

П. И. ПЕЧИН, Л. И. ФИЛИППОВ

РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ



П. И. ПЕНИН, Л. И. ФИЛИППОВ

РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

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



МОСКВА «РАДИО И СВЯЗЬ» 1984

ББК 32.81

П25

УДК 621.396.43(075)

Пенин П. И., Филиппов Л. И.
П25 Радиотехнические системы передачи информации: Учеб. пособие для вузов. — М.: Радио и связь, 1984. — 256 с.

В пер.: 90 к.

Изложены основы теории, принципы построения систем передачи непрерывной и дискретной информации. Для каждого вида информации приведены способы построения радиоприемных устройств, особенности искажающего влияния линий и аддитивных помех, а также типы сигналов и виды модуляции в передающих устройствах. Описаны принципы синтеза оптимальных приемных устройств, их инженерной реализации, а также основы построения Единой автоматизированной сети связи. Книга содержит задачи и контрольные вопросы.

Для студентов радиотехнических факультетов вузов.

П 2402020000-019
046(01)-84-112-84

ББК 32.81

6Ф0.1

Рецензенты: кафедра радиотехнических систем Ленинградского ордена Ленина электротехнического института имени В. И. Ульянова (Ленина), зав. кафедрой д-р техн. наук проф. Ю. М. Казаринов; кафедра радиосистем Горьковского политехнического института имени А. А. Жданова (зав. кафедрой д-р техн. наук проф. Ю. С. Лёзин).

Редакция литературы по конструированию
и технологии производства радиоэлектронной аппаратуры

374034

~~10 04485~~

ПЕРЕОБЛІК 1988



ІНТБ ВЛІ
г. ВІННИЦА

Издательство «Радио и связь», 1984

ПРЕДИСЛОВИЕ

В настоящее время курс «Радиотехнические системы передачи информации» играет важную роль при подготовке радиоинженеров. В ряде вузов накоплен опыт чтения этого курса и сложились различные точки зрения на его содержание, структуру, последовательность изложения и взаимосвязь с другими радиотехническими дисциплинами.

Основная цель данного пособия — помочь студентам быстрее освоить основы теории и техники передачи информации и облегчить им изучение многочисленной литературы по этим вопросам при самостоятельной работе.

При подготовке и написании этого учебного пособия авторы столкнулись с рядом трудностей, обусловленных как различными точками зрения на его содержание и построение, так и многообразием вопросов, относящихся к передаче информации.

В основу изложения положен принцип разделения материала по виду передаваемых сообщений. Сначала рассматриваются системы передачи непрерывных сообщений (гл. 4, 5), а затем дискретных (гл. 6—9). Этим основным разделам предшествует изложение некоторых общих вопросов теории систем связи (гл. 1—3). Завершающие главы (гл. 10, 11) посвящены рассмотрению вопросов построения сетей связи, включая единую автоматизированную сеть страны.

Материал, относящийся к передаче как непрерывных, так и дискретных сообщений, излагается по единой методике: сначала рассматриваются вопросы формирования сигналов в передающей части системы, а затем их приема. По методическим соображениям авторы сочли целесообразным рассмотреть физические линии связи (среды распространения сигналов) и их математические модели в главах, относящихся к передаче дискретных сообщений.

В нашей стране быстро развивается Единая автоматизированная сеть связи страны (ЕАСС), в которой наряду с традиционными радиотехническими системами передачи применяются кабельные, волноводные и иные системы. На этом основании авторы истолковали термин «радиотехнические системы» более широко, поэтому в пособии затронут ряд вопросов, которые имеют отношение к любым системам передачи информации.

При подготовке пособия авторы стремились изложить не только наиболее важные результаты, но и способы их получения. Ограниченный объем книги не позволил, однако, реализовать этот принцип в полной мере и некоторые результаты пришлось привести без доказательства, ограничившись лишь их обсуждением.

Для закрепления материала и проверки его усвоения в конце каждой главы приводятся контрольные вопросы и задачи.

Книга может быть полезной не только студентам, но и инженерам, работа которых связана с передачей и обработкой различной информации.

Работа по подготовке пособия была распределена следующим образом. Гл. 1—5 написаны П. И. Пениным, гл. 6, 7, 9—11—Л. И. Филипповым. (Гл. 11 написана по материалам, любезно предоставленным В. М. Дмитриченко.) Гл. 8 и заключение написаны авторами совместно.

Авторы выражают признательность Ю. М. Казаринову, Л. С. Гуткину, Ю. С. Лёзину, В. И. Крамущенко, Л. Я. Новосельцеву, В. Н. Смирнову и другим преподавателям кафедр радиотехнических систем МЭИ, ГПИ, ЛЭТИ, МАИ за труд по ознакомлению с рукописью, ее деловую критику и многие полезные замечания.

ГЛАВА I. ОБЩИЕ СВЕДЕНИЯ О СИСТЕМАХ ПЕРЕДАЧИ ИНФОРМАЦИИ

1.1. ВВЕДЕНИЕ

В нашей стране уделяется большое внимание ускоренному развитию и совершенствованию всех видов связи.

В соответствии с Основными направлениями экономического и социального развития СССР на 1981—1985 годы и на период до 1990 года, принятыми XXVI съездом КПСС, намечен ряд первоочередных задач в области развития связи: повышение качества и расширение услуг всех видов связи; продолжение работ по созданию Единой автоматизированной сети связи страны (ЕАСС); значительное повышение уровня автоматизации линий связи; развертывание работ по организации общегосударственной системы передачи данных (ОГС ПД); дальнейшее развитие радиовещания и телевидения; широкое использование искусственных спутников Земли для организации различных видов связи и решения других важных задач.

Связь в широком смысле представляет собой передачу различного вида сообщений из одного или нескольких пунктов в другой или в ряд других пунктов. *Сообщения* содержат некоторые сведения (информацию), которые для разных получателей могут представлять различную ценность в зависимости от их смыслового содержания. В теории и технике связи семантическая (смысловая) особенность сообщений не учитывается и задачей средств связи является только передача сообщений в определенное место, поскольку оценка смыслового содержания полученных сообщений — дело самого получателя¹.

Теория и техника передачи информации складывались в течение многих лет и в настоящее время продолжают быстро развиваться. Это развитие происходит в условиях усиливающегося взаимодействия между различными областями науки и техники (радиотехники, физики, вычислительной техники, автоматики и т. д.). Такой процесс, являющийся объективным проявлением особенностей современной научно-технической революции, приводит к ломке многих традиций в инженерном мышлении и способствует развитию системного подхода к решению задач, относящихся к проблеме передачи информации.

Современные системы передачи информации — весьма сложные устройства, и их создание связано с большими затратами средств. Поэтому при проектировании систем нужно учитывать перспективы развития техники связи и предусматривать возмож-

¹ Далее в этой книге термины «связь», «передача сообщений» и «передача информации» считаются эквивалентными.

ность использования таких систем не только автономно, но и в составе сложных информационных комплексов, сетей связи и т. п.

Усложнение и совершенствование систем связи и расширение областей их применения заставляют уделять большое внимание вопросам стандартизации и унификации параметров систем не только в пределах одной страны, но и в международном масштабе. Разработкой рекомендаций и предложений в области передачи информации занимается ряд организаций: международный консультативный комитет по телефонии и телеграфии (МККТТ); международный консультативный комитет по радиосвязи (МККР); международная организация по стандартизации (МОС) и другие. В деятельности этих организаций наша страна принимает самое активное участие.

1.2. СИСТЕМА СВЯЗИ. ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ

Системой связи (системой передачи информации) называют функционально объединенную совокупность устройств и звеньев, с помощью которой можно передавать сообщения из одного пункта в другой. Система связи состоит из трех основных частей: передающего устройства, физической линии связи (среды) и приемного устройства (источник сообщений и их получатель в состав системы не входят).

Сообщениями могут быть речь, музыка, тексты, цифровые данные, чертежи, команды управления, результаты измерений параметров в телеметрии и т. д. При передаче такие сообщения преобразуют в электрические колебания с помощью различных преобразователей: микрофонов, передающих телевизионных трубок, датчиков, магнитофонов и т. п.¹

Все сообщения можно разделить на *непрерывные* и *дискретные*.

Источник может создавать множество отличающихся друг от друга сообщений. Каждое конкретное передаваемое сообщение является одним из множества возможных. Поскольку на приемной стороне системы неизвестно заранее, какое из возможных сообщений будет передано, удобно все их рассматривать как реализации $x(t)$ некоторого случайного процесса $X(t)$. Такой статистический подход к передаче сообщений представляет одну из характерных особенностей современной теории связи.

Для уяснения некоторых важных особенностей передачи сообщений рассмотрим обобщенную структурную схему системы связи (рис. 1.1).

Непосредственная передача сообщений по линии связи возможна лишь в некоторых случаях (например, в проводных и кабельных линиях). В большинстве же случаев передаваемые сообщения необходимо преобразовывать в сигналы, которые выпол-

¹ Далее везде под источниками сообщений будем понимать источники первичных сообщений разной природы вместе с их преобразователями в электрическую форму.

