на $e^{j2\pi(f_0+f_r)t}$, а (7.58) — на $e^{j2\pi f_r t}$.

При этом получим следующие соотношения:

$$U_\Sigma e^{j2\pi(f_0+f_r)t} = (U_0 e^{j2\pi f_0 t}) e^{j2\pi f_r t} + (K e^{j2\pi f_r t}) U_1 e^{j2\pi f_0 t}, \quad (7.69)$$

$$K e^{j2\pi f_r t} = -\gamma \mathbf{M} \left\{ U_\Sigma e^{j2\pi(f_0+f_r)t} (U_1 e^{j2\pi f_0 t})^* / 2 \right\}. \quad (7.70)$$

На вход гетеродинного автокомпенсатора (см. рис. 7.40, б) поступают напряжения помехи с комплексными амплитудами U_0 и U_1 на несущей частоте f_0 . Оба напряжения в соответствующих смесителях CM1 и CM2 преобразуются на суммарную несущую частоту $f_0 + f_r$. При преобразовании напряжения основного канала используется гетеродинное напряжение с неуправляемыми, а компенсационного — с управляемыми амплитудой $|K|$ и начальной фазой $\arg K$. В результате такого преобразования обеспечиваются требуемые для компенсации амплитуда и начальная фаза компенсирующего напряжения $K U_1 e^{j2\pi(f_0+f_r)t}$.

Суммарное напряжение частоты $f_0 + f_r$ поступает на выход автокомпенсатора и в цепь корреляционной обратной связи. В ней вырабатывается гетеродинное напряжение $Ke^{j2\pi f_r t}$ согласно (7.70). Для этого выходное напряжение $U_\Sigma e^{j2\pi(f_0 + f_r)t}$ гетеродинируется входным напряжением компенсационного канала $U_1 e^{j2\pi f_0 t}$. На выходе смесителя СМЗ выделяется напряжение разностной частоты $(f_0 + f_r) - f_0 = f_r$, начальная фаза которого определяется разностью начальных фаз перемножаемых колебаний, что учитывается в (7.53) знаком комплексного сопряжения $(U_1 e^{j2\pi f_0 t})^*$.

Таким образом, особенностями обработки в гетеродинном компенсаторе являются:

— введение комплексного весового коэффициента K путем гетеродина-
мирования с соответствующим регулированием в смесителе СМ2;

— вычисление корреляционного момента выходного и компенсационного напряжения путем гетеродинирования как некоторой усредненной комплексной амплитуды колебаний гетеродинной частоты.

При этом, как отмечалось выше, усреднение обеспечивается путем накопления (интегрирования) колебаний узкополосной колебательной системой с большой памятью, например, кварцевым резонатором.

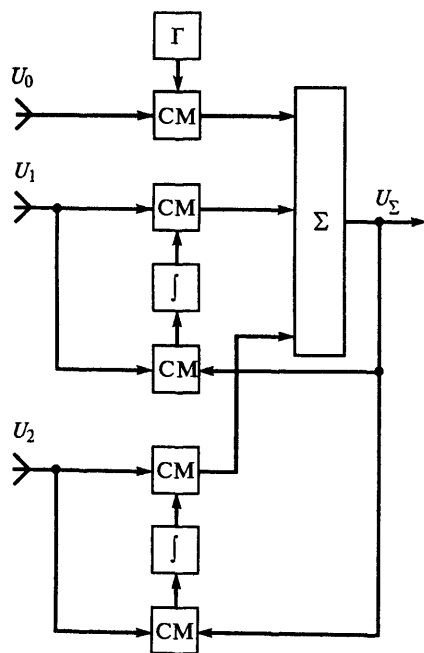


Рис. 7.41. Двухканальный гетеродинный автокомпенсатор

Гетеродинный автокомпенсатор, как и квадратурный, декоррелирует выходное напряжение U_{Σ} по отношению к напряжению вспомогательного канала U_1 . Коэффициент подавления одной помехи одноканальным гетеродинным автокомпенсатором также определяется выражением (7.68).

Заметим, что эффективная компенсация одновременно действующих M помех осуществляется при наличии не менее M независимых вспомогательных каналов приема со своими антеннами. Так, при наличии двух разнесенных в пространстве источников помех необходимо использовать как минимум двухканальный автокомпенсатор. На рис. 7.41 показан пример структурной схемы двухканального гетеродинного автокомпенсатора.

Рассмотрим особенности построения автокомпенсаторов при наличии

мощных протяженных полезных сигналов. Как отмечалось выше, анализ работы автокомпенсаторов проводился для случая полного отсутствия полезного сигнала на их входах. При наличии достаточно мощного и длительного полезного сигнала как на основном, так и на вспомогательном входах АК последний может осуществлять компенсацию и полезного сигнала (даже при наличии угловых различий между целью и постановщиком помехи). Если же отраженные от цели импульсы имеют небольшой уровень и малую длительность по сравнению с временем настройки цепей корреляционной обратной связи (КОС) автокомпенсатора, то цепи адаптации практически не реагируют на полезный сигнал и компенсируется только помеха (при наличии соответствующих угловых различий). В противном случае (когда велики мощность и длительность сигнала) необходимо принимать специальные меры для исключения его из компенсационного канала и устранения влияния этого сигнала на цепи КОС. Эти меры могут быть основаны на использовании *пространственных, временных и частотных различий* сигнала и помехи. Кроме того, могут использоваться алгоритмические методы [40].

При использовании пространственных различий задача решается путем формирования провалов в ДН компенсационных антенн на ожидаемое направление приема полезного сигнала, т. е. в области главного лепестка ДН антенны основного канала (см. рис. 7.42, а). Применительно к антенной

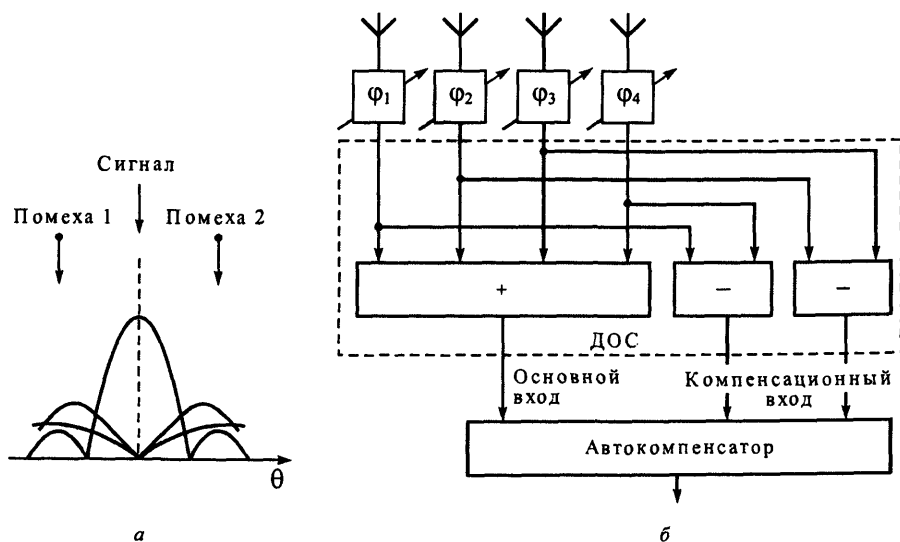


Рис. 7.42. Устранение полезного сигнала из цепей КОС при использовании пространственных различий:

а — диаграммы направленности антенн основного и компенсационных каналов,
б — устройство компенсации помех с фазированной антенной решеткой

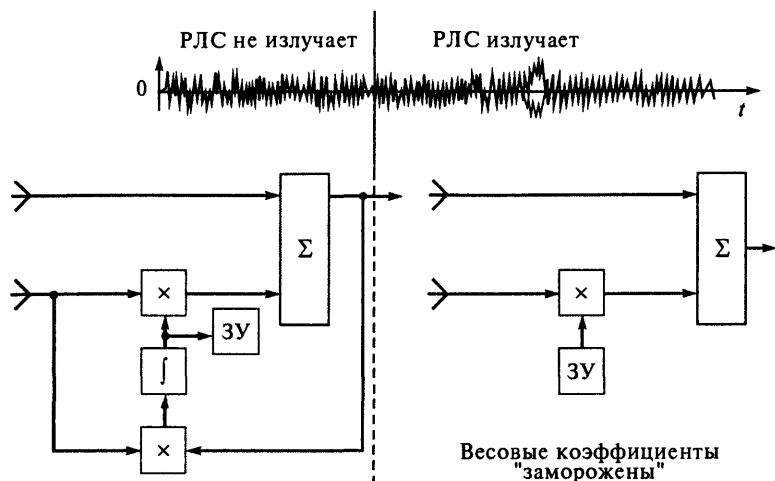


Рис. 7.43. Устранение полезного сигнала из цепей КОС при использовании временных различий

решетке (см. рис. 7.42, б) такие диаграммы можно сформировать, например, путем суммирования (основная ДН) и вычитания (компенсационные ДН) напряжений с выходов соответствующих элементов АР. Диаграммообразующая схема (ДОС) может быть построена на основе, например, преобразований Фурье или Адамара.

При использовании временных различий весовые коэффициенты могут формироваться в моменты отсутствия приема полезных сигналов (например, РЛС не излучает). Полученные весовые коэффициенты «замораживаются», т. е. сохраняются на определенное заданное время, например на период зондирования (рис. 7.43).

Использование частотных различий возможно в тех случаях, когда помеха имеет составляющие частотного спектра $N(f)$, не совпадающие с частотным спектром $G(f)$ полезного сигнала. Эти составляющие (с помощью соответствующих фильтров) могут быть выделены и использованы для управления цепями адаптации (рис. 7.44).

Однако описанные выше специальные меры не могут устранить компенсацию сигнала наряду с компенсацией помехи при отсутствии пространственных различий между целью и источником помехи. Создаваемый с помощью описанного выше автокомпенсатора (квадратурного или гетеродинного) провал ДН в направлении на источник помехи приводит к исключению приема с этого направления полезного сигнала. Поэтому при отсутствии угловых различий между целью и источниками помех для выделения сигнала на фоне помех необходимо использовать различия по другим параметрам, например по поляризации или по доплеровской частоте [33, 35].

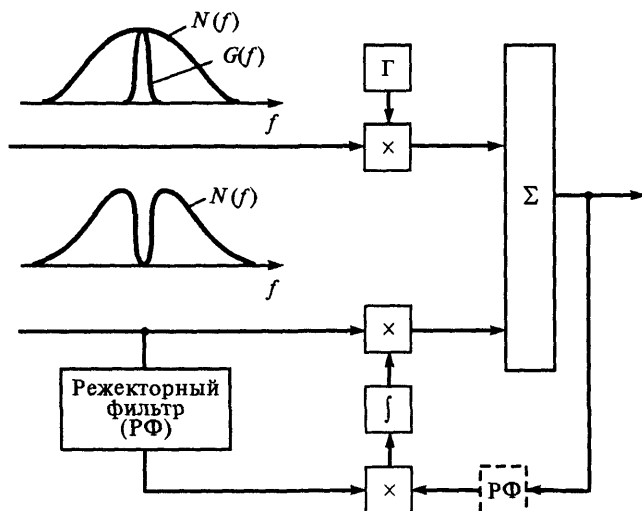


Рис. 7.44. Устранение полезного сигнала из цепей КОС при использовании частотных различий

Рассмотрим принцип поляризационной адаптивной компенсации помех. Для адаптивной компенсации помех в ряде случаев целесообразно использовать поляризационные различия сигнала и помехи. Это представляет интерес, например, для случая, когда помеха, перекрывающая сигнал в частотной и временной областях, воздействует по главному лепестку диаграммы направленности приемной антенны, и поэтому ни пространственная, ни частотная или временная селекции помех невозможны.

Автоматическая подстройка поляризации может осуществляться за счет использования корреляционного автокомпенсатора (квадратурного или гетеродинного) и двухканальной антенны. Конструкция антенны такова, что обеспечивается прием колебаний с двумя взаимно ортогональными (например, линейными) поляризациями. На рис. 7.45, а представлена упрощенная схема автоматического поляризационного селектора, построенного на базе квадратурного автокомпенсатора. Такой автокомпенсатор обеспечивает подавление помех от одного источника, действующего в главном лепестке диаграммы направленности приемной антенны РЛС [33, 35, 40].

Пусть помеха N имеет, например, регулярную эллиптическую поляризацию. В этом случае между ортогональными составляющими N_x и N_y имеется фазовый сдвиг φ_n (значение φ_n зависит от вида поляризации). Автокомпенсатор выравнивает амплитуды и фазы входных колебаний и осуществляет компенсацию помехи (см. рис. 7.45, б). Полезный сигнал имеет другую поляризацию и, следовательно, другой фазовый сдвиг между его ортогональными составляющими. Колебания сигнала можно разложить на две со-

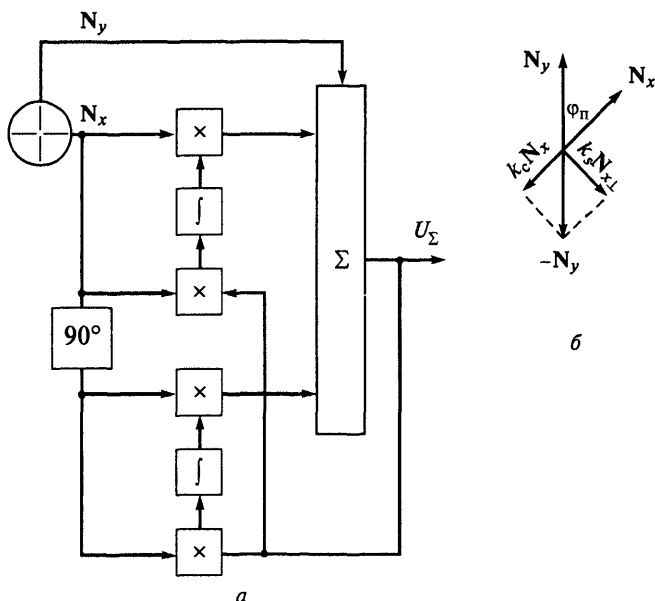


Рис. 7.45. Пояснение принципа поляризационной адаптивной компенсации помех:

а — поляризационный автокомпенсатор, *б* — векторная диаграмма

ставляющие: имеющую поляризацию помехи и ортогональную ей. Первая составляющая сигнала будет подавлена поляризационным автокомпенсатором, а для другой составляющей этот селектор будет оптимально настроенным. Эта составляющая и создает соответствующий выходной эффект. Таким образом, если сигнал и помеха имеют различную *регулярную* поляризацию, то в результате работы автокомпенсатора антенная система настраивается на поляризацию, которая ортогональна поляризации помехи.

Следует, однако, заметить, что применение активных помех с *хаотической* поляризацией существенно затрудняет возможность поляризационной селекции.

7.9. Методы защиты от пассивных маскирующих помех

7.9.1. Основные различия сигналов целей и пассивных маскирующих помех

Сигналы, отраженные от целей, и пассивные маскирующие помехи в общем случае имеют различные статистические характеристики. Для сигналов и помех, распределенных по нормальному закону, эти различия сводят-

ся к различиям их корреляционных матриц, которые, в свою очередь, обусловлены различиями некоторых физических характеристик целей и отражателей, создающих пассивную помеху. К числу этих различий можно отнести следующие [31].

1. *Различие мешающих отражателей и целей по характеру распределения в пространстве.* Цель обычно близка к сосредоточенному объекту, мешающие отражатели распределены в пространстве. Повышая разрешающую способность по координатам и сокращая при этом размеры разрешаемого объема (во всяком случае, до размеров, превышающих размеры самолета), можно добиться улучшения наблюдаемости сигнала на фоне пассивных помех.

2. *Различия в поляризации* отраженных сигналов наблюдаются, если пассивная помеха создается, например, гидрометеорами (дождь, тучи), состоящими из мелких капель, имеющих форму шара. Если гидрометеоры облучаются колебаниями с круговой поляризацией, то они отражают колебания также с круговой поляризацией, но с обратным (в направлении распространения волны) вращением плоскости поляризации. Если приемная антенна не воспринимает колебания с такой поляризацией, она, тем не менее, может принимать колебания от целей, обладающих несимметрией структуры.

3. *Различия в скорости перемещения мешающих отражателей и цели.* Скорость перемещения наземных мешающих отражателей относительно наземной радиолокационной станции близка к нулю, в то время как представляющие практический интерес цели перемещаются с достаточно большой скоростью.

Если пассивная помеха создается противорадиолокационными отражателями, то эти отражатели, будучи сброшены с самолета, быстро приобретают скорость, близкую к скорости ветра. Поскольку скорость ветра не постоянна по высоте, имеет место разброс скоростей противорадиолокационных отражателей. Тем не менее, различия в радиальных скоростях целей и отражателей имеются и могут быть использованы для селекции. Селекцию по скорости (или по эффекту движения цели) называют *селекцией движущихся целей* (СДЦ).

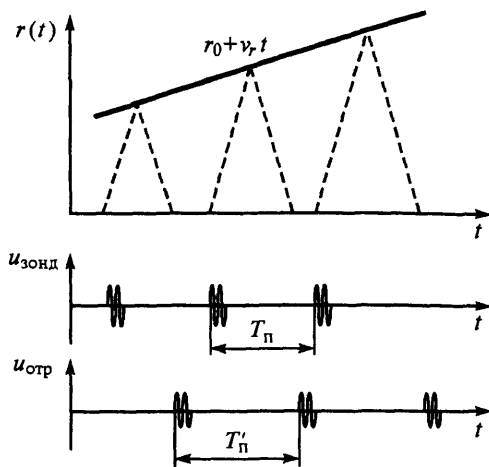


Рис. 7.46. Пояснение трансформации периодической последовательности радиоимпульсов, отраженных движущейся блестящей точкой

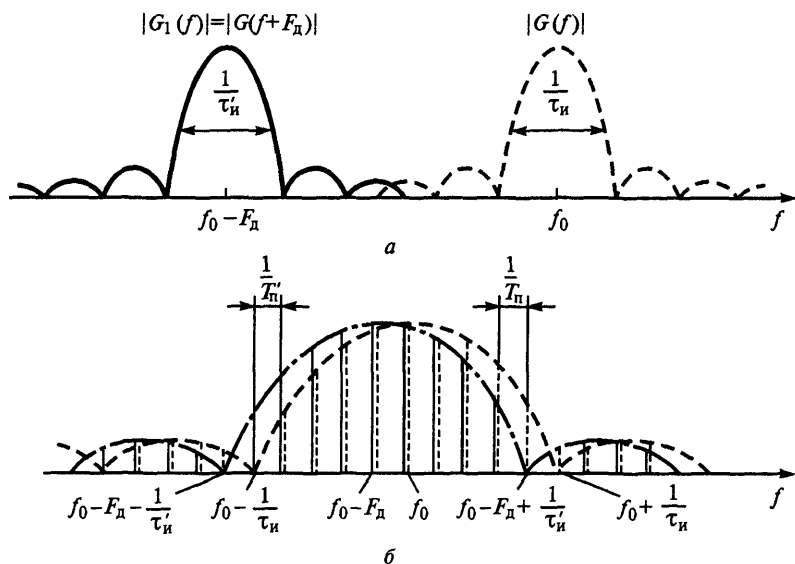


Рис. 7.47. Трансформация спектров сигналов, отраженных от движущейся цели:

a — одиночного радиоимпульса, b — периодической последовательности радиоимпульсов

В основе СДЦ лежит явление деформации структуры сигнала при отражении от движущейся цели. Это явление иллюстрируется на рис. 7.46 для последовательности коротких радиоимпульсов. Можно убедиться в том, что при движении цели от радиолокатора происходит увеличение длительности всего импульса и периода высокочастотных колебаний в $K_t = (c + v_r)/(c - v_r) \approx (1 + 2v_r/c)$ раз (см., например, [31, 43, 73]). Для импульсной последовательности (см. рис. 7.46) происходит также растяжение промежутка между двумя соседними импульсами в указанное число раз, так что этот промежуток получает приращение $\Delta T_n = 2v_r T_n/c$. На рис. 7.47, a показано соответствующее изменение спектра для случая одиночного радиоимпульса, а на рис. 7.47, b — для случая, приведенного на рис. 7.46 (в предположении, что все радиоимпульсы на этом рисунке представляют вырезку из одной синусоиды, а последовательность импульсов периодическая). В каждом из указанных случаев растяжению по оси времени в $(c + v_r)/(c - v_r)$ раз соответствует сжатие по оси частот также в $(c + v_r)/(c - v_r)$ раз. Если ширина спектра, как это показано на рис. 7.47, a , мала по сравнению с несущей, то деформация амплитудно-частотного спектра сводится к смещению его на некоторую доплеровскую частоту

$F_d = 2v_r/\lambda$. Подобный эффект показан на рис. 7.47, б для гребенчатого спектра периодической последовательности радиоимпульсов. Численно изменение промежутка между импульсами за счет скоростной деформации сигнала невелико. Например, для $T_n = 10^{-3}$ с, $v_r = 150$ м/с и $c = 3 \cdot 10^8$ м/с оно составляет 10^{-9} с, т. е. величину одного порядка с периодом колебаний высокой частоты. Это значит, что деформацию сигнала можно заметить лишь по изменению фазы колебаний высокой частоты. Чтобы использовать эту возможность, предъявляют достаточно жесткие требования к фазовой структуре высокочастотных колебаний, т. е. к их когерентности.

Различают несколько видов обеспечения когерентности колебаний [31].

1. Истинная внутренняя когерентность достигается тем, что колебания создаются стабильным задающим генератором, после которого стоит усилитель мощности с устойчивой фазовой характеристикой.

2. Эквивалентная внутренняя когерентность достигается тем, что генератор с самовозбуждением вырабатывает последовательность импульсов постоянной несущей частоты со случайными начальными фазами. Начальная фаза каждого зондирующего импульса запоминается на время приема отраженных сигналов до следующего зондирования. Путем соответствующей обработки эта фаза исключается и принимаемые колебания оказываются практически такими же, как и в случае истинной когерентности.

3. Внешняя когерентность достигается тем, что информация о случайной начальной фазе зондирующего импульса извлекается из приходящих от пассивных отражателей колебаний.

Будем полагать далее, что имеет место истинная внутренняя когерентность.

Если вторичные излучатели, имеющие разные скорости, разрешаются по дальности и угловым координатам, то независимо от вида когерентности задача селекции по скорости состоит в выявлении временных (фазовых) или спектральных различий для каких-либо участков пространства. Например, для сигнала, показанного на рис. 7.47, следует определить, имеется ли смещение спектра на доплеровскую частоту, соответствующую радиальной скорости цели $F_{Дц}$.

Значительно сложнее задача селекции движущихся целей в том случае, когда цель и мешающие отражатели находятся в одном элементарном разрешаемом объеме. В этом случае необходимо улучшить условия обнаружения сигнала от цели на фоне пассивной помехи за счет имеющих место временных и спектральных отличий. Последнее достигается путем *режекции* (подавления) колебаний помехи и накопления сигнала.

Принцип режекции колебаний помехи состоит в том, что обеспечиваются условия ее подавления, например спектральные составляющие помехи вырезаются (рис. 7.48). Накопление сигнала состоит в том, что спектральные составляющие сигнала обрабатываются с целью наилучшего выделения сигнала на фоне шумов и пассивной помехи. Поскольку требования

режекции и накопления могут оказаться противоречивыми, для выяснения их оптимального соотношения при селекции движущихся целей может быть использована формула оптимальной фильтрации сигнала на фоне небелого шума, строгий вывод которой приведен в гл. 3. Далее рассмотрим другой способ получения этой формулы и ее приложение для синтеза устройств оптимальной обработки сигналов на фоне пассивных помех.

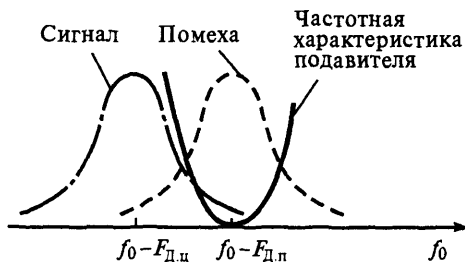


Рис. 7.48. Пояснение принципа режекции помехи

7.9.2. Оптимальное обнаружение сигнала на фоне пассивной помехи в виде стационарного небелого шума

Небелый шум, как известно, характеризуется неравномерным распределением спектральной плотности мощности по оси частот. Такое же неравномерное распределение спектральной плотности мощности характерно и для пассивной помехи (в том числе с наложенным внутренним шумом). В самом деле, если пассивная помеха в некотором импульсном объеме образована отражателями, имеющими различные радиальные скорости v_{r_i} (соответствующие доплеровским частотам $F_{Д.п_i}$), то спектральная плотность мощности суммарной помехи будет определяться выражением вида

$$N(f) = \sum_i k_i G_n(f - F_{Д.п_i}) + N_0, \quad (7.71)$$

где k_i — коэффициент пропорциональности, зависящий от числа отражателей в разрешаемом объеме, имеющих радиальную скорость v_{r_i} ; $G_n(f)$ — спектральная плотность помехи для неподвижных отражателей с учетом обзора по угловой координате; N_0 — спектральная плотность мощности белого шума.

При определенных условиях максимум спектральной плотности соответствует средней доплеровской частоте $F_{Д.п.ср}$. Ширина спектра зависит от

степени разброса скоростей и ширины пика спектральной плотности зондирующего сигнала.

Поскольку число отражателей и распределение скоростей могут меняться от одного разрешаемого объема к другому, в случае последовательного обзора по дальности и азимуту пассивную помеху следует считать нестационарной. Тем не менее, при изучении возможности селекции в пределах каждого разрешаемого объема нестационарность не существенна, и помеху в первом приближении можно заменить стационарным небелым шумом, подобным тому, который формируется при прохождении белого шума через фильтр [31, 74].

Получим формулу оптимальной фильтрации сигнала на фоне небелого (коррелированного) стационарного шума. Предполагая, что спектральная плотность $N(f)$ нигде не обращается в нуль, примем, что шум со спектральной плотностью мощности $N(f)$ и сигнал со спектральной плотностью напряжения $G(f)$ пропущены через фильтр с частотной характеристикой $K_0(f)$ (рис. 7.49). Амплитудно-частотную характеристику этого предварительного фильтра выберем так, чтобы спектральная плотность мощности помехи на его выходе $N(f)|K_0(f)|^2$ не зависела от частоты:

$$N(f)|K_0(f)|^2 = \text{const} = C,$$

или

$$|K_0(f)| = \sqrt{\frac{C}{N(f)}}. \quad (7.72)$$

Если все значения спектральной плотности $N(f)$ конечны, то предварительная фильтрация не приведет к потере каких-либо спектральных составляющих, т. е. любую составляющую спектра можно восстановить при последующей фильтрации. Поскольку шум на выходе предварительного фильтра стал белым, то оптимальное обнаружение осуществляется путем известной процедуры фильтрации на фоне белого шума (см. гл. 3). Такой фильтрации должен быть подвергнут полезный сигнал с выхода предварительного фильтра, имеющий комплексный амплитудно-частотный спектр $G(f)K_0(f)$. Поэтому оптимальная частотная характеристика последующего фильтра с точностью до постоянного множителя выражается формулой

$$K_1(f) = K_{\text{опт}}(f) = (G(f)K_0(f))^* e^{-j2\pi f t_0}. \quad (7.73)$$

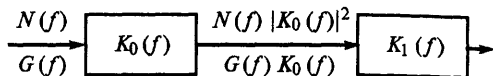


Рис. 7.49. Пояснение вывода формулы оптимальной фильтрации сигнала на фоне стационарного небелого шума

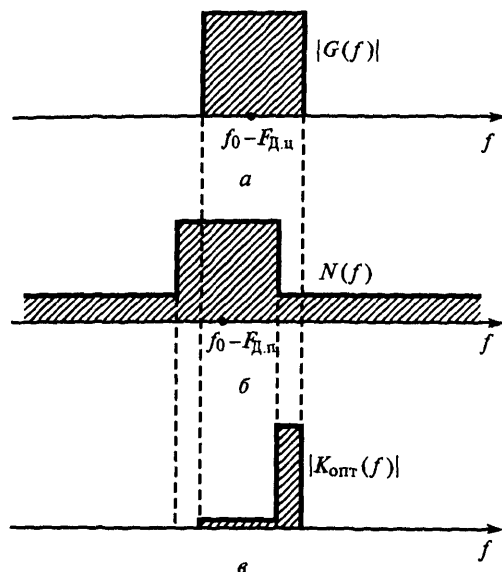


Рис. 7.50. Пояснение формулы оптимальной фильтрации:

a — амплитудно-частотный спектр сигнала,
б — спектральная плотность мощности помехи,
в — амплитудно-частотная характеристика оптимального фильтра

Тогда оптимальную частотную характеристику устройства обработки сигнала на фоне небелого шума в целом можно определить следующими выражениями:

$$K_{\text{опт}}(f) = K_0(f) K_{1\text{опт}}(f) = \\ = |K_0(f)|^2 G^*(f) e^{-j2\pi f t_0}$$

или

$$K_{\text{опт}}(f) = C \frac{G^*(f)}{N(f)} e^{-j2\pi f t_0}. \quad (7.74)$$

Полученная формула оптимальной частотной характеристики для случая небелого шума является обобщением формулы для случая белого шума. Последнюю получаем из соотношения (7.74), полагая $N(f) = N_0 = \text{const}$.

Из формулы (7.74) следует, что наряду с накоплением сигнала необходимо проверить режекцию помехи. Это иллюстрируется на рис. 7.50, где показаны амплитудно-частотный спектр сигнала $|G(f)|$, спектральная плотность мощности небелого шума $N(f)$ и амплитудно-частотная характеристика оптимального фильтра $|K_{\text{опт}}(f)|$. Эта характеристика свидетельствует о необходимости подавления спектральных составляющих шума с наибольшей интенсивностью, хотя при этом одновременно подавляются и отдельные спектральные составляющие сигнала. Результирующую амплитудно-частотную характеристику $|K_{\text{опт}}(f)|$ (см. рис. 7.50, *в*) можно представить как произведение двух амплитудно-частотных характеристик: $K_1(f) = |G(f) K_0(f)|$ обеспечивает оптимальную обработку сигнала на фоне белого шума, $|K_0(f)|$ — режекцию (подавление) спектральных составляющих принимаемых колебаний.

На рис. 7.51 иллюстрируется применение полученных в настоящем параграфе формул к случаю оптимального приема когерентной пачки периодически следующих импульсов при наличии пассивной помехи и внутреннего шума. На рис. 7.51, *a* изображен амплитудно-частотный спектр сиг-

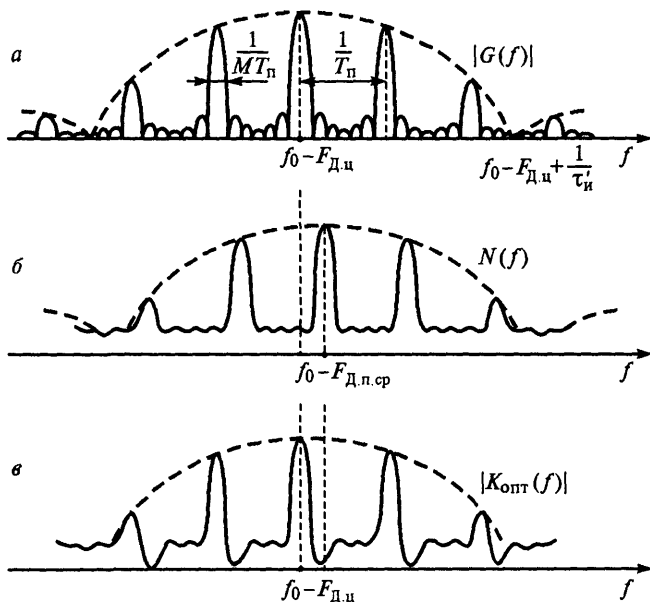


Рис. 7.51. Амплитудно-частотный спектр пачки радиоимпульсов:

а — отраженных от движущейся цели, *б* — спектральной плотности мощности помехи, *в* — амплитудно-частотная характеристика оптимального фильтра

нала $|G(f)|$, на рис. 7.51, *б* — спектральная плотность мощности помехи $N(f)$. Представленная кривая $N(f)$ получена в результате сдвига кривой спектральной плотности зондирующего сигнала на среднюю доплеровскую частоту помехи $F_{Д.п.ср}$, наложения составляющих внутреннего шума N_0 , а также учета размытия гребенчатого спектра помехи за счет разброса скоростей отражателей. На рис. 7.51, *в* представлена амплитудно-частотная характеристика оптимального фильтра, соответствующая формуле (7.74).

Условия оптимальной обработки могут быть реализованы, если последовательно включены согласованный (оптимальный) фильтр для одиночного импульса (ОФОИ) пачки, гребенчатый фильтр накопления (ГФН), например, в виде линии задержки — рециркулятора и, наконец, гребенчатый фильтр подавления (ГФП) гребней спектра помехи (см. рис. 7.52, *а*). Соответствующие частотные характеристики показаны на том же рисунке. Первые два фильтра (рис. 7.52, *а*) обеспечивают оптимальную обработку импульсов пачки на фоне белого шума, последний — режекцию помехи (порядок включения фильтров ГФП и ГФН может быть изменен, так как произведение амплитудно-частотных характеристик при этом не меняется). Фильтр накопления настроен на скорость цели (с помощью, например, уст-

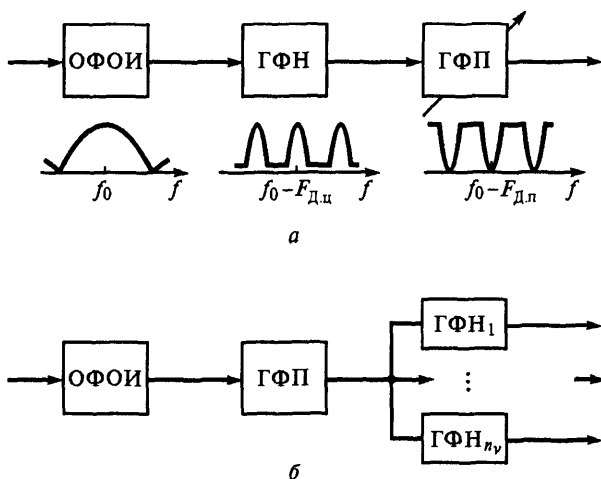


Рис. 7.52. Схема оптимальной фильтрации сигнала на фоне небелого шума:

a — с известной, *б* — с неизвестной частотой сигнала

ройства автосопровождения по скорости), а фильтр подавления — на скорость помехи. Если настройка на скорость цели не производится, то ГФН должен быть многоканальным по скорости цели (см. рис. 7.52, б). С выхода фильтров напряжение подается на детектор [31, 40].

На рис. 7.53 показана видоизмененная схема обработки, в которой когерентный накопитель (в общем случае многоканальный) заменен некогерентным (последетекторным). При использовании этой схемы не требуется настройка накопителя на скорость цели, достаточно настроить гребенчатый фильтр подавления на среднюю скорость помехи. Частотная характеристика додетекторных каскадов определяется в данном случае по формуле

$$K_{\text{опт}}(f) \equiv \frac{|G_{\text{н}}(f)|}{N(f)},$$

т. е. соответствует обнаружению одиночных радиоимпульсов некогерентной пачки на фоне небелого шума. Гребенчатые фильтры подавления и накопления могут быть сравнительно просто реализованы при аналоговой (на промежуточной частоте) и при цифровой обработке. Оптималь-

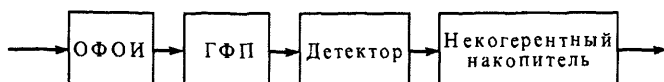


Рис. 7.53. Схема когерентного подавления пассивной помехи с некогерентным междупериодным накоплением сигнала

ная обработка сигналов на фоне пассивных помех может быть проведена и на видеочастоте.

Качество (эффективность) защиты от пассивных помех может оцениваться теми же показателями, которые используются при оценке эффективности защиты от активных помех. Однако на практике получили распространение свои специфические показатели. К ним, в частности, относится *коэффициент подпомеховой видимости*

$$K_{п.в} = \frac{(P_c/P_n)_{\text{ВЫХ}}}{(P_c/P_n)_{\text{ВХ}}}, \quad (7.75)$$

где P_c и P_n — мощности соответственно сигнала и помехи.

Коэффициент $K_{п.в}$ показывает, во сколько раз улучшается отношение сигнал—помеха при использовании схем помехозащиты. Если положить, что коэффициент передачи схемы по сигналу равен единице, то соотношение (7.75) преобразуется к виду

$$K_n = \frac{P_{п.ВХ}}{P_{п.ВЫХ}}.$$

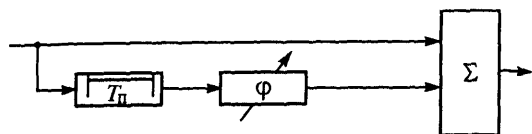
Этот показатель называется *коэффициентом подавления* помехи. В современных РЛС обеспечиваются значения K_n , превышающие 50 дБ [92].

7.9.3. Череспериодное вычитание — способ создания гребенчатых фильтров подавления

Принципы построения входящих в состав оптимального фильтра (см. рис. 7.52) оптимального (согласованного) фильтра одиночного импульса (ОФОИ) и когерентного накопителя (ГФН) были рассмотрены в гл. 3. Напомним лишь, что когерентные накопители могут быть реализованы в виде аналоговых и цифровых рециркуляторов, многоотводных линий задержки с сумматором. В этой связи основное внимание сосредоточим далее на принципах построения и особенностях ГФП помехи.

Простейший гребенчатый фильтр подавления выполняется по схеме однократного череспериодного вычитания (ЧПВ) (рис. 7.54). Линия задержки осуществляет задержку точно на период посылки зондирующих сигналов T_n . Фазовращатель вносит сдвиг по фазе на угол φ . Найдем частотную характеристику цепи как отношение выходного напряжения ко входному, если последнее представляет комплексную гармонику произвольной частоты $\exp(j2\pi ft)$:

$$K(f) = 2j \sin(\pi f T_n + \varphi/2) \exp[-j(\pi f T_n + \varphi/2)]. \quad (7.76)$$



Отсюда находим амплитудно-частотную характеристику (АЧХ) цепи

Рис. 7.54. Схема однократного череспериодного вычитания

$$K(f) = 2|\sin(\pi f T_n + \varphi/2)|, \quad (7.77)$$

графически представленную на рис. 7.55. Из формулы (7.77) следует, что АЧХ имеет минимумы (провалы) через интервалы $1/T_n$, положение которых можно изменять, подбирая необходимую величину φ . Можно подобрать такое положение провалов, чтобы они совпадали с положением гребней спектра пассивной помехи, определяемые радиальной скоростью движения мешающих отражателей.

Полезные сигналы, отраженные от цели, также являются когерентными в смежных периодах зондирования и, проходя через устройство когерентного подавления помехи, искажаются. В этой связи наряду с АЧХ используют *амплитудно-скоростную характеристику (АСХ)*, определяющую зависимость амплитуды сигнала на выходе устройства ЧПВ от разности скоростей цели и источников помехи. Амплитудно-скоростная характеристика $K(\Delta v_r)$ устройства ЧПВ (рис. 7.56) по форме совпадает с его амплитудно-частотной характеристикой (см. рис. 7.55).

Радиальные скорости цели, для которых амплитуда сигнала на выходе устройства ЧПВ обращается в нуль (см. рис. 7.56), называют *слепыми скоростями*. Значения слепых скоростей определяются соотношением

$$v_{r\text{ сл}}^{(n)} = v_{r\text{ п.ср}} + n \frac{\lambda}{2T_n}, \quad (7.78)$$

где $v_{r\text{ п.ср}}$ — средняя радиальная скорость перемещения источников пассивных помех, $n = 0, \pm 1, \pm 2, \dots$

Если устройство ЧПВ настроено на подавление помех от местных не-

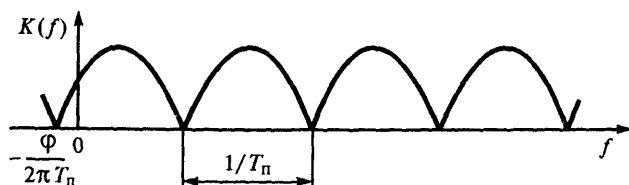


Рис. 7.55. Амплитудно-частотная характеристика устройства однократного ЧПВ

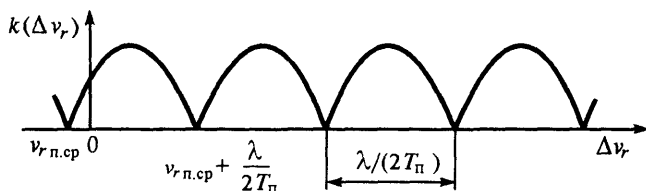


Рис. 7.56. Амплитудно-скоростная характеристика устройства однократного ЧПВ

подвижных предметов ($v_{r п.сп} = 0$), то соотношение (7.78) преобразуется к виду

$$v_{r сл}^{(n)} = n \frac{\lambda}{2T_{п}}, \quad n = 0, \pm 1, \pm 2, \dots \quad (7.79)$$

В этом случае слепая скорость цели — это такая радиальная ее скорость, при которой цель за время, равное периоду повторения импульсов T , проходит расстояние, кратное половине длины волны РЛС. При таком изменении расстояния фаза сигнала от импульса к импульсу не изменяется, так как, набег фаз на удвоенном пути $2r = 2(n\lambda/2)$ определяется соотношением

$$\varphi = k \cdot 2r = \frac{2\pi}{\lambda} \frac{2n\lambda}{2} = 2\pi n,$$

при выполнении которого обеспечивается полная компенсация сигнала наряду с пассивной помехой.

Наиболее распространенным способом устранения слепых скоростей является изменение (вобуляция) периода повторения зондирующих импульсов.

Наряду с АЧХ и АСХ для описания устройства ЧПВ используют характеристику, представляющую собой зависимость отношения сигнал—помеха на его выходе (или коэффициента подпомеховой видимости) от разности скоростей цели и источников помехи. Эта характеристика одновременно учитывает прохождение и помехи, и сигнала через устройство ЧПВ.

Для более качественного приближения частотной характеристики к оптимальной на практике используют устройства многократного вычитания. На рис. 7.57 показан пример устройства двукратного вычитания, исполь-

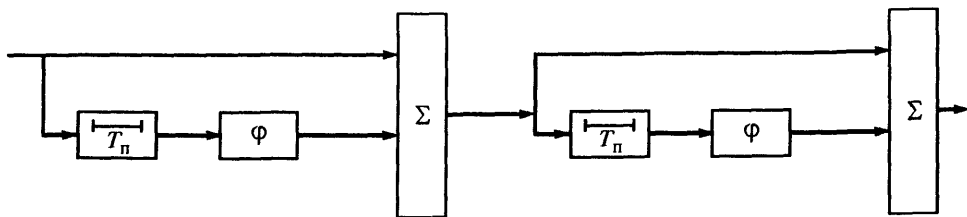


Рис. 7.57. Схема устройства двукратного ЧПВ

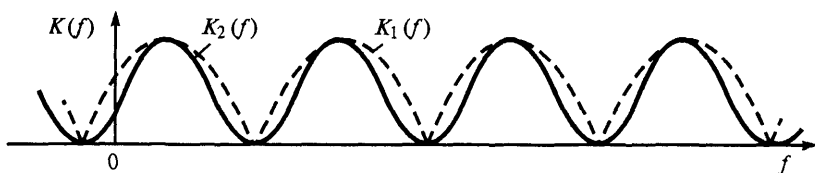


Рис. 7.58. Амплитудно-частотная характеристика устройства двукратного ЧПВ

зующего последовательное включение двух устройств однократного ЧПВ. Амплитудно-частотную характеристику $K_2(f)$ устройства двукратного вычитания (рис. 7.58) можно получить, перемножая амплитудно-частотные характеристики (7.77) устройств однократного череспериодного вычитания:

$$K_2(f) = K_1^2(f) = 4 \sin^2(\pi f T_n + \varphi/2). \quad (7.80)$$

Как отмечалось выше, для характеристики эффективности устройств СДЦ можно использовать коэффициент подавления помехи. Определяя мощности помехи на входе $P_{\text{вх}} = \sigma_{\text{вх}}^2 = \mathbf{M}\{u_{\text{вх}}^2\}$ и выходе $P_{\text{вых}} = \mathbf{M}\{[u_{\text{вх}}(t) - u_{\text{вх}}(t - T_n)]^2\}$ устройства однократного ЧПВ, после преобразований получим выражение для коэффициента подавления помехи:

$$K_n = \frac{1}{2(1 - r(T_n))}, \quad (7.81)$$

где

$$r(T) = \frac{1}{\sigma_{\text{вх}}^2} \mathbf{M}\{u_{\text{вх}}(t)u_{\text{вх}}(t - T_n)\}, \quad (7.82)$$

— коэффициент междупериодной корреляции флуктуации пассивной помехи.

Для устройства двукратного вычитания коэффициент подавления определяется в виде

$$K_{n2} = \frac{1}{2(3 - 4r(T_n) + r(2T_n))}, \quad (7.83)$$

где $r(2T_n)$ — коэффициент, учитывающий корреляцию флуктуаций помехи за два периода повторения импульсов.

На рис. 7.59 приведены зависимости коэффициента подавления помехи с корреляционной функцией (КФ) гауссовского вида от количества ступеней вычитания ($n_{\text{ст}}$) с учетом внутриприемных шумов. Анализ этих зависимостей показывает, что на практике целесообразно использовать не более

4–6 ступеней вычитания, поскольку при дальнейшем увеличении количества ступеней прирост величины коэффициента подавления незначителен [40].

Коэффициенты подпомеховой видимости устройств однократного и двукратного вычитания, определяемые с учетом соотношения (7.75), имеют вид

$$K_{п.в1} = \frac{2 \sin^2 \pi \Delta F_{\text{д}} T_{\text{п}}}{1 - r(T_{\text{п}})}, \quad (7.84)$$

$$K_{п.в2} = \frac{8 \sin^4 \pi \Delta F_{\text{д}} T_{\text{п}}}{3 - 4r(T_{\text{п}}) + r(2T_{\text{п}})}. \quad (7.85)$$

Отсюда следует, что коэффициент подпомеховой видимости зависит как от коэффициента междупериодной корреляции флуктуаций помехи, так и от разности доплеровских частот сигнала и помехи $\Delta F_{\text{д}} = F_{\text{д.с}} - F_{\text{д.п}}$.

В системах СДЦ необходима адаптация к случайной средней скорости радиального перемещения по отношению к РЛС мешающих, например, дипольных отражателей («компенсация ветра»). В устройствах ЧПВ это может быть осуществлено путем автоматизированного или неавтоматизированного управления фазовращателем (см. рис. 7.54) или частотой колебаний гетеродина, предшествующего устройству череспериодной компенсации. Эта задача решается также в автокомпенсаторах пассивных помех. На рис. 7.60 представлены схемы однократного квадратурного (см. рис. 7.60, а) и гетеродинного (см. рис. 7.60, б) автокомпенсаторов пассивных помех. Коэффициент подавления однократного компенсатора определяется в виде

$$K_{\text{п}} = \frac{1}{1 - r^2(T_{\text{п}})}. \quad (7.86)$$

На рис. 7.61 показана схема гетеродинного автокомпенсатора с двукратным вычитанием. Автокомпенсаторы автоматически отслеживают положение, форму и глубину провалов АСХ в соответствии с динамикой изменения помехи [40].

В заключение отметим особенности цифровой обработки сигналов на фоне пассивной помехи. В этом случае алгоритмы СДЦ могут быть реализованы с помощью спецпроцессоров. Так, при цифровой обработке во временной области для адаптации к пассивным помехам могут быть использо-

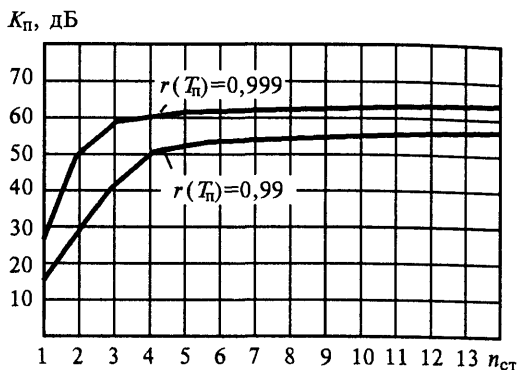


Рис. 7.59. Зависимости величины коэффициента подавления помехи с КФ гауссовского вида от количества ступеней вычитания с учетом внутримприемных шумов и коэффициента междупериодной корреляции помехи

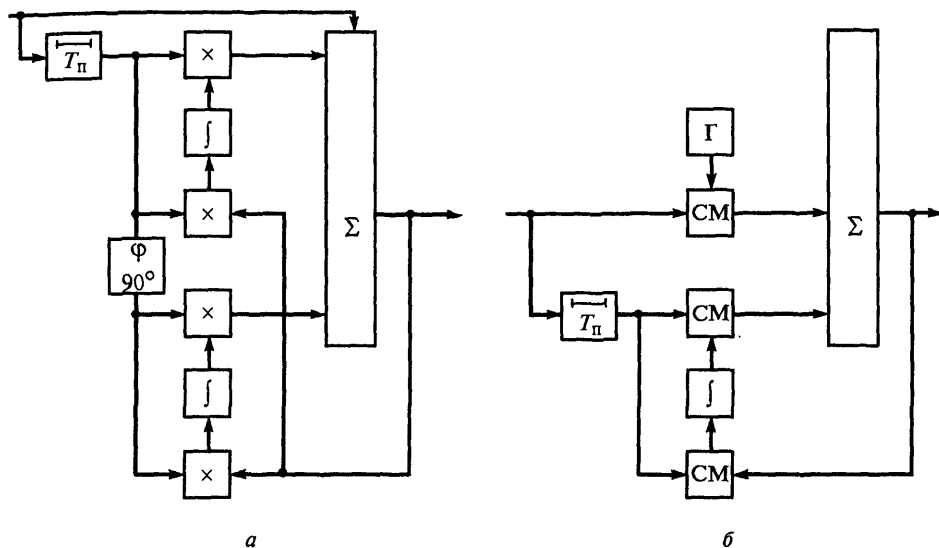


Рис. 7.60. Автокомпенсаторы пассивных помех:

а — квадратурный, б — гетеродинный

ваны оптимальные алгоритмы оценивания корреляционной матрицы (векторного вектора), рассмотренные, например, в [34, 35, 40, 74]. Не исключается также и применение квазиоптимальных методов оценивания на базе корреляционных автокомпенсаторов.

При цифровой обработке в частотной области с помощью БПФ формируется набор доплеровских каналов. При этом каналы, настроенные на доплеровскую частоту помехи, из обработки исключаются.

7.10. Методы стабилизации уровня ложных тревог

Несмотря на использование в современных РЛС разнообразных систем помехозащиты, уровень помех и некомпенсированных остатков помех (нежелательные эхо-сигналы от дождя и местных предметов и нежелательные сигналы от других источников излучений) часто превышают уровень шума приемника. Эти источники помех могут полностью забить экран индикатора РЛС или перегрузить ЭВМ, осуществляющую решения типа «да—нет» с целью выявления эхо-сигналов, отраженных от целей, подлежащих обнаружению или сопровождению [73, 87].

Дискретный процесс решения этой задачи обычно связан с установлением порогов обнаружения на входе и выходе цифрового вычислителя. В каждой

точке можно определить вероятность обнаружения желательной цели и вероятность ложной тревоги из-за шума или одного из упомянутых выше источников помех. Оператор, наблюдающий экран индикатора кругового обзора (ИКО), принимает в некотором смысле сходное решение, и поэтому понятие *ложных тревог* применимо к большинству радиолокационных систем. И оператор, и цифровой вычислитель будут стремиться поддерживать частоты ложных тревог на некотором приемлемом постоянном уровне путем регулирования усиления в приемном тракте РЛС при изменении интенсивности помех. Приемники, обладающие этим свойством, называются *приемниками с постоянной частотой ложных тревог*, или ПЧЛТ-приемниками.

Методы стабилизации уровня ложных тревог основаны на использовании амплитудных или фазовых различий между полезными сигналами и помехами. Рассмотрим эти методы [73, 87].

Методы амплитудного различения. Для изменения коэффициента усиления приемного тракта РЛС широко используются усилители с регулируемым усилением, но их возможности по поддержанию ПЧЛТ ограничены. Так, система временной регулировки усиления (ВРУ) меняет усиление в соответствии со средним ожидаемым уровнем отражений от дождя или местных предметов на различных дальностях, но отклонения от этих ожидаемых уровней или воздействие активных источников помех могут привести к нарушению нормальной работы РЛС. Система автоматической регулировки усиления (АРУ) будет стремиться поддерживать уровень помех постоянным, но такая система всегда инерционна и медленно реагирует на изменение уровня помех. Например, системы автоматической регулировки усиления по уровню шума (ШАРУ) производят стробирование шума в нерабочих интервалах между излученным и принимаемым импульсами и приводят обычно к неправильному решению о том, что данный уровень помех характерен для всех дальностей.

Если селектируемая цель по своим физическим размерам мала по сравнению с размерами элемента разрешения РЛС по дальности, а мешающие сиг-

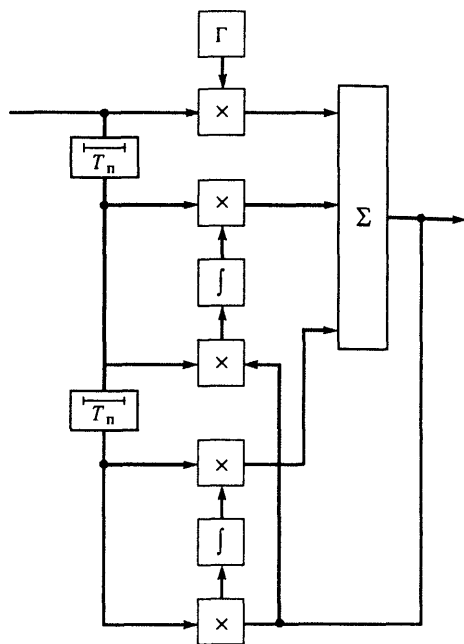


Рис. 7.61. Гетеродинный автокомпенсатор с двукратным вычитанием

налы распределены по протяженной зоне, большей по сравнению с размером элемента разрешения, то знание этих особенностей цели и помех можно также использовать для их различения. Оптимизация приемника для данной длительности импульса может состоять в применении фильтра высоких частот (схема с малой постоянной времени) или дискриминатора по длительности. Такой дискриминатор запрещает прохождение сигналов, обладающих меньшей или большей длительностью, чем ожидаемый сигнал. Поскольку пространственное распределение большинства мешающих отражений содержит много сигналов, близких по длительности к полезным, этот метод в отдельности также не обеспечивает достаточно эффективного подавления помехи.

Более широкое применение находит система, включающая логарифмический усилитель и схему с малой постоянной времени (МПВ) или дискриминатор по длительности. На выходе логарифмического обнаружителя напряжение от подобных пространственно распределенных помех от местных предметов можно разделить на две составляющие. Одна составляющая быстро меняется с дальностью, и ее распределение вероятности и скорость флуктуаций хорошо согласуются с распределением вероятности и скоростью флуктуации шума независимо от величины сигналов, отраженных от местных предметов. Известно [73], что среднеквадратическое значение выбросов на выходе логарифмического обнаружителя для сигналов, подчиняющихся рэлеевскому распределению в основной полосе частот, равно $a\pi\sqrt{24}$, где a — крутизна логарифмической характеристики, В/дБ; эта составляющая имеет ПЧЛТ-характеристику для помехи данного типа. Другая составляющая напряжения меняется более медленно в зависимости от дальности, представляя среднюю интенсивность помех от местных предметов или мешающих эхо-сигналов на любой дальности; возможность получения ПЧЛТ-характеристики зависит от того, в какой степени эта составляющая может быть подавлена фильтром.

В приемнике с логарифмическим обнаружителем, имеющим малую постоянную времени (логарифмический МПВ-приемник), используют фильтр простейшей формы и дифференцирующую цепь с малой постоянной времени. Логарифмический МПВ-приемник оказывается эффективным лишь в случае, когда помеха является «равномерной» (по дальности) в течение времени, большем постоянной времени схемы МПВ. Неравномерность в пространстве пассивной помехи любого вида ограничивает эффективность подобного устройства. В этом случае используют систему: логарифмический усилитель — дискриминатор длительности импульсов (ДДИ). Система lg—ДДИ позволяет режектировать импульсные помехи и мешающие отражения, существенно отличающиеся по длительности от зондирующих импульсов [87].

Методы фазового различения. В ПЧЛТ-приемниках другого класса вся информация об уровне эхо-сигналов полностью утрачивается вследст-

вие использования ограничителей с уровнем ограничения, лежащим значительно ниже нормального уровня шумов [77]. В этих приемниках различение полезных эхо-сигналов и помех осуществляется исключительно по изменению во времени фазовой структуры напряжения на выходе ограничителя на основе того, насколько хорошо она коррелирована с фазовым кодом излученного импульса.

В тех РЛС, которые излучают последовательность простых импульсов, характеристикой, позволяющей различать полезный эхо-сигнал от шума или помех от местных предметов, является мгновенная частота, которая остается постоянной на протяжении импульса независимо от значения несущей частоты. Для улучшения этой способности различения и использования других преимуществ сжатия импульсов часто применяют более сложные фазовые коды излучаемых сигналов. На этом общем принципе основано широкое многообразие аппаратных реализаций, каждая из которых обладает теми или иными преимуществами при подавлении помех определенных типов и недостатками в некоторых других помеховых ситуациях. Тем не менее, всем им присущи определенные общие свойства, на которых целесообразно остановиться.

В отличие от методов амплитудного различения методы фазового различения не связаны с проблемой скорости реакции на изменяющуюся интенсивность помех. Они чувствительны только к тому, насколько хорошо эхо-сигнал соответствует заданной внутриимпульсной фазовой структуре, и для них допустимы быстрые изменения уровня сигнала в широких пределах.

Режим ограничения разрушает некоторую часть информации, и в результате качество системы несколько ухудшается. Потери качества относятся к идеальному ограничителю, т. е. такому ограничителю, который осуществляет нормирование амплитуды, не искажая фазовых соотношений. Если фазовый сдвиг в ограничителе изменяется с амплитудой, то ограничитель преобразует амплитудный шум в фазовый. Если изменение амплитуды в пределах 1 дБ (12%) преобразуется в эквивалентное изменение фазы в пределах 0,12 радиана (7°), то амплитудный шум на входе создает фазовый шум на выходе, имеющий эквивалентную энергию, и вероятность обнаружения уменьшается. Чтобы эти потери были незначительными, крутизна амплитудно-фазового преобразования не должна превышать 2° на 1 дБ.

Увеличение произведения полосы пропускания на длительность сигнала с целью сокращения потерь из-за ограничения принимаемых колебаний приводит и к другим последствиям. Чем больше полоса пропускания приемника (до перехода его в режим ограничения) превышает ширину спектра излучаемого сигнала, тем более подвержен приемник влиянию помех не на основной частоте приема или мешающих сигналов, частота которых медленно изменяется. С другой стороны, если необходимо обеспечить хо-

рошие импульсные характеристики приемника, то нужно улучшить его способность к быстрому восстановлению после воздействия импульса или организованной помехи, частота которой быстро изменяется.

По мере возрастания длительности эхо-сигнала в ограниченном каскаде в результате фазового кодирования РЛС становится более уязвимой для эхо-сигналов, частично перекрывающихся по времени. Вследствие эффекта подавления слабого сигнала сильным вероятность обнаружения слабого быстро уменьшается при увеличении относительного их перекрытия, если после звена согласованной фильтрации установлен обнаружитель с фиксированным порогом обнаружения. Уменьшение вероятности обнаружения связано с подавлением и слабого сигнала, и шума в чашеобразной области, охватывающей временную длительность сильного сигнала. Влияние этой чашеобразной помеховой области можно ослабить, если после приемника с фазовым различением включить логарифмический обнаружитель ПЧЛТ. Последний эквивалентен по своему действию системе АРУ, изменяющей усиление, чтобы удалить «черный провал» или чашеобразную область, возникающую из-за ограничения.

7.11. Основные характеристики некоторых современных РЛС

7.11.1. Радиолокационные станции управления воздушным движением

В системах УВД используют различные наземные радиолокаторы (РЛ) и радиолокационные комплексы (РЛК), в том числе *трассовые* (ТРЛ и ТРЛК), *аэродромные* (АРЛ и АРЛК), *посадочные, вторичные* (с активным ответом), *обзора летного поля*. С развитием межгосударственных рейсов требования к ним согласуются в Международной авиационной организации ИКАО. В трассовых радиолокаторах часто используется диапазон длин волн 23 см, в аэродромных — диапазоны 10 и 23 см [104].

Трассовый и аэродромный радиолокационные комплексы ТРЛК-11, АРЛК-11 (Россия) систем УВД используют в качестве источника радиолокационной информации в автоматизированных и неавтоматизированных системах УВД на трассах и в аэродромной зоне. Включают первичный и вторичный (с активным ответом) радиолокаторы, аппаратуру обработки информации, систему передачи данных. Аппаратно унифицированы, аппаратура размещена в передвижных кузовах-фургонах. Внешний вид антенного поста ТРЛК-11 с защитным укрытием приведен на рис. 1 цветной вклейки. Максимальные дальности обнаружения цели с эффективной площадью 3 м^2

при условных вероятностях обнаружения не ниже 0,8 и ложной тревоги не выше 10^{-6} составляют 350 км (ТРЛК-11) и 150 км (АРЛК-11). Секторы обзора по азимуту составляют 360° , по углу места $0,5...45^\circ$, по высоте 20 тыс. м (ТРЛК-11) и 12 тыс. м (АРЛК-11). Ошибки измерения: по дальности 300 м и 150 м, по азимуту 9 мин в обоих случаях. Разрешающие способности: по дальности 800 м и 500 м для первичных каналов, 1000 м для вторичных каналов; по азимуту $1,5^\circ$ для первичных каналов, 3° для вторичных каналов. Скорости обзора — 6 и 12 об/мин соответственно. Коэффициенты подавления отражений от местных предметов — 40 и 45 дБ. Средние мощности передатчиков первичных радиолокаторов составляют $3,6 \text{ кВт} \times 2$ и $1,5 \text{ кВт} \times 2$. Потребляемые мощности от сети — 100...150 кВт и 90...120 кВт (в зависимости от состава оборудования). При перерыве питания на 1 с включается дизель-электростанция [105].

7.11.2. Радиолокационные станции обнаружения, наведения и целеуказания на средних и больших высотах

Радиолокационная станция 55Ж6-3 (Россия) — транспортируемая наземная трехкоординатная РЛС обнаружения и сопровождения воздушных целей *метрового* диапазона. Имеет отдельные антенные решетки для азимутальных и угломестных локационных измерений. Последняя закреплена вдоль мачты с тросовыми растяжками, устойчивой к ветровым нагрузкам и гололеду. Пределы работы составляют: по азимуту 360° , по дальности 1200 км, по высоте 75 км, по углу места 16° . Дальность обнаружения истребителя — 300...400 км на высотах 10...20 км. Точность измерения дальности — 400 м, азимута — $0,4^\circ$, высоты — 750 м. Скорость обзора — 6 об/мин. Время развертывания — 22 ч. Подавление помех от местных предметов — 45 дБ. Предусмотрена защита от помех в условиях организованного противодействия. Потребная мощность электропитания — 100 кВт, обслуживающий персонал — 4 чел. Внешний вид РЛС показан на рис. 2 цветной вклейки [93].

Радиолокационная станция «Противник-ГЕ» (Россия) — мобильная трехкоординатная РЛС обнаружения, наведения и целеуказания дециметрового диапазона с цифровой пространственно-временной обработкой сигналов. Имеет плоскую фазированную антенную решетку с цифровым преобразованием сигнала на антенне. Зона обнаружения: по дальности 10...400 км, по высоте до 200 км, по азимуту 360° , по углу места до 45° . Дальность обнаружения воздушной цели с эффективной площадью рассеивания $\sigma_{\text{ц}} = 1,5 \text{ м}^2$ равна 340 км. Ошибки измерения: по дальности не более 100 м, по высоте не более 450 м, по азимуту не более 12 мин, по углу места не более 10 мин. Разрешающие способности: по азимуту $2,5^\circ$, по углу места $1,6^\circ$. Скорость обзора — 6 и 12 об/мин. Число выдаваемых трасс целей за обзор

равно 150, число классов распознаваемых целей — 8. Коэффициент подавления отражений от местных предметов — не менее 50 дБ. Обеспечиваются: картографирование местных предметов, адаптивное подавление боковых лепестков диаграммы направленности антенны. Среднее время наработки на отказ — 600 ч. Потребляемая мощность — 100 кВт, средняя генерируемая мощность — 12 кВт, пиковая мощность генерируемых импульсов — 500 кВт, коэффициент шума приемника — 3 дБ. Время развертывания — не более 0,7 ч. Обслуживаемый персонал — 3 чел. Внешний вид РЛС показан на рис. 3 цветной вклейки [93].

Радиолокационная станция 64Л6Е или «Гамма-С1Е» (Россия) — мобильная трехкоординатная РЛС обнаружения, наведения и целеуказания *сантиметрового* диапазона с цифровой обработкой информации. Имеет плоскую одномерную ФАР, вращающуюся по азимуту. Зона обнаружения: по дальности 10...300 км (10...400 км в дополнительном режиме), по высоте до 30 км, по азимуту 360°, по углу места $-2...+30^\circ$ и $-2...+55^\circ$. Ошибки измерения: по дальности 50 м, по азимуту 15 мин, по углу места 10...15 мин, по высоте 400 м. Разрешающие способности: по дальности 250 м, по азимуту 1,4°. Скорость обзора — 6 об/мин. Число выдаваемых трасс целей за обзор равно 100. Среднее время наработки на отказ — 400 ч. Автоматизированная система контроля и диагностика обеспечивают среднее время восстановления — 0,5 ч. Время включения — 5 мин (экстренное — 3 мин). Время развертывания (свертывания) — 40 мин. Число основных транспортных единиц равно двум. Подавление отражений от местных предметов — 45 дБ. Средняя генерируемая мощность — 10...12 кВт. Потребляемая мощность — 70...90 кВт (лето-зима). Предусмотрено сопряжение со средствами защиты от противорадиолокационных ракет. Внешний вид РЛС показан на рис. 4 цветной вклейки [92, 93].

7.11.3. Радиолокационные станции обнаружения маловысотных целей

Радиолокационная станция 39Н6Е или «Каста-2Е2» (Россия) — мобильная, трехкоординатная, твердотельная РЛС *дециметрового* диапазона с цифровой обработкой информации, с антенной в виде усеченного параболоида. Зона обнаружения: по дальности 5...150 км, по азимуту 360°, по углу места 25°, по высоте 6 км. Дальности обнаружения целей, летящих на высоте 100 м, — 41 км и 55 км в зависимости от высоты подъема антенны (14 м и 50 м), а целей, летящих на высоте 1000 м, — 95 км. Ошибки измерения: по дальности 100 м, по азимуту 40 мин, плоскостных координат 900 м, по скорости 20 м/с. Различают три уровня высот полета целей: 0...2 км; 2...4 км; более 4 км. Разрешающие способности: по дальности 300 м, по азимуту 5,5°.

Скорости обзора — 6 и 12 об/с. Число сопровождаемых трасс — не менее 50. Среднее время наработки на отказ — 700 ч, среднее время восстановления доведено до 0,3 ч за счет автоматизированного допускового контроля с выдачей данных о местонахождении отказавших и запасных элементов. Продолжительность непрерывной работы — до 20 суток. Время включения составляет 3,3 мин, время развертывания — 20 мин. Подавление отражений от местных предметов — более 50 дБ. Приняты меры защиты от импульсных и шумовых активных помех: ручная или автоматическая перестройка рабочих частот в полосе частот 10 %, автоматическая смена М-кодов зондирующих импульсов с числом дискрет 127 и 255; стабилизация уровня ложных тревог; критериальная обработка принятых сигналов в сочетании с изменением частоты следования импульсов. Потребляемая мощность — не более 23 кВт. Внешний вид показан на рис. 5 цветной вклейки [92, 93].

7.11.4. Радиолокационные станции наведения зенитных управляемых ракет

Радиолокационная станция 30Н6Е (Россия) — многофункциональный радиолокатор подсвета и наведения (РПН), трехкоординатный, сантиметрового диапазона длин волн, с ФАР, когерентно-импульсный, используется в мобильном зенитном ракетном комплексе С-300 ПМУ-1. Обеспечивает автоматическое обнаружение и захват аэродинамических и баллистических целей, одновременное сопровождение шести целей, пуск и наведение 12 зенитных управляемых ракет (ЗУР) по шести целям, опознавание «свой—чужой». Может вести работу автономно или по внешнему целеуказанию. Дальность обнаружения целей с эффективной площадью 1 м^2 составляет 140 км с вероятностями правильного обнаружения 0,9 и ложной тревоги 10^{-6} . Максимальная скорость целей — 2800 м/с. Ошибки измерения: по дальности 3...5 м, по азимуту и углу места $0,02^\circ$. Разрешающие способности: по дальности 100 м, по скорости 5 м/с, по азимуту $1,3^\circ$ и по углу места $1,2^\circ$. Сектор обзора (угол места \times азимут) для обнаружения: аэродинамических целей на средних и больших высотах соответственно $14^\circ \times 64^\circ$ и $5^\circ \times 64^\circ$; маловысотных целей $1^\circ \times 90^\circ$; баллистических целей $10^\circ \times 32^\circ$. По данным целеуказания РПН осуществляет обнаружение цели в секторе допосиска ($4^\circ \times 4^\circ$, $2^\circ \times 2^\circ$), захват цели и переход на автоматическое сопровождение. Импульсная мощность — 75 кВт, средняя мощность — 10 кВт. Коэффициент шума приемника равен 7—9.

В антенной системе применяется поэлементное фазирование элементов ФАР, ширина диаграммы направленности по азимуту — $0,9^\circ$, по углу места — $0,8^\circ$. Используется квазинепрерывные зондирующие сигналы с

частотой повторения до 100 кГц. Потребляемая мощность — 130 кВт, масса РПН, включая оборудованное шасси, — 45500 кг, в том числе антенного поста — 11500 кг. Габариты РПН — $14,5 \times 3,14 \times 3,8$ м³. Внешний вид РПН показан на рис. 6 цветной вклейки [93, 106].

Радиолокационная станция 9С32 (Россия) — многоканальная станция наведения ракет (МСНР), трехкоординатная, сантиметрового диапазона, с ФАР, когерентно-импульсная, используется в самоходной зенитной ракетной системе С-300В. Проводит секторный поиск аэродинамических и баллистических целей, одновременно сопровождает до 12 целей и обеспечивает, кроме того, пуск и наведение 12 ЗУР по шести целям. Предусмотрены режимы работы:

- а) по целеуказанию (ЦУ) командного пункта системы;
- б) автономной работы (АР).

Сектор наблюдения целей в режиме ЦУ: 5° по азимуту и 6° по углу места, в режиме АР: $\pm 30^\circ$ по азимуту и $0^\circ \dots 18^\circ$ по углу места. Установка азимутов проводится путем поворота ФАР. Дальности обнаружения и автосопровождения самолетов-истребителей, летящих на высоте 5 км, — соответственно 150 км и 140 км в режиме ЦУ и 140 км и 130 км в режиме АР. Ошибки измерения в режиме автосопровождения самолетов-истребителей: $5 \dots 25$ м по дальности, $0,3 \dots 1,5$ м/с по скорости, $0,72 \dots 7,2$ мин по азимуту и углу места. Разрешающие способности: 100 м по дальности, 5 м/с по скорости, 1° по азимуту и углу места. Импульсная мощность — 150 кВт, средняя мощность — $10 \dots 13$ кВт, чувствительность приемника — 10^{-17} Вт. Антенная система включает:

а) пассивную ФАР, обеспечивающую формирование луча шириной 1° в обеих плоскостях и его фазовое управление;

- б) дополнительные антенны квадратурных автокомпенсаторов помех.
- Типы зондирующих сигналов:

а) *квазинепрерывный* с немодулированными и ЛЧМ радиоимпульсами в пачке — применяется для работы по данным ЦУ и автономной работы;

б) *импульсный с ЛЧМ* — применяется только в режиме автономного поиска.

Обработка квазинепрерывного сигнала — корреляционно-фильтровая с накоплением в узкополосных фильтрах промежуточной частоты. Формирование и обработка импульсного ЛЧМ сигнала — на дисперсионных линиях задержки [107]. Внешний вид зенитной ракетной системы С-300В показан на рис. 7 цветной вклейки [108].

7.11.5. Радиолокационные станции и комплексы разведки на поле боя

Радиолокационная станция «Кредо-1Е» (Россия) — РЛС обнаружения на поле боя средней дальности, в том числе для корректировки стрельбы

артиллерии. Дальность обнаружения солдат составляет 15...16 км, техники — до 40 км, корректировки стрельбы — до 13 км. Применение цветного дисплея позволяет наносить картографическую обстановку, выделять медленно и быстро движущиеся и неподвижные объекты различными цветами. Внешний вид РЛС показан на рис. 8 цветной вклейки [108].

Радиолокационная станция «Зоопарк-1» (Россия) — автоматизированный мобильный радиолокационный комплекс. Определяет координаты огневых позиций противника и корректирует стрельбу собственной артиллерии. Дальности разведки: минометов — 15 км, артиллерии — 10 км, реактивных систем залпового огня (РСЗО) — 20 км, тактических ракет — 35 км, приборная дальность обнаружения — 45 км. Сектор сканирования луча антенны — 60° по азимуту, 40° по углу места. Число одновременно сопровождаемых целей равно 12, пропускная способность — 20 траекторий/мин. Живучесть комплекса обеспечивается за счет кратковременности излучения, быстрой смены несущих частот, мер защиты от преднамеренных помех и воздействия электромагнитного импульса. Внешний вид комплекса показан на рис. 9 цветной вклейки [108, 109].

7.11.6. Радиолокационные станции подповерхностного зондирования

Георадар «Зонд» — наземный широкополосный (диапазон рабочих частот — 20...120 МГц) транспортируемый радиолокатор, обслуживаемый в труднодоступных районах одним оператором. Выявляет структурные неоднородности почвы: водоносные слои до глубин 15...30 м, слои торфа до глубин 0,5...10 м, подземные коммуникации, закрытые трубы, археологические объекты. Обеспечивает запись подпочвенных профилей на самописец и их наблюдение на цветном дисплее [6].

7.11.7. Радиолокационные станции противоракетной обороны

Радиолокационная станция «ДОН» (Россия) — многофункциональная РЛС сантиметрового диапазона с дальностью действия в несколько тысяч километров, предназначенная для обнаружения и сопровождения баллистических целей и наведения на них противоракет. Размещена в сооружении, имеющем вид усеченной пирамиды длиной и шириной по 100 м и высотой 45 м. На четырех сторонах последней установлены антенные решетки: круговые приемные — диаметром 16 м и прямоугольные передающие — 7×8 м². Приемные решетки могут формировать независимые группы лучей для моноимпульсных измерений. Передающее устройство генерирует импульсные сигналы различной длительности с различной внутриимпульсной модуля-

(ЛЧМ и фазовой манипуляцией). Высокоскоростная ЭВМ обеспечивает цифровую обработку сигналов на промежуточной частоте. Внешний вид РЛС «Дон» показан на рис. 10 цветной вклейки [93, 110, 111].

7.11.8. Корабельные РЛС

Радиолокационная станция AN/SPY-1A (США) — многофункциональная РЛС десятисантиметрового диапазона зенитной ракетной системы вооружения «Иджис» (США). Обеспечивает поиск, обнаружение и сопровождение 250...300 воздушных и надводных целей, способна наводить до 18 ЗУР, выдает данные об общей тактической обстановке в радиусе более 200 миль от корабля в его информационный центр. Дальность обнаружения высотных воздушных целей — 320 км, целей с высокой радиолокационной заметностью — 450 км. Дальность обнаружения маловысотных целей — 80...82 км (углы места от 0° до 4...5°). Ширина диаграммы направленности антенны ~1°. Ошибка углового сопровождения воздушной цели при ее скорости $M = 1$ и перегрузке $1g$ составляет 0,02...0,04 от ширины диаграммы направленности антенны. Частота обновления данных о воздушных целях — 1...15 Гц. Помехозащищенность обеспечивается перестройкой рабочей частоты, большой импульсной мощностью излучаемых сигналов 4 МВт, узкой диаграммой направленности антенны, возможностью быстрого перехода к режиму радиомолчания с последующим возобновлением работы. Частоты следования импульсов: 600 ± 100 Гц и 1430 ± 100 Гц длительности импульсов 0,4 мкс и 40 Гц при длительностях импульсов 20 и 40 мкс. Используется 4 плоских ФАР. Управление работой станции и необходимые расчеты осуществляются ЭВМ [112].

7.11.9. Авиационные (самолетные) РЛС

Радиолокационная станция «Жук» (Россия) — многофункциональный, многорежимный, всепогодный *трехсантиметровый* радиолокатор для самолетов МиГ-29, МиГ-25, Су-27 и их модификаций с антенной типа ФАР и цифровой обработкой сигналов. В режиме «воздух—воздух» обеспечивает: сопровождение 10 целей, одновременное наведение управляемого оружия на 2—4 цели, различные режимы воздушного боя, автоматическое огибание рельефа местности при полете на малых высотах. В режиме «воздух—земля» обеспечивает картографирование земной поверхности реальным лучом диаграммы направленности, доплеровское облучение этого луча, синтезирование апертуры и ее искусственную остановку, обнаружение и сопровождение движущихся наземных целей, определение высоты и скорости полета самолета. Дальность обнаружения целей в передней полусфере — 80 км,

зона обзора по азимуту $\pm 20^\circ$, $\pm 60^\circ$, $\pm 90^\circ$. Импульсная мощность РЛС — 5 кВт, средняя мощность — 1 кВт. Среднее время наработки на отказ — 120 ч, масса — 250 кг, объем — 800 дм³. Внешний вид РЛС показан на рис. 11 цветной вклейки [113].

Авиационный комплекс радиолокационного дозора и наведения А-50Э (Россия) включает самолет А-50Э (модификация Ил-76МД) и радиотехнический комплекс (РТК). Предназначен для обнаружения и определения государственной принадлежности (по принципу «свой—чужой») воздушных и морских целей, выдачи информации о них на командные пункты АСУ, наведения истребителей на воздушные цели и самолетов фронтовой авиации на наземные и морские цели. Основой РТК является трехкоординатная импульсно-доплеровская РЛС с цифровой СДЦ и внефюзеляжной антенной системой («гриб»). Дальность обнаружения воздушных целей: бомбардировщиков типа В-52 в режиме «выше горизонта» — 650 км, низколетящих истребителей типа МиГ-29 — до радиогоризонта, крылатых ракет (с эффективной площадью 1 м²) в режиме квазинепрерывного излучения — 215 км. Дальность обнаружения надводных кораблей (с эффективной площадью более 250 м²) — до радиогоризонта. Ошибки измерения координат сопровождаемых воздушных целей: плоскостных — до 2 км (относительно самолета А-50Э на дальности до 300 км); высоты — до 2,5 км (для дальности от 100 км до 200 км). Ошибка пеленгации постановщиков активных помех, работающих на частотах РЛС, — не более 0,75°. Число одновременно сопровождаемых воздушных целей — до 300, средний коэффициент проводки сопровождения воздушных целей равен 0,74. Число наводимых на воздушные цели истребителей — до 30. Дальность обмена телекодовой информацией с командным пунктом через специальный наземный приемопередающий центр — до 350 км по ММ-ДМВ радиолинии, до 2000 км по спутниковой радиолинии, до 2000 км по КВ радиолинии. Дальность симплексной связи между самолетом А-50Э и взаимодействующими истребителями по МВ-ДМВ радиолинии — до 450 км. Продолжительность патрулирования на удалении 1000 км от аэродрома взлета при взлетной массе 190 тонн — до 4 ч. Экипаж: летный — 4 чел., РТК — 11 чел. Внешний вид авиационного комплекса дозора и наведения показан на рис. 12 цветной вклейки [93].

Контрольные вопросы

1. Сформулируйте определение радиолокации.
2. Перечислите информационные задачи радиолокации.
3. Поясните основные виды радиолокации.
4. Приведите структурную схему простейшего совмещенного импульсного радиолокатора.
5. Каким образом измеряются дальность, радиальная скорость и угловые координаты целей в радиолокационных станциях?

6. Какие признаки используются для классификации РЛС?
7. Какие задачи решают наземные, корабельные, авиационные РЛС и РЛС космического базирования?
8. Какие характеристики РЛС относятся к тактическим, а какие к техническим?
9. Приведите обобщенную структурную схему РЛС. Какие системы, входящие в состав РЛС, являются основными, какие считаются дополнительными?
10. Как определяется эффективная площадь рассеяния целей?
11. От каких факторов зависит эффективная площадь рассеяния целей?
12. Перечислите основные направления снижения эффективной площади рассеяния воздушных целей.
13. Какие изменения происходят с зондирующим сигналом при отражении его от движущейся точечной цели?
14. Как определяются характеристики вторичного и обратного вторичного излучения целей?
15. Какие радиолокационные сигналы называют когерентными?
16. Дайте общую характеристику современных и дискретных (цифровых) методов обработки сигналов.
17. Поясните понятие группового запаздывания, а на этой основе понятия дисперсионной характеристики фильтра и эффекта сжатия ЛЧМ сигналов в согласованном фильтре.
18. Опишите аналоговые методы фильтровой согласованной обработки ЛЧМ сигналов.
19. Укажите особенности фильтров сжатия на поверхностных акустических волнах.
20. Поясните аналоговые методы фильтровой согласованной обработки фазоманипулированных сигналов и когерентных пачек радиоимпульсов с неизвестными временем запаздывания и доплеровской частоты.
21. Поясните особенности активных и пассивных помех, маскирующих и имитирующих помех.
22. Перечислите основные меры, позволяющие обеспечить электромагнитную совместимость радиолокационных систем.
23. Поясните способы создания активных и пассивных маскирующих помех.
24. Перечислите основные направления защиты РЛС от маскирующих активных помех.
25. Поясните методы некогерентной и когерентной компенсации маскирующих активных помех.
26. Поясните работу одноканального квадратурного автокомпенсатора.
27. Поясните работу одноканального гетеродинного автокомпенсатора.
28. Поясните особенности построения автокомпенсаторов при наличии мощных протяженных полезных сигналов.
29. Поясните принцип поляризационной адаптивной компенсации помех.
30. Приведите основные отличия сигналов целей и пассивных помех.
31. Поясните особенности практической реализации оптимального обнаружителя когерентной пачки радиоимпульсов на фоне пассивной помехи в виде стационарного небелого шума.
32. Поясните принцип построения гребенчатых фильтров подавления.
33. Поясните особенности построения гетеродинного автокомпенсатора пассивных помех.

8. СПУТНИКОВЫЕ РАДИОНАВИГАЦИОННЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ НА ИХ ОСНОВЕ

Описаны основные особенности построения и функционирования спутниковых радионавигационных систем (СРНС) второго поколения, а также реализованные на их основе информационные технологии. Дан краткий обзор сведений об основных элементах СРНС: орбитальной группировке навигационных космических аппаратов, наземном контрольно-измерительном комплексе и аппаратуре потребителей. Изложены основные принципы навигационно-временных определений, применяемые в этой аппаратуре. Проанализированы факторы, ограничивающие точность измерений, и методы ее повышения. Приведены примеры современных образцов аппаратуры и систем, реализующих рассмотренные принципы.

8.1. Введение

Одним из наиболее заметных и практически важных научно-технических достижений последнего десятилетия XX в. стало создание спутниковых радионавигационных систем (СРНС) второго поколения ГЛОНАСС (РФ) и GPS (США). Введенные в эксплуатацию в начале 90-х годов, эти системы первоначально предназначались для решения задач точного местоопределения высокодинамичных объектов военного назначения. Однако полученный в процессе испытаний и эксплуатации СРНС опыт показал, что создаваемое этими системами глобальное навигационно-временное поле позволяет обеспечивать не только высокоточную навигацию подвижных средств, но и решать на принципиально новой основе широкий круг других прикладных задач, которые при проектировании СРНС не рассматривались. Как следствие, возникли и быстро развиваются многочисленные приложения спутниковой радионавигации в различных сферах: от геофизических исследований до автоматического управления карьерными механизмами и автотранспортом.

В настоящее время более 200 фирм во всем мире выпускают сотни тысяч экземпляров аппаратуры потребителей (АП) СРНС различных клас-

сов, в том числе недорогие малогабаритные приборы для индивидуальных пользователей. В результате АП СРНС, наряду со средствами мобильной связи и другими новейшими достижениями радиоэлектроники, стала не только средством сугубо профессиональной деятельности, но и предметом личного пользования, что многократно расширяет круг ее потребителей.

В данной главе описаны основные особенности построения и функционирования СРНС второго поколения, а также реализованные на их основе информационные технологии.

Термин «навигация» (от лат. *navigo* — *плыву на судне*), появившийся в русском языке в петровскую эпоху, первоначально относился к морскому судовождению. В настоящее время этот термин охватывает широкий круг приложений, связанных с местоопределением различных объектов и обеспечением их движения по заданным траекториям. Близкие как по замыслу, так и по методам решения задачи возникают при топогеодезической привязке неподвижных объектов.

Источниками информации при решении навигационных задач могут быть различные факторы и явления естественного и искусственного происхождения: астрономические данные, параметры и законы изменения магнитного, гравитационного и электромагнитного полей, инерциальные явления в движущихся массах и т. п. Системы, не использующие для решения навигационных задач искусственные поля и данные, созданные внешними (не бортовыми) источниками, относятся к классу *автономных*, в противном случае система называется *неавтономной*. *Радионавигационными* в настоящее время принято называть системы, в которых для решения задачи местоопределения используются радиосигналы внешнего искусственного источника с известными координатами. Таким образом, радионавигационные системы относятся к классу неавтономных.

Принципиальная возможность использования для целей навигации радиосигналов стала очевидной практически одновременно с первыми удачными опытами по радиосвязи. Первые нашедшие практическое применение радионавигационные системы использовали принцип пеленгации источника радиоизлучения (радиомаяка) с помощью направленных антенн. Вначале пеленгация осуществлялась по максимуму диаграммы направленности простейших рамочных антенн, затем стали применять многоэлементные или сканирующие антенны, позволявшие реализовать более точный метод равносигнальной зоны.

Радиопеленгаторные системы получили широкое применение для судовождения, а затем и в аэронавигации. В настоящее время в мире используются десятки тысяч судовых и аэронавигационных радиомаяков, в том

числе с активным ответом, позволяющих определить не только направление на маяк, но и дальность до него.

Ограничения, присущие радиомаячным системам, использующим амплитудные методы пеленгации, обусловлены тем обстоятельством, что погрешность местоопределения в них быстро возрастает с увеличением расстояния до источника сигнала. Действительно, даже при допущении, что угловая погрешность пеленгации источника не зависит от дальности до него, соответствующая линейная погрешность возрастает пропорционально расстоянию. На практике угловая погрешность также возрастает вследствие уменьшения мощности сигнала, которая падает пропорционально квадрату дальности. В итоге при расстояниях порядка сотен километров и более точность амплитудного метода оказывается неприемлемо низкой.

Поэтому в 30-е годы XX в. были начаты разработки радионавигационных систем (РНС), использующих фазовые (радиоинтерферометрические) методы пеленгации. Большой вклад в развитие этого направления внесли советские ученые Л.И. Мандельштам, Н.Д. Папалекси и Е.Я. Щеголев. Аппаратура, реализующая предложенный и запатентованный указанными авторами «способ определения расстояния между двумя пунктами при помощи электромагнитных волн», основанный на измерении в точке наблюдения разности фаз сигналов, принимаемых от двух источников, нашла практическое применение. Так, в Великобритании была создана система Декка, обеспечивавшая судовождение при удалении от берега до 500 км. В частности, система Декка использовалась для управления кораблями и судами, участвовавшими в боевой операции по высадке десанта союзных войск в Нормандии.

Однако в полном смысле массовое применение РНС началось в период Второй мировой войны с созданием в США разностно-дальномерной системы Лоран-А. (Сходная по принципу действия система Консол была в этот же период создана в Германии, однако после завершения войны ее функционирование и дальнейшее развитие были прекращены.)

Принцип действия РНС Лоран-А состоит в измерении в точке приема разности времени прихода импульсных радиосигналов, моменты излучения которых двумя или более передающими станциями синхронизированы. Линией положения (геометрическим местом точек на поверхности Земли, для которых разность времени прихода пары сигналов постоянна) является гипербола, в фокусах которой расположены передающие станции, поэтому РНС данного класса иногда называют *гиперболическими*. Поскольку координаты передающих станций известны потребителю, для определения своего местоположения он должен, вычислив разность времени прихода сигнала от двух пар передающих станций, найти точку пересечения соответствующих двух гипербол.

Цепочка станций Лоран-А состояла из трех передающих станций: ведущей и двух ведомых. Принцип синхронизации системы предусматривал, что ведомые станции излучали сигналы в момент поступления на них импульса ведущей станции. Таким образом, устранялась неоднозначность измерений, связанная с тем, что в общем случае гиперболы могут иметь несколько точек пересечения. Станции располагались на расстоянии в несколько сот километров и использовали пять фиксированных частотных каналов в диапазоне средних волн 154...172 м. Дальность действия системы Лоран-А в зависимости от времени суток составляла 1000...2500 км, а погрешность местоопределения была 1...10 км. Для повышения точности использовался метод, основанный на приеме сигналов более чем двух пар станций и выборе гипербол, пересекающихся под углом, близким к прямому.

Дальнейшим развитием данного направления РНС явилась система Лоран-С. Ее основное отличие от системы Лоран-А состоит в том, что для измерения разности времени прихода сигналов в ней используется не только огибающая принятого импульсного сигнала, но и фаза колебаний его несущей частоты (100 кГц). Учет фазовой информации позволил уменьшить погрешность местоопределения примерно до 100 м. Дальность действия Лоран-С при высоте подъема антенн около 200 м достигает 4000 км. До появления спутниковой радионавигации система Лоран являлась наиболее массовым средством радионавигации, и до настоящего времени приемоиндикаторы этой системы входят в штатный комплект навигационного оборудования большинства надводных и воздушных судов. В СССР была создана существующая до настоящего времени гиперболическая система навигации Чайка, аналогичная Лоран-С.

Наряду с импульсно-дальномерными системами в послевоенные годы продолжали развиваться фазовые радионавигационные системы: уже упоминавшаяся выше Декка и разработанная в конце 60-х годов XX в. глобальная система Омега. Особенность последней состояла в том, что синхронизация всех передающих станций осуществлялась от единого эталона всемирного времени, а это обеспечило независимость их работы и повысило точность фазовых измерений. В состав системы Омега вошли восемь станций, расположенных в США, Японии, Норвегии, Аргентине, Либерии, на Гавайских островах и о. Реюньон. Высота мачт антенных систем этих станций составляет 350...400 м, что обеспечило дальность действия системы до 10 тыс. км. В это же время для обеспечения судовождения в арктических районах в СССР была введена в строй фазовая радионавигационная система высокой точности (РСВТ) с дальностью действия в пределах 1000 км, а также система дальнего действия РСДН-20 (Маршрут), работающая в сверхдлинноволновом диапазоне.

Принципы навигационных измерений, статистические методы обработки сигналов и другие технические решения, разработанные примени-

тельно к РНС наземного базирования явились научно-техническим фундаментом для проектирования спутниковых радионавигационных систем, в которых носителем источника навигационного сигнала является искусственный спутник Земли (далее всюду — навигационный космический аппарат, НКА). Возможность использования в качестве источника навигационного сигнала объекта, движущегося со скоростью порядка нескольких километров в секунду, базируется на том, что орбита НКА и параметры его движения могут прогнозироваться и контролироваться с весьма высокой точностью (см. далее), т. е. известны в любой момент времени.

Первые работы в области навигационного использования ИСЗ были опубликованы в 1957 г., одновременно с запуском первого ИСЗ. В 1958—1959 гг. были проведены работы, определившие технический облик первого поколения СРНС, который был реализован в советской низкоорбитальной системе Цикада. Сходные решения были использованы в американской низкоорбитальной СРНС Транзит.

Система Цикада включала в себя 4 НКА, система Транзит — 6 НКА, высота орбиты НКА этих систем составляла около 1000 км, период обращения — около 100 мин. При таких параметрах зона радиовидимости НКА для наземного потребителя имеет радиус порядка 2000 км, а время пребывания НКА в этой зоне — 5...15 мин, перерыв между сеансами наблюдения отдельных спутников лежит в пределах от 35 мин (в приполярных зонах) до 90 мин (вблизи экватора).

Суть используемого в этих системах метода навигационных определений, получившего название *интегрального доплеровского*, сводится к следующему. Пусть в точке приема аппаратура потребителя определяет доплеровский сдвиг несущей частоты сигнала НКА F_d и вычисляет интеграл

$$\int_{t_1}^{t_2} F_d(t) dt = -\frac{f_0}{c} \int_{t_1}^{t_2} v(t) dt = \frac{f_0}{c} (D(t_1) - D(t_2)) = \frac{f_0}{c} \Delta D,$$

где $t_2 - t_1$ — фиксированный интервал времени, f_0 — известная несущая частота сигнала НКА, c — скорость света.

Таким образом, значение интеграла от доплеровской частоты на интервале $\Delta t = t_2 - t_1$ пропорционально разности дальностей до НКА в моменты времени t_1 и t_2 . Поскольку координаты НКА на орбите в отсчетные моменты t_1 и t_2 известны (они передаются с НКА потребителю в составе так называемого *навигационного сообщения*; подробнее см. § 8.2), на основании значений ΔD можно рассчитать поверхности положения потребителя относительно НКА, имеющие вид гиперболоида. Получив второе значение ΔD_2 , относящееся к интервалу времени $\Delta t = t_3 - t_2$, рассчитывают вторую анало-

гичную поверхность положения, при этом положение наблюдателя определяется точкой пересечения этих гиперболоидов с земной поверхностью. Полученная точка, вообще говоря, не является единственной, т. е. решение навигационной задачи не является однозначным, но с учетом имеющихся у потребителя априорных данных о его местоположении может быть выбрано единственное правдоподобное решение. Поскольку период передачи навигационного сообщения равен 2 мин, за время одного сеанса измерений (10...15 мин) могут быть получены дополнительные поверхности положения, которые затем усредняются для повышения точности местоопределения. Успешный опыт эксплуатации СРНС Транзит и Цикада подтвердил перспективность спутниковой радионавигации как основной линии развития радионавигации в целом. В то же время стали очевидными принципиальные недостатки СРНС первого поколения.

Первый недостаток состоял в том, что наличие перерывов между сеансами наблюдения НКА исключало возможность местоопределения в произвольный момент времени или непрерывного местоопределения в течение длительного интервала времени. Уменьшить интервал между сеансами, увеличив количество НКА, было нельзя, поскольку все НКА использовали идентичные сигналы и при их одновременном наблюдении возникали взаимные помехи.

Второй недостаток связан с тем, что при интегральном доплеровском методе местоопределения необходимо результаты измерений, соответствующие различным моментам времени, приводить к одному моменту. Технические средства, существовавшие на момент создания СРНС первого поколения, обеспечивали точность синхронизации бортового эталона времени НКА и опорного генератора аппаратуры потребителя, соответствующую погрешности местоопределения неподвижного объекта порядка 50 м (СКО). Для движущихся объектов точность интегрального доплеровского метода измерений существенно зависит от точности определения собственной скорости движения потребителя. Так, погрешность измерения собственной скорости, равная 0,5 м/с, приводит к погрешности местоопределения порядка 500 м, т. е. по сравнению с неподвижным потребителем погрешность возрастает примерно в 10 раз.

Таким образом, СРНС первого поколения не удовлетворяли требованиям непрерывного высокоточного местоопределения динамических объектов в любой точке Земли с выдачей результатов измерений в реальном масштабе времени. Рассматривавшиеся на определенном этапе варианты расширения возможностей этих систем за счет увеличения числа спутников при одновременном увеличении высоты орбиты, а также за счет добавления к доплеровскому каналу дальномерного канала не обеспечивали радикального улучшения ситуации, поскольку перспективная СРНС должна предоставлять потребителю возможность, в любой момент времени определить три

координаты, вектор скорости и точное время. Для реализации такой возможности необходимо проводить измерения не менее чем по четырем спутникам, что накладывает соответствующее условие на структуру орбитальной группировки НКА: в любой момент времени в любой точке земной поверхности должна быть обеспечена возможность наблюдения не менее четырех НКА, при этом общее число НКА в группировке должно быть минимальным.

Для решения поставленной задачи при проектировании СРНС второго поколения были избраны среднеорбитальные НКА с высотой орбиты порядка 20 тыс. км (дальнейшее увеличение высоты орбиты нецелесообразно, так как не приводит к расширению зоны видимости НКА). Период обращения НКА при такой высоте орбиты равен примерно 12 ч. В этом случае, для того чтобы гарантировать в любой точке Земли одновременное наблюдение не менее 4 НКА, в составе орбитальной группировки должно быть не менее 18 НКА, однако для повышения точности и надежности навигационных измерений было решено увеличить это число до 24. Кроме изменения структуры и геометрии орбитальной группировки, в СРНС второго поколения с самого начала были заложены средства, обеспечивающие прецизионную (с точностью до единиц наносекунд) взаимную синхронизацию бортовых шкал (эталонов) времени НКА. Большое внимание было уделено также развитию средств высокоточного определения и прогнозирования параметров орбит (эфемерид) НКА. Достаточно сказать, что для достижения требуемой точности прогноза эфемерид должны учитываться такие факторы, как световое давление на НКА, влияние релятивистских эффектов на гравитационное поле Земли, неравномерность вращения Земли и ее полюсов, а также наличие реактивных сил, связанных с газоотделением материалов покрытия НКА. Решение перечисленных, а также и многих других второстепенных задач позволило создать, ввести в эксплуатацию и предоставить в пользование всему населению Земли две среднеорбитальные СРНС второго поколения: российскую, получившую название ГЛОНАСС (Глобальная навигационная спутниковая система) и американскую, первоначально названную NAVSTAR (Navigation Satellite Time and Ranging), а в настоящее время обычно именуемую GPS (Global Positioning System).

8.2. Основные принципы построения и функционирования СРНС

8.2.1. Обобщенная структурная схема СРНС

Принципиальная особенность СРНС, состоящая в том, что роль опорных точек в них играют НКА, совершающие орбитальное движение, определяет необходимость изменения структуры СРНС по сравнению с РНС наземного базирования. В отличие от последних, содержащих в качестве ос-

новых звеньев только аппаратуру базовых станций и АП, в составе СРНС необходима подсистема, обеспечивающая контроль траекторий движения НКА, дистанционный мониторинг исправности бортовой аппаратуры и точности бортовых эталонов времени, управление режимами работы бортовой аппаратуры, составом, объемом и скоростью передачи служебной информации и пр. Поэтому структура СРНС содержит три основные подсистемы:

- подсистему (орбитальную группировку) НКА;
- подсистему контроля и управления (контрольно-измерительный комплекс, КИК);
- подсистему аппаратуры потребителей.

Кроме указанных подсистем в структуру СРНС входят так называемые *средства функциональных дополнений*, т. е. специальная наземная аппаратура, используемая для обеспечения потребителям в определенном регионе или локальной области дополнительных возможностей, например повышения точности измерений за счет использования дифференциальных методов измерений (см. § 8.9).

Общие принципы взаимодействия отдельных подсистем и функциональных дополнений СРНС иллюстрируются структурной схемой, показанной на рис. 8.1.

Отличительные особенности построения подсистем НКА и КИК, присущие СРНС ГЛОНАСС и GPS, описаны ниже. Принципы построения и функционирования третьей подсистемы СРНС — аппаратуры потребителей — подробно рассматриваются в § 8.5.

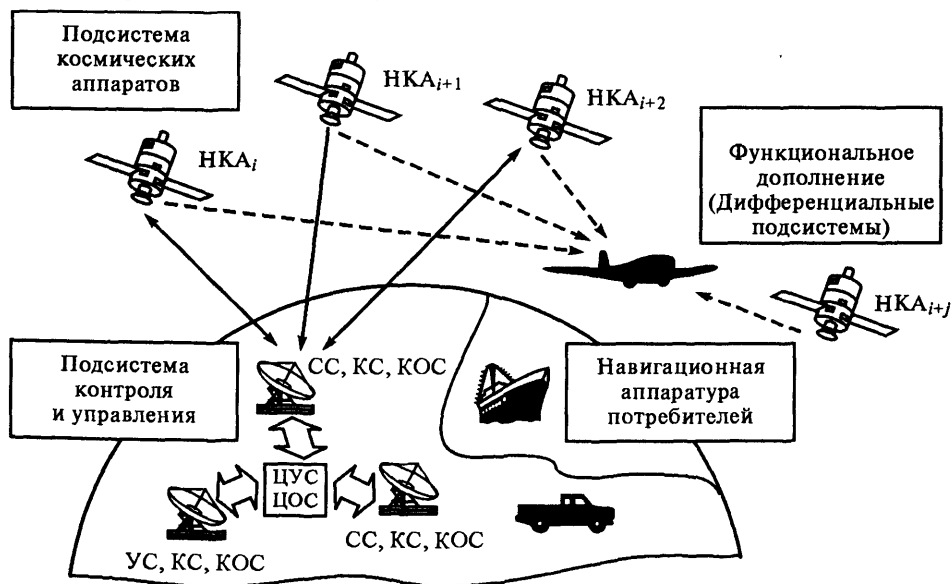


Рис. 8.1. Структурная схема СРНС

8.2.2. Подсистема НКА

Основное назначение НКА — формирование и излучение жестко синхронизированных сигналов, которые используются потребителями для навигационных определений, а назначение КИК — контроль функционирования бортовых систем и организация управления их работой. Соответственно, в состав бортовой аппаратуры входят передатчики и антенны навигационных сигналов и телеметрической информации, антенны и приемники данных и команд, передаваемых от КИК, бортовой эталон времени и частоты, блоки ориентирования, источники питания и различное вспомогательное оборудование.

Сигналы, излучаемые НКА обеих рассматриваемых СРНС, лежат в коротковолновой части дециметрового диапазона волн. При этом с целью уменьшения погрешностей в измерениях, возникающих при распространении сигналов от НКА через ионосферу (см. § 8.6), в СРНС ГЛОНАСС и GPS организовано излучение навигационных сигналов в двух полосах частот: 1,25 и 1,6 ГГц. Минимальный (для окологоризонтных спутников) энергетический потенциал сигналов, излучаемых НКА, составляет -160 дБ/Вт, что позволяет использовать в АП слабонаправленные антенны и умеренно сложные приемники. В обеих СРНС сигналы НКА имеют правую круговую поляризацию.

Навигационные сигналы НКА модулируются по фазе двумя видами псевдослучайных последовательностей (ПСП):

- широкополосной (длительность символа 0,1 или 1 мкс), используемой в АП для дальномерных измерений;

- узкополосной (длительность символа 20 мс), предназначенной для передачи потребителям *навигационного сообщения* (НС), т. е. массива навигационной и служебной информации, включающей в себя координаты НКА и их производные (эфемеридная информация), частотно-временные поправки для бортового эталона времени и частоты данного НКА, прогноз погрешностей, связанных с условиями распространения радиоволн, данные об исправности аппаратуры и т. п.

Отметим, что информация для НС, в основном, формируется на КИК, а затем транслируется на НКА для использования бортовыми системами и передачи потребителю.

Более подробно тип, структура и другие особенности сигналов СРНС, тесно связанные с используемыми методами навигационных измерений, описаны в соответствующих параграфах.

Структура и параметры орбитальной группировки НКА выбраны так, чтобы была обеспечена заданная кратность глобального покрытия поверхности Земли зонами видимости НКА, т. е. возможность круглосуточного

наблюдения в любой точке Земли числа НКА, необходимого для навигационных измерений (для используемых в СРНС ГЛОНАСС и GPS псевдодалномерных методов измерений минимально необходимое число НКА равно четырем; подробнее об этом см. § 8.4). Кроме того, учитывались соображения, связанные с геометрическим фактором точности (см. § 8.6), минимизацией интерференции сигналов от разных НКА в точке приема, и ряд других.

Рассмотрим более подробно структуры орбитальных группировок ГЛОНАСС и GPS.

Полная орбитальная группировка в российской среднеорбитальной СРНС второго поколения ГЛОНАСС содержит 24 штатных НКА, размещенных на трех круговых орбитах по восемь НКА в каждой (см. рис. 8.2).

Наклонение орбитальных плоскостей составляет $64,8^\circ$, долготы восходящих узлов (точек пересечения экваториальной плоскости с орбитой, при движении НКА с юга на север) различаются на 120° . В каждой орбитальной плоскости восемь НКА разнесены по аргументу широты на 45° , а аргументы широты НКА в каждой орбитальной плоскости сдвинуты относительно соседней на $\pm 15^\circ$ (см. рис. 8.3, а). Высоте круговых орбит, равной 19 100 км над поверхностью Земли, соответствует номинальный период обращения НКА $T = 11 \text{ ч } 15 \text{ мин } 44 \text{ с}$ [115, 116], не кратный периоду суточного вращения Земли. Орбиты с таким периодом, называемые *несинхронными*, менее чувствительны к возмущениям, обусловленным неоднородностью гравитационного поля Земли, по сравнению с синхронными круговыми орбитами ($T = 12 \text{ ч } 00 \text{ мин}$), используемыми в системе GPS (см. далее). Действительно, след синхронной орбиты НКА на поверхности Земли раз в сутки проходит через одни и те же точки, поэтому картина возмущения орбиты неоднородностями гравитационного поля будет повторяться для каждого

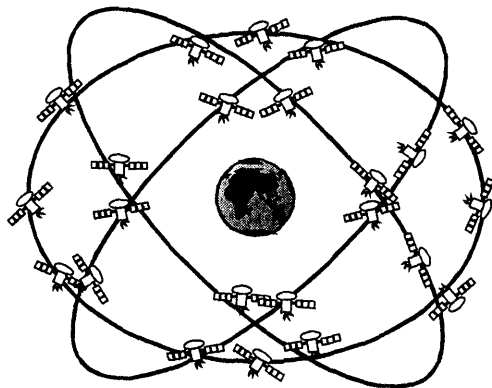


Рис. 8.2. Сеть НКА ГЛОНАСС на этапе полного развертывания

НКА, при этом заметно отличаясь от картины возмущений для всех других НКА. След же несинхронной круговой орбиты, благодаря вращению Земли, медленно (с периодом 8 суток) перемещается относительно поверхности Земли, при этом возмущения орбит для всех НКА в среднем будут практически одинаковы [116].

Полномасштабные летные испытания СРНС ГЛОНАСС начались в 1982 г. запуском спутника «Космос-1413». Первоначально

разрабатывавшаяся как система военного назначения, СРНС ГЛОНАСС в 1999 г. получила статус системы двойного (военного и гражданского) применения, ответственными за использование, поддержание и развитие которой являются Министерство обороны РФ и Российское авиационно-космическое агентство. Соответствующими постановлениями Правительства РФ от 1995 и 1999 гг. определено, что СРНС ГЛОНАСС в стандартном режиме можно бесплатно использовать в гражданских, коммерческих и научных целях любыми потребителями. Основным документом, устанавливающим взаимоотношения потребителей с системой, является Интерфейсный контрольный документ ГЛОНАСС. Развертывание орбитальной группировки до полного состава — 24 НКА — завершилось в 1995 г., однако по причине выработки ресурса НКА на начало 2003 г. действующий состав орбитальной группировки ГЛОНАСС составлял 10 НКА. В ближайшие годы планируется произвести запуск необходимого количества НКА и таким образом полностью восстановить и в дальнейшем поддерживать работоспособность системы [114].

Согласно разработанной программе модернизации СРНС ГЛОНАСС для восполнения орбитальной группировки будут применяться модернизированные НКА (ГЛОНАСС-М), основные отличия которых от НКА первой модификации состоят в следующем [117].

1. Используемый в настоящее время бортовой эталон времени, имеющий СКО среднесуточных значений частоты порядка $(2...4)10^{-13}$, будет заменен на более совершенный, имеющий СКО не более 10^{-13} , что

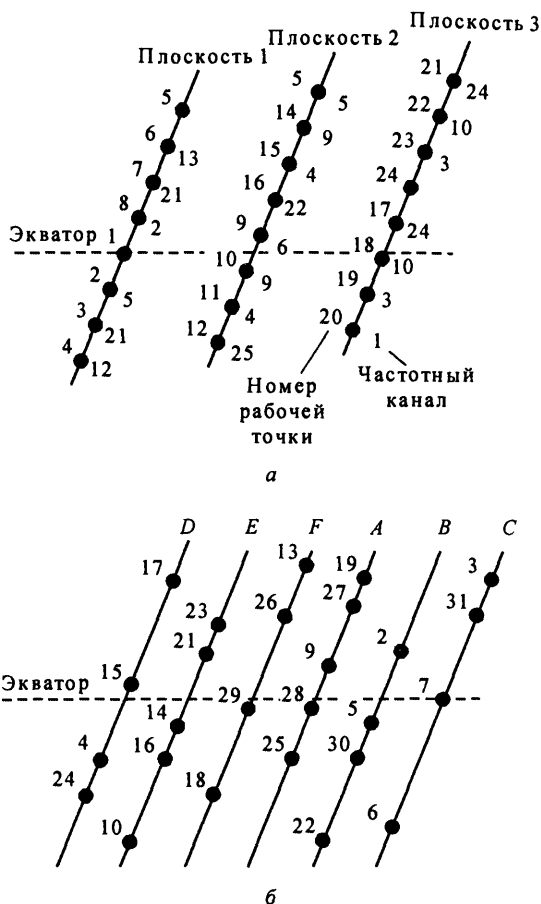


Рис. 8.3. Размещение НКА СРНС по орбитам:

а — ГЛОНАСС; б — GPS

позволит (при соответствующем снижении методических погрешностей) обеспечить величину погрешности полусуточного прогноза бортовой шкалы времени в пределах 5 нс (СКО), а суточного прогноза — в пределах 7 нс (СКО).

2. Открытый (коммерческий) дальномерный код будет передаваться как в диапазоне $L_1 = 1,6 \dots 1,62$ ГГц, так и в диапазоне $L_2 = 1,25 \dots 1,27$ ГГц (в настоящее время сигнал диапазона L_2 модулируется только высокоточным дальномерным кодом, не доступным для коммерческого использования; подробнее об этом см. § 8.3). В состав навигационного сообщения будет включена информация о разнице аппаратурных задержек дальномерных кодов в диапазонах L_1 и L_2 , а также признак ожидаемой секундной коррекции шкал времени.

3. Будут приняты меры по снижению уровня возмущений орбиты НКА.

Перечисленные и другие изменения позволят повысить точность и надежность навигационных измерений СРНС ГЛОНАСС.

В США разработки концепции среднеорбитальной СРНС второго поколения были начаты в 70-е годы XX в. Полномасштабные работы по созданию и испытанию СРНС, в то время называвшейся NAVSTAR, начались в 1978 г. запуском четырех НКА. Полное развертывание СРНС, в настоящее время официально именуемой GPS, завершилось в 1995 г.

Так же, как и СРНС ГЛОНАСС, GPS имеет статус системы двойного применения, за ее военное использование ответственность несет Министерство обороны, а за коммерческое — Министерство транспорта США. Основным документом, регламентирующим использование системы, является Интерфейсный контрольный документ GPS [118], распространение которого возложено на компанию ARINC.

Штатная орбитальная группировка СРНС GPS состоит из 24 основных НКА, расположенных на 6 круговых орбитах. Дополнительно на некоторых орбитах может находиться один резервный НКА, предназначенный для сохранения параметров системы при выходе из строя основных НКА (см. рис. 8.3, б). Наклонение орбитальных плоскостей составляет 55° , долготы восходящих узлов различаются на 60° . Высоте орбит около 20 тыс. км соответствует период обращения $T = 12$ ч 00 мин, т. е. орбиты НКА GPS являются *синхронными*.

В процессе создания и эксплуатации СРНС GPS НКА подвергались неоднократной модернизации. Опытные НКА №1—12 (Блок-1) были сконструированы так, чтобы обеспечивать в течение 3—4 дней заданную точность местоопределения без контакта с контрольно-измерительным комплексом. Навигационные космические аппараты, предназначенные для регулярной эксплуатации, получили наименование Блок-II, Блок-IIA, Блок-IIR.

Навигационный космический аппарат Блок-II (№ 13—21) должен был обеспечивать местоопределение без контакта с КИК в течение 14 дней.

Перспективный НКА Блок-IIR (№ 41—66) должен обеспечивать местоопределение без контакта с КИК по крайней мере в течение 14 дней при работе совместно с НКА Блок-IIА и в течение, как минимум, 180 дней с работающей системой автономной навигации (Autonav) при работе только спутников Блок-IIR. Для этого на бортах НКА Блок-IIR размещаются средства межспутниковой связи, обмена данными и измерения взаимных расстояний, позволяющие проводить автономную синхронизацию бортовой ШВ и автономное уточнение параметров орбит. Именно эти средства обеспечивают автономную работу СРНС в течение 180 дней без существенного ухудшения точности определения координат (СКО не более 16 м) с сохранением характеристик по устойчивости функционирования. Заметим, что режим Autonav может быть реализован лишь после запуска необходимого числа НКА Блок-IIR.

8.2.3. Контрольно-измерительный комплекс

Подсистема КИК решает следующие основные задачи:

- контроль параметров движения НКА, прогнозирование параметров орбит (эфемерид) и передача эфемеридной информации на борт НКА для ввода в бортовой компьютер и включения в навигационное сообщение, передаваемое потребителям;

- определение и прогноз для каждого НКА отклонений бортовой шкалы времени, формируемой бортовым эталоном, относительно общесистемной ШВ, формируемой наземным эталоном, и передача их на борт для коррекции бортовой ШВ и включения в навигационное сообщение;

- мониторинг качества радионавигационных сигналов, излучаемых НКА, с целью выявления возможных неисправностей и отклонений в работе бортовой аппаратуры;

- радиотелеметрический контроль и диагностика состояния бортовой аппаратуры НКА;

- программное и командное управление полетом НКА и обеспечение функционирования бортовых систем.

Для решения перечисленных и других задач в состав КИК входят следующие элементы:

- центр управления системой;
- центральный синхронизатор;
- станции слежения;
- радиотехнические контрольные станции;
- квантово-оптические контрольные станции;
- аппаратура контроля поля.

Остановимся более подробно на методах решения задач эфемеридного и частотно-временного обеспечения, которые непосредственно влияют на точность навигационных измерений с помощью сигналов СРНС ГЛОНАСС [116].

Эфемеридное обеспечение включает процедуры определения и прогноза параметров движения НКА и трансляцию соответствующей информации на борт НКА для передачи в составе навигационного сообщения.

Частотно-временное обеспечение включает процедуры определения, прогноза и передачи на борт НКА частотно-временных поправок бортовой ШВ относительно ШВ системы.

Определение и прогноз параметров движения НКА осуществляет Баллистический центр системы на основе результатов траекторных измерений дальности и радиальной скорости НКА, поступающих от сети наземных радиотехнических «запросных» контрольно-измерительных станций (КИС). При расчетах используются данные не менее трех контрольно-измерительных станций, расположенных на территории России (западная, центральная, восточная) в пределах 50...60° северной широты. При таком расположении станций на каждом витке орбиты НКА наблюдается при углах возвышения не менее 5° в течение сеансов длительностью 1...5 ч.

Сформированная в Баллистическом центре эфемеридная информация передается на борт НКА через сеть контрольно-измерительных станций ежесуточно. Среднеквадратическая погрешность эфемеридных данных для наименее благоприятных условий составляет по высоте 5 м, вдоль орбиты 20 м, по нормали к плоскости орбиты 10 м. При благоприятных условиях погрешности по высоте и вдоль орбиты приблизительно в два раза меньше.

Подсистема частотно-временного обеспечения содержит:

- наземный центральный синхронизатор на основе водородного стандарта частоты с относительным среднеквадратическим отклонением среднесуточных значений частоты не более $3 \cdot 10^{-14}$;
- радиотехническую «беззапросную» измерительную станцию;
- радиотехническую «запросную» измерительную станцию.

Центральный синхронизатор формирует ШВ системы и опорные сигналы для «беззапросной» измерительной станции.

Принцип определения поправок бортовой ШВ состоит в следующем [117]. На «беззапросной» измерительной станции для i -го НКА производится измерение сдвига по времени τ_i широкополосной ПСП, модулирующей навигационные радиосигналы 1600 МГц и 1250 МГц, относительно опорного сигнала. На основании этих измерений определяют *псевдодальность* $\hat{R}_i = c\tau_i$ этого НКА, отличающуюся от истинной дальности R_i на неизвестную, но постоянную за время определения навигационных параметров величину $\delta R_i = c\delta\tau_i$, где $\delta\tau_i$ — сдвиг ШВ i -го НКА относительно ШВ системы.

«Запросная» измерительная станция, вычисляя интервал времени τ_{iz} между излучением запросного сигнала и приемом сигнала бортового ответчика i -го НКА, определяет расстояние до НКА $R_{iz} = c\tau_{iz}$, погрешность которой не зависит от бортовой ШВ. Результаты одновременных измерений беззапросной и запросной измерительных станций поступают в Баллистический центр, где определяется значение сдвига ШВ i -го НКА относительно ШВ системы: $\delta\tau_i = (\hat{R}_i - R_{iz})/c$.

Опыт эксплуатации СРНС ГЛОНАСС показал, что при собственной среднесуточной нестабильности бортовой ШВ порядка $3 \cdot 10^{-13}$ погрешность (СКО) взаимной синхронизации бортовых ШВ любых двух НКА равна 20 нс (более подробно об этом см. § 8.6).

Контрольно-измерительный комплекс GPS состоит из сети наземных станций слежения, расположенных по всему миру. Сеть включает главную (ведущую) станцию, контрольные станции и три земные станции ввода данных на НКА. Главная станция контроля и управления находится на авиабазе Фалкон (Шривер) ВВС США в районе г. Колорадо-Спрингс, штат Колорадо. Основу главной станции составляет центр управления с вычислительным комплексом и средства передачи данных на земную станцию связи с НКА [117]. Контрольные станции в настоящее время размещены на атолле Диего-Гарсиа (архипелаг Чагос в Индийском океане), на о. Вознесения (в Атлантическом океане), на Гавайях и атолле Кваджалейн (в Тихом океане); одна контрольная станция совмещена с главной станцией [117]. Тот факт, что станции расположены сравнительно равномерно по Земному шару вблизи экватора, создает благоприятные условия для наблюдений НКА.

Перечисленные станции принимают сигналы спутников GPS и осуществляют специальные прецизионные измерения дальности до НКА, по которым на главной станции осуществляют точные расчеты параметров орбит, ионосферной модели и корректирующих поправок для бортовых часов. Для точного определения орбит НКА используют беззапросный метод. По соответствующим измерениям задержки дальномерных сигналов НКА осуществляется точное определение параметров орбит и параметров движения НКА. Через земные станции связи (атолл Диего-Гарсиа, о. Вознесения; атолл Кваджалейн) главная станция передает на борт каждого НКА эти параметры совместно с данными обработки метеорологической информации, позволяющей уточнить параметры модели тропосферы. Производится также мониторинг состояния НКА и управление их работой. Сегмент управления устанавливает шкалу времени GPS, которая привязана к шкале времени UTC (шкала Универсального координированного времени), поддерживаемой Военно-морской обсерваторией США. Начало отсчета времени установлено в полночь с 5 на 6 января 1980 г. Самой крупной единицей времени GPS является одна неделя, которая состоит из 604 800 с. Отсчет времени в GPS может

отличаться от UTC, поскольку ШВ GPS является непрерывной, а ШВ UTC может корректироваться на целое число секунд. В результате между ними имеется некоторое постоянно растущее расхождение. Передаваемые с НКА навигационные данные содержат следующую информацию о расхождениях шкал времени:

- накопленное расхождение в целых секундах между GPS и UTC;
- данные для вычисления текущих поправок к бортовой ШВ НКА.

Точность этих данных такова, что погрешность алгоритмической привязки шкалы времени GPS к UTC находится в пределах 90 нс (СКО).

8.3. Структура сигналов и сообщений СРНС

При выборе типов и параметров сигналов, используемых в СРНС, необходимо учитывать комплекс требований и условий, подчас противоречащих друг другу. Прежде всего эти сигналы должны обеспечивать высокую точность измерения времени прихода (задержки) сигнала и его доплеровской частоты и высокую вероятность правильного декодирования навигационного сообщения. Для того чтобы сигналы отдельных спутников надежно различались АП, они должны иметь низкий уровень взаимной корреляции. Кроме того, сигналы СРНС должны наилучшим образом использовать отведенную полосу частот при малом уровне внеполосного излучения, обладать высокой устойчивостью к преднамеренным и непреднамеренным помехам различных видов. Выполнить эти требования можно только при использовании сложных сигналов, база которых (произведение полосы на длительность) превышает единицу. В качестве таковых в обеих рассматриваемых СРНС используют сигналы с псевдослучайной фазовой манипуляцией. Имеющиеся различия в видах модулирующих ПСП, методе разделения сигналов НКА и способах передачи навигационного сообщения описаны далее.

8.3.1. Структура навигационных сигналов и навигационных сообщений ГЛОНАСС

В спутниковой радионавигационной системе ГЛОНАСС применен частотный метод разделения сигналов различных НКА: каждый из них использует свою пару литерных частот, одна из которых принадлежит диапазону L_1 , другая — диапазону L_2 [114]. На этапе проектирования для СРНС ГЛОНАСС был принят следующий способ определения номинальных значений несущих частот в верхнем (L_1) и нижнем (L_2) диапазонах:

$$\begin{aligned} f_{1,k} &= f_{1,0} + k\Delta f_1, & f_{1,0} &= 1602,0000 \text{ МГц}, & \Delta f_1 &= 0,5625 \text{ МГц}, \\ f_{2,k} &= f_{2,0} + k\Delta f_2, & f_{2,0} &= 1246,0000 \text{ МГц}, & \Delta f_2 &= 0,4375 \text{ МГц}, \\ & & f_{1,k} / f_{2,k} &= 9/7, \end{aligned}$$

где k — литера (условный порядковый номер) пары несущих частот $f_{1,k}$ и $f_{2,k}$ в диапазонах L_1 и L_2 . При этом крайним значениям литер $k = 1$ и $k = 24$ соответствовали следующие значения несущих частот:

$$\begin{aligned} f_{1,1} &= 1602,5625 \text{ МГц}, & f_{1,24} &= 1615,5000 \text{ МГц}, \\ f_{2,1} &= 1246,4375 \text{ МГц}, & f_{2,24} &= 1256,5000 \text{ МГц}. \end{aligned}$$

Однако, поскольку в непосредственной близости от рабочих частот ГЛОНАСС располагаются рабочие частоты систем спутниковой радиосвязи (полосы частот 1559,0...1610,0 МГц и 1215,0...1260,0 МГц), воздушной радионавигации (полоса частот 1535,0...1559,0 МГц), а также выделенная для нужд радиоастрономии полоса частот 1610,6...1613,8 МГц, план использования радиочастот в интересах ГЛОНАСС в дальнейшем пересматривался.

В частности, было учтено, что даже при полном составе орбитальной группировки (24 НКА) для большинства потребителей (сухопутных, морских, воздушных) в зоне радиовидимости не может находиться более 12 НКА. Поэтому в целях сокращения ширины диапазона, занимаемого сигналами СРНС ГЛОНАСС, в 1993 г. было принято решение для взаимно антиподных (находящихся в диаметрально противоположных точках орбиты) НКА использовать одинаковые литерные частоты, число которых в каждом диапазоне частот сокращено до 12. При этом в 1998—2005 гг. будут использоваться литеры несущих частот $k = 1, \dots, 12$, а с 2005 г. литеры $k = -7, \dots, 4$. Это обстоятельство необходимо иметь в виду, если АП используется на других космических аппаратах, высота орбиты которых позволяет одновременно «видеть» взаимно антиподные НКА. Для разделения сигналов в этих ситуациях необходимо использовать пространственную (за счет формы диаграммы направленности антенны АП) и доплеровскую селекцию сигналов. Возможность последней следует из того факта, что для наземных объектов доплеровский сдвиг, обусловленный собственным движением НКА, может иметь максимальные значения ± 5 кГц, а для низкоорбитальных космических объектов — до ± 40 кГц. В аппаратуре потребителей поиск несущей частоты каждого НКА осуществляется в полосе порядка ± 500 Гц относительно центральной частоты, прогнозируемой с учетом доплеровского сдвига. Таким образом, в АП космических объектов может быть реализована эффективная доплеровская селекция сигналов от всех радиовидимых НКА, в том числе взаимно антиподных НКА с одинаковыми литерными частотами.

Сигнал, излучаемый в диапазоне L_1 , подвергается относительной фазовой манипуляции (ОФМ) на π двумя ПСП: ПСП1 и ПСП2. Первая ПСП образуется сложением по модулю 2 трех двоичных сигналов (рис. 8.4):

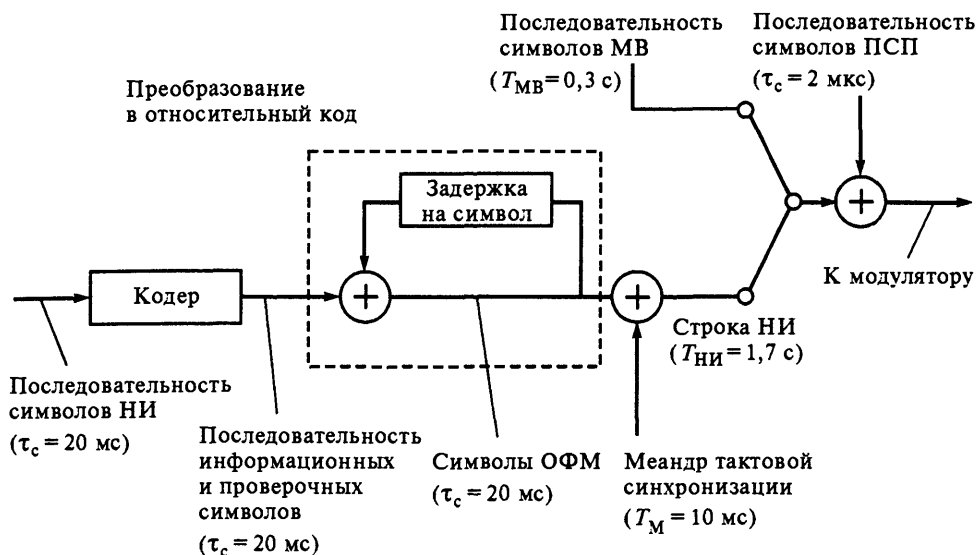


Рис. 8.4. Принцип формирования модулирующей последовательности

— дальномерного кода стандартной точности (СТ), являющегося последовательностью максимальной длины (М-последовательностью) с тактовой частотой $f_T = 511 \text{ кГц}$ и периодом 1 мс (иногда для обозначения этого кода используется аббревиатура ПТ-код, т. е. код пониженной точности);

— двоичного сигнала навигационной информации (НИ) с тактовой частотой 50 Гц , передаваемого в виде строк длительностью 2 с ;

— сигнала тактовой синхронизации в виде меандра с частотой 100 Гц .

Образующий полином СТ-кода имеет вид $G_d(x) = 1 + x^5 + x^9$.

Схема формирования дальномерного кода показана на рис. 8.5. Основным элементом этой схемы являются: регистр сдвига генератора ПСП, формирователь синхроимпульсов, а также триггеры синхронизации для формирования синхроимпульсов с периодом $T_{СИ} = 1 \text{ мс}$ и меандра тактовой синхронизации.

В каждой двухсекундной строке навигационного сообщения на интервале времени $T_{НИ} = 1,7 \text{ с}$ передаются 85 двоичных символов длительностью $\tau_c = 20 \text{ мс}$ каждый. В конце двухсекундной строки передается метка времени (МВ), представляющая собой укороченную на один символ 31 -символьную М-последовательность длительностью $T_{МВ} = 0,3 \text{ с}$. В приемнике с помощью меандра осуществляется символьная синхронизация для МВ, а с помощью МВ — строчная и символьная синхронизация НИ. Фронты меандра и границы символов МВ и НИ синхронизированы с высокой точностью (см. рис. 8.6).

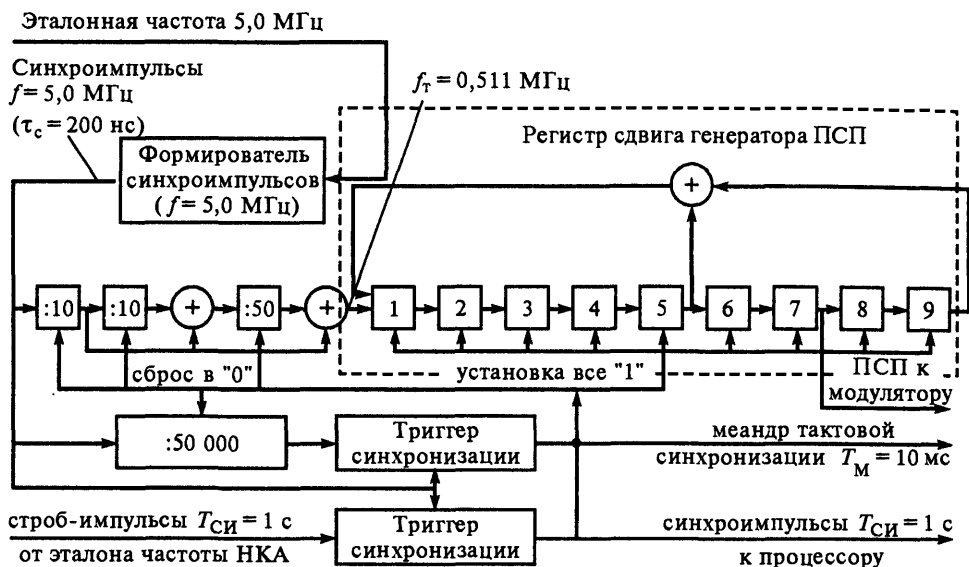


Рис. 8.5. Схема формирования дальномерного кода

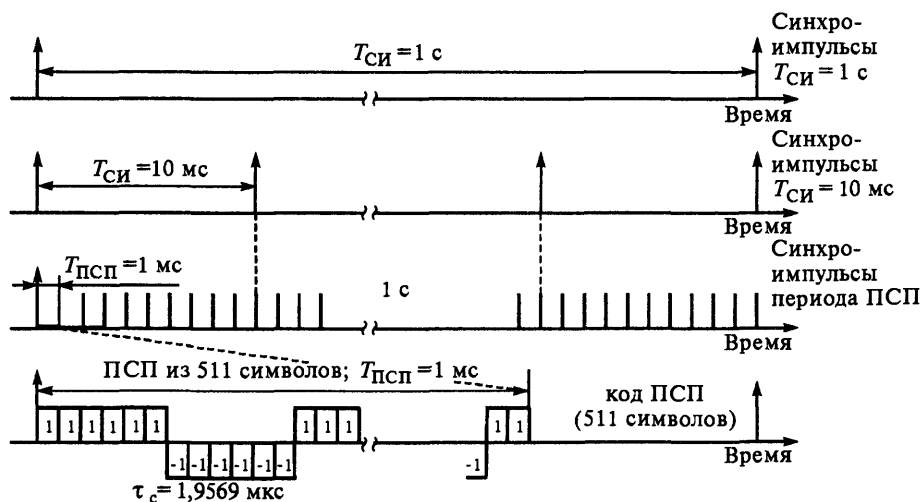


Рис. 8.6. Временная диаграмма сигналов СРНС ГЛОНАСС

Вторая ПСП формируется по аналогичному принципу, однако в ней используется М-последовательность с тактовой частотой 5,11 МГц и периодом 1 мс, образующая дальномерный код высокой точности (ВТ). Использование ВТ-кода регламентируется МО РФ, поэтому он защищен от несанкционированного доступа и его структура официально не опубликована. Ор-

тогональность, т. е. отсутствие взаимного влияния при передаче ПСП1 и ПСП2 обеспечивается благодаря их сдвигу по фазе несущей на 90° .

Спектр мощности навигационного радиосигнала практически полностью определяется тактовой частотой ПСП, которая на несколько порядков превышает тактовую частоту НИ. Соответственно, ширина основного лепестка огибающей спектра мощности радиосигнала, модулированного СТ-кодом равна 1,022 МГц, а ВТ-кодом — 10,22 МГц.

Навигационная информация структурирована в виде строк, кадров и суперкадров. Строка, как уже говорилось, имеет длительность 2 с (вместе с МВ). Поскольку используется относительная фазовая манипуляция, первый символ каждой строки является начальным («холостым»). Последние восемь символов в каждой строке являются проверочными символами кода Хемминга, позволяющими исправлять одиночный ошибочный символ и обнаруживать два ошибочных символа в строке. Кадр (длительность 30 с) содержит 15 строк, суперкадр (длительность 2,5 мин) содержит 5 кадров.

В составе каждого кадра передается полный объем оперативной НИ и часть альманаха системы. Полный альманах передается в пределах суперкадра.

Альманах системы содержит:

- параметры орбиты, номер пары несущих частот и поправку к бортовой ШВ для каждого штатного НКА орбитальной группировки (24 НКА);
- время, к которому относится информация альманаха;
- поправку к ШВ системы относительно ШВ страны (погрешность поправки не более 1 мкс).

Альманах системы необходим в АП для планирования сеанса навигации при выборе оптимального созвездия НКА и для прогноза доплеровского сдвига несущей частоты.

Оперативная НИ в кадре относится к НКА, излучающему навигационный радиосигнал, и содержит: признаки достоверности НИ в кадре; время начала кадра t_k ; эфемеридную информацию (ЭИ) — координаты и производные координат НКА в прямоугольной геоцентрической системе координат на момент времени t_0 ; частотно-временные поправки (ЧВП) на момент времени t_0 в виде относительной поправки к несущей частоте навигационного радиосигнала и поправки к бортовой ШВ НКА; время t_0 , к которому «привязаны» ЭИ и ЧВП, кратное 30 мин, от начала суток. Оперативная НИ используется потребителем в процессе определения собственных координат и вектора скорости.

В диапазоне L_2 используемые в настоящее время НКА первой модификации излучают однокомпонентный сигнал, образуемый посредством фазовой манипуляции несущей М-последовательностью ВТ-кода (см. выше), без передачи НИ. На навигационном космическом аппарате второй модифи-

рации излучаемый в диапазоне L_2 навигационный сигнал 1250 МГц будет так же, как сигнал диапазона L_1 , содержать две ортогональных компоненты: М-последовательность с тактовой частотой 511 кГц (СТ-код) и М-последовательность с тактовой частотой 5,11 МГц (ВТ-код). Передача НИ в диапазоне L_2 не планируется.

8.3.2. Структура навигационных сигналов и навигационных сообщений GPS

В отличие от СРНС ГЛОНАСС, в СРНС GPS используется метод кодового разделения сигналов, излучаемых на общих для всех НКА несущих частотах $L_1 = 1572,42$ МГц и $L_2 = 1227,6$ МГц. Сигналы на частоте L_1 модулируются двумя ортогональными (сдвинутыми относительно фазы несущей на 90°) ПСП, образованными суммированием по модулю 2 дальномерного кода и информационной последовательности, содержащей навигационное сообщение и другую служебную информацию. Дальномерные коды для каждого НКА являются индивидуальными, причем их структура специально выбрана из условия малого уровня их взаимной корреляции.

Основным видом дальномерного кода GPS является точный код (Р-код), представляющий собой ПСП с тактовой частотой 10,23 МГц и периодом 7 суток; Р-код GPS предназначен для санкционированного доступа, поэтому его структура не публикуется. Для навигационного космического аппарата Блок-II и его последующих модификаций предусмотрен режим дополнительной защиты точного кода, при котором вместо Р-кода используется так называемый Р(У)-код, получаемый дополнительным нелинейным преобразованием Р-кода. Несанкционированное использование такого кода при отсутствии ключевой информации практически невозможно, поскольку в настоящее время не существует вычислительных средств, способных расшифровать его.

Для коммерческого несанкционированного использования предоставлен «короткий» С/А-код пониженной точности (от англ. Clear Acquisition), используемый также для синхронизации аппаратуры потребителя, работающей по Р- или Р(У)-кодам. Код С/А представляет собой код Голда, образуемый из двух 10-разрядных ПСП с образующими полиномами

$$G_1(x) = 1 + x^3 + x^{10}, \quad G_2(x) = 1 + x^2 + x^3 + x^6 + x^8 + x^9 + x^{10}.$$

Код, соответствующий i -му НКА, образуется суммированием по mod2 упомянутых ПСП, причем вторая ПСП имеет индивидуальный для данного НКА сдвиг на фиксированное число символов (длительность символа 1 мкс).

Необходимо отметить, что на протяжении ряда лет C/A-код подвергался дополнительному зашумлению, ограничивающему точность измерений с помощью коммерческой АП (так называемый режим ограниченного доступа S/A, от англ. Selective Availability). Режим S/A был отключен 1.05.2000 г., однако МО США оставляет за собой право по собственному усмотрению вновь вводить его, чтобы исключить использование коммерческой АП для высокоточных навигационных определений в военных целях.

Вспомогательная частота L_2 модулируется только закрытым P-кодом.

Состав информации, передаваемой в навигационном сообщении GPS, сходен с составом НИ ГЛОНАСС: в нем содержится информация об эфемеридах НКА, альманахах созвездия НКА, частотно-временные поправки, метки времени, параметры ионосферной модели, сведения о работоспособности бортовой аппаратуры НКА и др.

Информационная последовательность передается кадрами емкостью 1500 бит и длительностью 30 с. Один кадр делится на субкадры (строки) длительностью 6 с, содержащие 10 слов по 30 бит каждое. В одном кадре передается 1/25 всего альманаха. Поэтому передача всей информации альманаха занимает 12,5 мин. Этот массив информации объемом 37500 бит называют *суперкадром*.

8.4. Методы измерения навигационных параметров

8.4.1. Основные понятия

Основным содержанием задачи навигационно-временных определений (НВО) в СРНС является определение *вектора состояния* потребителя. В качестве компонент этого вектора обычно рассматривают пространственные координаты x, y, z потребителя, временную поправку τ_n шкалы времени потребителя относительно системной шкалы, а также составляющие вектора скорости $\dot{x}, \dot{y}, \dot{z}$. Существенно то, что перечисленные величины не могут быть непосредственно измерены радиотехническими методами. Поэтому в АП СРНС реализован косвенный метод: измеряются некоторые параметры принятого радиосигнала, а именно время его прихода t , фаза ϕ или разность фаз $\Delta\phi$, доплеровский сдвиг частоты F_d , которым сопоставляют функционально связанные с ними величины, а именно дальность R до НКА, его радиальную скорость V_r , разность дальностей ΔR до нескольких НКА, угол визирования ϕ , угловую скорость Ω , по которым и рассчитываются искомые компоненты вектора состояния.

Измеряемые в интересах решения навигационной задачи параметры радиосигнала получили название *радионавигационных параметров* (РНП), а

сопоставляемые им параметры вектора состояния потребителя — *навигационных параметров* (НП). Соответствующие функциональные зависимости между ними получили название *навигационных функций*. Очевидно, что вектор состояния потребителя может быть полностью определен, если его размерность не превышает числа независимых измерений РНП.

В спутниковых радионавигационных системах ГЛОНАСС и GPS используются, в основном, дальномерный и радиально-скоростной методы измерений (см. далее). В этом случае радионавигационными параметрами являются время τ распространения (задержка) сигнала на трассе НКА—потребитель и доплеровское смещение частоты F_d . Соответствующая им дальность НКА—потребитель R и радиальная скорость потребителя относительно НКА v_r являются навигационными параметрами, а связывающие их навигационные функции в данном случае имеют вид $R = c\tau$, $v_r = F_d\lambda$. В общем случае вид навигационных функций определяется многими факторами: видом НП, параметрами движения НКА и потребителя, выбранной системой координат и т. д.

Геометрическое место точек пространства с одинаковым значением навигационного параметра называют *поверхностью положения*. Пересечение двух поверхностей положения определяет *линию положения*, т. е. геометрическое место точек, соответствующих двум разным значениям навигационных параметров. Местоположение потребителя определяется координатами точки пересечения трех поверхностей положения или двух линий положения. Если в силу неоднозначности измерений линии положения пересекаются более чем в одной точке, то необходимо использовать дополнительную поверхность положения или априорную информацию, позволяющую отделить истинное положение от ложных.

8.4.2. Методы определения координат по сигналам НКА

Дальномерный метод. Наиболее простой дальномерный метод навигационных определений основан на измерениях дальности R_i между i -м НКА и потребителем. В этом методе навигационным параметром является дальность R_i , а поверхностью положения — сфера радиуса R_i с центром, расположенным в центре масс i -го НКА. Уравнение сферы имеет вид

$$R_i = \sqrt{(x_{ci} - x)^2 + (y_{ci} - y)^2 + (z_{ci} - z)^2}, \quad 1 \leq i \leq n. \quad (8.1)$$

Здесь x_{ci} , y_{ci} , z_{ci} — известные в момент измерения координаты i -го НКА (с учетом его перемещения за время распространения сигнала); x , y , z — координаты потребителя.

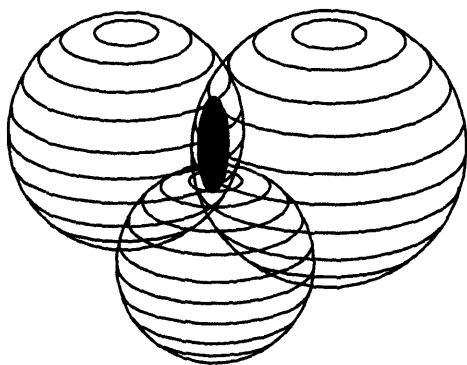


Рис. 8.7. Поверхности положения для дальномерного метода

такой системы уравнений возникает проблема неоднозначности определения координат потребителя, устраняемая с помощью известной потребителю дополнительной информации (априорные координаты потребителя, его радиальная скорость и т. д.).

В выражении (8.1) неявно подразумевается, что все входящие в него величины относятся к одному и тому же моменту времени. Однако координаты НКА привязаны к бортовой ШВ, а потребитель измеряет задержку сигнала и определяет свои координаты в своей ШВ. Если шкалы времени НКА и потребителя идеально синхронизированы, то проблем не возникает. При наличии расхождения τ_n шкал времени возникает смещение $\delta R = c\tau_n$ измеренной дальности относительно истинной и точность определения местоположения потребителя падает, поэтому недостатком метода является необходимость очень точной привязки шкал времени НКА и потребителя. Уменьшить влияние этого фактора можно, установив у потребителя высокостабильный эталон времени (частоты) и периодически проводя его калибровку по бортовой ШВ. Однако применение в АП дорогостоящих высокостабильных эталонов времени экономически не оправдано, а проблема создания относительно дешевых высокостабильных эталонов времени (частоты) в настоящее время не решена, поэтому более широко применяют псевдодальномерный метод.

Псевдодальномерный метод. Под псевдодальностью от i -го НКА до потребителя понимают измеренную дальность \hat{R}_i этого НКА, отличающуюся от истинной дальности R_i на неизвестную, но постоянную за время определения навигационных параметров величину δR . Таким образом, псевдодальность до i -го НКА можно описать формулой

$$\hat{R}_i = \sqrt{(x_{ci} - x)^2 + (y_{ci} - y)^2 + (z_{ci} - z)^2} + c\tau_n, \quad (8.2)$$

Местоположение потребителя, т. е. координаты x, y, z , определяют как координаты точки пересечения трех поверхностей положения, другими словами, трех сфер (см. рис. 8.7). Поэтому для реализации дальномерного метода необходимо измерить дальности (8.1) до трех НКА, т. е. $i = 1, 2, 3$. Таким образом, для дальномерного метода навигационная функция представляет собой систему из трех уравнений вида (8.1). Ввиду нелинейности

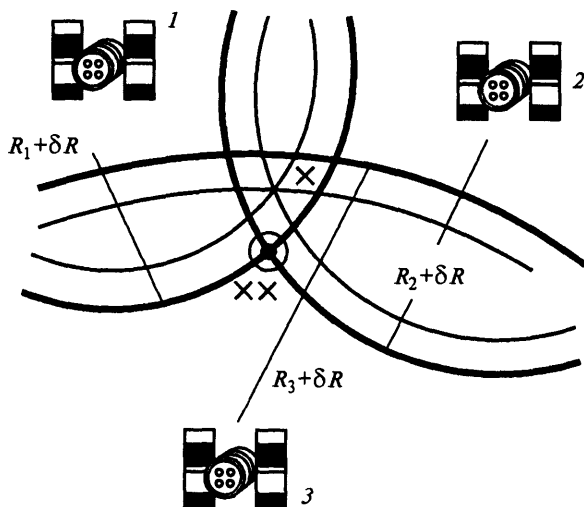


Рис. 8.8. Принцип псевдодальномерных измерений при определении положения на плоскости

где τ_n — смещение временной шкалы потребителя относительно системного времени.

В псевдодальномерных методах, основанных на измерениях псевдодальностей, в качестве навигационного параметра выступает \hat{R}_i . Поверхностью положения по-прежнему является сфера с центром в точке центра масс НКА, но радиус этой сферы отличается на неизвестную величину δR (см. рис. 8.8). Измерение псевдодальностей до трех НКА приводит к системе трех уравнений с четырьмя неизвестными x , y , z , δR . В решении этой системы уравнений возникает неопределенный параметр δR , и для устранения возникшей неопределенности необходимо провести дополнительное измерение, т. е. измерить псевдодальность до четвертого спутника. Полученная таким образом система четырех уравнений имеет точное решение, и, следовательно, местоположение потребителя при измерениях псевдодальностей определяется как точка пересечения четырех поверхностей положения. Именно необходимость нахождения в зоне видимости четырех НКА в значительной степени определяет структуру и параметры орбитальной группировки НКА.

Псевдодальномерный метод не накладывает жестких ограничений на значение погрешности $\delta R = \sigma_{\tau_n}$ (погрешности временной шкалы) и позволяет одновременно с определением местоположения вычислять отклонение шкалы времени потребителя.

Разностно-дальномерный метод. Метод основан на измерении разности дальностей от потребителя до одного или нескольких НКА. По своей сути этот метод аналогичен псевдодальномерному, так как его целесообразно использовать только при наличии в дальномерных измерениях неизвестных сдвигов δR , т. е. когда фактически проводятся измерения псевдодальностей. Разностно-дальномерный метод использует три разности $\Delta \hat{R}_{ij} = \hat{R}_i - \hat{R}_j$ до четырех НКА, так как при постоянстве δR за время навигационных определений разности псевдодальностей равны разностям истинных дальностей, для определения которых требуется лишь три независимых уравнения. Навигационным параметром является $\Delta \hat{R}_{ij}$.

Поверхности положения в разностно-дальномерном методе определяются из условия $\Delta \hat{R}_{ij} = \text{const}$ и представляют собой поверхности двухполостного гиперболоида вращения, фокусами которого являются координаты опорных точек i и j (центров масс i -го и j -го НКА). Расстояние между этими опорными точками называют *базой* измерительной системы. Если расстояния от опорных точек (НКА) до потребителя велики по сравнению с размерами базы, то гиперболоид вращения в окрестности точки потребителя практически совпадает со своей асимптотой — конусом, вершина которого расположена в середине базы.

Точность определения координат потребителя разностно-дальномерным методом не отличается от точности определения псевдодальномерным методом [115, 116]. Недостатком разностно-дальномерного метода является то, что в нем не может быть измерено смещение δR , а следовательно, и смещение шкалы времени потребителя.

Радиально-скоростной (доплеровский) метод. Метод основан на измерении трех радиальных скоростей перемещения потребителя относительно трех НКА. Физической основой радиально-скоростного метода является зависимость радиальной скорости точки относительно НКА от координат и относительной скорости НКА. Дифференцируя (8.1) по времени, получаем

$$\dot{R}_i = [(x_{ci} - x)(\dot{x}_{ci} - \dot{x}) + (y_{ci} - y)(\dot{y}_{ci} - \dot{y}) + (z_{ci} - z)(\dot{z}_{ci} - \dot{z})] / R_i, \quad i = 1, 2, 3. \quad (8.3)$$

Здесь компоненты $\{(\dot{x}_{ci} - \dot{x}), (\dot{y}_{ci} - \dot{y}), (\dot{z}_{ci} - \dot{z})\}$ характеризуют вектор относительной скорости; R_i — дальности до i -го НКА.

Из соотношения (8.3) следует, что для определения компонент $\{\dot{x}, \dot{y}, \dot{z}\}$ вектора скорости потребителя необходимо знать векторы координат

$\{x_{ci}, y_{ci}, z_{ci}\}$ и скорости $\{\dot{x}_{ci}, \dot{y}_{ci}, \dot{z}_{ci}\}$ трех НКА, а также координаты потребителя (x, y, z) . Последние можно получить, если измерить радиальные скорости \dot{R}_i в течение некоторого времени Δt , а затем вычислить интеграл

$$N_i = \int_{t_1}^{t_2} \dot{R}_i(t) dt = R_i(t + \Delta t) - R_i(t), \quad (8.4)$$

где $t_2 = t_1 + \Delta t$.

В результате интегрирования определяются «новые эквивалентные» измерения N_i , соответствующие разностно-дальномерному методу с той лишь разницей, что разности дальностей формируются для одного и того же НКА, но в различные моменты времени.

Действительно, если за время Δt перемещением потребителя можно пренебречь, то из (8.4) следует

$$N_i = \left\{ [x_{ci}(t + \Delta t) - x]^2 + [y_{ci}(t + \Delta t) - y]^2 + [z_{ci}(t + \Delta t) - z]^2 \right\}^{0.5} - \left\{ [x_{ci}(t) - x]^2 + [y_{ci}(t) - y]^2 + [z_{ci}(t) - z]^2 \right\}^{0.5} \quad (8.5)$$

Отсюда видно, что, зная координаты трех НКА в моменты времени t и $t + \Delta$ и проведя измерения N_i , можно определить координаты точки потребителя.

Недостатком данного метода измерения координат потребителя является невозможность их проведения в реальном масштабе времени. Кроме того, в средневысотных СРНС реализация радиально-скоростного метода затруднена ввиду медленного изменения радиальной скорости, что приводит к малым значениям разностей N_i в (8.4). Это обусловило применение данного метода в таких СРНС только для определения составляющих скорости потребителя.

Другой возможностью получения информации о координатах (x, y, z) является использование радиально-скоростного метода совместно с одним из описанных выше вариантов дальномерного метода.

Недостатком радиально-скоростного метода при определении скорости потребителя является необходимость наличия высокостабильного эталона частоты, так как любая нестабильность частоты приводит к неконтролируемому изменению доплеровского смещения частоты, а следовательно, к дополнительным ошибкам измерения составляющих скорости потребителя.

Псевдорadiaльно-скоростной метод. Псевдорadiaльно-скоростной (псевдодоплеровский) метод позволяет определять вектор скорости потребителя в присутствии неизвестного смещения частоты сигнала, например ввиду нестабильности эталона частоты. При наличии такого смещения $\delta \dot{R}_i$ выражение для радиальной скорости можно представить в виде двух слагаемых:

$$\hat{\dot{R}}_i = \dot{R}_i + \delta\dot{R}_i = [(x_{ci} - x)(\dot{x}_{ci} - \dot{x}) + (y_{ci} - y)(\dot{y}_{ci} - \dot{y}) + (z_{ci} - z)(\dot{z}_{ci} - \dot{z})]/R_i + \delta\dot{R}_i. \quad (8.6)$$

По своей сущности метод аналогичен псевдодальномерному методу определения координат потребителя. Для нахождения вектора скорости потребителя $\{\dot{x}, \dot{y}, \dot{z}\}$ и поправки $\delta\dot{R}_i = \lambda\Delta F_{\text{Д}i}$ необходимо провести измерения по четырем НКА и решить систему четырех уравнений вида (8.6). Для ее решения потребуются знания дальностей R_i и координат (x, y, z) потребителя. Эта информация может быть получена, например, из псевдодальномерных измерений (8.2).

Разностно-радиально-скоростной метод. Сущность данного метода заключается в определении трех разностей $\Delta\dot{R}_{ij} = \dot{R}_i - \dot{R}_j$ двух радиальных скоростей НКА. При этом разности можно вычислять относительно одного или относительно различных НКА. По существу, при вычислении разностей могут использоваться и псевдорадиальные скорости $\hat{\dot{R}}_i$, так как при таком вычитании компенсируется неизвестное смещение $\delta\dot{R}_i$ (в предположении, что это смещение одинаково для различных спутников). Навигационные параметры

$$\Delta\dot{R}_i = [(x_{ci} - x)(\dot{x}_{ci} - \dot{x}) + (y_{ci} - y)(\dot{y}_{ci} - \dot{y}) + (z_{ci} - z)(\dot{z}_{ci} - \dot{z})]/R_i - \\ - [(x_{cj} - x)(\dot{x}_{cj} - \dot{x}) + (y_{cj} - y)(\dot{y}_{cj} - \dot{y}) + (z_{cj} - z)(\dot{z}_{cj} - \dot{z})]/R_j.$$

Поверхности положения представляют собой поверхности тела вращения, фокусами которого являются координаты центров масс i -го и j -го НКА.

Так же, как и для «дальномерных» методов, точность определения составляющих вектора скорости в разностно-радиально-скоростном методе совпадает с точностью определения тех же составляющих в псевдорадиально-скоростном методе.

Достоинством разностно-радиально-скоростного метода является его нечувствительность к нестабильностям эталонов частоты и другим неконтролируемым смещениям частоты, а его недостатком — невозможность оценки нестабильности эталонов частоты.

Комбинированные методы. Помимо перечисленных основных методов определения компонент вектора потребителя возможны комбинированные методы, использующие кроме СРНС дополнительные измерители координат, имеющиеся у потребителя. Так, в дальномерном методе при наличии у потребителя измерителя высоты h можно вместо измерений трех дально-

стей до НКА ограничиться измерением двух дальностей. В этом случае система уравнений, описывающая навигационную функцию, будет включать два уравнения вида (8.1), а третье уравнение использует независимые измерения высоты:

$$(R_3 + h)^2 = x^2 + y^2 + z^2,$$

где R_3 — радиус Земли.

Другой аспект использования комбинированных методов заключается в замене совокупности одновременных измерений на комбинацию одновременных и последовательных измерений или на совокупность только последовательных измерений, например определение координат потребителя разностно-скоростным методом (8.5). В качестве другого примера можно привести псевдодальномерный метод, который можно реализовать, заменив четыре одновременных измерения по четырем НКА на два последовательных измерения по двум НКА или на четыре последовательных измерения до одного НКА.

Аналогичные комбинации возможны и для других методов.

8.5. Навигационная аппаратура потребителя

8.5.1. Обобщенная структурная схема АП

Подсистема АП предназначена для приема сигналов от НКА, измерения навигационных параметров, выделения служебной информации (включающей эфемериды и временные поправки) и решения задачи НВО. По измеренным НП — псевдодальности и радиальной псевдоскорости — относительно четырех НКА определяются три пространственные координаты потребителя, три составляющие его скорости и поправки к фазе и частоте его бортового генератора.

В состав АП входят антенное устройство (антенна, блоки предварительного усилителя и управления антенной), приемное устройство (блоки преобразования и усиления, поиска сигналов, навигационных измерений и выделения навигационного сообщения, а также блок опорного генератора (ОГ) и синтезатора частот), вычислительное устройство (процессор и блок связей — интерфейс), блок управления и индикации и блок питания. В зависимости от назначения и конструктивного исполнения конкретных моделей современной АП в ее составе могут отсутствовать антенное устройство, блок питания, блок управления и индикации. Обязательными функциональными блоками являются приемное устройство и навигационный процессор.

Как уже было отмечено (см. § 8.4), для однозначного решения навигационной задачи необходимо получить оценки РНП по сигналам, как минимум, четырех НКА. В первых образцах АП, использовавших одноканальные приемники, эта задача решалась методом последовательной настройки приемника на соответствующие литерные частоты ГЛОНАСС, либо методом перебора опорных последовательностей дальномерных кодов GPS. Однако оперативность и точность таких измерений не удовлетворяет современным требованиям. Поэтому в настоящее время приемники АП строятся исключительно как многоканальные, причем число параллельных каналов, реализующих процедуру согласованной фильтрации, колеблется от 6–12 в несложных GPS-приемниках, до 24 и более в совмещенных (ГЛОНАСС + GPS) образцах АП.

Современная АП оснащена высокопроизводительными процессорами, позволяющими полностью автоматизировать выполнение всех основных операций: поиска и обнаружения сигналов выбранного созвездия НКА, слежения за сигналами и измерения РНП, приема и декодирования служебной информации, учет в измеренных РНП различных поправок (релятивистской, тропосферной, временной и т. п.), решения основной навигационно-временной задачи и сопутствующих сервисных задач, контроля навигационных решений, контроля работоспособности аппаратуры, регистрации и индикации результатов решений.

Кроме того, в зависимости от типа и назначения АП на ее вычислительные средства могут возлагаться различные сервисные задачи. Например, для самолетной аппаратуры к числу сервисных задач относятся: хранение координат промежуточных пунктов маршрута и промежуточных аэропортов, расчет и хранение параметров линии заданного пути, расчет ортодромических координат (главных и частных), расчет времени полета и оставшегося расстояния до очередного промежуточного пункта, его азимута и т. д.

С позиций современной теории оптимальной фильтрации решаемая в АП задача получения оценки вектора состояния потребителя может рассматриваться в виде единого алгоритма обнаружения—оценивания. Однако для упрощения АП и программного обеспечения задачу получения оценок вектора потребителя разбивают на два этапа обработки [116, 117]: *первичную* и *вторичную*. На этапе первичной обработки решаются задачи поиска и обнаружения сигналов, слежения за ними, фильтрации (оценки) РНП сигнала, приема и декодирования служебной информации. На этапе вторичной обработки с использованием полученных на первом этапе оценок РНП и соответствующих навигационных функций решается задача НВО, т. е. вычисляются оценки вектора состояния потребителя. В состав программного обеспечения вторичной обработки входят и блоки управления первичной

обработкой, а также блоки для ввода и вывода необходимой информации и для решения сервисных задач.

8.5.2. Структура информационного обмена между элементами АП

На рис. 8.9 приведена обобщенная схема обмена информацией между основными элементами АП: процессорами первичной и вторичной обработки, интерфейсом внешних устройств — и различными потребителями информации.

На вход процессора первичной обработки поступают сигналы НКА, принятые антенной, усиленные и преобразованные на промежуточную частоту в соответствующем радиочастотном блоке, а также сигналы от опорного генератора и (или) синтезатора. Из блока первичной обработки в блок вторичной передаются измеренные значения НП (квазидальности и радиальной квазискорости), отсчет системного времени спутника, а также строки служебной информации. С блока вторичной обработки (навигационного процессора) результаты НВО поступают на интерфейсный блок, с которого информация распределяется на пульт управления, другим бортовым системам, внешним потребителям.

Кроме того, навигационный процессор выдает в блок первичной обработки управляющую информацию: номера видимых НКА, прогнозируемые значения РНП, режимы и подрежимы работы, оценку точности целеуказания. В случае использования антенн с управляемым обзором необходимые сигналы передаются и на блок управления антенной.

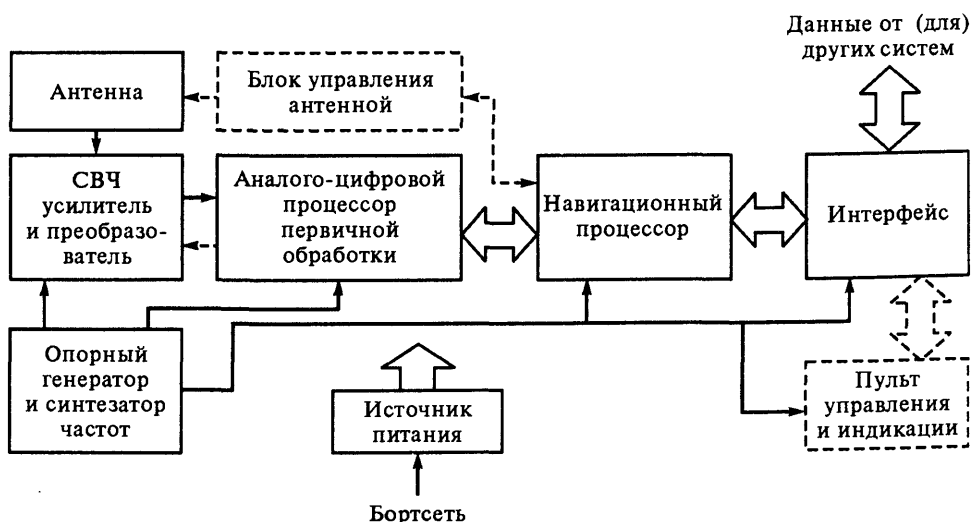


Рис. 8.9. Структурная схема АП

Пульт (при его наличии) используется для ввода в АП исходной информации, необходимой для организации навигационного сеанса: режимов работы, априорных координат, характерных точек маршрута и т. д.

Очевидно, что периодичности обращений к различным программам сильно различаются и наиболее часто используются блоки программ, обеспечивающие НВО. Кроме того, с определенной периодичностью проверяется видимость НКА, проводится перезапись эфемерид и выполняются другие процедуры. Порядок и периодичность работы различных блоков определяется программой-диспетчером.

Временная диаграмма программы диспетчера (см. рис. 8.10) распределяет блоки алгоритма и их части либо по параллельным ветвям вычислений, либо по одной последовательной ветви. На каждой из ветвей эти блоки размещаются так, чтобы реализовать требуемую периодичность расчетов при условии их максимального уплотнения по времени и соблюдения приоритетности отдельных операций. Основное ядро временной диаграммы реализует программу жесткого типа, однако обычно в ней предусматривается наличие определенных элементов адаптации к условиям навигационного сеанса.

Интерфейс с внешними потребителями обычно осуществляется через асинхронный последовательный порт RS232 или RS422. При этом используются различные протоколы обмена: NMEA 0183 (в корабельной АП),



Рис. 8.10. Алгоритм управления подрежимами в АП

RINEX (в геодезической АП), ГОСТ18977-79, ARINC 429 (для авиационной АП), бинарные (для межмашинного обмена) и ряд других. Наибольшее распространение получил протокол NMEA, современные версии которого предусматривают передачу потребителю не только результатов НВО, но и первичной информации, что позволяет реализовывать режимы относительных и дифференциальных измерений.