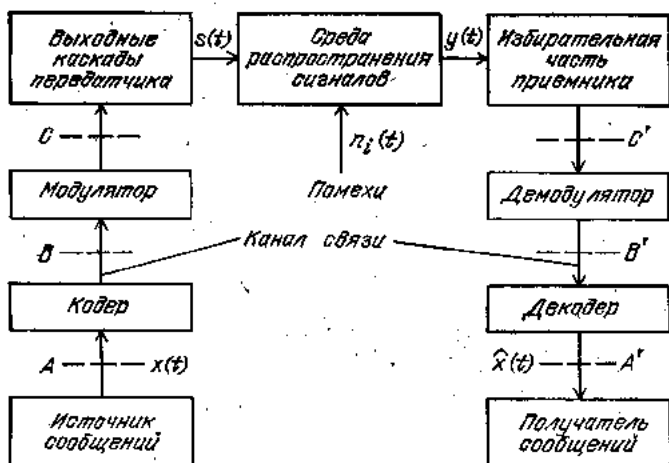


Рис. 1.1. Обобщенная структурная схема системы связи

няют роль переносчика информации по физическим линиям связи. Такое преобразование, выполняемое в передающем устройстве, в общем случае состоит из двух процедур: кодирования и модуляции.

Кодирование позволяет более эффективно использовать систему и уменьшить влияние различных помех и искажений на передачу сообщений. Это сложный процесс, состоящий из ряда операций:

- преобразование сообщений из одной формы в другую, например непрерывных в дискретные (натуральное кодирование);
- устранение естественной избыточности сообщений (эффективное кодирование);
- введение специально рассчитанной искусственной избыточности в сообщения (помехоустойчивое кодирование).

Устройства, которые выполняют эти задачи, называются **кодерами**. Кодеры можно построить либо для каждой из указанных выше операций отдельно, либо объединить их в единое устройство.

Реализация кодирования связана с существенным усложнением систем связи, и поэтому часто ограничиваются только натуральным кодированием. Однако в ряде случаев приходится применять и более сложное кодирование, особенно в телеметрии (эффективное кодирование) и в космических линиях дальней связи (помехоустойчивое кодирование).

Модуляция представляет собой операцию формирования сигналов, находящихся в однозначном соответствии с передаваемыми сообщениями (в общем случае закодированными в указанном выше смысле). Эта операция связана с изменением одного из параметров сигнала-переносчика в соответствии с передаваемым сообщением и выполняется с помощью **модулятора**. Сигна-

лом-переносчиком (несущей) обычно являются периодические колебания.

В выходных каскадах передатчика модулированные сигналы усиливаются и с помощью согласующих устройств (антенн) вводятся (излучаются) в физическую линию связи (среду распространения).

Совокупность операций, связанных с преобразованием передаваемых сообщений $x(t)$ в сигналы $s(t)$, называется *способом формирования сигналов*.

При распространении в реальных средах сигналы не только ослабляются, но и искажаются, так как характеристики таких сред неидеальны.

Кроме сигналов $s(t)$, прошедших среду, на вход приемника системы связи попадают различные аддитивные помехи $n_i(t)$: сигналы других систем, шумы среды и т. п. В результате на его входе действует случайное колебание

$$y(t) = s(t) + \sum_i n_i(t). \quad (1.1)$$

Прием сигналов сводится к выполнению ряда операций над колебанием $y(t)$, в результате которых на выходе приемника должно быть сформировано сообщение, соответствующее переданному. Последовательность этих операций такова: после предварительной фильтрации и усиления в избирательных каскадах приемника принятое колебание $y(t)$ демодулируется и декодируется. Заметим, что эти операции, выполняемые в *демодуляторе* и *декодере*, обратны операциям в кодере и модуляторе.

Совокупность операций, связанных с преобразованием принятой смеси сигнала и помех в сообщение, называется *способом приема*.

Из-за воздействия помех, влияния среды распространения, неидеальности и нестабильности характеристик системы полного соответствия между переданным $x(t)$ и принятым $\hat{x}(t)$ сообщениями может и не быть, т. е. сообщение воспроизводится с некоторой ошибкой, величина которой зависит от уровня сигнала и помех в канале и их статистических характеристик, а также от качества процедур формирования сигналов и их приема.

При рассмотрении вопросов передачи информации часто пользуются термином *канал связи*, понимая под этим часть системы между некоторыми точками (сечениями), взятыми в передающей и приемных частях (например, между сечениями $A-A'$ или $B-B'$, или $C-C'$ на рис. 1.1). В зависимости от вида передаваемых сообщений и выбора места сечений канал может быть *дискретным* или *непрерывным*. В дискретном канале входные и выходные воздействия дискретны, а в непрерывном — непрерывны.

Такой подход удобен при теоретических исследованиях, так как в зависимости от специфики решаемой задачи позволяет пользоваться той или иной *моделью* (описанием) канала. Непрерывный канал обычно описывают занимаемой полосой частот и

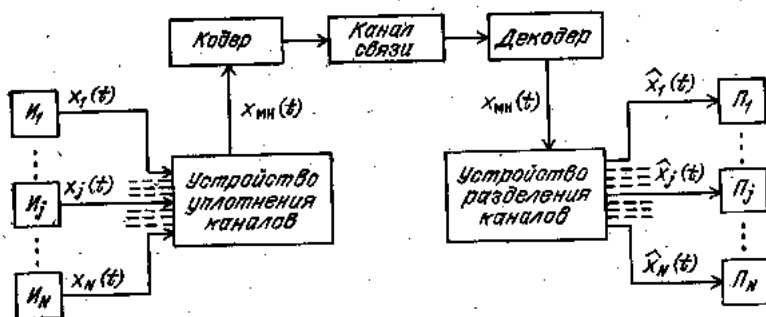


Рис. 1.2. Структурная схема многоканальной системы связи

видом действующих в нем сигналов и помех, а дискретный — стохастической матрицей трансформации входных символов в выходные. Эта матрица определяет все возможные вероятности переходов элементов входных сообщений в элементы выходных сообщений.

Часто системы связи используют для передачи сообщений не от одного источника, а от многих источников одновременно. Такие системы называют *многоканальными*. Для их реализации канал связи уплотняют с помощью устройств уплотнения (рис. 1.2). Суть этой операции состоит в том, что сообщения от многих источников в устройстве уплотнения подвергают преобразованиям, в результате которых на его выходе образуется сложное колебание $x_{mn}(t)$, содержащее в преобразованном виде все исходные сообщения:

$$x_1(t), x_2(t), \dots, x_j(t), \dots, x_N(t),$$

где $x_j(t)$ — сообщение от j -го источника; N — число источников.

Колебание $x_{mn}(t)$ называют *групповым сигналом*, а его получение — *уплотнением каналов*. При многоканальной передаче групповой сигнал поступает на вход кодера (рис. 1.2). Дальнейшие процедуры, связанные с формированием сигнала-переносчика информации и его приемом, в принципе не отличаются от рассмотренных ранее для одного источника. По сути дела групповой сигнал $x_{mn}(t)$ представляет собой некоторое сложное сообщение, которое можно назвать *многоканальным сообщением*.

Особенностью приемной части многоканальной системы связи является то, что она должна содержать устройство разделения каналов, позволяющее преобразовать принятое колебание $\hat{x}_{mn}(t)$ в отдельные принятые сообщения $\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_j(t), \dots, \hat{x}_N(t)$ и доставить их соответствующим получателям. Такое преобразование называют *разделением каналов*.

Важным обстоятельством при многоканальной передаче является то, что часть системы оказывается общей для всех источников и получателей информации. Эта часть называется *общим (групповым) трактом системы*. Неидеальность характеристик это-

го тракта приводит к появлению дополнительных искажений, проявляющихся во взаимном влиянии отдельных каналов друг на друга. Такое влияние ухудшает работу системы и снижает ее показатели.

1.3. ПОКАЗАТЕЛИ КАЧЕСТВА СИСТЕМ СВЯЗИ

Качество работы системы связи можно характеризовать рядом показателей. К наиболее важным из них относятся: *достоверность передачи сообщений, помехоустойчивость, скорость, передача и эффективность системы.*

Достоверность передачи сообщений. Передача сообщений сопровождается неизбежным действием помех и искажений, обусловленных неидеальностью характеристик системы связи. Поэтому принятое сообщение может отличаться от переданного. Степень соответствия принятого сообщения переданному называют *достоверностью (точностью) передачи.*

При передаче дискретных сообщений это соответствие количественно можно оценить отношением числа ошибочно принятых элементов сообщения $n_{\text{ош}}$ к общему числу переданных элементов $n_{\text{общ}}$:

$$k_{\text{ош}} = n_{\text{ош}}/n_{\text{общ}}. \quad (1.2)$$

Это отношение называется частотой ошибок (коэффициентом ошибок). При ограниченном времени передачи (конечном числе $n_{\text{общ}}$) величина $k_{\text{ош}}$ случайна. Если общее время передачи информации (длительность сеанса связи) значительно превышает длительность передачи отдельного элемента, а условия передачи остаются неизменными, то статистические характеристики случайной величины $k_{\text{ош}}$ весьма устойчивы, т. е. их изменения от сеанса к сеансу невелики. Поэтому в первом приближении величину $k_{\text{ош}}$ можно считать близкой к вероятности ошибки приема элемента сообщения $P_{\text{ош}}$. Оценка достоверности передачи дискретных сообщений вероятностью $P_{\text{ош}}$ получила широкое распространение из-за своей простоты и удобства. Однако необходимо помнить, что этот показатель является идеализированным и условия его применимости не всегда строго выполняются.

При передаче непрерывных сообщений различие между переданным сообщением $x(t)$ и принятым $\hat{x}(t)$ определяется величиной *случайной ошибки:*

$$e_x(t) = \hat{x}(t) - x(t). \quad (1.3)$$

Записав (1.3) в виде

$$\hat{x}(t) = x(t) + e_x(t), \quad (1.4)$$

ошибку $e_x(t)$ можно рассматривать как случайную помеху на выходе системы связи. Для ее полной характеристики нужно знать многомерную плотность распределения вероятностей. Во многих случаях это связано с большими трудностями и поэтому обычно

при оценке точности передачи непрерывных сообщений ограничиваются *критерием относительной среднеквадратической ошибки* (критерий СКО). В соответствии с этим критерием квадрат относительной среднеквадратической ошибки определяется величиной

$$\delta^2 = P_d/P_x = (P_n/P_c)_{\text{вых}}, \quad (1.5)$$

которая представляет собой отношение средней мощности помехи $P_n = P_n$ к средней мощности сообщения $P_c = P_x$ на выходе системы или на выходе отдельного канала при многоканальной передаче¹. Критерий СКО особенно целесообразно использовать в тех случаях, когда нежелательность ошибки возрастает с ее величиной. Кроме того, для практики очень ценным является свойство аддитивности критерия, суть которого состоит в том, что при действии независимых источников помех квадрат ошибки можно находить суммированием квадратов ее отдельных составляющих

$$\delta^2 = \sum_i \delta_i^2. \quad (1.6)$$

Физический смысл этого свойства состоит в том, что при одновременном действии нескольких независимых мешающих факторов средняя мощность помехи на выходе канала определяется суммой средних мощностей ее отдельных составляющих².

В некоторых случаях при оценке качества передачи непрерывных сообщений целесообразно помимо величины СКО вычислять также вероятность того, что абсолютная величина ошибки e_x не превысит заданного порогового значения Δ , т. е.

$$P(|e_x| \leq \Delta) = \int_{-\Delta}^{\Delta} p_1(e_x) d e_x, \quad (1.7)$$

где $p_1(e_x)$ — одномерная плотность распределения вероятности помехи.

Физически это условие соответствует вероятности отсутствия так называемых *аномальных ошибок*, т. е. таких ошибок, которые могут повлечь за собой существенные последствия для получателя (кратковременные мешающие воздействия, проявляющиеся на выходе канала как импульсная помеха).

Помехоустойчивость. Передача информации с требуемой достоверностью предполагает надежную работу системы связи. Это возможно, если система обладает высокой *надежностью*, т. е. способностью приборов длительно выполнять возложенные на них

¹ Далее везде под средней мощностью любого колебания понимается его средняя удельная мощность, т. е. мощность, развиваемая в среднем этим колебанием на активной нагрузке 1 Ом. Это позволяет упростить запись соответствующих выражений и не влияет на результаты рассмотрения.

² Строго говоря, свойство аддитивности критерия СКО выполняется только для линейных систем. Однако при условии, что $\delta^2 \ll 1$, это свойство можно считать справедливым и для систем связи, так как при этом влияние нелинейных элементов системы не приводит к появлению заметной зависимости между отдельными составляющими ошибки на выходе.

функции, и обеспечивает при этом необходимую *помехоустойчивость*, т. е. способность противостоять мешающему действию различных помех и искажений. Надежность и помехоустойчивость системы связи зависят от многих факторов: способов практической реализации системы, элементной базы, технологии изготовления аппаратуры, условий эксплуатации, принципов построения системы и т. д. Надежность системы количественно оценивается вероятностью того, что аппаратура в целом будет выполнять предназначенные ей функции в течение заданного времени при заданных условиях работы.

Для оценки помехоустойчивости системы связи применяют различные критерии. Наиболее распространенным из них является критерий, в соответствии с которым помехоустойчивость оценивается требуемым *отношением средних мощностей сигнала и помехи на входе* приемника системы $q_{вх} = (P_c/P_n)_{вх}$, обеспечивающим заданную достоверность передачи δ^2 или $P_{ош}$. Чем меньше требуется величина $q_{вх}$, тем выше помехоустойчивость системы.

Количественная оценка помехоустойчивости системы связи и сравнение различных систем по этому показателю производится с помощью зависимостей

$$P_{ош} = f(q_{вх}) \text{ или } \delta^2 = \psi(q_{вх}),$$

которые функционально связывают достоверность передачи информации с величиной отношения мощностей сигнала и помехи на входе приемника $q_{вх}$. Отыскание этих зависимостей представляет одну из главных задач теории помехоустойчивости.

Скорость передачи информации. Одни и те же сообщения можно передать с помощью различных систем связи с различной скоростью. Передача непрерывных сообщений обычно осуществляется в реальном масштабе времени, т. е. в темпе, определяемом естественной скоростью создания таких сообщений источником. Однако в некоторых случаях целесообразно сообщения от источника записать (например, на магнитофон или другое подобное устройство), а затем передать со скоростью, отличающейся (в большую или меньшую сторону) от скорости их создания.

Такой принцип передачи позволяет эффективнее использовать каналы связи, организовать более гибкую связь и получить ряд других преимуществ. Системы, в которых реализуется подобный принцип, называют системами с задержкой сообщений. К подобным системам можно отнести некоторые спутниковые и космические системы связи.

Численно *скорость передачи информации* \mathcal{R} определяется количеством информации, поступившей от отправителя к получателю за 1 с. *Количество информации* обычно измеряют в *двоичных единицах (битах)*¹. Поэтому скорость передачи определяет

¹ При необходимости используют и более крупные единицы измерения информации: килобиты (Кбит) и мегабиты (Мбит). В вычислительной технике часто применяют единицу, называемую *байтом* (один байт, как правило, принимают равным восьми битам).

ся числом бит в секунду. Эта скорость зависит от ряда факторов: статистических свойств передаваемых сообщений, характеристик канала связи, искажений, возникающих при передаче информации за счет неидеальности и нестабильности характеристик системы связи.

При передаче дискретной информации для характеристики быстрой действия аппаратуры, формирующей информационный поток в передающей части системы, пользуются понятием *техническая скорость передачи*. Она определяется количеством элементарных посылок дискретного сообщения, формируемых в единицу времени, и измеряется в бодах, т. е.

$$R_T = 1/\tau_0, \quad (1.8)$$

где τ_0 — длительность передачи одной элементарной посылки дискретного сообщения.

Предельные возможности скорости передачи информации по каналу связи удобно оценивать величиной *пропускной способности канала*. Пропускная способность канала C численно определяется максимальным количеством информации, которое может быть передано по нему в 1 с. Пропускная способность является одной из наиболее важных предельных характеристик канала связи.

Эффективность системы связи. Передача информации с требуемой достоверностью и заданной скоростью еще не может служить достаточным основанием для оценки качества работы рассматриваемой системы связи. Одни и те же достоверность и скорость передачи информации можно получить, применяя разные системы. Желательно иметь такую систему связи, которая обеспечивала бы передачу информации с заданной точностью и скоростью наиболее экономно, т. е. с наименьшими «затратами». Под *затратами* можно понимать мощность сигнала, полосу частот, стоимость аппаратуры, ее габаритные размеры, массу и т. д., которые необходимы для передачи определенного количества информации (например, одной двоичной единицы). Совокупность свойств, характеризующих экономичность системы с точки зрения указанных затрат, называют *эффективностью системы*. Поэтому величину затрат можно рассматривать как меру эффективности системы.

Для сравнения различных систем связи по эффективности нужно выбрать совокупность затрат различного вида, ввести меру их количественной оценки и указать критерии, в соответствии с которыми можно принять решение о предпочтительности одной системы по сравнению с другой. Как правило, выбор критериев приходится делать с учетом ряда противоречивых требований: с одной стороны, критерии должны учитывать особенности работы конкретной системы связи и предъявляемые к ней требования, а с другой — быть простыми, удобными и пригодными для различных систем передачи информации. При выборе критериев необходимо также учитывать определенные, заранее устанавливаемые

ограничения на некоторые параметры и характеристики системы. В одних системах, например в системах дальней космической связи, ограничивающим фактором первостепенной важности является величина средней или пиковой мощности передатчика, а в других — например в системах передачи данных по стандартным телефонным каналам, в системах коротковолновой радиосвязи и других — используемая полоса частот. Имеются системы, в которых необходимо учитывать не один, а ряд ограничивающих факторов (габаритные размеры, массу, излучаемую мощность и т. п.).