8.5.3. Принципы и устройства первичной обработки навигационной информации

Отношение сигнал—шум на входе приемника АП при ширине спектра СТ-кода сигнала ГЛОНАСС (С/А-кода сигнала GPS) порядка 1 МГц имеет величину порядка $-150...-160$ дБ. В то же время для получения удовлетворительных точностей оценок РНП необходимо, чтобы отношение сигнал—шум для одного измерения имело порядок $6...10$ дБ, что достигается согласованной фильтрацией. Согласованная фильтрация в АП реализуется сверткой на корреляторе принимаемого сигнала с эталонной ПСП дальномерного кода, иными словами, путем вычисления их взаимно корреляционной функции (ВКФ). При этом, для того чтобы такая процедура фильтрации действительно была близка к оптимальной (согласованной), необходимо, чтобы сдвиг по времени между началом опорной и выделяемой ПСП не превышал половины длительности τ_c элементарного символа (бита) дальномерного кода. Поскольку задержка τ принятого сигнала является одним из измеряемых РНП (см. выше), она априори неизвестна и может рассматриваться как случайная величина, равномерно распределенная в интервале $0 \leq \tau \leq N\tau_c$, где N — число символов дальномерного кода. Аналогичная априорная неопределенность существует относительно второго РНП — доплеровского сдвига несущей частоты принятого сигнала.

Эквивалентная ширина полосы пропускания систем фазовой автоподстройки частоты (ФАПЧ), обеспечивающих в АП слежение за сигналом, имеет порядок 1 Гц, в то время как априорный диапазон значений доплеровского сдвига F_d для наземных потребителей имеет, как уже указывалось, порядок ± 5 кГц, а для низкоорбитальных космических объектов — до ± 40 кГц.

Таким образом, для того чтобы следящие системы АП обеспечили измерения РНП с требуемой точностью, необходимо в начале каждого навигационного сеанса, а также в случае «срывов» слежения, выполнить процедуру поиска сигнала в пространстве неизвестных параметров (задержки и частоты), в ходе которой осуществляется грубая, но достаточная для дальнейшего захвата следящими фильтрами оценка этих параметров.

Поиск и фильтрация сигнала каждого НКА осуществляется независимо, поэтому далее ограничимся рассмотрением путей решения этой задачи по одному НКА.

8.5.4. Поиск сигналов по задержке и частоте

Задача поиска сигнала в АП СРНС полностью укладывается в рамки классической трактовки: оптимальная процедура состоит в поиске пары значений $\{F_d; \tau\}$, которой соответствует максимум апостериорной плотности вероятности (АПВ). Незвестные параметры F_d и τ считаются независимыми случайными величинами, имеющими непрерывную плотность распределения. Априорные распределения этих величин предполагаются равномерными в интервалах соответственно $0 \leq F_d \leq F_{d\max}$ и $0 \leq \tau \leq N\tau_c$. В этом случае оптимальные оценки, как известно, совпадают с максимально правдоподобными, а достаточной статистикой является величина, пропорциональная квадрату огибающей на выходе согласованного фильтра. При этом фаза принятого сигнала НКА также считается равномерно распределенной на интервале $(0, 2\pi)$, поэтому согласованная фильтрация сигналов осуществляется в квадратурах.

Однако реализация оптимальной процедуры поиска по непрерывному пространству параметров $\{F_d; \tau\}$ требует больших вычислительных мощностей, поэтому на практике в АП обычно используют квазиоптимальную процедуру, получаемую из оптимальной путем перехода от непрерывного к дискретному множеству этих параметров, т. е. считается, что параметры τ и F_d могут принимать дискретные значения $\tau_i, i = 1, \dots, N_\tau$, и $F_{dj}, j = 1, \dots, N_f$.

Для СРНС ГЛОНАСС диапазон доплеровских частот $\Delta F_d = -5 \dots 5$ кГц, а полоса доплеровских частот δF_d , соответствующая элементарной ячейке поиска, определяется полосой захвата системы частотной автоподстройки и составляет примерно 500 Гц, т. е. число анализируемых ячеек по частоте равно 20. Число ячеек поиска по задержке равно числу элементов дальномерного СТ-кода (элементов разрешения по дальности), т. е. 511. Таким образом, общее число анализируемых ячеек составляет 10220. Для каждой пары значений параметров $\{F_d; \tau\}$ рассчитывается решающая статистика, которая сравнивается с некоторым заранее заданным порогом, при превышении которого сигнал с данными параметрами считается обнаруженным и АП переходит в режим непрерывной фильтрации параметров F_d и τ . Время некогерентного накопления решающей статистики при анализе в одной элементарной ячейке составляет 1...4 мс, что соответствует 1...4 периодам СТ-кода.

Для ускорения поиска АП описанная процедура организуется как параллельно-последовательная, при этом число параллельно анализируемых

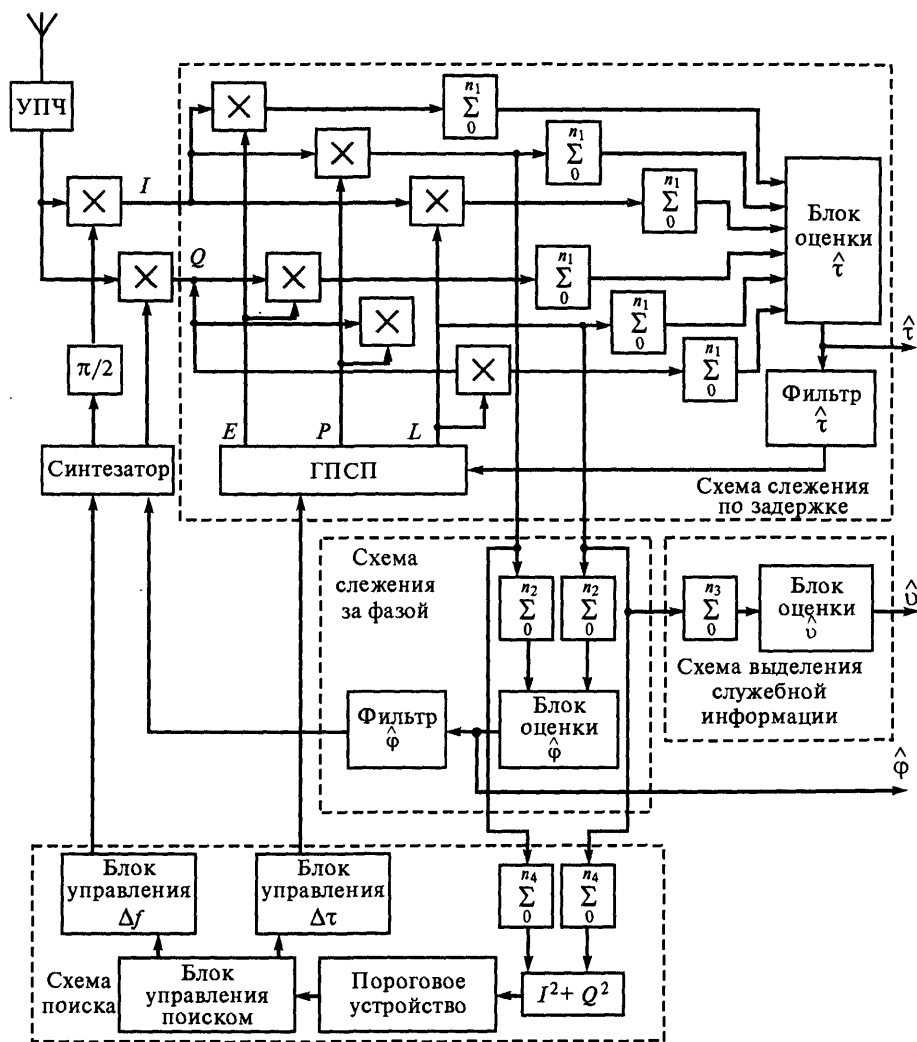


Рис. 8.11. Структурная схема каналов поиска и слежения АП

комбинаций значений F_d и τ зависит от требований к оперативности измерений и стоимости (сложности) АП.

Упрощенная схема устройства поиска приведена на рис. 8.11. Входной сигнал с антенны АП после понижения частоты до промежуточной и усиления поступает на умножители каналов формирования синфазной и квадратурной составляющих.

Процедура поиска сигнала j -го НКА начинается с того, что на синтезаторе частот устанавливается частота опорного колебания $f_{снj}$, удовлетво-

ряющая соотношению $f_0 + F_{dj} - f_{снj} = f_{пр}$, где F_d — прогнозируемое значение доплеровского сдвига, $f_{пр}$ — промежуточная частота. На выходе генератора кода ПСП формируется функция $h(t - \tau_j)$, соответствующая дальномерному коду j -го НКА, задержанному на величину τ_j , устанавливаемую блоком управления на основании прогноза положения НКА по данным альманаха.

Сигнал частоты $f_{снj}$ поступает на расщепитель фазы и затем на квадратурные перемножители, с помощью которых осуществляется перенос сигнала НКА на видеочастоту. Полученный сигнал, имеющий вид меандра со случайным периодом, перемножается в квадратурных корреляторах с опорным колебанием $h(t - \tau_j)$ от генератора ПСП. Квадраты значений выходных сигналов корреляторов суммируются и поступают на пороговое устройство, в котором принимается решение об обнаружении (необнаружении) сигнала. Данные об обнаружении сигнала передаются в блок управления поиском, где принимается решение о последующем поиске или переходе в режим фильтрации РНП и навигационной информации.

8.5.5. Фильтрация радионавигационных параметров и дешифрация навигационной информации

В отличие от этапа поиска сигнала, при фильтрации информационных параметров используется режим квазикогерентного накопления, реализуемый с помощью узкополосных следящих фильтров. Оптимальное решение этой задачи состоит в совместной фильтрации двух непрерывных РНП: задержки τ и доплеровского сдвига F_d частоты несущей сигнала — и одного дискретного параметра — фазы модулирующего сигнала, с помощью которого передается НИ. Однако для простоты во многих образцах современной АП взаимосвязью этих параметров пренебрегают и оценки задержки τ и доплеровского сдвига F_d формируются с помощью независимых дискриминаторов. При этом оценка доплеровского сдвига F_d формируется не непосредственно путем слежения за частотой несущей, а интегрированием оценки фазы несущей $\hat{\phi}$, вычисляемой соответствующим фазовым дискриминатором (см. схему слежения за фазой на рис. 8.11).

Отметим, что до начала слежения за фазой необходима более точная частотная автоподстройка для обеспечения попадания частоты сигнала в полосу захвата фазовой автоподстройки (порядка 50 Гц). Нетрудно показать [116], что функция частотного дискриминатора может быть реализована путем алгоритмической обработки отсчетов сигналов I и Q , полученных в смежные моменты времени n , $n-1$ в схеме слежения за фазой. При этом сигнал ошибки системы частотной автоподстройки u_f определяется по формуле

$$u_f = I_{n-1}Q_n - I_nQ_{n-1}.$$

Рассмотрим работу канала оценки задержки сигнала, который включает дискриминатор задержки, сглаживающий фильтр и ГПСП. В современных типах АП дискриминатор задержки обычно содержит три коррелятора (см. рис. 8.11). На опорный вход первого коррелятора опорный сигнал подается без дополнительного временного сдвига (P -канал), а на входы двух других — либо с опережением (E -канал), либо с запаздыванием (L -канал) относительно P -канала. Величина временного сдвига опережения—запаздывания обычно выбирается близкой к половине длительности символа ПСП дальномерного кода. Характеристика дискриминатора схемы слежения за задержкой формируется путем алгоритмической обработки квадратурных составляющих каналов E , P , L . Существуют и используются различные алгоритмы вычисления выходного сигнала u_τ дискриминатора, например

$$u_\tau = (I_E - I_L)I_P + (Q_E - Q_L)Q_P.$$

Иногда для упрощения используется двухканальная схема дискриминатора, для которой алгоритм вычисления выходного сигнала имеет вид

$$u_\tau = \sqrt{(I_E^2 + Q_E^2)} - \sqrt{(I_L^2 + Q_L^2)}.$$

С выхода дискриминатора сигнал рассогласования u_τ через сглаживающий фильтр поступает на ГПСП для управления временным положением кодовой последовательности в P -канале. Структура сглаживающего фильтра определяется принятой моделью изменений задержки.

Для выделения навигационной информации, передаваемой путем низкочастотной ОФМ (см. § 8.3), используется P -канал одного из квадратурных корреляторов. Алгоритм дешифрации НИ сводится к определению знака суммы M отсчетов квадратурной составляющей:

$$\text{sign}\left(\sum_0^M I_n\right),$$

где M — число отсчетов на интервале времени, равном длительности бита НИ (20 мс).

Пример структурной схемы канала дешифрации НИ для СРНС ГЛОНАСС приведен на рис. 8.12. Блок выделения синхросимволов, фиксируя моменты смены полярности напряжения на выходе P -канала, выделяет импульсы символьной синхронизации бидвоичного кода частоты 100 Гц, образующегося на НКА путем суммирования по модулю 2 информационной последовательности НИ и вспомогательного меандра частоты 100 Гц

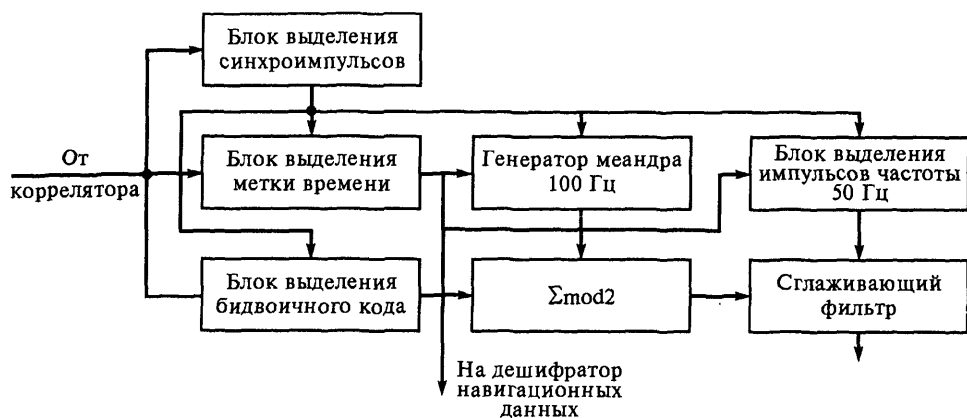


Рис. 8.12. Схема канала дешифрации НИ

(см. § 8.3). Эти импульсы используются для символьной синхронизации блоков выделения метки времени, формирования вспомогательного меандра частоты 100 Гц, выделения бидвоичного кода, а также блока выделения импульсов символьной частоты 50 Гц информационной последовательности.

В блоке выделения метки времени после сглаживания (интегрирования) шумов формируется импульс, синхронный с задним фронтом тридцатого символа кода метки времени и совпадающий с двухсекундной меткой.

В результате суммирования по модулю 2 вспомогательного меандра и бидвоичного кода восстанавливаются символы НИ, которые после дополнительного сглаживания на интервале времени 20 мс поступают в блок выделения НИ, а затем вместе с метками времени — на навигационный процессор.

8.5.6. Вторичная обработка навигационной информации

Основной задачей, решаемой на этапе вторичной обработки, является задача НВО, т. е. определение на основании полученных на первом этапе оценок РНП и соответствующих навигационных функций вектора состояния потребителя. Кроме того, программа вторичной обработки содержит блоки управления первичной обработкой, вводом и выводом необходимой информации, а также программы для решения сервисных задач.

Перечислим основные этапы решения навигационно-временной задачи.

1. Выбор рабочих НКА, т. е. определение номеров радиовидимых в течение ближайшего интервала времени НКА. В качестве исходных данных используется информация из альманаха СРНС.

2. Прогноз ожидаемых значений РНП и подготовка целеуказания для поиска сигналов НКА. В качестве исходных данных используют результаты

прогноза положения НКА, а также априорную информацию о положении потребителя (при наличии таковой).

3. Краткосрочный прогноз эфемерид для предельно точных расчетов значений координат и составляющих вектора скорости НКА на момент измерений. Исходные данные для задачи — номера рабочих НКА и моменты времени измерений, а также оперативная служебная информация по НКА на ближайший узловой момент времени.

4. Формирование массива измерений. Проводится с темпом считывания измерений квазидальностей и радиальной квазискорости соответствующих схем слежения за параметрами радиосигнала. При решении этой задачи измеренные значения НП корректируются с учетом поправок из служебной информации НКА и данных двухчастотных измерений для компенсации ионосферной погрешности; могут использоваться также данные дифференциальной коррекции (ДК) (см. § 8.7).

5. Собственно задача НВО, т. е. расчет пространственных координат и составляющих вектора скорости и определение текущего времени в системах отсчета СРНС и потребителя. Для решения этой задачи используются исходные данные в виде уточненных измеренных РНП, результатов краткосрочного прогноза эфемерид, данные других средств навигации.

6. Прием и обработка НИ. Задача обеспечивает формирование и обновление данных альманаха, эфемерид и других поправок, передаваемых в кадре служебной информации за период сеанса для каждого рабочего НКА.

В состав сервисных задач ВО входят, как правило, запись массивов измерений НП и координат, вывод визуальной информации о траектории объекта, расчет различных вспомогательных данных, например времени входа в заданный район, контроль траектории в заданных пределах и т. п. Особо следует выделить задачу априорной и апостериорной оценки точности полученных определений, что позволяет оценить качество навигационного обеспечения сеанса измерений.

Поскольку минимальное число НКА, при котором навигационно-временная задача имеет единственное решение, равно четырем (см. § 8.4), одномоментную выборку измерений РНП, содержащую не менее четырех независимых значений псевдодальностей (псевдоскоростей), принято называть *полной*. Кроме одномоментных выборок могут использоваться результаты разновременных измерений РНП по каждому из НКА рабочего созвездия, т. е. выборки нарастающего объема. Отметим, что современная АП, как правило, использует измерения всех радиовидимых НКА, при этом измерения по различным НКА могут быть несинхронными.

Рассмотрим основы статистической обработки полной выборки измерений применительно к методу наименьших квадратов.

Координаты НКА и потребителя при решении навигационно-временной задачи удобно описывать в геоцентрической экваториальной прямоугольной системе с центром в центре масс Земли. В этом случае определение координат в АП по данным полной выборки измерений псевдодальностей сводится к решению системы нелинейных уравнений вида (8.2).

При линейризации такой системы в окрестностях расчетных значений определяемых параметров формируется градиентная (4×4) -матрица вида

$$\mathbf{G} = \begin{bmatrix} \frac{\partial \hat{R}_1}{\partial x} & \frac{\partial \hat{R}_1}{\partial y} & \frac{\partial \hat{R}_1}{\partial z} & \frac{\partial \hat{R}_1}{\partial c\tau_n} \\ \frac{\partial \hat{R}_2}{\partial x} & \frac{\partial \hat{R}_2}{\partial y} & \frac{\partial \hat{R}_2}{\partial z} & \frac{\partial \hat{R}_2}{\partial c\tau_n} \\ \frac{\partial \hat{R}_3}{\partial x} & \frac{\partial \hat{R}_3}{\partial y} & \frac{\partial \hat{R}_3}{\partial z} & \frac{\partial \hat{R}_3}{\partial c\tau_n} \\ \frac{\partial \hat{R}_4}{\partial x} & \frac{\partial \hat{R}_4}{\partial y} & \frac{\partial \hat{R}_4}{\partial z} & \frac{\partial \hat{R}_4}{\partial c\tau_n} \end{bmatrix}, \quad (8.7)$$

где

$$\frac{\partial \hat{R}_i}{\partial x} = \frac{(x_{ci} - x)}{R_i}, \quad \frac{\partial \hat{R}_i}{\partial y} = \frac{(y_{ci} - y)}{R_i}, \quad \frac{\partial \hat{R}_i}{\partial z} = \frac{(z_{ci} - z)}{R_i}, \quad \frac{\partial \hat{R}_i}{\partial c\tau_n} = -1.$$

Решение системы линеаризованных уравнений методом последовательных приближений Ньютона проводится по формуле

$$\mathbf{g}_k = \mathbf{g}_{k-1} + \mathbf{G}_{k-1}^{-1} \Delta R_{k-1}, \quad (8.8)$$

где $\mathbf{g} = [\Delta x \ \Delta y \ \Delta z \ c\Delta\tau_n]^T$ — вектор оцениваемых параметров; $\Delta R = [\hat{R}_1 - R_1 \ \hat{R}_2 - R_2 \ \hat{R}_3 - R_3 \ \hat{R}_4 - R_4]^T$ — разница измеренных и расчетных дальностей от потребителя до НКА.

Начальные условия итерационного алгоритма определяют исходя из имеющейся у потребителя априорной информации. Заканчивается итерационный процесс при достижении заданной погрешности значений определяемых параметров.

Практические алгоритмы определения координат объекта по измерениям четырех псевдодальностей могут отличаться используемыми численными методами решения системы уравнений и, главным образом, методами обращения матрицы (8.7).

Если число НКА превышает число неизвестных, возникает избыточность системы. В данном случае при использовании метода наименьших квадратов уравнение (8.8) приобретает вид

$$\mathbf{g}_k = \mathbf{g}_{k-1} + \left(\mathbf{G}_{k-1}^T \mathbf{G}_{k-1} \right)^{-1} \mathbf{G}_{k-1}^T \Delta \mathbf{R}_{k-1}.$$

При этом матрица (8.7) имеет размерность $4 \times N$, где N — число независимых измерений РНП.

Рассмотренный выше алгоритм получен в предположении того, что измерения по всем НКА приведены к одному моменту времени. На практике вследствие конечной точности сведения шкал времени НКА и АП это условие нарушается, однако для низкодинамичных объектов возникающая по указанной причине погрешность может считаться пренебрежимо малой. В случае динамичного объекта для высокоточного измерения и интерполяции координат необходима фильтрация высокочастотных составляющих погрешностей, которая должна выполняться с учетом модели динамики объекта, позволяющей прогнозировать текущие значения скорости, ускорения, а также использовать разновременные измерения.

В качестве навигационного фильтра, учитывающего модель динамики объекта, обычно используется дискретный фильтр Калмана или его модификации [115, 116, 119] (см. гл. 6). Такой фильтр, как известно, является оптимальным при условии, что модели системы и измерителя линейны, а случайные возмущения и ошибки распределены по гауссовскому закону.

В режиме навигации без использования информации от внешних навигационных датчиков для высокодинамичных объектов выработка текущих значений вектора состояния проводится с помощью оцененных фильтром значений скорости и ускорения. Вектор состояния в этом случае включает одиннадцать компонент: три пространственные координаты, две оценки погрешностей опорного генератора АП (фазы и частоты), три составляющие вектора скорости, три составляющие вектора ускорения. Третьи производные перемещения объекта по времени и члены более высоких порядков трактуются как возмущающие силы в уравнениях погрешностей.

Опыт построения АП на основе рассмотренных принципов показал, что обеспечиваемая с их использованием погрешность навигационных определений оказывается достаточной для большинства приложений. Более сложные алгоритмы целесообразно использовать только в специальной АП, предназначенной для прецизионных измерений.

8.6. Точность навигационно-временных определений в СРНС

На точность навигационно-временных определений с помощью СРНС влияет множество факторов. Они связаны с характеристиками используемых сигналов, среды распространения, с особенностями построения АП и используемых в ней алгоритмов определения РНП и НП и т. д. Рассмотрим основные источники погрешностей этих измерений применительно к псевдодальномерному методу навигационно-временных определений (см. § 8.4).

В зависимости от локализации источника составляющие дальномерной погрешности измерений СРНС можно разделить на три категории:

- 1) погрешности, вносимые на НКА или КИК СРНС;
- 2) погрешности, вносимые на трассе распространения сигнала от НКА до АП;
- 3) аппаратные погрешности АП.

Ниже рассмотрены основные причины возникновения погрешностей, относящихся к перечисленным категориям.

8.6.1. Погрешности, вносимые на НКА или КИК СРНС

Погрешности, связанные с функционированием бортовой аппаратуры НКА и КИК СРНС обусловлены, в основном, несовершенством частотно-временного и эфемеридного обеспечения. Погрешности частотно-временного обеспечения возникают вследствие несовершенства аппаратуры, а также процедур сверки бортовой ШВ с системным эталоном. Они проявляются, например, в смещении фаз излучаемых дальномерных кодов и ме-ток времени, что приводит к погрешностям измерения расстояния до спутника и расчета его координат. По результатам многолетних наблюдений СКО сдвига бортовой ШВ через 2 часа после коррекции составляет примерно 9 нс, а через сутки после коррекции достигает 25,4 нс для цезиевых и 108 нс для рубидиевых эталонов.

Командно-измерительный комплекс СРНС корректирует бортовой ШВ таким образом, чтобы СКО ее сдвига относительно системной ШВ не превосходила 10 нс. При этом погрешность бортовой ШВ обусловлена двумя основными факторами:

- собственными погрешностями измерений, выполняемых «беззапросным» и «запросным» методами;
- погрешностью используемой модели прогноза расхождения бортовой ШВ относительно системной ШВ.

Погрешности беззапросных измерений псевдодальности составляют 3...5 м, погрешности запросных измерений дальности не превышают 1...2 м, соответственно, методическая погрешность измерений сдвига бортовой ШВ

относительно ШВ системы составляет в среднем 5 нс. Погрешность прогноза бортовой ШВ на 12 ч при использовании простейшей (линейной) модели расхождения составляет в среднем 14 нс [115].

Кроме указанных факторов должны учитываться релятивистские и гравитационные эффекты: поперечный доплеровский и гравитационный сдвиги частоты, а также эффекты, связанные с пересчетом моментов излучения и приема сигналов, вращением используемых систем координат и др. [115]. Для компенсации влияния указанных эффектов частоту бортового эталона несколько уменьшают относительно номинала (на $4,55 \cdot 10^{-3}$ Гц в СРНС GPS) [114, 118].

Погрешности эфемеридного обеспечения вызваны неточностью расчета параметров орбит НКА на КИК и непрогнозируемыми отклонениями реальной орбиты НКА относительно экстраполированной. Составляющие вектора этой погрешности лежат в пределах 0,6...10 м [116]. Порождаемая эфемеридными погрешностями составляющая эквивалентной погрешности дальности (ЭПД) равна примерно 1 м (СКО). Суммарная составляющая ЭПД, вносимая на КИК и НКА (включая погрешности ШВ НКА, эфемеридные и другие компоненты), в СРНС ГЛОНАСС не превышает 9,2 м (наиболее пессимистическая оценка) [116].

Существует еще один вид погрешности дальномерных измерений — фазовый сдвиг (групповая задержка) навигационного сигнала относительно бортового эталона времени в аппаратуре спутника [114]. Информация о систематической составляющей содержится в НИ НКА (параметры коррекции бортовой шкалы времени), случайная составляющая, не превышающая 3 нс (СРНС GPS), входит в дальномерную погрешность.

8.6.2. Погрешности, вносимые на трассе распространения сигнала от НКА до АП

Погрешности, возникающие на трассе распространения, в основном обусловлены рефракцией сигналов в атмосфере (тропосфере и ионосфере) Земли, а также особенностями интерференции сигналов НКА, связанной с их многолучевым распространением в точке расположения антенны АП.

Основная составляющая тропосферной погрешности навигационных определений в СРНС обусловлена рефракцией в тропосфере, связанной с неоднородностями ее диэлектрической проницаемости. Дополнительная задержка сигнала НКА в тропосфере может достигать 8...80 нс (экспериментальные данные для СРНС GPS) [115, 116]. Поскольку в диапазоне СРНС величина задержки в тропосфере не зависит от частоты (дисперсия сигнала отсутствует), измерить эту задержку с помощью двухчастотного сигнала невозможно. Однако значение тропосферной погрешности зависит от вза-

имных координат НКА и АП, температуры, давления и влажности воздуха. Указанные параметры достаточно точно известны и хорошо прогнозируются, поэтому в АП используется метод расчета тропосферной задержки, позволяющий уменьшить эту погрешность до единиц наносекунд [115, 116].

Значение тропосферной рефракции, пропорциональное длине пути сигнала в тропосфере, максимально при малых углах возвышения НКА над горизонтом. Поэтому для уменьшения атмосферных погрешностей в АП измерения проводят только для тех НКА, угол возвышения которых превышает некоторое значение («угол маски»), устанавливаемое потребителем. Обычно этот угол составляет $5...10^\circ$.

Ионосферная рефракция сигнала НКА, вызванная различиями диэлектрической проницаемости слоев, расположенных на разных высотах, а также локальными неоднородностями ионосферы, изменяется в широких пределах в зависимости от района Земли, где размещена АП, времени суток, года, солнечной и геомагнитной активности, и составляет $5...500$ нс [115, 116]. Среднее значение этой погрешности для углов возвышения, близких к 90° , составляет $5...10$ нс в ночное время и $30...50$ нс — в дневное. При углах возвышения около 15° эти величины возрастают в 2–3 раза. Таким образом, учитывать эти погрешности необходимо только в тех случаях, когда суммарная погрешность местоопределения АП не должна превышать $5...10$ м.

Ионосфера, в отличие от тропосферы, является диспергирующей средой, поэтому для оценки задержки сигнала в ней может быть использован метод двухчастотных измерений, обеспечивающий остаточную погрешность порядка $1...2$ м. Кроме того, для определения и учета ионосферной погрешности в современной АП применяют метод моделирования условий на трассе распространения сигналов НКА и метод избыточных одночастотных измерений.

Многолучевость распространения сигналов НКА в точке расположения антенны АП возникает из-за их переотражений от земной и морской поверхностей и близлежащих объектов, например элементов конструкции носителя АП. Уровень отраженного сигнала может быть соизмеримым с прямым сигналом, а его задержка относительно прямого сигнала для НКА, находящегося в зените, может составлять от единиц до сотен микросекунд (при расположении АП на борту самолета); при небольших углах возвышения НКА это значение уменьшается на порядок. Многолучевость приводит к существенным искажениям полезного сигнала и к погрешностям в схемах слежения за задержкой, частотой и фазой. Легко видеть, что для снижения влияния переотраженных сигналов желательно использовать сигналы с возможно более узким главным пиком ВКФ и низким уровнем ее боковых лепестков. С этой точки зрения, в СРНС предпочтительно использовать модулирующие ПСП с максимальной базой, например P(Y)-код GPS или BТ-код ГЛОНАСС. Их применение позволяет снизить погрешности, связанные с

многолучевостью, в среднем до 1...3 м (СКО) [116]. При использовании в условиях многолучевости относительно узкополосных сигналов стандартной точности (C/A GPS или СТ ГЛОНАСС) применяются специальные методы коррекции импульсной характеристики коррелятора, позволяющие снизить вероятность захвата схемами слежения пиков ВКФ, обусловленных приемом отраженных сигналов.

8.6.3. Погрешности аппаратуры потребителей

Источниками дальномерных погрешностей в АП являются схемы слежения за задержкой огибающей (ССЗ) и несущей (ЧАП и ФАПЧ) сигналов НКА.

Типовую некогерентную схему слежения за задержкой, использующую огибающую спутникового сигнала (двоичную псевдослучайную последовательность), можно характеризовать шумовой погрешностью с СКО [115, 116, 119]

$$\sigma = \tau_3 \left[\frac{k_1 \Pi_{\text{ССЗ}}}{P_c / N_0} + \frac{k_2 \Pi_{\text{ССЗ}} \Pi_{\text{ПЧ}}}{(P_c / N_0)^2} \right]^{0,5},$$

где τ_3 — длительность элемента кода ПСП; k_1, k_2 — постоянные коэффициенты, зависящие от выбранной схемы слежения; $\Pi_{\text{ССЗ}}, \Pi_{\text{ПЧ}}$ — односторонняя ширина полосы замкнутой схемы слежения и тракта ПЧ соответственно; P_c / N_0 — отношение мощности сигнала к спектральной мощности шума на входе АП.

При использовании в современной АП шага временной дискретизации, равного $\frac{1}{64} \tau_3$, соответствующие погрешности измеряемой дальности со-

ставляют 0,27 м для Р-кода и 2,66 м для C/A-кода GPS. С учетом этого суммарное значение флуктуационной составляющей дальномерной погрешности АП (СКО) равно приблизительно 0,2...1 м для Р-кода и 1,5...10 м для C/A-кода GPS. Динамическую составляющую дальномерной погрешности АП можно оценить как $\sigma_{\text{дин}}(R) = \frac{1,12 \Delta a \lambda}{4 \Pi_{\text{ССЗ}}^2}$, где Δa — скорость изменения

доплеровского сдвига частоты [119]. Это соотношение справедливо для установившегося режима схемы слежения второго порядка при квадратичном законе изменения задержки сигнала.

Для оценки точности измерения приращения дальности с использованием фазы несущей справедливо выражение [119]

$$\sigma_\phi(R) = \left[\frac{\lambda^2 \Pi_{\text{ФАПЧ}}}{(2\pi)^2 P_c / N_0} \right]^{0,5}.$$

Типовые значения этой погрешности составляют примерно 1,6 мм (СКО) для Р-кода и 1,2 мм для С/А-кода GPS.

Суммарное значение аппаратурной составляющей дальномерной погрешности имеет вид $\sqrt{\sigma_{\phi}^2(R) + \sigma_{\text{дин}}^2(R)} = 0,1\tau$, при полосе $\Pi_{\text{ССЗ}} = 3$ Гц и $P_c / N_0 = 30$ дБ/Гц.

С учетом полученных выше оценок вклада различных факторов в погрешность дальномерных измерений легко рассчитать, что суммарное значение ЭПД составляет 5 м для Р-кода и 15 м для С/А-кода GPS, при этом ее флуктуационная составляющая лежит в диапазоне 4...6,5 м.

8.6.4. Геометрический фактор

При фиксированных значениях погрешностей измерения псевдодальностей существенное влияние на точность местоопределения оказывает взаимное положение АП и рабочих НКА, которое определяет углы пересечения поверхностей положения. Минимальным погрешностям местоопределения для дальномерных методов соответствует такое расположение НКА, при котором в точке расположения АП поверхности положения пересекаются под прямым углом (см. рис. 8.13). Однако выполнить это условие при одновременном наблюдении более трех НКА невозможно.

Количественной характеристикой погрешности местоопределения, связанной с особенностями пространственного положения НКА и АП, служит так называемый геометрический фактор Γ_{Σ} , или коэффициент геометрии [115]. В англоязычной литературе используется обозначение GDOP (от англ. Geometrical delusion of precision — геометрический фактор ухудшения точности).

Для того чтобы уточнить математический смысл понятия геометрического фактора, представим уравнение (8.8) в виде зависимости вектора δ_0 погрешности определяемых параметров от вектора $\delta_{\text{и}}$ погрешности измеряемых НП

$$\delta_0 = \mathbf{g}_k - \mathbf{g}_{k-1} = \mathbf{G}_{k-1}^{-1} \Delta R_{k-1} = \mathbf{G}^{-1} \delta_{\text{и}}. \quad (8.9)$$

С учетом (8.9) получим корреляционную матрицу ошибок навигационных определений потребителя в виде

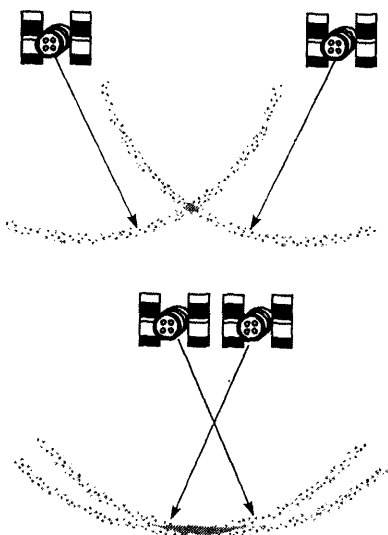


Рис. 8.13. Влияние положения НКА на точность местоопределения

$$\mathbf{K}_0 = \mathbf{M}\{(\delta_0 - m_0)(\delta_0 - m_0)^T\} = \mathbf{G}^{-1}\mathbf{K}_n(\mathbf{G}^{-1})^T = (\mathbf{G}^T\mathbf{K}_n^T\mathbf{G})^{-1}, \quad (8.10)$$

где $\mathbf{K}_n = \mathbf{M}\{(\delta_n - m_n)(\delta_n - m_n)^T\}$ — корреляционная матрица погрешностей измерения НП (псевдодальностей); m_0, m_n — векторы математических ожиданий погрешностей δ_0, δ_n соответственно.

Предположим, что составляющие вектора m_n равны нулю. Тогда $m_0 = 0$ и дисперсии параметров $(\sigma_x^2, \sigma_y^2, \sigma_z^2, \sigma_\tau^2)$, т. е. диагональные члены матрицы \mathbf{K}_0 , полностью определяют погрешности навигационного сеанса. Анализ выражения (8.10) показывает, что соотношение погрешностей определения вектора потребителя и измеряемых НП зависит только от вида матрицы градиентов \mathbf{G} , т. е. от геометрии взаимного положения НКА и АП.

Как правило, Γ_Σ вводится для случая, когда погрешности измерения НП, т. е. элементы матрицы \mathbf{K}_n , равновелики и некоррелированы. В этом случае соотношение (8.10) можно представить в виде

$$\mathbf{K}_0 = (\mathbf{G}^T\mathbf{K}_n^T\mathbf{G})^{-1} = \sigma_n^2(\mathbf{G}^T\mathbf{G})^{-1}.$$

Тогда, выразив дисперсию сеанса навигационных определений через след матрицы \mathbf{K}_0 :

$$\sigma_\Sigma^2 = (\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_\tau^2) = \text{tr}(\mathbf{K}_0),$$

можно представить геометрический фактор в виде коэффициента

$$\Gamma_\Sigma = [\text{tr}(\mathbf{G}^T\mathbf{G})^{-1}]^{0,5} = (\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_\tau^2)^{0,5} / \sigma_n.$$

При выборе орбитальных параметров НКА на этапе разработки СРНС, а также при формировании рабочего созвездия НКА в малоканальной аппаратуре коэффициент геометрии Γ_Σ является основным критерием.

Иногда удобно рассматривать отдельно значения геометрических факторов, характеризующих точность определения различных компонент вектора состояния: соответствующий геометрический фактор для пространственных координат обозначают Γ_Π (англ. PDOP — Position DOP), для плановых (горизонтальных) координат — Γ_Γ (англ. HDOP — Horizontal DOP), для высоты (вертикальной составляющей) — $\Gamma_\text{В}$ (англ. VDOP — Vertical DOP), для временного параметра — Γ_τ (англ. TDOP — Time DOP). Между этими коэффициентами и суммарным геометрическим фактором существует простая связь:

$$\Gamma_\Sigma^2 = \Gamma_\Pi^2 + \Gamma_\tau^2 = \Gamma_\Gamma^2 + \Gamma_\text{В}^2 + \Gamma_\tau^2.$$

Если в навигационном сеансе используется более четырех НКА, выражение для расчета геометрического фактора преобразуется к виду

$$\Gamma_{\Sigma}^2 = \text{tr}((\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T ((\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T)^T).$$

В [116] показано, что минимальное значение $\Gamma_{\Pi} = 1,5$ достигается в случае, когда потребитель находится в центре правильного тетраэдра. Соответственно, для наземного потребителя минимальное значение $\Gamma_{\Sigma} = 1,732$ достигается тогда, когда один НКА находится в зените, а три других равномерно расположены в горизонтальной плоскости, т. е. когда объем тетраэдра также максимальный.

8.7. Дифференциальная коррекция и относительные измерения в СРНС

Как следует из материалов предыдущего параграфа, современная АП СРНС в стандартном режиме измерений может обеспечить эквивалентную погрешность дальности на уровне 5...10 м (СКО), что позволяет потребителю при типовых значениях геометрического фактора определять абсолютные координаты с точностью до 15...60 м (СКО). Для многих приложений: навигации в открытом море, полетов на высотах более 1000 м, слежения за автотранспортными средствами и так далее — такая точность вполне достаточна. Однако для решения таких задач, как судовождение в проливах, каналах, акваториях портов, автоматическая инструментальная посадка воздушных судов, геодезическая привязка различных объектов, часто требуется более высокая точность местоопределения. Одним из наиболее эффективных путей решения указанных проблем является применение методов *относительных измерений* (ОИ) или методов *дифференциальной коррекции* (ДК).

Основная идея указанных методов состоит в совместной обработке результатов измерений НП или координат, полученных по крайней мере в двух разнесенных точках рабочей зоны СРНС. Принципиальная возможность повысить таким способом точность НВО обусловлена наличием в ЭПД коррелированной по времени и пространству составляющей, которая при использовании дифференциальных методов устраняется (при сильной корреляции — практически полностью, при слабой — частично). Необходимо отметить, что одновременно возрастает дисперсия некоррелированной (шумовой) составляющей погрешности дифференциальных НП, поскольку эта погрешность равна сумме дисперсий шумовых ошибок измерений, подвергаемых дифференциальной (разностной) обработке.

Различие методов ДК и ОИ состоит в том, что первый имеет своей конечной целью определение координат точки нахождения потребителя в национальной или глобальной геодезической системе. При этом геодезические координаты и прочие параметры вектора состояния другой точки считают априори известными с высокой степенью точности, принимают за эталонные и используют для вычисления поправок к результатам НВО потребителя. В отличие от ДК методом ОИ решается задача определения взаимного положения двух объектов (например, воздушных или надводных судов). В такой постановке выбор системы координат, в которой фиксируют координаты обоих объектов, не является принципиальным так же, как и требование точной геодезической привязки координат одного из них.

Далее рассматриваются основные особенности и способы реализации алгоритмов ДК и ОИ. (Отметим, что с учетом общности идеи и используемого математического аппарата в ряде работ методы ДК и ОИ описывают в рамках единого *дифференциального метода*).

8.7.1. Принцип дифференциальной коррекции

Комплекс технических средств для реализации ДК (см. рис. 8.14), являющийся функциональным дополнением СРНС, принято называть *дифференциальной подсистемой* (ДПС). Основными составными частями подсистемы являются: контрольно-корректирующая станция (ККС), канал передачи корректирующей информации (КИ), аппаратура приема и обработки КИ

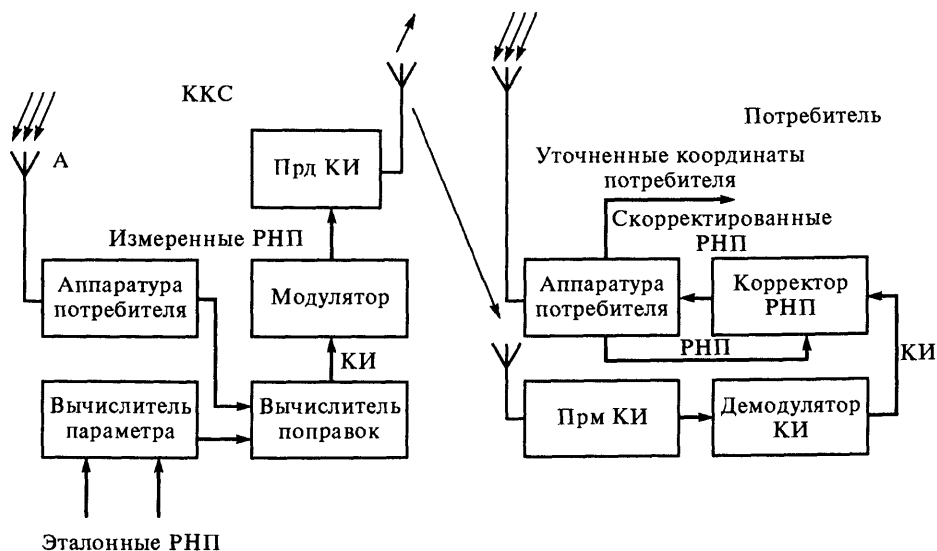


Рис. 8.14. Схема реализации метода дифференциальной коррекции

в АП. На контрольно-корректирующей станции, собственные геодезические координаты которой известны с высокой точностью и используются в качестве эталонных, размещается прецизионная АП с малым уровнем случайной погрешности измерения РНП. С помощью этой АП определяют текущие значения РНП, которые сравнивают с эталонными. Разница между текущими измерениями и эталонными значениями используется для формирования поправок, передаваемых потребителю в составе КИ. Переданная потребителю КИ используется для внесения поправок в вектор состояния, полученный его АП.

Эффективность ДК определяется погрешностями АП эталонной точки и объекта, расстоянием между ними, а также видом зависимости коэффициента корреляции ЭПД от расстояния и времени. Теоретический радиус корреляции ЭПД СРНС достигает 2000 км [117], но на практике эффективным считается использование ДК при удалениях АП от эталонной точки не более чем на 500 км.

8.7.2. Параметры корректирующей информации

По составу параметров корректируемой в АП информации могут быть выделены два метода:

- 1) метод дифференциальной коррекции координат потребителя;
- 2) метод дифференциальной коррекции измеряемых потребителем НП.

Метод дифференциальной коррекции координат предусматривает, что КИ формируется путем сопоставления эталонных координат ККС с координатами, вычисленными в ее АП. Полученные дифференциальные поправки (как разности эталонных и измеренных координат ККС) передаются в составе КИ потребителям, которые могут уточнить по ним свои координаты. Данный метод сравнительно прост, так как не меняет основной алгоритм навигационных определений АП объекта, но налагает существенные ограничения на состав рабочих созвездий ККС и АП и используемых в них алгоритмов навигационных определений. Действительно, поправки к координатам, вычисляемые на ККС по ее рабочему созвездию, могут учитываться только при условии, что АП работает по этому же созвездию и использует идентичные алгоритмы навигационных определений. При этом переносить поправки, найденные по одному созвездию, на результаты определений, полученные по другому созвездию, недопустимо, поскольку в этом случае использование КИ приводит не к уменьшению, а к увеличению погрешности навигационных определений. По этой причине в современной практике этот метод используется все реже.

Суть метода дифференциальной коррекции навигационных параметров состоит в том, чтобы сообщать потребителям набор поправок к измере-

ниям НП по всем НКА, которые потенциально может использовать их АП, предоставив возможность каждому потребителю выбирать рабочее созвездие и количество используемых для определения НКА. Иными словами, в данном методе вычисляются поправки к измеренным параметрам (например, к псевдодальности) по всем НКА, находящимся в зоне радиовидимости АП на ККС. Для этого одновременно с измерением псевдодальностей до всех НКА находят их расчетные значения, используя данные эфемерид. Результаты расчетов принимают за истинные псевдодальности, поскольку их вычисления проводят по эталонным координатам ККС. Разности расчетных и измеренных значений в качестве поправок к псевдодальномерным измерениям передаются в составе КИ всем потребителям. При этом отсутствуют ограничения на выбор алгоритмов функционирования АП, поскольку потребитель может выбрать оптимальное для него созвездие и число НКА в нем, а измеренные в его АП псевдодальности корректируются с помощью поправок, относящихся к выбранным НКА.

8.7.3. Прямой и инверсный ДМ

В зависимости от места использования КИ в подсистеме потребителей навигационной информации различают *прямой* и *инверсный* дифференциальные методы.

Прямой метод предусматривает использование КИ непосредственно в навигационной аппаратуре потребителя для уточнения параметров его вектора состояния. Этот метод получил наибольшее распространение, поэтому в большинстве типов АП имеется интерфейс для приема КИ в соответствующем формате, определенном международными или национальными стандартами.

Инверсный метод используется в тех случаях, когда необходимо уточнить вектор состояния объекта не в точке его нахождения, а в некотором другом пункте, на котором осуществляется контроль положения объекта или управление его движением. При этом с АП контролируемого объекта на пункт управления (контроля) транслируют либо результаты первичных измерений, либо текущие координаты, дополненные информацией о составе рабочего созвездия, по которому они вычислены. В пункте управления (контроля) полученные данные подвергают коррекции с использованием КИ, поступающей от ККС, которая может располагаться как в самом пункте управления, так и находиться в другом месте. Уточненные в результате коррекции навигационные данные затем могут ретранслироваться на объект или использоваться информационными системами пункта управления как в реальном времени, так и в режиме постобработки.

8.7.4. Методы относительных измерений

Наиболее простой и наглядный вариант реализации метода ОИ с использованием СРНС сводится к следующему. Пусть на двух объектах, разнесенных в пространстве, установлены комплекты АП, которые по одним и тем же созвездиям НКА в совпадающие моменты времени определяют свои геоцентрические координаты $\{x_1, y_1, z_1\}, \{x_2, y_2, z_2\}$. Вычисленные координаты по соответствующему каналу связи передаются с одного объекта на другой. На каждом объекте вычитание одноименных компонент этих векторов положения позволяет определить проекции базовой линии, соединяющей эти два объекта: $\Delta x = x_1 - x_2$, $\Delta y = y_1 - y_2$, $\Delta z = z_1 - z_2$, расстояние между объектами $D = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$, а также углы α, β, γ , характеризующие направление базовой линии в пространстве:

$$\alpha = \arccos \frac{\Delta x}{D}, \quad \beta = \arccos \frac{\Delta y}{D}, \quad \gamma = \arccos \frac{\Delta z}{D}.$$

Однако в таком предельно упрощенном виде метод ОИ в настоящее время применяется редко, поскольку выигрыш в точности, получаемый при использовании только координатной информации, не окупает усложнения системы. Современные высокоточные системы ОИ работают по принципу совместной обработки РНП измеренных АП объектов, при этом наряду с дальномерными измерениями используются измерения фазы несущей частоты сигналов НКА (см. § 8.6), что требует точной частотной и временной синхронизации АП объектов. Для исключения начального рассогласования и дрейфа шкалы времени АП используют *разностные* методы. Рассмотрим принцип применения этих методов на примере обработки фазовых измерений.

Принципиальная особенность фазовых измерений состоит в том, что, поскольку интервал однозначного определения фазы равен $(-\pi, \pi)$, измерения разности фаз между двумя АП являются однозначными только при длине базовой линии $l \leq \lambda/2$. Если длина базовой линии превосходит указанную величину, измерения содержат неизвестный параметр M_{φ_j} — целое число длин волн в разности расстояний от j -го НКА до антенн АП (далее — *параметр фазовой неоднозначности*).

Фаза сигнала j -го НКА, принятого потребителем в момент времени τ , относительно фазы сигнала ОГ АП может быть представлена в виде [120, 121]

$$\varphi(\tau) = \varphi_n^0 - \varphi_j^0 + f_j \left(\frac{R_{\varphi_j}(\tau)}{c} + \tau_j - \tau_n - \tau_{ji} + \tau_{jt} + \tau_{\text{прм}} \right) + M_{\varphi_j} + e_{\varphi_j}, \quad (8.11)$$

где $\varphi_{\Pi}^0, \varphi_j^0$ — начальные фазы сигналов генераторов АП и j -го НКА; f_j — несущая частота сигнала j -го НКА; $R_{\Pi j}$ — расстояние между j -м НКА и потребителем; τ_j — расхождение между шкалой времени j -го НКА и системной шкалой времени; τ_{Π} — расхождение между шкалой времени потребителя и системной шкалой времени; τ_{ji}, τ_{jT} — задержки, вызванные ионосферной и тропосферной рефракциями; $\tau_{\text{прм}}$ — задержка сигнала в радиочастотном тракте АП; $e_{\Pi j}$ — шумовая ошибка измерения псевдодоплеровской фазы.

Если прием сигналов j -го НКА ведется одновременно двумя потребителями Π_1 и Π_2 , то для исключения начальной фазы сигнала генератора j -го НКА и расхождения между шкалой времени j -го НКА и системной шкалой времени сформируем первую разность псевдодоплеровских фаз $\Delta_{1,2} = \varphi_1 - \varphi_2$, используя выражение (8.11):

$$\Delta_{1,2}(\tau) = \delta_{1,2}(\tau_0) + f_j \left(\frac{R_{\Pi 1j}(\tau) - R_{\Pi 2j}(\tau)}{c} - \Delta\tau_{\Pi} - \Delta\tau_{ji} + \Delta\tau_{jT} + \Delta\tau_{\text{прм}} \right) + \Delta M_j + \Delta e_j, \quad (8.12)$$

где $\delta_{1,2}(\tau_0)$ — расхождение фаз сигналов генераторов потребителей Π_1 и Π_2 ; $\Delta\tau_{\Pi} = \tau_{\Pi 1} - \tau_{\Pi 2}$ — разность расхождений шкал времени потребителей Π_1 и Π_2 и системной шкалы времени; $\Delta\tau_{ji} = \tau_{1,ji} - \tau_{2,ji}$, $\Delta\tau_{jT} = \tau_{1,jT} - \tau_{2,jT}$ — разности задержек сигналов, принятых потребителями Π_1 и Π_2 , обусловленные ионосферной и тропосферной рефракциями; $\Delta\tau_{\text{прм}} = \tau_{\text{прм}1} - \tau_{\text{прм}2}$ — разность задержки сигналов в радиочастотных трактах потребителей Π_1 и Π_2 ; $\Delta M_j = M_{\Pi 1,j} - M_{\Pi 2,j}$ — первая разность значений параметров фазовой неоднозначности; $\Delta e_j = e_{\Pi 1,j} - e_{\Pi 2,j}$ — шумовая ошибка измерения псевдодоплеровской фазы.

В соотношении (8.12) кроме координат потребителей Π_1 и Π_2 неизвестными являются значения $\delta(\tau_0)$ и $\Delta\tau_{\Pi} + \Delta\tau_{ji} - \Delta\tau_{jT} - \Delta\tau_{\text{прм}}$, одинаковые для всех первых разностей.

Для измерений по НКА СРНС GPS исключить значения расхождения начальных фаз и разностей задержек и уходов можно с помощью вычисления вторых разностей, т. е. вычитанием первой разности по любому НКА из всех остальных. Для СРНС ГЛОНАСС необходимо учитывать различие несущих частот сигналов НКА, т. е. в общем случае вторые разности псевдодоплеровских фаз можно записать в виде

$$\nabla \Delta_{jk} = f_j \frac{R_{n_1 j}(\tau) - R_{n_2 j}(\tau)}{c} - f_k \frac{R_{n_1 k}(\tau) - R_{n_2 k}(\tau)}{c} - \\ - (f_j - f_k)(\Delta\tau_n - \Delta\tau_{jn} + \Delta\tau_{jt} + \Delta\tau_{прм}) + \nabla \Delta M_j, \quad (8.13)$$

где $\nabla \Delta M_j = \Delta M_j - \Delta M_k$ — вторая разность параметров фазовой неоднозначности.

В работе [128] предложен алгоритм оценки величины $\Delta\tau_n + \Delta\tau_{jn} - \Delta\tau_{jt} - \Delta\tau_{прм}$ по измерениям псевдодалности, поэтому можно считать, что вторые разности фаз являются функцией только координат потребителей и параметров фазовой неоднозначности.

Неоднозначность фазовых измерений можно исключить, сформировав третьи разности, используя вторые разности, относящиеся к разным моментам времени, при условии непрерывности слежения приемниками за сигналами НКА между этими моментами времени:

$$\Delta \nabla \Delta_{ij}(\tau_1, \tau_2) = \nabla \Delta_{jk}(\tau_2) - \nabla \Delta_{jk}(\tau_1) = \\ = f_j \left(\frac{R_{n_1 j}(\tau_2) - R_{n_2 j}(\tau_2)}{c} - \frac{R_{n_1 j}(\tau_1) - R_{n_2 j}(\tau_1)}{c} \right) - \\ - f_k \left(\frac{R_{n_1 k}(\tau_2) - R_{n_2 k}(\tau_2)}{c} - \frac{R_{n_1 k}(\tau_1) - R_{n_2 k}(\tau_1)}{c} \right). \quad (8.15)$$

В формуле (8.15) неизвестными являются координаты потребителей P_1 и P_2 . Следовательно, погрешности, определяемые расхождением фаз и частот генераторов потребителей и НКА, оказываются скомпенсированными. Вместе с тем, число полученных таким образом измерений в четыре раза меньше первоначального числа разностей фаз. Отметим, что аналогичный результат при том же общем количестве измерений может быть получен, если включить неизвестные частотные и фазовые разности в число параметров, определяемых при решении уравнения соответствующей размерности. В настоящее время существуют и другие методы разрешения неоднозначности (см. § 8.8).

Основными источниками погрешностей определения компонент базовой линии являются [121] погрешности измерения относительной задержки сигналов, нестабильность генераторов, некомпенсированные остаточные погрешности, обусловленные ионосферной и тропосферной рефракциями; погрешности, обусловленные неточностью знания положения НКА. Вклад случайных и систематических компонент этих погрешностей в формирование погрешностей определения длины базовой линии характеризуется данными, представленными в табл. 8.1.

Таблица 8.1

Влияние различных источников на точность относительных определений

Источник погрешностей	Погрешность определения компонент базовой линии, см	
	СКО	систематическая
Измерение фазовой задержки на интервале 4 с	0,2...0,5	—
Относительная нестабильность генератора НКА	0,3	0,3
Тропосферная рефракция (после калибровки)	1,5	1,5
Ионосферная рефракция (после калибровки)	1	1
Неточность знания положения НКА	1	1

В наземных приложениях методы ОИ используются, в основном, для определения положения пунктов с погрешностью менее 0,1 м (СКО) в интересах создания геодезических сетей различного назначения. Достижимая при этом точность слежения за фазой несущей на уровне единиц миллиметров позволяет проводить измерения базовых линий до 1000 км с точностью на уровне нескольких единиц сантиметров.

В настоящее время применяют два основных класса методов ОИ: *статические*, при которых относительное положение АП в процессе измерений остается неизменным, и *кинематические*, допускающие возможность изменения этого положения [121]. Фундаментальным правилом кинематических методов является то, что все АП в этом режиме должны непрерывно сопровождать не менее четырех общих НКА в течение всего сеанса, в том числе во время передвижений. До недавнего времени высокоточные геодезические измерения, требующие использования сложных алгоритмов для обработки больших массивов данных, выполнялись в режиме постобработки.

Наиболее современным и перспективным из кинематических методов ОИ является режим измерений и обработки их результатов в реальном времени RTK (от англ. Real Time Kinematic). Особенности этого режима являются наличие радиоканала для передачи результатов первичных измерений по дальномерному коду и фазе несущей на пункт обработки со скоростью не менее 4,8 кбит/с и использование специальных алгоритмов обработки, обеспечивающих разрешение фазовой неоднозначности за минимальное время (не более 5 с).

8.8. Угломерная навигационная аппаратура

За время, прошедшее с момента создания СРНС второго поколения, сфера их применения постоянно расширялась, и сегодня они используются для решения целого ряда прикладных задач, которые на этапе создания

СРНС не ставились. Одним из ярких примеров такого рода приложений стали построенные по принципу фазового интерферометра навигационные приемники, позволяющие практически мгновенно фиксировать не только местоположение и скорость объекта, но и значения углов, определяющих его ориентацию в пространстве. Далее излагаются основные принципы построения и функционирования угломерной АП и описаны некоторые варианты ее практической реализации.

8.8.1. Параметры угловой ориентации объектов

Задача определения угловой ориентации некоторого объекта с помощью сигналов СРНС обычно рассматривается в следующей постановке.

Пусть с объектом жестко связана прямоугольная система координат, которая далее именуется *объектовой* (на практике часто в качестве осей объектовой системы рассматриваются строительные оси объекта). Тогда пространственное положение объекта полностью определяется углами ориентации осей объектовой системы координат, относительно осей топоцентрической системы координат, начало которой совпадает с началом координат объектовой системы. (Существует несколько вариантов определения топоцентрической системы координат; ниже предполагается, что ось x_T направлена на север (по истинному меридиану), ось y_T — вверх по местной вертикали, а ось z_T дополняет систему до правой системы координат и направлена на восток по горизонтали.) Наиболее часто ориентация объектовой системы координат относительно топоцентрической системы координат описывается с помощью направляющих косинусов или углов Эйлера.

Если использовать направляющие косинусы, ориентация одной системы относительно другой описывается ортогональной матрицей, элементами которой являются девять значений направляющих косинусов осей объектовой системы. Достоинство этого варианта — максимальная простота перехода из одной системы в другую. Недостаток такого описания состоит в том, что между элементами этой матрицы имеются дополнительные связи, учет которых затрудняет обработку измерений.

Описание ориентации с помощью углов Эйлера от этого недостатка свободно, поскольку указанные углы независимы и могут рассматриваться по отдельности. Это достоинство вполне окупает некоторое усложнение выражений, описывающих связь между системами координат.

В настоящее время в технике используется несколько вариантов названий углов Эйлера: в морской навигации ориентацию корпуса корабля описывают углами крена, дифферента и азимута; в аэронавигации используются понятия крена, тангажа (эквивалент дифферента) и курсового угла

(азимута); в артиллерии и смежных областях угол поворота объекта относительно поперечной оси называют углом места (возвышения) и т. д.

Ниже для описания конечного результата определения пространственной ориентации объекта используются углы Эйлера в топоцентрической системе координат: крен, тангаж, азимут. Однако при рассмотрении алгоритмов определения пространственной ориентации предпочтение отдается описанию с помощью направляющих косинусов как менее громоздкому.

8.8.2. Принцип определения угловой ориентации объектов по сигналам СРНС

Общий принцип определения угловой ориентации объекта по сигналам СРНС состоит в следующем. Пусть по результатам стандартного сеанса местоопределения рассчитаны геоцентрические координаты фазового центра антенны АП x_0, y_0, z_0 , а из эфемеридной информации известны координаты i -го НКА x_{ci}, y_{ci}, z_{ci} . Этой информации достаточно для того, чтобы рассчитать в *геоцентрической системе* координат направляющие косинусы линии визирования i -го НКА через фазовый центр антенны АП:

$$C_x = \frac{x_{ci} - x_0}{R_i}; \quad C_y = \frac{y_{ci} - y_0}{R_i}; \quad C_z = \frac{z_{ci} - z_0}{R_i}. \quad (8.16)$$

Здесь R_i — дальность от объекта до i -го НКА (см. формулу (8.1)).

Затем необходимо радиотехническим способом (подробнее об этом см. далее) измерить углы между осями объектовой системы и направлением на этот же НКА, т. е. получить значения направляющих косинусов линии визирования НКА в объектовой системе. Однако определить однозначно углы Эйлера объектовой системы координат относительно геоцентрической системы координат по данным визирования только одного НКА невозможно, поскольку остается еще одна степень свободы: возможность поворота объектовой системы вокруг линии визирования НКА. Для однозначного определения углов пространственной ориентации объектовой системы необходимо провести описанные расчеты и измерения не менее чем по двум НКА. Значения направляющих косинусов двух неколлинеарных векторов (линий визирования НКА), рассчитанные по результатам навигационного сеанса в геоцентрической и измеренные радиотехническим методом в объектовых системах координат, полностью определяют углы Эйлера для перехода от одной системы координат к другой. Затем полученные углы ориентации в геоцентрической системе можно пересчитать для топоцентрической системы и в ней определить интересующие нас углы ориентации объекта: азимут, крен и тангаж.

8.8.3. Радиointерферометрический метод измерения угловых координат

Итак, для определения угловой ориентации объекта с использованием сигналов СРНС необходимо в системе координат, жестко связанной с объектом, определить радиотехническим методом углы визирования нескольких (не менее двух) НКА. Эта задача может быть решена двумя методами: радиопеленгационным и радиointерферометрическим.

Радиопеленгационный метод предполагает использование остронаправленной антенны, позволяющей известными амплитудными методами (пеленгация по максимуму сигнала или методом равносигнальной зоны) определить направление на источник сигнала (НКА). Точность данного метода зависит от размера апертуры антенны; в частности, для того чтобы при описанных выше параметрах СРНС ГЛОНАСС и GPS обеспечить погрешность определения углов визирования источника сигнала порядка 10 угловых минут, необходима зеркальная антенна или фазированная антенная решетка с апертурой порядка 3×3 м. Очевидно, что разместить антенны с такими размерами на подвижных объектах сложно.

Поэтому все современные образцы угломерной АП применяют интерферометрический (фазовый) принцип измерений. В этом случае используется несколько (не менее трех) слабонаправленных и поэтому сравнительно небольших (с линейным размером порядка λ) антенн, которые располагаются на объекте таким образом, чтобы расстояние (база) между ними было примерно равно размеру апертуры антенны, обеспечивающей ту же точность при амплитудном методе.

Как известно, интерферометрический принцип определения направления на источник сигнала основывается на том, что разность фаз сигналов $\Delta\varphi_i$, принимаемых антеннами, разнесенными на расстояние l , пропорциональна косинусу угла θ_i между базой интерферометра и направлением на НКА:

$$\Delta\varphi_i = \frac{2\pi}{\lambda_i} l \cos\theta_i. \quad (8.17)$$

Таким образом, радионавигационным параметром в данном случае является фазовый сдвиг $\Delta\varphi$, а навигационным параметром — угол θ .

Принципиальная особенность интерферометрического метода состоит в том, что поскольку интервал однозначного определения фазы равен $(-\pi, \pi)$, измерения разности фаз в интерферометре являются однозначными только при длине базовой линии $l \leq \lambda/2$. В общем случае косинус угла между базой интерферометра и направлением на i -й НКА и измеряемая интерферометром разность фаз сигнала в разнесенных антеннах связаны соотношением

$$\cos \theta_i = \frac{\lambda_i}{l} \left(M_i + \frac{\Delta \varphi_i}{2\pi} \right)_i,$$

где M_i — неизвестное целое число циклов фазы несущей частот i -го НКА (параметр фазовой неоднозначности); $\Delta \varphi_i$ — измеренный фазовый сдвиг, $0 \leq \Delta \varphi_i \leq 2\pi$; l — длина базовой линии интерферометра.

Вследствие многозначности измерений в интерферометре невозможно точно определить целое число M_i :

— фазовых циклов в полной разности фаз;

— длин волн в разности расстояний от i -го НКА до антенн интерферометра.

Как следствие, результату фазовых измерений будет сопоставлен ансамбль из $2M_i + 1$ значений $\cos \varphi_i$, из которых лишь одно соответствует истинному. Поэтому процедура устранения (разрешения) неоднозначности фазовых измерений является обязательной составной частью алгоритма функционирования угломерной АП. Важно отметить, что после того, как неоднозначность устранена, она никак не влияет на точность решения задачи собственно угловых измерений, что позволяет рассматривать две эти задачи независимо. Поэтому в настоящем параграфе при рассмотрении алгоритма определения угловой ориентации будем полагать, что неоднозначность фазовых измерений отсутствует или устранена. Методы разрешения неоднозначности фазовых измерений рассмотрены в п. 8.8.4.

Учитывая свойства скалярного произведения векторов и тот факт, что разность расстояний от антенн до НКА равна проекции базы на линию визирования НКА, можно записать основное уравнение связи между рассчитанными в ГЦСК направляющими косинусами линии визирования i -го НКА, измеренным с помощью интерферометра косинусом угла между базой интерферометра и направлением на тот же НКА и искомыми направляющими косинусами базы интерферометра в геоцентрической системе координат:

$$\cos \theta = \mathbf{r}_0 \mathbf{l}_0 = C_x \cos \beta_x + C_y \cos \beta_y + C_z \cos \beta_z,$$

где \mathbf{l}_0 — единичный вектор базы интерферометра; $\cos \beta_x$, $\cos \beta_y$, $\cos \beta_z$ — направляющие косинусы базы интерферометра в ГЦСК; \mathbf{r}_0 — единичный вектор линии визирования НКА; C_x , C_y , C_z — направляющие косинусы вектора \mathbf{r}_0 в геоцентрической системе координат (см. формулу (8.16)).

Как уже говорилось, для однозначного определения углов пространственной ориентации объектовой системы необходимо провести измерения не менее чем по двум НКА. Кроме того, поскольку с помощью одного вектора невозможно задать все три угла пространственной ориентации, необ-

ходимо жестко связать с объектом по крайней мере два неколлинеарных вектора, т. е. использовать антенную систему из двух интерферометров. Таким образом, система уравнений, определяющая параметры угловой ориентации объекта, будет иметь вид

$$\begin{aligned} C_{x_i} \cos \beta_{x_1} + C_{y_i} \cos \beta_{y_1} + C_{z_i} \cos \beta_{z_1} &= \cos \theta_{i_1}, \\ C_{x_i} \cos \beta_{x_2} + C_{y_i} \cos \beta_{y_2} + C_{z_i} \cos \beta_{z_2} &= \cos \theta_{i_2}. \end{aligned} \quad (8.18)$$

Данная система для $i = 1, 2$ содержит четыре уравнения с шестью неизвестными — направляющими косинусами баз интерферометров в геоцентрической системе координат. Чтобы найти однозначные решения, эту систему дополняют известным уравнением связи между направляющими косинусами вектора

$$\cos^2 \beta_x + \cos^2 \beta_y + \cos^2 \beta_z = 1$$

и выражением для угла γ между базами интерферометров (угол γ определяется конструкцией антенной системы и считается известным):

$$\cos \beta_{x_1} \cos \beta_{x_2} + \cos \beta_{y_1} \cos \beta_{y_2} + \cos \beta_{z_1} \cos \beta_{z_2} = \cos \gamma.$$

Решая систему (8.18) известными методами, можно однозначно определить ориентацию баз интерферометров в геоцентрической системе координат. Поскольку положение этих баз в объектовой системе координат жестко определено, становится возможным найти ориентацию объектовой системы в геоцентрической системе координат.

Далее, используя полученные по результатам навигационного сеанса географические координаты объекта (широту B и долготу L), при помощи известной матрицы преобразования координат из ГЦСК в ТЦСК:

$$\begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = \begin{bmatrix} -\sin B \cos L & \sin B \sin L & \cos B \\ \cos B \cos L & \cos B \sin L & \sin B \\ -\sin L & \cos L & 0 \end{bmatrix} \begin{bmatrix} x_g \\ y_g \\ z_g \end{bmatrix},$$

можно рассчитать направляющие косинусы осей x_o, y_o, z_o объектовой системы координат, а затем определить углы Эйлера (азимут, крен и тангаж) с помощью следующих выражений:

$$\alpha = \arctg \frac{\cos \beta_{z_r x_o}}{\cos \beta_{x_r x_o}},$$

$$\psi = \operatorname{arctg} \frac{\cos \beta_{y_t x_0}}{\sqrt{\cos^2 \beta_{x_t x_0} + \cos^2 \beta_{z_t x_0}}},$$

$$\theta = \operatorname{arctg} \frac{\cos \beta_{y_t y_0}}{\cos \beta_{x_t y_0} \cos \beta_{z_t x_0} - \cos \beta_{x_t x_0} \cos \beta_{z_t y_0}} = \operatorname{arctg} \frac{\cos \beta_{y_t y_0}}{\cos \beta_{y_t z_0}}.$$

(В данных выражениях считается, что ось x_0 объектовой системы координат направлена вдоль строительной оси объекта, ось y_0 — вверх, а ось z_0 — поперек строительной оси.)

8.8.4. Методы разрешения фазовой неоднозначности при интерферометрических измерениях в спутниковой навигации

Методы разрешения неоднозначности фазовых измерений в угломерной АП основаны на использовании дополнительной информации, получаемой либо благодаря избыточности числа НКА в рабочем созвездии и их перемещению в процессе наблюдений, либо с помощью относительного перемещения элементов интерферометра в процессе калибровки, либо путем приема сигналов НКА на двух частотах, либо многошкальным (нониусным) методом, т. е. при использовании дополнительных антенн, расстояние между которыми меньше половины длины волны. Возможно также привлечение данных, получаемых от источников, использующих другие физические принципы определения угловой ориентации, например гироскопических или магнитных компасов.

В связи с быстрым ростом мощности вычислительных средств АП предпочтение отдают вычислительным методам, использующим процедуры оптимизации поиска решений в пространстве целых чисел, которые обеспечивают эффективное решение проблемы разрешения неоднозначности при относительно простой аппаратной реализации антенных устройств и приемников. При использовании этих методов неизвестные значения параметра фазовой неоднозначности рассматриваются как дискретные случайные переменные с конечным числом реализаций. В программной части должна быть реализована система параллельных фильтров Калмана, где каждый фильтр моделирует отдельную реализацию неизвестного параметра. Для принятия решения о том, какая реализация истинная, могут использоваться известные статистические методы: максимального правдоподобия, наименьших квадратов и др. На основе этих методов специально для разрешения неоднозначности в спутниковых навигационных приемниках разработан λ -метод [121, 122], в котором использованы эффективные процедуры декорреляции фазовых измерений и сокращения области поиска, что позволяет значительно уменьшить объем вычислений.

Заметим, что полная разность фаз сигналов, принимаемых антеннами интерферометра, зависит не только от их взаимного расположения (геометрии антенной системы), но и от ряда других факторов: неидентичности геометрических и электрических длин фидеров, неопределенности положения фазовых центров антенн, задержек в аппаратуре, систематических погрешностей фазометров АП и т. п. Неопределенность этих параметров усложняет процедуру разрешения неоднозначности и понижает итоговую точность угловых измерений.

Устранить указанную неопределенность можно с помощью дополнительных измерений по избыточным спутникам и решения системы уравнений более высокого порядка, чем (8.18). Такая процедура калибровки антенно-фидерной системы обычно входит в пакет программного обеспечения угломерной АП, при этом возможны два варианта ее проведения. Если антенно-фидерная система имеет жесткую конструкцию, обеспечивающую ее неизменность в процессе работы, то калибровку целесообразно провести до начала работы, при этом полученные в результате калибровки параметры запоминаются как константы, используемые при решении системы уравнений (8.18). Если же в процессе работы возможны изменения каких-либо параметров, то они включаются в систему как неизвестные и определяются многократно. Очевидно, что в последнем случае необходимо в течение всего сеанса иметь возможность привлекать измерения по избыточным НКА. Возможен также комбинированный вариант, когда результаты начальной калибровки уточняются по каждому новому НКА, включаемому в рабочее созвездие.

Очевидно, что если некоторые параметры антенной системы, например длины баз, известны, то они могут быть использованы в качестве начальных условий процедуры устранения неоднозначности, что упрощает калибровку.

8.8.5. Факторы, ограничивающие точность угломерной АП СРНС

Как следует из формулы (8.17), теоретически угловая чувствительность интерферометра неограниченно возрастает при увеличении базы l . На практике точность угломерной АП определяется рядом факторов, поэтому ее зависимость от длины базы имеет более сложный характер. В частности, существенными являются нарушения фазовой структуры сигнала, возникающие в случаях, когда сигнал в точке приема является суперпозицией колебаний, приходящих непосредственно от НКА, и колебаний, переотраженных близлежащими объектами (многолучевое распространение). Кроме того, на точность измерений влияет зависимость положения фазового центра приемных антенн от угла падения и частоты сигнала (последняя проявляет-

ся при навигационно-временных определениях по сигналам СРНС ГЛОНАСС, использующей частотное разделение), а также аппаратно-алгоритмический шум фазометрического тракта приемника, который зависит от качества аппаратуры, особенностей динамики объекта и способов ее учета в используемых алгоритмах сглаживания. Оценочные значения указанных и некоторых других составляющих погрешности измерения углов при длине базы интерферометра порядка 2 м приведены в табл. 8.2 [123].

Таблица 8.2

**Типичные значения составляющих погрешности измерения углов
по сигналам СРНС**

Источник погрешности	СКО, угл. мин
1. Шумовая погрешность фазометра (при полосе дискриминатора, соответствующей ускорению 2g)	1...10
2. Изменение положения фазового центра	0,5...5
3. Многолучевое распространение	0,5...5
4. Погрешность вычисления координат НКА и потребителя	менее 0,015
5. Ионосферная и тропосферная рефракция	менее 0,015

Заметим, что интерферометрический метод по своей сути является дифференциальным, поэтому атмосферная погрешность, обусловленная неопределенностью задержки сигналов в атмосфере, не влияет на точность угловых измерений, поскольку при размерах базы в пределах 100 м она оказывается полностью коррелированной для всех антенн.

Результаты отечественных и зарубежных исследований позволяют сделать вывод, что зависимость флуктуационной составляющей погрешности угловых измерений от базы достаточно близка к теоретической (линейной) при длине базы в пределах 10 м. В табл. 8.3 приведен пример такой зависимости для АП TANS Vector (см. ниже) [124].

Таблица 8.3

**Зависимость погрешности угловых измерений
от длины базовой линии интерферометра**

Длина базовой линии, м	СКО единичного измерения, угл. мин, при частоте измерений 10 Гц
0,5	30...35
1,0	13...15
2,0	8...10
4,0	4...5
8,0	2...3

Дальнейшее увеличение базы при классической схеме, когда сигналы всех антенн передаются по кабелям в общий приемник, по-видимому, нецелесообразно из-за роста влияния таких факторов, как многолучевое распространение, потери в кабелях, нестабильность их электрической длины и т. п. Более рациональным в этом случае является метод «третьих разностей», когда каждая антенна снабжается своим комплектом АП, в которой определяются вторые разности фаз сигналов, принимаемых от всех спутников, а затем в центральном процессоре вычисляются третьи разности фаз, что позволяет устранить ошибку, возникающую от расхождения опорных генераторов приемников (см. п. 8.7.4). В работе [126] приводятся данные о том, что в ходе экспериментов на гидрографическом судне при максимальной длине базы порядка 30 м СКО угловых измерений по методу «третьих разностей» составило несколько десятков угловых секунд. Тот же порядок погрешностей получен в ходе испытаний (на самолете) построенной по классической схеме аппаратуры типа 3DF фирмы Ashtech (США) при длине базы, равной 40 м [127].

8.8.6. Особенности реализации угломерной АП

Рассмотрим два примера серийной угломерной АП, реализующей рассмотренные принципы измерения пространственной ориентации объектов.

Аппаратура TANS Vector, разработанная фирмой Trimble (США) в начале 90-х годов, представляет собой шестиканальный четырехантенный интерферометр, работающий по сигналам C/A-кода СРНС GPS [125]. Структурная схема приемника TANS Vector приведена на рис. 8.15. Как следует из рисунка, каждая из четырех антенн интерферометра: master (ведущая) и три slave (ведомых) — поочередно (с частотой 1 кГц) подключается с помощью мультиплексора ко входу приемника. Сигнал с выхода мультиплексора после переноса на видеочастоту и аналого-цифрового преобразования проходит на процессор первичной обработки.

В аппаратуре TANS Vector использован косвенный метод измерения разности фаз, состоящий в следующем. При подключении к входу мультиплексора ведущей антенны в процессоре первичной обработки через демultipлексор и фильтр канала слежения за фазой замыкается кольцо обратной связи. На выходе коррелятора формируется сигнал рассогласования, пропорциональный разности фаз сигналов на выходах ведущей антенны и опорного синтезатора. Сигнал рассогласования является управляющим для подстройки синтезатора, который, таким образом, запоминает фазу сигнала, принимаемого ведущей антенной во время ее подключения ко входу приемника.

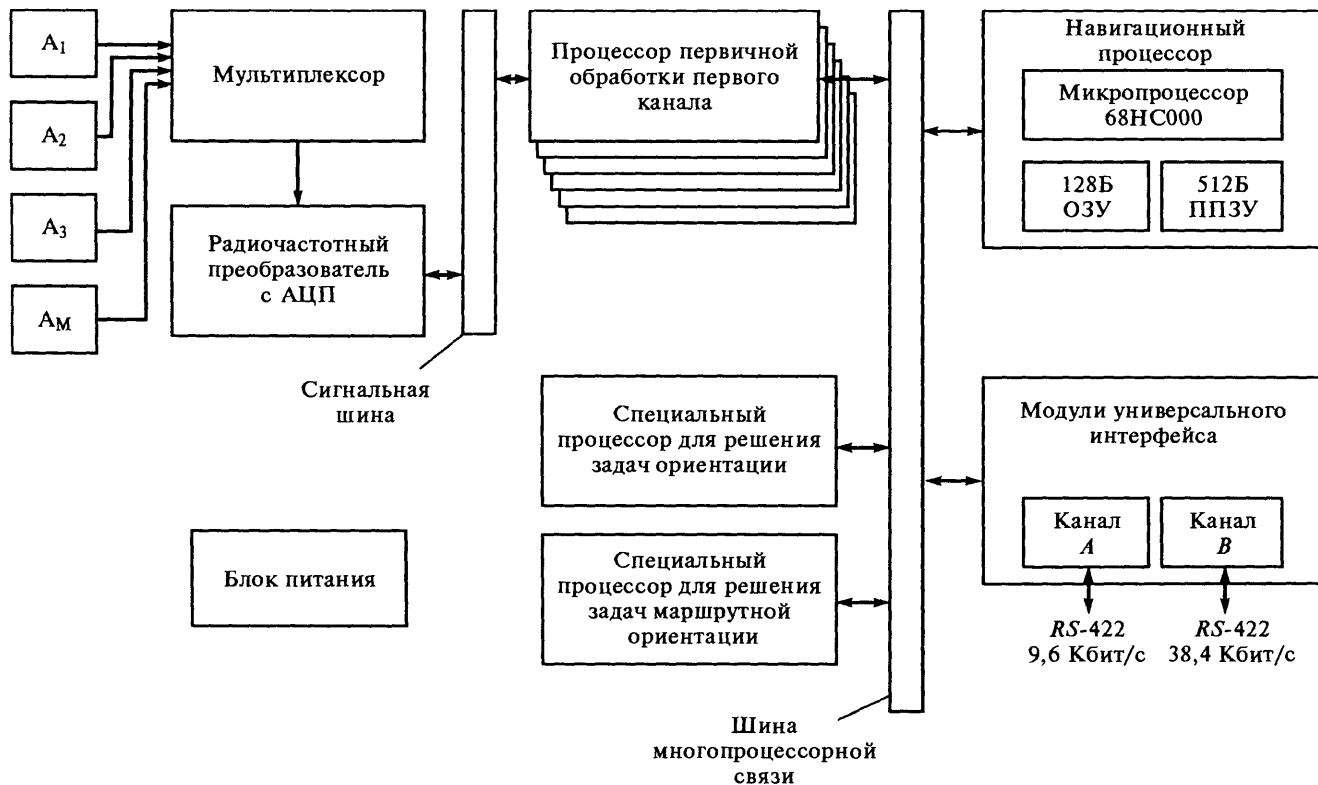


Рис. 8.15. Структурная схема угломерной АП TANS Vector

При подключении других антенн система работает как измеритель разностей фаз, сравнивая фазу принимаемого сигнала и фазу опорного сигнала с выхода синтезатора. Отказ от прямого измерения разности фаз сигналов ведущей и ведомых антенн снижает точность фазовых измерений из-за влияния нестабильности частоты синтезатора, а также ошибок слежения. Однако такое решение позволяет упростить схему и устранить ошибки, связанные с различием фазовых характеристик трактов усиления высокой частоты.

Дальнейшая обработка, включающая выделение служебной информации, вычисление координат спутников, решение навигационных задач определения положения, скорости и ориентации объекта, выполняется двумя блоками: навигационным процессором и специальным процессором определения ориентации (Attitude Processor).

Процесс первоначальной калибровки, основанный на учете собственного движения НКА, при полном отсутствии априорной информации об абсолютном и относительном расположении антенн в пространстве занимает 8...10 ч. Если потребителю известны приближенные значения углов азимута, крена, тангажа и координат, а также конфигурация антенного поля и длина фидеров, то они могут быть введены в АП, что уменьшит размерность пространства возможных решений и, как следствие, ускорит калибровку. После проведения калибровки, недостающие параметры конфигурации системы запоминаются и в дальнейшем используются для составления уравнений связи при разрешении неоднозначности. Как следует из табл. 8.2, погрешность угловых измерений АП TANS Vector при базе длиной 2 м составляет около 10 угл. мин (СКО).

Аппаратура МРК-11. Российская угломерная АП МРК-11, разработанная в 1995—1996 гг., представляет собой трехантенный интерферометр, рассчитанный на прием сигналов СРНС ГЛОНАСС и GPS. Возможность работы по сигналам двух СРНС увеличивает число одновременно наблюдаемых НКА, что позволяет применять алгоритмы, использующие измерения по избыточным НКА. Кроме того, повышается надежность измерений в условиях ограниченной радиовидимости НКА.

Структурная схема АП МРК-11 приведена на рис. 8.16. Особенностью построения данной АП является использование принципа кодового разделения сигналов НКА, поступающих с выходов трех антенн. После объединения этих сигналов в сумматоре вся их дальнейшая обработка осуществляется единым радиотрактом, что позволяет исключить систематические погрешности, обусловленные неидентичностью приемных каналов. При этом преобразование сигналов на первую ПЧ и их суммирование проводится в выносном блоке, конструктивно объединенном с антенной платформой, что позволяет использовать для передачи сигналов в приемовычислительный

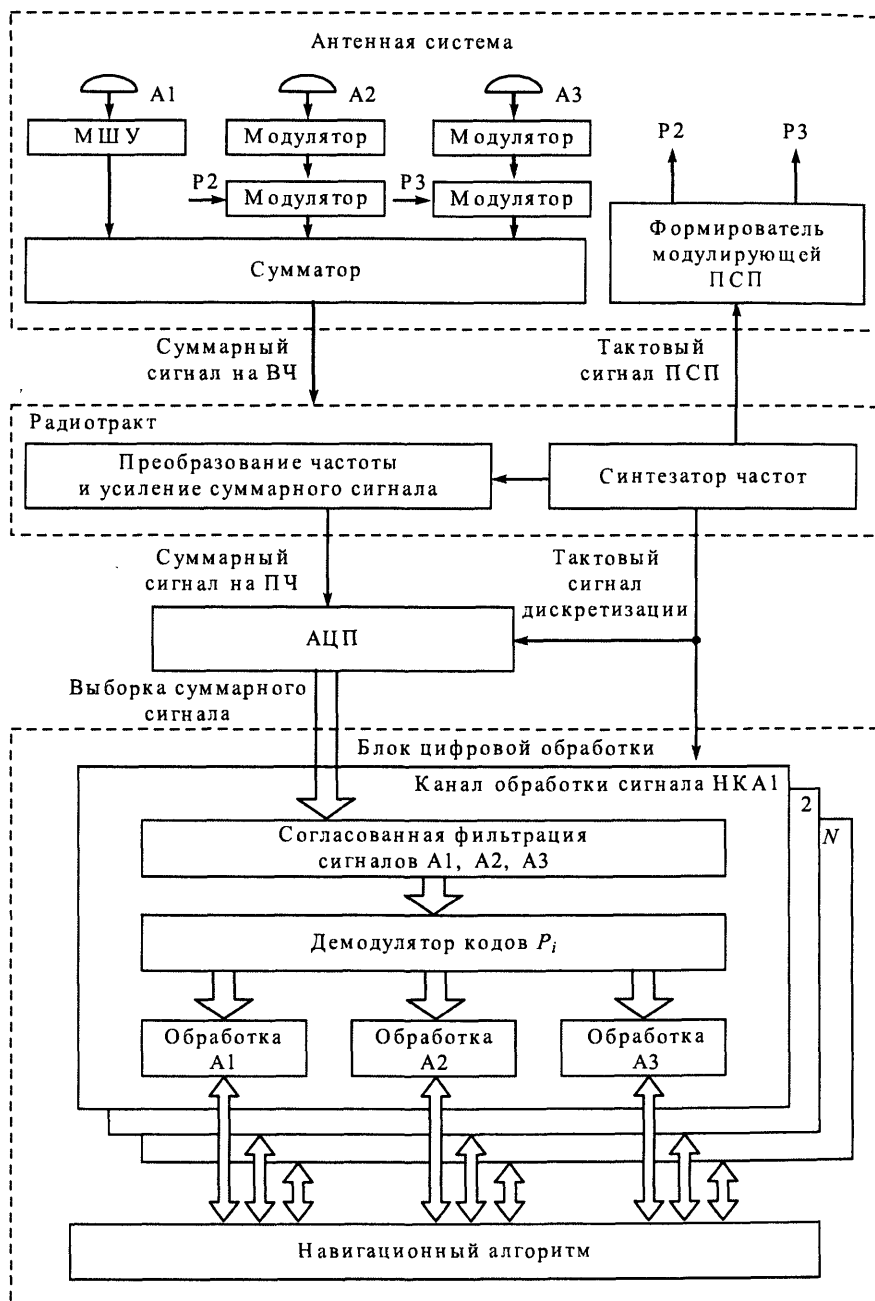


Рис. 8.16. Структурная схема угломерной АП МРК-11

блок один общий фидер. Вся последующая обработка: разделение литерных частот НКА ГЛОНАСС, свертка ПСП, поиск и сопровождение сигналов НКА по частоте и фазе — проводится в цифровой части АП. Производительность тракта цифровой обработки обеспечивает возможность одновременных измерений РНП по сигналам девяти НКА.

Для разрешения начальной фазовой неоднозначности в МРК-11 применен комбинированный метод, использующий избыточные измерения не менее чем по четырем НКА, причем учитывается их собственное движение. Данный метод позволяет в несколько раз сократить время начальной калибровки на объекте по сравнению с АП TANS Vector.

При расстоянии между антеннами МРК-11 (базе интерферометра) не менее 2 м СКО определения углов пространственной ориентации так же, как и у АП TANS Vector, составляет менее 10 угл. мин.

8.9. Информационные технологии на основе СРНС

Уникальные возможности СРНС второго поколения, общедоступность их сигналов, а также последовательное внедрение в АП СРНС достижений современной интегральной микроэлектроники способствовали созданию новых информационных технологий, нашедших применение в различных областях техники (коммерческий транспорт, геодезия, мобильная связь и др.). Наиболее эффективные технологии реализуются на основе интеграции СРНС и их подсистем с другими информационными системами (спутниковой и мобильной связи, радиомаячными, радиолокационными, радиовещательными и т. п.).

В зависимости от решаемой задачи и степени использования информации СРНС такие интегрированные системы могут быть разделены, с некоторой долей условности, на следующие классы:

- 1) функциональные дополнения СРНС, предназначенные прежде всего для повышения точности и целостности НВО, проводимых в интересах различных потребителей;
- 2) системы, в которых НВО СРНС являются основным источником информации;
- 3) комплексированные системы, в которых НВО СРНС обрабатываются совместно с данными, полученными от других источников.

Далее, в соответствии с приведенной классификацией, приведем краткий обзор современных информационных технологий, в той или иной форме использующих данные, полученные с помощью СРНС.

8.9.1. Функциональные дополнения СРНС

В настоящее время описаны следующие разновидности функциональных дополнений (ФД) СРНС:

- дифференциальные подсистемы (ДПС), которые в зависимости от размера зоны покрытия разделяются на локальные, региональные, широкозонные;

- псевдоспутники;

- ретрансляторы сигналов СРНС.

Рассмотрим принципы функционирования и основные особенности перечисленных ФД СРНС.

Локальными обычно называют ДПС с максимальной дальностью действия передатчика КИ до 200 км. Дифференциальные подсистемы этого класса широко используются для обеспечения мореплавания в прибрежных и проливных зонах, аэронавигации, системах контроля наземного транспорта, а также для геодезических, землемерных и других работ (подробнее об этих и других приложениях см. п. 8.9.2).

В морских локальных ДПС для передачи КИ обычно используют морские радиомаяки, работающие в диапазоне от 283,5 до 325,0 кГц. Для передачи КИ в этих системах используется фазовая манипуляция несущей радиомаяка с минимальным фазовым сдвигом (MSK). Этот вид модуляции выбран по той причине, что он не создает помех приемникам, использующим традиционные методы пеленгации. Скорость передачи данных в диапазоне средних волн составляет от 25 до 200 бод, на практике ее выбирают либо 100 бод, либо 200 бод. Наиболее широко в современных типах АП используется формат КИ, определяемый документом 194-93/SC104-STD, принятым Специальным комитетом 104 Морской радиотехнической комиссии (RTCM SC-104).

Поскольку сигнал радиомаяка распространяется в режиме поверхностной волны (с огибанием земной поверхности) на расстояния, значительно превышающие радиогоризонт, фактическая дальность действия радиомаяка над морской поверхностью обычно составляет от 350 до 500 км. Использование КИ на этих дальностях обеспечивает погрешность позиционирования от 0,5 до 5 м (СКО). Поскольку антенны радиомаяков, как правило, ненаправленные, сигнал, распространяющийся над поверхностью суши, может быть доступен пользователям СРНС на земле, в воздухе и во внутренних водах.

Еще одна область использования локальных ДПС — диспетчерский контроль и управление движением автотранспорта. Корректирующая информация для этой категории потребителей передается по УКВ радиовещательных каналов с использованием метода модуляции поднесущей. В

настоящее время большая часть территории Западной Европы покрыта полем локальных ДПС, использующих именно этот способ передачи КИ.

В авиации локальные ДПС используются для решения следующих задач:

- обеспечение в сложных метеоусловиях инструментального захода на посадку на дальностях порядка 50 км от начала взлетно-посадочной полосы;
- аэронавигационное обеспечение местных авиалиний;
- навигационно-временное обеспечение систем автоматических зависимых наблюдений (см. п. 8.9.2).

Авиационные локальные ДПС используют для передачи КИ УКВ диапазон частот 112...118 МГц с применением 8-кратной ОФМ. Формат передачи сообщений базируется на стандарте RTCA/DO-217. При передаче КИ по двенадцати НКА объем сообщения составляет 664 бита. Для обеспечения инструментальной посадки в более сложных условиях рассматриваются другие варианты радиоканала, в том числе диапазона 5...5,25 ГГц, а также совместное использование локальных и широкозонных ДПС (см. далее). Работа в таких условиях регламентируется стандартами RTCA/DO-245 и RTCA/DO-246.

Региональные ДПС отличаются от локальных прежде всего радиусом действия, который составляет 500...2000 км. В состав региональных ДПС входят до нескольких десятков ККС, а также системы контроля целостности информации. Так, американские региональные ДПС Starfix и SkyFix используют для формирования КИ данные 60 наземных ККС, размещенных по всему миру. Для передачи КИ используют 4 геостационарных ИСЗ ИНМАРСАТ, канал передачи данных которых работает в диапазоне *L*, скорость передачи информации составляет от 600 до 2400 бит/с, формат сообщения соответствует стандарту RTCM SC-104.

Согласно опубликованным данным на расстояниях порядка 1000 км указанные региональные ДПС обеспечивают точность местоопределения порядка 1 м (СКО), а на расстояниях до 2000 км — примерно 3 м (СКО).

Несколько другой подход к проблеме передачи КИ в региональных ДПС применен в проекте Eurofix, разрабатываемом совместно специалистами России и ряда стран Западной Европы. Здесь предлагается использовать в качестве средства передачи КИ передающие станции РНС типа Лоран-С (Чайка). Эти станции работают в диапазоне 100 кГц и имеют радиус действия порядка 1000 км. Скорость передачи данных в таких системах составляет от 15 до 30 бит/с, а сообщение, содержащее КИ для каждого НКА, имеет длину 45 бит, т. е. передача полного пакета КИ для всех НКА может занимать 10 с и более. Для увеличения скорости передачи до 175 бит/с предлагается использовать импульсно-фазовую модуляцию. Основным достоинством такой системы является низкая стоимость, поскольку применяется уже существующее оборудование, а также лучшее, чем для УКВ диапазона, распространение сигнала в

горной местности и в условиях городской застройки. Расчетная точность местоопределения для системы Eurofix имеет порядок 5м (СКО).

Широкозонные ДПС имеют ту особенность, что дифференциальные поправки в них рассчитываются не для вектора состояния конкретной ККС, находящейся в некоторой точке земной поверхности, а для вектора положения и погрешностей бортовой ШВ каждого НКА. Эти поправки формируются на основе информации от нескольких контрольных (мониторинговых) станций, разнесенных на большие расстояния (до нескольких тысяч километров). Принципы, положенные в основу широкозонных ДПС, позволяют охватить дифференциальным полем значительную территорию (существенно превышающую суммарную зону действия отдельных ККС, входящих в состав такой ДПС) при использовании минимального числа ККС и наличии определенной свободы выбора мест их размещения.

В состав каждой широкозонной ДПС должны входить [117]:

- главная ККС;
- 3–5 периферийных мониторинговых станций;
- наземные и спутниковые радиоканалы доведения КИ до потребителей в пределах рабочей зоны ДПС;
- система передачи данных от мониторинговых станций на ККС.

В настоящее время функционируют две широкозонные ДПС: WAAS (от англ. Wide Area Augmentation System), созданная по заказу Федеральной авиационной администрации и обслуживающая территорию США, и EGNOS (от англ. European Geostationary Navigation Overlay Service), обслуживающая территорию Европы.

В состав этих широкозонных ДПС входят геостационарные ИСЗ типа ИНМАРСАТ-3 для ретрансляции КИ и проведения дополнительных дальномерных измерений по ретранслированным сигналам *L*-диапазона.

Дифференциальные подсистемы на основе псевдоспутников. Наряду с традиционными все более широкое распространение приобретает схема построения ДПС с использованием так называемых *псевдоспутников* (Pseudolite). В настоящее время этим термином определяется комплекс технических средств, располагаемый в эталонной точке на земной поверхности и обеспечивающий реализацию двух основных функций:

- функции собственно псевдоспутника, т. е. формирование и излучение дальномерных сигналов, структура и несущая частота которых аналогичны сигналам НКА;
- функции ККС, т. е. формирование КИ и включение ее в состав навигационного сообщения, передаваемого потребителю.

Включение псевдоспутника в состав рабочего созвездия бортовой АП позволяет улучшить качество навигационных определений: уменьшить геометрический фактор и повысить надежность и достоверность местоопреде-

лений. Кроме того, при использовании псевдоспутника не нужен дополнительный канал связи для передачи КИ, а сам псевдоспутник может выполнять функции локального средства контроля для оценки работоспособности бортовой АП.

Ретрансляторы сигналов СРНС. Еще одним перспективным вариантом ФД СРНС являются системы, использующие принцип ретрансляции (переизлучения) сигналов НКА. Основным элементом рассматриваемого ФД является специальное устройство (ретранслятор), осуществляющее прием сигналов СРНС, их преобразование и излучение с полным или частичным сохранением информации, содержащейся в сигналах НКА. Сигналы, излучаемые ретранслятором, могут использоваться потребителем для НВО как наряду с сигналами, принятыми непосредственно от НКА, так и независимо от последних. Например, ретранслятор, координаты которого известны с высокой точностью, может выполнять функцию псевдоспутника. Сеть таких ретрансляторов позволяет дополнять (или даже в значительной степени заменять) в локальной зоне орбитальную группировку НКА, что весьма важно в условиях ограниченной радиовидимости, например в горных условиях. Кроме того, результаты НВО, выполняемых в АП по ретранслированным сигналам, позволяют при известных координатах ретранслятора рассчитать дифференциальные поправки непосредственно в самой АП и таким образом повысить точность НВО по сигналам НКА СРНС.

Еще одно применение принципа ретрансляции связано с использованием сигналов СРНС для внешнетраекторных измерений в процессе испытаний объектов авиационной и ракетно-космической техники.

Особенность проблемы внешнетраекторных измерений высокодинамичных объектов состоит в том, что применение стандартной АП для этих целей часто неэффективно из-за аппаратурных и алгоритмических ограничений рабочего диапазона значений скорости и ускорения, времени «горячего» старта и перезахвата рабочего созвездия, недостаточной точности измерения малых относительных расстояний, низкого (порядка 1 с) темпа обновления информации и ряда других факторов.

Ретрансляционные методы свободны от этих ограничений, поскольку обработка сигнала на борту объекта сводится к преобразованию спектра принятого сигнала, т. е. является практически безынерционной операцией. Навигационно-временные определения, осуществляемые на приемном пункте, могут проводиться как в реальном времени, так и в режиме постобработки с использованием высокопроизводительных вычислительных средств и специального программного обеспечения, что позволяет повысить их точность и достоверность по сравнению с вариантом использования бортовой АП. Поэтому системы внешнетраекторных измерений, построенные по принципу ретрансляции сигналов СРНС, представляют большой практический интерес.

8.9.2. Транспортные информационно-управляющие системы, использующие сигналы СРНС

Типичным примером современных информационных технологий, основанных на использовании данных НВО СРНС, могут служить транспортные информационно-управляющие системы (ТИСУ, англ. ITS — Intelligent Transportation Systems), аэронавигационные системы автоматических зависимых наблюдений (АЗН, англ. ADS — Automatic Dependent Surveillance), морские автоматические информационно-идентификационные системы (АИИС) [117, 129]. В основу построения перечисленных систем положен единый принцип: на борту подвижного объекта (морского или воздушного судна, другого транспортного средства) устанавливается АП, позволяющая определять местонахождение объекта. Данные НВО, а также другая информация от бортовых систем передаются на диспетчерские центры по соответствующему радиоканалу (линии связи). Далее дается краткий обзор состояния и перспектив развития перечисленных систем.

Аэронавигационные системы автоматических зависимых наблюдений (АЗН) рассматриваются Международной организацией гражданской авиации как один из важнейших информационных каналов, который наряду с первичными и вторичными радиолокаторами будет составлять основу перспективных комплексов управления воздушным движением. Принцип АЗН предусматривает установку на борту летательного аппарата оборудования, позволяющего определять его координаты по сигналам СРНС и постоянно передавать эти данные по радиоканалу на наземные диспетчерские центры, а также остальным участникам воздушного движения (см. ниже).

В настоящее время наиболее широко применяется система ACARS (от англ. Airline Communication Addressing and Reporting System). Аппаратура ACARS, используемая более чем на 7000 воздушных судах, может работать в диапазонах КВ, а также в метровом и дециметровом диапазонах волн (последний используется для организации спутниковой связи через систему ИНМАРСАТ). Система обеспечивает передачу речи в аналоговой форме, а также передачу цифровых данных со скоростью 2400 бит/с. Структура сообщения при передаче цифровых данных в системе ACARS определяется стандартом ARINC-618, при этом используется самоуправляемый доступ к радиоканалу с контролем несущей CSMA (от англ. Carrier Sense Multiple Access).

В качестве дальнейшего развития технологии АЗН Комитетом ИКАО по будущим аэронавигационным системам рассматривается концепция VDL (от англ. VHF digital link). Эта концепция предусматривает ряд режимов (версий): VDL-1, VDL-2, VDL-3, VDL-4, отличающихся скоростью передачи информации, используемыми видами модуляции, протоколами доступа к каналам передачи данных и т. п.

Наибольшие возможности для организации различных вариантов передачи данных предусмотрены в режиме VDL-4, который реализует метод самоорганизующегося многостанционного доступа с временным разделением каналов STDMA (от англ. Self Time Division Multiple Access). При этом каждый абонент передает свою информацию в течение временного интервала (слота), выделяемого ему системой по предварительному запросу. Время начала и конца каждого слота жестко синхронизировано с использованием системного времени СРНС. Для увеличения скорости передачи информации при ограниченной полосе используемых частот (12,5 кГц) применяется 8-кратная ОФМ. Основным приложением VDL-4 является так называемый вещательный режим АЗН (АЗН-В), при котором сообщения, передаваемые с борта каждого воздушного судна, могут приниматься всеми другими воздушными судами, что позволяет организовать высокoeffективную систему предупреждения столкновений в воздухе, а также использовать самолеты, находящиеся в зоне радиосвязи с диспетчером, для ретрансляции сообщений, переданных с бортов тех воздушных судов, непосредственная связь которых с диспетчером в данный момент отсутствует или затруднена. Общее число воздушных судов, охваченных системой АЗН-В, может достигать десятков тысяч.

Морские автоматические информационно-идентификационные системы в соответствии с резолюцией ИМО (от англ. International Maritime Organization) должны решать три основные задачи:

- обмен навигационными данными между судами в целях предупреждения их столкновения в море;
- передача с судна навигационных данных в береговую службу управления движением судов для обеспечения более точной и надежной их проводки в зоне действия системы;
- передача данных о судне и его грузе в береговые службы при нахождении судна в районах с обязательным оповещением.

Для решения указанных задач на судне устанавливается АП СРНС и отдельный приемопередатчик (транспондер АИИС). С начала 2003 г. согласно требованиям Международной конвенции транспондеры АИИС должны быть установлены на всех морских судах [129]. Особенность принятой концепции организации связи в АИИС состоит в том, что правила многостанционного доступа к каналам передачи данных для судов, находящихся в открытом море и в прибрежной зоне, различаются: в открытом море транспондеры всех судов равноправны и используется протокол многостанционного доступа STDMA, сходный с протоколом VDL-4; в прибрежной зоне функция организации распределения каналов связи между судовыми транспондерами возлагается на береговые транспондеры. При этом так же, как и в системе АЗН-В, сигналы СРНС используются не

только для определения местоположения, но и для синхронизации транспондеров.

Наземные транспортные информационно-управляющие системы решают следующие задачи:

- обеспечение безопасности движения;
- организация перевозок пассажиров и грузов в процессе хозяйственной деятельности;
- слежение за особо ценными и экологически опасными грузами;
- защита от угона, поиск угнанных средств и т. д.

Для передачи данных с транспортного средства используются различные каналы связи: выделенные, транкинговые, спутниковые или сотовые (в последнем случае могут применяться как основные каналы, так и каналы службы коротких сообщений). Диспетчерские центры таких систем строятся на основе геоинформационных систем, обеспечивающих отображение транспортного средства на фоне электронной карты, а также поддержку ряда технологических сервисных функций (контроль датчиков охранной сигнализации, контроль режимов работы двигателя, контроль температуры в кузове рефрижератора и т. п.)

Наряду с диспетчерскими системами данные от бортовой АП транспортного средства все чаще используют в системах информационной поддержки работы водителя. В частности, в оборудование современных автомобилей все чаще включаются дисплеи, на которых могут отображаться текущее положение транспортного средства на карте, информация о ситуации на дорогах и другая информация, получаемая по радиовещательным каналам.

Во многих странах начаты работы по применению спутниковых навигационных технологий на железнодорожном транспорте в целях повышения безопасности движения, оптимизации режимов ведения поезда, контроля местоположения и скорости поезда на электронной карте диспетчерского центра железной дороги, автоматической регистрации графика движения.

В состав программного обеспечения терминала машиниста таких систем входит база данных, содержащая информацию обо всех особенностях пути и маршрутное задание машиниста, например сведения о координатах границ блок-участков, данные о профиле пути, ограничения скорости на перегонах, координаты, названия и схемы станций, расстояния до мест ограничений и повышенного внимания. Комплексное использование данных приемника СРНС и информации базы данных позволяет осуществлять непрерывное и надежное сопровождение локомотива в условиях местности со сложным рельефом, приводящим к потере радиовидимости НКА в таких местах, как туннели, ущелья и др.

8.9.3. Вторичные эталоны времени и частоты на основе АП СРНС

Потребность в недорогих, компактных, автономных эталонах времени и частоты весьма велика в различных областях науки и техники, где требуется высокоточное измерение абсолютных и относительных временных интервалов: радиофизике, оптической и радиоастрономии, радиоинтерферометрии, ядерной физике, установках по исследованию кратковременных процессов, а также в различных технических системах и комплексах (системы электроснабжения, многопозиционные радиотехнические системы, полигонные комплексы траекторных измерений, системы связи со сложными сигналами, сети мобильной связи и т. п.). Общеизвестно, что в настоящее время наиболее экономичное и изящное решение данной проблемы базируется на использовании сигналов СРНС. Кратковременная относительная нестабильность периода этих колебаний имеет порядок единиц пикосекунд (10^{-12} с), а максимальная долговременная погрешность не превышает единиц микросекунд (подробнее об этом см. § 8.6). С учетом всех ошибок, связанных с формированием, излучением, распространением, приемом и обработкой сигналов, погрешность измерения временных интервалов с помощью стандартных радионавигационных приемников составляют десятки наносекунд.

В настоящее время ведущие отечественные и зарубежные фирмы приступили к выпуску специальных спутниковых приемников (GPS Clock), предназначенных для точного измерения временных интервалов и синхронизации внешних устройств. Местный опорный генератор таких приемников с помощью схемы фазовой автоподстройки частоты синхронизируется сигналом СРНС. Благодаря применению в таких устройствах прецизионных кварцевых генераторов удается обеспечить высокие метрологические характеристики этого типа АП.

Например, предел допускаемой абсолютной погрешности воспроизведения опорной шкалы времени по сигналам СРНС в АП такого класса не превышает ± 50 нс при условии точной топопривязки его антенны. При этом предел допускаемой абсолютной погрешности взаимной синхронизации шкал времени двух образцов АП обычно не превышает ± 20 нс при работе по сигналам нескольких НКА и, соответственно, ± 10 нс при приеме сигналов одного (общего) НКА.

Стоимость таких спутниковых синхронизаторов лежит в пределах 2...10 тыс. долларов США, что позволяет широко использовать их в различных приложениях. В частности, широкое применение нашли приборы синхронизации сетей сотовой связи стандарта CDMA (IS-95), пейджинговых сетей и т. п. [130].

Типичные характеристики стабильности частоты АП для синхронизации представлены в табл. 8.4.

Таблица 8.4

Наименование нормированной метрологической характеристики	Интервал времени измерения		
	1 с	1 ч	12 ч
Предел допускаемой среднеквадратической относительной погрешности формирования частоты при работе от внутреннего ОГ, не более, с	10^{-10}	$4 \cdot 10^{-11}$	10^{-11}
Предел допускаемой погрешности измерения относительного отклонения частоты внешнего ОГ, не более, с	$\pm 4 \cdot 10^{-11}$	$\pm 2 \cdot 10^{-11}$	$\pm 10^{-11}$

8.9.4. Использование СРНС в геодезии и для мониторинга деформации земной поверхности

Благодаря высокой точности, оперативности, независимости от погодных условий, отсутствию требований прямой видимости между определяемыми точками методы измерений с использованием фазы несущей сигналов СРНС находят все более широкое применение в инженерной геодезии, геодинاميке и других областях техники, где требуется высокоточное оперативное определение взаимного положения и перемещений объектов или элементов их конструкций, а также в области контроля геометрии крупных инженерных сооружений.

Новое поколение геодезической специализированной АП существенно отличается от предыдущих моделей приемников. Основных отличий два: изменение архитектуры сигнальных процессоров и их ПО с целью минимизации влияния эффекта многолучевости, а также внедрение RTK-алгоритмов обработки фазовых измерений в качестве основных режимов работы навигационного процессора. В результате этого на практике стало возможно достижение следующих уровней погрешности относительного позиционирования антенн двух геодезических АП в реальном времени:

— $(1 + L \cdot 2 \cdot 10^{-6})$ см по координатам и $(2 + L \cdot 2 \cdot 10^{-6})$ см по высоте 1 раз в секунду с задержкой в получении скорректированных измерений менее 0,4 с, где L — расстояние между АП;

— $(3 + L \cdot 2 \cdot 10^{-6})$ см по координатам и $(5 + L \cdot 2 \cdot 10^{-6})$ см по высоте 5 раз в секунду с задержкой в получении скорректированных измерений менее 0,2 с.

В качестве примеров практического применения вышеуказанной АП можно указать:

- контроль положения секций мостов при их монтаже;
- контроль деформаций и положения ниток трубопроводов во время их укладки в траншею или на дно моря;

— мониторинг деформаций протяженных конструкций, например контроль прогиба мостов под нагрузкой.

Другим характерным примером является мониторинг различных машин и механизмов, работающих в открытых карьерах. Обычно в районе карьера или горных разработок организуется локальная ДПС, передающая поправки к измерениям по фазе в реальном времени для реализации режима РТК. В зависимости от технических требований к точности позиционирования того или иного механизма и типа используемой АП в режиме РТК достигаются погрешности ОИ от единиц сантиметров (при контроле положения ковша экскаватора относительно кузова грузовика) до метра (при контроле положения машин на территории карьера).

Использование фазовых относительных измерений по сигналам НКА эффективно и в тех случаях, когда контролируемые перемещения не велики по размерам и происходят достаточно медленно, например геодинамические деформации земной коры. В последнем случае необходимо контролировать относительные перемещения менее 10 мм на больших расстояниях (как правило, более 100 км), поэтому обработку результатов первичных измерений проводят по большой выборке (записи в течение суток и более). В настоящее время ряд пунктов контроля по всей Земле обменивается такими данными в рамках Международной программы контроля тектонических разломов и плит [117]. В указанных примерах использование радиоканала для передачи данных измерений в реальном времени необязательно и можно применять хорошо отработанные методы постобработки статических измерений по фазе сигналов НКА.

8.9.5. Комплексируемые системы навигации

Типичным примером систем, использующих данные СРНС совместно с данными, полученными от других источников, могут служить современные интегрированные навигационные комплексы воздушных и морских судов.

До появления СРНС наиболее совершенными средствами навигации и определения пространственной ориентации мобильных объектов самых различных классов — от самолетов до подводных лодок — являлись гироскопические и инерциальные навигационные системы (ИНС). Ряд несомненных достоинств таких систем, важнейшим из которых является их автономность (независимость от полей внешних источников), определяет необходимость их дальнейшего развития и применения совместно с АП СРНС.

Совместное применение (комплексирование) АП СРНС и ИНС не только обеспечивает необходимое резервирование источников информации, но и позволяет значительно уменьшить итоговую погрешность измерений,

поскольку физическая природа, а соответственно, корреляционные свойства погрешностей для этих систем существенно различаются.

Действительно, погрешности ИНС в основном связаны с внешними магнитными и гравитационными возмущениями, механическими характеристиками чувствительного элемента, потерями на трение и т. д. Как следствие, высокочастотная составляющая этих погрешностей невелика по сравнению с соответствующими погрешностями СРНС. Однако обусловленная накоплением систематических ошибок инфранизкочастотная составляющая погрешностей ИНС (дрейф) с увеличением времени наблюдения нарастает, что приводит к необходимости систематической юстировки (переустановки) системы.

Напротив, погрешности НВО СРНС характеризуются относительно более высоким уровнем высокочастотного шума, однако этот шум имеет нулевое среднее, и эффект накопления низкочастотной составляющей погрешностей в АП СРНС практически отсутствует. Таким образом, совместная обработка (фильтрация) результатов измерений ИНС и АП СРНС позволяет реализовать высокоточную «бездрейфовую» систему навигации и пространственной ориентации объектов, в которой мгновенная погрешность почти полностью определяется ИНС, а долговременная — АП СРНС.

В зависимости от организации алгоритмов совместной фильтрации измерений датчиков различной физической природы выделяют следующие схемы комплексирования[117]: разомкнутую, слабосвязанную, сильносвязанную и глубокоинтегрированную. В частности, в разомкнутой схеме АП СРНС и ИНС функционируют независимо, а совместная фильтрация начинается только на уровне координат. Напротив, глубокоинтегрированная схема предполагает наличие единого навигационного фильтра первичных измерений комплексированных датчиков.

Отметим, что комплексирование ИНС с АП СРНС позволяет существенно снизить требования к классу ИНС, а следовательно, и стоимость последней, поскольку именно долговременная погрешность является фактором, в значительной степени определяющим цену ИНС.

Контрольные вопросы

1. Каковы основные принципы построения и функционирования СРНС? Какие основные подсистемы включает в себя обобщенная структурная схема СРНС? Каковы структура и параметры орбитальной группировки? Из чего состоит подсистема КИК?
2. Каковы общие свойства и различия структуры сигналов и сообщений СРНС ГЛОНАСС и GPS?

3. Что такое навигационные и радионавигационные параметры, навигационные функции?
4. Чем различаются дальномерный, псевдодальномерный и разностно-дальномерный методы? В чем суть радиально-скоростного метода?
5. Как выглядит обобщенная структурная схема аппаратуры потребителя? Какие основные функции она выполняет?
6. Опишите принципы и устройства первичной обработки навигационной информации. Как осуществляется поиск сигналов по задержке и частоте?
7. Как производится фильтрация радионавигационных параметров и дешифрация навигационной информации?
8. Какие задачи решаются при вторичной обработке навигационной информации?
9. Какова точность навигационно-временных определений в СРНС? Перечислите факторы, влияющие на точность.
10. Назовите источники погрешностей НКА, КИК и погрешностей на трассе распространения.
11. Назовите источники погрешностей АП. Что такое геометрический фактор?
12. Каковы принципы дифференциальных и относительных измерений в СРНС?
13. Какие параметры корректирующей информации используются на практике?
14. Чем различаются прямой и инверсный ДМ?
15. Какие методы относительных измерений используют в СРНС?
16. Что такое угломерная навигационная аппаратура? Какими параметрами описывается угловая ориентация объектов? В чем состоит принцип определения угловой ориентации объектов по сигналам СРНС?
17. Опишите радиоинтерферометрический метод измерения угловых координат.
18. Какие методы разрешения фазовой неоднозначности при интерферометрических измерениях используются в спутниковой навигации?
19. Какие факторы ограничивают точность угломерной НАП СРНС? Какова достижимая точность угловых измерений?
20. Что такое функциональное дополнение СРНС? Какие виды ФД в настоящее время используют?
21. Опишите локальные, региональные и широкозонные ДПС.
22. Опишите основные виды транспортно-информационных систем с использованием СРНС.
23. Какие преимущества дает комплексирование спутниковых и других навигационных систем?

9. РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Рассмотрены основы теории и принципы построения систем передачи дискретной информации, модели сообщений и каналов, основные информационные характеристики, вопросы выбора сигналов и способы их обработки как в одноканальных, так и в многоканальных системах. Проанализированы помехоустойчивость и основные направления повышения эффективности радиотехнических систем передачи информации. Большое внимание уделено вопросам помехоустойчивого кодирования, синхронизации цифровых систем передачи информации, а также перспективным системам мобильной связи.

9.1. Структурная схема и основные характеристики цифровых радиотехнических систем передачи информации

Современные системы передачи сообщений по радиоканалам используют, в основном, цифровые методы. В цифровых системах передачи информации (ЦСПИ) сообщение (речь, текст, подвижное и неподвижное изображение, данные и т. д.) преобразуются в последовательность цифр, которые записываются в виде последовательностей символов (кодовых слов) и передаются по радиоканалу. Преимуществом ЦСПИ является высокое качество передачи информации. Универсальность формы сигнала позволяет объединять их в более крупные системы и комплексы. Примером являются цифровые сети связи с интеграцией служб, в которых каждый абонент сети (вместо привычного аналогового телефонного канала связи с полосой пропускания 3,1 кГц) на первом этапе их развертывания будет иметь два канала со скоростью передачи 64 кбит/с и один со скоростью 16 кбит/с, а на втором этапе — высокоскоростные каналы со скоростью передачи от единиц до сотен мегабит в секунду. Во многих развитых странах эта задача частично или полностью уже решена. Цифровыми информационными сетями являются сотовые системы подвижной связи стандартов D-AMPS, GSM и CDMA, а также практически все современные спутниковые системы связи. Особенно-

стью построения цифровой информационной сети является использование в них пакетной передачи (информационная сеть с коммутацией сообщений), при которой цифровой поток от одного источника разбивается на пакеты, передаваемые независимо в соответствии с определенным протоколом взаимодействия. Такой режим позволяет повысить пропускную способность сети при заданном качестве и числе каналов.

Основным элементом цифровой информационной сети является ЦСПИ, включающая в свой состав совокупность технических средств между источником сигнала (ИС) и ближайшим получателем сигнала (ПС) и обеспечивающая так называемый физический уровень. Для работы в составе цифровой информационной сети ЦСПИ дополняется устройством управления (контроллером), поддерживающим канальный и транспортный уровень в соответствии с протоколом взаимодействия.

Внешними характеристиками вход—выход ЦСПИ являются качество и скорость передачи информации [11]. Качество передачи характеризуется вероятностью ошибки $P_{\text{ош}}$ в приеме одного символа переданного сообщения, а для каналов с переменными параметрами дополнительно еще и надежностью по помехоустойчивости, под которой понимают вероятность $P(P_{\text{ош}} \leq P_{\text{доп}})$ выполнения неравенства $P_{\text{ош}} \leq P_{\text{доп}}$, где $P_{\text{доп}}$ — допустимая вероятность ошибки.

Различают техническую и информационную скорости передачи. Техническая скорость — это количество посылок (символов), передаваемых в секунду: $R = 1/T_c$ (бод), где T_c — длительность посылки, а информационная — количество двоичных единиц информации в секунду I (бит/с), переданных по каналу связи. Различие указанных скоростей связано с необходимостью передачи дополнительной служебной информации, возможностью использования многопозиционного кода и потерями информации в канале. Если потерь и служебной информации в канале нет, то $I = R \log_2 m$, где m — основание кода. Для двоичных каналов без потерь эти скорости совпадают, поэтому часто используют только термин «скорость передачи информации».

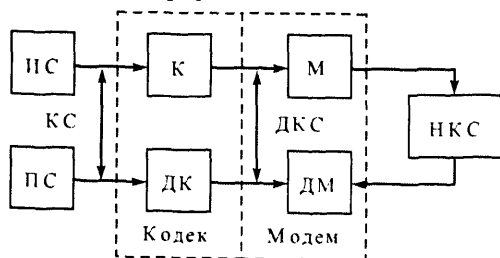


Рис. 9.1. Структурная схема одноканальной ЦСПИ

Структурная схема одноканальной ЦСПИ приведена на рис. 9.1. Она включает непрерывный канал связи (НКС), модулятор—демодулятор (модем) и кодер—декодер (кодек).

Если первичный сигнал источника является аналоговым, то он на входе претерпевает аналого-цифровое преобразование, а

на выходе обратное цифрово-аналоговое преобразование (ЦАП).

Кодек преобразует одну кодовую последовательность символов в другую, при этом может изменяться их тактовая частота, основание, структура и разрядность кодовых слов. Основными задачами, которые решает кодек, являются следующие:

- устранение избыточности в получаемом от источника сигнале с целью снижения его технической скорости при фиксированном основании кода (в пределах до его производительности — экономное кодирование);
- внесение избыточности для повышения достоверности передачи сообщения по каналу связи (помехоустойчивое кодирование);
- изменение структуры сигнала для затруднения несанкционированного извлечения информации из принятого цифрового сигнала (секретное или криптостойкое кодирование).

Модем преобразует последовательность символов в последовательность радиосигналов, согласованных по своим параметрам с радиоканалом. В простейшем случае на передающей стороне каждому символу из алфавита с основанием m ставится в соответствие радиосигнал определенной формы, а на приемной стороне проводится обратная операция. Процессы модуляции и демодуляции могут быть и более сложными. Например, на передающей стороне группе символов ставится в соответствие радиосигнал и тем самым, по сути, изменяется основание кода. На приемной стороне демодулятор может работать с так называемым мягким решением, когда принятому радиосигналу не ставится в соответствие определенный символ из того же алфавита, а вычисляется только его апостериорная вероятность или принимается решение о его ненадежности (стирание символа). Необходимо отметить, что в последнее время наметились тенденции к совмещению функций помехоустойчивого кодера и модема, появился термин «сигнально-кодовые конструкции». Этому в существенной мере способствует высокий уровень информационной технологии, при котором решение сложнейших вычислительных операций в реальном масштабе времени не вызывает каких-либо трудностей.

Непрерывный канал связи содержит выходные каскады передатчика, передающую антенну, среду распространения с источником помех, приемную антенну и входные каскады приемника, обеспечивающие усиление и при необходимости гетеродинирование сигнала (смещение спектра сигнала по оси частот, как правило, в более низкую область), с возможно малым собственным шумом.

Для передачи радиосигналов по НКС требуются определенные затраты энергии и полосы частот. Чем меньше их значения, тем, при прочих равных условиях, эффективнее система в целом.

Оценим эффективность ЦСПИ для простейшей модели канала, в котором сигнал не искажается, а на входе приемника на него накладывается помеха в виде белого гауссовского шума. Эффективность определяется степенью совершенства модема и кодека и характеризуется двумя показателями: коэффициентом использования полосы частот $\gamma = F/R$, где R — производительность источника, а F — полоса частот НКС, и коэффициентом использования энергии сигнала $\beta = E/N_0$, где E — энергия сигнала, затрачиваемая для передачи одного бита информации, а N_0 — спектральная плотность мощности шума. Чем меньше коэффициенты γ и β , тем эффективнее система. Связь между этими коэффициентами следует из фундаментальной теоремы Шеннона о пропускной способности НКС, которую, не претендуя на строгость выкладок, нетрудно вывести, зная информационные характеристики непрерывных сигналов [129].

Сигнал на выходе НКС в полосе частот F представляет аддитивную смесь полезного сигнала и шума и характеризуется энтропией или средним количеством информации в секунду. Чтобы достичь максимального значения пропускной способности канала C , естественно потребовать максимальность энтропии полезного сигнала, что будет, если он имеет гауссовский закон распределения и постоянную спектральную плотность мощности в полосе $(0 \dots F)$. Поскольку в соответствии с теоремой Котельникова такой сигнал полностью определяется $2F$ независимыми отсчетами в секунду, дифференциальная энтропия смеси сигнала и шума будет равна $\log \sqrt{2\pi e(P_c + P_{\text{ш}})}$, где P_c и $P_{\text{ш}}$ — мощности сигнала и шума соответственно.

Учитывая, что часть информации при действии шума теряется, среднее количество информации в секунду (пропускная способность канала) на выходе НКС равна

$$C = 2F \left(\log_2 \sqrt{2\pi e(P_c + P_{\text{ш}})} - \log_2 \sqrt{2\pi e P_{\text{ш}}} \right) = F \log_2 (1 + P_c / P_{\text{ш}}). \quad (9.1)$$

Пропускная способность характеризует предельные возможности, которые можно достичь, оптимизируя модем и кодек. Из теоремы Шеннона нетрудно установить связь между показателями γ и β . Если достигается предельная скорость передачи по каналу, то $R = C$. Используя формулу (9.1), после несложных преобразований получаем

$$\beta = \gamma(2^{1/\gamma} - 1).$$

Зависимость, представленная на рис. 9.2, называется *границей Шеннона* для идеального гауссовского канала и характеризует геометрическое место точек, находящихся выше кривой, координаты которых соответствуют параметрам реальных систем. Чем ближе точка к кривой, тем совершеннее

модем и кодек системы. Улучшение одного из параметров приводит к неизбежному ухудшению другого. Поэтому условно ЦСПИ можно разделить на два класса: эффективно использующие полосу частот (область II) и эффективно использующие энергетику канала (область I). Можно указать следующие пути решения задач,

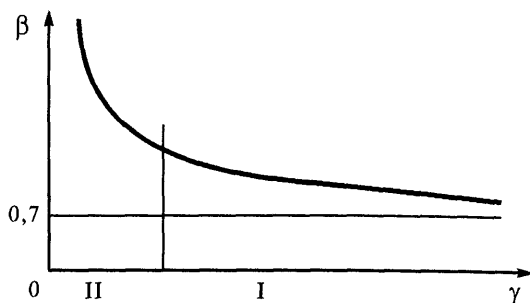


Рис. 9.2. Граница Шеннона

связанных с повышением γ - и β -эффективностей. Поскольку минимальная ширина спектра сигнала, передаваемого по каналу, определяется, в основном, длительностью посылок, то уменьшить коэффициент γ можно, объединив группу символов источника в один символ из алфавита с большим основанием, т. е. используя многопозиционное кодирование. Именно так эта задача решается в компьютерных модемах, использующих стандартные аналоговые телефонные каналы. Чтобы уменьшить коэффициент β , к группе информационных посылок для повышения помехоустойчивости добавляются дополнительные символы, что при фиксированной скорости требует уменьшения их длительности и, как следствие, расширения спектра сигнала и увеличения γ . На практике эти процедуры могут носить более сложный характер, но принципы решения указанных задач остаются такими же.

Анализируя рис. 9.2, можно сделать один очень важный вывод: улучшая один из показателей, ухудшаем другой. Например, для достижения скорости передачи 30 кбит/с в телефонном канале с полосой 3,1 кГц необходимо иметь отношение мощности сигнала к мощности шума, примерно равное 100, т. е. очень качественный канал связи, а рекламируемые модемы со скоростью передачи 56 кбит/с можно использовать только в идеальных каналах без помех.

9.2. Каналы связи

Каналы связи можно классифицировать по различным показателям. В теории передачи сигналов каналы классифицируют по характеру сигналов на входе и выходе. Различают *непрерывные*, *дискретные* и *дискретно-непрерывные* каналы. В непрерывных каналах сигналы на входе и выходе непрерывны по уровням, в дискретных — дискретны, а в дискретно-непрерывных — на входе дискретны (непрерывны), а на выходе непрерывны (дискретны).

Возможна также классификация каналов по назначению радиотехнических систем передачи информации (РСПИ) (телеграфные, телефонные, телевизионные, телеметрические и др.), по виду физической среды распространения (проводные, кабельные, волноводные и др.) и по диапазону используемых ими частот. К радиодиапазону относят частоты в пределах $30 \dots 3 \cdot 10^{12}$ Гц, что соответствует длинам волн от 10^8 м до 0,1 мм. Кроме радиодиапазона, в настоящее время широко используется оптический диапазон волн. В силу дискретного характера электромагнитного излучения в оптическом диапазоне волн такие каналы принято называть квантовыми. Данные о радиодиапазонах приведены в табл. 1.1. Далее более подробно, чем в гл. 1, рассматриваются особенности распространения и использования радиоволн различных диапазонов.

По способу распространения радиоволн различают каналы с открытым и закрытым распространением. В каналах с закрытым распространением электромагнитная энергия распространяется по направляющим линиям (кабельные, проводные, волноводные СВЧ тракты и др.). Для таких каналов характерны малый уровень помех и постоянство параметров сигнала, что позволяет передавать информацию с высокой скоростью и достоверностью.

В диапазонах ОНЧ и НЧ на небольших расстояниях поле в месте приема создается в результате дифракционного огибания волнами выпуклой поверхности Земли. На больших расстояниях радиоволны распространяются в своеобразном сферическом волноводе, внутренняя стенка которого образуется поверхностью Земли, а внешняя — ионосферой. Такой механизм распространения радиоволн позволяет принимать сигналы в любой точке Земли, причем параметры принятых сигналов отличаются достаточно высокой стабильностью. Особенностью этих диапазонов является также способность волн проникать в толщу Земли и воды на глубину в десятки метров. Принципиальным недостатком таких каналов являются ограниченная полоса частот (единицы герц) и очень большие линейные размеры антенных устройств, соизмеримых с длиной волны, составляющей километры. Сверхдлинные волны находят применение в навигации и передаче ограниченного объема информации на подводные объекты.

В распространении волн диапазона ВЧ участвует ионосфера. Однако если волны, длиннее 1 км, отражаются от нижнего ее слоя практически зеркально, то декаметровые волны достаточно глубоко проникают в ионосферу, что приводит к эффекту многолучевости, когда в точку приема приходит одновременно несколько сигналов с разными временами запаздывания. Многолучевость может носить дисперсионный или дискретный характер. Дисперсия (рассеяние) сигнала определяется отражением радиоволн от некоторого объема ионосферы, а дискретная многолучевость — отражением от разных слоев ионосферы. Поскольку глубина проникновения в ионосфе-

ру зависит от длины волны, то для передачи информации между двумя пунктами можно указать оптимальную рабочую частоту, на которой связь будет наиболее надежной (максимум мощности принимаемого сигнала, минимум эффекта многолучевости). Значения оптимальной рабочей частоты рассчитывают для определенных трасс и времени связи. Для этого составляют долговременные и кратковременные прогнозы по данным мировой сети станций ионосферного зондирования. Декаметровые волны широко применяются для глобальной связи и радиовещания. С их помощью можно передавать информацию сравнительно большого объема в пределах всего земного шара при ограниченной мощности передатчика и небольших по размеру антеннах. Полоса частот передаваемых сигналов в декаметровом канале не превышает 10 кГц.

До появления спутниковых систем связи этот диапазон был единственно пригодным для организации связи между двумя любыми пунктами на Земле без промежуточной ретрансляции. Однако эффект глобального распространения коротких волн имеет и свою отрицательную сторону: в точке приема могут появиться сильные помехи от дальних радиостанций.

Гектометровые волны днем распространяются как земные, а ночью — как ионосферные. Дальность распространения земной волны над сушей не превышает 500 км, а над морем — 1000 км. Диапазон СЧ широко используется в радиовещании, связи и радионавигации.

Волны диапазона частот выше 30 МГц слабо дифрагируют и поэтому распространяются в пределах прямой видимости. Расстояние прямой видимости (радиогоризонт) в километрах по поверхности Земли примерно равно $5\sqrt{A+B}$, где A и B — высоты передающей и приемной антенн в метрах. Если предположить, что $A = 30$ м, а $B = 3$ м, то радиогоризонт равен 30 км и для всех радиоволн, имеющих путь распространения менее 30 км, потери сигнала не будут связаны с кривизной поверхности Земли. Для всех остальных лучей возникает дополнительное затухание, обусловленное экранированием сигнала земной поверхностью. Однако резкое уменьшение уровня сигнала на расстояниях, превышающих прямую видимость, обуславливает снижение уровней помеховых сигналов и возможность использования одинаковых частот для организации связи в зонах, удаленных на расстояния, превышающие радиогоризонт. Некоторого увеличения дальности можно достичь, применив поднятые антенны, а для организации связи на расстояния, превышающие прямую видимость, — ретрансляцию сигналов. Системы с ретрансляцией сигналов называются радиорелейными линиями. Одним из основных достоинств высокочастотных диапазонов является большой частотный ресурс, что позволяет создавать радиосистемы передачи информации с высокой скоростью передачи и радиосети с большим числом одновременно работающих радиостанций. В

последние 30 лет диапазон ОВЧ и СВЧ нашел широкое применение для создания сетей мобильной связи.

Стремление увеличить дальность радиолинии в этих диапазонах без промежуточной ретрансляции нашло свое решение в РСПИ, использующих рассеяние радиоволн на неоднородностях тропосферы, ионосферы и метеорных следах. Однако такие системы по качеству передачи информации не могут конкурировать с радиорелейными линиями того же диапазона, поэтому их имеет смысл применять тогда, когда ретрансляция сигналов по тем или иным причинам невозможна.

Стремление увеличить ширину полосы частот канала, а также повысить пространственную селекцию сигналов использованием остронаправленных антенн при их ограниченных размерах привело к освоению диапазона миллиметровых волн. Основной особенностью их является сильное поглощение в дожде и тумане, что ограничивает их применение в наземных системах большой дальности. Однако в космических и спутниковых системах они весьма перспективны.

Новую эру в освоении высокочастотной области радиодиапазона для средств связи открыл запуск искусственных спутников Земли (ИСЗ). Обычно ИСЗ находятся на высоте 500...40 000 км от поверхности Земли и поэтому обеспечивают радиосвязь между земными станциями, удаленными на расстояния до 10...17 тыс. км. Линия спутниковой связи состоит из двух оконечных земных станций и одного или нескольких спутников-ретрансляторов, обращающихся вокруг Земли по заданным орбитам.

Из всего многообразия орбит ИСЗ особый интерес представляет экваториальная круговая орбита, удаленная от поверхности Земли на расстояние около 36 тыс. км (стационарная экваториальная орбита). Когда направление движения ИСЗ по такой орбите совпадает с направлением вращения Земли, спутник будет казаться наземному наблюдателю неподвижным (стационарный спутник). При использовании трех стационарных спутников, расположенных в экваториальной плоскости через 120° по дуге, оказывается принципиально возможным организовать глобальную систему связи. Максимальный, от горизонта до горизонта, обзор земной поверхности одним ИСЗ, или, иначе говоря, максимальное расстояние вдоль поверхности Земли между двумя станциями, будет практически составлять 15...17 тыс. км. Существенные преимущества стационарной орбиты заключаются в снижении требований к системам слежения за спутником, сведении к минимуму доплеровских сдвигов частоты сигналов, что упрощает приемное устройство при большом обзоре поверхности Земли. Недостатком стационарной орбиты является плохой охват приполярных зон. Поэтому в России для систем связи широко применяются сильно вытянутые эллиптические орбиты с большой полуосью до пяти земных радиусов с эксцентриситетом 0,8—0,9 и углом

наклона примерно 65° . Три спутника, выведенные через равномерные интервалы времени на аналогичные эллиптические орбиты, восходящие узлы которых смещены относительно друг друга на 120° , могут обеспечить круглосуточную непрерывную связь между земными станциями, расположенными в Северном полушарии Земли, на глобальные расстояния. Другая возможность в создании глобальных РСПИ заключается в использовании «созвездий» спутников, орбиты которых выбраны так, что над любой точкой на поверхности Земли находится по крайней мере один спутник.

Выбор рабочих частот для линии радиосвязи через ИСЗ определяется следующими факторами: условиями распространения и поглощения радиоволн, уровнем внешних помех, принимаемых антенной, техническими средствами (коэффициент шума приемного устройства, ширина лепестка диаграммы направленности антенны, точность ориентации и т. п.), взаимными помехами между системами связи через ИСЗ и другими службами, работающими в смежных или совмещенных диапазонах частот. Ограничение диапазона частот снизу определяется экранирующим действием ионосферы, а сверху — поглощением в тропосфере. Эти два фактора предопределили диапазон рабочих частот 40 МГц...40 ГГц. В настоящее время наибольшее использование находит диапазон частот 1...12 ГГц.

9.3. Модели каналов связи

9.3.1. Физическая модель непрерывного канала связи

Физическая модель непрерывного канала связи включает в свой состав технические средства, расположенные между выходом модулятора и входом демодулятора (рис. 9.3).

Проходя по непрерывному каналу связи, сигнал претерпевает ряд изменений. Эти изменения сводятся к ослаблению, искажению сигнала и наложению на него помех. В отдельных случаях искажению подвергается смесь сигнала и помех, например во входных цепях приемника или при ретрансляции в радиорелейных линиях. Для анализа системы важно знать характер искажений и уметь их моделировать. Реальные искажения имеют достаточно сложный характер. Однако для решения большинства задач непрерывный канал можно смоделировать в виде последовательно включенных линейных инерционных и нелинейных безынерционных четырехполюсников, обуславливающих соответственно линейные и нелинейные искажения сигнала.

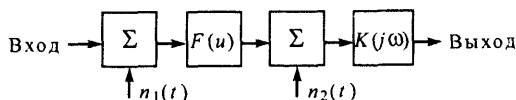


Рис. 9.3. Модель непрерывного канала связи

лов (см. рис. 9.3). Помехи могут накладываться на сигнал в любой точке цепи. Несмотря на кажущуюся простоту такой модели канала, нахождение отклика на ее выходе в тех случаях, когда помеха действует на входе нелинейного звена, является сложной математической задачей. Поэтому часто при решении подобных задач обращаются к различным методам моделирования.

Линейные искажения проявляются в изменении спектра (корреляционной функции) сигналов и помех. В зависимости от того, каковы эти искажения: регулярны или случайны, различают соответственно каналы с детерминированными или случайными линейными искажениями. Детерминированные линейные искажения в реальных каналах связаны с наличием частотно-избирательных цепей (фильтров во входных каскадах приемника и в выходных каскадах передатчика, коаксиальных и волноводных трактов, антенн и т. д.).

Случайные линейные искажения определяются средой распространения и связаны с прохождением сигнала от передающей антенны к приемной антенне разными путями (лучами). Этот эффект называется рассеянием сигнала. Различают два вида рассеяния сигнала: дискретное, когда запаздывание между сигналами в соседних лучах принимает конкретное значение (многолучевый канал), и дисперсное, когда запаздывание между соседними лучами бесконечно мало, а число лучей бесконечно велико.

Характер рассеяния сигнала определяется диапазоном используемых частот и типом системы. Если раньше типичными каналами с рассеянием сигналов являлись тропосферный и ионосферный, в которых связь за пределами прямой видимости достигалась рассеянием радиоволн на неоднородностях тропосферы и ионосферы и для которых были найдены способы борьбы с многолучевостью, то в последние десятилетия с развитием мобильных систем связи, действующих в условиях городской застройки, борьба с многолучевостью приобрела еще большую актуальность. Искажения сигналов, особенно в мобильных системах, имеют достаточно сложный характер. Однако даже упрощенные модели каналов позволяют разобраться в характере искажений сигналов и находить способы повышения качества передачи информации по каналам с рассеянием.

Пусть в точку приема приходят сигналы, переотраженные от совокупности бесконечно малых по размеру отражателей, размещенных в некотором пространстве, которые случайным образом перемещаются, сохраняя постоянным, в среднем, объем занимаемого пространства. Тогда на входе приемника будем иметь сумму сигналов с разной амплитудой и временем прихода, которые, в свою очередь, случайно изменяются с некоторой скоростью. Максимальную разность времен прихода сигналов называют *временем рассеяния сигнала*, *расширением задержки* или *памятью канала*. Естественно, что эта величина также носит случайный характер, но можно указать

ее среднее значение. Рассмотрим характер искажения гармонического сигнала при прохождении по такому каналу. Сигнал на входе приемника является суммой синусоид со случайными амплитудами и фазами.

Если число переотраженных сигналов велико, то по центральной предельной теореме теории вероятностей суммарные ортогональные сигналы будут иметь нормальные законы распределения амплитуд, а результирующий сигнал будет иметь случайную огибающую и фазу, изменяющиеся соответственно по рэлеевскому и равномерному законам.

Скорость изменения огибающей (замираний сигнала) определяется доплеровским сдвигом по частоте при движении отражателей. Обычно в мобильных системах связи переотраженные сигналы действуют на фоне достаточно мощного прямого сигнала. Тогда результирующий сигнал будет иметь райсовский закон распределения огибающей.

Если число лучей ограничено, например только два, то результирующий сигнал будет представлять биения колебаний последних. Вероятностные характеристики огибающей и фазы результирующего сигнала будут определяться соответствующими характеристиками амплитуд и фаз суммируемых сигналов.

Для сигналов с фиксированной шириной спектра F_c канал с рассеянием можно представить как фильтр со случайно изменяющимися во времени параметрами.

В общем случае сигнал на выходе линейного канала с изменяющимися параметрами можно найти, используя интеграл Дюамеля

$$u(t) = \int_0^t h(t, \tau) s(t - \tau) d\tau,$$

где $h(t, \tau)$ — импульсная характеристика канала. Таким образом, для оценки линейных искажений необходимо знать функцию $h(t, \tau)$ или связанную с ней преобразованием Фурье комплексную частотную характеристику $K(j\omega, t)$.

Решение задач анализа и синтеза устройств обработки сигналов существенно упрощается, если перейти к дискретной модели канала и сигналов. Дискретное представление математических моделей каналов основывается на конечном времени t_n рассеяния сигнала (памяти канала), определяемом протяженностью импульсной характеристики $h(t, \tau)$, и конечной ширине спектра передаваемого сигнала F_c . Формальным способом введения дискретной модели может быть разложение функции в ряды Котельникова, Фурье и т. п. Если полоса частот сигнала, передаваемого по каналу, ограничена интервалом $F_c = F_v - F_n$, где F_v и F_n — верхняя и нижняя частоты спектра сигнала, то достаточно рассматривать функцию $K(j2\pi f, t)$ переменной f только в интервале F_c . При этом импульсную характеристику $h(t, \tau)$ можно представить в виде ряда Котельникова для сигнала с полосовым спектром:

$$h(t, \tau) = \sum_{i=-\infty}^{\infty} h_i(t, \tau) = \sum_{i=-\infty}^{\infty} H(i, t) \frac{\sin[\pi F_c (\tau - i/F_c)]}{\pi F_c (\tau - i/F_c)} \cos \left[\pi (F_b + F_n) \left(\tau - \frac{i}{F_c} \right) + \varphi(i, t) \right], \quad (9.2)$$

где $H(i, t)$ — значения огибающей импульсной характеристики при $\tau = i/F_c$; $\varphi(i, t)$ — значения фазы.

Физическая модель канала (рис. 9.4), построенная в соответствии с формулой (9.2), содержит линию задержки с L отводами, усилители, комплексный коэффициент которых $h(i, t) = H(i, t) \exp(j\varphi(i, t))$ может изменяться, и сумматор. Временная задержка между соседними отводами равна $1/F_c$.

В частотной области модель канала можно построить, предположив конечность времени τ_n рассеяния сигнала. Тогда функция $K(j\omega, t)$ переменной $f = \omega/(2\pi)$ может быть задана комплексными значениями: $K(i, t) = K(j2\pi i \Delta f, t)$, где $\Delta f = 1/\tau_n$. Дискретная модель канала содержит набор полосовых фильтров с примыкающими частотными характеристиками, полоса пропускания каждого из которых равна $1/\tau_n$, и усилителей с управляемыми комплексными коэффициентами передачи $K(i, t)$ (рис. 9.5). Величину $1/\tau_n = F_k$ иногда называют *полосой когерентности*. Гармонические сигналы с разном по частоте, превышающим $1/\tau_n$, будут иметь некоррелированные случайные огибающую и фазу. Этот параметр определяет и характер замираний сигнала. Если ширина спектра передаваемого по каналу сигнала меньше F_k , то все спектральные составляющие сигнала изменяются одновременно и такие замирания называются *общими*. В том случае, когда $F \gg F_k$, отдельные участки спектра сигнала изменяются независимо и замирания называются *селективными*.

Необходимо знать характер изменения комплексных коэффициентов передачи

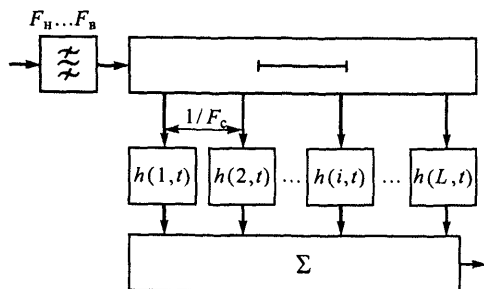


Рис. 9.4. Модель канала с рассеянием для сигналов с ограниченной ширины спектра

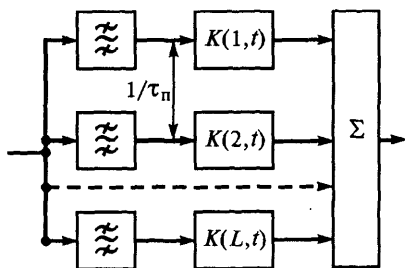


Рис. 9.5. Модель канала с ограниченным временем рассеяния сигнала

$$h(i, t) = H_d(i, t) + jH_m(i, t) = H(i, t) \exp(j\varphi(i, t)), \\ K(i, t) = K_d(i, t) + jK_m(i, t) = K(i, t) \exp(j\psi(i, t))$$

в каждой ветви. Если рассеивающий объем состоит из большого числа независимых отражателей, то по центральной предельной теореме теории вероятностей коэффициенты при действительной и мнимой частях будут гауссовскими независимыми случайными величинами с нулевыми математическими ожиданиями и дисперсиями, равными σ_i^2 . Тогда модули $H(i, t)$, $K(i, t)$ и фазы $\varphi(i, t)$, $\psi(i, t)$ будут подчиняться соответственно закону Рэлея и равномерному закону. В тех случаях, когда кроме рассеянной составляющей канал имеет и регулярную, модули коэффициентов передачи будут подчиняться закону Райса.

Информацию о динамике изменения коэффициента передачи дает корреляционная функция или спектральная плотность мощности флуктуации этого коэффициента.

Время корреляции τ_k или ширина спектра флуктуации $F_{\text{фл}}$ характеризует скорость изменения параметров канала. Например, для коротковолнового канала ширина спектра флуктуации составляет 0,1...1 Гц. В мобильных системах, где диапазон используемых частот много больше даже в предположении равенства скорости перемещения отражателей, спектр флуктуаций оказывается значительно шире.

В системах передачи дискретной информации рассеяние во времени сигнала приводит также к эффекту *межсимвольной интерференции*, заключающемуся в наложении следующих друг за другом посылок. Это имеет место, если длительность передаваемых посылок оказывается соизмеримой с временем рассеяния сигнала. Чтобы избежать этого вида искажения, приходится снижать скорость передачи в канале.

Нелинейные искажения возникают в результате прохождения сигнала по звеньям с нелинейной амплитудной характеристикой $F(u)$. Поскольку среда распространения, как правило, линейна, то нелинейные искажения обусловлены техническими устройствами, входящими в канал связи. Часто они возникают в ретрансляторах радиорелейных линий, в которых для получения максимальной мощности излучения передатчики умышленно переводят в режим работы с ограничением сигнала. Это имеет место, например, в спутниковых ретрансляторах.

Для узкополосных радиосигналов

$$s(t) = A(t)\cos(\omega_0 t + \varphi(t)) = A(t)\cos(\theta(t))$$

сигнал на выходе нелинейного звена является периодической функцией θ и может быть представлен в виде ряда Фурье от аргумента θ :

$$s_{\text{вых}}(t) = g_0(A) + g_1(A) \cos(\theta(t)) + g_2(A) \cos(2\theta(t)) + \dots$$

Поскольку приемное устройство обычно содержит на входе полосовой фильтр, пропускающий только спектральные составляющие в области несущей частоты ω_0 , то сигнал в полосе пропускания такого фильтра будет определяться выражением

$$s_{\omega_0}(t) = g_1(A) \cos(\omega_0 t + \varphi(t)),$$

где

$$g_1(A) = \frac{1}{\pi} \int_0^{2\pi} F(A \cos \theta) \cos(\theta) d\theta$$

— преобразование Чебышева первого порядка характеристики $F(u)$, которое определяет огибающую выходного сигнала в основной полосе частот.

Таким образом, нелинейные искажения сигнала сводятся к появлению новых спектральных составляющих на частотах $n\omega_0$, $n = 0, 2, 3, \dots$, и изменению огибающей $A(t)$. Точки перехода через нуль сигнала с частотой ω_0 не изменяют своего положения на оси времени.

Картина искажения сигнала существенно усложняется, когда одновременно с полезным сигналом $s(t)$ действуют другие сигналы или помехи. В этом случае на сигнал воздействуют еще и комбинационные составляющие, обусловленные взаимодействием сигнала и помех на нелинейном элементе. Это приводит к потере мощности полезного сигнала и к дополнительным помехам. Подавление полезного сигнала на нелинейности, которое обычно оценивают уменьшением отношения сигнал—шум в децибелах, зависит от формы кривой $F(u)$ и вида помеховых сигналов.

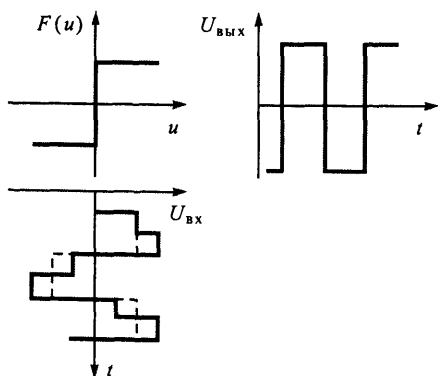


Рис 9.6. Диаграмма подавления слабого сигнала сильным на нелинейном элементе (штриховой линией обозначен сильный сигнал)

Особый интерес представляет так называемый предельный ограничитель, для которого $F(u) = \text{sign}(u)$. Пусть на его входе действуют два сигнала с разными амплитудами (рис. 9.6), один из которых полезный, а другой мешающий. На выходе ограничителя будем иметь либо только полезный сигнал, либо только мешающий, в зависимости от соотношения амплитуд. Таким образом, сильный сигнал полностью подавляет слабый сигнал. При других формах сигнала и помехи степень подавления имеет конечное значение. Например, если входной полезный сигнал явля-

ется узкополосным радиосигналом, то при любом виде модуляции степень подавления его сильным синусоидальным мешающим сигналом составляет около 6 дБ. Для помехи, являющейся суммой гармонического сигнала и гауссовской помехи, коэффициент подавления полезного сигнала $K_{\text{под}}$ можно рассчитать по формуле [13]

$$K_{\text{под}} = \frac{\pi}{4}(1 + \alpha) \left[\exp\left(-\frac{\alpha}{2}\right) I_0\left(\frac{\alpha}{2}\right) \right]^2,$$

где α — отношение мощности синусоидальной составляющей помехи к флуктуационной; I_0 — модифицированная функция Бесселя первого рода нулевого порядка от аргумента $\alpha/2$.

Из рис. 9.7 видно, что в предельном ограничителе подавление гармонического сигнала будет наибольшим при воздействии гармонической помехи и наименьшим при воздействии гауссовской помехи.

Ошибки, возникающие при приеме сообщений, в значительной степени определяются видом и интенсивностью помех, действующих в канале. В зависимости от места нахождения источника помех различают *внутренние* и *внешние* помехи. Внутренние помехи возникают в самой системе. К ним относятся шумы входных каскадов приемника, приемной антенны, линий канализации сигнала и электрические сигналы, попадающие в приемник по внутренним цепям вследствие плохого экранирования или развязки между каскадами. Последний вид помех связан с ошибками в конструкции и по возможности должен быть устранен. Внутренний шум, обусловленный хаотическим движением носителей зарядов, принципиально неустраним, хотя может быть в значительной степени ослаблен применением качественных узлов и деталей, а также снижением рабочей температуры.

Различают *тепловой* и *дробовый* шумы. Тепловой шум обусловлен тепловым движением носителей заряда, приводящим к появлению случайной разности потенциалов. Он представляет собой гауссовский случайный процесс с нулевым средним и спектральной плотностью мощности

$$N_0(f) = \frac{hf}{\exp(hf/kT^\circ) - 1},$$

где $h = 6,6 \cdot 10^{-34}$ Дж · с — постоянная Планка; $k = 1,38 \cdot 10^{-23}$ Дж/град — постоянная Больцмана; T° — абсолютная температура источника шума; f — частота.

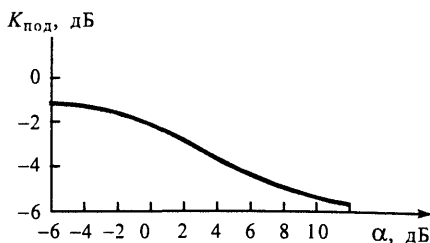


Рис 9.7. Зависимость коэффициента подавления узкополосного радиосигнала суммой синусоидальной и гауссовской помех от отношения мощностей этих помех

В диапазоне частот, в котором работают радиосистемы, выполняется условие $hf \ll kT^\circ$, и поэтому $N_0(f) = kT^\circ = N_0$, Вт/Гц. Таким образом, тепловой шум можно рассматривать как белый с односторонней спектральной плотностью $N_0 = kT^\circ$.

В реальных системах полоса частот пропускания ограничена и мощность шума определяется равенством $P_{\text{ш}} = N_0 F$.

Шумы электровакуумных и полупроводниковых приборов (дробовые шумы) обусловлены дискретной природой носителей заряда. Статистические характеристики дробового шума такие же, как у теплового.

Внешние помехи возникают из-за различных электромагнитных процессов, происходящих в атмосфере, ионосфере, космическом пространстве, а также излучения земной поверхности (*естественные помехи*). Кроме того, они создаются различными радиостанциями (*станционные помехи*), промышленными установками, медицинской аппаратурой, электрическими двигателями и т. п. В зависимости от диапазона частот и условий, в которых работает система передачи информации (СПИ), преобладает тот или иной вид помех.

Атмосферные помехи возникают в результате различных электрических процессов, происходящих в земной атмосфере. Наиболее мощным источником являются электрические грозовые разряды, которые приводят к излучению электромагнитной энергии практически во всем радиочастотном диапазоне. Максимум излучения разряда приходится на полосу частот 5...30 кГц. Интенсивность поля помех, создаваемых электрическими разрядами, в пределах прямой видимости уменьшается обратно пропорционально частоте.

Для диапазона частот выше 30 МГц заметными становятся помехи, связанные с источниками, находящимися в пределах нашей Галактики и вне ее (*космические шумы*). Причиной возникновения этих помех является тепловое излучение межзвездных газов, Солнца, радиозвезд. Большинство известных радиозвезд находятся в пределах нашей Галактики, и их излучение во много раз превышает по интенсивности излучение тепловых источников. Интенсивность космических шумов так же, как и внутренних, оценивается шумовой температурой.

Земная поверхность, как и всякое нагретое тело, излучает электромагнитные волны. Они могут попадать в антенну по основному или боковым лепесткам диаграммы направленности. Мощность этих шумов в значительной степени определяется положением и формой диаграммы направленности, а также температурой и электрическими характеристиками земной поверхности. По своим статистическим характеристикам они аналогичны тепловому шуму.

Промышленные помехи создаются различным электрооборудованием промышленных предприятий, транспорта, линиями электропередач и дру-

гими электроустановками. В большинстве случаев они представляют собой последовательности импульсов с постоянным или переменным периодом следования. Распространение промышленных помех происходит в основном земной волной, однако часто они передаются по линиям связи, электропередач, железнодорожным линиям и т. п. Уровень промышленных помех зависит от места расположения приемника относительно промышленных объектов.

Одним из распространенных видов внешних помех являются помехи от посторонних радиостанций. Насыщенность радиосредствами (радиосвязь, радиолокация, радионавигация и т. п.) и, следовательно, загрузка радиодиапазонов таковы, что весьма часто помехи от посторонних радиосредств превышают прочие виды помех. Станционные помехи обусловлены целым рядом причин: нарушением регламента распределения рабочих частот, недостаточной стабильностью генераторов и плохой фильтрацией гармоник сигнала, нелинейными искажениями в канале, ведущими к перекрестным помехам. Снизить уровень станционных помех можно с помощью организационно-технических мероприятий. Это направление в радиоэлектронике последнее время усиленно развивается и называется *электромагнитная совместимость радиоэлектронных средств*.

Станционные помехи присутствуют практически во всех диапазонах, и особенно в коротковолновом, где из-за ионосферного распространения радиоволн часто складываются благоприятные условия для прохождения радиоволн от посторонних, очень далеко расположенных передатчиков, работающих на той же частоте. Появление станционных помех в полосе принимаемого сигнала, их уровень и амплитуда являются, как правило, случайными процессами. Если число помех, попадающих в полосу сигнала, велико, то по центральной предельной теореме теории вероятностей мгновенные значения результирующего сигнала будут подчиняться гауссовскому закону. В то же время изменение загрузки канала во времени и по частоте приводит к тому, что станционная помеха оказывается нестационарным случайным процессом. Упрощенную физическую модель образования станционных помех при высокой загрузке канала можно представить в виде последовательно включенных генератора белого шума и фильтра с частотной характеристикой, изменяющейся случайным образом во времени.

Спектральную плотность мощности помех $N(f, t)$ (рис. 9.8) как случайный процесс можно достаточно полно охарактеризовать плотностью вероятности $w_{i,f}(N)$ и корреляцион-

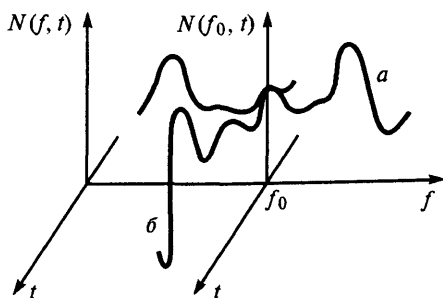


Рис. 9.8. Изменение спектральной плотности помех по частоте и во времени

ными функциями флуктуаций во временной и частотной областях $R_N(\tau)$ и $R_N(\nu)$. На практике распространена модель с логонормальным распределением помех в частотной и временной областях. Параметрами корреляционных функций являются интервал корреляции во времени τ_k и интервал корреляции по частоте F_n .

Если число станционных помех, попадающих в полосу сигнала, ограничено, то рассмотренная модель не всегда применима. В этом случае сигнал на входе приемника приходится представлять в виде суммы полезного сигнала и ограниченного числа аддитивных помех с известными или неизвестными статистическими характеристиками:

$$u(t) = s(t) + n(t) + \sum_{i=1}^k \gamma_i(t),$$

где $\gamma_i(t) = A_i(t) \cos(\omega_{cp}t + \theta_i(t))$.

Огибающая $A_i(t)$ и фаза $\theta_i(t)$ помехи могут быть как случайными, так и детерминированными процессами.

9.3.2. Математическая модель непрерывного канала связи

Математическая модель непрерывного канала связи устанавливает соответствие между сигналами на выходе и входе канала. Она должна по возможности точно описывать основные особенности реального канала и в то же время быть достаточно простой для получения конечных результатов при анализе и синтезе систем передачи. Рассмотрим наиболее простые и часто встречающиеся модели непрерывных каналов связи [130].

Идеальный канал без помех вносит детерминированные искажения, связанные с изменением амплитуды и временного положения сигнала. Переданный сигнал может быть полностью восстановлен на приемной стороне в новом временном отсчете. Эта модель используется для описания каналов с закрытым распространением малой протяженности (кабель, провод, волновод, световод и т. д.).

Канал с гауссовским белым шумом представляет собой идеальный канал, в котором на сигнал накладывается помеха

$$u(t) = \mu s(t - \tau) + n(t).$$

Коэффициент передачи μ и время запаздывания τ постоянны и известны в точке приема. Такая модель, например, соответствует радиоканалам, работающим в пределах прямой видимости.

Гауссовский канал с неопределенной фазой сигнала отличается от предыдущего тем, что фаза несущего колебания в точке приема предполагается

случайной с плотностью распределения $w(\varphi)$ в интервале $-\pi \leq \varphi \leq \pi$. Эта неопределенность вызвана двумя причинами: отсутствием устройств оценки и предсказания фазы или ошибками в оценке фазы при их работе. Важно знать скорость флуктуации фазы. В дискретных системах различают каналы с быстрыми флуктуациями, когда интервал их корреляции меньше длительности посылки, и с медленными, когда это условие не выполняется. При медленных флуктуациях фаза несущего колебания за время посылки практически не изменяется.

Гауссовский канал с неопределенной амплитудой и фазой сигнала вносит в сигнал наряду с флуктуациями фазы и флуктуации амплитуды, которые связаны с изменением во времени по случайному закону коэффициента передачи μ . Как и в предыдущем случае, флуктуации могут быть быстрыми и медленными. Для определения модели канала необходимо задать плотность распределения $w(\mu)$ и корреляционную функцию флуктуаций $R_\mu(\tau)$.

Гауссовский канал с линейными искажениями приводит к изменению формы сигнала из-за наличия избирательных цепей. В общем случае линейные искажения носят случайный характер. Частотная характеристика канала $K(j\omega, t)$ неравномерна в полосе частот сигнала F_c и изменяется во времени, а импульсная характеристика $h(t, \tau)$ имеет длительность τ_n (время памяти канала), превышающую величину $1/F_c$. Такая модель полезна при анализе систем, использующих, например, каналы с рассеянием сигнала. Сигнал на выходе канала с линейными искажениями описывается формулой

$$u(t) = \int_0^t h(t, \tau) s(t - \tau) d\tau + n(t).$$

В радиосистемах передачи дискретной информации, когда время памяти канала τ_n соизмеримо с длительностью посылки T_c (а тем более превышает ее), имеет место межсимвольная интерференция, которая проявляется в наложении друг на друга соседних посылок. Одной из причин возникновения межсимвольной интерференции является увеличение скорости передачи при ограниченной полосе пропускания канала.

Гауссовский канал с нелинейными искажениями сигнала, как и в предыдущем случае, предполагает существование аддитивной помехи в виде гауссовского белого шума, однако смесь сигнала и помехи, проходя по каналу, претерпевает нелинейные искажения так, что сигнал на входе приемника имеет вид $u(t) = F[s(t) + n(t)]$, где $F[\cdot]$ — амплитудная характеристика нелинейного звена канала.

Возможно дальнейшее усложнение модели с нелинейными искажениями, если предположить наличие в канале еще и линейных искажений, вызванных частотно-избирательными звеньями системы.