Выбрать критерии, универсальные для любых систем передачи информации, практически невозможно. Поэтому наряду с некоторыми сравнительно общими критериями иногда приходится пользоваться и более частными критериями, приспособленными для данного конкретного класса систем.

Наиболее широкое применение в теории связи и в прикладных задачах получили критерии, в соответствии с которыми система связи оценивается величиной затрат на передачу двоичной единицы информации при заданной точности приема сообщений. Такие критерии называют критериями удельных затрат. В соответствии с ними более эффективной считается та система, которая характеризуется меньшими удельными затратами.

Показатели качества работы любой системы связи определяются как принципами и методами, положенными в основу ее построения, так и конкретными способами и техническими особенностями реализации. Общее число показателей, характеризующих работу системы связи, может быть значительным.

Основное внимание в этой книге будет уделяться показателям, характеризующим принципы и методы построения систем. Из всех видов затрат ниже будут учитываться только удельные энергетические затраты и удельные затраты полосы.

Удельные энергетические затраты характеризуют величиной

$$\beta_E = E_0/N_0, \quad (1.9)$$

где E_0 — энергия сигнала на входе приемника, затраченная на передачу одного бита информации с заданной точностью; N_0 — спектральная плотность помехи на входе приемника (энергетический спектр помехи).

Удельные затраты полосы характеризуют величиной

$$\beta_{\Delta f} = \Delta f_{\text{экв}}/\mathcal{R}, \quad (1.10)$$

где \mathcal{R} — скорость передачи информации (бит/с); $\Delta f_{\text{экв}}$ — эквивалентная полоса пропускания приемника системы связи (полоса частот, занимаемая каналом связи), соответствующая ширине спектра используемого сигнала.

Значения β_E и $\beta_{\Delta f}$ можно рассматривать как показатели работы системы связи, оцениваемой в соответствии с критериями удельных затрат.

1.4. КЛАССИФИКАЦИЯ СИСТЕМ И ЛИНИЙ ПЕРЕДАЧИ ИНФОРМАЦИИ

Системы передачи информации характеризуются большим разнообразием областей применения, принципов и методов построения, видов передаваемых сообщений и т. д. В связи с этим они могут быть классифицированы по различным признакам. Обычно учитываются:

область применения (телефония, передача дискретных сообщений, телеметрия, управление, телевидение и т. д.);

форма представления передаваемых сообщений (дискретная или непрерывная);

вид сигналов (непрерывное или импульсное излучение, класс используемых сигналов);

диапазон рабочих частот и абсолютная ширина полосы частот, занимаемой системой;

вид службы связи (стационарные или подвижные линии связи, ведомственная связь и др.);

принцип уплотнения и разделения каналов при многоканальной передаче.

Независимо от приведенной классификации все системы передачи информации удобно разделить на две группы.

Системы со свободным распространением сигналов между передатчиком и приемником. К этой группе относятся радиосистемы, оптические (лазерные), акустические и гидроакустические системы. Характерной особенностью таких систем связи является рассеяние энергии при распространении в среде, пропорциональное квадрату дальности связи.

Системы с направляемым распространением сигналов. Сюда входят проводные, кабельные, волноводные и световодные системы связи. Характерной особенностью таких систем является принудительная канализация сигналов в заданном направлении с помощью направляющих устройств (кабелей, волноводов и т. п.). Энергия сигналов в таких системах практически не рассеивается и лишь поглощается (затухает) в направляющих устройствах.

В системах с направляемым распространением сигналов можно обеспечить высокую стабильность параметров, и следовательно, передавать информацию с высокой достоверностью. Большим достоинством таких систем является почти идеальное решение проблемы электромагнитной совместимости, поскольку взаимное влияние различных систем друг на друга практически отсутствует, даже если эти системы находятся рядом. Кабельные, волноводные и световодные системы связи обладают очень большой пропускной способностью и могут обеспечить передачу огромных потоков информации, измеряемых сотнями и тысячами мегабит в секунду. Однако эти системы связи дороги и их применение в ряде случаев экономически нецелесообразно, особенно на участках большой протяженности, где для компенсации затухания сигнала в линии необходимо ставить значительное число промежуточных усилительно-ретрансляционных пунктов.

Среди систем с направляемым распространением сигналов наиболее широко используются кабельные, особенно в экономически развитых районах с высокой плотностью населения. Ведутся интенсивные исследования, как в нашей стране, так и за рубежом (Япония, ФРГ, США), по использованию волноводных и световодных линий связи [1].

Системы связи со свободным распространением сигналов более подвержены действию различных внешних факторов (помех, изменяющихся условий распространения сигналов и т. д.). Кроме того, в этих системах часто возникает проблема электромагнитной совместимости систем, работающих в близких диапазонах частот. Несмотря на это, системы со свободным распространением сигналов (особенно радиосистемы) в настоящее время применяются очень широко. Такие системы экономичны и гибки с точки зрения организации связи между различными пунктами (в ряде случаев их применение является единственной возможностью реализовать связь в труднодоступных местах).

По условиям работы все системы связи со свободным распространением сигналов можно разбить на две группы: системы с постоянными параметрами и системы со случайными параметрами.

К системам связи с постоянными параметрами относятся системы, в которых за время сеанса связи параметры сигнала при прохождении через среду распространения не претерпевают сколько-нибудь существенных случайных изменений, т. е. практически остаются неизменными (за исключением, может быть, начальной фазы сигнала). В таких системах основное влияние на качество передачи информации оказывают флуктуационные шумы, а также нестабильность параметров и неидеальность характеристик аппаратуры. К ним можно с некоторым приближением отнести многие реальные системы, работающие в дециметровом и сантиметровом диапазонах волн в пределах прямой видимости (радиорелейные линии с промежуточными ретрансляторами; спутниковые линии связи; космические линии связи).

К системам связи со случайными параметрами относятся системы, в которых за время сеанса связи параметры сигнала при прохождении через среду претерпевают существенные случайные изменения. Эти изменения могут быть вызваны наличием неоднородностей в среде; перемещением положения диаграммы направленности антенны передатчика или приемника при движении объектов, между которыми устанавливается связь, и ряд других трудно учитываемых факторов.

Среди систем со случайными параметрами и свободно распространяющимися сигналами можно выделить системы, работающие либо с прямой, либо с отраженно-рассеянной волной.

Условия работы систем со случайными параметрами и прямой волной характеризуются тем, что существует прямая видимость между передатчиком и приемником, а сигнал распространяется в условиях случайных изменений параметров среды. При-

мером подобных систем являются радиорелейные линии связи, работающие во время выпадения осадков, наземные линии оптической связи, гидроакустические линии.

Системы, использующие отраженно-рассеянные волны, функционируют в условиях, когда прямая видимость между передатчиком и приемником отсутствует. Принимаемые сигналы в таких системах являются результатом рассеяния передаваемых сигналов некоторым объемом среды с неоднородными свойствами или отражений от границ раздела сред с различными свойствами. К подобным системам относятся тропосферные, ионосферные и коротковолновые линии связи. Сигналы в таких линиях претерпевают глубокие замирания (случайные изменения уровня), при этом время прохождения сигналов в среде (время запаздывания) становится случайным.

1.5. ОСНОВНЫЕ ВИДЫ РАДИОТЕХНИЧЕСКИХ СИСТЕМ ПЕРЕДАЧИ ИНФОРМАЦИИ

Необходимость надежной и высококачественной передачи значительных потоков информации обусловила широкое использование в современных радиотехнических системах передачи информации радиосигналов метрового, дециметрового и сантиметрового диапазонов волн. Как известно, характер распространения сигналов существенно зависит от длины их волны. При определенных условиях волны длиной $\lambda = 3-10$ м могут хорошо отражаться от ионосферы, что позволяет принимать их на весьма больших расстояниях (до 1800—2000 км). Волны короче 3 м практически распространяются в основном только в пределах прямой видимости. Для волн короче 3 м характерно значительное поглощение их энергии в атмосфере, особенно при неблагоприятных метеорологических условиях (туман, осадки, пыль и др.).

Наиболее широкое применение в радиотехнических системах находят волны от 3 м до 1 см ($100 \text{ МГц} \leq f \leq 30 \text{ ГГц}$): Важными особенностями таких радиоволн являются:

значительный частотный диапазон, позволяющий разместить большое количество различных систем и использовать в них широкополосные виды модуляции;

распространение в пределах геометрической видимости, позволяющее уменьшить взаимное влияние различных систем и использовать одни и те же диапазоны частот в системах, расположенных вне зон видимости;

малый уровень промышленных и атмосферных помех по сравнению с внутренними флуктуационными шумами системы;

возможность реализации сравнительно малогабаритных антенных устройств с узкой диаграммой направленности. Это позволяет существенно ослабить мешающее влияние сторонних радиосистем, а также уменьшить мощность передатчика, необходимую для нормальной работы системы.