Линейный канал со сложной аддитивной помехой характеризуется тем, что на сигнал могут действовать помехи любого вида: сосредоточенные по спектру, во времени, гауссовские, негауссовские и т. д. Модель помех можно определить, указав способ вычисления многомерной плотности распределения вероятностей. Эта модель наиболее полно отображает реальный шум в каналах связи, однако редко используется в силу сложности. Проще задать модель сложных аддитивных помех в виде небелого гауссовского шума с изменяющейся во времени и по частоте спектральной плотностью $N(f, t)$, характеризуемой как случайный процесс с плотностью распределения $w(N)$ и корреляционными функциями во временной $R_N(\tau)$ и частотной $R_N(\nu)$ областях.

9.3.3. Математические модели дискретных каналов связи

Дискретно-непрерывный канал имеет дискретный вход и непрерывный выход. Примером такого канала является канал, образованный совокупностью технических средств между выходом кодера канала и входом демодулятора. Для его описания необходимо знать алфавит входных символов α_r , $r = 1, 2, \dots, m$; вероятности появления символов алфавита $p(\alpha_r)$, $r = 1, 2, \dots, m$; полосу пропускания непрерывного канала F_k , входящего в рассматриваемый канал; плотности вероятности $w(u|\alpha_r)$ появления сигнала $u(t)$ на выходе канала при условии, что передавался символ α_r .

Зная вероятности $p(\alpha_r)$ и условные плотности распределения вероятностей $w(u|\alpha_r)$, можно найти апостериорные вероятности

$$p(\alpha_r | u) = \frac{p(\alpha_r)w(u|\alpha_r)}{\sum_{r=1}^m p(\alpha_r)w(u|\alpha_r)}, \quad r = 1, 2, \dots, m,$$

по которым, как правило, и принимается решение о переданном символе.

Ширина спектра сигнала $u(t)$ не может превышать значения F_k . Поэтому в соответствии с теоремой Котельникова (см. гл. 2) его можно представить совокупностью $M = 2F_k T_c$ отсчетов, где T_c — длительность сигнала. Соответственно, условные плотности вероятности $w(u|\alpha_r)$, $r = 1, 2, \dots, m$, можно задать как M -мерные плотности вероятности совокупности M отсчетов сигнала $u(t)$.

В тех случаях, когда сигнал $u(t)$ является аддитивной смесью полезного сигнала $s_r(t)$ с известными параметрами, несущего информацию о символе α_r , и шума $n(t)$, M -мерная плотность вероятности $w_M(u_1, u_2, \dots, u_M|\alpha_r)$ будет полностью определяться M -мерной плотностью вероятности шума, т. е.

$$w_M(u_1, u_2, \dots, u_M | \alpha_r) = w_M[(u_1 - s_1^r), (u_2 - s_2^r), \dots, (u_M - s_M^r)] = w_M[n_1, n_2, \dots, n_M],$$

где u_i , s_i^r и n_i — отсчеты соответственно сигналов $u(t)$, $s_r(t)$ и шума $n(t)$ в момент t_i . При независимых отсчетах шума

$$w_M(u_1, u_2, \dots, u_M | \alpha_r) = \prod_{i=1}^M w(n_i).$$

Если плотность вероятности $w(u|\alpha_r)$ для любого сочетания $u(t)$ и α_r не зависит от времени, то канал называется *стационарным*.

Если выполняется условие $w_M(u | X_k^r, X_{k-1}^r, \dots, X_{k-N}^r) = w(u | X_k^r)$,

где $X_{k-N}^r, \dots, X_{k-1}^r, X_k^r$ — последовательность передаваемых символов, то такой канал называется *каналом без памяти*.

Реальные каналы являются обычно нестационарными и обладают памятью. Тем не менее, модель дискретно-непрерывного стационарного канала с памятью часто применяется благодаря ее простоте.

Дискретный канал имеет дискретный вход и дискретный выход. Примером такого канала является канал, образованный совокупностью технических средств между выходом кодера канала и выходом демодулятора. Для описания дискретного канала необходимо знать алфавит входных символов α_r , $r = 1, 2, \dots, m$, их вероятности появления $p(\alpha_r)$, скорость R передачи символов, алфавит символов на выходе канала y_j , $j = 1, 2, \dots, n$, и значения *переходных вероятностей* $p(y_j|\alpha_r)$, $j = 1, 2, \dots, n$, $r = 1, 2, \dots, m$, появления символа y_j при условии передачи символа α_r .

Первые две характеристики определяются свойствами источника сообщений, скорость R — полосой пропускания непрерывного канала, входящего в состав дискретного канала, объем алфавита выходных символов — алгоритмом работы решающей схемы, переходные вероятности $p(y_j|\alpha_r)$ — характеристиками непрерывного канала.

Заметим, что в общем случае в дискретном канале объемы алфавитов входных и выходных символов не совпадают. Примером может быть канал со стиранием. Алфавит на его выходе содержит один добавочный символ по сравнению с алфавитом на входе. Этот добавочный символ (*символ стирания*) появляется на выходе канала тогда, когда анализируемый сигнал нельзя с большой вероятностью отождествить ни с одним из передаваемых символов. Стирание символов при применении соответствующего помехоустойчивого кода позволяет повысить помехоустойчивость.

Зная вероятности $p(\alpha_r)$ и $p(y_j|\alpha_r)$, $r = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, можно вычислить апостериорные вероятности

$$p(\alpha_r | y_j) = \frac{p(\alpha_r) p(y_j | \alpha_r)}{\sum_{r=1}^m p(\alpha_r) p(y_j | \alpha_r)}, \quad r = 1, 2, \dots, m, \quad j = 1, 2, \dots, n,$$

того, что при принятом символе y_j был передан символ α_r . Вероятности $p(\alpha_r)$ и $p(\alpha_r | y_j)$ позволяют определять полную вероятность ошибки в канале (или полную вероятность правильного приема) и информационные характеристики дискретного канала.

Дискретный канал называется *стационарным*, если переходные вероятности $p(y_j | \alpha_r)$, $j = 1, 2, \dots, n$, $r = 1, 2, \dots, m$, не зависят от времени. Дискретный канал называется *без памяти*, если переходные вероятности $p(y_j | \alpha_r)$, $j = 1, 2, \dots, n$, $r = 1, 2, \dots, m$, не зависят от того, какие символы передавались и принимались ранее.

Если в стационарном дискретном канале алфавиты на входе и выходе совпадают и

$$p(\alpha_r | y_j) = \begin{cases} P_{\text{ош}} & \text{для } j \neq r, \\ 1 - (m-1)P_{\text{ош}} & \text{для } j = r, \end{cases}$$

то такой канал называется *симметричным*.

Математическая модель канала должна обеспечивать возможность нахождения основных характеристик потока ошибок. К ним относятся: вероятность ошибки в приеме символа $P_{\text{ош}}$; распределение вероятностей $P_n(r)$ появления r ошибок в блоке длины n ; распределение длин интервалов между соседними ошибками; распределение длин серий ошибок и т. п.

Модель должна быть простой и удобной для проведения расчетов. В то же время она должна достаточно точно описывать реальный канал, т. е. находиться в хорошем соответствии с экспериментальными данными. Наиболее простой является модель стационарного симметричного канала без памяти. В таком канале ошибки возникают независимо друг от друга, т. е. между ошибками отсутствуют статистические связи. Вероятность ошибки $P_{\text{ош}}$ при передаче любого символа одинакова и не меняется во времени. Стационарный симметричный канал без памяти полностью описывается вероятностью $P_{\text{ош}}$. Распределение числа ошибок в нем подчиняется биномиальному закону:

$$P_n(r) = C_n^r P_{\text{ош}}^r (1 - P_{\text{ош}})^{n-r}, \quad (9.3)$$

где n — число символов в блоке; r — число ошибочных символов.

Зная вероятность ошибки $P_{\text{ош}}$ и используя выражение (9.3), можно найти все необходимые характеристики, в частности вероятность правиль-

ного приема блока из n символов $P_n(0) = (1 - P_{\text{ош}})^n$; вероятность приема блока, содержащего хотя бы одну ошибку, $P_n(r \geq 1) = 1 - P_n(0)$; вероятность появления в блоке l и более ошибок:

$$P_n(r \geq l) = \sum_{r=l}^n C_n^r P_{\text{ош}}^r (1 - P_{\text{ош}})^{n-r}.$$

Большинство реальных каналов имеют «память», которая проявляется в том, что вероятность ошибки в символе зависит от того, какие символы передавались до него и как они были приняты.

При рассеянии сигнала приходящая на вход приемника посылка является суммой некоторого числа предыдущих посылок с соответствующими весовыми коэффициентами. Поэтому вероятность ошибки в последующем символе будет зависеть от характера передаваемой информации за время рассеяния сигнала. Например, при чередовании посылок разных частот ошибка будет больше, чем в последовательности, состоящей из посылок одной частоты. Если меняется длительность отдельных мешающих воздействий, например в результате общих замираний сигнала или изменения уровня помех, то ошибки будут группироваться в пачки. Вероятность ошибки при приеме символа в этом случае зависит от того, была ошибка в предыдущем символе или не была. Простой моделью двоичного симметричного канала с памятью является канал, который может находиться в одном из двух состояний: $d = 0$ и $d = 1$. В обоих состояниях возможны независимые ошибки с вероятностями p_0 и p_1 , где нижние индексы указывают на состояние канала.

Одним из распространенных методов описания дискретного канала с памятью, связанной с межсимвольными искажениями, является использование аппарата цепей Маркова (посимвольное описание). В этом случае последовательность состояний двоичного канала рассматривается как N -связная двоичная цепь Маркова, а значения символов на каждой позиции — как состояние цепи, где N — число символов, на которое распространяется память канала.

9.4. Передача и прием дискретных сообщений

9.4.1. Модуляция и демодуляция. Модемы

Алгоритм преобразования информационных посылок источника в радиосигналы и обратно определяется типом канала связи и выбранным критерием эффективности системы. Задача поиска оптимального алгоритма преобразования основывается на использовании метода проверки статистических гипотез при различении сигналов, рассмотренного в гл. 3.

Основными характеристиками модема являются помехоустойчивость, которая количественно оценивается связью вероятности ошибки с отношением сигнал—шум, и эффективность использования полосы частот канала.

В гауссовском канале с полностью известными параметрами сигналов оптимальным устройством обработки является коррелятор или согласованный фильтр, а вероятность ошибки при различении сигналов зависит от отношения сигнал—шум, коэффициента корреляции между сигналами и не зависит от их формы. Поэтому для повышения эффективности использования полосы частот канала целесообразно использовать простые сигналы.

Если система передачи двоичная, то максимальная помехоустойчивость достигается, если коэффициент корреляции между сигналами равен -1 , т. е. для противоположных сигналов. Зависимость вероятности ошибки от отношения сигнал—шум имеет вид

$$P_{\text{ош}} = 1 - \Phi\left(\sqrt{(1 - r_{ij})h^2}\right),$$

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{z^2}{2}\right) dz,$$

$$r_{ij} = \frac{1}{E} \int_0^{T_c} s_i(t) s_j(t) dt.$$

Для двоичных систем передачи информации фазовая модуляция (ФМ) позволяет наиболее эффективно использовать полосу частот канала, т. е. передавать сообщение в отведенной полосе частот с максимальной скоростью, однако это требует обязательного измерения в демодуляторе фазы несущей частоты посылок. Если фаза несущей по каким-то причинам не измеряется, то переходят к некогерентному приему (см. гл. 3).

Максимальной помехоустойчивости при некогерентном приеме можно достичь, если сигналы будут ортогональными в усиленном смысле:

$$\int_0^{T_c} s_i(t) s_j(t) dt = 0, \quad \int_0^{T_c} s_i(t) \hat{s}_j(t) dt = 0,$$

где $\hat{s}_j(t)$ — преобразование Гильберта от сигнала $s_j(t)$.

Вероятность ошибки при приеме двоичных сигналов с одинаковой энергией, ортогональных в усиленном смысле, вычисляется по формуле

$$P_{\text{ош}} = 0,5 \exp(-h^2/2).$$

Получить ортогональные сигналы можно с помощью амплитудной модуляции (АМ) или частотной модуляции (манипуляции) (ЧМ). Для амплитудной модуляции таких сигналов два, а для частотной число сигналов $m \gg 2$, при этом сдвиг по частоте между посылками должен быть кратен $1/T_c$, где T_c — длительность посылки. Спектр ЧМ сигнала можно представить как сумму спектров АМ сигналов с разными несущими. Естественно, эффективность использования полосы канала для ЧМ хуже, чем для ФМ. Некоторого сужения спектра ЧМ сигнала достигают использованием модуляции без разрыва фазы, когда частоты посылок формируют от одного генератора, сохраняя их когерентность. Тогда происходит сложение не амплитудных спектров, а комплексных с учетом их фазовой структуры, что и приводит к такому изменению спектра, при котором полоса сигнала уменьшается.

Эффективность использования полосы частот канала в РСПИ является важной характеристикой, поскольку именно она в существенной степени определяет стоимость передачи. Для повышения частотной эффективности применяют многопозиционные амплитудно-модулированные, фазомодулированные и амплитудно-фазомодулированные сигналы. При этом блоку двоичных символов источника ставится в соответствие радиосигнал из ансамбля сигналов, отличающихся друг от друга либо амплитудой, либо фазой, либо амплитудой и фазой одновременно. Во сколько раз увеличивается длительность радиосигнала, во столько и сужается его ширина спектра. Число сигналов в ансамбле должно быть равно 2^n , где n — число двоичных символов в блоке.

Амплитудно-фазоманипулированные (АФМ) сигналы имеют вид

$$s_i(t) = \sqrt{2/T_c} a_i \cos(\omega_0 t) - \sqrt{2/T_c} b_i \sin(\omega_0 t), \quad 0 \leq t \leq T_c, \quad i = 1, 2, \dots, m, \quad (9.4)$$

или

$$s_i(t) = \sqrt{2/T_c} U_i \cos(\omega_0 t + \psi_i), \quad i = 1, 2, \dots, m, \quad (9.5)$$

где $U_i = \sqrt{a_i^2 + b_i^2}$, $\psi_i = \arctg(b_i/a_i)$.

Используя функции $\sqrt{2/T_c} \cos(\omega_0 t)$ и $\sqrt{2/T_c} \sin(\omega_0 t)$ в качестве базисных, сигнал $s_i(t)$ в соответствии с соотношениями (9.4) и (9.5) можно рассматривать либо как двумерный вектор с координатами a_i и b_i в декартовой системе, либо как вектор с амплитудой U_i и фазой ψ_i в полярной системе координат.

В принципе, для каждого числа m можно построить бесконечно большое число ансамблей АФМ сигналов. Поэтому существенной является задача нахождения оптимальных ансамблей.

Пользуясь геометрической трактовкой, каждому сигналу $s_i(t)$ можно поставить в соответствие некоторую область пространства сигналов, кото-

рую обычно называют *собственной областью*, или *областью правильного приема*. Для дискретных сообщений вероятность правильного приема есть вероятность попадания конца вектора принятого сигнала в собственную область S_i передаваемого сигнала $s_i(t)$.

При равновероятной передаче сообщений оптимизация ансамбля заключается в таком размещении сигнальных точек, при котором собственные области S_i , $i = 1, 2, \dots, m$, примерно одинаковы и имеют максимальный объем (площадь). Задача в общем случае сводится к плотнейшей укладке сфер одинакового радиуса в m -мерном пространстве сигналов или, как в рассматриваемом случае, к плотнейшей укладке окружностей на плоскости. При этом центры сфер или окружностей соответствуют сигнальным точкам.

Большинство известных ансамблей АФМ сигналов найдены эвристическим методом. На рис. 9.9, *а* показаны ансамбли сигналов на основе так называемой треугольной сети для $m = 3, 7, 19$. Сигнальные точки лежат в вершинах правильных треугольников, а собственные области сигналов (за исключением периферийных) имеют вид правильных шестиугольников. На рис. 9.9, *б* представлены ансамбли сигналов на основе квадратной сети, а на рис. 9.9, *в* — различные варианты круговых расположений сигнальных точек. В последнем случае ансамбли обозначаются как (i_1, i_2, \dots, i_k) , где i_j — величина, равная числу сигнальных точек на j -й окружности. Радиусы окружностей r_j , $j = 1, 2, \dots, k$, или отношения радиусов должны быть заданы.

Класс АФМ сигналов включает в себя сигналы с m -ичной фазовой манипуляцией, которые имеют вид

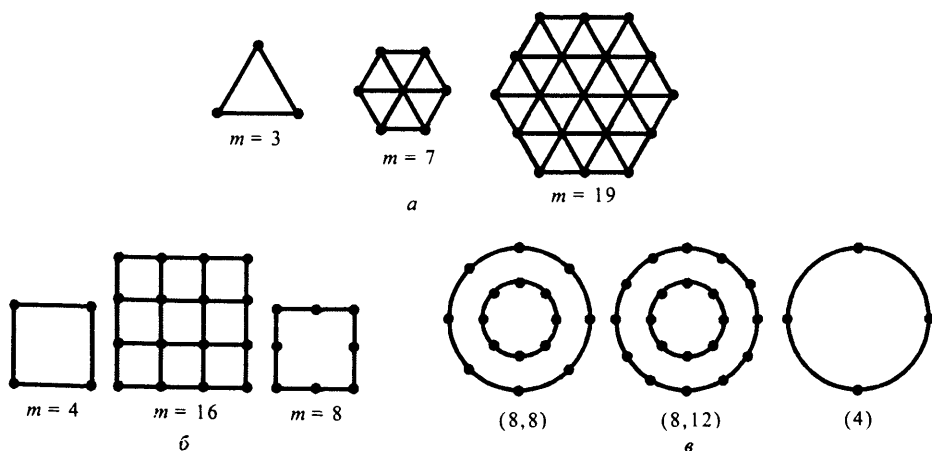


Рис. 9.9. Ансамбли АФМ сигналов на основе:

а — треугольной сети; *б* — квадратной сети; *в* — круговой сети

$$s_i(t) = A_0 \cos(\omega_0 t + 2\pi i / m),$$

$$i = 1, 2, \dots, m.$$

Они образуют круговую сеть с равномерным распределением точек по окружности.

Методы формирования АФМ сигналов зависят от вида ансамбля, требований к точности и быстродействию модуляторов. При задании сигналов в декартовой системе координат (рис. 9.10) передаваемые двоичные символы поступают на цифроаналоговые преобразователи (ЦАП) блоками длиной $k = \log_2 m$. Вырабатываемые ЦАП сигналы a_i и b_i модулируют квадратурные составляющие несущего колебания в балансных модуляторах (БМ).

В схеме формирования АФМ сигналов, заданных в полярной системе координат (рис. 9.11), ЦАП вырабатывает из двоичных символов сигналы, используемые для модуляции несущей последовательно в фазовом (ФМ) и амплитудном (АМ) модуляторах.

Существуют и другие методы формирования АФМ сигналов. Вычисление средней вероятности ошибки

$$P_{\text{ош}} = \sum_{i=1}^m P_{\text{ош}}(s_i) P_{\text{ош}}(s_i)$$

при использовании АФМ сигналов в общем случае является весьма громоздким, что обусловлено необходимостью перебора всех собственных областей S_i , $i = 1, 2, \dots, m$, которые, как правило, имеют различную форму и сложную конфигурацию.

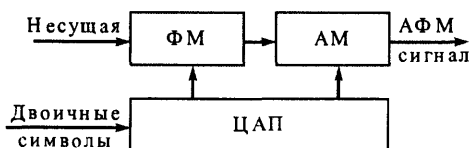


Рис. 9.11. Структурная схема формирователя АФМ сигналов, заданных в полярной системе координат

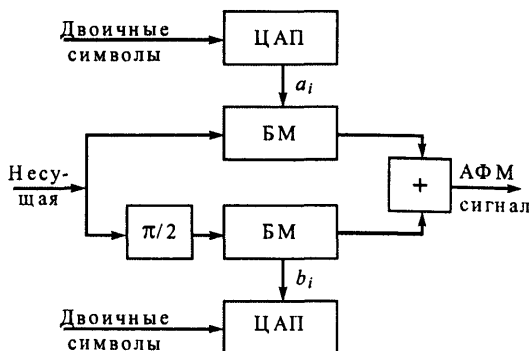


Рис. 9.10. Структурная схема формирователя АФМ сигналов, заданных в декартовой системе координат

Решение задачи упрощается при больших отношениях сигнал—шум. При этом можно воспользоваться верхней границей для вероятности ошибки

$$P_{\text{ош}}(s_l) \leq \sum_{\substack{j=1 \\ j \neq l}}^m P_{\text{ош}}(s_j | s_l)$$

или

$$P_{\text{ош}} \leq (m-1) \max P_{\text{ош}}(s_j | s_l).$$

При работе системы в условиях действия гауссовского белого шума с односторонней спектральной плотностью N_0 вероятность ошибки, выраженная (см. § 3.4) через расстояние $d(s_i, s_j)$, находится по формуле

$$P_{\text{ош}}(s_j | s_l) = 1 - \Phi[d(s_i, s_j) / \sqrt{2N_0}].$$

Тогда

$$P_{\text{ош}}(s_l) \leq \sum_{\substack{j=1 \\ j \neq l}}^m \left\{ 1 - \Phi \left[\frac{d(s_l, s_j)}{\sqrt{2N_0}} \right] \right\}.$$

Используя асимптотическое представление интеграла вероятности, можно записать

$$P_{\text{ош}}(s_l) \approx \sum_{\substack{j=1 \\ j \neq l}}^m \frac{\exp(-d^2(s_l, s_j)/(4N_0))}{\sqrt{2\pi} d(s_l, s_j) / \sqrt{2N_0}}.$$

Соответственно, средняя вероятность ошибки имеет вид

$$P_{\text{ош}} \approx \sum_{l=1}^m \sum_{\substack{j=1 \\ j \neq l}}^m \frac{\sqrt{2N_0} \exp(-d^2(s_l, s_j)/(4N_0))}{\sqrt{2\pi} d(s_l, s_j)} p(s_l),$$

что дает удовлетворительную точность при $P_{\text{ош}} < 0,01$.

Расчеты, проведенные в [131], показывают, что при $m \geq 8$ системы с АФМ сигналами обладают более высокой помехоустойчивостью, чем m -ичные системы с фазовой манипуляцией. Например, при $p_{\text{ош}} = 10^{-5}$ и $m = 8$ проигрыш в средней энергии системы с фазовой манипуляцией по сравнению с системой, использующей оптимальный ансамбль сигналов, составляет 1,7 дБ, при $m = 16$ —4,3 дБ, при $m = 32$ —7,1 дБ, при $m = 64$ —10,1 дБ, при $m = 128$ —13,1 дБ. Анализируя эти результаты, можно сделать вывод, что многие из известных ансамблей АФМ сигналов, построенных на основе треугольной и квадратной сетей, и ансамблей с круговым расположением сигнальных точек практически обеспечивают одинаковую помехоустойчивость. По крайней мере, могут быть построены различные типы систем АФМ сигналов, проигрыш которых в средней энергии по сравнению с оптимальными системами не будет превышать 0,5 дБ. Это позволяет выбирать сигналы, для которых реализация модулятора и демодулятора не вызывает трудностей.

При выборе ансамбля сигналов необходимо иметь в виду следующее. Все многопозиционные сигналы можно разделить на два класса. К одному

из них принадлежат сигналы, для которых характерно, что с увеличением объема ансамбля m растет энергетическая эффективность, но при этом расширяется полоса частот, занимаемая сигналами (снижается частотная эффективность). К этому классу относятся ортогональные, биортогональные и симплексные сигналы. При $m \gg 1$ они обеспечивают практически одинаковую помехоустойчивость и являются наилучшими. В то же время их полосы частот по сравнению с двоичными противоположными сигналами шире соответственно в $m/\log_2 m$, $m/(2\log_2 m)$ и $(m-1)/\log_2 m$ раз при той же скорости передачи информации.

К другому классу принадлежат сигналы, для которых с увеличением объема ансамбля m расстояние между сигналами уменьшается (снижается энергетическая эффективность), а полоса частот, занимаемая сигналами, не увеличивается (повышается частотная эффективность). К этому классу относятся АФМ сигналы. Очевидно, что применение АФМ сигналов требует обеспечения линейности и стабильности параметров приемопередающего тракта.

9.4.2. Передача и прием дискретных сообщений в каналах с замираниями

В реальных радиоканалах действуют аддитивные помехи, порождаемые внешними источниками, по своим свойствам отличающиеся от модели гауссовского белого шума, а также случайные искажения сигнала. Виды помех и искажений весьма разнообразны, и учесть все их одновременно при проектировании СПИ не представляется возможным. Однако для каждого диапазона частот можно указать наиболее характерные ситуации, составить математическую модель канала и провести оптимизацию параметров сигналов и алгоритмов их обработки.

Рассмотрим сначала канал с белым шумом и общими замираниями, которые проявляются в изменении уровня сигнала на входе приемника. Если скорость изменения коэффициента передачи канала μ мала по сравнению со скоростью передачи посылок ($\tau_k \gg T_c$), то за время длительности посылки условия приема сигнала практически не меняются и решающая схема, оптимальная для канала с постоянными параметрами, сохраняет свою оптимальность и в данном случае. Однако достоверность принимаемых символов будет меняться во времени в зависимости от μ . Поэтому можно ввести условную вероятность ошибки $P_{\text{ош}}(\mu)$. Учитывая, что коэффициент μ принимает случайные значения, качество передачи информации можно задавать средней вероятностью ошибки $\bar{P}_{\text{ош}}(\mu)$ и надежностью по помехоустойчивости $P(P_{\text{ош}} \leq P_{\text{доп}})$, характеризующей вероятностью непревышения $P_{\text{ош}}(\mu)$ допустимого значения $P_{\text{доп}}$.

Оценим, как влияют общие замирания на помехоустойчивость и надежность для двоичной СПИ. Вероятность ошибки при приеме информации является функцией отношения $h^2 = E/N_0$ и коэффициента взаимной корреляции сигналов: $P_{\text{ош}} = \varphi(h^2; r_{1,2})$. Вид функции $\varphi(h^2; r_{1,2})$ определяется способом обработки сигналов (когерентная, некогерентная и т. д.). Среднюю вероятность ошибки при медленных общих замираниях можно оценить, усредняя $P_{\text{ош}}(\mu)$ по закону распределения $w(\mu)$:

$$\bar{P}_{\text{ош}}(\mu) = \bar{P}_{\text{ош}}(h) = \int_0^{\infty} P_{\text{ош}}(\mu) w(\mu) d\mu = \int_0^{\infty} P_{\text{ош}}(h) w(h) dh.$$

Тогда для канала с рэлеевскими замираниями при некогерентном приеме ортогональных сигналов с активной паузой

$$\bar{P}_{\text{ош}}(h) = \int_0^{\infty} \frac{1}{2} \exp\left(-\frac{h^2}{2}\right) \frac{2h}{h_{\text{ср}}^2} \exp\left(-\frac{h^2}{h_{\text{ср}}^2}\right) dh = \frac{1}{2 + h_{\text{ср}}^2},$$

где $h_{\text{ср}}^2$ — среднее значение $h^2 = (\mu^2 / \mu_{\text{ср}}^2) h_{\text{ср}}^2$.

Сравнивая формулы вероятности ошибки в каналах с замираниями и без замираний, видим, что для получения одинаковой вероятности ошибки, например 10^{-4} , необходимо увеличить мощность сигнала в канале с замираниями в 600 раз. Заметим, что поскольку зависимость вероятности ошибки от параметра h^2 — монотонно убывающая функция экспоненциального характера, то любые флуктуации h относительно среднего значения $h_{\text{ср}}$ приведут к увеличению вероятности ошибки независимо от вида распределения $w(h)$, т. е. $\bar{P}_{\text{ош}}(h) \geq P_{\text{ош}}(h_{\text{ср}})$.

Вероятность ошибки $\bar{P}_{\text{ош}}(h)$ недостаточно полно характеризует качество приема, особенно при передаче сообщений, длительность которых соизмерима с интервалом корреляции замираний. В этом случае вероятность правильного приема сообщений в различных сеансах будет разной. В такой ситуации часто пользуются понятием надежности по помехоустойчивости $P(P_{\text{ош}} \leq P_{\text{доп}})$. Если известны характер замираний $w(h)$, способ приема (функциональная связь между $P_{\text{ош}}$ и h) и значение $P_{\text{доп}}$, то можно записать

$$P(P_{\text{ош}} \leq P_{\text{доп}}) = P(h \leq h_{\text{доп}}) = \int_{h_{\text{доп}}}^{\infty} w(h) dh.$$

Для канала с рэлеевскими замираниями

$$P(P_{\text{ош}} \leq P_{\text{доп}}) = \int_{h_{\text{доп}}}^{\infty} \frac{2h}{h_{\text{ср}}^2} \exp\left(-\frac{h^2}{h_{\text{ср}}^2}\right) dh = \exp\left(-\frac{h_{\text{доп}}^2}{h_{\text{ср}}^2}\right). \quad (9.6)$$

Учитывая, что при некогерентном приеме двоичных символов $P_{\text{доп}} = 0,5 \exp(-h_{\text{доп}}^2/2)$, получаем

$$P(P_{\text{ош}} \leq P_{\text{доп}}) = (2P_{\text{доп}})^{2/h_{\text{ср}}^2}.$$

При низкой достоверности и надежности принимаемой информации в канале с замираниями требуются специальные меры для их повышения. Увеличение мощности передатчика, как видно из соотношений (9.6), неэффективно.

Для повышения помехоустойчивости СПИ в канале с медленными замираниями применяют кодирование, разнесенный прием и передачу с переменной скоростью. Поскольку интервал корреляции медленных замираний τ_k значительно больше длительности посылки T_c , то уровень сигнала на входе приемника с большой вероятностью сохраняется постоянным для серии посылок. Поэтому существует тенденция к группированию ошибок в пакеты. Дискретный канал, в котором ошибки группируются в пакеты, обладает памятью, и для повышения достоверности передачи в нем могут применяться помехоустойчивые коды с перемежением символов.

Принципиально можно избавиться от флуктуаций уровня сигнала на входе решающей схемы, если длительность сигнала сделать обратно пропорциональной коэффициенту передачи канала по мощности. Действительно, на выходе согласованного фильтра отношение сигнал—шум имеет вид

$$q = \frac{\mu^2}{\mu_{\text{ср}}^2} \frac{2\bar{P}_c T_c}{N_0},$$

где \bar{P}_c — средняя мощность сигнала.

Если теперь предположить, что $T_c = T_{\text{ср}} \mu_{\text{ср}}^2 / \mu^2$, то q станет величиной постоянной. Система, использующая такой принцип работы, должна быть адаптивной с переменной длительностью посылки (рис. 9.12).

В состав такой РСПИ входит прямой канал, по которому передается информация от источника к потребителю, и обратный для получения информации о требуемой скорости передачи. На приемной стороне в анализа

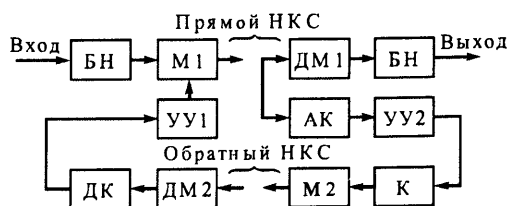


Рис. 9.12. Структурная схема РСПИ с переменной скоростью передачи

торе канала (АК) оценивается текущее отношение h^2 и определяется требуемое значение скорости передачи, которое кодируется в кодере (К), модулирует высокочастотную несущую в модуляторе М2 и по обратному каналу передается на передающую сторону, где эта информация демодулируется в демодуляторе ДМ2, декодируется в декодере ДК и управляет работой модулятора М1, обеспечивая необходимую скорость передачи. Поскольку в большинстве случаев источник выдает информацию с постоянной скоростью, а передача ее по каналу осуществляется с переменной скоростью, то на входе модулятора М1 и выходе демодулятора ДМ1 прямого канала включают буферные накопители (БН), которые выполняют согласование источника и канала. Для управления работой модема служат устройства УУ1 и УУ2. Эффективность работы РСПИ с адаптацией по скорости передачи определяется тем, насколько синхронно будет изменяться T_c и коэффициент передачи μ . Очевидно, что запаздывание информации о параметрах прямого канала должно быть значительно меньше интервала корреляции замираний. В реальных системах для облегчения работы модема системы синхронизации модема скорость передачи посылок по прямому каналу изменяют дискретно. Предельным случаем дискретизации является передача информации «пакетами» в моменты «хорошего» состояния канала, что имеет, например, место при использовании *метеорного* канала.

Адаптивная система позволяет повысить помехоустойчивость, однако требует наличия обратного канала, что не всегда может быть обеспечено. Кроме того, она оказывается практически неработоспособной при замираниях, интервал корреляции которых меньше или соизмерим с запаздыванием информации по обратному каналу.

Эффективным способом повышения качества передачи по каналам с замираниями является разнесенный прием. Сущность его заключается в том, что демодулятор принимает решение о переданном символе по нескольким сигналам, несущим одну и ту же информацию. Основным условием для эффективного приема разнесенных сигналов является независимость замираний в каналах. Разнесенный прием подобен резервированию устройств, обладающих конечной надежностью. Если, например, надежность передачи в одном канале равна $P_1 (P_{\text{ош}} \leq P_{\text{доп}})$, то вероятность того, что $P_{\text{ош}}$ будет меньше $P_{\text{доп}}$ в одном из L каналов, определяется формулой

$$P_L(P_{\text{ош}} \leq P_{\text{доп}}) = 1 - [1 - P_1(P_{\text{ош}} \leq P_{\text{доп}})]^L.$$

При низкой надежности $P_1 \leq 1/L$ разнесенный прием позволяет ее повысить примерно в L раз.

Следует различать способы разнесения и объединения сигналов. Разнесение сигналов заключается в организации каналов, в которых полностью отсутствует или мала степень корреляции уровней сигналов. В РСПИ при-

меняют следующие способы разнесения сигналов: по времени (он сводится к повторению сигнала через промежутки времени, превышающие интервал корреляции замираний); по частоте (сигнал дублируется по многим частотным каналам); прием на различные антенны, разнесенные в пространстве или расположенные в одном месте, но с узкими диаграммами направленности, позволяющими разделить сигналы по углу прихода; прием на две антенны, расположенные в одном месте, но принимающие электромагнитные волны различной поляризации (поляризационное разнесение).

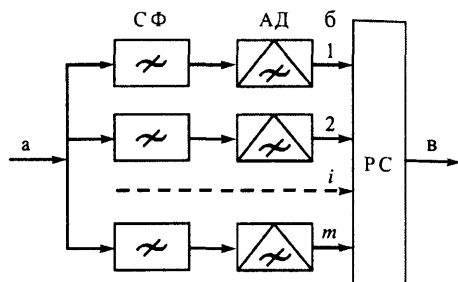


Рис. 9.13. Схема объединения каналов при разнесенном приеме

Поскольку демодулятор дискретных сигналов (рис. 9.13) в общем случае содержит набор согласованных фильтров (СФ), амплитудные детекторы (АД) и решающую схему (РС), то сигналы можно объединять на входе согласованных фильтров, на входе решающей схемы и на ее выходе. Соответственно, различают разнесенный прием с когерентным сложением сигналов (точка а), с последетекторным некогерентным сложением сигналов (точки б) и с дискретным сложением сигналов (точка в). В первых двух случаях для получения максимального отношения сигнал—шум на входе решающей схемы необходимо складывать сигналы с весовыми коэффициентами, определяемыми $\mu_i(t)$. Действительно, нет необходимости складывать сигналы в каналах, в которых на данный момент присутствует только шум или очень слабый сигнал. Кроме того, при когерентном сложении фазы всех суммируемых сигналов должны быть одинаковы. Поэтому для реализации разнесенного приема необходимо измерять уровень и фазу сигналов, приходящих по разным каналам, осуществлять их коррекцию и только потом суммировать. Значение h^2 при оптимальном разнесенном приеме с когерентным сложением равно сумме значений h_i^2 в каждой ветви:

$$h^2 = \sum_{i=1}^L \frac{\mu_i^2(t)}{\mu_1^2 \text{ cp}} h_{1 \text{ cp}}^2 = \frac{1}{L} \sum_{i=1}^L \frac{\mu_i^2(t)}{\mu_1^2 \text{ cp}} h_{\text{cp}}^2 = \sum_{i=1}^L h_i^2,$$

где $h_{1 \text{ cp}}^2 = h_{\text{cp}}^2 / L$.

Величина $\frac{1}{L} \sum_{i=1}^L \frac{\mu_i^2(t)}{\mu_1^2 \text{ cp}}$ случайная. Плотность распределения ее опреде-

ляется видом замираний коэффициента передачи отдельного канала $w(\mu)$ и

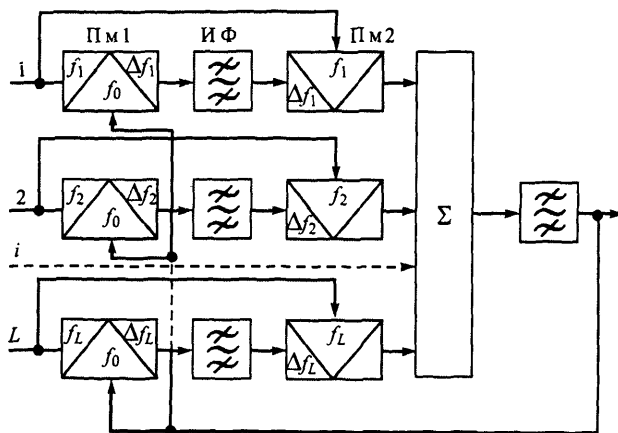


Рис. 9.14. Схема объединения каналов при разнесенном приеме с когерентным сложением сигналов

числом ветвей разнесения L . Ее среднее значение равно 1 независимо от вида распределения $w(\mu)$, а дисперсия обратно пропорциональна \sqrt{L} . Следовательно, с увеличением числа ветвей разнесения при фиксированной суммарной мощности сигнала, передаваемого по всем каналам, дисперсия флуктуаций на входе решающей схемы уменьшается и при $L \rightarrow \infty$ стремится к нулю. На практике ограничиваются числом ветвей разнесения 3...6, поскольку уже при таких значениях удается получить помехоустойчивость, близкую к предельной.

Разнесенный прием с когерентным сложением сигналов требует оценки комплексных коэффициентов передачи по отдельным каналам. Существует ряд оригинальных схем для когерентного сложения разнесенных сигналов.

Рассмотрим работу одной из них (рис. 9.14) при отсутствии шума. В каждой i -й ветви разнесения действует полезный сигнал $u_i(t) = \mu_i(t)s_{ir}(t)$. Предположим, что разнесение частотное, а модуляция фазовая, хотя для рассматриваемой схемы это принципиального значения не имеет. Тогда $u_i(t) = \mu_i(t)\{x_r\} \cos(\omega_0 t + \varphi_i)$, где $\{x_r\}$ — последовательность символов сообщения, $x_r = \pm 1$. В результате когерентного сложения должны получить сигнал

$$u_{\Sigma}(t) = \sum_{i=1}^L \mu_i^2(t) \{x_r\} \cos(\omega_0 t).$$

При большом количестве ветвей разнесения, как было уже показано, этот сигнал должен иметь практически постоянную амплитуду. Используя его в качестве сигнала гетеродина, получаем на выходе перемножителя $\Pi м 1$ в каждом канале в качестве одного из компонентов немодулированный сиг-

нал вида $u_{оп}(t) = \mu_i(t) \cos((\omega_i - \omega_0)t + \varphi_i)$, который несет в себе информацию о коэффициенте передачи канала $\mu_i(t)$ и о фазе сигнала φ_i . В измерительном фильтре (ИФ), который обычно строят на базе системы фазовой автоподстройки, сигнал $u_{оп}(t)$ выделяется из смеси и затем используется как сигнал гетеродина при втором преобразовании. Тогда на выходе перемножителя Пм2 одним из компонентов будет сигнал $\mu_i^2\{x_r\} \cos(\omega_0 t)$, который, собственно, и требовалось получить на входе сумматора (Σ) в каждом канале.

Схему объединения каналов можно существенно упростить, если при суммировании использовать только один сигнал с максимальной амплитудой. Такой метод разнесенного приема носит название *автовывбор ветви с наибольшим сигналом*. Схема (рис. 9.15) содержит коммутатор каналов (Ком), который управляется решающим устройством. Ветви с наибольшим сигналом определяют измерители коэффициентов передачи μ . Поскольку при автовывборе теряется часть энергии принимаемого сигнала, помехоустойчивость будет ниже, чем при сложении разнесенных сигналов.

Дискретное сложение сигналов реализуется наиболее просто. В этом случае решение о переданном символе принимается методом мажоритарного сложения. Для однозначного принятия решения необходимо, чтобы число ветвей разнесения было нечетным: $L = 2n - 1$, где $n = 1, 2, \dots$. Ошибка при дискретном сложении возникает в том случае, если число ошибочно принятых символов превысит $n - 1$. Вероятность такого события имеет вид

$$\bar{P}_{ош}(h) = \sum_{i=n}^{2n-1} C_L^i P_{ош}^i(h) (1 - P_{ош}(h))^{2n-1-i}.$$

В отличие от схемы разнесенного приема с когерентным сложением сигналов, где увеличение L при фиксированной суммарной средней мощности принимаемого по всем каналам сигнала приводит к монотонному уменьшению средней вероятности ошибки, во всех остальных схемах объединения каналов эта зависимость носит экстремальный характер. Это связано с тем, что при увеличении L уменьшается средняя мощность сигнала в каждом канале и, соответственно, растет вероятность ошибок при приеме. При превышении L оптимального значения $L_{опт}$ рост вероятности ошибки на символ может превысить выигрыш от применения разнесения и в итоге привести к снижению помехоустойчивости в целом.

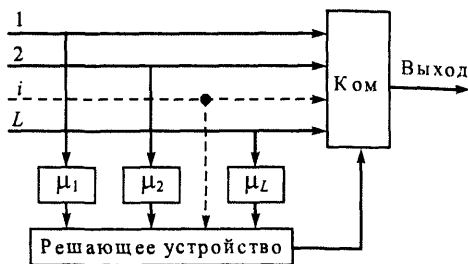


Рис. 9.15. Схема объединения каналов при разнесенном приеме с автовывбором

9.4.3. Передача и прием дискретных сообщений в каналах с небелым шумом

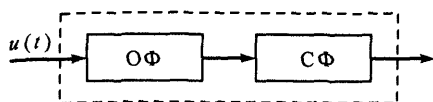


Рис. 9.16. Структурная схема устройства оптимальной обработки в канале с небелым шумом

сигнала должен быть адаптивным. Для этого необходимо, чтобы обесцарапывающий фильтр в схеме на рис. 9.16 был перестраиваемым [11, 131].

Решение задачи реализации ОФ, а также анализа работы РСПИ в условиях действия нестационарного небелого шума существенно упрощается при дискретном представлении спектральной плотности мощности помех. В этом случае спектральную плотность мощности можно записать в виде

$$N(f, t) = \sum_{i=0}^{\infty} N(iF_n, t) \Pi(f - iF_n),$$

где

$$\Pi(f - iF_n) = \begin{cases} 1 & \text{при } (iF_n - 0,5F_n) \leq f \leq (iF_n + 0,5F_n); \\ 0 & \text{при других } f, \end{cases}$$

а F_n — интервал корреляции функции $N(f, t)$ в частотной области.

Схема обесцарапывающего фильтра (рис. 9.17, а) содержит группу полосовых фильтров с полосой пропускания каждого F_n . Частотные характеристики этих фильтров перекрывают всю ширину спектра сигнала $F_c = F_B - F_n$ (рис. 9.17, б).

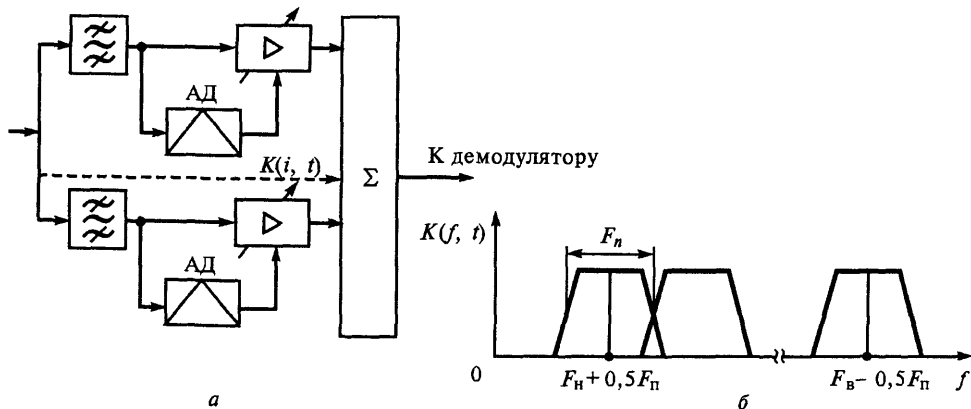


Рис. 9.17. Схема обесцарапывающего фильтра (а) и его частотная характеристика (б)

На выходе каждого из них оценивается уровень помехи, в соответствии с которым затем изменяется коэффициент передачи усилителя $K(i, t) = K / \sqrt{N(iF_n, t)}$.

Поскольку схема ОФ—СФ — основное звено оптимального демодулятора, работающего в условиях небелого шума, оценим, как меняется отношение сигнал—шум на ее выходе при изменении параметров сигналов и помех. Пусть сигнал имеет амплитудный спектр $S(f)$, а помеха — спектральную плотность $N(f, t)$. Тогда на выходе схемы ОФ—СФ отношение сигнал—шум определяется выражением

$$q_{\text{ОФ-СФ}}^2(t) = \left[\int_0^\infty S^2(f) \frac{1}{N(f, t)} df \right]^2 / \int_0^\infty S^2(f) \frac{1}{N(f, t)} df = \int_0^\infty \frac{S^2(f)}{N(f, t)} df. \quad (9.7)$$

Отметим здесь, что отношение сигнал—шум на выходе СФ имеет вид

$$q_{\text{СФ}}^2(t) = \left[\int_0^\infty S^2(f) df \right]^2 / \int_0^\infty S^2(f) N(f, t) df.$$

В зависимости от соотношения ширины спектра сигнала F_c и интервала корреляции помехи в частотной области F_n можно рассмотреть два случая: $F_n \geq F_c$ и $F_c > F_n$. Для РСПИ, в которых $F_n > F_c$, нет необходимости включать обесцараивающий фильтр, так как спектральная плотность $N(f, t)$ практически постоянна в пределах полосы частот F_c . При случайном изменении $N(f, t)$ отношение сигнал—шум — величина случайная и записывается в виде

$$q_{\text{СФ}}^2(t) = \frac{\int_0^\infty S^2(f) df}{N(f_n, t)} = \frac{2E}{N(f_n, t)},$$

где f_n — центральная частота в спектре сигнала. Соответственно, будет меняться и вероятность ошибки при приеме информации. Поэтому здесь, как и при замираниях сигнала, можно ввести понятие средней вероятности ошибки при приеме $\bar{P}_{\text{ош}}(N)$ и надежности по помехоустойчивости $P(P_{\text{ош}} \leq P_{\text{доп}})$. Для их определения необходимо знать плотность вероятности $w_{f_n}(N)$, а также способ обработки сигнала при приеме и вид модуляции, т. е. функциональную связь значения вероятности ошибки с параметром q . Однако даже при известном распределении $w_{f_n}(N)$ нахождение $P_{\text{ош}}(N)$ встречает существенные математические трудности. Здесь, как и во многих других случаях, целесообразно воспользоваться численными методами или метода-

ми математического моделирования. Значительно проще вычислить $P(P_{\text{ош}} \leq P_{\text{доп}})$. Для этого надо задать $P_{\text{доп}}$, рассчитать допустимое значение спектральной плотности помехи $N_{\text{доп}}$ и затем, зная распределение $w_{f_n}(N)$, определить

$$P(P_{\text{ош}} \leq P_{\text{доп}}) = \int_0^{N_{\text{доп}}} w_{f_n}(N) dN.$$

Перейдем теперь к случаю $F_c > F_n$. Здесь функция $N(f, t)$ существенно неравномерна в пределах полосы сигнала. На выходе ОФ—СФ при равномерном спектре сигнала $S(f)$ отношение сигнал—шум имеет вид

$$q_{\text{ОФ-СФ}}^2(t) = \frac{\int_0^\infty S^2(f) df}{\int_0^\infty S^2(f) N(f, t) df} = \frac{2E}{k_2 - k_1 + 1} \sum_{i=k_1}^{k_2} \frac{1}{N(iF_n, t)},$$

где $k_1 = [F_n/F_n]$ и $k_2 = [F_n/F_n]$ — целые числа.

На выходе согласованного фильтра

$$q_{\text{СФ}}^2(t) = \frac{\left(\int_0^\infty S^2(f) df \right)^2}{\int_0^\infty S^2(f) N(f, t) df} = \frac{2E(k_2 - k_1 + 1)}{\sum_{i=k_1}^{k_2} N(iF_n, t)}.$$

Для оценки влияния ширины спектра сигнала на отношение сигнал—шум необходимо проанализировать, как меняется характер распределения случайных величин

$$a(t) = \frac{1}{k_2 - k_1 + 1} \sum_{i=k_1}^{k_2} \frac{1}{N(iF_n, t)} \quad \text{и} \quad b(t) = \frac{k_2 - k_1 + 1}{\sum_{i=k_1}^{k_2} N(iF_n, t)}.$$

Точный количественный анализ помехоустойчивости РСПИ требует знания конкретных законов распределения помех в частотно-временной области и может быть выполнен численными методами или моделированием на ЭВМ. Однако качественные результаты, позволяющие оценить эффективность того или иного способа построения РСПИ, при некоторых упрощениях можно получить, не переходя к точным расчетам. Так, с увеличением ширины спектра сигнала ($F_c \rightarrow \infty$, $k_2 - k_1 \rightarrow \infty$) для $a(t)$ и $b(t)$ соответственно имеем

$$a = \lim_{k_2 - k_1 \rightarrow \infty} a(t) = \overline{1/N(f, t)}; \quad b = \lim_{k_2 - k_1 \rightarrow \infty} b(t) = \overline{1/N(f, t)}.$$

Можно показать, что при любом законе распределения $w_{f_n}(N)$ верно неравенство $\left[\overline{1/N} \right] \geq \overline{[1/N]}$, причем разность $a - b$ тем больше, чем больше дисперсия флуктуаций величины $N(f, t)$. Проведенный качественный анализ позволяет сделать важный вывод: в канале с небелым шумом для повышения помехоустойчивости целесообразно использовать широкополосные сигналы, включая обесцараживающий фильтр на входе демодулятора.

Адаптивный прием широкополосных сигналов с использованием ОФ для повышения помехоустойчивости РСПИ в каналах с сосредоточенными по спектру помехами не всегда применим. Основными причинами этого являются сложности технической реализации ОФ и трудности формирования и обработки широкополосных сигналов при ширине спектра сигнала выше десятков мегагерц. От этих недостатков свободны РСПИ с перестройкой рабочей частоты (ПРЧ) (рис. 9.18). Несущая частота передатчика изменяется дискретно по программе в широких пределах. Приемник перестраивается синхронно с изменением частоты принимаемого сигнала.

В зависимости от скорости переключения несущей частоты различают системы с медленной перестройкой, когда время работы на одной частоте T_4 много больше длительности посылки T_c ($T_4 \gg T_c$), и быстрой, когда $T_4 \ll T_c$. Системы с быстрой перестройкой и когерентным накоплением элементов дискретного частотного сигнала обладают такой же потенциальной помехоустойчивостью, что и системы с широкополосными сигналами, однако реализация обесцараживающего фильтра упрощается. Это связано с тем, что в результате преобразования сигнала в приемнике на выходе фильтра смесителя с полосой, определяемой длительностью элемента ПРЧ сигнала, получаем последовательность радиоимпульсов на промежуточной частоте, амплитуды которых соответствуют уровням смеси сигнала и шума на принимаемых частотах. Таким образом, сигналы ПРЧ как бы трансформируются из спектральной области во временную, в результате чего помехи, сосредоточенные по спектру, преобразуются в помехи, сосредоточенные во времени. Введя теперь схему автоматической регулировки усиления с коэффициентом передачи, обратно пропорциональным значению спектральной плотности мощности помех на данной частоте, получаем в итоге схему (рис. 9.19),

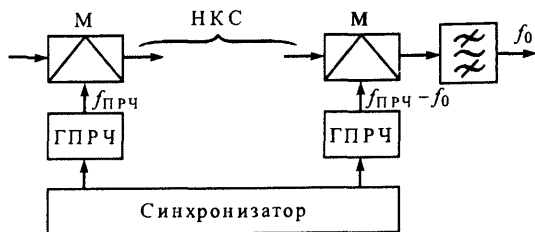


Рис. 9.18. Структурная схема РСПИ с перестройкой частоты

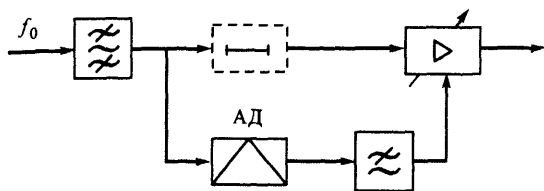


Рис. 9.19. Основное звено обеляющего фильтра системы с ПРЧ

по своим характеристикам эквивалентную обеляющему фильтру. На данной частотной позиции уровень помех измеряется путем амплитудного детектирования смеси. Постоянная времени фильтра нижних частот (ФНЧ) должна быть много меньше длительности

элемента сигнала, чтобы схема автоматической регулировки усиления (АРУ) успевала отработать изменение уровня помех. Однако для повышения помехозащищенности АРУ ее выбирают примерно равной длительности элемента сигнала ПРЧ. При этом в канал обработки вводят задержку, чтобы совпадали элемент ПРЧ сигнала и напряжение АРУ. Полосу пропускания усилителя промежуточных частот (УПЧ) для когерентного ПРЧ сигнала выбирают исходя из ширины спектра информационного сигнала.

Передавать информацию в РСПИ с когерентными сигналами ПРЧ можно с помощью модуляции любого вида, в том числе и ФМ, при которой обеспечивается максимальная помехоустойчивость.

Трудности реализации РСПИ с «быстрыми» ПРЧ сигналами связаны с ограниченными возможностями создания синтезаторов частот, в которых сохраняется когерентность радиоимпульсов, формируемых на разных частотах. Поэтому в настоящее время часто используют системы с некогерентными ПРЧ сигналами. Такие сигналы не позволяют реализовать когерентное сложение элементов, что влечет за собой энергетические потери. Кроме того, возникают ограничения в выборе вида модуляции. Фазовые методы модуляции здесь неприменимы. Для получения ортогональных сигналов с любым основанием m можно использовать частотную манипуляцию или частотно-временное кодирование внутри посылки.

Рассмотрим, как формируются ортогональные ПРЧ сигналы (рис. 9.20). Посылку длительностью T_c разбивают на N временных интервалов, в течение которых передается один элемент сигнала на частоте f_i . Для m возможных сигналов последовательность частот за время T_c должна быть своя, причем такая, чтобы наложение частотно-временных матриц не давало совпадений. Получить ортогональные ПРЧ сигналы можно и с помощью частотной манипуляции, смещая результирующий спектр посылки на фиксированные частотные интервалы Δf .

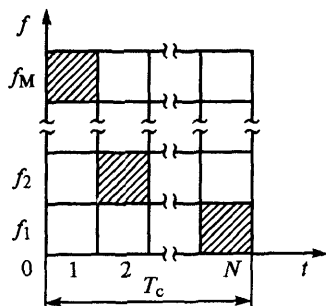


Рис. 9.20. Диаграмма формирования сигналов с ПРЧ

Некогерентная обработка ПРЧ сигнала отличается от когерентной тем, что суммирование элементов осуществляется после амплитудного детектирования. Это приводит к энергетическим потерям, которые зависят от отношения $h_0^2 = E_0/N_0$ для элемента сигнала. Для $h_0^2 \gg 1$ потерь практически нет.

Рассмотрим теперь систему с «медленной» ПРЧ. Особенность ее работы заключается в том, что отношение сигнал—помеха на входе решающей схемы зависит от частоты несущей. Поэтому вероятность ошибки при приеме будет меняться. Дискретный канал приобретает «память» с интервалом корреляции ошибок, равным продолжительности работы на одной частоте при условии, что скорость изменения помеховой обстановки меньше скорости переключения частот. В противном случае отношение сигнал—помеха может изменяться за время работы на одной частоте и «память» канала будет определяться скоростью изменения спектральной плотности помех.

Поскольку в этом случае имеем канал с переменными параметрами, качество передачи можно характеризовать средней вероятностью ошибки и надежностью по помехоустойчивости. Методика их оценки такая же, как и для узкополосной системы с произвольно выбранной несущей частотой. Основная сложность при расчетах заключается в определении плотности вероятностей $w(h)$ и усреднении условной вероятности ошибки $p_{\text{ош}}(h)$ по этому закону.

Таким образом, неравномерность спектральной плотности помех по-разному влияет на помехоустойчивость систем с ПРЧ при медленном и быстром переключении частот. Для систем с быстрым переключением и когерентным накоплением элементов сигнала помехоустойчивость растет с увеличением степени неравномерности помех по спектру, а для систем с медленным переключением это, как видно на примере одиночной помехи, не всегда выполняется. Можно существенно повысить помехоустойчивость систем с медленной ПРЧ, если применить кодирование, эффективное для дискретного канала с памятью.

Теперь вернемся к выражению (9.7) для определения отношения сигнал—помеха на входе решающей схемы. Предположим, что сведения о спектральной плотности помех $N(f, t)$ известны на передающей стороне. Тогда

очевидно, что при фиксированной энергии посылки $E = \frac{1}{2} \int_0^\infty S^2(f) df$ значение $q_{\text{сф}}$ будет наибольшим при минимальном знаменателе $\int_0^\infty S^2(f) N(f, t) df$. Это

условие выполняется, если

$$S^2(f) = 2E\delta(f-f_0); \quad N(f_0, t) = \min,$$

где $\delta(f-f_0)$ — дельта-функция.

Таким образом, для получения максимального отношения сигнал—помеха на входе решающей схемы спектр сигнала должен быть предельно узким и иметь несущую частоту, на которой в момент передачи мощность помех минимальна. Для реализации такого метода повышения качества передачи в условиях небелого шума необходим обратный канал, по которому должна передаваться информация о помеховой обстановке в месте приема. В принципе, достаточно передать значение частоты, соответствующее минимуму помех, или при наборе дискретных значений частот — номер рабочей частоты.

Прямой и обратный каналы НКС1 и НКС2 содержат РСПИ с адаптивным переключением несущей частоты (рис. 9.21). В прямом канале передается информация от источника к потребителю, в обратном — значение оптимальной рабочей частоты. Для определения рабочей частоты в приемнике имеется анализатор канала (АК), в котором измеряется спектральная плотность помех в отведенном для передачи диапазоне частот. Решающее устройство (РУ) определяет значение оптимальной рабочей частоты, которое кодируется в кодере (К) и передается по обратному каналу. Таким образом, реализуется практически синхронный переход с частоты на частоту при изменении помеховой обстановки. При идеальном обратном канале, по которому информация передается без задержки и ошибок, отношение сигнал—помеха q на выходе СФ будет равно $2E/N_{\min}$, где N_{\min} — минимальное значение спектральной плотности помех в отведенном для передачи диапазоне частот. В канале, где спектральная плотность помех изменяется, величина q будет случайной и для определения средней вероятности ошибки необходимо

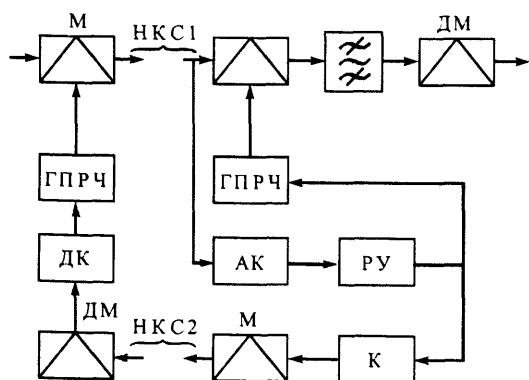


Рис. 9.21. Структурная схема РСПИ с адаптивной перестройкой рабочей частоты

знать закон распределения $w(q)$. Для этого нужно определить плотность распределения максимума случайной величины $1/N(f)$ в частотной области при заданном интервале наблюдений. Решение задачи несколько упрощается при переходе к дискретной модели помех, действующих в канале. При этом достаточно найти плотность распределения максимума конечного числа независимых отсчетов, равного

F_K/F_n , где F_n — интервал корреляции помех по частоте. Для количественной оценки можно рекомендовать методы математического моделирования.

Значительно проще оценить надежность по помехоустойчивости при адаптивном выборе оптимальной рабочей частоты. В этом случае необходимо определить отношение сигнал—помеха или значение спектральной плотности помех $N_{\text{доп}}$, обеспечивающие требуемую достоверность. Затем, зная плотность распределения спектральной плотности помех $w(N)$, найдем вероятность $P(N \leq N_{\text{доп}})$, характеризующую ситуацию, когда уровень помех на произвольно выбранной частоте меньше допустимого. При адаптивном выборе частоты из набора, содержащего M фиксированных частот, надежность по помехоустойчивости определяется формулой

$$P_M(N \leq N_{\text{доп}}) = 1 - [1 - P(N \leq N_{\text{доп}})]^M.$$

В реальных ситуациях несущая частота в прямом канале может не совпадать с оптимальным значением. Это отличие обусловлено в основном двумя причинами: ошибками, вносимыми анализатором канала, и потерей информации в обратном канале. Для определения оптимальной частоты необходимо затратить определенное время. Если спектральная плотность помех меняется, то ее вид будет воспроизводиться с некоторой погрешностью. Используя оптимальные методы фильтрации, основанные на экстраполяции и интерполяции результатов наблюдения, можно уменьшить погрешности измерения, но исключить их полностью нельзя.

Потеря информации в обратном канале связана с ее задержкой и ошибками при передаче. Обратный канал в рассмотренной РСПИ работает в менее благоприятных условиях, чем прямой. Это определяется случайным выбором несущей частоты. В момент передачи мощность помех в обратном канале может быть недопустимо большой и, как следствие, достоверность передаваемой информации низкой. Для повышения достоверности передачи в обратном канале можно, например, использовать широкополосные сигналы с применением оптимальной обработки их в приемнике.

9.5. Помехоустойчивое кодирование и декодирование

9.5.1. Принципы построения кодеков

Кодированием и декодированием в широком смысле называют любое преобразование сообщения в сигнал и, обратно, сигнала в сообщение путем установления взаимного соответствия. Преобразование следует считать *оптимальным*, если в итоге производительность источника и пропускная способность непрерывного канала окажутся равными, т. е. полностью исполь-

зуются возможности канала. К сожалению, такая постановка задачи не дает ответа на вопрос: что для этого надо делать? Поэтому это преобразование и разбивают на два этапа, а именно на этап модуляции—демодуляции, который позволяет перейти от непрерывного канала (радиоканала) к дискретному, и этап кодирования—декодирования в узком смысле, где все операции совершаются над последовательностью символов. Однако и само кодирование—декодирование имеет два противоположных по своим функциям этапа: устранение избыточности в получаемом от источника сигнале (экономное кодирование) и внесение избыточности в передаваемый по каналу цифровой сигнал (помехоустойчивое или избыточное кодирование) для повышения достоверности передаваемой информации.

Экономное кодирование направлено на то, чтобы передаваемый дискретный сигнал имел максимальную энтропию (максимальное количество информации на символ). Тогда для его передачи по радиоканалу с выбранным модемом потребуется минимальная полоса частот. Не рассматривая методы экономного кодирования, укажем только основные свойства дискретного сигнала, в котором полностью устранена избыточность, — это равная вероятность и независимость их появления в последовательности. В этом случае среднее количество информации на символ равно $\log_2 m$.

Примером экономного кодирования является передача речевого сигнала по цифровым каналам. Если ориентироваться только на смысловое (информационное) содержание, то можно перейти к передаче текста со скоростью 5...10 букв в секунду. С учетом объема алфавита в двоичном канале это потребует скорость передачи 25...50 бит/с. Если устранить избыточность, связанную с неодинаковой вероятностью их появления и их корреляцией в тексте, то, как показывают расчеты, скорость передачи может быть уменьшена до 10 бит/с. Если предавать речь в цифровой форме, используя аналогово-цифровое преобразование, ориентируясь только на ширину спектра и динамический диапазон, то скорость потока двоичных символов составит 32...64 кбит/с. Такая колоссальная избыточность привела к необходимости разработки специальных кодеков речевых сигналов, называемых *вокодерами*, которые нашли применение при передаче речи в цифровой форме по радиоканалам. Например, в сотовых системах мобильной связи стандарта GSM скорость передачи составляет 8,5 кбит/с, причем сохраняется не только смысловое содержание, но и индивидуальные особенности говорящего. Подобные кодеки находят применение и при передаче подвижных и неподвижных изображений в цифровых телевизионных каналах.

При помехоустойчивом кодировании в поток передаваемых символов вводятся дополнительные (избыточные) символы для исправления возникающих на приемной стороне ошибок. Это требует увеличения скорости

передачи по каналу, что при выбранном типе модема эквивалентно расширению полосы частот сигнала и уменьшению энергии посылки. Поэтому может возникнуть правомерный вопрос о целесообразности использования вообще избыточного кодирования. На этот вопрос дает ответ теорема Шеннона о пропускной способности непрерывного канала связи, из которой следует, что пропускная способность непрерывного канала увеличивается с расширением его полосы, но при оптимальном в широком смысле кодировании. Поэтому следует ожидать повышения достоверности передачи при заданной скорости и отношении сигнал—шум в канале при внесении избыточности. Однако ответа на вопрос «каким должен быть оптимальный код?» нет, да и, наверное, для сообщения, не фиксированного по длительности, не может быть. Тем не менее, избыточное кодирование широко используется для повышения верности передачи особенно в последние десятилетия, когда проблема создания сложных вычислительных устройств в малых габаритах практически решена.

Рассмотрим принципы кодирования на примере двоичного канала. Допустим, что источник обладает максимальной производительностью. Тогда обязательным условием внесения избыточности является увеличение числа переданных посылок за единицу времени по сравнению с их числом, поступающим от источника. Как вносятся избыточные символы, определяет тип кода. Наиболее просто предположить, что группе из k символов источника ставится в соответствие n символов, передаваемых по каналу. Такой код называется *блочным* и записывается условно как (n, k) -код. Возможны непрерывные коды, которые характеризуются тем, что операции кодирования и декодирования производятся над непрерывной последовательностью символов без разбиения ее на блоки.

Рассмотрим принципы помехоустойчивого кодирования на примере блочного двоичного кода как наиболее простого [11, 132]. Если к символам источника добавляются избыточные символы, то код называют *систематическим*. Если группе информационных символов ставится в соответствие новая группа символов, передаваемая по каналу, в которой информационных символов в явном виде нет, то код называется *неразделимым*. Теперь ответим на основной вопрос: как строить кодек, чтобы при фиксированной избыточности $\chi = r/n$, где r — число проверочных символов для делимого кода, достоверность передачи была бы максимальной? При передаче безыбыточным примитивным кодом с числом разрядов k в каждом слове все $N_p = 2^k$ комбинаций являются разрешенными и ошибка хотя бы в одном символе приводит к тому, что одна разрешенная комбинация переходит в другую и происходит ошибка в приеме сообщения.

Введение избыточных символов приводит к тому, что полное число комбинаций увеличивается и становится равным $N = 2^n$, причем часть из них

$N - N_p$ являются запрещенными и могут возникать только тогда, когда в канале происходят ошибки. Этот факт положен в основу обнаружения и исправления ошибок.

Введем понятие кодового расстояния. Предварительно отметим, что для оценки отличия одной кодовой комбинации от другой можно использовать *расстояние Хэмминга* $d(N_i, N_j)$, определяемое числом разрядов, в которых одна кодовая комбинация отличается от другой. Для двоичного кода

$$d(N_i, N_j) = \sum_{k=1}^n b_{ik} \oplus b_{jk},$$

где b_{ik} и b_{jk} — k -е символы кодовых комбинаций N_i и N_j соответственно; \oplus — символ суммирования по модулю 2. Наименьшее расстояние Хэмминга для данного кода называется *кодovým расстоянием*. В дальнейшем его будем обозначать через d .

При независимых ошибках в канале корректирующую способность кода удастся выразить через кодовое расстояние. Пусть имеется код с $d = 1$. Учитывая, что искажение одного символа изменяет расстояние Хэмминга на одну единицу, при применении кода с $d = 1$ обнаруживаются не все одиночные ошибки. Для того чтобы код мог обнаруживать любую одиночную ошибку, необходимо обеспечить кодовое расстояние, равное двум. Рассуждая аналогичным образом, получаем, что для обнаружения всех ошибок кратности l требуется код с

$$d \geq l + 1.$$

Для исправления всех ошибок некоторой кратности требуется большее кодовое расстояние, чем для их обнаружения. Если кратность исправляемых ошибок равна l , то кодовое расстояние должно удовлетворять условию

$$d \geq 2l + 1.$$

Очевидно, что кодовое расстояние между разрешенными комбинациями можно сделать тем больше, чем больше избыточность. Однако при этом уменьшается длительность посылок и возрастает вероятность ошибок при их приеме. Поэтому вводят понятие эффективности избыточного кодирования как отношение вероятностей ошибочного приема кодовой комбинации из k информационных символов при передаче их примитивным и избыточным кодами:

$$\xi = P(k) / P(n, k).$$

Если код примитивный, то ошибка возникает, если хотя бы в одном символе при приеме произошла ошибка. Вероятность такого события равна

$$P(k) = 1 - (1 - P_{\text{ош.п}})^k \approx kP_{\text{ош.п}},$$

где $P_{\text{ош.п}}$ — вероятность ошибки в приеме одного символа при передаче сообщения примитивным кодом.

Для избыточного кода ошибка в приеме кодовой комбинации будет иметь место тогда, когда число ошибок превысит исправляющую способность кода t_n и ее вероятность

$$P(n, k) = \sum_{i=t_n+1}^n C_n^i P_{\text{ош.и}}^i (1 - P_{\text{ош.и}})^{n-i}$$

где $P_{\text{ош.и}}$ — вероятность ошибки в приеме одного символа при передаче избыточным кодом.

Различие в $P_{\text{ош.п}}$ и $P_{\text{ош.и}}$ определяется уменьшением длительности посылки при передаче избыточным кодом, пропорциональной отношению k/n . Величины $P_{\text{ош.п}}$ и $P_{\text{ош.и}}$ могут быть найдены, если известен вид модуляции и демодуляции, отношение P_e/N_0 и длительность посылок источника.

Таким образом, задача построения кода с заданной корректирующей способностью сводится к обеспечению необходимого кодового расстояния при введении избыточности. При этом желательно, чтобы число используемых проверочных символов было минимальным. К сожалению, задача определения минимального числа проверочных символов, необходимых для обеспечения заданного кодового расстояния, не решена. Имеется лишь ряд оценок для максимального кодового расстояния при фиксированных n и k , которые часто используются при выяснении того, насколько построенный код близок к оптимальному.

Можно показать [133], что если существует блочный линейный код (n, k) , то для него справедливо неравенство

$$r \geq \log_2 \left(\sum_{i=0}^{\lfloor (d-1)/2 \rfloor} C_n^i \right), \quad (9.8)$$

называемое *верхней границей Хэмминга*, где $\lfloor (d-1)/2 \rfloor$ означает целую часть числа $(d-1)/2$.

Граница Хэмминга (9.8) близка к оптимальной для кодов с большими значениями k/n . Для кодов с малыми значениями k/n более точной является *верхняя граница Плоткина*:

$$r \geq 2d - 2 - \log_2(d).$$

Можно также показать, что существует блочный линейный код (n, k) с кодовым расстоянием d , для которого справедливо неравенство

$$r \leq \log_2 \left(\sum_{i=0}^{d-2} C_n^i \right),$$

называемое *нижней границей Варшавова—Гильберта*.

Таким образом, границы Хэмминга и Плоткина являются необходимыми условиями существования кода, а граница Варшавова—Гильберта — достаточным. Эти границы позволяют оценить эффективность блочных кодов и целесообразность их применения.

9.5.2. Линейные блочные коды

Линейный (n, k) -код можно получить из k линейно независимых кодовых комбинаций путем их посимвольного суммирования по модулю 2 в различных сочетаниях. Исходные линейно независимые кодовые комбинации называются *базисными*.

Представим базисные кодовые комбинации в виде $(n \times k)$ -матрицы

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ g_{k1} & g_{k2} & \cdots & g_{kn} \end{bmatrix}. \quad (9.9)$$

В теории кодирования матрица \mathbf{G} называется *порождающей*. Тогда процесс кодирования заключается в выполнении операции

$$\mathbf{V} = \mathbf{A}\mathbf{G},$$

где \mathbf{A} — вектор размерности k , соответствующий сообщению; \mathbf{V} — вектор размерности n , соответствующий кодовой комбинации.

Таким образом, порождающая матрица (9.9) содержит всю необходимую для кодирования информацию. Она должна храниться в памяти кодирующего устройства. Для двоичного кода объем памяти равен $k \times n$ двоичных символов. При табличном задании кода кодирующее устройство должно запоминать $n \cdot 2^k$ двоичных символов.

Две порождающие матрицы, отличающиеся друг от друга только порядком расположения столбцов, задают коды, которые имеют одинаковые расстояния Хэмминга между соответствующими кодовыми комбинациями, а следовательно, одинаковые корректирующие способности. Такие коды называются *эквивалентными*.

В качестве базисных комбинаций часто выбирают кодовые комбинации, содержащие по одной единице среди информационных символов. При этом порождающую матрицу (9.9) удастся записать в канонической форме:

$$\mathbf{G} = \|\mathbf{I}, \mathbf{P}\| = \begin{bmatrix} 100\dots 00 & g_{1k+1} & \dots & g_{1n} \\ 010\dots 00 & g_{2k+1} & \dots & g_{2n} \\ \dots & \dots & \dots & \dots \\ 000\dots 01 & g_{kk+1} & \dots & g_{kn} \end{bmatrix}, \quad (9.10)$$

где \mathbf{I} — единичная $(k \times k)$ -подматрица; \mathbf{P} — $(k \times (n - k))$ -подматрица проверочных символов, определяющая свойства кода. Матрица (9.10) задает систематический разделимый код. Можно показать, что для любого линейного кода существует эквивалентный систематический код.

Линейный (n, k) -код может быть задан так называемой *проверочной* матрицей \mathbf{H} размерности $r \times n$. При этом некоторая комбинация \mathbf{B} принадлежит коду только в том случае, если вектор \mathbf{B} ортогонален всем строкам матрицы \mathbf{H} , т. е. если выполняется равенство

$$\mathbf{B}\mathbf{H}^T = 0, \quad (9.11)$$

где T — символ транспонирования матрицы.

Поскольку равенство (9.11) справедливо для любой кодовой комбинации, то

$$\mathbf{G}\mathbf{H}^T = 0.$$

Каноническая форма матрицы \mathbf{H} имеет вид

$$\mathbf{H} = \|\mathbf{P}^T, \mathbf{I}\| = \begin{bmatrix} g_{1k+1} & g_{2k+1} & \dots & g_{kk+1} & 100\dots 00 \\ g_{1k+2} & g_{2k+2} & \dots & g_{kk+2} & 010\dots 00 \\ \dots & \dots & \dots & \dots & \dots \\ g_{1n} & g_{2n} & \dots & g_{kn} & 000\dots 01 \end{bmatrix}, \quad (9.12)$$

где \mathbf{P}^T — подматрица, столбцами которой служат строки подматрицы \mathbf{P} из соотношения (9.10); \mathbf{I} — единичная $(r \times r)$ -подматрица.

Подставляя (9.12) в (9.11), можно получать $n - k$ уравнений вида

$$b_{k+j} \oplus \sum_{i=1}^k \oplus g_{ik+j} b_i = 0, \quad j = 1, 2, \dots, n - k, \quad (9.13)$$

которые называются *уравнениями проверки*. Из (9.13) следует, что проверочные символы кодовых комбинаций линейного кода образуются различными линейными комбинациями информационных символов. Единицы в любой j -й строке подматрицы \mathbf{P}^T , входящей в проверочную матрицу (9.12), указывают, какие информационные символы участвуют в формировании j -го проверочного символа.

Очевидно, что линейный (n, k) -код можно построить, используя уравнения проверки (9.13). При этом первые k символов кодовой комбинации информационные, а остальные $n - k$ символов — проверочные, образуемые в соответствии с формулой (9.13).

С помощью проверочной матрицы сравнительно легко можно построить код с заданным кодовым расстоянием. Это построение основано на следующей теореме: кодовое расстояние линейного (n, k) -кода равно d тогда и только тогда, когда любые $d - 1$ столбцов проверочной матрицы этого кода линейно независимы, но некоторые d столбцов проверочной матрицы линейно зависимы.

Заметим, что строки проверочной матрицы линейно независимы. Поэтому проверочную матрицу можно использовать в качестве порождающей для некоторого другого линейного кода $(n, n - k)$, называемого *двойственным*.

Кодирующее устройство для линейного (n, k) -кода (рис. 9.22) состоит из k -разрядного сдвигающего регистра и $r = n - k$ блоков сумматоров по модулю 2. Информационные символы одновременно поступают на вход регистра и на выход кодирующего устройства через коммутатор К. С поступлением k -го информационного символа на выходах блоков сумматоров в соответствии с уравнениями (9.13) формируются проверочные символы, которые затем последовательно поступают на выход кодера.

Процесс декодирования сводится к выполнению операции

$$\mathbf{S} = \tilde{\mathbf{B}}\mathbf{H}^T,$$

где \mathbf{S} — вектор размерности $n - k$, называемый *синдромом*; $\tilde{\mathbf{B}}$ — вектор принятой кодовой комбинации.

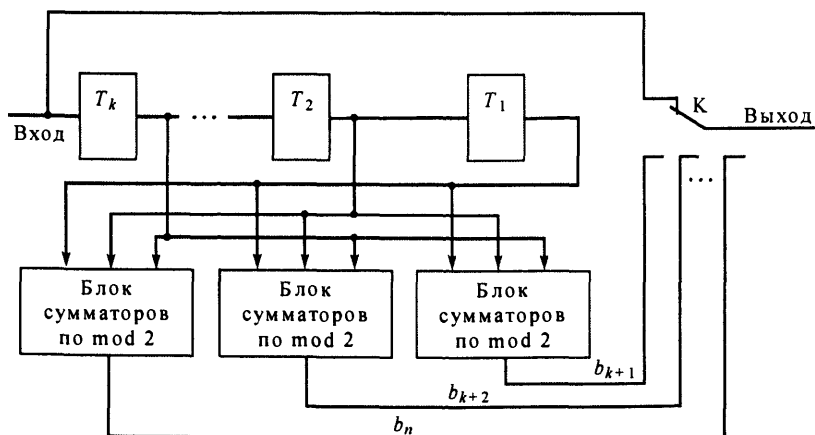


Рис. 9.22. Структурная схема кодера линейного кода

Если принятая кодовая комбинация $\tilde{\mathbf{B}}$ совпадает с одной из разрешенных \mathbf{B} (это имеет место тогда, когда либо ошибки в принятых символах отсутствуют, либо при действии помех одна разрешенная кодовая комбинация переходит в другую), то

$$\mathbf{S} = \tilde{\mathbf{B}}\mathbf{H}^T = 0.$$

В противном случае $\mathbf{S} \neq 0$, причем вид синдрома зависит только от вектора ошибок \mathbf{e} . Действительно,

$$\mathbf{S} = \tilde{\mathbf{B}}\mathbf{H}^T = (\mathbf{B} \oplus \mathbf{e})\mathbf{H}^T = \mathbf{e}\mathbf{H}^T,$$

где \mathbf{B} — вектор, соответствующий передаваемой кодовой комбинации. При $\mathbf{S} = 0$ декодер принимает решение об отсутствии ошибок, а при $\mathbf{S} \neq 0$ — о наличии ошибок. Число различных синдромов, соответствующих разным сочетаниям ошибок, равно $2^{n-k} - 1$. По конкретному виду синдрома можно в пределах корректирующей способности кода указать на ошибочные символы и исправить их.

Декодер линейного кода (рис. 9.23) состоит из k -разрядного сдвигающего регистра, $n - k$ блоков сумматоров по модулю 2, схемы сравнения, анализатора ошибок и корректора. Регистр служит для запоминания информационных символов принятой кодовой последовательности, из которых в блоках сумматоров формируются проверочные символы. Анализатор оши-

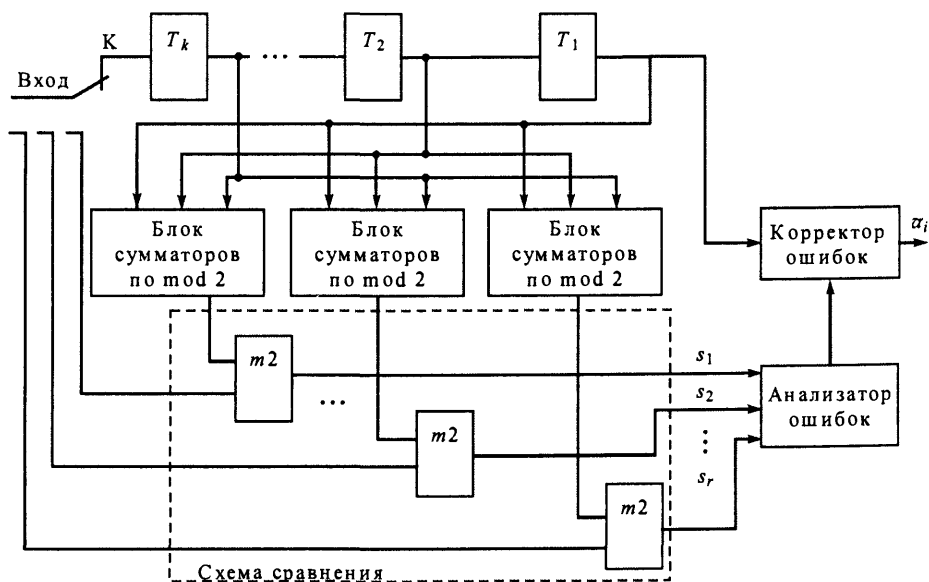


Рис. 9.23. Структурная схема декодера линейного кода

бок по конкретному виду синдрома, получаемого в результате сравнения формируемых на приемной стороне и принятых проверочных символов, определяет места ошибочных символов. Исправление информационных символов проводится в корректоре.

Заметим, что в общем случае при декодировании линейного кода с исправлением ошибок в памяти декодера должна храниться таблица соответствий между синдромами и векторами ошибок, содержащая 2^{n-k} строк. С приходом каждой кодовой комбинации декодер должен перебрать всю таблицу. При небольших значениях $n - k$ эта операция не вызывает затруднений. Однако для высокоэффективных кодов длиной n , равной нескольким десяткам, разность $n - k$ принимает такие значения, что перебор таблицы из 2^{n-k} строк оказывается практически невозможным. Например, для кода (63, 51), имеющего кодовое расстояние $d = 5$, таблица состоит из $2^{12} = 4096$ строк. При заданных значениях n и k существует $2^{k(n-k)}$ линейных кодов. Задача заключается в выборе наилучшего (с позиции того или иного критерия) кода. Следует заметить, что до сих пор общие методы синтеза оптимальных линейных кодов не разработаны.

Циклические коды относятся к классу линейных систематических. Поэтому для их построения, в принципе, достаточно знать порождающую матрицу.

Можно указать другой способ построения циклических кодов, основанный на представлении кодовых комбинаций многочленами $b(x)$ вида

$$b(x) = b_{n-1} x^{n-1} \oplus b_{n-2} x^{n-2} \oplus \dots \oplus b_1 x \oplus b_0,$$

где b_i , $i = 0, 1, \dots, n - 1$, — символы кодовой комбинации. Над данными многочленами можно проводить все алгебраические действия с учетом того, что сложение здесь осуществляется по модулю 2.

Каждый циклический код (n, k) характеризуется так называемым *порождающим* многочленом. Им может быть любой многочлен $p(x)$ степени $n - k$, который делит без остатка двучлен $x^n \oplus 1$. Циклические коды характеризуются тем, что многочлены $b(x)$ кодовых комбинаций делятся без остатка на $p(x)$. Поэтому процесс кодирования сводится к нахождению по известным многочленам $a(x)$ и $p(x)$ многочлена $b(x)$, делящегося на $p(x)$, где $a(x)$ — многочлен степени $k - 1$, соответствующий информационной последовательности символов.

Очевидно, что в качестве многочлена $b(x)$ можно использовать произведение $a(x)p(x)$. Однако при этом код оказывается несистематическим, что затрудняет процесс декодирования. Поэтому на практике, в основном, применяется следующий метод нахождения многочлена $b(x)$.

Умножим многочлен $a(x)$ на x^{n-k} и полученное произведение разделим на $p(x)$. Пусть

$$a(x)x^{n-k} = m(x)p(x) \oplus c(x), \quad (9.14)$$

где $m(x)$ — частное, а $c(x)$ — остаток. Поскольку операции суммирования и вычитания по модулю 2 совпадают, то выражение (9.14) перепишем в виде

$$a(x)x^{n-k} \oplus c(x) = m(x)p(x). \quad (9.15)$$

Из соотношения (9.15) следует, что многочлен $a(x)x^{n-k} \oplus c(x)$ делится на $p(x)$ и, следовательно, является искомым.

Многочлен $a(x)x^{n-k}$ имеет следующую структуру: первые $n - k$ членов низшего порядка равны нулю, а коэффициенты остальных совпадают с соответствующими коэффициентами информационного многочлена $a(x)$. Многочлен $c(x)$ имеет степень меньше $n - k$. Таким образом, в найденном многочлене $b(x)$ коэффициенты при x в степени $n - k$ и выше совпадают с информационными символами, а коэффициенты при остальных членах, определяемых многочленом $c(x)$, совпадают с проверочными символами.

В соответствии с формулой (9.15) процесс кодирования заключается в умножении многочлена $a(x)$ на x^{n-k} и нахождении остатка от деления $a(x)x^{n-k}$ на $p(x)$ с последующим его сложением по модулю 2 с многочленом $a(x)x^{n-k}$.

Операции умножения и деления многочленов легко осуществляются линейными цепями на основе сдвигающих регистров [133]. В качестве примера на рис. 9.24, а представлена схема умножения многочлена $b(x)$ степени $n = 6$ на многочлен $f(x) = x^3 \oplus x \oplus 1$ по модулю $x^7 \oplus 1$. Нетрудно убедиться, что после семи тактов в регистре записывается многочлен $b(x)f(x) \bmod (x^7 \oplus 1)$. При делении многочлена $b(x)$ степени $n = 6$ на многочлен $f(x) = x^3 \oplus x^2 \oplus 1$ (рис. 9.24, б) после семи тактов в регистре оказывается записанным остаток от деления.

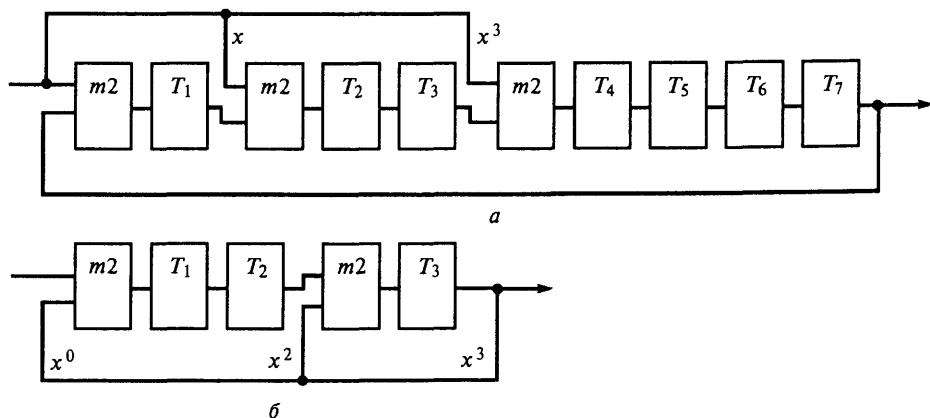


Рис. 9.24. Схемы умножения (а) и деления (б) многочленов (частный случай)

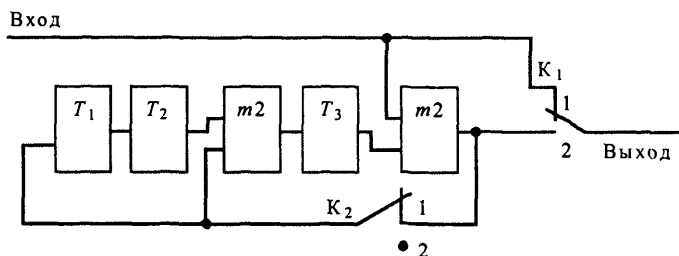


Рис. 9.25. Структурная схема кодера циклического кода с порождающим многочленом $p(x) = x^3 \oplus x^2 \oplus 1$

На основе приведенных схем умножения и деления многочленов строятся кодирующие устройства для циклических кодов. На рис. 9.25 в качестве примера приведена схема кодера для кода (7, 4) с порождающим многочленом $p(x) = x^3 \oplus x^2 \oplus 1$. В исходном состоянии ключи K_1 и K_2 находятся в положении 1. Информационные символы поступают одновременно на вход канала и на выход ячейки x^3 сдвигающего регистра (это соответствует умножению многочлена $a(x)$ на x^3). В течение четырех тактов происходит деление многочлена $a(x)x^3$ на многочлен $p(x) = x^3 \oplus x^2 \oplus 1$. В результате в регистре записывается остаток, представляющий собой проверочные символы. Ключи K_1 и K_2 перебрасываются в положение 2, и в течение трех последующих тактов содержащиеся в регистре символы поступают в канал.

Циклический код может быть задан *проверочным* многочленом $h(x)$: кодовая комбинация B принадлежит данному циклическому коду, если $b(x)h(x) = 0 \bmod (x^n \oplus 1)$. Проверочный многочлен связан с порождающим соотношением

$$h(x) = (x^n \oplus 1)/p(x).$$

Задание кода проверочным многочленом эквивалентно заданию кода системой проверочных уравнений (9.13). Характерной особенностью циклического кода является то, что все проверочные уравнения можно получить из одного путем циклического сдвига индексов символов, входящих в исходное уравнение. Например, для кода (7, 4) с порождающим многочленом $p(x) = x^3 \oplus x^2 \oplus 1$ проверочный многочлен имеет вид $h(x) = x^4 \oplus x^3 \oplus x^2 \oplus 1$. Проверочные уравнения получаются из условия

$$b(x)h(x) = 0 \bmod (x^7 \oplus 1).$$

Осуществив умножение и приравняв коэффициенты при x^4 , x^5 и x^6 нулю, получим следующие уравнения, разрешенные относительно проверочных символов:

$$\begin{aligned} b_2 &= b_6 \oplus b_4 \oplus b_3, \\ b_1 &= b_5 \oplus b_3 \oplus b_2, \\ b_0 &= b_4 \oplus b_2 \oplus b_1. \end{aligned} \quad (9.16)$$

В качестве примера на рис. 9.26 показана схема кодера циклического кода (7,4), задаваемого проверочным многочленом $h(x) = x^4 \oplus x^3 \oplus x^2 \oplus 1$, или, что то же самое, проверочными соотношениями (9.16). В исходном состоянии ключ находится в положении 1. В течение четырех тактов импульсы поступают в регистр, после чего ключ переводится в положение 2. При этом обратная связь замыкается. Начиная с пятого такта, формируются проверочные символы в соответствии с соотношениями (9.16). После седьмого такта все проверочные символы оказываются сформированными, ключ вновь переключается в положение 1. Кодер готов к приему очередного сообщения. Символы кодовой комбинации поступают в канал, начиная с пятого такта.

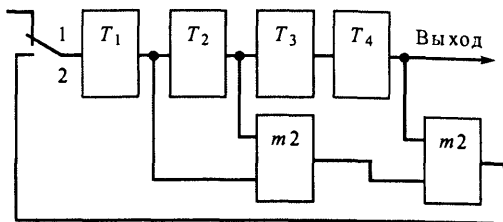


Рис. 9.26. Структурная схема кодера циклического кода, задаваемого проверочным многочленом $h(x) = x^4 \oplus x^3 \oplus x^2 \oplus 1$

Корректирующая способность кода зависит от порождающего многочлена $p(x)$. Поэтому его выбор очень важен при построении циклического кода. Необходимо помнить, что степень порождающего многочлена должна быть равна числу проверочных символов. Кроме того, многочлен $p(x)$ должен делить двучлен $x^n \oplus 1$.

Обнаружение ошибок при использовании таких кодов заключается в делении многочлена $b'(x) = b(x) + e(x)$, соответствующего принятой комбинации $\tilde{\mathbf{B}} = \mathbf{B} \oplus \mathbf{e}$, на $p(x)$. Если остаток $s(x)$ оказывается равным нулю, то считается, что ошибки нет, в противном случае фиксируется ошибка.

Пусть необходимо построить код, обнаруживающий все одиночные ошибки. В этом случае многочлен ошибок имеет вид $e(x) = x^i$, где $i = 0, 1, \dots, n - 1$. Решение задачи заключается в нахождении такого многочлена $p(x)$, чтобы многочлен $e(x)$ не делился на $p(x)$. Наиболее простым, удовлетворяющим этому требованию, является многочлен $p(x) = x \oplus 1$.

Аналогично можно построить код, обнаруживающий ошибки большей кратности.

Многочлен $s(x) = (b(x) \oplus e(x)) \bmod (p(x)) = e(x) \bmod (p(x))$ зависит только от многочлена ошибок $e(x)$ и играет ту же роль, что и вектор-синдром. Поэтому в принципе ошибки можно исправлять на основе таблицы соответствий между $e(x)$ и $s(x)$, хранящейся в памяти декодера, как при линейных нециклических кодах. Однако свойство цикличности позволяет существенно упростить процедуру декодирования.

Один из алгоритмов исправления ошибок основан на следующих свойствах синдрома циклического кода. Пусть имеется циклический код с кодовым расстоянием d , исправляющий все ошибки до кратности $l = [(d-1)/2]$ включительно, где $[(d-1)/2]$ — целая часть числа $(d-1)/2$. Тогда можно показать [133], что:

- если исправляемый вектор ошибок искажает только проверочные символы, то вес синдрома будет меньше или равен l , а сам синдром будет совпадать с вектором ошибок;

- если вектор ошибки искажает хотя бы один информационный символ, то вес синдрома будет больше l ;

- если $s(x)$ — остаток от деления многочлена $b(x)$ на $p(x)$, то остатком от деления многочлена $b(x)x^i$ на $p(x)$ является многочлен $s(x)x^i \bmod [p(x)]$, другими словами, синдром некоторого циклического сдвига многочлена $b(x)$ является соответствующим циклическим сдвигом синдрома исходного многочлена, взятого по модулю $p(x)$.

В качестве примера на рис. 9.27 представлена схема декодера для кода $(7, 4)$ с порождающим многочленом $p(x) = x^3 \oplus x^2 \oplus 1$. Код имеет кодовое расстояние $d = 3$, что позволяет ему исправлять все однократные ошибки.

Принятая кодовая комбинация одновременно поступает в буферный регистр сдвига, служащий для запоминания кодовой комбинации и ее циклического сдвига, и на устройство деления на многочлен $p(x)$ для вычисления синдрома. В исходном состоянии ключ находится в положении 1. После семи тактов буферный регистр оказывается загруженным, а в регистре устройства деления будет вычислен синдром. Если вес синдрома больше единицы, то декодер начинает проводить циклические сдвиги комбинации в буферном регистре при отсутствии новой комбинации на входе и одновременно вычислять их синдромы $s(x)x^i \bmod (p(x))$ в устройстве деления. Если

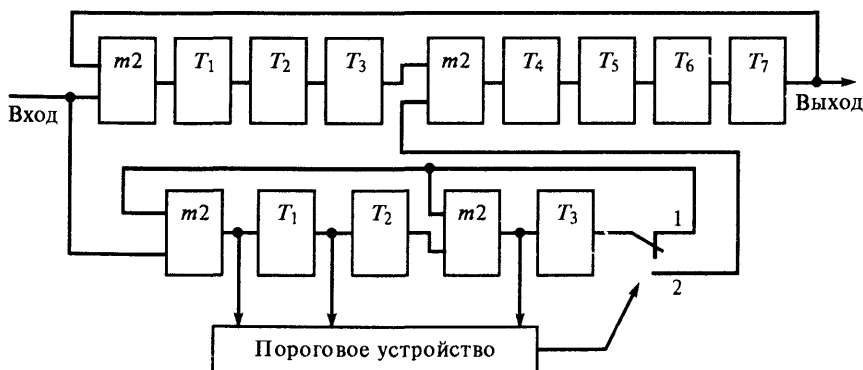


Рис. 9.27. Структурная схема декодера циклического кода с порождающим многочленом $p(x) = x^3 \oplus x^2 \oplus 1$

на некотором i -м шаге вес синдрома окажется меньше двух, то ключ переходит в положение 2, обратные связи в регистре деления разрываются. При последующих тактах ошибки исправляются путем подачи содержимого регистра деления на вход сумматора по модулю 2, включенного в буферный регистр. После семи тактов работы декодера в автономном режиме исправленная комбинация в буферном регистре возвращается в исходное положение (информационные символы будут занимать старшие разряды).

Существуют и другие, более универсальные, алгоритмы декодирования.

К циклическим кодам относятся коды Хэмминга, которые являются примерами немногих известных совершенных кодов. Они имеют кодовое расстояние $d = 3$ и исправляют все одиночные ошибки. Длина кода выбирается из условия $2^{n-k} - 1 = n$, которое имеет простой смысл: число различных ненулевых синдромов равно числу символов в кодовой последовательности. Так, существуют коды Хэмминга ($2^r - 1$, $2^r - r - 1$), в частности коды (7, 4), (15, 11), (31, 26), (63, 57) и т. д.

Заметим, что ранее использованный многочлен $p(x) = x^3 \oplus x^2 \oplus 1$ является порождающим для кода Хэмминга (7, 4).

Среди циклических кодов широкое применение нашли коды Боуза—Чоудхури—Хоквингема (БЧХ). Можно показать, что для любых целых положительных чисел m и $l < n/2$ существует двоичный код БЧХ длины $n = 2^m - 1$ с кодовым расстоянием $d \geq 2l + 1$, причем число проверочных символов $n - k \leq ml$.

Для кодов БЧХ умеренной длины и ФМ при передаче символов можно добиться значительного выигрыша (4 дБ и более) [133]. Он достигается при скоростях $1/3 \leq k/n \leq 3/4$. При очень высоких и очень низких скоростях выигрыш от кодирования существенно уменьшается.

Иногда целесообразно использовать коды с несколько худшей корректирующей способностью по сравнению с лучшими известными кодами, но простые в реализации. К ним относятся коды, допускающие *мажоритарное декодирование*. Оно основано на возможности для некоторых циклических кодов выразить каждый информационный символ с помощью Q различных линейных соотношений. Решение о значении символа принимается по мажоритарному принципу. Для исправления всех ошибок до кратности l включительно необходимо иметь $2l + 1$ независимых соотношений.

В некоторой области значений параметров мажоритарные коды имеют корректирующую способность, незначительно уступающую корректирующей способности кодов БЧХ. В то же время их реализация сравнительно проста.

Покажем принцип мажоритарного декодирования на примере кода (7, 3) с проверочной матрицей

$$\mathbf{H} = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_7 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (9.17)$$

Рассматриваемый код является циклическим с порождающим многочленом $p(x) = x^4 \oplus x^3 \oplus x^2 \oplus 1$. Он имеет кодовое расстояние $d = 4$.

Используя матрицу (9.17), можно записать следующие соотношения для символа b_1 :

$$b_1 = b_3 \oplus b_4 = b_2 \oplus b_6 = b_5 \oplus b_7. \quad (9.18)$$

С учетом (9.18) в декодере имеется возможность четырьмя разными способами вычислить первый информационный символ:

$$b_1^I = \tilde{b}_1; \quad b_1^{II} = \tilde{b}_3 \oplus \tilde{b}_4; \quad b_1^{III} = \tilde{b}_2 \oplus \tilde{b}_6; \quad b_1^{IV} = \tilde{b}_5 \oplus \tilde{b}_7, \quad (9.19)$$

где $\tilde{b}_1 \tilde{b}_2 \dots \tilde{b}_7$ — принятая кодовая комбинация.

При отсутствии ошибок $b_1^I = b_1^{II} = b_1^{III} = b_1^{IV}$, т. е. все проверочные соотношения (9.19) дают один и тот же результат. При наличии одного ошибочного символа три проверочных соотношения дают правильное значение, а соотношение, в котором участвует ошибочный символ, дает неверный результат. Принимая решение по мажоритарному принципу, декодер выдает правильный символ b_1 .

Пусть ошибочно приняты два символа. Если они входят в различные проверочные соотношения, то два проверочных соотношения дадут значение 1, а два других — значение 0. В этом случае декодер выдает сигнал отказа от декодирования. Если оба искаженных символа входят в одно проверочное соотношение, то все четыре проверки дают один и тот же результат. Декодер выдает правильный символ b_1 .

Аналогично определяются остальные информационные символы. Проверочные соотношения для символов b_2 и b_3 получаются из (9.19) циклической перестановкой:

$$\begin{aligned} b_2^I &= \tilde{b}_2; & b_2^{II} &= \tilde{b}_4 \oplus \tilde{b}_5; & b_2^{III} &= \tilde{b}_3 \oplus \tilde{b}_7; & b_2^{IV} &= \tilde{b}_6 \oplus \tilde{b}_1, \\ b_3^I &= \tilde{b}_3; & b_3^{II} &= \tilde{b}_5 \oplus \tilde{b}_6; & b_3^{III} &= \tilde{b}_4 \oplus \tilde{b}_1; & b_3^{IV} &= \tilde{b}_7 \oplus \tilde{b}_2. \end{aligned}$$

Схема декодера (рис. 9.28) состоит из сдвигающего регистра, сумматоров по модулю 2 и мажоритарного элемента М. Простота ее обусловлена

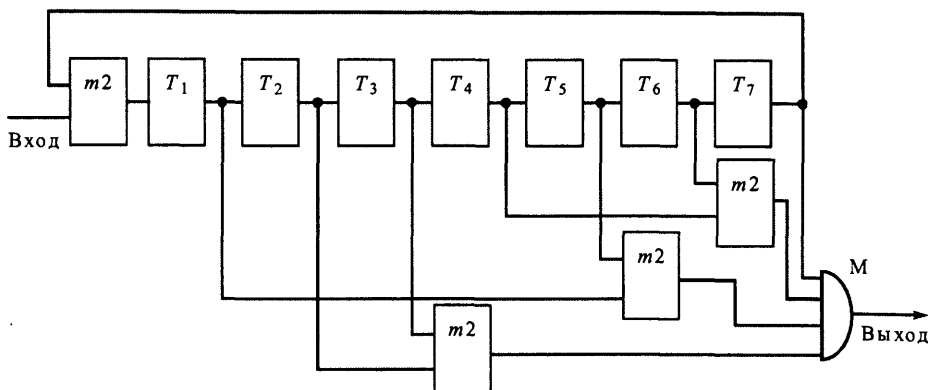


Рис. 9.28. Структурная схема декодера циклического мажоритарного кода (7, 3)

тем, что в данном случае каждый символ кодовой комбинации участвует в одном проверочном соотношении. Код, для которого выполняется это условие, называется кодом с *разделенными проверками*.

Мажоритарное декодирование возможно и тогда, когда один и тот же символ участвует в нескольких проверочных соотношениях. Однако алгоритм декодирования усложняется.

Идея построения *итеративных кодов* заключается в следующем. Информационные символы записываются в виде таблицы из k_1 столбцов и k_2 строк. К каждой строке таблицы дописываются $n_1 - k_1$ проверочных символов в соответствии с некоторым кодом (n_1, k_1) . Затем к каждому из n_1 столбцов полученной таблицы добавляют $n_2 - k_2$ проверочных символов в соответствии с некоторым кодом (n_2, k_2) . Таким образом строится код длиной $n = n_1 n_2$ с числом информационных символов $k = k_1 k_2$.

Можно показать, что для полученного двумерного итеративного кода кодовое расстояние d равно $d_1 d_2$, где d_1 и d_2 — кодовые расстояния для кодов (n_1, k_1) и (n_2, k_2) соответственно.

Кодовая комбинация двумерного итеративного кода обычно передается последовательно по строкам, начиная с первой. Соответственно, декодирование ведется сначала по строкам, а затем, после приема всего двумерного блока, — по столбцам.

Проиллюстрируем построение кодовой комбинации двумерного итеративного кода. Пусть информационные символы записаны в виде таблицы

1	1	1	0
1	0	1	0
0	1	0	0
1	1	0	1

В качестве кодов (n_1, k_1) и (n_2, k_2) будем использовать коды с проверкой на четность. Тогда кодовая комбинация будет иметь вид

1	1	1	0	1
1	0	1	0	0
0	1	0	0	1
1	1	0	1	1
1	1	0	1	1

Легко показать, что кодовое расстояние этого кода равно 4. Код исправляет все однократные ошибки. Их координаты определяются по номерам строк и столбцов, в которых не выполняется проверка на четность. Одновременно код обнаруживает все двухкратные ошибки.

Итеративные коды характеризуются большой длиной, большим кодовым расстоянием и сравнительно простой процедурой декодирования. Недостатком их является малая скорость k/n при заданной исправляющей способности.

Каскадные коды получаются комбинированием двух или более кодов и в некоторой степени похожи на итеративные. Кодирование осуществляется следующим образом [133]. Множество $k_1 k_2$ информационных символов (в дальнейшем предполагают, что они двоичные) разбивается на k_2 подблоков по k_1 символов. Каждый подблок из k_1 символов рассматривается как символ из алфавита объемом 2^{k_1} . Затем k_2 подблоков кодируются кодовыми комбинациями внешнего кода (рис. 9.29), составленными из n_2 подблоков по k_1 двоичных символов. Наконец, каждый из n_2 подблоков кодируется кодовыми комбинациями внутреннего кода (n_1, k_1) . Полученное множество n_2 кодовых слов внутреннего (n_1, k_1) -кода является кодовым словом каскадного $(n_1 n_2, k_1 k_2)$ -кода. Обычно в качестве внешнего используют код Рида—Соломона с основанием 2^{k_1} , обеспечивающий максимальное кодовое расстояние при заданных n_2 и k_1 , $n_2 < 2^{k_1}$, а в качестве внутреннего — двоичный (n_1, k_1) -код.

Декодирование осуществляется следующим образом. Сначала декодируется внутренний код. При этом получается n_2 подблоков, содержащих по k_1 символов, которые декодируются внешним кодом. В результате на выходе внешнего декодера появляются k_2 подблоков по k_2 символов.

Декодирование двумя отдельными декодерами позволяет существенно снизить сложность по сравнению с той, которая потребуется для получения

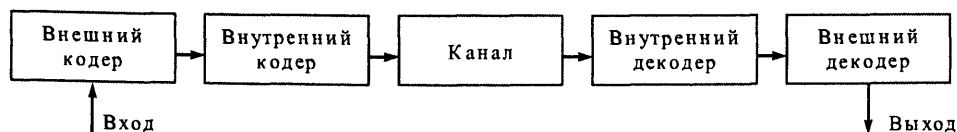


Рис. 9.29. Схема каскадного кодирования

той же вероятности ошибки при одном уровне кодирования.

Каскадные коды, как и итеративные, имеют большую длину и большое кодовое расстояние. Во многих случаях они являются наилучшими среди блочных кодов. В частности, для двоичного симметричного канала при любой скорости передачи, не превосходящей пропускной способности канала, существует каскадный код, при котором вероятность ошибки может быть сколь угодно мала.

9.5.3. Непрерывные (сверточные) коды

Сверточный код — это линейный рекуррентный код. В общем случае он образуется следующим образом. В каждый i -й тактовый момент времени на вход кодирующего устройства поступает k_0 символов сообщения $a_{i1}a_{i2}\dots a_{ik_0}$. Выходные символы $b_{i1}b_{i2}\dots b_{in_0}$ формируются с помощью рекуррентного соотношения из K символов сообщения, поступивших в данный и предшествующие тактовые моменты времени:

$$b_{im} = \sum_{v=0}^{K/K_0-1} \oplus \sum_{j=1}^{k_0} \oplus c_{jvm} a_{(i-v)j}, \quad m = 1, 2, \dots, n_0,$$

где c_{jvm} — коэффициенты, принимающие значения 0 или 1. Символы сообщения, из которых формируются выходные символы, хранятся в памяти кодирующего устройства. Величина K называется *длиной кодового ограничения*. Она показывает, на какое максимальное число выходных символов влияет данный информационный символ, и играет ту же роль, что и длина блочного кода. Сверточный код имеет избыточность $\chi = 1 - k_0/n_0$ и обозначается k_0/n_0 .

Типичные параметры сверточного кода: $k_0, n_0 = 1, 2, \dots, 8$; $k_0/n_0 = 1/4, \dots, 7/8$; $K = 3, \dots, 10$ [133].

Кодирующее устройство сверточного кода может быть реализовано с помощью сдвигающего регистра и сумматоров по модулю 2. Для схемы, показанной на рис. 9.30, на каждый символ сообщения вырабатываются два символа, которые последовательно во времени через коммутатор подаются в канал. Выходные символы являются линейными функциями поступающего информа-

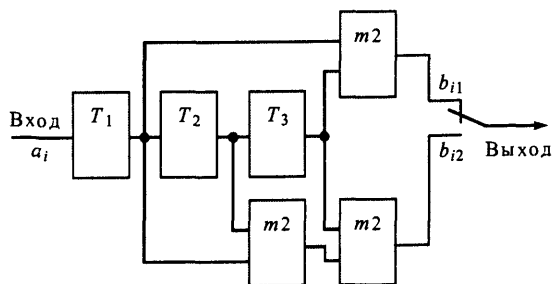


Рис. 9.30. Структурная схема кодера сверточного кода ($k_0/n_0 = 1/2, K = 3$)

ционного символа и комбинации, записанной в первых двух разрядах регистра (логического состояния регистра). Связь между ячейками сдвигающего регистра и сумматорами по модулю 2 удобно описывать порождающими многочленами $q_j(x)$, $j = 1, 2, \dots, n_0$. Для рассматриваемого случая $q_1(x) = x^2 \oplus 1$ (описывает связи верхнего сумматора) и $q_2(x) = x^2 \oplus x \oplus 1$ (описывает связи нижнего сумматора). Наличие члена x^i , $i = 0, 1, 2, \dots$, в порождающем многочлене означает, что $(i + 1)$ -й разряд регистра сдвига соединен с сумматором. Счет разрядов регистра ведется слева направо.

Сверточный код получается разделимым, если в каждый тактовый момент k_0 выходных символов совпадают с символами сообщения. На практике обычно используются несистематические сверточные коды.

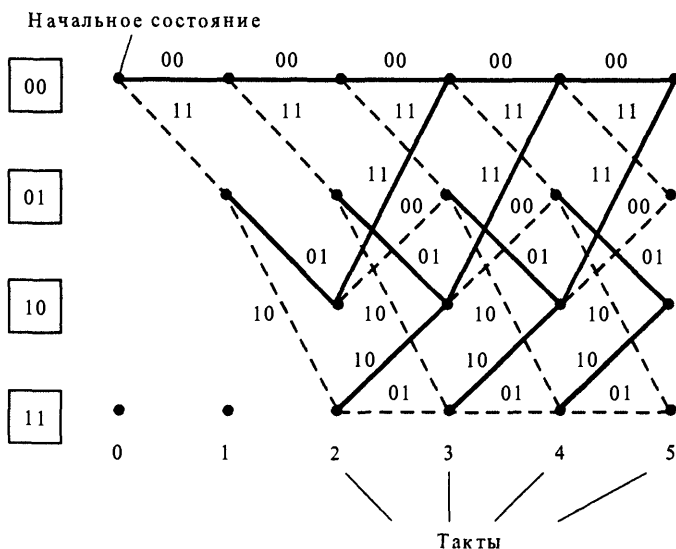
Различают *прозрачные* и *непрозрачные* сверточные коды. Первые характеризуются свойством инвариантности по отношению к операции инвертирования кода, которое заключается в следующем: если значения символов на входе кодера поменять на противоположные, то выходная последовательность символов также инвертируется. Соответственно, декодированная последовательность символов будет иметь такую же неопределенность в знаке, что и принятая последовательность символов, а следовательно, неопределенность знака последовательности можно устранить после декодирования сверточного кода (рис. 9.31). Указанное свойство прозрачных кодов особенно важно для СПИ, использующих противоположные фазоманипулированные сигналы, которым свойственно явление обратной работы.

Для непрозрачного кода неопределенность знака последовательности символов придется устранять до сверточного декодирования, что приводит к увеличению вероятности ошибок. Нетрудно показать, что сверточный код будет прозрачным, если каждый его порождающий многочлен содержит нечетное число членов.

Помимо рассмотренного способа задания сверточного кода, возможны и другие. В частности, выходные символы можно рассматривать как свертку импульсной характеристики кодера с информационной последовательностью (отсюда происходит название кода).



Рис. 9.31. Схема СПИ при использовании прозрачных сверточных кодов

Рис. 9.32. Решетчатая диаграмма для кода $1/2$ с $K=3$

Для пояснения процессов кодирования и декодирования часто используют решетчатую диаграмму, представляющую собой одно из возможных изображений кодового дерева. Такая диаграмма для кодера на рис. 9.32 состоит из узлов и ветвей (ребер). Число ветвей, исходящих из узла, равно основанию кода. Число узлов равно 2^{K-1} . Единичному символу сообщения приписываются штриховые линии, а нулевому — сплошные. Выходные символы записываются над ветвями. Надписи около узлов характеризуют логическое состояние кодирующего устройства. Каждой информационной последовательности символов соответствует определенный путь (определенная траектория) на диаграмме. Кодовая последовательность формируется путем считывания комбинаций над ветвями при прослеживании данного пути. Соответственно, процесс кодирования заключается в выборе одного из путей диаграммы.

Корректирующая способность сверточного кода зависит от так называемого *свободного расстояния* $d_{\text{св}}$, которое, по существу, содержит ту же информацию о коде, что и кодовое расстояние для блочных кодов. Оно определяется как минимальный вес (минимальное число единиц) пути на решетчатой диаграмме, начинающегося и заканчивающегося в нулевом узле. Например, для кода $k_0/n_0 = 1/2$, $K=3$ имеем $d_{\text{св}} = 5$.

В табл. 9.1 приведены порождающие многочлены оптимальных сверточных кодов с относительной скоростью передачи $1/2$ и кодовым ограничением длины $3 \dots 8$, а также значения свободных расстояний этих кодов.

Таблица 9.1

Порождающие многочлены оптимальных сверточных кодов

Длина кодового ограничения	Порождающие многочлены	Свободное расстояние
3	$q_1(x) = 1 \oplus x \oplus x^2$ $q_2(x) = 1 \oplus x^2$	5
4	$q_1(x) = 1 \oplus x \oplus x^2 \oplus x^3$ $q_2(x) = 1 \oplus x \oplus x^3$	6
5	$q_1(x) = 1 \oplus x \oplus x^2 \oplus x^4$ $q_2(x) = 1 \oplus x^3 \oplus x^4$	7
6	$q_1(x) = 1 \oplus x \oplus x^2 \oplus x^4 \oplus x^5$ $q_2(x) = 1 \oplus x \oplus x^5$	8
7	$q_1(x) = 1 \oplus x \oplus x^2 \oplus x^4 \oplus x^6$ $q_2(x) = 1 \oplus x^2 \oplus x^3 \oplus x^5 \oplus x^6$	10
8	$q_1(x) = 1 \oplus x \oplus x^2 \oplus x^3 \oplus x^4 \oplus x^7$ $q_2(x) = 1 \oplus x^2 \oplus x^5 \oplus x^6 \oplus x^7$	10

Сверточные коды можно декодировать различными методами. Различают декодирование с вычислением и без вычисления проверочной последовательности.

Декодирование с вычислением проверочной последовательности применяется только для систематических кодов. Оно ничем не отличается от соответствующего метода декодирования блочных кодов. На приемной стороне из принятых информационных символов формируют проверочные символы по тому закону, что и на передающей стороне, которые затем сравнивают с принимаемыми проверочными символами. В результате сравнения образуется проверочная последовательность, которая при отсутствии ошибок состоит из одних нулей. При наличии ошибок на определенных позициях последовательности появляются единичные символы. Закон формирования проверочных символов выбирается таким образом, чтобы по структуре проверочной последовательности можно было определить искаженные символы.

Декодирование по принципу максимума правдоподобия сводится к задаче отождествления принятой последовательности с одной из 2^N возможных, где N — длина информационной последовательности. Решение принимается в пользу той кодовой последовательности, которая в меньшем количестве позиций отличается от принятой. Метод применим для любого сверточного кода. Однако при больших значениях N он практически не реализуем из-за необходимости перебора 2^N возможных кодовых последовательностей. Существенное упрощение процедуры декодирования по максимуму правдоподобия предложил А. Витерби [13]. Характерной особенностью

его метода является то, что на каждом шаге декодирования запоминается только 2^{K-1} наиболее правдоподобных путей. Осуществляется это следующим образом. Для определенности будем рассматривать код $k_0/n_0 = 1/2$.

Пусть начальное состояние кодирующего устройства известно. Из анализа решетчатой диаграммы следует, что в любой i -й узел на любом l -м такте из начального состояния ведут несколько путей, которым соответствуют определенные кодовые последовательности. Из всех путей выберем тот, которому соответствует кодовая последовательность $B_i(l)$, отличающаяся от принятой $\tilde{B}(l)$ меньшим числом символов. Этот путь называется *выжившим*. Обозначим расстояние Хэмминга между последовательностями $B_i(l)$ и $\tilde{B}(l)$ через $d_i(l)$. Припишем i -му узлу вес, равный $d_i(l)$. Подобную процедуру проделаем для всех остальных узлов. Возьмем любой k -й узел решетчатой диаграммы в следующий тактовый момент. Он связан с двумя предшествующими узлами, например с i -м и j -м, ветвями ik и jk соответственно (см. рис. 9.32). Для нахождения правильного пути в узел k вычислим величины $d_i(l) + \Delta d_{ik}$ и $d_j(l) + \Delta d_{jk}$, где Δd_{ik} и Δd_{jk} — приращения расстояний Хэмминга при продолжении путей $B_i(l)$ и $B_j(l)$ в узел k . Эти приращения находятся по принятому на $(l+1)$ -м шаге сегменту последовательности и символам, соответствующим ветвям ik и jk . Если $d_i(l) + \Delta d_{ik} < d_j(l) + \Delta d_{jk}$, то ветвь ik считается истинной и записывается в память декодирующего устройства. Ветвь jk и все ей предшествующие отбрасываются. Аналогичные операции проделывают и для остальных узлов. В результате на $(l+1)$ -м шаге в памяти декодирующего устройства будет храниться 2^{K-1} путей.

Исследования показывают [13], что через $l_0 \approx 5K$ тактов в кодовом дереве остается лишь один путь с минимальным весом. Поэтому решение о том, какое сообщение передавалось в некоторый $(m+1)$ -й тактовый момент, можно принимать на $(m+l_0)$ -м такте. Для уменьшения объема памяти декодирующего устройства отрезки, по которым приняты решения, стираются. Для этого же из весов всех узлов кодового дерева периодически вычитают минимальный на данном такте вес.

Для пояснения работы декодера, реализующего алгоритм Витерби, рассмотрим следующий пример. Пусть используется ранее рассмотренный код $k_0/n_0 = 1/2$ при $K = 3$, $q_1(x) = 1 \oplus x^2$ и $q_2(x) = 1 \oplus x \oplus x^2$. Предположим, что передавалась нулевая последовательность, а принятая последовательность имеет вид 10 00 10 00 00 Работу декодера иллюстрируют диаграммы (рис. 9.33, a — d), где числа в узлах характеризуют значения $d_i(l)$.

На 3-м такте работы декодера (рис. 9.33, b) каждый из путей, «выживших» на предыдущем такте (рис. 9.33, a), раздваивается. Общее число путей становится равным 8. Декодер сравнивает метрики для пар путей, ведущих в каждый узел, и из каждой пары оставляет лишь лучший. Вновь

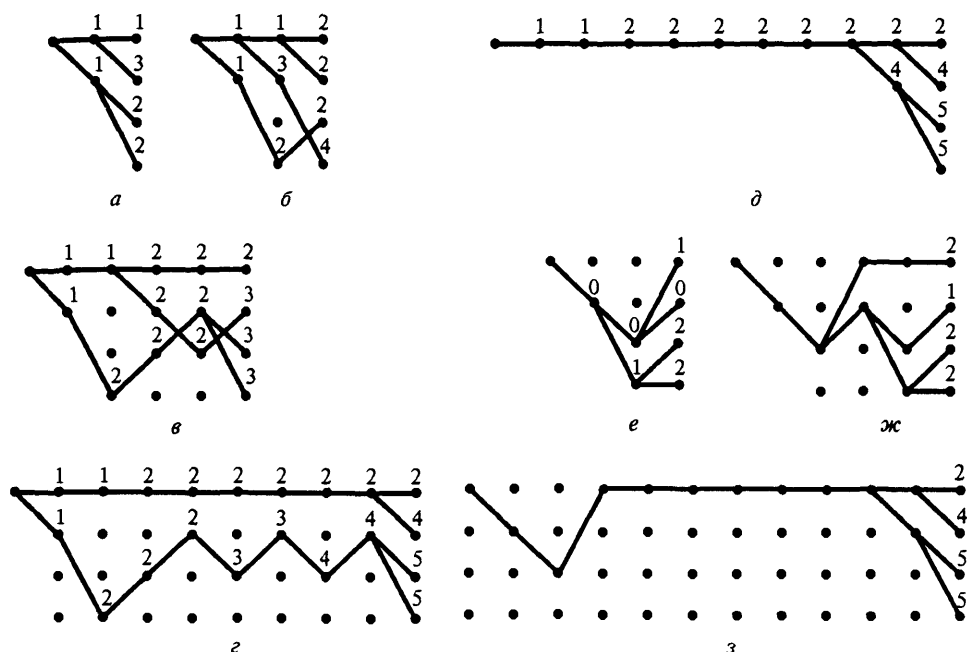


Рис. 9.33. Диаграмма декодирования исправляемой (а—д) и неисправляемой (ж—з) комбинаций ошибок с помощью алгоритма Витерби

число путей оказывается равным 4. Этот процесс повторяется при каждом приеме новой ветви (рис. 9.33, в, г, д). Заметим, что на 5-м такте (см. рис. 9.33, в) метрика нулевого пути оказывается наилучшей.

«Выживающие» пути могут отличаться друг от друга в течение многих тактов. Однако на 10-м такте (см. рис. 9.33, д) первые восемь ветвей всех «выживших» путей совпадают. В этот момент согласно алгоритму Витерби принимается решение о переданных символах, так как все «выжившие» пути приходят из одного узла.

Глубина (число тактов), на которой происходит слияние путей, является случайной величиной, зависящей от ошибок в принятой последовательности, и заранее не может быть вычислена. Поэтому обычно устанавливают некоторую фиксированную глубину декодирования. При этом каждый раз, обрабатывая новую ветвь, декодер выдает самый старый символ «выжившего» пути с наилучшей метрикой.

В рассмотренном случае принятая последовательность содержала два ошибочных символа, а код имел свободное расстояние, равное 5, что позволило ему исправить эти ошибки. Предположим, что принятая последовательность содержит 3 ошибочных символа и имеет вид 11 01 00 00 Такая комбинация ошибок оказывается *неисправляемой*. Заметим, что на 3-м

такте (рис. 9.33, е) удаляется правильный путь, что неизбежно приводит к ошибке. Кроме того, все «выжившие» пути имеют одну и ту же первую ветвь. На 5-м такте все выжившие пути имеют одинаковые первые две ветви (рис. 9.33, ж), а на 11-м такте — первые девять ветвей (рис. 9.33, з). Хотя при декодировании возникла ошибка, выбранный ошибочный путь отличается от правильного лишь на коротком отрезке, состоящем из трех ветвей. Информационная последовательность, соответствующая выбранному пути, имеет вид 1000, ..., т. е. содержит один ошибочный символ. Рассмотренный пример описывает типичное поведение ошибочных последовательностей при использовании алгоритма сверточного декодирования Витерби.

Алгоритм Витерби обладает рядом преимуществ. При небольших значениях длины кодового ограничения декодирующее устройство оказывается достаточно простым, реализуя в то же время высокую помехоустойчивость. Так, исследования показывают, что применение сверточных кодов с $K = 3, 5$ и 7 при фиксированной вероятности ошибки $P_{\text{ош}} = 10^{-5}$ позволяет получить энергетический выигрыш $4...6$ дБ по сравнению с системой, использующей ФМ сигналы без кодирования. Важным преимуществом по сравнению с методом последовательного декодирования является фиксация числа вычислительных операций на один декодированный символ. Декодирование по методу Витерби особенно перспективно в каналах с независимыми ошибками.

Декодер Витерби (рис. 9.34) состоит из синхронизатора, устройства управления и тактирования, устройства для вычисления метрик ветвей, устройства для обновления и хранения метрик ветвей, устройства для обновления и хранения гипотетических информационных последовательностей и решающего устройства.

Устройство хранения и обновления метрик путей осуществляет сложение метрик ветвей с хранящимися метриками путей, проделявает необходимые сравнения и запоминает новые метрики путей.

Устройство хранения и обновления гипотетических информационных последовательностей может быть выполнено на сдвигающих регистрах, в

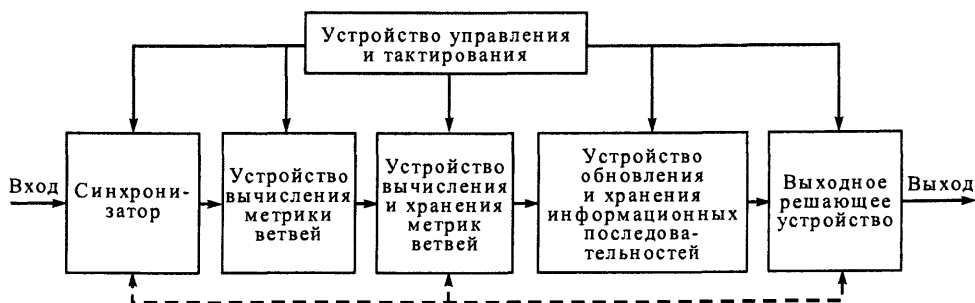


Рис. 9.34. Структурная схема декодера, реализующего алгоритм Витерби

каждом из которых хранится полная информационная последовательность символов, соответствующая одному из «выживших» путей. Их число равно числу узлов. После обработки новой ветви регистры обмениваются содержимым в соответствии с тем, какие последовательности «выживают» при сравнении. В последнюю ячейку каждого регистра поступает новый информационный символ, а самый старый символ каждого регистра поступает в выходное решающее устройство.

Выходное решающее устройство принимает решение о переданных информационных символах. Наилучшие результаты получаются, когда в качестве переданного информационного символа берется наиболее старый символ в последовательности с наименьшей метрикой. Иногда используют мажоритарный принцип: за переданный информационный символ берут чаще всего встречающийся символ из самых старых символов всех последовательностей. Устройство управления и тактирования задает необходимый ритм работы декодера.

Во многих системах, кроме основного (прямого) канала, с помощью которого сообщение передается от источника к потребителю, имеется обратный канал для вспомогательных сообщений, которые позволяют улучшить качество передачи сообщений по прямому каналу.

Наиболее распространены системы с обратной связью, в которых для обнаружения ошибок применяют избыточные коды. Такие системы называются системами с *решающей обратной связью*, или системами с *переспросом*. В качестве кодов часто используют коды с проверкой на четность, простейшие итеративные коды, циклические коды и др. Они позволяют хорошо обнаруживать ошибки при сравнительно небольшой избыточности и простой аппаратной реализации.

Передаваемое сообщение кодируется избыточным кодом. Полученная комбинация передается потребителю и одновременно запоминается в накопителе-повторителе. Принятая последовательность символов декодируется с обнаружением ошибок. Если при этом ошибки не обнаружены, то сообщение поступает потребителю. В противном случае сообщение бракуется и по обратному каналу передается специальный сигнал переспроса. По этому сигналу проводится повторная передача забракованной кодовой комбинации, которая извлекается из накопителя-повторителя.

Можно показать, что если в обратном канале ошибки отсутствуют, то остаточная вероятность ошибочного приема кодовой комбинации имеет вид

$$P_{\text{ост}} = P_{\text{н.о}} / (1 - P_{\text{о.ош}}),$$

где $P_{\text{н.о}}$ — вероятность необнаруженной ошибки (вероятность того, что переданная кодовая комбинация перешла в другую разрешенную); $P_{\text{о.ош}}$ — вероятность обнаружения ошибки (вероятность того, что вместо переданной

кодовой комбинации принята какая-либо запрещенная кодовая комбинация). Вероятности $P_{н.о}$ и $P_{о.ош}$ можно найти, если известны свойства канала и задан код. Соответственно, эквивалентная вероятность ошибки имеет вид

$$P_z \approx P_{н.о}/k(1 - P_{о.ош}),$$

где k — число информационных символов в кодовой комбинации. Среднее число передач одного сообщения определяется следующим образом:

$$Q_n = 1/(1 - P_{о.ош}).$$

Несмотря на то, что обратный канал можно сделать весьма помехоустойчивым (обычно скорость передачи информации в обратном канале значительно меньше, чем в прямом), тем не менее, существует конечная вероятность того, что сигнал переспроса будет принят как сигнал подтверждения, и наоборот. В первом случае сообщение не поступает потребителю, а во втором случае оно поступает дважды. Одним из средств борьбы с ошибками в обратном канале, приводящими к потере сообщения, является использование несимметричного правила декодирования, при котором вероятность ошибочного приема сигнала переспроса существенно меньше вероятности ошибочного приема сигнала подтверждения. Например, сигнал переспроса передается кодовой комбинацией из n единичных символов, а сигнал подтверждения — комбинацией из n нулей. При приеме кодовой комбинации, содержащей хотя бы одну единицу, решение принимается в пользу сигнала переспроса. Очевидно, что в этом случае вероятность ошибочного приема сигнала переспроса можно получит сколь угодно малой.

Для того чтобы к потребителю не поступали лишние сообщения, обусловленные ошибочным приемом сигналов подтверждения, передаваемые кодовые комбинации либо снабжаются номерами, либо дополняются опознавательными символами, по которым можно узнать, передается кодовая комбинация в первый раз или она повторяется. При этом принятая повторная комбинация при отсутствии сигнала переспроса стирается и не поступает потребителю. Возможны и другие способы борьбы с ошибками такого рода.

Системы с решающей обратной связью весьма эффективны в случае каналов с замираниями. При ухудшении состояния канала увеличивается частота переспроса (уменьшается скорость передачи информации), но вероятность ошибочных сообщений, поступающих потребителю, практически не увеличивается. При улучшении состояния канала частота переспроса уменьшается. Таким образом, система как бы автоматически приспосабливается к состоянию канала связи, используя все его возможности в отношении передачи информации.

Следует заметить, что применение решающей обратной связи, конечно, не увеличивает пропускной способности прямого канала, но позволяет

более простыми средствами по сравнению с длинными кодами приблизить скорость передачи информации к пропускной способности канала.

9.5.4. Сигнально-кодовые конструкции. Прием сигналов в целом

Как известно [131], многопозиционные сигналы, такие как сигналы многократной ФМ и сигналы АФМ, обеспечивают высокую удельную скорость передачи информации (высокую частотную эффективность) при уменьшении энергетической эффективности, а помехоустойчивые коды позволяют повышать энергетическую эффективность при снижении удельной скорости передачи. Сочетание методов многопозиционной модуляции и помехоустойчивого кодирования дает возможность повысить либо энергетическую эффективность без уменьшения частотной, либо частотную эффективность без снижения энергетической, а в ряде случаев — оба параметра. Задача заключается в формировании таких сигнальных последовательностей, которые можно достаточно плотно разместить в многомерном пространстве (для обеспечения высокой частотной эффективности) и в то же время разнести на достаточно большие расстояния (для обеспечения высокой энергетической эффективности). Такие последовательности, построенные на базе помехоустойчивых кодов и многопозиционных сигналов с плотной упаковкой, называются *сигнально-кодовыми конструкциями*.

В качестве помехоустойчивого кода обычно используются каскадные, итеративные и сверточные коды, а в качестве многопозиционных сигналов — сигналы многократной ФМ и сигналы АФМ.

Для согласования кода двоичного помехоустойчивого кода и модема многопозиционных сигналов используется манипуляционный код, при котором большему расстоянию по Хэммингу между кодовыми комбинациями соответствует большее расстояние между соответствующими им сигналами. Этому требованию частично удовлетворяет код Грея. Возможны и другие способы такого преобразования.

На рис. 9.35 показана структурная схема одной из возможных систем с многоуровневой ФМ и помехоустойчивым кодированием. Сформированные на выходе помехоустойчивого кодера комбинации преобразуются в коде Грея в последовательность кодовых комбинаций длины m , которые и определяют начальную фазу радиоимпульса фиксированной длительности на выходе фазового модулятора. На приемной стороне принятый сигнал сначала синхронно детектируется фазовым модулятором. Полученная при этом последовательность символов преобразуется декодерами Грея и помехоустойчивого кода в сообщение.

Применение сигнально-кодовых конструкций позволяет существенно приблизиться к границе эффективности, определяемой пропускной способностью канала.

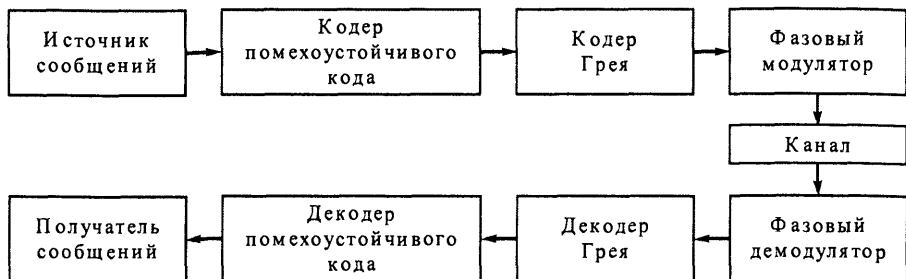


Рис. 9.35. Структурная схема СПИ при использовании сигнально-кодовых конструкций

До сих пор предполагалось, что кодовые комбинации принимаются посимвольно, т. е. на приемной стороне вначале выносятся решение о каждом символе кодовой комбинации, а затем по совокупности n принятых символов принимается решение о том, какая кодовая комбинация была передана.

При избыточных кодах такая двухэтапная процедура принятия решения оказывается неоптимальной. Объясняется это тем, что процесс демодуляции является необратимой операцией и может сопровождаться потерей информации. Действительно, после принятия решения о символе ни соответствующий элемент сигнала, ни фактическое значение результата обработки этого символа (значение апостериорной вероятности или функции правдоподобия) в дальнейшем процессе приема (при декодировании) не принимаются во внимание. В то же время их учет мог бы привести к уменьшению вероятности ошибочного декодирования кодовой комбинации.

Вся информация, содержащаяся в принимаемом сигнале, будет наиболее полно использована, если отказаться от посимвольного приема и демодулировать кодовую комбинацию в целом.

Для идеализированного двоичного канала с постоянными параметрами и помехой типа гауссовского белого шума оптимальный алгоритм «приема в целом» совпадает с алгоритмом (3.42), если все 2^k возможных сигналов в точности известны, или с алгоритмом (3.15), если начальная фаза сигнала, соответствующего кодовой комбинации, случайна, но сохраняется в процессе приема всей кодовой комбинации. При этом сигнал $s_i(t)$ понимают как сигнал, соответствующий всей кодовой комбинации.

Можно показать, что при использовании кода с избыточностью помехоустойчивость «приема в целом» выше помехоустойчивости поэлементного приема с исправлением ошибок, однако уступает помехоустойчивости поэлементного приема с обнаружением ошибок и переспросом по обратному каналу. При использовании кода без избыточности «прием в целом» не имеет преимуществ по сравнению с поэлементным приемом.

В общем случае вычислить вероятность ошибочного приема кодовой комбинации трудно. Однако иногда, например при использовании ортогональных, биортогональных и симплексных кодов, эту вероятность можно выразить через интегралы, которые можно определить численными методами.

Недостатком «приема в целом» является то, что он требует значительно более сложной аппаратуры по сравнению с поэлементным приемом. В частности, для его реализации требуется 2^k корреляторов. Очевидно, что при достаточно эффективном коде (такой код является длинным) «прием в целом» технически нереализуем. Например, если используется (n, k) -код с $k = 10$, то демодулятор, реализующий «прием в целом», будет состоять из 1024 корреляторов или согласованных фильтров.

В связи с трудностями построения оптимального демодулятора для «приема в целом» большое внимание уделяется алгоритмам приема, которые не используют всю информацию о принятом сигнале, но допускают меньшие потери по сравнению с поэлементным приемом. Такие алгоритмы являются двухэтапными, как и при поэлементном приеме. Однако на первом этапе решение о переданном символе не принимается, а запоминаются значения напряжений на выходах корреляторов или согласованных фильтров, предназначенных для приема различных символов, из которых составляются кодовые комбинации. Такой вид решения называется «мягким». Как известно, эти напряжения пропорциональны логарифму функций правдоподобия и несут информацию о степени соответствия принятого сигнала тому или иному символу. Их использование при дальнейшей обработке (декодировании) и позволяет получить лучшие результаты по сравнению с поэлементным приемом.

В реальных системах выходные напряжения обычно квантуются и представляются числами, т. е. вместо оптимального аналогового декодирования по максимуму правдоподобия используют цифровое декодирование. Цифровое декодирование уже при восьми уровнях квантования практически дает такие же результаты, что и аналоговое декодирование [133]. В то же время оно значительно проще в реализации.

Существуют и другие методы приема, занимающие промежуточное положение между поэлементным приемом и «приемом в целом», например прием по наиболее надежным символам. В его основу положен тот факт, что при применении кода с кодовым расстоянием d любую его комбинацию можно декодировать, если «стереть» $d-1$ символ. Устройство приема состоит из двух решающих схем. Первая из них вычисляет апостериорные вероятности и принимает предварительно решение о переданном символе. Полученная последовательность символов подается на вторую решающую схему, куда также поступает информация об апостериорных вероятностях.

Декодирование выполняется по $n - d + 1$ наиболее надежным (имеющим большие значения апостериорной вероятности) символам.

Описанный метод дает лучшие результаты, чем поэлементный прием, так как в нем частично используется информация об апостериорных вероятностях, но уступает приему в целом, так как информация о $d - 1$ менее надежных символах не используется.

9.6. Многоканальные и многоадресные системы

9.6.1. Принципы многостанционного доступа

Существующие в настоящее время и проектируемые СПИ должны обеспечивать одновременной связью большое число стационарных и подвижных объектов, произвольно расположенных на некоторой территории. Вследствие этого перспективными являются *многостанционные* СПИ. В таких системах необходимо осуществлять многостанционный доступ (МСД) в общий частотный канал, при котором корреспонденты передают (и принимают) независимо друг от друга информацию тогда, когда в этом возникает необходимость. Многостанционные СПИ играют основную роль при построении систем связи с подвижными объектами (самолетами, кораблями, автомобилями). Многостанционный доступ в общий частотный канал является наиболее целесообразным методом построения спутниковой системы связи — важнейшего звена Единой автоматизированной сети связи. В состав таких систем (рис. 9.36) входит обычно l корреспондентов, каждый из которых является источником дискретной или непрерывной информации (ИИ _{i} , где $i = 1, 2, \dots, l$). Сообщение каждого корреспондента в передатчике П преобразуется в сигнал $s_i(t)$. Однако системы МСД имеют ряд существенных отличий от многоканальных систем. Так, групповой сигнал $s_{\Sigma}(t)$ образуется в результате сложения радиосигналов корреспондентов непосредственно в канале, отсутствует временная синхронизация источников информации, уровни принимаемых сигналов могут существенно отличаться, например в силу разных протяженностей трасс распространения.

Наряду с помехами канала в системах МСД действуют специфические для этих систем искажения, связанные с влиянием сигналов корреспондентов друг на друга при выделении их из группового сигнала (*межстанционные помехи*). Уменьшение этого влияния и соответственно ослабление искажений передаваемых сообщений могут быть достигнуты правильным выбором сигналов $s_i(t)$ и методов их выделения из группового сигнала.

В системах МСД все сигналы $s_j(t)$, $j = 1, 2, \dots, k$, называемые *адресными*, могут или заранее распределяться и закрепляться за конкретными

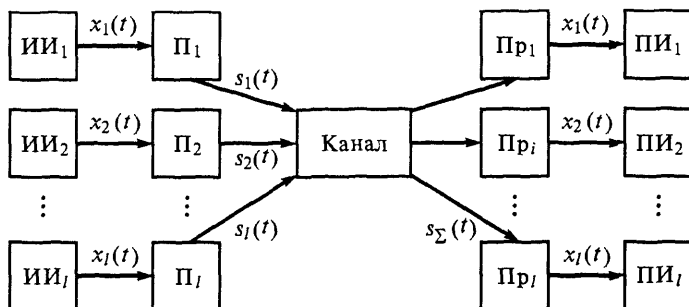


Рис. 9.36. Структурная схема многостанционной РСПИ

корреспондентами, или выделяться только на время сеанса связи, после которого эти сигналы используются другими корреспондентами системы. Метод распределения сигналов между корреспондентами определяется взаимодействием станций в системе МСД и активностью корреспондентов.

Активность корреспондента в системе МСД характеризуется вероятностью p_a передачи им информации в некоторый момент времени. В зависимости от типа системы значение p_a может меняться в широких пределах, однако в большинстве случаев $p_a \ll 1$. Вероятность того, что в данный момент времени активными являются l_a корреспондентов, определяется биномиальным законом, который записывается в следующем виде:

$$p(l_a) = C_{l_a \max}^{l_a} p_a^{l_a} (1 - p_a)^{l_a \max - l_a}, \quad (9.20)$$

где биномиальный коэффициент $C_{l_a \max}^{l_a}$ — число сочетаний из $l_a \max$ по l_a ; $l_a \max$ — общее число корреспондентов.

Среднее значение числа активных корреспондентов с учетом распределения (9.20) будет

$$\bar{l}_a = p_a l_a \max,$$

а дисперсия

$$D_a = p_a (1 - p_a) l_a \max.$$

При $p_a \ll 1$ дисперсия числа активных корреспондентов будет $D_a \approx p_a l_a \max$, а отношение $\sqrt{D_a} / \bar{l}_a = 1 / \sqrt{p_a l_a \max}$.

По характеру организации совместной работы станций различают системы МСД с *ограниченным доступом* (контролируемые) и со *свободным доступом* корреспондентов в общий частотный канал (неконтролируемые).

В неконтролируемых системах МСД адресные сигналы $s_j(t)$ жестко закрепляются за определенными корреспондентами. Это обеспечивает воз-

возможность связи каждой пары корреспондентов независимо от других. В таких системах число сигналов k примерно равно общему числу корреспондентов.

В контролируемых системах МСД сигналы не закрепляются жестко за корреспондентами, а выделяются им по мере необходимости выхода на связь. Число сигналов в контролируемой системе МСД может быть гораздо меньше общего числа корреспондентов, обслуживаемых системой: $k \ll l_{a \max}$. Объясняется это тем, что в контролируемых системах учитывается статистика работы отдельных корреспондентов, т. е. тот факт, что корреспонденты в системе передают информацию не непрерывно и число активных корреспондентов l_a обычно значительно меньше $l_{a \max}$. В контролируемой системе МСД можно выбирать $k \leq l_a$, т. е. существенно меньшим, чем в неконтролируемой системе МСД. При больших $l_{a \max}$ эта особенность контролируемых систем МСД, обеспечивающая упрощение выбора систем адресных сигналов $s_j(t)$, $j = 1, 2, \dots, k$, и уменьшение влияния сигналов корреспондентов друг на друга, делает такие системы предпочтительнее неконтролируемых. Однако необходимость вводить в систему устройство контроля за состоянием загрузки канала и распределения свободных адресных сигналов между корреспондентами значительно усложняет реализацию и снижает надежность работы системы. Кроме того, существует вероятность, что число корреспондентов, которым в данный момент необходимо передать информацию (потенциальных активных корреспондентов), окажется больше k , и некоторые корреспонденты будут ожидать очереди, что в ряде случаев оказывается недопустимым.

Корреспонденты системы МСД могут вести обмен информацией либо непосредственно друг с другом — прямое объединение корреспондентов в сеть (рис. 9.37, а), либо через центральную станцию (ЦС), где собираются сигналы всех корреспондентов сети. Связь между корреспондентами осуществляется по радиусам от корреспондента K_i на ЦС и от ЦС к другому корреспонденту K_j — радиальное объединение (рис. 9.37, б).

Преимуществом систем МСД с радиальным объединением в сеть является простота реализации контролируемых систем. В таких системах обязательно должна иметься ЦС, снабжающая корреспондентов информацией о загрузке канала и выделяющая им свободные

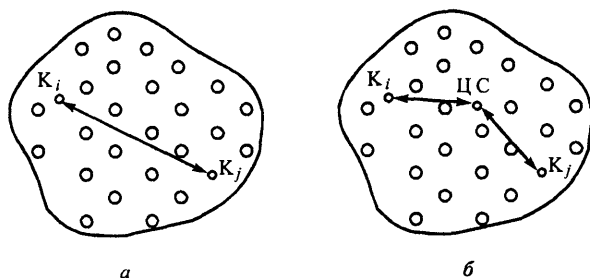


Рис. 9.37. Способы объединения корреспондентов в сеть многостанционных РСПИ

адресные сигналы. Контролируемые системы МСД с радиальным объединением корреспондентов имеют меньшую надежность, чем системы с прямым объединением, так как выход из строя ЦС приводит к потере работоспособности всей системы.

Уплотнение общего частотного канала и разделение сигналов корреспондентов на приемной стороне основываются на теории линейного разделения сигналов. Поэтому в системах МСД, как и в многоканальных системах, применяется частотное, временное и кодовое уплотнение.

Системы МСД с временным уплотнением требуют синхронизации по времени. Поскольку корреспонденты расположены произвольно на определенной территории или даже перемещаются в пространстве, требования к устройствам синхронизации в системах МСД выше, а сами эти устройства гораздо сложнее, чем в многоканальных системах.

Кодовое уплотнение позволяет создавать как синхронные, так и асинхронные системы МСД. Достоинством синхронных систем является возможность достижения полной ортогональности адресных сигналов. В асинхронных системах не требуется синхронизация по времени между сигналами корреспондентов. Однако при асинхронной работе передатчиков в каждом приемнике при разделении сигналов возникают межстанционные помехи, что является основным недостатком данных систем. Этот недостаток не снижает того интереса, который проявляется к подобным системам в связи с возможностью не зависимой друг от друга во времени работы корреспондентов. Поскольку форма каждого сигнала является адресом корреспондента, которому предназначена заключенная в этом сигнале информация, такие системы называют *асинхронными адресными* (ААС).

9.6.2. Системы с временным разделением

В системах многостанционного доступа с временным разделением (МДВР) каждый корреспондент передает (или принимает) информацию в течение специально для него отведенных интервалов времени. Метод МДВР широко распространен в спутниковых системах связи, представляющих собой сети с радиальным объединением корреспондентов. Ретранслятор на спутнике поочередно предоставляется для передачи сигналов каждой земной станции системы МСД. Чтобы исключить наложение сигналов различных станций друг на друга из-за ошибок временной синхронизации, между ними предусматриваются *защитные* временные интервалы. Принцип формирования группового сигнала на входе ретранслятора (РТР) при работе передатчиков трех земных станций (ЗС) поясняется на рис. 9.38, где τ_i — время, в течение которого каждая ЗС излучает свой сигнал; T_i — период

следования этих сигналов; τ_3 — защитный интервал; τ_c — длительность сигнала, обеспечивающего синхронизацию в системе МДВР.

Важной характеристикой системы МДВР является эффективность использования ретранслятора по времени:

$$\eta_{РТР} = 1 - \frac{\tau_c + (l_a + 1)\tau_a}{T_i},$$

где l_a — число ЗС, сигналы которых передаются за интервал времени T_i (число активных корреспондентов).

Чем больше $\eta_{РТР}$, тем лучше используется ретранслятор и тем совершеннее построена система МДВР, поскольку основная часть времени расходуется в ней на передачу полезной информации. Если принять, что величина τ_c , определяющая надежность синхронизации в системе МДВР, не может быть изменена, то для увеличения $\eta_{РТР}$ необходимо уменьшать τ_3 или увеличивать T_i . Это может быть достигнуто двумя путями: совершенствованием методов построения системы синхронизации при МДВР и использованием передачи информации с изменением масштаба времени.

В отличие от многоканальных СПИ с временным уплотнением в системах МДВР групповой сигнал образуется только на входе РТР. Следовательно, момент включения передатчика каждой ЗС должен определяться на основе точного знания как расстояния от этой станции до спутника, так и параметров движения спутника. Все это требует в системе МДВР высокоточной синхронизации всех ЗС.

Существует целый ряд вариантов построения систем синхронизации при МДВР, отличающихся способом обмена информацией для установле-

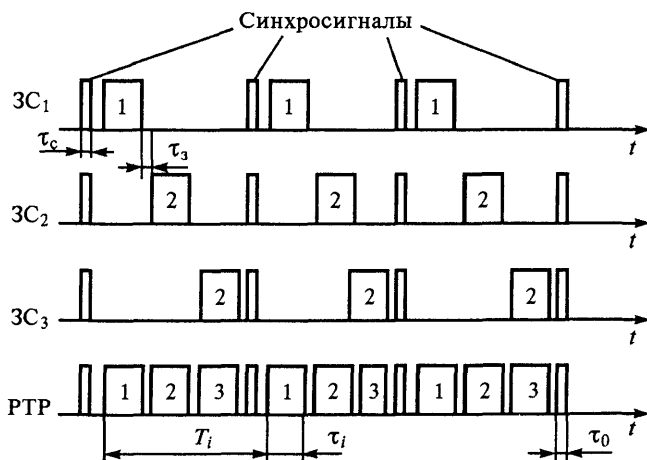


Рис. 9.38. Диаграммы формирования группового сигнала в системе МДВР

ния и поддержания синхронизации и сложностью аппаратурной реализации. Простейшим является программный метод синхронизации в сочетании с односторонним способом обмена синхросигнацией, при котором применяются активный РТР и пассивные в режиме синхронизации ЗС. Суть метода заключается в следующем. Ретранслятор излучает последовательность синхросигналов, которые принимаются всеми ЗС. На ЭВМ, входящей в состав каждой ЗС, рассчитывается задержка излучения своего информационного сигнала относительно синхронизирующего, принятого от РТР, с учетом запаздывания сигнала при его распространении. Время задержки должно быть таким, чтобы переданный земной станцией сигнал попал на временную позицию, выделенную ему в составе группового сигнала, образуемого на входе РТР.

Рассмотренный метод иллюстрируется временными диаграммами, изображенными на рис. 9.39. Пусть система МДВР работает таким образом, что сигнал ЗС₁ в адрес ЗС₂ должен приходить на шестую временную позицию после принятого ею синхросигнала РТР. Необходимо определить момент начала передачи ЗС₁ в адрес ЗС₂. Из рис. 9.39, а следует, что этот сигнал должен быть переизлучен РТР также спустя шесть временных интервалов T_i после синхросигнала. Из рис. 9.39, б видно, что если задержка в

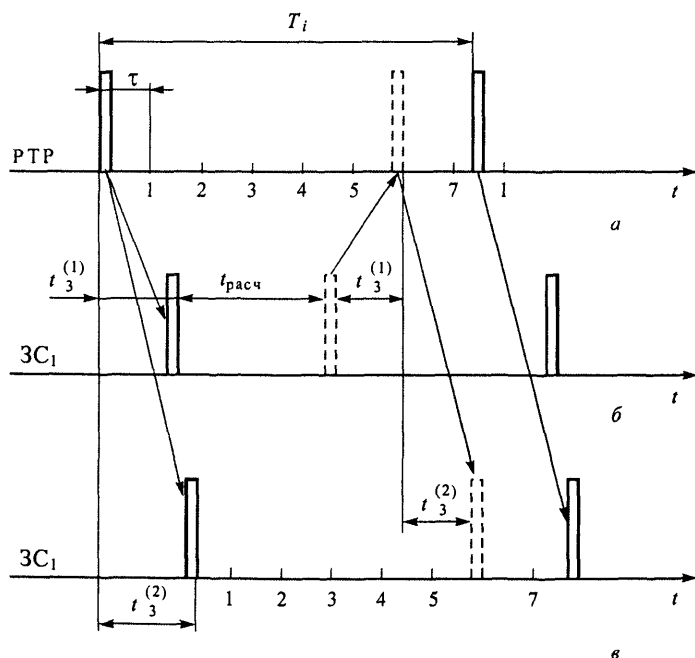


Рис. 9.39. Диаграмма, поясняющая программный метод синхронизации при МДВР

распространении сигнала от РТР к ЗС₁ составляет $t_3^{(1)}$, то после прихода синхросигнала к ЗС₁ ее информационный сигнал должен излучаться через время $t_{\text{расч}} = 6\tau_i - t_3^{(1)}$. В этом случае, как следует из рис. 9.39, в, независимо от времени задержки при распространении от РТР до ЗС₂ приходящий в ее адрес сигнал будет находиться на шестой временной позиции.

Величина $t_{\text{расч}}$ определяется с помощью ЭВМ ЗС₁ по известным T_i и τ_i и прогнозируемому (с учетом заложенных в блоке памяти ЭВМ параметров орбиты спутника) значению $t_{3.\text{пр}}^{(1)}$.

Существенным недостатком программного метода синхронизации является его малая точность. Это объясняется тем, что даже для геостационарных спутников под воздействием целого ряда случайных возмущающих факторов имеет место довольно значительное отклонение параметров орбиты от прогнозируемых. Для грубых расчетов можно полагать, что корректируемый геостационарный спутник может находиться в области пространства размером $25 \times 25 \times 75$ км с диагональю 83 км. Следовательно, неопределенность априорной информации о времени запаздывания сигнала составляет около 300 мкс. Это приводит к необходимости периодической коррекции данных о параметрах орбиты на всех ЗС системы МДВР, а также требует увеличения защитных интервалов между ретранслируемыми сигналами, что снижает эффективность использования РТР по времени (уменьшает $\eta_{\text{РТР}}$).

9.6.3. Системы с частотным разделением

В системах многостанционного доступа с частотным разделением (МДЧР) сигналу каждого корреспондента предоставляется отдельная полоса частот. Число этих полос определяется шириной общей полосы частот, выделяемой системе МДЧР. При таком методе все сигналы корреспондентов имеют одинаковую форму и могут передаваться одновременно и непрерывно. Значения несущих частот передатчиков станций в системе выбирают так, чтобы между спектрами отдельных сигналов оставались защитные интервалы для уменьшения межстанционных помех.

Для организации связи в системе МДЧР может использоваться так называемый метод *приемной волны*. Это означает, что каждому приемнику присваивается определенная несущая частота (волна). Передатчики перестраиваются по всему диапазону в зависимости от номера частоты корреспондента, с которым они хотят связаться.

Основными достоинствами систем МДЧР являются: простота реализации и возможность совместимости с существующими РСПИ, а также отсутствие необходимости синхронизации работы станций, входящих в сис-

тому. Эти достоинства способствуют широкому распространению МДЧР в системах многостанционного доступа, поскольку позволяют применять при построении систем имеющийся парк радиостанций практически без существенных изменений или использовать при разработках готовые технические решения. В результате значительно ускоряется процесс создания систем и снижается их стоимость.

Однако системе МДЧР присущ целый ряд существенных недостатков. К ним относятся:

- плохое использование частотного диапазона, отведенного системе, при малой активности отдельных корреспондентов;

- уменьшение числа возможных рабочих частот в отведенном диапазоне, связанное с необходимостью введения защитных частотных интервалов между соседними сигналами;

- трудности обеспечения одновременной работы нескольких близко расположенных станций без значительных взаимных помех даже при наличии защитных частотных интервалов.

Эти недостатки можно частично устранить при организации многостанционного доступа с использованием ЦС. В этом случае можно перейти к выделению частот по требованиям, поступающим на ЦС от радиостанций, выходящих на связь, и увеличить тем самым число обслуживаемых корреспондентов при том же количестве имеющихся рабочих частот. Кроме того, можно применить регулировку мощности передатчиков радиостанций по командам, поступающим от ЦС, для приблизительного уравнивания сигналов на входах приемников корреспондентов и уменьшения таким образом межстанционных помех в системе.

Однако использование ЦС в ряде случаев делает систему многостанционного доступа менее гибкой, так как необходима постоянная привязка работы всех корреспондентов к ЦС.

9.6.4. Асинхронные адресные системы

Создание ААС связано с применением в качестве адресных сигналов системы сложных сигналов с базой $B \gg 1$, обладающих свойством делимости при взаимном наложении в частотно-временной области. Структура таких сложных сигналов определяет при прочих равных условиях (ширине спектра сигналов, числе активных корреспондентов l_a , скорости передаваемой информации, расположении станций на местности) уровень и характер межстанционных помех и поэтому может быть положена в основу классификации ААС. По этому признаку обычно различают системы с частотно-временным кодированием (ЧВК) и с фазо-кодовой модуляцией (ФКМ).

Асинхронные адресные системы можно использовать для передачи как непрерывной, так и дискретной информации, причем в первом случае применяют различные виды модуляции как импульсной (ВИМ, ИКМ, дельта-модуляция и др.), так и аналоговой (частотная, фазовая).

В ААС с частотно-временным кодированием для разделения сигналов корреспондентов можно использовать время-интервальные и частотно-временные адресные коды. В первом случае коды различных адресов отличаются друг от друга интервалами между импульсами (рис. 9.40, а). Во втором случае дополнительным признаком кодообразования является частота заполнения импульсов. Данный код удобно изображать на частотно-временной плоскости в виде частотно-временной матрицы (ЧВМ).

Частотно-временная матрица (рис. 9.40, б) имеет размер F_{Σ} по частоте и T_c по длительности, где F_{Σ} определяется полосой, выделяемой для работы системы, а T_c — длительностью кодируемых двоичных символов. Временной интервал T_c разбивается на N дискретных интервалов, а полоса F_{Σ} — на M частотных неперекрывающихся подканалов. Длительность каждого дискретного интервала — $T_0 = T_c/N$, полоса частотного подканала — $F_0 = F_{\Sigma}/M$.

При образовании адресных кодовых комбинаций любая комбинация состоит из n импульсов, расположенных на различных дискретных интервалах, каждый из которых передается на одной из m неповторяющихся частот, соответствующих частотным подканалам матрицы. База такого сигнала при $F_0 \approx 1/T_0$

$$B_{\text{ЧВК}} = mF_0nT_0 \approx mn.$$

Максимальная база адресного сигнала

$$B_{\text{ЧВК max}} = F_{\Sigma}T_c = MF_0NT_0 \approx MN$$

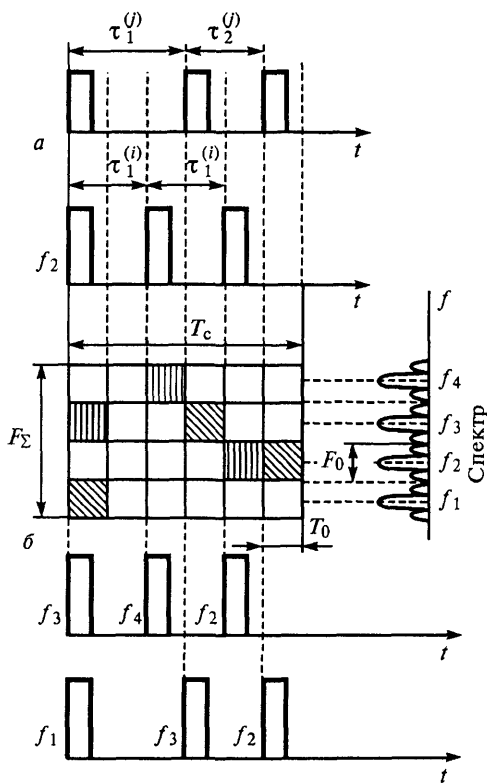


Рис. 9.40. Структура сигналов в ААС с частотно-временным кодированием при использовании:

а — время-интервального кода; б — частотно-временного кода

или при $M = N$ (квадратная ЧВМ)

$$B_{\text{ЧВК max}} = N^2.$$

Как видно из рис. 9.40, время-интервальные коды можно рассматривать как частный случай частотно-временных адресных кодов при $M = 1$.

Общее возможное число адресов (кодových комбинаций) для случая, когда имеется только один частотный подканал ($M = 1$), вычисляется по формуле

$$k^{(1)} = C_{N-1}^{n-1} = \frac{(N-1)!}{(n-1)!(N-n)!}.$$

При этом считается, что первый импульс всегда фиксирован на первом дискретном интервале ЧВМ.

При частотно-временном кодировании ($M > 1$) общее число кодových комбинаций определяется следующим образом:

$$k^{(1)} = C_M^n \sum_{i=0}^{n-1} C_n^i (N-1), \quad (9.21)$$

где $C_M^n = M!/(n!(M-n)!)$ — число сочетаний из M частотных подканалов по числу импульсов в группе n ; C_n^i — число сочетаний из числа импульсов в группе по числу возможных положений на временной оси матрицы.

Формула (9.21) справедлива в предположении, что все импульсы данной группы имеют различное частотное заполнение. Из всей совокупности адресов, определяемой (9.21), для практических целей наиболее удобны лишь те, у которых все временные интервалы между импульсами с одинаковыми частотами заполнения отличаются друг от друга для всех сигналов, используемых в системе. Такие адреса называют *рациональными*. Для частотно-временного кодирования ($M > 1$) число рациональных адресов

$$k_{\text{рац}} < \frac{2N-1}{C_n^2} C_M^2.$$

При использовании рациональных кодов в качестве адресных уменьшается вероятность перехода одной адресной комбинации в другую из-за действия межстанционных помех. Действительно, при работе l_a корреспондентов на вход каждого приемного устройства одновременно поступает несколько независимых последовательностей радиоимпульсов. В этой совокупности импульсов только небольшая часть, а именно те, которые соответствуют коду данного корреспондента, являются рабочими. Остальные импульсы от $l_a - 1$ корреспондентов являются мешающими и формируют межстанционную помеху. При большом числе корреспондентов всегда воз-

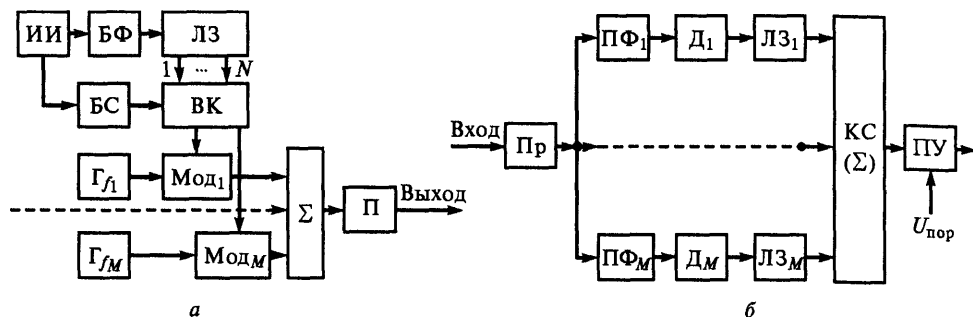


Рис. 9.41. Структурные схемы устройств формирования (а) и обработки (б) частотно-временных сигналов в ААС

можны такие комбинации мешающих импульсов, которые в случайном сочетании могут образовать ложный адресный код данного корреспондента. Вероятность его образования тем меньше, чем в большем числе параметров отличаются друг от друга используемые адресные коды. Именно рациональные адреса обеспечивают наилучшим образом выполнение этого условия.