г. ЕННННЦА



Передача информации в дециметровом или сантиметровом диапазоне волн на большие расстояния обеспечивается радиорелейными линиями с промежуточными ретрансляционными пунктами, а также спутниковыми системами связи, в которых ретранслятор размещен на искусственном спутнике Земли. В этих системах осуществляется прием прямой волны. Наряду с такими системами получили также некоторое распространение тропосферные, ионосферные и метеорные системы связи, в которых используются отраженно-рассеянные волны.

Остановимся кратко на характеристиках некоторых видов радиотехнических систем и соответствующих им линий связи.

Радиорелейные системы связи. Эти системы относятся к одному из основных видов современных средств радиосвязи в нашей стране, и их развитию уделяется большое внимание. Они обеспечивают высококачественную и надежную передачу сообщений различного вида: телефонных, телеграфных, факсимильных, телевизионных, цифровых данных.

В радиорелейных системах обычно используют сигналы сантиметрового диапазона волн, которые распространяются только в пределах прямой видимости. Поэтому для организации связи на большие расстояния радиорелейные линии связи (РРЛ) приходится выполнять в виде цепочки *ретрансляторов*, в каждом из которых ослабленные сигналы усиливаются и передаются к следующему ретранслятору.

При двусторонней (дуплексной) связи на каждом ретрансляторе необходимо иметь две приемно-передающие станции, осуществляющие передачу и прием на разных частотах. Такое двухчастотное построение РРЛ позволяет существенно ослабить влияние передатчика на приемник, так как частоты приема и передачи разнесены на несколько сотен мегагерц. Недостатком двухчастотной системы является то, что в приемник одного направления связи может попасть сигнал с противоположного направления. Для уменьшения этого эффекта необходимо значительно (на 70—80 дБ) подавлять боковые и задние лепестки диаграмм направленности антенн, а также размещать все ретрансляторы не на одной прямой линии, а «змейкой».

С целью повышения технико-экономических показателей РРЛ и увеличения ее пропускной способности на каждом ретрансляторе обычно устанавливают несколько комплектов приемно-передающей аппаратуры для разных рабочих частот. Это позволяет образовать несколько независимых широкополосных высокочастотных радиоканалов. Такой радиоканал принято называть высокочастотным стволом или *радиостволом* РРЛ. Современные магистральные РРЛ имеют до восьми радиостволов (обычно один-два ствола используют для автоматического резервирования рабочих стволів).

В стационарных РРЛ применяют остронаправленные антенны с коэффициентами направленности 30—40 дБ и более. Это позволяет обеспечивать надежную связь при малых мощностях пере-

датчиков ретрансляторов (не превышающих нескольких ватт или даже долей ватт).

В настоящее время на отечественных магистральных линиях широко используются радиорелейные системы «Восход» и «Дружба», полностью удовлетворяющие современным требованиям. Эти системы позволяют передавать в одном стволе до 1920 телефонных каналов или одну телевизионную программу и несколько каналов звукового сопровождения, обеспечивая дальность связи до 12,5 тыс. км; они работают на частотах 3400—3900 МГц («Восход») и 5670—6170 МГц («Дружба»).

Наличие двух современных радиорелейных систем с различными диапазонами позволяет организовать гибкие и эффективные линии и сети связи.

Большое значение в развитии и совершенствовании линий связи имеет разработанный в СССР комплекс унифицированных радиорелейных систем (КУРС). Этот экономичный высококачественный и надежный комплекс содержит четыре системы связи, работающие в диапазонах 2, 4, 6, и 8 ГГц. Диапазоны 4 и 6 ГГц используются для магистральных линий, а диапазон 2 и 8 ГГц — для организации внутренних линий связи местного значения. Все системы комплекса построены из унифицированных блоков, выполненных на современной элементной базе. Наряду с передачей непрерывных сообщений в комплексе КУРС предусмотрена передача цифровой информации со скоростями 2,048 и 8,448 Мбит/с [1, 2].

Спутниковые системы связи. Искусственные спутники Земли (ИСЗ) открыли исключительно широкие возможности для развития связи. Размещение *активного ретранслятора* на ИСЗ позволяет организовать связь на большие расстояния не только между фиксированными пунктами, но и между подвижными объектами (судами, самолетами, космическими кораблями) [1, 2].

Характерной особенностью спутниковых систем связи является принцип *многостанционного доступа*, суть которого состоит в том, что ретранслятор спутника может использоваться одновременно несколькими наземными станциями, т. е. спутниковый канал связи является общим для нескольких одновременно работающих линий связи.

Разработан ряд методов многостанционного доступа, основанных на частотном, временном, кодовом и пространственном уплотнении линий связи, использующих спутниковый канал.

В настоящее время наиболее распространен многостанционный доступ с частотным уплотнением. При этом методе каждой наземной передающей станции выделяется своя несущая частота и соответствующая полоса пропускания. При одновременной работе нескольких станций в ретрансляторе спутника могут возникать нелинейные искажения, которые приводят к взаимным (перекрестным) помехам между станциями. Для снижения уровня этих помех необходимо уменьшать мощность наземных передатчиков спутниковой системы связи, что приводит к уменьше-

нию отношения сигнала к шуму на входе приемника линии связи. Поэтому при частотном уплотнении стремятся применять малощумящие приемники с порогопонижающими (квазиоптимальными) схемами демодуляторов. Однако даже при оптимизации параметров влияние нелинейных искажений в ретрансляторе может при частотном уплотнении приводить к энергетическому проигрышу на 5—7 дБ по сравнению с доступом с временным уплотнением.

При доступе с временным уплотнением каждой наземной станции выделяется определенный временной интервал для передачи, и все станции работают на одной и той же частоте несущей. Эффективность использования ретранслятора при этом существенно ухудшается, а перекрестные помехи значительно ослабляются. Несмотря на усложнение наземного оборудования и необходимость синхронизации работы наземных станций, метод доступа с временным уплотнением уже находит заметное применение в практике. Это объясняется тем, что он более эффективен, чем метод с частотным уплотнением, и хорошо сочетается с цифровыми системами связи, которые интенсивно развиваются в последние годы.

Кроме указанных выше в спутниковой связи находит также применение многостанционный доступ с кодовым уплотнением (уплотнение по форме). Для такого метода характерно применение сложных (широкополосных) сигналов. Передаваемая с каждой наземной станции информация представляется в цифровой форме и кодируется в виде псевдослучайных последовательностей, частотный спектр которых занимает всю полосу пропускания ретранслятора. На приемной станции спутниковой системы связи имеются образцы используемых последовательностей, которые позволяют с помощью корреляторов выделять только «свои» сигналы из входной смеси шума и сигналов других передающих станций. Доступ с кодовым уплотнением позволяет создавать спутниковые системы связи с повышенной помехоустойчивостью и скрытностью работы.

Развитию спутниковых систем связи в нашей стране уделяется очень большое внимание. С помощью ИСЗ «Молния» создана разветвленная сеть телевизионного вещания «Орбита», которая обслуживает десятки миллионов жителей Средней Азии, Сибири и Дальнего Востока. Большую роль в укреплении и развитии культурных связей между СССР и социалистическими странами играет система «Интерспутник».

В современных спутниковых системах связи используются диапазоны 4 и 6 ГГц, мощности наземных передатчиков составляют несколько киловатт (3—10 кВт), мощности активных ретрансляторов — десятки и сотни ватт, полосы пропускания — десятки и сотни мегагерц [1, 3].

Усовершенствования систем спутниковой связи ведутся в направлении повышения надежности и увеличения пропускной способности за счет применения более высокочастотных диапазонов волн (вплоть до миллиметровых).

Космические системы связи. В связи с развитием ракетно-космической техники и непрерывным расширением программ космических исследований перед теорией и техникой радиосвязи возник ряд новых и сложных задач:

обеспечение надежной связи при дальности в сотни миллионов километров;

передача с космических аппаратов большого объема разнообразной информации о состоянии и работе бортовых систем и агрегатов, а также о различных процессах, происходящих в космическом пространстве и на поверхности исследуемых планет;

управление режимом работы космических аппаратов с помощью передачи командной информации с Земли.

В настоящее время космические системы обеспечивают связь с ИСЗ, орбитальными космическими станциями, с автоматическими станциями для исследования Луны и планет солнечной системы.

Как правило, космические системы связи являются совмещенными, т. е. позволяют одновременно решать задачи связи, телеметрии, измерения параметров движения (траекторные измерения), управления. К системам космической связи предъявляются очень высокие требования. Они должны надежно работать в течение длительного времени; обеспечивать высокую точность передачи информации; обладать высокой энергетической эффективностью; иметь малые габаритные размеры и массу.