Оптимальная обработка адресных кодовых последовательностей на фоне флуктуационных шумов должна проводиться согласованным фильтром. Создать такой фильтр сложно, поэтому обычно в реальных приемниках его заменяют близким к оптимальному устройством, так называемым *дешифратором*. Один из применяющихся методов формирования и обработки адресных частотно-временных сигналов в ААС представлен структурной схемой на рис. 9.41, где изображены передающая (рис. 9.41, а) и приемная (рис. 9.41, б) части аппаратуры.

Дискретная информация от источника информации (ИИ) поступает на блок формирования импульсов (БФ) и блок синхронизации (БС), синхронизирующий работу ААС. Видеоимпульсы с БФ подаются на линию задержки (ЛЗ) с N отводами через интервалы T_0 . Видеоимпульсы с отводов ЛЗ через временной коммутатор (ВК) идут на модуляторы (Мод), где заполняются сигналами генераторов (Γ) с частотами f_1, \dots, f_M в соответствии с выбранным для корреспондента кодом. С выходов модуляторов радиоимпульсы через сумматор (Σ) поступают в передатчик (П), в котором осуществляется перенос их спектров в область частот, отведенную для группового сигнала, и излучение. Пройдя по каналу, адресный сигнал попадает в приемное устройство (Пр), обрабатывается по высокой частоте, а затем поступает на полосовые фильтры (ПФ), каждый из которых настроен на одну из центральных частот радиоимпульсов кодовой комбинации. Напряжения с выходов фильтров через демодуляторы (Д) и линии задержки (ЛЗ) подаются на входы или каскада совпадения (КС), или сумматора (Σ). Решение о наличии

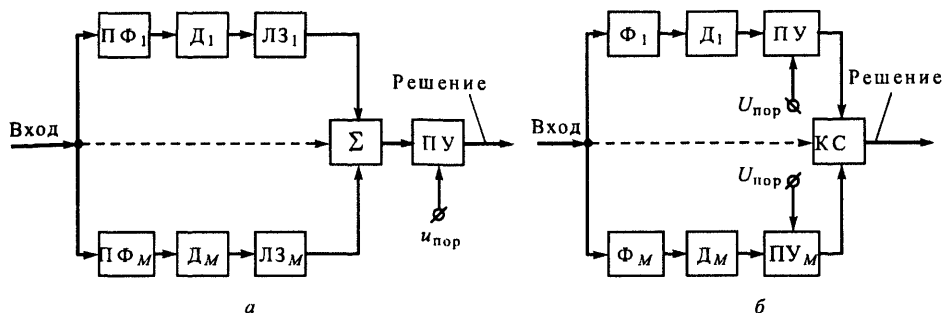


Рис. 9.42. Структурные схемы устройств обработки частотно-временных сигналов со сложением демодулированных частотных составляющих (а) и с каскадом совпадения (б)

или отсутствии кодовой комбинации, переносящей информацию корреспонденту K_i , выносится путем сравнения выходных сигналов этих устройств с порогом в пороговом устройстве (ПУ).

Приведенная структурная схема обработки адресных кодов называется *схемой с раздельной обработкой ортогональных составляющих*. На практике распространены две разновидности этой схемы, отличающиеся методами объединения демодулированных частотных составляющих адресного кода: схема со сложением (рис. 9.42, а) и схема с пороговыми устройствами на выходах демодуляторов и каскадом совпадения (рис. 9.42, б). В

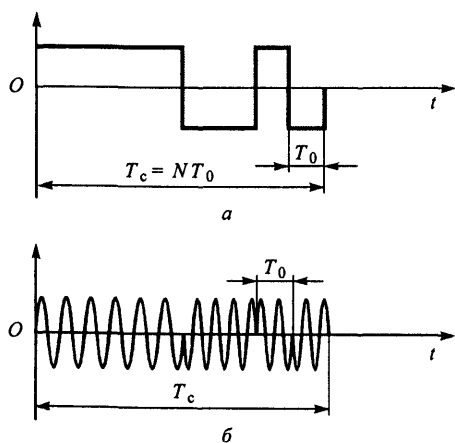


Рис. 9.43. Структура сигналов в ААС со сложными фазоманипулированными сигналами:

а — модулирующего; б — фазоманипулированного

первой сигналы с выходов детекторов через линии задержки подаются на сумматор, по выходному сигналу которого принимается решение о наличии информационной единицы или нуля. Во второй к выходам детекторов подключены пороговые устройства, а решение принимается в каскаде совпадения.

Сложные фазоманипулированные сигналы в ААС состоят из элементарных импульсов, имеющих одинаковую несущую частоту и отличающихся по какому-либо параметру, например по фазе. Фаза изменяется по закону некоторого модулирующего кода (рис. 9.43, а), причем наиболее распространена двухфазная

манипуляция со сдвигом фазы на 180° (рис. 9.43, б). Если определить полосу сигнала на рис. 9.43, а известным соотношением $F_c \approx 1/T_0$, то при длительности сигнала $T_c = NT_0$ его база $B \approx N$, где N — число символов в модулирующей кодовой последовательности.

Системой сигналов называется множество сигналов, определяемых единым правилом построения (алгоритмом). Возможное число адресных сигналов k представляется как объем системы сигналов. Принято сравнивать объем системы сложных сигналов с базой B . Различают малые системы сигналов ($k \approx \sqrt{B} \ll B$), нормальные ($k \approx B$) и большие ($k \gg B$). Большинство известных систем сигналов являются малыми или нормальными.

Сигналы, входящие в систему, должны обеспечивать минимально возможный уровень взаимных помех, который для систем сложных фазоманипулированных сигналов зависит от вида модулирующей кодовой последовательности. Среди множества кодовых последовательностей особый интерес для применения в ААС рассматриваемого типа представляют линейные рекуррентные последовательности максимальной длины, или сокращенно М-последовательности, формируемые с помощью весьма простых генераторов на основе сдвигающих регистров с линейными обратными связями. Они обладают рядом важных свойств, позволяющих формировать на их основе квазиортогональные системы сигналов, характеризующиеся достаточно слабыми взаимными помехами.

В передатчике ААС со сложными фазоманипулированными (ФМ) сигналами, предназначенной для передачи дискретных сообщений (рис. 9.44, а), от источника информации (ИИ) последовательность символов 1 и 0 со скоростью $R = 1/T_c$ (рис. 9.45, а) поступает на вход фазового модулятора (ФМн). На второй вход ФМн подается кодовый сигнал $\{a_i\}$ (рис. 9.45, б) от генератора кода (ГК). Этот сигнал имеет длительность T_c и число импульсов N . Работой ГК и ИИ управляет синхронизатор С, который формирует необходимые сигналы управления и тактовые частоты. В модуляторе (Мод) по закону кодовой последовательности осуществляется фазо-

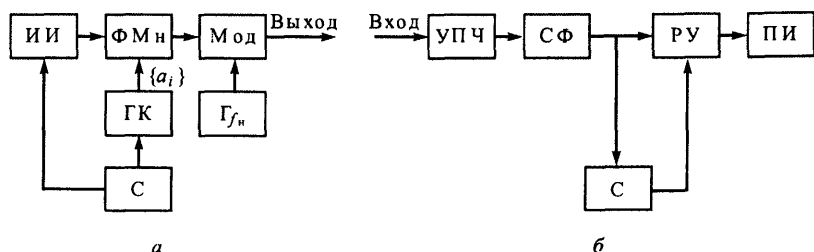


Рис. 9.44. Структурные схемы передатчика (а) и приемника (б) ААС со сложными фазоманипулированными сигналами

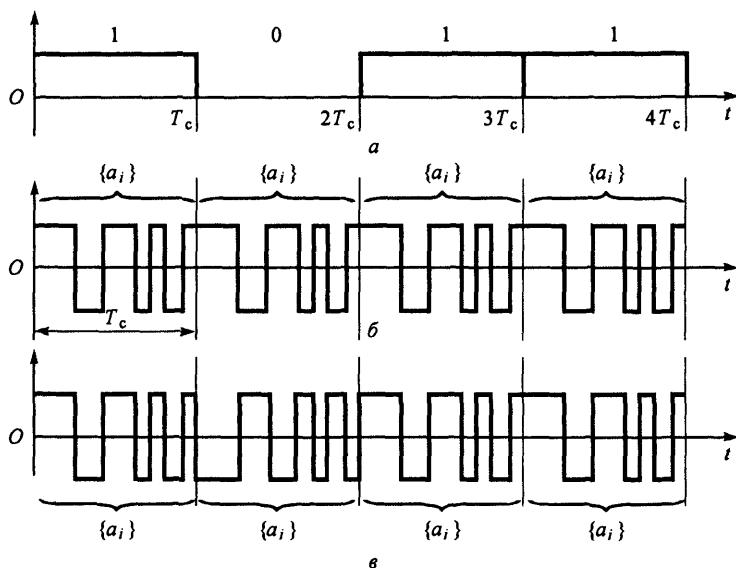


Рис. 9.45. Диаграммы, поясняющие принцип передачи двоичной информации в ААС со сложными фазоманипулированными сигналами:

a — последовательность информационных символов; *б* — кодовый сигнал; *в* — модулированный сигнал

вая манипуляция несущей частоты, вырабатываемого генератором Γ_{f_n} (рис. 9.45, *в*). В приемнике (рис. 9.44, *б*) сигнал переносится на промежуточную частоту, усиливается в усилителе промежуточной частоты (УПЧ) и обрабатывается согласованным фильтром (СФ). Сигнал с выхода СФ поступает на синхронизатор (С) и решающее устройство (РУ). Синхронизатор осуществляет поиск ФМ сигнала по времени и управляет режимом работы решающего устройства. После вхождения в синхронизм на выходе РУ появляется информационная последовательность в виде двоичных символов, которая выдается получателю информации (ПИ).

При работе ААС возникают *межстанционные* (или *системные*) *помехи*. При проектировании системы следует так выбирать ее параметры, чтобы уменьшить такие помехи до допустимого уровня. При этом надо учитывать целый ряд сопутствующих факторов и в первую очередь динамический диапазон сигнала.

Одновременно передающие информацию корреспонденты могут находиться на различных расстояниях друг от друга. При работе передатчиков с постоянной мощностью, рассчитанной на предельные дальности связи в ААС, на входы всех приемников будут приходить сигналы, значительно от-

личающиеся по своей интенсивности. Разброс интенсивностей принято характеризовать динамическим диапазоном сигналов. При малом динамическом диапазоне сигналов и одновременной работе большого числа корреспондентов межстанционная помеха, равная сумме сложных сигналов от отдельных корреспондентов, по своим статистическим характеристикам близка к гауссовскому случайному процессу, т. е. к шуму в полосе частот, занимаемой сигналами ААС.

Рассмотрим более подробно влияние межстанционных помех на помехоустойчивость ААС при малом динамическом диапазоне и большом числе одновременно работающих корреспондентов l_a . Пусть ширина спектра сигналов ААС равна F_Σ , а мощности сигналов всех активных корреспондентов на входе j -го приемника одинаковы и равны P_c . В этом случае мощность полезного сигнала равна P_c , а мощность помехи $\dot{l}^* P_c$, где $\dot{l}^* = l_a - 1$. Допустим, что спектральная плотность мощности помехи постоянна в пределах общей полосы частот:

$$N_n = \dot{l}^* P_c / F_\Sigma. \quad (9.22)$$

Предположим также, что число слагаемых в помехе $\dot{l}^* \gg 1$. Поэтому можно допустить, что помеха по своим статистическим свойствам приближается к гауссовскому случайному процессу. Таким образом, сделанные предположения позволяют считать в первом приближении помеху гауссовским случайным процессом с равномерной спектральной плотностью мощности (9.22).

Как известно, помехоустойчивость когерентного и некогерентного приема полностью определяется отношением

$$h^2 = \frac{P_c T_c}{N_n} = \frac{P_c T_c F_\Sigma}{\dot{l}^* P_c} = \frac{B}{\dot{l}^*}, \quad (9.23)$$

где $B = F_\Sigma T_c$ — база сигнала. Из (9.23) следует, что для надежной передачи информации ($h^2 \gg 1$) необходимо, чтобы база B была много больше числа мешающих корреспондентов \dot{l}^* , т. е. чтобы $\dot{l}^* / B \ll 1$. При передаче информации с заданной достоверностью (при $h^2 = \text{const}$) формула (9.23) позволяет найти необходимое отношение \dot{l}^* / B . Из (9.23) следует, что при известном числе активных корреспондентов $l_a = \dot{l}^* + 1$ увеличить отношение сигнал—помеха можно только за счет увеличения базы B .

При записи формулы (9.23) не учитывался собственный шум приемника, являющийся гауссовским стационарным случайным процессом с равномерной спектральной плотностью N_0 . С его учетом результирующая спектральная плотность равна $N_n + N_0$, а отношение

$$h^2 = P_c T_c (P_c \dot{l}^* / F_\Sigma + N_0)^{-1}. \quad (9.24)$$

Обозначая $h^2_0 = P_c T_c / N_0$, из (9.24) получаем

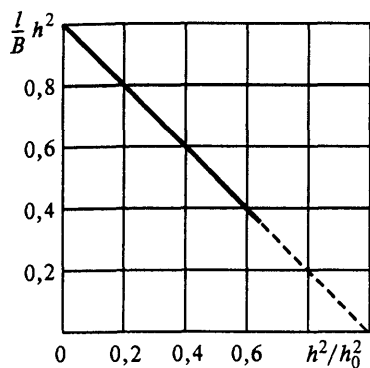


Рис 9.46. Зависимость числа мешающих корреспондентов в ААС от достоверности приема и уровня шума в канале связи

формулы (9.23). При возрастании шума и при $h^2 = \text{const}$ допустимое число мешающих корреспондентов уменьшается.

В тех случаях, когда на входе приемника присутствует одна или несколько ($l^* \leq 2 \dots 4$) стационарных помех, их воздействие на достоверность приема отличается от гауссовского белого шума. Степень различия зависит от коэффициента взаимной корреляции r_{ji} используемых адресных сигналов. Чем он больше, тем сильнее влияет один сигнал на другой.

$$h^2 = (l^*/B + 1/h_0^2)^{-1}. \quad (9.25)$$

В пределе $h^2 = h_0^2$ при $B \rightarrow \infty$.

Из соотношения (9.25) можно найти допустимое число мешающих корреспондентов при заданной достоверности (задано требуемое значение h^2) и для определенного уровня шума (известно значение h_0^2). Преобразуя (9.25), находим, что

$$h^2 l^*/B = 1 - h^2/h_0^2. \quad (9.26)$$

Зависимость (9.26) представлена на рис. 9.46. Если шума нет ($N_0 \rightarrow 0$, $h_0^2 \rightarrow \infty$), то $l^* h^2/B = 1$ и относительное число мешающих корреспондентов равно $l^*/B = 1/h^2$, что совпадает с результатом, получаемым из

9.7. Синхронизация в системах передачи дискретной информации

9.7.1. Принципы построения и основные характеристики систем синхронизации

Система синхронизации в РСПИ в общем случае должна определять следующие синхропараметры сигнала:

- фазу высокочастотного несущего колебания (фазовая синхронизация ФС);
- временные границы принимаемых посылок (тактовая синхронизация ТС);
- моменты времени, соответствующие началу кодовых слов (цикловая синхронизация ЦС);
- моменты времени, соответствующие началу и концу групповых сигналов в многоканальной СПИ (кадровая синхронизация — КС);
- начало и конец передаваемого сообщения.

В подавляющем большинстве случаев сигналы тактовой, кадровой и цикловой синхронизации связаны по фазе между собой (синхронны). Частота повторения кодовых слов $f_{\text{ц}}$ находится делением тактовой частоты $f_{\text{т}}$ на число разрядов в кодовом слове ($f_{\text{ц}} = f/n$), частота повторения кадров — делением частоты повторения кодовых слов на число кодовых слов в кадре ($f_{\text{к}} = f_{\text{ц}}/k_{\text{сл}}$).

Все устройства синхронизации (УС), входящие в систему синхронизации, можно разделить на два принципиально различных типа. Первый тип УС служит для синхронизации отсчетов времени (фазовая и тактовая синхронизация). С их помощью формируются временные шкалы. Второй тип устройств служит для устранения неоднозначности отсчетов времени при определении начала слова, кадра и сообщения. Устройства синхронизации отсчетов времени должны функционировать непрерывно, отслеживая изменение фазы входного колебания, а функции устройств устранения неоднозначности отсчетов времени сводятся к периодическому, а иногда и к однократному фазированию.

Принципиальное различие указанных двух типов УС, естественно, сказывается на методах поиска и оценки синхропараметров, а также на характеристиках качества их работы. Качество работы УС в общем случае должно определяться степенью соответствия фазы входного колебания и колебания местного генератора. До начала работы неопределенность в оценке фазы φ принимаемых синхроколебаний может быть задана плотностью распределения $w_{\text{вх}}(\varphi)$. Если нет дополнительной информации о параметре φ , то логично предположить, что закон распределения $w_{\text{вх}}(\varphi)$ является равномерным на интервале $[-\pi, \pi]$. В результате работы системы синхронизации неопределенность уменьшается, причем происходит это поэтапно. Сначала при когерентном приеме осуществляется ФС, затем ТС и только потом устраняется неоднозначность отсчетов в устройствах ЦС и КС.

Погрешности синхронизации отсчетов времени при случайных внешних воздействиях могут быть заданы законом распределения $w(\Delta\varphi)$, а скорость их изменения — корреляционной функцией $R_{\varphi}(\tau)$ или спектром флуктуации $G_{\varphi}(f)$. Чаще всего полагают закон распределения $w(\Delta\varphi)$ гауссовским с дисперсией σ^2 .

Удобной количественной характеристикой оценки погрешности синхронизации является вероятность попадания фазы колебания местного генератора в некоторую область допустимых значений $\Delta\varphi_{\text{сх}}$. Эту область можно определить как *область синхронизма*. Если $\Delta\varphi_{\text{сх}}$ задана, то можно ввести еще ряд показателей качества работы УС. Такими показателями следует считать время достижения синхронизма $T_{\text{сх}}$ (длительность переходных процессов до достижения области синхронизма), вероятность срыва синхронизма и время поддержания синхронизма при пропадании сигнала на входе.

При наличии на входе сигнала решение о синхронизме может быть правильным или ошибочным. Поэтому качество синхронизации следует характеризовать вероятностью ложного синхронизма $P_{л.с}$ при заданном времени анализа T_a .

В ряде РСПИ важно знать начало передаваемого сообщения. Для этого в начале сеанса связи передается специальный сигнал (преамбула), по которому оценивается факт передачи сообщения и его временное положение. В этом случае возможны ошибки следующих видов: ложное обнаружение сигнала с вероятностью $P_{л.}$, пропуск сигнала с вероятностью $P_{п.с}$ и ложная синхронизация, когда выносится решение, что сигнал есть, но его временное положение оценивается неправильно с вероятностью $P_{л.с}$.

Обрабатывая входной сигнал, система синхронизации получает информацию о синхропараметрах. В простейшем случае, когда в спектре принимаемого сигнала содержится составляющая требуемой частоты и фазы (например, при ФС), обработка сводится к фильтрации синхроколебания. Если в спектре сигнала отсутствуют составляющие, несущие информацию о синхропараметрах, то сигнал предварительно подвергается в приемнике нелинейным преобразованиям.

В общем случае УС должно содержать входной преобразователь (ВП), в выходном сигнале которого содержится колебание требуемой частоты и фазы. Этим колебанием синхронизируется местный управляемый генератор (УГ). Для уменьшения дисперсии флуктуаций фазы, возникающих из-за действия помех, синхросигнал предварительно фильтруется полосовым фильтром.

Различают УС *разомкнутые* (рис. 9.47, а), в которых синхроколебание фильтруется полосовым фильтром (аналоговым или цифровым), и *замкнутые* (рис. 9.47, б), построенные на базе систем фазовой автоподстройки частоты (ФАПЧ). В замкнутых системах колебание синхронизируемого УГ в фазовом детекторе (ФД) сравнивается по фазе с входным колебанием. Затем сигналом рассогласования УГ управляется по частоте так, чтобы свести это рассогласование к минимуму. Сглаживание сигнала рассогласования осуществляется низкочастотным фильтром (ФНЧ). Принципиальным достоинст-

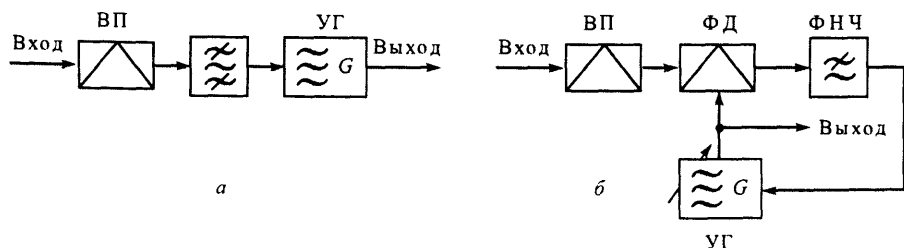


Рис. 9.47. Структурные схемы разомкнутой (а) и замкнутой (б) систем фазовой синхронизации

вом системы ФАПЧ является то, что она представляет собой следящий фильтр, в котором шумовая полоса $F_{\text{ш}}$ может быть достаточно узкой независимо от диапазона изменения частоты входного колебания.

Работа отдельных ступеней системы синхронизации по-разному сказывается на качестве работы РСПИ в целом. Погрешности в работе устройства цикловой синхронизации (УЦС) и устройства кадровой синхронизации (УКС) приводят к неправильному приему всего сообщения или части его, а устройства фазовой синхронизации (УФС) и устройства тактовой синхронизации (УТС) — к снижению его достоверности. Оценим влияние погрешностей УФС и УТС на достоверность принимаемого сообщения. УФС входит в состав когерентного демодулятора и обеспечивает совпадение по фазе напряжения местного генератора и несущей частоты. Погрешности $\Delta\varphi$ УФС приводят к уменьшению отношения сигнал—шум на выходе устройства обработки:

$$q^2(\Delta\varphi) = (2E/N_0) \cos^2(\Delta\varphi) = q^2 \cos^2(\Delta\varphi).$$

Поэтому, оценивая достоверность принимаемого сообщения с учетом ошибок ФС, можно ввести условную вероятность ошибки $P_{\text{ош}}(\Delta\varphi)$. Если скорость флуктуации ошибок ФС мала (интервал корреляции случайного процесса $\Delta\varphi(t)$ много больше длительности посылок, $\tau_k \gg T_c$), то среднюю вероятность ошибок можно найти по формуле

$$\bar{P}_{\text{ош}}(\Delta\varphi) = \int_{-\pi}^{\pi} P_{\text{ош}}(\Delta\varphi) w(\Delta\varphi) d(\Delta\varphi). \quad (9.27)$$

Сложность зависимости вероятности ошибок $P_{\text{ош}}$ от погрешности $\Delta\varphi$ не позволяет получать точной формулы для $\bar{P}_{\text{ош}}(\Delta\varphi)$ (рис. 9.48, а). Поэтому пользуются либо приближенными соотношениями, найденными при тех или иных упрощениях, либо численными методами с привлечением ЭВМ.

Влияние ошибок ФС на помехоустойчивость можно оценить по приближенной формуле, справедливой для систем с фазовой модуляцией любой кратности m :

$$\bar{P}_{\text{ош}}(\Delta\varphi) \approx 1 - \Phi \left[\frac{\pi}{m} \frac{h}{\sqrt{1 + 2\sigma_{\varphi}^2 h^2}} \right]. \quad (9.28)$$

Значение σ_{φ} на рис. 9.48, а и в формуле (9.28) задано в радианах.

Погрешности тактовой синхронизации τ приводят к тому, что момент отсчета напряжения на выходе коррелятора или согласованного фильтра не соответствует моменту окончания посылки. Тогда при смене символов на входе решающей схемы будут накапливаться сигналы от посылок разного знака и отношение сигнал—шум q^2 уменьшится:

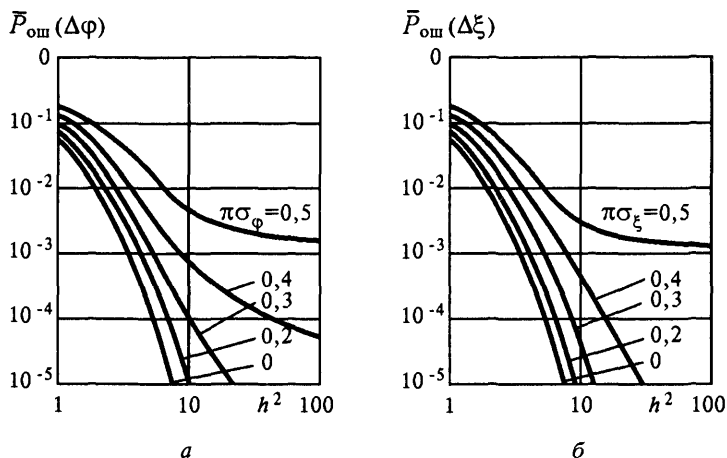


Рис. 9.48. Зависимости вероятности ошибки когерентного приема противоположных двоичных сигналов при наличии погрешностей фазовой (а) и тактовой (б) синхронизации (законы распределения $w(\varphi)$ и $w(\xi)$ гауссовские)

$$q^2(\xi) = \frac{2E}{N_0} \left(1 - 2 \frac{|\tau|}{T_c} \right) = q^2 (1 - 2|\xi|).$$

Введя условную вероятность ошибки $P_{\text{ош}}(\xi)$ и зная закон изменения $w(\xi)$, можно вычислить среднюю вероятность ошибки $\bar{P}_{\text{ош}}(\xi)$ аналогично (9.27). Как и в предыдущем случае, количественные результаты можно получить численными методами. Ошибки тактовой синхронизации, в отличие от фазовой, будут влиять только на прием чередующихся символов. Для случайной последовательности двоичных символов вероятность смены знака равна 0,5, а средняя вероятность ошибки (рис. 9.48, б)

$$\bar{P}_{\text{ош}}(\xi) = 0,5[P_{\text{ош}}(\xi) + P_{\text{ош}}],$$

где $P_{\text{ош}}(\xi)$ — вероятность ошибки при идеальной тактовой синхронизации.

Интересно сравнить степень влияния на помехоустойчивость ошибок фазовой и тактовой синхронизации при заданном отношении сигнал—шум в канале. Для этого допустим, что тактовая частота формируется из опорной, равной частоте несущей, путем деления на k_d . Тогда флуктуации фазы тактовой частоты будут меньше флуктуаций фазы опорной частоты в k_d раз. Из этого следует, что в режиме слежения, когда неоднозначность отсчета фазы тактовой частоты устранена, на помехоустойчивость будут влиять ошибки устройства фазовой синхронизации. Это утверждение справедливо и тогда, когда опорная и тактовая частоты формируются от разных генераторов.

9.7.2. Фазовая синхронизация модемов

Система ФАПЧ (см. рис. 9.47, б) является основным звеном устройств синхронизации отсчетов времени. Она в том или ином виде входит в УФС и УТС демодулятора и служит для фильтрации синхроколебания. Работу ФАПЧ можно характеризовать режимами слежения и захвата. В режиме слежения частоты входного и опорного сигналов одинаковы. С фазового детектора снимается напряжение, пропорциональное $\cos(\varphi)$, где φ — разность фаз между входным и опорным сигналами. Этим напряжением управляется по частоте (фазе) УГ так, чтобы свести рассогласование к $\pi/2$. Включив фазовращатель на $\pi/2$, можно компенсировать постоянный фазовый сдвиг между входным и опорным сигналами. В установившемся режиме значение φ определяется параметрами ФАПЧ, начальной разностью частот между входным и опорным сигналами и уровнем шумов на входе. Если систему ФАПЧ рассматривать как четырехполюсник, в котором выходным сигналом являются колебания УГ, то по своим свойствам она подобна следящему полосовому фильтру. Характеризовать работу такого следящего фильтра можно следующими параметрами:

- шумовой полосой $F_{\text{ш}}$ (полосой пропускания);
- полосой удержания F_y (максимальным изменением частоты входного колебания, при котором ФАПЧ, находясь в режиме слежения, сохраняет работоспособность);
- полосой захвата F_z (максимальной расстройкой между частотой входного колебания и частотой управляемого генератора, при которой после включения ФАПЧ начинает следить за фазой входного колебания, входит в синхронизм);
- временем ввода в синхронизм $T_{\text{сх}}$, которое характеризует продолжительность переходного процесса от момента включения сигнала на входе до момента установления равенства частот входного сигнала и сигнала управляемого генератора.

Анализ работы системы ФАПЧ, особенно в режиме захвата и при воздействии помех, представляет сложную математическую задачу. Сложность ее решения определяется тем, что поведение ФАПЧ описывается нелинейным дифференциальным уравнением. Обозначим частоты входного и опорного сигналов соответственно через $\omega_0 + \Delta\omega_{\text{вх}}(t)$ и $\omega_0 + \Delta\omega_{\text{ур}}(t)$, где ω_0 — центральная частота, а $\Delta\omega_{\text{вх}}(t)$ и $\Delta\omega_{\text{ур}}(t)$ — изменение ее во времени для входного и опорного сигналов. Тогда разность частот входного сигнала и опорного можно записать в виде

$$\Delta\omega(t) = \Delta\omega_{\text{вх}}(t) - \Delta\omega_{\text{ур}}(t).$$

Рассматривая $\Delta\omega_{\text{вх}}(t)$ как входное воздействие на систему, а $\Delta\omega(t) = \partial\varphi(t)/\partial t$ как реакцию системы, можно записать дифференциальное уравнение, описывающее ее поведение:

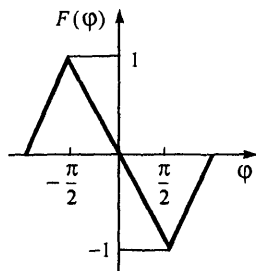


Рис. 9.49. Характеристика фазового детектора

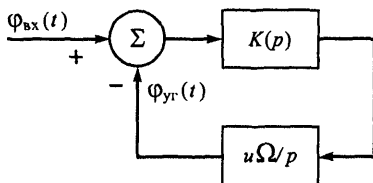


Рис. 9.50. Линейаризованная модель системы ФАПЧ

$$\partial\varphi/\partial t + u\Omega F(\varphi)K(p) = \Delta\omega_{\text{вх}}(t),$$

где φ — текущая разность фаз между сигналами управляемого генератора и входным; u — напряжение входного сигнала; $F(\varphi)$ — нормированная характеристика фазового детектора; $K(p)$ — коэффициент передачи фильтра в операторной форме; Ω — коэффициент передачи управляемого генератора, характеризующий приращение частоты управляемого генератора при подаче на его вход постоянного напряжения один вольт.

Характеристика фазового детектора $F(\varphi)$ нелинейна, периодична и определяется формой опорного и входного сигналов. Если аппроксимировать $F(\varphi)$ полигональной кривой (рис. 9.49), то $F(\varphi) = -(2/\pi)\varphi$, $-\pi/2 < \varphi < \pi/2$. При малых внешних воздействиях, когда частоты входного и опорного сигналов совпадают, а фазовое рассогласование не превышает $\pi/2$, систему ФАПЧ можно заменить эквивалентной линейаризованной моделью (рис. 9.50). Управляемый генератор, для которого частотный сдвиг выходного сигнала пропорционален входному сигналу, а фазовый — интегралу от него, выполняет роль интегрирующего звена с коэффициентом передачи в операторной форме $u\Omega/p$.

Передаточная функция замкнутой линейаризованной системы ФАПЧ $H(p)$ связана с передаточной функцией фильтра системы $K(p)$ в операторной форме соотношением

$$H(p) = \frac{\varphi_{\text{уг}}(p)}{\varphi_{\text{вх}}(p)} = \frac{2u\Omega K(p)/\pi}{(2/\pi)u\Omega K(p) + p},$$

где $\varphi_{\text{уг}}(p)$ и $\varphi_{\text{вх}}(p)$ — запись сигналов $\varphi_{\text{уг}}(t)$ и $\varphi_{\text{вх}}(t)$ в операторной форме.

Анализ этого выражения позволяет сделать вывод, что передаточная функция $H(p)$ определяется как внутренними параметрами ФАПЧ (Ω , $K(p)$), так и внешними (u). Зависимость $H(p)$ от u нежелательна, и для ее устра-

нения или ослабления применяют автоматическую регулировку усиления (АРУ). В дальнейшем будем полагать $u = 1$. Выбор оптимальной формы передаточной характеристики $K(p)$ и коэффициента Ω является основной задачей при проектировании ФАПЧ. Для определения $K(p)$ достаточно задать $H(p)$ и затем через нее выразить $K(p)$:

$$K(p) = \frac{p H(p)}{(2/\pi) \Omega (1 - H(p))}.$$

Поскольку ФАПЧ в УС выполняет роль полосового фильтра с частотной характеристикой $H(j\omega)$, то важнейшим показателем качества ее работы является *шумовая полоса*

$$F_{\text{ш}} = \frac{1}{2\pi} \int_0^{\infty} |H(j\omega)|^2 d\omega.$$

Выбрав соответствующим образом $K(p)$ и Ω , можно получить требуемое значение шумовой полосы. Однако при этом необходимо помнить, что фильтр, реализующий характеристику $H(p)$, должен быть устойчивым. Стремление сузить шумовую полосу приводит к уменьшению полосы захвата. Доказано, что полоса захвата в системе ФАПЧ с любым ФНЧ не может быть больше шумовой полосы. Уменьшение шумовой полосы влечет за собой также увеличение времени ввода в синхронизм $T_{\text{сх}}$. Для ФАПЧ без ФНЧ шумовая полоса, полоса захвата и полоса удержания равны между собой:

$$F_{\text{ш}} = F_{\text{з}} = F_{\text{у}} = \frac{1}{2\pi} \int_0^{\infty} \frac{(2\Omega/\pi)^2}{(2\Omega/\pi)^2 + \omega^2} d\omega = \frac{\Omega}{2\pi}.$$

Используя пропорционально интегрирующий фильтр (рис. 9.51), у которого

$$F_{\text{ш}} = \frac{\Omega}{2\pi} \frac{\pi + 2\Omega T_1^2/T_2}{\pi + 2\Omega T_1},$$

где $T_1 = R_2C$, $T_2 = (R_1 + R_2)C$, можно получить достаточно узкую шумовую полосу, если соответствующим образом выбрать параметры R_1 , R_2 , C . Время ввода в синхронизм в этом случае $T_{\text{сх}} \approx 3,5\Delta f^2/F_{\text{ш}}^3$, где Δf — расстройка между частотой опорного и управляемого генераторов.

Противоречие между улучшением фильтрующих свойств ФАПЧ и уменьшением полосы захвата может быть разрешено, если применить до-

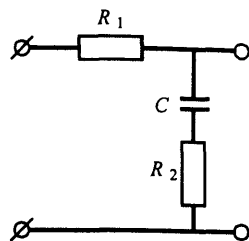


Рис. 9.51. Электрическая схема пропорционального интегрирующего устройства

полнительные поисковые процедуры или изменить параметры ФНЧ (адаптация) на этапе ввода в синхронизм. В первом случае на вход УГ следует подать пилообразное напряжение, которое изменяет его частоту. В определенный момент разность частот входного и опорного сигналов оказывается такой, что происходит захват и система входит в синхронизм. Амплитуда пилообразного напряжения должна обеспечивать перестройку генератора в требуемом диапазоне частот, а его период должен быть достаточен для завершения переходных процессов в системе ФАПЧ. При этом методе можно время поиска $T_{п.с}$ ориентировочно оценить по формуле $T_{п.с} \approx \pi \Delta f / F_{ш}^2$. Поисковая процедура ввода в синхронизм особенно эффективна, когда $\Delta f \gg F_{ш}$.

Добиться уменьшения $T_{сх}$ можно также изменением параметров ФАПЧ Ω и $K(p)$. На этапе ввода в синхронизм можно расширить полосу пропускания ФАПЧ или увеличить коэффициент передачи Ω . После того как поиск будет завершен, эти параметры должны обеспечивать нужное качество слежения. Выбор оптимального режима требует моделирования системы ФАПЧ.

Системы передачи дискретной информации, в приемном тракте которых для демодуляции сигнала используется когерентное опорное колебание на несущей частоте, получили название *когерентных*. В качестве демодуляторов в них применяется согласованный фильтр или коррелятор, а основной вид модуляции — фазовый. Опорное напряжение несущей частоты формируется в устройстве фазовой синхронизации (УФС), реализуемом, как правило, на базе ФАПЧ.

Спектр сигнала $A_0 \sin(\omega_0 t + \{x_i\} \psi + \varphi(t))$ при равновероятной передаче символов x_i и $\psi = 2\pi/m$ не содержит дискретной составляющей на частоте ω_0 . Один из способов ее восстановления основан на том, что в спектре передаваемого радиосигнала остается составляющая достаточной мощности на частоте несущей. Для этого девиацию делают равной $(2\pi - \varphi)/m$. Другой способ предусматривает использование сигнала с полностью подавленной несущей. Чтобы при этом восстановить несущую, сигнал подвергают нелинейным преобразованиям. Рассмотрим способы ФС и оценим их целесообразность в тех или иных конкретных ситуациях.

Спектр сигнала с неполностью подавленной несущей состоит из дискретной и непрерывной составляющих, т. е. смешанный. Для двоичной ФМ сигнала дискретная составляющая на частоте ω_0 равна $A^2_0 \cos^2 \varphi / 2$. Непрерывная часть спектра связана со случайным чередованием нулей и единиц в сообщении и имеет вид

$$G(\omega) = 2A_0^2 T_c \sin^2 \left(\frac{\psi}{2} \right) \frac{\sin^2((\omega - \omega_0)T_c/2)}{((\omega - \omega_0)T_c/2)^2}.$$

Для двоичной ФМ с произвольной девиацией фазы ψ при идеальном канале синхронизации вероятность ошибки

$$P_{\text{ош}} = 1 - \Phi\left(\sqrt{h^2(1-r_{12})}\right), \quad (9.29)$$

где $r_{12} = \frac{1}{E_c} \int_0^{T_c} s_1(t)s_2(t)dt = \cos(\psi)$ — коэффициент взаимной корреляции

сигналов $s_1(t)$ и $s_2(t)$. Из (9.29) следует, что вероятность ошибки оказывается минимальной при $\psi = \pi$. Однако в этом случае в спектре сигнала будет отсутствовать дискретная составляющая на частоте ω_0 . Соотношение между мощностью дискретной составляющей и мощностью непрерывной составляющей зависит от девиации фазы ψ . Изменяя значение ψ , можно добиться оптимального режима работы системы передачи в целом. Под оптимальным режимом следует понимать такой, для которого при прочих равных условиях достигается минимальное значение вероятности ошибки.

В канале синхронизации дисперсия флуктуации фазы имеет вид

$$\sigma_\varphi^2 \approx \frac{F_{\text{ш}}(N_0 + G(f_0))}{P_c \cos^2(\psi/2)} = \frac{N_0 F_{\text{ш}}}{P_c \cos^2(\psi/2)} + F_{\text{ш}} T_c \text{tg}^2\left(\frac{\psi}{2}\right).$$

Влиянием ошибок синхронизации на вероятность ошибки $\bar{P}_{\text{ош}}(\Delta\varphi)$, как ранее указывалось, можно пренебречь при $\sigma_\varphi \leq 0,2 \dots 0,3$. Поэтому, если допустить, что на передачу информации отводится большая часть мощности сигнала, то рассматриваемый способ формирования синхросигнала можно

применять при $Q = \frac{1}{F_{\text{ш}} T_c} \geq 10$. Это означает, что полоса пропускания

фильтра $F_{\text{ш}}$ в канале синхронизации должна быть существенно меньше ширины спектра информационного сигнала. В противном случае помеха, создаваемая ФМ сигналом, делает прием неэффективным. Оптимальное соотношение между мощностью синхросигнала $P_{\text{сч}}$ и полной мощностью сигнала P_c с учетом ошибок, вносимых системой синхронизации, можно оценить по формуле

$$\left(\frac{P_{\text{сч}}}{P_c}\right)_{\text{опт}} = \left(h^2 - \frac{1}{4}\right)^{-1} \left(\left(1 + \frac{h^2(Q+2)}{(h^2 - 1/4)^2}\right)^{1/2} - 1 \right)^{-1}.$$

Заметим, что когда значение Q велико по сравнению с h^2 , то

$$(P_{\text{сч}}/P_c)_{\text{опт}} \approx 1/\sqrt{h^2 Q}.$$

Использование части мощности сигнала на синхронизацию приводит к потере помехоустойчивости. Этого можно избежать, если выделить синхросигнал непосредственно из принимаемого путем нелинейного преобразования. В зависимости от этой процедуры различают три вида схем:

- с квадратичной нелинейной обработкой сигнала (схемы Пистолькорса, Сифорова);
- ФАПЧ с квадратурными каналами (схема Костаса);
- с обратной связью по решению.

Для схемы с квадратичной обработкой (рис. 9.52, а) получаемый на выходе нелинейного элемента сигнал будет содержать гармонику частоты $2f_0$, которую можно отфильтровать полосовым фильтром (схема Пистолькорса) или ФАПЧ (схема Сифорова). Разделив эту частоту на 2, получим частоту f_0 . Поскольку во всяком нелинейном элементе при малом отношении сигнал—помеха подавляет сигнал, то перед устройством возведения в квадрат желательно включить фильтр.

Принципиальный недостаток этой схемы, как, впрочем, и всех других, — неоднозначность оценки фазы: изменение фазы входного сигнала на π не отражается на фазе выходного сигнала. Следствием этого является так называемый эффект *обратной работы* (инверсия принятой последовательности символов). Для устранения обратной работы используется относительная фазовая модуляция.

В схеме Костаса (рис. 9.52, б) входной сигнал раскладывается на две квадратурные составляющие: $A_0 x_i \cos(\varphi(t))$ и $A_0 x_i \sin(\varphi(t))$, которые затем в

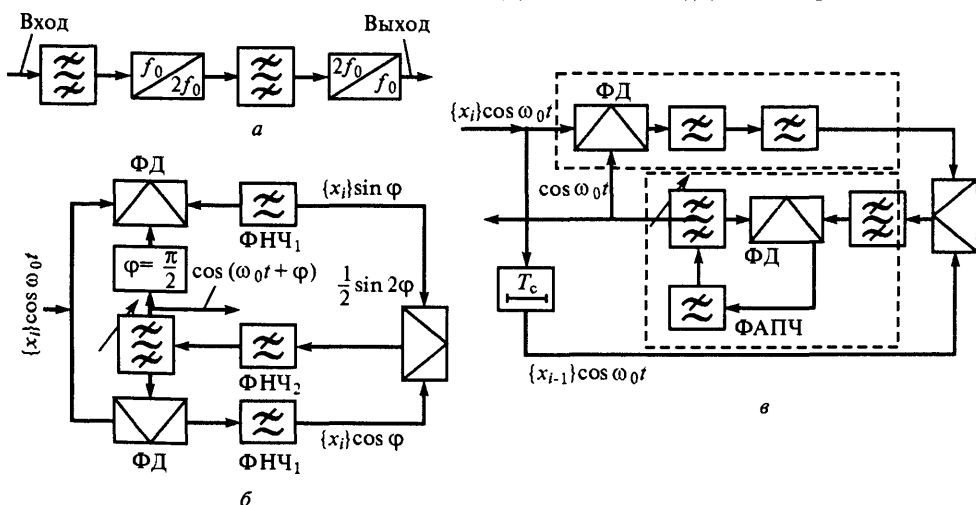


Рис. 9.52. Структурные схемы устройств фазовой синхронизации с квадратичной нелинейной обработкой модулированного сигнала (а), с квадратурными каналами (б) и обратной связью по решению (в)

результате умножения дают напряжение, пропорциональное $A_0 \sin(2\varphi(t))$. Этим напряжением, предварительно отфильтрованным низкочастотным фильтром ФНЧ₂ с характеристикой $K_2(f)$, генератор управляется по частоте так, чтобы свести к минимуму фазовое рассогласование $\varphi(t)$. Оптимальная форма частотной характеристики ФНЧ₁ $K_1(f)$, минимизирующая погрешности измерения фазы $\varphi(t)$, определяется спектральными плотностями полезного сигнала на входе фильтра $G(f)$ и шума $N(f)$. При белом шуме

$$K_1(f) = \frac{G(f)}{G(f) + 0,5N_0}, \quad (9.30)$$

где

$$G(f) = P_c T_c \left(\frac{\sin(\pi f T_c)}{\pi f T_c} \right)^2.$$

Из (9.30) следует, что при малой спектральной плотности мощности шума фильтр ФНЧ₁ практически не нужен, а при большой он должен быть согласован с входным сигналом. Для этих двух крайних случаев дисперсия флуктуации фазы $\sigma_\varphi^2 \approx 1/(Qh^2)$ и $\sigma_\varphi^2 \approx 3/(4Qh^4)$ и определяется отношением сигнал—шум и параметром Q . Схема с квадратичной обработкой и схема Костаса при отсутствии частотных нестабильностей несущей обладают одинаковой точностью оценки фазы φ . Однако схема Костаса проще и технологичнее, так как в ней отсутствуют полосовые фильтры. Кроме того, при частотных сдвигах несущей, например обусловленных эффектом Доплера, точность оценки фазы в этой схеме выше, потому что фильтрация сигнала осуществляется после отслеживания частоты в более узкой полосе. Неоднозначность отсчета фазы в схеме сохраняется.

Дальнейшего улучшения характеристик УФС можно достичь, если при формировании опорного сигнала использовать демодулированные послышки. Такое устройство, называемое УФС с обратной связью по решению (рис. 9.52, в), позволяет получать минимальную дисперсию ошибок в оценке фазы, так как все операции над сигналом при малой вероятности ошибки на символ в схеме линейны. Задержанный принимаемый сигнал перемножается с восстановленными послылками, и затем получаемый гармонический сигнал несущей частоты фильтруется с помощью системы ФАПЧ. В момент включения УФС, когда фазы опорного и принимаемого сигналов не совпадают, послышки восстанавливаются с большими искажениями, но по мере уменьшения фазового рассогласования их достоверность растет и схема входит в режим синхронизма. Как и все предыдущие схемы, УФС с обратной связью по решению обладает неоднозначностью отсчета фазы.

9.7.3. Тактовая синхронизация модемов

Для обеспечения оптимального приема дискретных сигналов необходима тактовая синхронизация (ТС) демодулятора приемника относительно потока поступающих на вход посылок. Тактовые импульсы (ТИ), временное положение которых совпадает с моментами окончания посылок, управляют работой интеграторов при корреляционной обработке сигнала или используются для снятия отсчета напряжения с выхода согласованных фильтров. При неоптимальном приеме ТИ используются при регенерации посылок. Помимо этого, ТС необходима тогда, когда квазисинхронные потоки символов разных источников объединяются в один поток.

Поскольку при случайном характере передаваемой информации спектр радиосигнала сплошной и расположен в области несущей частоты, то он не содержит составляющей тактовой частоты. Поэтому для обеспечения ТС сигнал должен быть соответствующим образом обработан. Необходимо отметить, что информацию о тактовой частоте в СПИ с простыми сигналами можно выделить только из сигнала, в котором модулирующие посылки меняют свое значение. Сигнал, модулированный посылкой одного знака, информации о тактовой частоте не несет. Чтобы предотвратить появление длинных последовательностей одного знака, часто используют специальные устройства рандомизации потока. Например, в кодере СПИ с ОФМ выходные символы y_i связаны с входными x_i соотношением $y_i = y_{i-1} \oplus x_i$, при этом последовательности одного знака преобразуются в меандр.

При создании устройств тактовой синхронизации (УТС) необходимо найти алгоритм, обеспечивающий наилучшую (в смысле выбранного критерия) оценку временного положения сигнала. Из теории оценок известно, что эта задача сводится к определению максимума функции правдоподобия $\Lambda(\tau)$. Максимум функции $\Lambda(\tau)$ можно найти устройством с параллельным анализом на интервале неопределенности $(0, T_c)$ или с последовательным (следающие УТС). Первый тип устройств позволяет определять τ за минимальное время, однако из-за сложности реализации применяется редко.

В следающих УТС в произвольной точке вычисляется значение производной функции правдоподобия $\left. \frac{\partial \Lambda(\tau)}{\partial \tau} \right|_{\tau=\tau_i}$ (иногда еще дополнительно значение $\Lambda(\tau_i)$), а затем по этому значению в решающем устройстве оценивается наиболее вероятное положение максимума $\Lambda(\tau_i) = \max$. Следующее вычисление проводится в точке, которая позволяет оценивать положение максимума с наибольшей достоверностью. Ею могла бы являться координата максимума функции $\Lambda(\tau)$. Однако система слежения в этой точке оказывается нечувствительной к изменению временного положения входного

сигнала, так как $\left. \frac{\partial \Lambda(\tau)}{\partial \tau} \right|_{\tau=\tau_i} = 0$. Поэтому целесообразно следить за точкой,

где производная $\partial \Lambda(\tau)/\partial \tau$ и значение $\Lambda(\tau)$ достаточно большие. Если передаваемый сигнал известен, то определение $\Lambda(\tau)$ заключается в нахождении модуля функции взаимной корреляции принимаемого сигнала и опорного. В СПИ это принципиально невозможно, так как передаваемая информация носит случайный характер. При этом оптимальный алгоритм вычисления $\Lambda(\tau)$ оказывается слишком сложным и его целесообразно применять лишь для получения оценок потенциально достижимой точности измерения фазы.

На практике используют квазиоптимальные алгоритмы, реализуемые на базе демодулятора посылок без синхронизации. Сигналы с выходов согласованных фильтров детектируются и их разность затем подается на решающую схему (РС). Момент смены знака содержит информацию о фазе тактовой частоты. В качестве примера рассмотрена работа демодулятора двоичных ЧМ сигналов (рис. 9.53, а). В отсутствие шумов сформированные импульсы ТС имеют постоянный временной сдвиг $T_c/2$ относительно тактовых импульсов посылок (рис. 9.53, б). При действии шумов их временное положение изменяется. Дисперсия флуктуаций определяется отношением h^2 , видом модуляции и способом обработки. При когерентной обработке

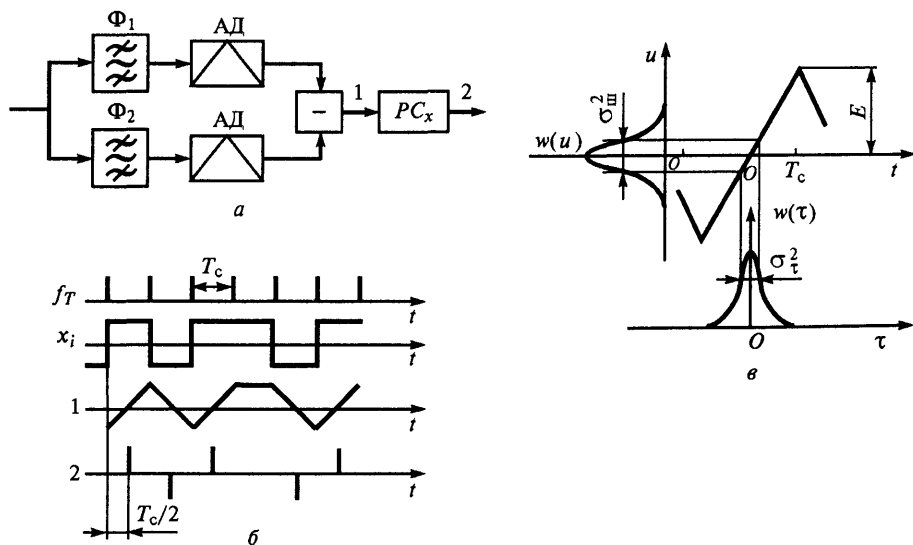


Рис. 9.53. Структурная схема демодулятора двоичных частотно-манипулированных сигналов (а), эпоры напряжений в различных точках ее, иллюстрирующие формирование тактовых импульсов (б), и определение дисперсии флуктуаций фазы импульсов тактовой синхронизации (в)

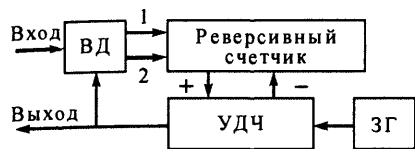


Рис. 9.54. Структурная схема устройства тактовой синхронизации с дискретным управлением

Повышения точности окончательную оценку фазы тактовой частоты проводят по ряду измерений в следящем фильтре. Для этого во временном (фазовом) дискриминаторе (ВД) сравнивают последовательности сформированных и опорных импульсов. Напряжение с выхода дискриминатора определяется разностью фаз. Опорный генератор управляется по фазе так, чтобы свести это рассогласование к минимуму.

По своей структуре алгоритму работы и характеристикам следящее УТС подобно ФАПЧ. Принципиальное отличие здесь заключается в форме фазовой характеристики и во входном сигнале, представляющем собой случайную последовательность импульсов, статистические характеристики которой определяются передаваемой информацией.

Существуют различные способы практической реализации следящих УТС, из которых наиболее распространены устройства с дискретным управлением (рис. 9.54). Последние позволяют получать высокую точность слежения и могут быть реализованы на современной элементной базе с цифровыми методами обработки сигналов. Принцип работы УТС с дискретным управлением основан на смещении фазы сигнала, формируемого управляемым делителем частоты (УДЧ) при добавлении или исключении одного импульса на его входе. Точность управления фазой (шаг подстройки) определяется коэффициентом деления $K_{\text{дел}}$. В качестве фильтра в цепи обратной связи используется интегратор, выполненный на основе реверсивного счетчика. В реверсивном счетчике вычисляется разность числа импульсов, поступивших с выходов ВД 1 и 2. Если она превышает емкость счетчика $N_{\text{сч}}$, то на соответствующем выходе формируется команда, которая в УДЧ приводит к смещению фазы опорного сигнала в требуемую сторону. Емкость реверсивного счетчика $N_{\text{сч}}$ определяет число импульсов, по которому выносится оценка о знаке рассогласования, и, соответственно, помехоустойчивость УТС.

При проектировании дискретного УТС важно правильно выбрать значения $N_{\text{сч}}$ и $K_{\text{дел}}$. Зная $K_{\text{дел}}$ и длительность посылки $T_{\text{с}}$, можно определить шаг подстройки $\Delta t = T_{\text{с}}/K_{\text{дел}}$. Время, через которое произойдет коррекция фазы опорного сигнала, определяется в виде $T_{\text{к}} = 2N_{\text{сч}}T_{\text{с}}$, где коэффициент 2

амплитуда сигнала в точке 1 при отсутствии шума равна E , а флуктуации имеют гауссовский закон распределения $w(u)$ с дисперсией $EN_0/2$. Поэтому, как это следует из рис. 9.53, в, при больших значениях $h^2 = E/N_0$ флуктуации фазы сформированных импульсов ТС также будут подчиняться гауссовскому закону $w(\tau)$ с дисперсией $\sigma_{\tau}^2 \approx T_{\text{с}}^2/h^2$. Для повы-

учитывает случайный характер появления импульсов на входе ВД. Устойчивую работу УТС можно обеспечить только тогда, когда шаг подстройки $T_{\text{кор}}$ достаточен для компенсации временной расстройки ΔT , возникающей за время между двумя подстройками. Причиной появления этой расстройки является нестабильность тактовой частоты $K_f = \frac{\Delta f_{\text{т}}}{f_{\text{т}}} = \frac{\Delta T}{T_{\text{с}}}$. Коэффициент K_f

характеризует скорость «скольжения» частот. Следовательно, предельное значение допустимой относительной нестабильности частот удовлетворяет неравенству $K_f \leq \Delta t / T_{\text{к}}$. Это неравенство является условием синхронизма и позволяет связать основные параметры УТС между собой:

$$K_f \leq \frac{2}{K_{\text{дел}} N_{\text{сч}}} = \frac{2\Delta t}{T_{\text{с}} N_{\text{сч}}}.$$

Важный параметр УТС — время ввода в синхронизм $T_{\text{сх}}$, поскольку оно определяет длительность вхождения в связь. При расчете $T_{\text{сх}}$ будем предполагать, что $K_f = 0$, а временной сдвиг наибольший и равен $T_{\text{с}}/2$. Тогда получим

$$T_{\text{сх}} = \frac{T_{\text{с}}}{2\Delta t} T_{\text{к}} = K_{\text{дел}} N_{\text{сч}} T_{\text{с}}.$$

Последнее выражение позволяет найти параметры устройства тактовой синхронизации с дискретным управлением.

9.7.4. Цикловая и кадровая синхронизация

Устройства цикловой синхронизации (УЦС) предназначены для определения начала кодовых слов. Поскольку при передаче сообщения безызыбычным кодом последовательность символов случайна и информации о начале и конце кодовых слов не несет, то предпринимают специальные меры для внесения этой информации. Цикловую синхронизацию обеспечивают либо с помощью специальных синхросигналов, либо с помощью внутренней избыточности кодовых слов. В обоих случаях цикловая синхронизация реализуется за счет снижения скорости передачи информации.

В качестве циклового синхросигнала можно использовать периодически повторяющиеся от слова к слову сосредоточенные или распределенные синхрогруппы. На приемной стороне синхросигнал, генерируемый местным генератором, сравнивается со входной последовательностью символов при различных взаимных временных положениях. Совпадение символов принимаемого и опорного сигналов фиксируется как режим синхронизма. Этот

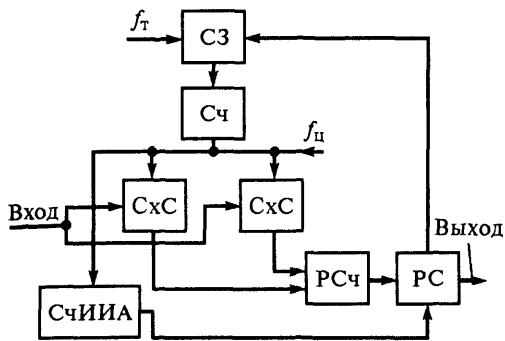


Рис. 9.55. Структурная схема устройства цикловой синхронизации, реализующего шаговый поиск односимвольного синхросигнала

вального синхросигнала в начале каждого кодового слова или группы кодовых слов (рис. 9.55). Импульсы с частотой $f_{\text{ц}}$ (частотой следования слов), формируемые с помощью счетчика (Сч), подаются на схемы совпадения (СхС), на другие входы которых поступают кодовые символы с выхода регенератора посылок. В зависимости от знака этих символов на реверсивный счетчик (РСч) поступают импульсы по одному из двух входов. Счетчик импульсов интервала анализа (СчИИА) определяет отрезок времени длиной в несколько слов, через который число, записанное в реверсивный счетчик, сравнивается с порогом в решающей схеме (РС). Если порог не превышен, то в РС формируется строб, которым с помощью схемы запрета (СЗ) вычеркивается один из подаваемых на счетчик тактовых импульсов (частоты $f_{\text{т}}$) и точка анализа синхронизации в кодовом слове смещается на один символ.

Таким образом, в зависимости от принимаемого решения устройство цикловой синхронизации находится либо в режиме поиска синхронизации, либо в установившемся режиме контроля за появлением символов синхросигнала на синхропозиции цикла.

Важной характеристикой УЦС является среднее время установления синхронизма T'_n (время поиска). Процесс поиска позиции, на которой передается синхросигнал, продолжается до тех пор, пока единица не повторится на проверяемой позиции требуемое число раз l на интервале анализа в g циклов. Величины l и g , по которым принимается решение о наличии циклового синхронизма, зависят от вероятности появления символа 1 на информационных позициях цикла, вероятности искажения синхросимволов $P_{\text{ош}}$ и заданных характеристик принятия решения. Используя известные формулы для обнаружения двоичных

режим можно обнаружить, просматривая все временные позиции одновременно или последовательно. Выбор того или иного метода передачи и обработки цикловой синхроинформации при реализации УЦС определяется необходимостью быстрого обеспечения синхронизма, высокой помехоустойчивости, минимального объема синхроинформации в цикле, а также простоты реализации УЦС.

Наиболее просто реализуется ЦС при передаче относим-

сигналов, можно определить вероятности ложного синхронизма $P_{л.с}$ и пропуска синхронизма $P_{п.с}$:

$$P_{л.с} = n \sum_{i=l}^g C_g^i P^i (1-P)^{g-i}, \quad (9.31)$$

$$P_{п.с} = 1 - \sum_{i=l}^g C_g^i (1-P_{ош})^i P_{ош}^{g-i}, \quad (9.32)$$

где n — число символов в кодовом слове.

Для равновероятного появления единичных и нулевых символов на информационных позициях цикла $P_{л.с} = \frac{n}{2^n} \sum_{i=l}^g C_g^i$. Если значения $P_{л.с}$ и $P_{п.с}$ заданы, то, решая систему уравнений, можно определить минимальное значение g_{\min} и соответствующий ему оптимальный порог $l_{\text{опт}}$, а затем вычислить среднее время поиска

$$\bar{T}_n = \frac{(n-1)^2}{2} T_c g. \quad (9.33)$$

Система уравнений (9.31) и (9.32) решается только численными методами. Определение характеристик УЦС существенно упрощается, если заданы \bar{T}_n и $P_{л.с}$. Тогда по формулам (9.33) и (9.31) можно определить требуемый объем выборки g и порог l , а затем из формулы (9.32) найти вероятность пропуска синхронизма $P_{п.с}$.

Основой для построения УЦС, использующих синхронизирующие свойства кодов, является то, что при отсутствии цикловой синхронизации вероятность появления обнаруживаемой избыточным (помехоустойчивым) кодом ошибки значительно больше, чем при синфазной работе. Если на приемной стороне СПИ декодер, проверяя правильность поступающих на него кодовых слов, установит, что число искаженных кодовых слов превысит пороговое значение, то управляющее устройство УЦС переключится в режим поиска циклового синхронизма. В этом режиме управляющее устройство дискретно (на один такт за каждый цикл) изменяет момент начала записи в декодер принятого кодового слова. Как только слово будет записано правильно (от начала до конца), обнаружение ошибок прекращается и на выходе декодера сформируется импульс, блокирующий работу управляющего устройства. Если число правильно принятых слов превысит соответствующее пороговое значение, то УЦС выйдет из режима поиска.

В дальнейшем наличие цикловой синхронизации будет проверяться по правильности приема информации.

Подобный способ фазирования можно реализовать в СПИ, где для обнаружения ошибок используется код, пригодный для синхронизации. Пригодным кодом для синхронизации считается такой, у которого вероятность появления разрешенной комбинации в последовательности из символов, входящих в два соседних кодовых слова (пересечение двух слов), очень мала. Например, при передаче сообщений вида

$$\overbrace{b_1^{(i)}, b_2^{(i)}, \dots, b_j^{(i)}, \dots, b_n^{(i)}}^{\text{1-е слово}} \overbrace{b_1^{(i+1)}, b_2^{(i+1)}, \dots, b_j^{(i+1)}, \dots, b_n^{(i+1)}}^{\text{2-е слово}},$$

пересечение двух слов

вероятность появления разрешенной комбинации в последовательности символов $a_1^{(i)}, \dots, a_n^{(i)}, b_1^{(i+1)}, b_2^{(i+1)}, \dots, b_j^{(i+1)}$ должна быть очень малой. Код, у которого эта вероятность равна нулю, называется кодом «без запятой». Ограничения при использовании рассмотренного метода цикловой синхронизации связаны с большой избыточностью самосинхронизирующихся кодов.

В многоканальных РСПИ с временным уплотнением каналов необходимо передавать информацию о начале и конце кадра. Это обеспечивается устройством кадровой синхронизации (УКС). При наличии цикловой синхронизации для синхронизации кадров используется одно определенное кодовое слово, передаваемое в начале кадра. Кадровое синхрослово (КСС) по своей структуре должно существенно отличаться от всех возможных кодовых комбинаций, образуемых при передаче дискретной информации, и обеспечивать наилучшие условия его поиска и обнаружения в информационной последовательности символов даже при наличии искажений принимаемых посылок. Если цикловой синхронизации нет, то неопределенность возрастает, и тогда для обеспечения этого условия под КСС должна выделяться довольно значительная часть кадра (иногда до десяти и более процентов от общего числа символов в кадре). Это объясняется тем, что к помехоустойчивости кадровой синхронизации многоканальных СПИ предъявляются высокие требования, так как нарушение синхронизма по кадрам влечет за собой потерю связи во всех каналах системы.

Для выделения КСС в приемнике СПИ используется дискретный согласованный фильтр (ДСФ), настроенный на КСС. В него поочередно записываются принимаемые кодовые слова и в момент превышения выходным напряжением порога выделяется импульс кадровой синхронизации. Используя повторяемость импульсов синхронизации кадров, можно, накапливая их, увеличить помехоустойчивость УКС.

Во избежание ложного выделения КСС из информационной последовательности посылок необходимо предъявлять определенные требования к

структуре и регулярности его повторения. Желательно, чтобы КСС или вообще в информационном сигнале не встречалось, или его появление в нем было маловероятным. Если информационный сигнал кодируется безыбыточным кодом и источник информации может выдавать все кодовые слова с равной вероятностью, то единственным отличием КСС от информационного кодового слова может быть регулярность его появления на одних и тех же позициях кадра. Когда повторение одних и тех же кодовых слов в информационном сигнале имеет высокую вероятность, синхросигнал образуют путем передачи двух чередующихся КСС. Например, в одном кадре посылается заданное КСС, а во втором инверсное ему, в третьем кадре снова повторяется первоначальное КСС и т. д. Инверсное КСС образуется заменой всех единиц нулями, а нулей единицами.

Перспективным методом кадровой синхронизации является метод, при котором специальные кодовые группы используются для синхронизации как кадров, так и слов. В этом случае для синхронизации СПИ тратится наименьшее число посылок в кадре. Однако требования к длине и, главным образом, к структуре КСС существенно ужесточаются. Действительно, в рассматриваемой ситуации УКС должно выделить КСС на основе анализа всей поступающей на его вход последовательности символов, в которой теперь уже не известны границы отдельных кодовых слов.

Для уменьшения времени установления кадровой (и одновременно цикловой) синхронизации необходимо, чтобы вероятность ложного появления КСС в принимаемой последовательности посылок была минимальна. Вероятность ложного фазирования определяется длиной (числом разрядов) и структурой КСС, а также числом информационных посылок в передаваемом сообщении.

Определим основные требования к длине и структуре КСС, если цикловой синхронизации нет. Предположим, что КСС состоит из n разрядов a_1, a_2, \dots, a_n . Тогда анализируемая кодовая комбинация будет иметь вид

$$a_1, a_2, \dots, a_n, b_1^{(i)}, b_2^{(i)}, \dots, b_n^{(i)}, b_1^{(i+1)}, b_2^{(i+1)}, \dots,$$

где $b_j^{(i)}$ и $b_j^{(i+1)}$ — информационные разряды i - и $(i+1)$ -й кодовых комбинаций.

В процессе скользящего посимвольного поиска ложное выделение КСС может быть как из информационной последовательности $b_1^{(i)}, b_2^{(i)}, \dots, b_n^{(i)}$, так и из последовательности, полученной на пересечении КСС и информационного слова $a_j, a_{j+1}, \dots, b_1^{(i)}, b_2^{(i)}, \dots, b_j^{(i)}$. Вероятность ложного выделения КСС из последовательности информационных посылок определяется длиной КСС и числом информационных посылок, а вероятность ложного выделения КСС из последовательности, получаемой на пересечении КСС и информационно-

го слова, зависит также и от структуры КСС. Специальным выбором структуры КСС можно добиться того, чтобы эта вероятность была достаточно малой. Кодовые последовательности, удовлетворяющие этому требованию, должны иметь корреляционные функции с низким уровнем боковых выбросов. В качестве КСС в настоящее время широко используются М-последовательности, формируемые генераторами, включающими в себя сдвигающие регистры, охваченные логической обратной связью.

Как было сказано, вероятность выделения КСС из информационной последовательности определяется числом посылок в n КСС и числом слов в кадре $k_{\text{сл}}$. Определим требуемую длину n КСС в зависимости от $k_{\text{сл}}$ при использовании в качестве обнаружителя дешифратора. Поскольку вероятность появления 0 и 1 на любой позиции одинакова и равна 0,5, то при длине КСС в n посылок вероятность ложного синхронизма по кадрам при одном анализе определяется следующим образом:

$$P_{\text{л.с}}(1) = 1/2^n.$$

Вероятность ложной синхронизации и вероятность пропуска синхросигнала при просмотре кадра определяются соотношениями

$$\begin{aligned} P_{\text{л.с}} &= 1 - (1 - P_{\text{л.с}}(1))^{k_{\text{сл}}} \approx k_{\text{сл}} / 2^n, \\ P_{\text{п.с}} &= 1 - (1 - P_{\text{ош}})^n \approx n P_{\text{ош}}. \end{aligned}$$

Ложное фазирование приводит к ошибочному приему всей информации, поэтому с точки зрения уменьшения вероятности ошибки за счет фазирования значение n должно быть как можно больше. Однако при длинном КСС увеличивается избыточность кадра, а следовательно, ухудшается использование пропускной способности СПИ. Поэтому необходимое значение n следует выбирать с учетом двух важных характеристик: требуемой достоверности и эффективности использования пропускной способности системы.

При достаточной длине КСС, когда требуемая вероятность $P_{\text{л.с}}$ обеспечивается с запасом, можно снизить порог l в решающей схеме, т. е. отказаться от дешифратора. В этом случае характеристики обнаружения можно рассчитать по формулам, аналогичным (9.31) и (9.32):

$$\begin{aligned} P_{\text{л.с}} &\approx \frac{k_{\text{сл}}}{2^n} \sum_{i=l}^n C_n^i, \\ P_{\text{п.с}} &= 1 - \sum_{i=l}^n C_n^i (1 - P_{\text{ош}})^i P_{\text{ош}}^{n-i}. \end{aligned}$$

При больших значениях n пользоваться этими формулами трудно. Поэтому, переходя при $n \rightarrow \infty$ от биномиального закона к гауссовскому, можно записать

$$P_{\text{л.с}} = 1 - \Phi\left(\frac{l}{\sqrt{n}}\right),$$

$$P_{\text{п.с}} = 1 - \Phi\left(\frac{n(1 - 2P_{\text{ош}}) - l}{2\sqrt{nP_{\text{ош}}(1 - P_{\text{ош}})}}\right).$$

В ряде СПИ цикловую и кадровую синхронизацию обеспечивают передачей в начале сообщения специального синхросигнала, называемого *командой фазового пуска*. Такой режим оправдан, когда длительность передаваемого сообщения ограничена и потеря даже части сообщения недопустима. Поскольку в точке приема сведения о начале сообщения могут быть весьма ограниченными, то для его надежного определения команда фазового пуска должна быть достаточно продолжительной. Это требование противоречит простоте реализации устройства обработки сигнала фазового пуска, память которого не может быть, по крайней мере, меньше числа символов сигнала. Выход здесь находят в использовании специальных сигналов и устройств их обработки. Если каждый символ команды фазового пуска связан с группой предыдущих рекуррентным правилом, то, приняв правильно группу символов, можно воссоздать весь сигнал и определить его начало и конец. Генерируя таким образом сигнал, на приемной стороне с помощью корреляционного приемника можно с высокой достоверностью определить его совпадение по фазе с принимаемым сигналом и тем самым перепроверить первоначально принятое по группе символов решение. В качестве рекуррентных последовательностей применяют М-последовательности, а сам метод назван *методом последовательной оценки*.

9.7.5. Синхронизация модемов с широкополосными сигналами

Оптимальный алгоритм обработки широкополосных сигналов (ШПС) при синхронизации не имеет каких-то принципиальных отличий от алгоритма обработки простых сигналов. Для оценки времени запаздывания следует использовать метод максимального правдоподобия. Если помеха представляет собой гауссовский случайный процесс с равномерной спектральной плотностью, то максимум функции правдоподобия совпадает с максимумом модуля взаимной корреляционной функции принимаемого и опорного сигналов $R(\tau)$. Если для передачи символов сообщения используют сигналы разной формы, то число каналов вычисления взаимной корреляционной функции должно быть равно основанию кода.

При наличии помех положение максимума ВКФ становится случайным и его определение сопровождается ошибками. Ошибки могут быть двух видов: нормальными, когда их значение не превышает ширины пика

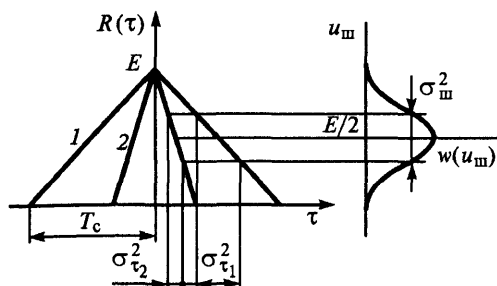


Рис. 9.56. Диаграммы, поясняющие принципы оценки точности определения временного положения простого (1) и сложного (2) сигналов

ной функцией и отношением сигнал—шум на выходе СФ. Идеальными сигналами для синхронизации следует считать такие, у которых боковые выбросы корреляционной функции невелики, например периодические М-последовательности.

Оценим приближенно, как влияют нормальные и аномальные ошибки на точность синхронизации сложных сигналов с идеальными корреляционными функциями. Временное положение выходного сигнала СФ с наибольшей достоверностью можно оценить по положению точки, где крутизна $R(\tau)$ максимальна (рис. 9.56). Для сигналов с треугольной формой корреляционной функции крутизна постоянна и равна EB/T_c , где B — база сложного сигнала. Поэтому координата этой точки выбрана равной $E/2$. Флуктуации шума с законом распределения $w(u_{\text{ш}})$ и дисперсией $\sigma_{\text{ш}}^2 = EN_0/2$ приведут к погрешностям измерения временного положения с дисперсией

$$\sigma_{\tau}^2 \approx T_c^2 / (2h^2 B^2). \quad (9.34)$$

Этот результат практически совпадает с оценками, получаемыми при точном анализе.

Теперь оценим аномальные ошибки. На интервале T_c выходное напряжение СФ можно представить во времени B независимыми отсчетами, один из которых принадлежит пику корреляционной функции, а остальные $(B - 1)$ — шуму. Тогда для сигналов с идеальной формой КФ при вынесении решения о наличии пика в каждой позиции вероятность ложного синхронизма $P_{\text{л.с}}(1)$ численно равна вероятности ошибки при различении двух ортогональных сигналов. Поскольку аномальные ошибки могут возникнуть независимо в любой из $B - 1$ точек, то их результирующая вероятность имеет вид

$$P_{\text{ан}} = 1 - (1 - P_{\text{л.с}}(1))^{B-1} \approx (B - 1)P_{\text{л.с}}(1),$$

а дисперсия определяется формулой

$$\sigma_{\text{ан}}^2 = T_c^2/12.$$

Дисперсию ошибок синхронизации с учетом нормальных и аномальных ошибок можно записать в виде

$$\sigma_{\Sigma}^2 = P_{\text{ан}}\sigma_{\text{ан}}^2 + (1 - P_{\text{ан}})\sigma_{\tau}^2. \quad (9.35)$$

Подставив в (9.35) выражение для вероятности аномальных ошибок

$$P_{\text{ан}} \approx \frac{B-1}{2} \exp\left(-\frac{h^2}{2}\right)$$

и приняв во внимание, что в реальных ситуациях $p_{\text{ан}} \ll 1$, окончательно получим

$$\sigma_{\Sigma}^2 \approx T_c^2 \left(\frac{B-1}{24} \exp\left(-\frac{h^2}{2}\right) + \frac{1}{2h^2 B^2} \right).$$

С увеличением базы сигнала B и при фиксированном h^2 нормальные ошибки уменьшаются, а аномальные растут. Нетрудно найти значение $B_{\text{опт}}$, при котором достигается минимум дисперсии ошибки:

$$B_{\text{опт}} = \sqrt[3]{\frac{24}{h^2} \exp\left(\frac{h^2}{2}\right)}.$$

Например, для $h^2 = 2$ база $B_{\text{опт}} = 3$, а для $h^2 = 24$ база $B_{\text{опт}} = 50$.

Когда сложные сигналы обрабатываются в корреляторах, синхронизацию выполняют в два этапа. Сначала, изменяя временное положение опорного сигнала, последовательно просматривают область неопределенности и грубо находят положение центрального пика взаимной корреляционной функции (этап поиска), а затем следящим устройством уточняют положение его максимума (этап слежения). Центральный пик необходимо искать с шагом Δt , не превышающим $0,5/F_c$. При этом исключается возможность его пропуска в отсутствие помех. Общее время синхронизации складывается из времени поиска $T_{\text{п.с}}$ и слежения $T_{\text{сл}}$. Дисперсию ошибок слежения можно приближенно оценить по формуле (9.34), подставив $h^2 = P_c T_{\text{сл}}/N_0$.

При неподвижных относительно друг друга приемнике и передатчике и высокой стабильности генераторов опорных частот требуемая точность σ_{τ}^2 достигается соответствующим выбором времени слежения $T_{\text{сл}}$.

Этап поиска характеризуется продолжительностью $T_{\text{п.с}}$ и вероятностью его правильного завершения $P_{\text{прав}} = 1 - P_{\text{л.с}}$. Существующие процедуры

поиска можно условно разбить на два класса: с фиксированным временем анализа на каждой позиции и с переменным. Минимальное время поиска можно получить для переменного времени анализа в каждой позиции (последовательная процедура оценки). Реализуется этот метод наиболее просто для схемы принятия решения с двумя порогами u_{n1} и u_{n2} ($u_{n1} < u_{n2}$). Если при анализе в течение времени T_a превышен порог u_{n1} , но не превышен u_{n2} , то наблюдение повторяется. Поиск считается завершенным, если превышен порог u_{n2} .

Пропуск синхронизма с вероятностью $P_{п.с}$ приводит к повторному просмотру области временной неопределенности, при этом продолжительность поиска оказывается случайной величиной. Среднее время поиска с учетом пропуска находится по формуле

$$\bar{T}_{п.с} = MT_a \left(1 + 2 \sum_{i=1}^{\infty} P_{п.с}^i \right) = MT_a \left(1 + \frac{2P_{п.с}}{1 - P_{п.с}} \right),$$

где M — среднее число шагов при поиске без пропусков.

При выборе шага $\Delta t = 0,5/F_c$ число $M \approx B$. Зная связь между вероятностью ложной синхронизации в одной точке анализа $P_{л.с}(1)$, вероятностью пропуска $P_{п.с}$ и временем анализа T_a :

$$\frac{P_c T_a}{N_0} = \sqrt{\ln \left(\frac{1}{P_{л.с}(1)} \right)} + \sqrt{\frac{1}{P_{п.с}} - 1}, 4,$$

можно найти зависимость продолжительности поиска от времени анализа T_a при фиксированных значениях B и $P_{л.с}(1)$. Для определения оптимального значения времени анализа целесообразно использовать численные методы или моделирование процедуры поиска. В качестве ориентировочной оценки воспользуемся результатами расчета времени поиска для процедуры с однократным просмотром области неопределенности и принятием решения в точке с максимальным значением напряжения на выходе коррелятора. В этом случае вероятность ложной синхронизации на одном шаге определяется следующим образом:

$$P_{л.с}(1) \approx \frac{1}{2} \exp \left(-\frac{P_c T_{п.с}}{2N_0 B} \right),$$

откуда несложно найти время поиска

$$T_{п.с} = \frac{4N_0 B}{P_c} \ln \left(\frac{1}{2P_{л.с}(1)} \right). \quad (9.36)$$

Формула (9.36) наглядно характеризует влияние отношения сигнал—шум P_c/N_0 , вероятности ложной синхронизации $P_{л.с}(1)$ и степени сложности сигнала B на продолжительность поиска $T_{п.с}$.

Устройство слежения (рис. 9.57) по принципу работы подобно устройству тактовой синхронизации. Оно включает в свой состав дискриминатор (Д), определяющий рассогласование по времени между входным и опорным сигналами, устройство управления временным положением опорного сигнала (УУ) и низкочастотный фильтр $K(p)$ цепи обратной связи. Для получения дискриминационной характеристики $u_d(\Delta\tau) = K\Delta\tau$ (при малых значениях $\Delta\tau$), где $\Delta\tau$ — временное рассогласование между входным и опорными сигналами, определяется разность напряжений с выходов двух корреляторов (К), на которые поданы опорные сигналы со сдвигом по времени. Временной сдвиг $\Delta\tau_c$ чаще всего выбирают так, чтобы обеспечить максимальную крутизну дискриминационной характеристики, т. е. минимальную ошибку слежения.