Использование современной элементной базы (больших и сверхбольших интегральных схем и микропроцессоров) и методов передачи цифровой информации позволяет успешно решать такие задачи, которые совсем недавно казались нереальными.

Тропосферные, ионосферные и метеорные системы связи. Такие системы связи основаны на использовании сигналов, которые являются результатом отражения и рассеяния излученных сигналов некоторым объемом неоднородной случайной среды, выполняющим роль пассивного ретранслятора.

Рассеивающие слои в тропосфере находятся на высотах до 8—12 км, и расстояния между соседними станциями в тропосферных системах могут составлять 150—600 км и более. При передаче информации на большие расстояния в таких системах приходится использовать промежуточные приемно-передающие станции (так же как в радиорелейных линиях).

Значительное ослабление сигналов на интервале распространения и их глубокие замирания приводят к необходимости обеспечивать высокий энергетический потенциал системы связи¹. Это достигается за счет остронаправленных антенн (с коэффициентами направленности до 50—55 дБ), большой мощности передатчиков (до нескольких киловатт), использования маломощных приемников с эквивалентной шумовой температурой 50—70 К.

¹ Энергетическим потенциалом системы называют отношение излучаемой мощности сигнала к минимальной мощности сигнала на входе приемника, при которой обеспечивается нормальное функционирование системы.

Обычно тропосферные системы работают на частотах 0,3—5 ГГц. Замирания сигналов и эффект многолучевого распространения приводит к тому, что эффективно используемая полоса частот в тропосферных линиях связи (на один ствол) не превышает нескольких мегагерц (2—4 МГц). Такая полоса обеспечивает передачу до 120 аналоговых телефонных каналов в одном стволе. Для борьбы с замираниями в тропосферных системах применяется разнесенный прием.

Ионосферные системы связи основаны на использовании явления рассеяния сигналов неоднородностями ионосферы на высотах 75—95 км. Это позволяет обеспечить дальность связи до 900—2000 км. Наиболее часто применяемые частоты лежат в диапазоне 30—60 МГц.

Использование более высоких частот приводит к резкому уменьшению интенсивности рассеяния сигналов, а более низких частот — к появлению глубоких замираний и усилению эффекта многолучевого распространения.

Метеорные системы связи основаны на использовании отражений радиоволн метрового диапазона от ионосферных следов, оставляемых на высотах 80—120 км огромным количеством сгорающих там метеороидов. Концентрация ионосферных следов имеет нестационарный случайный характер и резко возрастает в некоторые моменты времени. Это приводит к «вспышкам» уровня сигналов, отраженных от следов. Особенностью метеорных систем связи является прерывистый режим их работы, и информация передается только во время «вспышек» уровня сигналов. Наличие «вспышек» устанавливается специальными устройствами, которые входят в состав системы [1].

Скорость передачи информации по ионосферным и метеорным линиям связи невелика и составляет 500—1000 Бод. Несмотря на это такие линии имеют очень важное значение для организации некоторых видов связи в труднодоступных местах и т. п. [1].

Коротковолновые системы связи. В этих системах связи используется декаметровый диапазон волн (частоты от 3 до 30 МГц). Особенностью волн этого диапазона является их распространение на значительные расстояния за счет многократного отражения от ионосферных слоев и поверхности Земли. Такая природа распространения приводит к многолучевости электромагнитного поля в точке приема. В результате принимаемые сигналы испытывают значительные изменения амплитуды и фазы, что существенно усложняет прием и ограничивает возможности передачи информации по коротковолновым линиям связи. В настоящее время эти линии используются для передачи телеграфных, телефонных и факсимильных сообщений. При передаче телеграфных сообщений используется в основном частотная манипуляция и применяется аппаратура с автоматическим запросом и исправлением ошибок. Это позволяет включать радиотелеграфные каналы в международную абонентскую телеграфную сеть ТЕЛЕКС.

Появление эффективных систем передачи информации (кабельных, радиорелейных, спутниковых) привело к уменьшению удельного веса коротковолновых систем, поскольку их качественные показатели и возможности ниже. Однако это не означает, что КВ радиолинии не перспективны. Ряд таких важных качеств КВ систем связи, как мобильность, простота организации связи в малонаселенных и труднодоступных районах, возможность дальней связи с подвижными объектами практически на любом расстоянии в пределах земного шара, а также применение КВ систем для частичного резервирования других систем связи, делает эти системы в ряде случаев незаменимыми.

Оптические (лазерные) системы связи. В оптических системах связи используются электромагнитные волны светового диапазона частот (10^{13} — 10^{15} Гц). Оптические системы со свободным распространением могут использоваться в разных средах (атмосфера, вода, космос). Достоинства оптических систем связи — возможность передачи огромных потоков информации и работа очень узким лучом, который обеспечивает высокий энергетический потенциал линии связи и хорошую скрытность передачи. Недостатками систем являются значительное затухание волн оптического диапазона в атмосфере и в воде; высокие требования к точности наведения и удержания луча на заданном направлении; сложность реализации связи с подвижными объектами.

Краткое рассмотрение различных видов радиотехнических систем связи показывает, что каждый из них имеет ряд определенных достоинств и недостатков. Умелое сочетание этих качеств позволяет решать самые разнообразные задачи передачи информации.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Объясните назначение основных частей системы связи и дайте расшифровку терминов: «система связи», «сообщение», «сигнал», «помеха», «среда распространения», «линия связи», «канал связи».
2. Что такое «способ формирования сигнала» и «способ приема»? Назовите основные операции, осуществляемые при передаче сообщений.
3. Назовите причины появления ошибок при передаче информации по каналам связи.
4. Что такое операция уплотнения канала и в чем ее сущность?
5. Что такое эффективность системы связи и как ее оценивать?
6. В чем различие между системами связи со свободным и направленным распространением сигналов? Приведите примеры подобных систем.
7. Какие каналы связи называют каналами со случайными параметрами? Приведите примеры таких каналов.
8. Почему в современных радиотехнических системах передачи информации наиболее широко используются дециметровый и сантиметровый диапазоны?

9. Назовите основные особенности радиорелейных, спутниковых и космических систем связи.
10. Перечислите основные особенности тропосферных, ионосферных, метеорных и коротковолновых систем связи.
11. Сравните системы с направляемыми и со свободно распространяющимися сигналами по гибкости организации связи, по взаимному влиянию работающих систем друг на друга, стабильности параметров и экономичности.

ГЛАВА 2. СООБЩЕНИЯ, СИГНАЛЫ И ПОМЕХИ

2.1. ВВЕДЕНИЕ

Основное внимание в этой главе уделено аналитическому описанию сообщений, сигналов и помех. Важность такого описания состоит в том, что оно служит основой для установления количественных соотношений между параметрами системы связи и показателями качества ее работы в конкретных условиях. Изучая свойства сообщений, сигналов и помех, можно установить, при каких условиях обеспечивается наиболее эффективное использование данной системы связи.

Помимо аналитического описания большое значение имеет геометрическая трактовка сообщений, сигналов и помех, позволяющая наглядно представить весьма сложный механизм взаимодействия сигналов и помех при приеме и перекинуть «мостик» между абстрактностью аналитического описания и физической сутью рассматриваемых процессов и явлений.

2.2. ОСНОВНЫЕ ВИДЫ СООБЩЕНИЙ И ИХ КРАТКАЯ ХАРАКТЕРИСТИКА

Источник может создавать множество свойственных ему сообщений. Поэтому каждое передаваемое сообщение представляет собой одно из множества возможных. Получателю неизвестно заранее, какое конкретное сообщение из этого множества передано.

Статистический подход к проблемам передачи сообщений — характерная особенность современной теории связи. В соответствии с этим работа источника информации характеризуется случайной функцией времени, определяемой множеством (совокупностью) детерминированных функций времени (реализаций) и распределением плотности вероятностей этого множества. Каждое конкретное сообщение при этом рассматривается как *реализация* такого случайного процесса. Если рассматриваемые процессы эргодические (или близки к этому), то изучение характеристик и свойств отдельной реализации позволяет получить достаточно верное представление о всем процессе. Возможность замены рассмотрения случайного процесса рассмотрением отдельной его реализации (т. е. детерминированной функции) очень привлекательна, так как теория детерминированных колебаний хорошо

разработана, проста и удобна для решения многих практических задач. Несмотря на то, что для реальных сообщений свойство эргодичности строго не выполняется, указанной идеализацией часто пользуются.

Сообщения, создаваемые различными источниками, можно разделить на непрерывные и дискретные.

Непрерывными называются сообщения, которые в определенных пределах могут принимать любые мгновенные значения, т. е.

$$x_{\min} \leq x(t) \leq x_{\max}. \quad (2.1)$$

Интервал $A_x = x_{\max} - x_{\min}$ называют размером или размахом шкалы мгновенных значений непрерывного сообщения. Типичными примерами непрерывных сообщений являются речевые, телевизионные и некоторые телеметрические сообщения. Частный случай непрерывных сообщений — параметры, для которых характерно, что на некотором интервале времени T они не изменяются и принимают постоянные, заранее неизвестные значения, или могут принять за это время несколько таких значений. Чаще всего сообщения в виде параметров встречаются в телеметрии и телеуправлении.