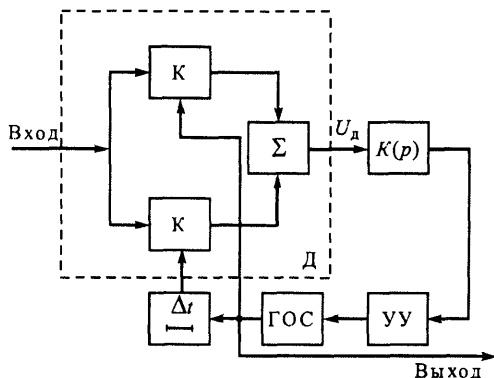


Рис. 9.57. Структурная схема устройства слежения за временным положением сложного сигнала

9.8. Перспективные системы передачи информации

9.8.1. Системы наземной подвижной связи

В последние годы во всем мире особенно быстрыми темпами развивается связь с подвижными объектами. Все большее распространение получают системы связи общего пользования, которые представляют абонентам, находящимся в движении, комплекс услуг: от образования диспетчерских и технологических служб отдельных ведомств до автоматического соединения с абонентами городской телефонной сети, а по междугородным и международным линиям связи — с абонентами других городов и стран.

Системы подвижной радиосвязи общего пользования до настоящего времени служили дополнением городских телефонных сетей, так как по сравнению с числом абонентов стационарных сетей число абонентов подвижной связи было сравнительно невелико. Однако если число абонентов телефонных сетей в промышленно развитых странах имеет тенденцию к

стабилизации, то число абонентов систем связи с подвижными объектами непрерывно растет и темпы роста их с каждым годом увеличиваются. В перспективе произойдет полное слияние сетей радиосвязи с подвижными объектами с проводными телефонными сетями в единую объединенную информационную сеть радиопроводной персональной связи.

Развитие систем подвижной радиосвязи общего пользования шло по двум направлениям: создание систем с большими зонами обслуживания (радиально-зоновых систем) и малыми, как правило, с «сотовой» структурой — сотовых систем подвижной связи (ССПС). Благодаря ряду преимуществ ССПС перед радиально-зоновыми системами они получили широкое внедрение в практику. Одно из главных преимуществ состоит в том, что они могут обеспечить высокую надежность и качество связи по всей рабочей зоне при значительно меньших мощностях передатчиков как центральных (базовых), так и абонентских станций (АС). Это объясняется тем, что в ССПС каждая базовая станция (БС) обслуживает небольшие по площади зоны — соты, на всей территории которых отношение сигнал — шум на входах приемников АС и БС, необходимое для обеспечения требуемого качества передачи сообщения, может быть получено при сравнительно малых мощностях передатчиков этих станций. В радиальных системах, где центральная станция (ЦС) обслуживает зону большой площади (радиусом 50...100 км), то же качество связи обеспечивается при большей мощности передатчиков. Другое важное достоинство ССПС — это экономия частотного ресурса вследствие повторного использования одних и тех же частот в различных зонах территории, обслуживаемой системой, удаленных друг от друга на достаточные расстояния.

Число абонентов ССПС зависит от пропускной способности системы связи между абонентскими и базовой станциями в пределах одной соты и числом БС. Последнее равно числу сот, которое при фиксированной площади территории обслуживания возрастает по квадратичному закону с уменьшением радиуса зоны (расчеты и практическое использование ССПС показывают, что радиус соты должен быть не менее 1,5 км, в противном случае усложняется процесс эстафетной передачи при движении объекта). С уменьшением радиуса рабочей зоны и, соответственно, мощности передатчиков АС и БС значительно снижается биологическое воздействие системы на живые организмы, улучшается электромагнитная совместимость (ЭМС) абонентов в ССПС и ЭМС между ССПС и другими системами, использующими соседние участки радиоспектра, а также снижаются стоимость и габаритные размеры абонентских станций.

Важная особенность ССПС заключается в возможности ее создания без значительных начальных капитальных затрат. Сначала ССПС развертывается с крупными рабочими зонами (радиусом примерно 10 км) и относи-

тельно небольшим числом абонентов. По мере роста числа заявок на АС размеры сот уменьшаются и увеличивается общее число абонентов. При этом постоянно наращивается объем типового оборудования БС и центров коммутации в результате доходов от использования ССПС действующими абонентами. Поэтому первоначальные капитальные вложения могут быть значительно меньше полных затрат, приходящихся на максимальное число абонентов.

В настоящее время в эксплуатации находятся ССПС нескольких поколений. Системы первого поколения (1G), использующие диапазон частот 450 МГц, характеризуются применением аналоговых методов передачи информации по радиоканалу (в большинстве случаев используют частотную модуляцию), крупными рабочими зонами и относительно небольшим числом абонентов (десятки и сотни тысяч). Для возможности увязки работы ССПС, существующих в различных странах и отличающихся тактико-техническими характеристиками и принципами управления, разработаны стандарты на эти системы в рамках Международного союза электросвязи. Один из них — NMT-450 — принят в России как Федеральный стандарт на аналоговые ССПС.

Ограниченная номенклатура предоставляемых абоненту услуг и незначительные возможности увеличения пропускной способности ССПС первого поколения привели к созданию цифровых ССПС второго поколения (2G), разработку принципов которых осуществляли на основе ряда стандартов. Так, в Европе группой экспертов по подвижной связи (Group Special Telephonic Mobile—GSM) создан общеевропейский стандарт цифровой ССПС — стандарт GSM, который принят в России в качестве Федерального стандарта на цифровые ССПС.

Отличительными характеристиками второго поколения ССПС являются освоение новых частотных диапазонов (900 и 1800 МГц), большая пропускная способность и большее число абонентов при малых зонах (радиусом 2...5 км), применение цифровых методов передачи информации как для радио, так и для проводных каналов, использование многостанционного доступа с временным разделением, высокое качество приема информации вследствие применения корректирующих кодов, автоматический поиск и вызов подвижных абонентов на территории обслуживания, динамическое управление и распределение каналов по территории обслуживания в зависимости от трафика и расположения абонентов.

Применение цифровых методов передачи информации позволило существенно увеличить число предоставляемых абоненту ССПС услуг. Кроме телефонных он сможет принимать телексные и факсимильные сообщения различного рода, графическую информацию (планы местности, графики движения и т. п.), медицинскую информацию и многое другое. Особенно

важной является возможность подвижного абонента осуществлять обмен данными, входя в сети ЭВМ.

Принципы построения цифровых ССПС позволили применить при организации сотовых сетей новые, более эффективные модели повторного использования частот, чем в аналоговых сетях. В результате без увеличения общей полосы частот системы связи значительно возросло число каналов на соту. В первую очередь, сказанное относится к стандарту GSM. Вид модуляции, способы кодирования и формирования сигналов в каналах связи, принятые в GSM, обеспечивают прием сигналов с отношением сигнал—помеха, равным примерно 9 дБ, в то время как в аналоговых системах тот же показатель равен 18 дБ. Поэтому передатчики базовых станций, работающие на совпадающих частотах, могут размещаться в более близко расположенных сотах без потери высокого качества приема сообщений. Эффективным способом снижения уровня помех является использование секторных антенн. При этом уровень излучения в направлении, противоположном направлению излучения сигнала, сокращается до минимума. Секторизация сот позволяет более часто повторно применять частоты в сотах при одновременном снижении уровня помех. С помощью секторных антенн и повторного использования частот можно увеличить емкость сетей цифровых ССПС стандарта GSM примерно на 40%.

Внедрение ССПС второго поколения стало новым этапом развития мобильной связи. Оно вызвало доселе невиданный рост популярности портативных мобильных телефонов и превратило подвижную связь в социальный фактор, оказывающий существенное влияние на образ жизни миллионов людей планеты.

Динамика роста абонентской базы продолжает превышать все самые смелые прогнозы: в отдельных странах она достигает нескольких миллионов абонентов в год.

Однако следует отметить, что стандарты второго поколения несовместимы между собой. Это обстоятельство создает для пользователей много проблем. Так, например, невозможно воспользоваться привычным сотовым телефоном и получить привычные услуги по приезду из США в Европу (и обратно). Поэтому одна из самых больших проблем, связанная с сегодняшними цифровыми подвижными системами, это имеющая место конкуренция несовместимых между собой стандартов в различных странах мира. В то время как Европа, часть Азии и стран Тихого океана являются приверженцами стандарта GSM, в США используется система PCS в диапазоне 1,9 ГГц с применением объединенных стандартов CDMAOne (или IS-95), D-AMPS (или IS-136) и GSM 1900 в зависимости от оператора. Япония также имеет свой собственный стандарт — PDC. Ни одна из этих систем не может взаимодействовать с другой системой. Кроме того, несмотря на многообразие и

высокое качество предоставляемых услуг, сотовые системы подвижной связи второго поколения не способны удовлетворять всем требованиям, предъявляемым к сетям подвижной связи. К ним, в первую очередь, относятся:

- *глобальная мобильность* (подвижные абоненты должны иметь возможность перемещаться без каких-либо ограничений и при этом иметь доступ к привычному набору услуг);

- *качество передачи речи* (оно должно соответствовать качеству передачи речевых сообщений стационарных телефонных сетей, что может быть обеспечено применением высокоэффективных вокодеров с адаптивной скоростью);

- *емкость сетей* (быстро растущая потребность в услугах подвижной связи и ограниченность частотного ресурса определяют повышенные требования к возможности перспективных ССПС обслуживать большое число абонентов на ограниченной территории);

- *высокоскоростная передача данных* (стремительный рост вычислительных сетей разного уровня, в частности Internet, появление новых приложений выдвигают в качестве одного из главных требований возможность передачи мультимедийного трафика).

Перечисленные требования не могут быть удовлетворены без радикального изменения радиоинтерфейсов и реализации дополнительных функций сетевого взаимодействия, что не может быть сделано в рамках существующих технологий сетей подвижной связи второго поколения.

Необходимость выработки единых глобальных подходов к построению ССПС третьего поколения (3G) побудила Международный союз электросвязи (МСЭ) к попытке создания единого стандарта для «Будущей наземной системы подвижной связи общего пользования», переименованной в дальнейшем в IMT-2000 (International Mobile Telecommunications — Международная подвижная связь), где число 2000 символически указывает используемый диапазон частот (2000 МГц). Однако при переходе от этапа разработки серии концептуальных рекомендаций к созданию конкретных спецификаций стало очевидно, что коммерческие интересы различных региональных производителей аппаратуры и операторов связи практически невозможно объединить в рамках единого стандарта, и была выдвинута «концепция семейства» систем третьего поколения, членами которого могут стать региональные и национальные стандарты, отвечающие ряду обязательных требований по их взаимной совместимости и обеспечению глобального роуминга.

Внедрение систем IMT-2000 положит начало совместному существованию ССПС второго и третьего поколений. В этот период будут постепенно вытесняться системы предыдущего поколения, в первую очередь, за пределы территорий с наивысшей плотностью абонентов, т. е. за пределы

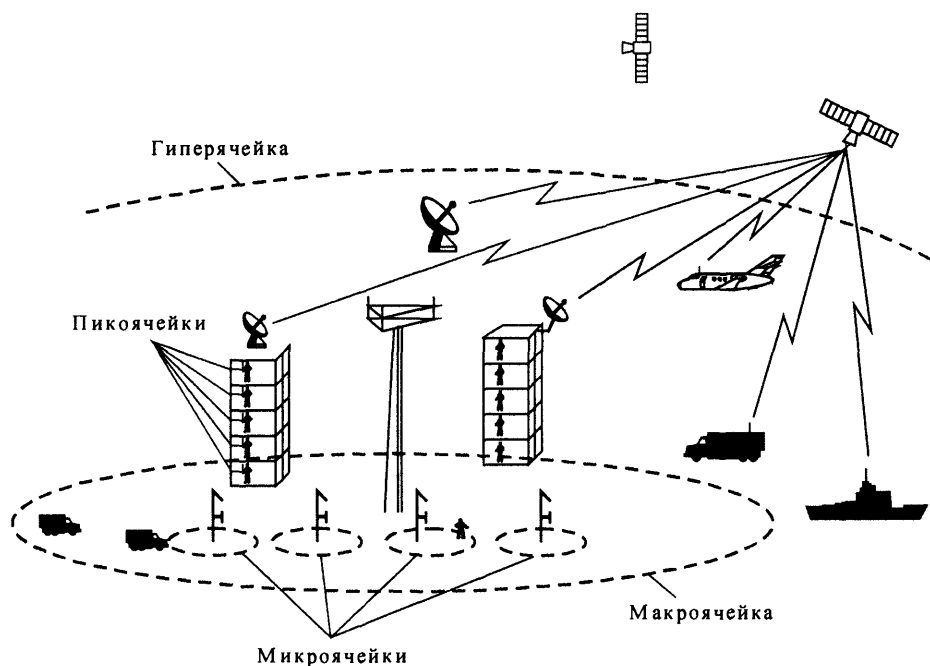


Рис. 9.58. Структура сотовой сети подвижной связи (ССПС) третьего поколения

мегаполисов. Переходный процесс может растянуться на годы, в течение которых будет происходить также дальнейшее развитие ССПС второго поколения.

Структура сотовой сети подвижной связи третьего поколения представлена на рис. 9.58. Она является иерархической. Внутри помещений организуются пикоячейки радиусом до 100 м с очень высокой пропускной способностью, определяемой большой плотностью абонентов на единицу обслуживаемого пространства. В общественных местах и пешеходных зонах создаются микроячейки с радиусом обслуживания до 1 км. Сотовые системы, обслуживающие пользователей, передвигающихся на автомобилях, оперируют макроячейками радиусом до нескольких десятков километров. Наконец, автомобильный и железнодорожный транспорт, воздушные, морские и речные суда, распределенные по территории с малой плотностью абонентов, обслуживаются спутниковой составляющей будущей системы с использованием гиперячеек радиусом до сотен и тысяч километров.

С учетом того, что для реализации такой системы необходима весьма широкая полоса частот, Всемирная административная конференция по Радио (WARC-92) в 1992 г. распределила для будущей системы сухопутной

подвижной связи общего пользования на всемирной основе полосы частот 1885...2025 МГц и 2110...2200 МГц, включая полосы 1980...2010 МГц и 2070...2200 МГц для ее спутниковой составляющей.

Серьезной проблемой при использовании иерархической сотовой структуры (от пико до гиперсот) является выбор способа радиодоступа к будущей сети подвижной радиосвязи. Предложенные технологии построения радиointерфейса базируются на конкурирующих методах многостанционного доступа — с временным (TDMA) и кодовым (CDMA) разделением.

Сторонниками метода TDMA являются производители и операторы систем стандартов GSM и D-AMPS (IS-136), получивших к настоящему времени наиболее широкое распространение во всем мире. Эволюционное развитие этих стандартов позволит создать системы подвижной связи, удовлетворяющие таким важным системным требованиям стандартов IMT-2000, как высокое качество связи, более высокая емкость, малая мощность передаваемого сигнала, высокая скорость передачи (от 144 кбит/с для быстро перемещающихся объектов и до 2 Мбит/с внутри зданий в интерактивном режиме — от базовой станции к абоненту). Такое улучшение характеристик ССПС может быть достигнуто применением в радиointерфейсах систем методов модуляции с высокой спектральной эффективностью (например, 8-фазной ФМ) более совершенных методов помехоустойчивого кодирования, а также усовершенствованных протоколов пакетной передачи данных. Таким образом, два основных современных стандарта на основе TDMA будут развиваться по единому сценарию, предоставляя услуги систем третьего поколения. При этом сохраняется возможность применения существующей инфраструктуры сетей GSM и IS-136 и уже используемых частотных диапазонов.

Реализация метода CDMA в виде стандартов на оборудование цифровых систем сотовой подвижной связи вызвала жаростные споры между сторонниками и противниками технологии, которые постепенно затихали по мере внедрения в мире реально действующих CDMA систем подвижной связи и абонентского радиодоступа. Сегодня системы CDMA прочно заняли свое место среди современных технологий радиосвязи.

Быстрое увеличение числа абонентов беспроводных сетей связи и рост потребности в высокоскоростных услугах передачи данных обеспечили методу CDMA возможность претендовать на роль базового метода радиодоступа в системах подвижной связи третьего поколения. Этот метод позволяет легче реализовать требования стандарта IMT-2000.

Первым CDMA стандартом явился IS-95, разработанный фирмой QUALCOMM. Системы этого стандарта появились на рынке в 1995 г. Одновременно фирма QUALCOMM предоставила лицензии на выпуск оборудования на базе стандарта IS-95 многим крупным производителям, в число которых вошли Lucent, LG, Motorola, NEC, Nortel, Samsung и другие компании.

Привлекательность систем, построенных на базе технологии CDMA с использованием широкополосных шумоподобных сигналов (ШПС), объясняется рядом присущих им достоинств [15, 135]. Одним из них является высокая помехоустойчивость при действии произвольных помех с конечной энергией (ограниченных по спектру). Другим важным достоинством технологии CDMA является способность эффективного функционирования в условиях многолучевого распространения сигнала за счет раздельной обработки принимаемых лучей с последующим их объединением, что позволяет добиться увеличения отношения сигнал—шум на выходе приемника.

Раздельная обработка лучей позволяет также реализовать надежную эстафетную передачу абонентской станции от одной базовой станции к другой, применяя так называемую «мягкую» передачу управления (soft handoff). Во время этой процедуры абонентская станция поддерживает связь одновременно с двумя (или тремя) базовыми станциями, что увеличивает вероятность нормального завершения переключения с одной БС на другую. В системах, базирующихся на других стандартах, АС сначала завершает обмен сообщениями с БС одной соты и лишь затем устанавливает связь с новой БС.

Применение технологии CDMA обеспечивает более высокую эффективность использования частотного спектра в ССПС по сравнению с той, которая достигается при частотном (FDMA) и временном (TDMA) разделении каналов.

Поскольку все БС работают на одной несущей частоте, т. е. фактор повторного использования частот равен 1, в такой системе не требуется частотное планирование. Это упрощает как начальное развертывание системы CDMA, так и ее последующее развитие. Вместе с тем при проектировании системы должен обеспечиваться определенный баланс мощности сигналов (планирование мощности), чтобы уровень взаимных помех в системе не превышал допустимого.

В связи с тем, что в системе не выделяется заранее определенный частотный (временной) ресурс для организации канала связи между базовой и абонентской станциями, система обладает свойствами эластичности. Благодаря этому обеспечивается возможность динамического перераспределения ресурсов, что является более простой задачей по сравнению с динамическим перераспределением частотных каналов.

Для ССПС с CDMA характерна повышенная конфиденциальность обмена сообщениями. В системе каждому абоненту присваивается ШПС, имеющий индивидуальную и достаточно сложную структуру.

Важное свойство CDMA заключается в уменьшении средней излучаемой мощности, что достигается благодаря эффективному управлению ею. Поэтому обеспечивается и хорошая электромагнитная совместимость с дру-

гими системами радиосвязи. Низкая средняя излучаемая мощность системы CDMA позволяет также минимизировать уровень электромагнитных излучений, воздействующих на человека, что снижает биологическую опасность системы. По тем же причинам маловероятно непреднамеренное воздействие систем CDMA на работу различного рода электронных устройств. Перечисленные достоинства технологии CDMA стали главным аргументом в борьбе за продвижение стандарта IS-95 на рынок ССПС.

Перечень услуг, предоставляемых оборудованием CDMA, постоянно расширяется. Вместе с тем базовые ограничения стандарта IS-95 (с полосой сигнала 1,25 МГц) не позволяют обеспечить в рамках сетей CDMA поддержку набора услуг, связанных с передачей мультимедиа трафика.

Сторонники метода CDMA предоставили в МСЭ целый ряд предложений по системам третьего поколения с кодовым разделением каналов [137]. Так, в Европе в качестве основы для стандарта UMTS выбраны две технологии радиодоступа: в режиме частотного дуплексного разноса будет использоваться широкополосная система CDMA (W-CDMA), а в режиме временного дуплексного разноса — гибридная широкополосная система с временным разделением сигналов (W-TDMA/CDMA). Соединенные Штаты представили в МСЭ несколько предложений. Среди них одним из основных является развитие персональной связи путем эволюционного развития стандарта IS-95 в стандарт cdma2000, отличающийся тем, что он обеспечивает обратную совместимость со стандартом IS-95.

Ассоциация радиопромышленности и бизнеса Японии ARIB, ответственная за разработку будущей системы связи, и Комитет по технологии электросвязи Японии планируют развертывание системы третьего поколения на основе системы W-CDMA. Ассоциация телекоммуникационных технологий Южной Кореи TTA также ориентирована на систему W-CDMA.

Остановимся на особенностях перечисленных вариантов.

Эволюция стандарта IS-95. В настоящее время ведутся работы по созданию перспективных версий этого стандарта. Наиболее существенным дополнением, предусмотренным в стандарте IS-95B, является увеличение верхней границы скорости передачи данных без изменения занимаемой полосы частот 1,25 МГц. Благодаря возможности объединения до восьми каналов трафика CDMA скорость передачи данных может достигать значений 115 кбит/с, причем производители оборудования смогут постепенно наращивать число объединяемых каналов. Так, компания QUALCOMM на начальном этапе предлагает использовать два или четыре объединенных канала в прямом направлении (от БС к АС), что обеспечит скорость 28,8 или 57,6 кбит/с соответственно, и один канал в обратном направлении. Этот выбор отражает несимметричную природу трафика, характерную для таких приложений, как электронная почта и доступ в сети Internet.

В последующей версии стандарта IS-95C предусматривается повышение частотной эффективности и емкости системы CDMA. Для этого используемый набор из 64 кодов Уолша будет дополнен группой из 64 кодов, передаваемых по квадратурному каналу. Несмотря на изменения, система сохранит обратную совместимость со стандартами IS-95A и B и будет занимать прежнюю полосу частот 1,25 МГц.

Новая версия — стандарт cdma2000 — еще больше расширит возможности по высокоскоростной передаче данных. Версия даст возможность организовать канал передачи данных с базовой скоростью 144 кбит/с, что достигается увеличением тактовой частоты «расширяющей спектр» псевдослучайной последовательности (ПСП) в 3 раза по сравнению со стандартом IS-95 ($3 \times 1,2288 \text{ МГц} = 3,6864 \text{ МГц}$). Возможно также дальнейшее увеличение скорости передачи в канале от БС к АС до 1 Мбит/с за счет объединения нескольких каналов трафика.

Система W-CDMA. Одним из предложений Европейского сообщества по стандарту третьего поколения является система W-CDMA, использующая для начала полосу частот ШПС 5 МГц с последующим расширением ее до 15 МГц. Европейский институт стандартов телекоммуникаций (ETSI) для предоставления нового варианта радиодоступа W-CDMA решил взять за основу базовой сети усовершенствованную коммутационную сеть GSM. Таким образом, система W-CDMA может сосуществовать с системой GSM и с помощью двухрежимных мобильных терминалов будет поддерживать полный роуминг и передачу управления от одной системы к другой. Использование двухрежимных терминалов на стадии внедрения W-CDMA даст абонентам возможность поддерживать связь с остальными пользователями GSM с самого начала внедрения.

Стандарт W-CDMA позволяет осуществлять комбинированные услуги на скорости передачи от 8 кбит/с до 384 кбит/с при полосе сигнала 5 МГц. Кроме того, можно комбинировать услуги с коммутацией пакетов и коммутацией каналов в одной и той же линии связи, обеспечивая, таким образом, предоставление действительно мультимедийных услуг. Также могут быть обеспечены услуги с различными требованиями к качеству передачи, например голосовая связь и пакеты данных.

Аналогично орган стандартизации Японии ARIB взял за основу ту же технологию W-CDMA, первоначально используя ширину полосы носителя 5 МГц. Кроме того, Япония и крупнейшая компания-оператор мобильной связи NTT DoCoMo решили использовать усовершенствованную базовую сеть GSM ETSI в качестве базовой сети третьего поколения.

В Корее принята широкополосная версия CDMA, называемая B-CDMA, разработанная компанией Samsung Electronics. В ее основе лежат те же ме-

тоды и алгоритмы преобразования сигналов, что и у CDMA (IS-95). Однако расширение спектра ШПС до 15 МГц позволило увеличить емкость системы, ее помехоустойчивость к внешним и внутрисистемным помехам, а также стойкость к интерференционным замираниям. Скорость передачи сообщений в системе В-CDMA может достигать 384 кбит/с.

Стандарт IMT-2000 охватывает широкий диапазон систем и применений от обычных сотовых сетей связи до персональной спутниковой телефонии и систем фиксированного радиодоступа, часто упоминаемых как Wireless Local Loop — беспроводная местная связь (WLL).

На начальной стадии постепенного внедрения систем третьего поколения будут применяться, как правило, мультирежимные переносные терминалы, предоставляющие возможность доступа к услугам 3G в местных сетях при сохранении нормальных функциональных возможностей систем второго поколения, когда пользователи, осуществляющие роуминг, попадают в зоны, где сети третьего поколения еще не функционируют. Таким образом, операторам будет предоставлено необходимое время для модернизации оборудования.

Внедрение систем третьего поколения окажет революционное влияние на использование сетей подвижной связи. Полностью цифровые системы 3G должны обеспечить пакетную передачу данных и полную совместимость с другими цифровыми устройствами, начиная от пультов игровых приставок и компьютеров и заканчивая цифровым телевидением и Internet.

Системы 3G должны (как уже указывалось выше) поддерживать скорости передачи данных до 2 Мбит/с. Это обстоятельство открывает широкий диапазон возможностей для пользователей новых служб: от быстрого доступа к сети Internet в реальном масштабе времени до проведения персональных видеоконференций с применением специально разработанных для этого телефонных трубок с большим экраном.

Европа быстро продвигается к своей готовности для внедрения следующего поколения подвижной связи. Здесь уже имеется четкий план реализации 3G, соответствующий рекомендациям, разработанным Европейской Комиссией и принятым Европейским Парламентом. В России ожидается внедрение систем третьего поколения не ранее 2004—2005 гг.

Дальнейшее расширение диапазона услуг подвижной связи станет возможным на основе систем стандартов четвертого поколения, важнейшим требованием при разработке которых будет увеличение скорости передачи до 10 Мбит/с. Достижение таких скоростей обеспечит возможность мобильному пользователю получения полного набора мультимедийных услуг с качеством, не уступающим тому, какое обеспечивается цифровыми сетями наземной связи.

9.8.2. Системы подвижной спутниковой связи

Следует учитывать, что в районах с низкой плотностью населения, в том числе в России, использование наземных сотовых радиосистем связи вне крупных городов будет крайне малоэффективно экономически. Поэтому возникает естественное предложение применять для этой цели спутниковые сети связи. Но энергетический баланс линий спутниковой связи до настоящего времени не позволял уменьшить абонентский терминал до размеров сотового телефона — трубки в руке [138].

Применение спутников на низких орбитах создает энергетические преимущества перед геостационарными спутниками и дает возможность организовывать сети подвижной связи с персональными телефонами, в которых используются ненаправленные антенны. Кроме того, запуск спутника на низкую орбиту проще и дешевле, для этого можно использовать ракеты средней мощности. Резко уменьшается запаздывание информационных сигналов в канале связи.

Однако имеется ряд недостатков, свойственных системам со спутниками на низких орбитах, в том числе:

- необходимость запуска большого числа спутников, даже если предполагается обслуживать ограниченную зону; затем эту группировку нужно поддерживать, заменяя вышедшие из строя спутники;

- необходимость организации большого числа соединений между зонами, для чего используются наземные или межспутниковые линии или геостационарные спутники;

- необходимость анализа и обработки сигналов на борту спутника; технологические сложности при создании бортового оборудования и абонентского терминала;

- сложность частотного совмещения с системами спутниковой связи, с геостационарными спутниками и наземными радиорелейными линиями.

Функции систем с низкоорбитальными спутниками достаточно широки. Такие системы подходят для радиоопределения (т. е. для определения местоположения объекта), пейджинга, электронной почты (с передачей сообщений с запоминанием; они ретранслируются в момент, удобный для приема на принимающей станции). Такие спутниковые системы удобны для организации технологической связи, особенно при передаче сообщений от диспетчеров к движущемуся средству. В отдельных случаях системы на низкоорбитальных спутниках можно использовать для общедоступной фиксированной связи с удаленными районами и передачи данных. Однако основное их применение — телефонная персональная связь с подвижным абонентом, находящимся вне зоны действия наземных сотовых систем.

С наземными сетями сотовой связи в зоне их действия спутниковые системы конкурировать не могут. Поэтому создаются абонентские терминалы — «трубки» двойного применения, способные работать как в наземной сети, так и через спутники.

Серьезной проблемой является выбор полос частот для фидерных линий, поскольку в этих полосах работают также системы фиксированной спутниковой службы с геостационарными спутниками, наземные радиорелейные линии и радиоастрономическая служба и др.

Полоса частот, выделенная для связи низколетящих спутниковых систем с подвижным абонентом, весьма ограничена (в диапазоне 1610...1626,5 МГц, т. е. всего 16,5 МГц). Это потребовало применения эффективных методов передачи и многократного использования полосы частот (с помощью узких лучей бортовых антенн).

Совмещение в терминале абонента всех функций станций спутниковой связи (аналого-цифровое преобразование, модуляция, демодуляция, усиление мощности и т. п.), да еще в двух вариантах — для наземной и спутниковой систем связи — представляет собой сложнейшую технологическую задачу, решаемую лишь при высокотехнологическом производстве на базе специализированных БИС и микропроцессоров.

Полосы частот, используемые системами персональной связи, показаны в табл. 9.2.

Создание ряда систем спутниковой связи на низколетящих спутниках началось как в России, так и за рубежом. Параметры основных систем на низких спутниках приведены в табл. 9.3.

На рис. 9.59 показана схема орбитальной группировки системы Iridium. Космический сегмент состоит из некоторого числа спутников, находящихся на негеостационарных орбитах. Спутники размещаются по несколько штук в некоторой плоскости (плане) таким образом, что, двигаясь в заданных плоскостях и последовательно сменяя друг друга, формируют заданную зону обслуживания. Требуемая зона обслуживания формируется узкими лучами антенных систем отдельных спутников, находящихся в различных орбитальных плоскостях.

Основными типами негеостационарных орбит, используемых в системах персональной связи, являются:

- низкие земные орбиты (LEO) высотой 700...1500 км;
- промежуточные круговые орбиты (МЕО) высотой порядка 10 000 км.

В наземный сегмент системы включаются:

- станции сопряжения, обеспечивающие взаимодействие системы с наземными сетями общего пользования (НСОП) при применении для каждой страны или региона конкретных систем нумерации и видов сигнализации;

Таблица 9.2

Полосы частот, используемые системами персональной связи

Система	Линия	Полоса частот
Iridium	Фидерная:	
	СС-ИСЗ-СС	19,3...19,6 ГГц
	Пользовательская:	
	АС-ИСЗ	1610,0...1626,5 МГц
	ИСЗ-АС	2483,5...2500,0 МГц
Globalstar	Фидерная:	
	СС-ИСЗ	5091,0...5250,0 МГц
	ИСЗ-СС	6700,0...7075,0 МГц
	Пользовательская:	
	АС-ИСЗ	1610,0...1626,5 МГц
	ИСЗ-АС	2483,5...2500,0 МГц
Inmarsat-ICO	Фидерная:	
	СС-ИСЗ	5091,0...5250,0 МГц
	ИСЗ-СС	6700,0...7075,0 МГц
	Пользовательская:	
	АС-ИСЗ	1980,0...2025,0 МГц
	ИСЗ-АС	2160,0...2200,0 МГц
Odyssey	Фидерная:	
	СС-ИСЗ	29,1...29,4 ГГц
	ИСЗ-СС	19,3...19,6 ГГц
	Пользовательская:	
	АС-ИСЗ	1610,0...1626,5 МГц
	ИСЗ-АС	2483,5...2500,0 МГц

— станции управления ИСЗ-ретрансляторами, включая станции управления сетью связи, а также станции телеметрии и передачи команд управления.

Для организации доступа абонентов в систему используют достаточно сложное сочетание нескольких видов многостанционного доступа (МД):

— с пространственным разделением (МДПР), организуемый за счет использования узких лучей, формируемых на ИСЗ;

— с временным разделением (МДВР), используемый в отдельном луче (cote), — системы Iridium, Inmarsat-ICO;

— с кодовым разделением (МДКР), используемый в каждом луче (cote), — системы Globalstar, Odyssey, «Сигнал»;

— с частотным разделением (МДЧР), используемый для смежных лучей (cot), — системы Iridium, Globalstar, «Сигнал».

Одной из основных проблем, возникающих при разработке систем персональной спутниковой связи, является организация соединений между абонентскими станциями, находящимися в зонах, формируемых различными ИСЗ.

Параметры систем спутниковой связи

Параметры системы	Indium	Globalstar	Inmarsat-ICO	Odyssey	«Сигнал»	«Гонец»
Число К/А	66	48	10	12	48	45
Число орбитальных плоскостей/число КА в плоскости	6/11	8/6	2/5	3/4	8/6	5/9
Тип орбиты	LEO	LEO	ICO	MEO	LEO	LEO
Высота орбиты, км	780	1389	10335	10354	1500	1400
Наклонение, град	86	52	45	55	74	83
Масса КА, кг (на орбите/стартовая)	317/689	250/450	1400/2200	1900	300	225
Зона обслуживания, град. (с.ш./ю.ш.)	0...90	0...72	0...90	0...90	27...90	0...90
Предоставляемые услуги	ТЛФ*, ПД**	ТЛФ, ПД	ТЛФ, ПД	ТЛФ, ПД	ТЛФ, ПД	ПД (эл. почта) ТЛФ (в пределах зоны)
Суммарная мощность передатчиков в направлении ИСЗ-АС, Вт	120	1000	600	200	150	40
Число лучей	48	16	85	61	11	1
Число каналов на КА	2500	2600	4000	—	220	—
Метод доступа абонентов	МДВР-МДЧР, МДПР	МДКР-МДЧР, МДПР	МДВР-МДПР	МДКР-МДПР	МДКР-МДЧР	МДВУ-МДЧР, МДВР
Способ связи между зонами	Межспутниковая	Через станцию сопряжения	Через станцию сопряжения	Через станцию сопряжения	Через станцию сопряжения	Перенос, память на спутнике
Число СС	25	200	12	7	6	—

* ТЛФ — телефония; ** ПД — передача данных



Рис. 9.59. Схема орбитальной группировки системы Iridium

В настоящее время эта задача решается двумя способами.

1. Использование межспутниковой связи (рис. 9.60 — система Iridium). Каждый ИСЗ имеет радиолинии связи с двумя ИСЗ, находящимися в той же орбитальной плоскости, и с двумя ИСЗ, находящимися в соседних орбитальных плоскостях. Сегмент управления сетью связи сообщает центральному процессору (ЦП) управления ИСЗ информацию о положении абонентской станции в той или иной зоне. ЦП выбирает необходимый маршрут прохождения информации по соответствующему межспутниковому каналу связи. Поэтому для сопряжения с наземными линиями связи требуется ограниченное количество станций сопряжения. Для линий межспутниковой связи применяются следующие антенные системы.

2. Использование наземных каналов связи между станциями сопряжения, находящимися в каждой зоне (системы Globalstar, Inmarsat-ICO, Odyssey, «Сигнал» — рис. 9.61). Сигнал вызова (или информационный сигнал) от абонентской станции через ИСЗ поступает на станцию сопряжения данной зоны. Поскольку в банке данных станции сопряжения хранится информация о зоне, в которой расположена вызываемая абонентская станция, станция сопряжения организует прохождение вызывного (или информационного) сигнала к соответствующей станции сопряжения по наземным каналам связи.

Система «Гонец» является наиболее простой системой персональной спутниковой связи и предназначена для нескольких режимов передачи сооб-

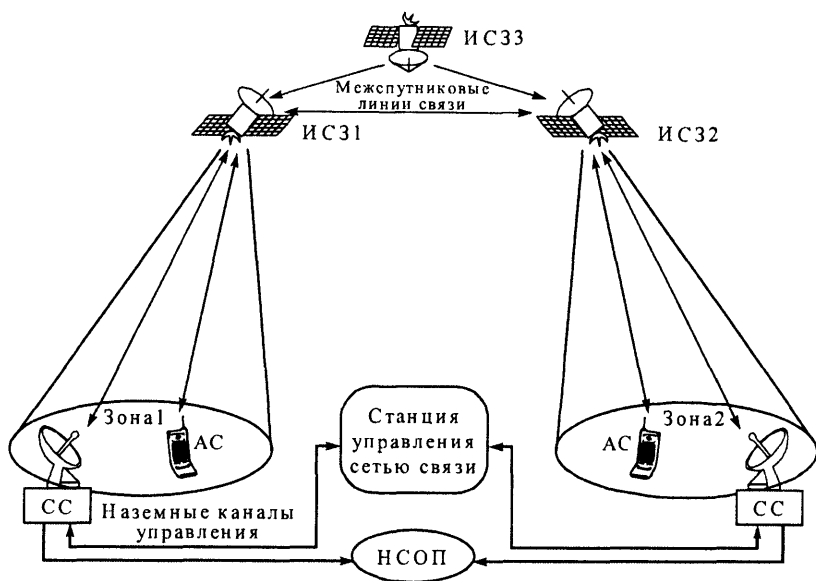


Рис. 9.60. Система с межспутниковыми связями

щений в зависимости от их объема и типа. При нахождении абонентов в зоне видимости одного ИСЗ обмен информацией (в том числе и телефонной) производится в реальном масштабе времени. Если вызывающая и вызываемая абонентские станции находятся в зонах радиовидимости различных ИСЗ, возможна передача сообщений только в режиме электронной почты — с запоминанием, хранением сообщений на борту ИСЗ и их передачей к вызываемой абонентской станции при пролете ИСЗ над регионом ее расположения.

Спутниковая связь является важнейшим элементом мобильных систем третьего поколения, так как обладает многими архитектурными и технологическими возможностями, например такими, как мультирегиональный и глобальный охват обслуживаемой территории, обеспечивающий универсальное решение задачи роуминга и доступа пользователей в удаленных и малонаселенных районах к услугам связи.

Спутниковый сегмент 3G следует рассматривать как развитие многих глобальных систем спутниковой связи, реализованных в 1998—2000 гг. (Iridium, Globalstar и др.) и предоставляющих услуги связи типа голосовой телефонии, факс-связи и низкоскоростной передачи данных (4,8...9,6 кбит/с).

Предложения по разработке проектов стандартов спутниковых систем в рамках S-IMT-2000 (Satellite IMT-2000) включают в себя целый ряд вариантов этих стандартов, среди которых важное место занимают стандарты на основе CDMA: SW-CDMA (ESA), SW-C/TDMA (ESA), SAT-CDMA (Корея).

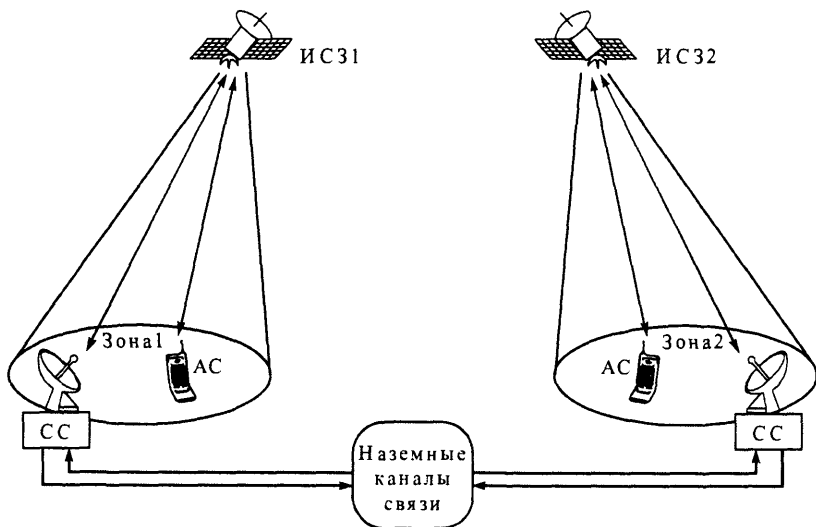


Рис. 9.61. Система с использованием наземных каналов связи

Несмотря на имеющиеся отличия, все перечисленные спутниковые системы третьего поколения должны удовлетворять ряду общих требований, включая [139]:

- освоение новых диапазонов частот, выделяемых с 2000 г. для спутниковых систем подвижной связи: 1980...2010 МГц (линия «Земля—спутник») и 2170...2200 МГц (линия «спутник—Земля»);

- предоставление уже на первой фазе развертывания широкого спектра услуг, включая передачу речи, данных и низкоскоростной мультимедийной информации со скоростями от 2,4 кбит/с до 64 (144) кбит/с в глобальной зоне обслуживания;

- возможность наращивания пропускной способности сети введением новых спутников, станций сопряжения и абонентских терминалов различных модификаций.

Рассмотрим особенности перечисленных выше вариантов.

Европейским космическим агентством (ESA) предложены два проекта спутниковых систем: первый из них основан на широкополосном кодовом разделении SW-CDMA (Satellite Wideband CDMA), а второй — на гибридном кодово-временном разделении каналов SW-C/TDMA (Satellite Wideband Code and Time Division Multiple Access). Основная идея проекта SW-CDMA — адаптация технологии системы W-CDMA, разработанной в рамках стандартов IMT-2000, применительно к спутниковой связи. В проекте SW-C/TDMA предполагается использовать два метода дуплексного разделения: FDD (двухчастотный дуплекс) и FTDD (комбинированный частотно-

временной дуплекс), что делает систему более гибкой и позволяет менять пропускную способность в зависимости от условий эксплуатации и видов обслуживания. Орбитальные группировки для рассматриваемых систем могут строиться на базе спутников, выведенных как на низкие (LEO), средние (MEO), так и на геостационарную (GEO) орбиты.

Максимальная скорость передачи достигает 64 кбит/с при полосе сигнала в радиоканале 2,5 или 5 МГц (тактовые частоты псевдослучайных последовательностей соответственно 2,048 и 4,096 МГц).

Проект спутниковой системы SATCDMA подготовлен Ассоциацией связанных технологий (ТТА) Кореи. Этот проект базируется на использовании CDMA с сигналами, полоса которых 10 МГц. Скорость передачи лежит в пределах от 9,6 до 144 кбит/с. В орбитальной группировке системы используются спутники на низких орбитах (LEO).

В заключение можно отметить, что оборудование на базе технологии CDMA позволяет создавать экономически эффективные системы связи различного назначения, обеспечивающие предоставление пользователям широкого спектра высококачественных услуг передачи речи и данных.

Контрольные вопросы

1. Поясните назначение отдельных узлов цифровой системы передачи информации.
2. Какие искажения сигнала и помехи возникают в непрерывном канале связи?
3. Дайте описание физической и математической моделей непрерывного канала связи.
4. Что такое пропускная способность непрерывного канала, дискретного канала и эффективность использования полосы частот и энергетики канала?
5. Каков алгоритм работы оптимального модема в гауссовом канале с полностью известными параметрами сигнала (когерентный прием) и в канале с неопределенной фазой сигнала (некогерентный прием)?
6. Какие виды модуляции следует применять для повышения эффективности использования полосы частот канала?
7. Как оценить качество приема дискретной информации в канале с медленными замираниями?
8. Назовите способы повышения достоверности передачи информации в канале с медленными замираниями.
9. Поясните алгоритм работы системы с переменной длительностью посылок.
10. Поясните алгоритм разнесенного приема с когерентным сложением, автовыбором, дискретным сложением.
11. Как влияет полоса частот сигнала на характер замираний в канале с рассеянием сигнала?
12. Как повысить достоверность передачи информации в канале с небелым шумом (в канале с сосредоточенными по спектру помехами)?
13. Поясните работу «обеляющего» фильтра в системах передачи с широкополосными сигналами.

14. В чем преимущества и недостатки систем с перестройкой частоты по сравнению с системами, использующими широкополосные сигналы?
15. Чем определяются энергетические потери при некогерентной обработке элементов сигнала с «быстрой» ПРЧ? Почему в системах с «медленной» ПРЧ необходимо использовать помехоустойчивое кодирование?
16. Поясните алгоритм функционирования системы с адаптивной ПРЧ.
17. Сравните между собой эффективность различных способов использования частотной избыточности для повышения помехоустойчивости в канале с небелым шумом.
18. Поясните принципы избыточного кодирования для повышения достоверности передачи информации.
19. Какими характеристиками кода определяется его обнаруживающая и исправляющая способность?
20. Дайте классификацию кодов и поясните принципы внесения избыточности.
21. Как оценить эффективность блочного кода?
22. Поясните алгоритм кодирования и декодирования линейных блочных кодов.
23. Поясните алгоритм кодирования и декодирования сверточных кодов.
24. Как используются блочные коды в системах с обратным каналом?
25. В чем основное различие между многостанционными и многоканальными РСПИ?
26. Чем оценивается активность корреспондентов в многостанционной РСПИ?
27. Какие способы построения систем МСД существуют?
28. Чем характеризуется эффективность систем МДВР и каковы пути ее улучшения?
29. В чем суть простейшего метода синхронизации при МДВР?
30. Назовите достоинства и недостатки систем МДЧР.
31. Перечислите синхропараметры сигнала дискретной РСПИ и назовите показатели качества работы системы синхронизации.
32. Поясните назначение отдельных узлов системы синхронизации.
33. Как влияют ошибки устройства фазовой и тактовой синхронизации на помехоустойчивость?
34. Поясните работу ФАПЧ и выбор ее основных параметров.

ЗАКЛЮЧЕНИЕ

В настоящем учебном пособии изложены основы статистической теории радиотехнических систем и рассмотрены наиболее важные информационные технологии, разработанные и применяемые в радиолокации, радионавигации и радиосвязи.

При создании книги авторы хорошо понимали, что охватить в одном пособии все многообразие вопросов, связанных с разработкой современных радиотехнических систем, практически невозможно. Это объясняется многоплановостью дисциплины «Радиотехнические системы», ограниченностью объема книги, постоянным развитием традиционных и появлением новых информационных технологий. Поэтому не все информационные технологии, представляющие интерес как для студентов и аспирантов, так и для разработчиков соответствующих систем, нашли должное отражение в пособии. Авторы старались восполнить пробелы ссылками на опубликованные работы, в которых заинтересованный читатель мог бы подробно ознакомиться с той или иной проблемой. Далее приведен краткий и, безусловно, не исчерпывающий перечень таких технологий.

1. Современные информационные технологии обработки сигналов немислимы без методов оптимизации в условиях априорной неопределенности, поскольку на практике часто приходится решать задачи приема сигналов, когда априорные сведения о статистических характеристиках сигналов и помех частично или полностью отсутствуют. В этих условиях используют адаптивные, непараметрические и робастные алгоритмы обработки сигналов, которые позволяют преодолеть априорную неопределенность, незначительно уступая по эффективности алгоритмам, синтезированным для случая наличия полной априорной информации.

2. Для решения многих радиолокационных задач, таких как первичная и вторичная обработка информации, успешно используются алгоритмы на основе нейрокомпьютерных технологий. Эти алгоритмы относятся к классу обучаемых, находят применение в задачах распознавания, управления и оптимизации, позволяют совмещать высокую степень адаптации к внешним условиям с устойчиво высоким качеством работы.

3. Неотъемлемой частью современной радиолокации являются информационные технологии, основанные на синтезе искусственного раскрыва антенны. Разрабатываемые самолетные и спутниковые радиолокационные станции бокового обзора, основанные на этих технологиях, позволяют получить изображения поверхности Земли с разрешающей способностью, близкой к разрешающей способности оптических средств. Такие радиолокационные станции используются для контроля за движением судов на морях и океанах; поиска различных объектов, терпящих бедствие; обнаружения нефтяных пятен; контроля за шоссейными и железными дорогами, мостами; исследования природных ресурсов и т. п.

4. Одним из важных направлений развития радиотехнических систем является подповерхностная радиолокация. Она используется для обнаружения под землей различных инженерных работ, построения геологических разрезов и профилей дна водоемов, определения положения уровня грунтовых вод и границ распространения полезных ископаемых и т. п.

5. Появившиеся в последнее время возможности широкого внедрения в системы радиосвязи быстродействующей микропроцессорной техники и современной элементной базы позволяют реализовать: новые принципы формирования приема и обработки сигналов, включая сложные шумоподобные сигналы; перестройку по псевдослучайному закону частоты несущей; помехоустойчивое кодирование. Все это позволяет обеспечить высокую помехоустойчивость систем передачи информации при функционировании в сложной помеховой обстановке.

6. Важным направлением развития радиосвязи является мобильная связь. Непрерывно расширяется объем и повышается качество предоставляемой ею услуг, делая их массовыми, общедоступными и приспособленными к индивидуальным потребностям пользователей.

Авторы надеются, что учебное пособие будет полезным при изучении таких дисциплин, как «Радиотехнические системы», «Радиолокационные системы», «Радионавигационные системы», «Радиотехнические системы передачи информации».

СПИСОК ЛИТЕРАТУРЫ

1. *Радиотехнические системы*: Учебник для вузов по спец. «Радиотехника» / Ю.П. Гришин, В.П. Ипатов, Ю.М. Казаринов и др.; Под ред. Ю.М. Казаринова. М.: Высш. шк., 1990.
2. *Пестряков В.Б., Кузенков В.Д.* Радиотехнические системы: Учебник для вузов. М.: Радио и связь, 1985.
3. *Чердынцев В.А.* Радиотехнические системы: Учеб. пособие для вузов. М.: Высш. шк., 1988.
4. *Дымова А.И., Альбац М.Е., Бонч-Бруевич А.М.* Радиотехнические системы: Учебник для вузов. М.: Сов. радио, 1975.
5. *100 лет радио*: Сб. статей / Под ред. В.В. Мигулина, А.В. Гороховского. М.: Радио и связь, 1995.
6. *Подповерхностная радиолокация* / М.И. Финкельштейн, В.И. Карпунин, В.А. Кутев и др.; Под ред. М.И. Финкельштейна. М.: Радио и связь, 1994.
7. *Гоноровский И.С.* Радиотехнические цепи и сигналы. М.: Радио и связь, 1986.
8. *Тихонов В.И.* Статистическая радиотехника. М.: Радио и связь, 1982.
9. *Зюко А.Г., Кловский Д.Д., Назаров М.В., Финк Л.М.* Теория передачи сигналов. М.: Радио и связь, 1986.
10. *Пенин П.И., Филиппов Л.И.* Радиотехнические системы передачи информации. М.: Радио и связь, 1984.
11. *Радиотехнические системы передачи информации*: Учеб. пособие для вузов / В.А. Борисов, В.В. Калмыков, Я.М. Ковальчук и др.; Под ред. В.В. Калмыкова. М.: Радио и связь, 1990.
12. *Дядюнов Н.Г., Сенин А.И.* Ортогональные и квазиортогональные сигналы. М.: Связь, 1977.
13. *Спилкер Дж.* Цифровая спутниковая связь: Пер. с англ./ Под ред. В.В. Маркова. М.: Связь, 1979.
14. *Питерсон У., Уэлдон Э.* Коды, исправляющие ошибки: Пер. с англ./ Под ред. Р.Л. Добрушина и С.И. Самойленко. М.: Мир, 1976.

15. *Варакин Л.Е.* Теория систем сигналов. М.: Сов. радио, 1978.
16. *Сарвате Д.В., Персли М.В.* Взаимно-корреляционные свойства псевдослучайных и родственных последовательностей // ТИИЭР. 1980. Т. 68. № 5. С. 59–90.
17. *Шумоподобные сигналы в системах передачи информации* / В.Б. Пестряков, В.П. Афанасьев, В.Л. Гурвиц и др.; Под ред. В.Б. Пестрякова. М.: Сов. радио, 1973.
18. *Быков В.В.* Цифровое моделирование в статистической радиотехнике. - М.: Сов. радио, 1971.
19. *Кузнецов А.А., Сенин А.И.* Моделирование радиотехнических систем: Учеб. пособие. В 2-х ч. Ч. 1, 2000. Ч. 2, 2001. М.: Изд-во МГТУ им. Н.Э. Баумана, 2000.
20. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров: Пер. с англ. / Под общей ред. И.Г. Арамановича. М.: Наука, 1973.
21. *Тихонов В.И.* Оптимальный прием сигналов. М.: Радио и связь, 1983.
22. *Левин Б.Р.* Теоретические основы статистической радиотехники. М.: Радио и связь, 1989.
23. *Акимов П.С., Сенин А.И., Соленов В.И.* Сигналы и их обработка в информационных системах: Учеб. пособие для вузов. М.: Радио и связь, 1994.
24. *Теория обнаружения сигналов* / П.С. Акимов, П.А. Бакут, В.А. Богданович и др.; Под ред. П.А. Бакута. М.: Радио и связь, 1984.
25. *Сосулин Ю.Г.* Теория обнаружения и оценивания стохастических сигналов. М.: Сов. радио, 1978.
26. *Кузнецов А.А., Сенин А.И., Соленов В.И., Шутко В.Н.* Обнаружение и различение сигналов: Учеб. пособие. К.: КМУГА, 1999.
27. *Вальд А.* Последовательный анализ: Пер с англ. / Под ред. Б.А. Севастьянова. М.: Физматгиз, 1960.
28. *Сосулин Ю.Г.* Теоретические основы радиолокации и радионавигации: Учеб. пособие для вузов. М.: Радио и связь, 1992.
29. *Сосулин Ю.Г.* Последовательное обнаружение сигналов: проблемы и перспективы // Радиотехника. 1998. № 10. С. 63–68.
30. *Обработка сигналов в радиотехнических системах* / А.Д. Далматов, А.А. Елисеев, А.П. Лукошкин и др.; Под ред. А.П. Лукошкина. Л.: Изд-во Ленингр. ун-та, 1987.
31. *Теоретические основы радиолокации:* Учебное пособие для вузов / Под ред. Я.Д. Ширмана. М.: Сов. радио, 1970.
32. *Ширман Я.Д., Манжос В.Н., Леховицкий Д.И.* Некоторые этапы развития и проблемы теории и техники разрешения радиолокационных сигналов // Радиотехника. 1997. №1. С. 31–42.

33. *Ширман Я.Д.* Разрешение и сжатие сигналов. М.: Сов. радио, 1974.
34. *Ширман Я.Д., Манжос В.Н.* Теория и техника обработки радиолокационной информации на фоне помех. М.: Радио и связь, 1981.
35. *Радиоэлектронные системы: основы построения и теория:* Справочник / Я.Д. Ширман, Ю.П. Лосев, Н.Н. Минервин и др.; Под ред. Я.Д. Ширмана. М.: ЗАО «Маквис», 1998.
36. *Слока В.К.* Схемы разрешения с применением режекции мешающих сигналов // Радиотехника и электроника. 1978. № 1. С. 57–66.
37. *Ван Трис Г.* Теория обнаружения, оценок и модуляции: В 3-х т. / Под ред. В.Т. Горяинова. Пер. с англ.; М.: Сов. Радио, 1972–1977. Т. 3. Обработка сигналов в радио- и гидролокации и прием случайных гауссовых сигналов на фоне помех, 1977.
38. *Куликов Е.И., Трифонов А.П.* Оценка параметров сигналов на фоне помех. М.: Сов. радио, 1978.
39. *Радиолокационные измерители дальности и скорости:* В 2-х т. Т. 1 / В.И. Меркулов, А.И. Перов, В.Н. Саблин и др.; Под ред. В.Н.Саблина. М.: Радио и связь, 1999.
40. *Алмазов В.Б., Белов А.А., Кокин В.Н., Рябуха В.П.* Теоретические основы радиолокации. Харьков: ХВУ, 1996.
41. *Кук Ч., Бернфельд М.* Радиолокационные сигналы: Пер. с англ. / Под ред. В.С. Кельзона. М.: Сов. радио, 1971.
42. *Черных М.М., Богданов А.В., Буров А.С.* и др. Анализ информационных свойств когерентных радиолокационных сигналов // Вестник МГТУ им. Н.Э. Баумана. Сер. «Приборостроение». 1999. № 4. С. 16–26.
43. *Теоретические основы радиолокации:* Учеб. пособие для вузов / А.А. Коростелев, Н.Ф. Клюев, Ю.В. Мельник и др.; Под ред. В.Е. Дулевича. М.: Сов. радио, 1978.
44. *Дрогалин В.В., Меркулов В.И., Родзивилов В.А.* и др. Алгоритмы оценивания угловых координат источников излучения, основанные на методах спектрального анализа // Радиолокация и радиометрия. 1999. Вып. 1. С. 57–64.
45. *Дрогалин В.В., Петухов В.Н., Родзивилов В.А., Чернов М.В.* Экспериментальные исследования алгоритма Прони—Берлекемпа // Радиотехника. 2001. Вып. 8. С. 42–49.
46. *Леонов А.И., Фомичев К.И.* Моноимпульсная радиолокация. М.: Радио и связь, 1984.
47. *Коростелев А.А.* Пространственно-временная теория радиосистем. М.: Радио и связь, 1987.
48. *Егоров Ю.М.* Комплексирование радиолокационных измерительных каналов на основе пространственно-временной обработки сигналов // Труды VII Международной научно-технической конференции «Ра-

- диолокация, навигация, связь» (24–26 апреля 2001 г.). Воронеж: ВГУ, 2001. Т. 3. С. 1498–1500.
49. *Пространственно-временная обработка сигналов* / И.Я. Кремер, А.И. Кремер, В.М. Петров и др.; Под ред. И.Я. Кремера. М.: Радио и связь, 1984.
50. *Wiener N.* Extrapolation, Interpolation and Smoothing of Stationary Time Series. N.Y.: Technology Press and Wiley, 1949.
51. *Колмогоров А.Н.* Интерполирование и экстраполирование стационарных случайных последовательностей // Изв. АН СССР. Сер. матем. 1941. Т. 5. № 1.
52. *Kalman R.E.* A new approach to linear filtering and prediction problems // J. of Basis Eng. Trans. Asme. Ser. D. 1960. Vol. 82. № 1. P 35–45.
53. *Kalman R.E., Busy R.S.* New results in linear filtering and prediction theory // J. of Basis Eng. Trans. Asme. Ser. D. 1961. Vol. 83. № 1. P. 95–108.
54. *Стратонович Р.Л.* Условные марковские процессы и их применение к теории оптимального управления. М.: Изд-во МГУ, 1966.
55. *Стратонович Р.Л.* Условное распределение коррелированных случайных точек и использование корреляций для оптимального выделения импульсного сигнала из шумов // Изв. АН СССР. Сер. Энергетика и автоматика. 1961. № 2.
56. *Ачкасов Ю.С.* Приложение теории апостериорного анализа потоков к задаче обнаружения траекторий // Изв. АН СССР. Техническая кибернетика. 1971. № 5. С. 157–166.
57. *Бакут П.А., Жулина Ю.В., Иванчук Н.А.* Обнаружение движущихся объектов / Под ред. П.А. Бакута. М.: Сов. радио, 1980.
58. *Ворочилин В.В., Слукин Г.П., Федоров И.Б.* Синтез алгоритмов совместного обнаружения—оценивания траекторий движущихся объектов на основе теории случайных потоков // Труды МВТУ им. Н.Э. Баумана. № 540. 1989. С. 95–128.
59. *Ворочилин В.В., Слукин Г.П., Федоров И.Б.* Синтез алгоритмов совместного обнаружителя—оценивания траекторий потоков случайных точек // Труды МГТУ им. Н.Э. Баумана. М.: Мир, 1990. С. 5–25.
60. *Vorothilun V.V., Slukin G.P., Fedorov I.B.* Algorithm synthesis for a detector—estimator of random point trajectories. Topics in Radioelectronic and Lazer System / Design Edited by I. Fedorov. Tokyo—London: CRC Press Boca Ration Ann Arbor, 1992.
61. *Фарина А., Стуер Ф.* Цифровая обработка радиолокационной информации. Сопровождение целей: Пер. с англ. М.: Радио и связь, 1993.
62. *Bar-Sholom Y., Xivo Rong Li.* Multitarget — Multisensor Tracking: Principles and Techniques. — YBS Publishing, 1995.

63. *Bar-Sholom Y., Xivo Rong Li. Estimation and Tracking: Principles, Techniques and Software.* – YBS Publishing, 1998.
64. *Научно-технические серии.* Радиолокация и радиометрия. № 1. 1999. Вып. 1: Радиолокационное сопровождение интенсивно маневрирующих воздушных объектов.
65. *Ильчук А.Р., Канащенков А.И., Слукин Г.П. и др.* Алгоритм автоматического радиолокационного сопровождения целей в режиме обзора // Радиотехника. 1999. № 11. С. 3–21.
66. *Слукин Г.П.* Синтез алгоритма траекторной обработки для плотных потоков помеховых отсчетов и разреженного потока траекторий // Вестник МГТУ им. Н.Э. Баумана. Сер. «Приборостроение». 2000. № 4. С. 27–40.
67. *Кузьмин С.З.* Цифровая радиолокация. Введение в теорию. Киев: Изд-во «Квіц», 2000.
68. *Большаков И.А.* Статистические проблемы выделения потока сигналов из шума. М.: Сов. радио, 1969.
69. *Справочник по теории автоматического управления* / Под ред. А.А. Кравцовского. М.: Наука, 1987.
70. *Жданюк Б.Ф.* Основы статистической обработки траекторных измерений. М.: Сов. радио, 1978.
71. *Усачев В.А., Федоров И.Б.* Отождествление потока объектов в системе измерителей // Изв. вузов. Сер. Радиоэлектроника. 1980. № 11. С. 32–37.
72. *Радиолокация.* Радиолокационная станция. Российская военная энциклопедия. В 6-ти т. Т. 6. М.: Воениздат, 2002.
73. *Сколник М.* Справочник по радиолокации в 4-х т.: Пер. с англ. / Под ред. К.Н. Трофимова. М.: Сов. радио, 1976–1979.
74. *Ширман Я.Д.* Теоретические основы радиолокации. Харьков: ВИРТА, 1984.
75. *Слока В.К.* Вопросы обработки радиолокационных сигналов. М.: Сов. радио, 1969.
76. *Черняк В.С.* Многопозиционная радиолокация. М.: Радио и связь, 1993.
77. *Лезин Ю.С.* Введение в теорию и технику радиотехнических систем: Учебное пособие для вузов. М.: Радио и связь, 1986.
78. *Основы загоризонтной радиолокации* / Под ред. А.А. Колосова. М.: Радио и связь, 1984.
79. *Бакулев П.А., Сосновский А.А.* Радиолокационные и радионавигационные системы: Учеб. пособие для вузов. М.: Радио и связь, 1994.
80. *Морская радиолокация* / Под ред. В.И. Винокурова. Л.: Судостроение, 1986.

81. *Радиолокационные станции обзора Земли* / Г.С. Кондратенков, В.А. Потехин, А.П. Реутов, Ю.А. Феоктистов; Под ред. Г.С. Кондратенкова. М.: Радио и связь, 1983.
82. *Леонов С.А.* Радиолокационные средства противовоздушной обороны. М.: Воениздат, 1988.
83. *Дмитриев Д.Д., Слукин Г.П., Федоров И.Б.* Обнаружение объектов в контролируемой зоне системой локационных станций // Вестник МГТУ им. Н.Э. Баумана. Сер. «Приборостроение». 1991. № 3. С. 87–92.
84. *Локационная системотехника* / Под ред. В.Б. Алмазова. Харьков: ВИРТА, 1993.
85. Computer simulation of aerial target radar scattering, recognition, detection and tracking // Ed. Ya. D. Shirman. London: Artech House Boston, 2002.
86. *Защита от радиопомех* / М.В. Максимов, М.П. Бобнев, Б.Х. Кривицкий и др.; Под ред. М.В. Максимова. М.: Сов. радио, 1976.
87. *Бакулев П.А., Степин В.М.* Методы и устройства селекции движущихся целей. М.: Радио и связь, 1986.
88. *Вакин С.А., Шустов Л.Н.* Основы радиопротиводействия и радиотехнической разведки. М.: Сов. радио, 1968.
89. *Аверьянов В.Я.* Разнесенные радиолокационные станции и системы. Минск: Наука и техника, 1978.
90. *Волжин А.Н., Сизов Ю.Г.* Борьба с самонаводящимися ракетами. М.: Воениздат, 1983.
91. *Федоров И.Б., Слукин Г.П., Нефедов С.И.* О возможности метровой радиолокации для контроля перемещений воздушных слоев при отсутствии облачности // Полет. 1999. № 3. С. 14–18.
92. *Всероссийский научно-исследовательский институт радиотехники. 75 лет.* М.: Финансово-промышленная группа «Оборонительные системы», 1996.
93. *Оружие России: Каталог. Т. V.* Вооружение и военная техника Войск противовоздушной обороны. М.: ЗАО «Военный парад», 1997.
94. *Кобак В.О.* Радиолокационные отражатели. М.: Сов. радио, 1975.
95. *Штагер Е.А.* Рассеяние радиоволн на телах сложной формы. М.: Радио и связь, 1986.
96. *Эффективная площадь отражения сложных радиолокационных объектов.* Тематический выпуск // ТИИЭР. 1989. Т. 77. № 5.
97. *Ахияров В.В., Слукин Г.П., Шустиков В.Ю.* Расчет отражательных характеристик идеально проводящих тел вращения в резонансной области методом интегральных уравнений // Антенны. 2001. Вып. 8 (54). С. 53–60.
98. *Конторов Д.С., Голубев-Новожилов Ю.С.* Введение в радиолокационную системотехнику. М.: Сов. радио, 1971.

99. *Волосатов А.Д., Николаев А.И.* Обработка многочастотных линейно-частотно-модулированных сигналов с использованием рециркулятора со смещающим гетеродином // *Радиотехника*. 1994. № 10. С. 58–60.
100. *Варакин Л.Е.* Теория сложных сигналов. М.: Сов. радио, 1970.
101. *Применение цифровой обработки сигналов*: Пер с англ. / Под ред. Э. Опенгейма. М.: Мир, 1980.
102. *Марпл С.П. мл.* Цифровой спектральный анализ и его применения: Пер с англ. М.: Мир, 1990.
103. *Ширман Я.Д., Красногоров С.И., Лебедев Е.П., Костин Г.А.* Первые отечественные исследования адаптации антенных систем к мешающим воздействиям // *Радиотехника*. 1989. № 11. С. 108–110.
104. *Давыдов П.С., Сосновский А.А., Хаймович И.А.* Авиационная радиолокация: Справочник. М.: Транспорт, 1984.
105. *Аэродромный радиолокационный комплекс АРЛК-11. Трассовый радиолокационный комплекс ТРЛК-11.* М.: НПО «Утес», 2000.
106. *Леманский А., Ненартович Н.* Современные средства ПВО // *Военный парад*. 1995. Март–апрель. С. 161–163.
107. *Петухов С.И., Шестов И.В.* История создания и развития вооружения и военной техники ПВО Сухопутных войск России. РАРАН. ЗЦНИИ МО РФ. М.: Изд-во «Вооружение, политика, конверсия», 1997.
108. *Оружие России*: Каталог. Т. I. Вооружение Сухопутных войск. М.: ЗАО «Военный парад», 1995.
109. *Зайцев Н.* Российские радары: на поле боя и охране рубежей // *Военный парад*. 1998. Вып. 3 (27). С. 60–61.
110. *Карпенко А.В.* Противоракетная и противокосмическая оборона // Приложение к военно-техническому сборнику. Вып. 4. Спб.: Невский бастион, 1998.
111. *Leonov A.L.* History of Nonopulde Radar in the USSR // *IEEE Systems*. 1998. № 5.
112. *Поярков Б., Юрин Ю.* Многофункциональная система оружия «Иджис» // *Зарубежное военное обозрение*. 1989. Вып. 10. С. 53–60.
113. *Канащенков А.П., Гуськов Ю.П.* Широкий спектр продукции корпорации «Фазотрон—НИИР» // *Военный парад*. 2001. № 4. С. 42–47.
114. *Глобальная спутниковая навигационная система ГЛОНАСС*. Интерфейсный контрольный документ: 4-я ред. М.: Изд-во КНИЦ ВКС РФ, 1998.
115. *Сетевые спутниковые радионавигационные системы* / В.С. Шебшаевич, П.П. Дмитриев и др. М.: Радио и связь, 1993.
116. *Глобальная спутниковая радионавигационная система ГЛОНАСС* / Под ред. В.Н. Харисова, А.И. Перова, В.А. Болдина. М.: ИПРЖР, 1998.

117. Соловьев Ю.А. Системы спутниковой навигации. М.: Эко-Трендз, 2000.
118. *Interface Control Document: NAVSTAR GPS Space Segment / Navigation User Interfaces (ICD-GPS-200.)*. Rockwell Int. Corp., 1987.
119. *Бортовые устройства спутниковой навигации* / И.В. Кудрявцев, И.Н. Мищенко, А.И. Волынкин и др.; Под ред. В.С. Шебшаевича. М.: Транспорт, 1988.
120. Манин А.П., Романов Л.М. Методы и средства относительных определений в системе NAVSTAR // *Зарубежная радиоэлектроника*. 1989. № 1. С. 33–45.
121. Генике А.А., Побединский Г.Г. Глобальная спутниковая система определения местоположения GPS и ее применений в геодезии. М.: Картогеоцентр—Геодезиздат, 1999.
122. Longe P. de, Tiberius C. The LAMBDA method for integer ambiguity estimation: implementation aspects // *Publication of the Delft Geodetic Computing Centre*. 1966. № 12.
123. Власов И.Б. Спутниковая радионавигационная аппаратура с угломерным каналом и ее применения // *Вестник МГТУ им. Н.Э. Баумана. Сер. «Приборостроение»*. 1997. № 4. С. 16–22.
124. Власов И.Б., Кириченко О.И., Копелович В.А. и др. Результаты летных экспериментов по исследованию точностных характеристик спутниковой навигационной аппаратуры с угломерным каналом // *Труды международной конференции «Спутниковые системы связи и навигации»*. Красноярск, 1997. Т. 1. С. 157–171.
125. Власов И.Б., Пудловский В.Б., Тарахнов С.Н. Точностные характеристики спутниковой навигационной аппаратуры с угломерным каналом // *Вестник МГТУ им. Н.Э. Баумана. Сер. «Приборостроение»*. 1997. № 1. С. 114–127.
126. Lachapelle G. Attitude Determination: System Implication and Innovative Application of Satellite Navigation // *Agard Lect. Ser.* 207. June 1996. P. 10.1–10.11.
127. Van Graas F., Braasch M.S. Real-time attitude and heading determination using GPS // *GPS world*. March 1992. P. 32–39.
128. Povaliaev A.A. Using single differences for relative positioning in GLONASS // *ION GPS*. 1997. P. 929–934.
129. Калмыков В.В., Сенин А.И. Основы теории информации: Учеб. пособие. М.: Изд-во МГТУ им. Н.Э. Баумана, 1992.
130. *Расчет помехоустойчивости систем передачи дискретных сообщений* / В.И. Коржик, Л.М. Финк, К.Н. Щелкунов; Под ред. Л.М. Финка. М.: Радио и связь, 1981.

131. *Помехоустойчивость и эффективность систем передачи информации* / А.Г. Зюко, А.И. Фалько, И.П. Панфилов и др.; Под ред. А.Г. Зюко. М.: Радио и связь, 1985.
132. *Помехоустойчивые коды: Учеб. пособие* / В.В. Калмыков, А.С. Косолапов, Ю.Н. Себекин и др.; Под ред. Сенина А.И. М.: Изд-во МГТУ им. Н.Э. Баумана, 1990.
133. *Кларк Дж. мл., Кейн Дж.* Кодирование с исправлением ошибок в системах цифровой связи: Пер. с англ. М.: Радио и связь, 1987.
134. *Системы фазовой синхронизации* / В.Н. Акимов, Л.Н. Белюстина, В.Н. Белых и др.; Под ред. В.В. Шахгильдяна и Л.Н. Белюстиной. М.: Радио и связь, 1982.
135. *Варакин Л.Е.* Системы связи с шумоподобными сигналами. М.: Радио и связь, 1985.
136. *Варакин Л.Е., Калмыков В.В., Трофимов Ю.К.* Подвижная связь общего пользования // 100 лет радио: Сб. статей / Под ред. В.В. Мигулина. М.: Радио и связь, 1995. С. 171–185.
137. *Варакин Л.Е., Калмыков В.В., Анфилофьев С.А.* На пути к системам подвижной связи третьего поколения в России // Мобильные системы. Спецвыпуск по стандарту CDMA. 1998. С. 13–15.
138. *Невдяев Л.М., Смирнов А.А.* Персональная спутниковая связь. М.: Эко-ТрендзКо, 1998.
139. *Горностаев Ю.М., Невдяев Л.М.* Новые стандарты широкополосной радиосвязи на базе технологий W-CDMA. М.: Межд. Центр НТИ, 1999.



Рис. 1. Внешний вид антенного поста ТРЛК-11

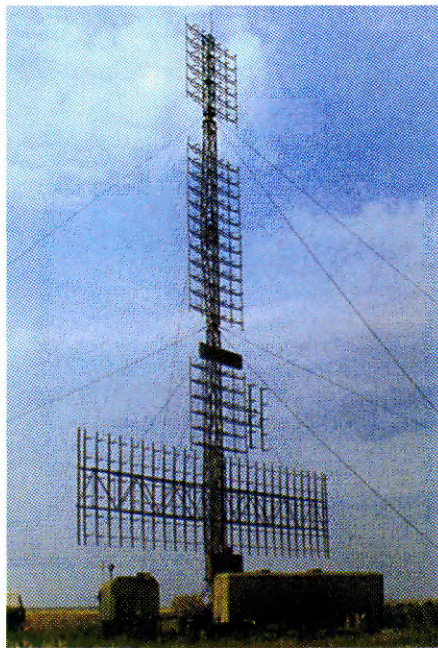


Рис. 2. Радиолокационная станция РЛС 55Ж6-3

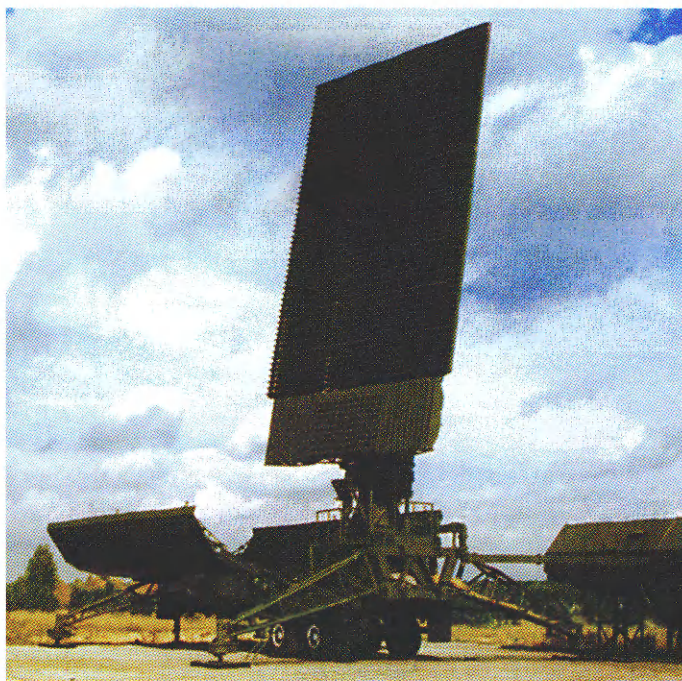


Рис. 3. Радиолокационная станция «Противник-ГЕ»



Рис. 4. Радиолокационная станция 64ЛЕ («Гамма-С1Е»)



Рис. 5. Радиолокационная станция 39Н6Е («Каста-2Е2»)



Рис. 6. Многофункциональный радиолокатор подсвета и наведения ЗРК С-300 ПМУ-1



Рис. 8. Наземная радиолокационная станция «Кредо-1Е»



Рис. 9. Автоматизированный мобильный радиолокационный комплекс «Зоопарк-1»

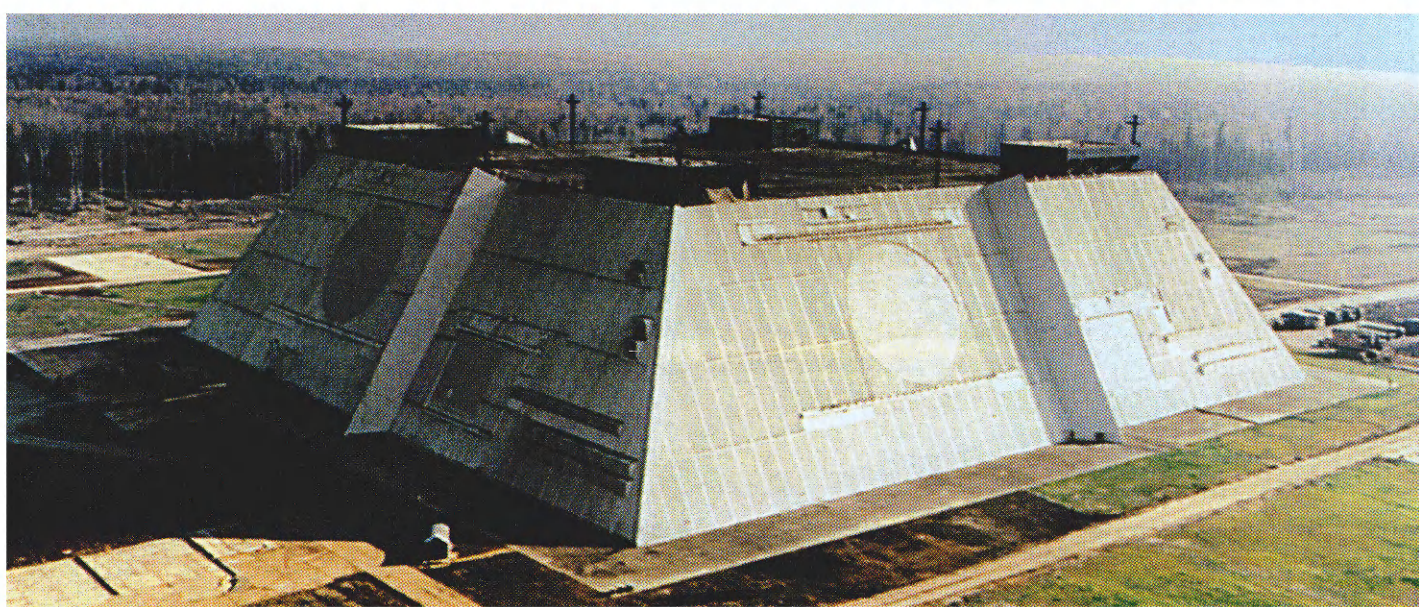


Рис.10. Многофункциональная радиолокационная станция «Дон»



Рис.12. Авиационный комплекс дозора и наведения А-50Э

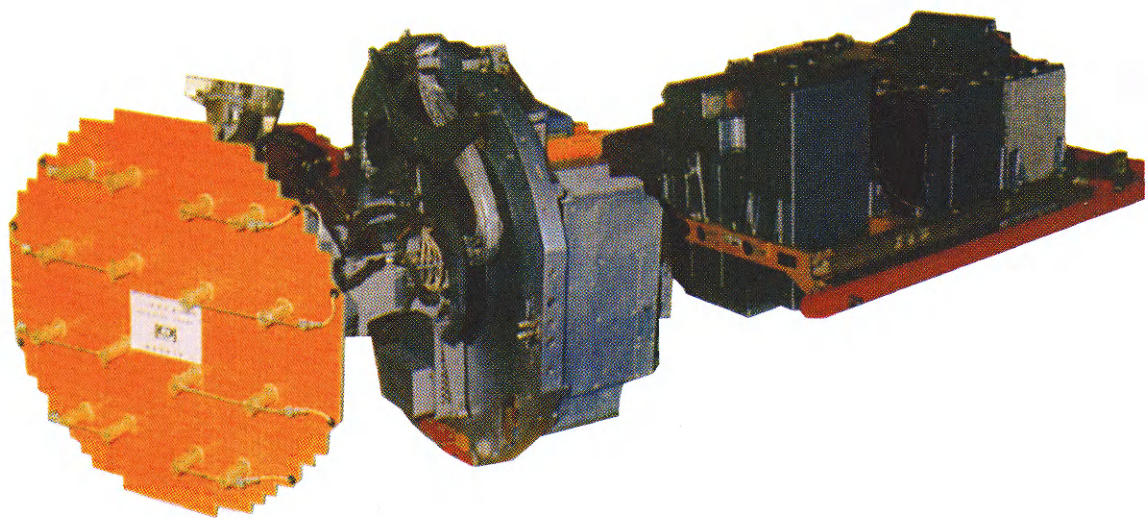


Рис.11. Бортовая радиолокационная станция «Жук»

Учебное издание

Информатика в техническом университете

Васин Валерий Анатольевич
Власов Игорь Борисович
Егоров Юрий Михайлович
Калмыков Вадим Валериевич
Кузнецов Алексей Алексеевич
Николаев Александр Иванович
Пудловский Владимир Борисович
Родзивилов Владимир Афанасьевич
Себекин Юрий Николаевич
Сенин Александр Иванович
Слукин Геннадий Петрович
Федоров Игорь Борисович

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В РАДИОТЕХНИЧЕСКИХ СИСТЕМАХ

Редактор *О.С. Воронина*
Художники *О.В. Левашова, Н.Г. Столярова, С.С. Водчиц*
Коллектив корректоров
Компьютерная верстка *С.Ч. Соколовского*

Подписано в печать 03.04.02. Формат 70×100/16. Печать офсетная.
Бумага офсетная. Гарнитура «Таймс». Печ. л. 42,0 + 0,5 цвет. вклейка.
Усл. печ. л. 55,25. Уч.-изд. л. 53,95.
Тираж 2000 экз. Заказ 7816

Издательство МГТУ им. Н.Э. Баумана.
105005, Москва, 2-я Бауманская, 5.

Отпечатано с оригинал-макета в ГУП ППП «Типография «Наука».
121099, Москва, Шубинский пер., 6.