Дискретными называются сообщения, образованные из отдельных элементов (символов, букв, цифр, импульсов), каждый из которых может принимать лишь конечное число M различных значений. Это число определяет *размер алфавита*, используемого для получения сообщений. Простейший алфавит — двоичный, он состоит из двух элементов (например, 0 и 1, или импульса и паузы и т. п.).

Различные дискретные сообщения, образованные одним и тем же источником, представляют собой последовательности элементов, отличающиеся друг от друга порядком следования элементов и их числом. Примерами дискретных сообщений являются телеграфные сообщения, команды в автоматизированных системах управления, числовые данные и т. п.

В настоящее время многие виды непрерывных сообщений также стремятся передавать в дискретной форме, поскольку это позволяет существенно повысить качество передачи и эффективность использования линий связи. Наиболее распространена *двоичная цифровая форма представления* дискретных сообщений, при которой каждый элемент алфавита дискретного сообщения заменяется определенной комбинацией двоичных элементов.

Достоинством цифровых сообщений является сравнительная простота их преобразований и обработки, а также возможность выполнения этих операций с высокой точностью.

Используя запоминающие устройства и различные скорости считывания, цифровые сообщения можно «сжимать» или «растягивать» во времени. Это позволяет передавать одни и те же сообщения с разной скоростью, что открывает широкие возможности для согласования различных каналов связи при их объединении в сети.

2.3. ДИСКРЕТИЗАЦИЯ НЕПРЕРЫВНЫХ СООБЩЕНИЙ ПО ВРЕМЕНИ

Непрерывные сообщения часто приходится преобразовывать в форму, более удобную для передачи по тем или иным каналам связи. Наиболее распространенным преобразованием непрерывных сообщений является их *дискретизация по времени*. Это преобразование применяется при импульсных методах передачи сообщений, многоканальной передаче с временными методами уплотнения и разделения, а также как промежуточная операция при преобразовании непрерывных сообщений в цифровую форму.

Дискретизация непрерывного сообщения $x(t)$ по времени состоит в замене его последовательностью мгновенных значений (отсчетов) $x(t_k)$, взятых в дискретные моменты времени $t_k = k\Delta t$, где $k=0, 1, 2, \dots$, (рис. 2.1). При такой замене из рассмотрения исключается все множество значений $x(t)$, находящихся внутри интервалов Δt . Если $\Delta t = \text{const}$, то дискретизация называется *равномерной* с периодом $T_n = \Delta t$. В общем случае моменты отсчетов t_k могут выбираться неравномерно, например, «сгущаться» и «разряжаться» пропорционально скорости изменения сообщения (*адаптивная дискретизация*). Равномерная дискретизация получила распространение как более простая. Технически дискретизация по времени реализуется с помощью ключевых устройств, управляемых периодической последовательностью коротких ($\tau \ll \Delta t$) прямоугольных импульсов (рис. 2.2). Такую операцию часто называют *импульсным преобразованием непрерывных сообщений*.

При дискретизации по времени возникают задачи, связанные с выбором интервала Δt и восстановлением исходного сообщения на выходе по его отсчетам. Сложность этих задач состоит в том,

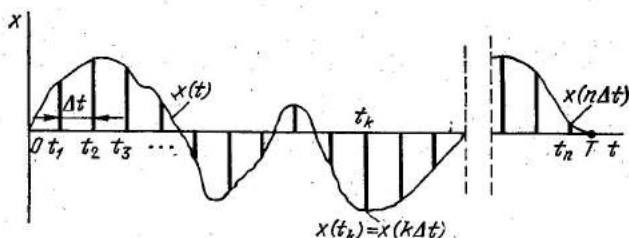


Рис. 2.1. Дискретизация непрерывного сообщения по времени

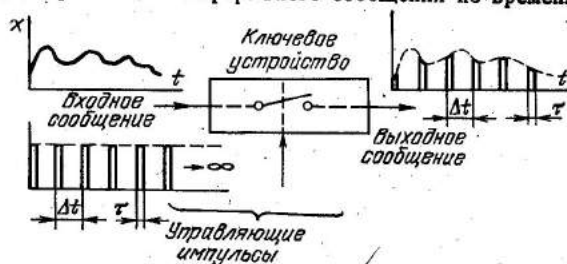


Рис. 2.2. Принцип технической реализации дискретизации по времени

что при их решении необходимо увязать свойства исходных непрерывных сообщений, способы их восстановления по отсчетам и требуемую точность восстановления.

В теории дискретизации особую важность имеет вопрос о максимальном (предельном) интервале $\Delta t = \Delta t_{\max}$, при котором еще имеется принципиальная возможность восстановить непрерывную функцию времени с заданной точностью по ее отсчетам.

Теория предельной дискретизации по времени применяется при анализе ряда важных вопросов передачи непрерывных сообщений, в частности исследовании вопросов устранения избыточности сообщений («сжатие информации»), определении пропускной способности каналов, помехоустойчивости приема и т. д. При построении этой теории необходимо использовать математическую модель непрерывных сообщений. В настоящее время наиболее широкое распространение получила модель сообщений с ограниченным спектром. Допустимость такой идеализации вытекает из того, что при передаче непрерывных сообщений по каналам связи всегда имеются неизбежные помехи и искажения. Так, как спектральная плотность $G_x(\Omega)$ любого реального сообщения уменьшается с ростом частоты, то, начиная с некоторой частоты $\Omega > \Omega_B = 2\pi F_B$, ее величина оказывается меньше спектральной плотности помех $G_B(\Omega)$ на выходе канала и, следовательно, вклад таких составляющих спектра сообщения в полезную информацию оказывается незначительным по сравнению с основной частью спектра (рис. 2.3). Ограничение спектра сообщения частотой Ω_B приводит к ошибке, относительная величина среднего квадрата которой равна

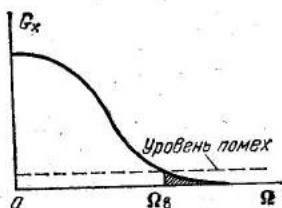


Рис. 2.3. Ограничение спектра сообщения

$$\delta_P^2 = \frac{\int_{\Omega_B}^{\infty} G_x(\Omega) d\Omega}{\int_0^{\infty} G_x(\Omega) d\Omega} = \frac{\Delta P_x}{P_x} \quad (2.2)$$

Эту величину можно трактовать как отношение средней мощности отброшенной части спектра («хвоста») ΔP_x к средней мощности всего сообщения P_x . Для непрерывных сообщений с ограниченным спектром справедлива теорема Котельникова:

Непрерывная функция времени $x(t)$ со спектром, ограниченным полосой частот от нуля до F_B , полностью определяется последовательностью своих мгновенных значений, взятых в моменты времени, отсчитываемые через интервалы $\Delta t = 1/2F_B$.

¹ Эта теорема рассматривалась рядом авторов. Однако применительно к проблемам передачи сообщений наиболее четко она была сформулирована в 1933 г. В. А. Котельниковым, поэтому в отечественной литературе ее вполне справедливо называют теоремой Котельникова. В иностранной литературе подобную теорему обычно называют теоремой Найквиста, или теоремой отсчетов.

Результатом доказательства теоремы является выражение, называемое рядом Котельникова:

$$x(t) = \sum_{k=-\infty}^{\infty} x(k \Delta t) \varphi_k(t), \quad (2.3)$$

где $\Delta t = \pi/\Omega_B = 1/2F_B$ — интервал дискретизации; $k=0 \pm 1; \pm 2, \dots$;

$$\varphi_k(t) = \frac{\sin \Omega_B(t - k \Delta t)}{\Omega_B(t - k \Delta t)}; \quad (2.4)$$

$x(k \Delta t)$ — значения функции $x(t)$ в точках отсчетов $t = k \Delta t$.

Функцию $\varphi_k(t)$ называют *функцией отсчетов*, так как для данного k она равна 1 только в той точке отсчета, где $t = k \Delta t$, а во всех других точках отсчетов $t = l \Delta t (l \neq k)$ она обращается в 0, т. е.

$$\varphi_k(t) = \begin{cases} 1 & \text{при } t = k \Delta t, \\ 0 & \text{при } t = l \Delta t (l \neq k). \end{cases}$$

Следовательно, значения функции в точках отсчета $x(k \Delta t)$ независимы друг от друга.

Ряд (2.3) представляет собой разложение $x(t)$ по системе ортогональных функций $\varphi_k(t)$. Нетрудно видеть, что это выражение дает решение как прямой (выбор интервала Δt), так и обратной задачи (восстановление непрерывной функции $x(t)$ по ее отсчетам $x(k \Delta t)$).

Теорема Котельникова опирается на модель непрерывной функции с ограниченным спектром и бесконечной протяженностью во времени. Все реальные сообщения имеют конечную протяженность во времени (финитны), поэтому модель с ограниченным спектром для них, строго говоря, некорректна.

Если в соответствии с теоремой Котельникова выполнить дискретизацию функции $x(t)$ конечной длительности T , то число отсчетов будет конечным и равно

$$N = T/\Delta t = 2 F_B T. \quad (2.5)$$

Величина N представляет собой число независимых отсчетов, которыми можно заменить непрерывную функцию времени с ограниченным спектром и конечной длительностью T . Величина N называется числом *степеней свободы* такой функции. В рассматриваемом случае ряд Котельникова содержит конечное число членов, т. е. дискретизация сопровождается ошибкой, средний квадрат которой можно записать в виде

$$\delta_{\Delta}^2 = \frac{\int_0^T \left[x(t) - \sum_{k=0}^{N-1} x(k \Delta t) \varphi_k(t) \right]^2 dt}{\int_0^T x^2(t) dt}. \quad (2.6)$$

Очевидно, что при увеличении длительности сообщения T (числа отсчетов N) ошибка уменьшается. Связь между N и δ_d исследована в работах [1, 2]. Было установлено, что среднеквадратическое значение ошибки дискретизации δ_d заключено в пределах

$$\delta_F < \delta_d < \sqrt{3} \delta_F, \quad (2.7)$$

где δ_F определяется выражением (2.2). При больших значениях N ($N > 10^3$) величина δ_d приближается к нижнему пределу¹.

Теорема Котельникова получила широкое применение в различных теоретических исследованиях как предельное соотношение. Для точного восстановления сообщения, дискретизированного по времени в соответствии с этой теоремой, нужно суммировать бесконечное число членов ряда (2.3). Такая процедура физически нереализуема.

В настоящее время в технических приложениях выбор интервала дискретизации Δt непрерывных сообщений определяется в значительной мере требованиями удобства и простоты восстановления сообщений по их отсчетам, однако при этом приходится увеличивать число отсчетов по сравнению с предельной дискретизацией. Действительно, при дискретизации по теореме Котельникова имеем

$$\Delta t = \Delta t_{\max} = 1/2 F_B; \quad N = N_{\min} = 2 F_B T. \quad (2.8)$$

При реальной (фактически используемой) дискретизации

$$\Delta t = \Delta t_{\text{факт}} = 1/\mu F_B; \quad N = N_{\text{факт}} = \mu F_B T, \quad (2.9)$$

где μ — коэффициент, удовлетворяющий условию $\mu > 2$. Из (2.8) и (2.9) следует, что при одинаковой точности восстановления

$$k_{\text{изб}} = (\Delta t_{\max} / \Delta t_{\text{факт}}) = (N_{\text{факт}} / N_{\min}) = \mu / 2 > 1. \quad (2.10)$$

Величину $k_{\text{изб}}$ называют *коэффициентом избыточности отсчетов*.

С инженерной точки зрения проблема дискретизации непрерывных колебаний по времени не имеет однозначного решения. При ее реализации приходится принимать компромиссное решение между простотой способов восстановления и приемлемой избыточностью отсчетов. Практика показывает, что во многих важных прикладных задачах достаточно выбрать значение $\mu \approx (2,5-5)$.

Среди различных способов на практике широко применяется восстановление (интерполяция) непрерывных сообщений по их отсчетам с помощью аналоговых фильтров нижних частот. Некоторые вопросы такого восстановления, а также выбора частоты дискретизации в зависимости от характеристик сообщения и параметров восстанавливающего фильтра будут рассмотрены в гл. 5 (см. § 5.6).

¹ При этих условиях более точное соотношение между значениями δ_d и δ_F может быть установлено на основе результатов работ Ландау и Поллака [1, 2].

2.4. КВАНТОВАНИЕ НЕПРЕРЫВНЫХ СООБЩЕНИЙ ПО УРОВНЮ

Суть *квантования* (*дискретизации*) по уровню состоит в замене континуума значений, которые может принимать непрерывное сообщение, дискретным множеством заранее установленных значений. Такая замена основывается на том, что передача информации по каналам связи всегда сопровождается действием помех и искажений. Это приводит к тому, что близкие друг к другу «похожие» непрерывные сообщения трудно различать при приеме. Появляется как бы некоторая зона неразличимости (неопределенности), в пределах которой **нельзя установить истинное значение сообщения.**

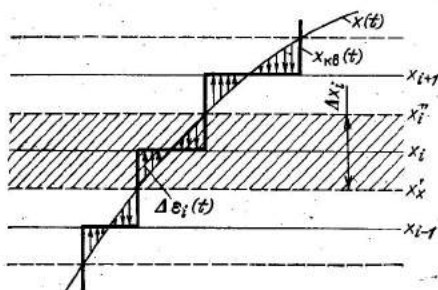


Рис. 2.4. Принцип квантования непрерывного сообщения по уровню

Если учесть заранее некоторые причины искажений (неидеальность характеристик общего тракта системы, нестабильность работы и т. п.), то можно предварительно, еще до передачи, преобразовать сообщения так, чтобы нежелательные факторы уже никакого влияния на них не оказывали. Характерной особенностью такого преобразования

является преднамеренное введение в передаваемые сообщения некоторой заранее запланированной ошибки (ошибки квантования).

Рассмотрим основные особенности квантования по уровню. Разобьем непрерывную шкалу мгновенных значений передаваемых сообщений с размером A_x на конечное число частей (квантов) $N_{кв}$. Полученная при этом дискретная шкала называется шкалой уровней квантования, а ее деления — *уровнями квантования*. Интервал между соседними уровнями квантования называется *шагом квантования* Δx . Величина Δx определяется допустимой зоной неразличимости. Квантование может быть *равномерным* (при шаге Δx постоянном по всей шкале) и *неравномерным* (при изменении шага от уровня к уровню по некоторому правилу, учитывающему статистику квантуемых сообщений). Равномерное квантование применяется чаще, так как его проще реализовать.

Для иллюстрации замены непрерывного сообщения квантованным по уровню рассмотрим рис. 2.4. На нем показана только часть шкалы квантования (три уровня) и введены следующие обозначения: x_{i-1} , x_i , x_{i+1} — уровни квантования; Δx_i — шаг квантования; $\Delta \varepsilon_i(t)$ — текущая ошибка квантования на i -м шаге.

Квантование осуществляют по правилу: мгновенные значения сообщения, заключенные между соседними уровнями, всегда относят к ближайшему из них. При таком правиле каждый уровень квантования должен находиться в середине зоны неразличимости, равной шагу квантования. На рис. 2.4 для наглядности зо-

на, соответствующая уровню x_i , заштрихована, а стрелками условно показано, как нужно относить данные мгновенные значения сообщения $x(t)$ к этому уровню.

В результате квантования непрерывное сообщение $x(t)$ заменяется дискретным сообщением $x_{кв}(t)$, которое имеет ступенчатую форму и может принимать только конечное число различных мгновенных значений, равное числу уровней квантования $N_{кв}$.

Текущая ошибка квантования $\varepsilon_{кв}(t)$ представляет собой разницу между $x(t)$ и $x_{кв}(t)$:

$$\varepsilon_{кв}(t) = x_{кв}(t) - x(t), \quad (2.11)$$

которую называют *шумом квантования*.

В пределах одного шага для i -го уровня мгновенное значение текущей ошибки лежит в интервале

$$-\Delta x_i/2 \leq \Delta \varepsilon_i(t) \leq \Delta x_i/2. \quad (2.12)$$

Исследования показывают, что при равномерном квантовании ($\Delta x_i = \Delta x = \text{const}$) дисперсия (средний квадрат) шума квантования по всем уровням равна [3]:

$$\overline{\varepsilon_{кв}^2} = \frac{1}{N_{кв}} \sum_{i=1}^{N_{кв}} \Delta \varepsilon_i^2 = \Delta x^2/12. \quad (2.13)$$

Эта величина определяет среднюю мощность шума квантования $P_{кв}$. Так как $\Delta x = A_x/N_{кв}$, то

$$\overline{\varepsilon_{кв}^2} = \Delta x^2/12 = A_x^2/12 N_{кв}^2. \quad (2.14)$$

Для оценки ошибки квантования удобно пользоваться относительной величиной

$$\delta_{кв}^2 = \overline{\varepsilon_{кв}^2} / \overline{x^2} = P_{кв}/P_x, \quad (2.15)$$

где $\overline{x^2}$ — дисперсия (средняя мощность) квантуемого сообщения.

Приняв во внимание (2.14) и (2.15), имеем:

$$\delta_{кв} = \frac{1}{2\sqrt{3}N_{кв}} (A_x^2/\overline{x^2})^{1/2} = \frac{1}{2\sqrt{3}N_{кв}} \frac{A_x}{x_{эфф}}, \quad (2.16)$$

где $x_{эфф}$ — эффективное значение сообщения $x(t)$.

Выражение (2.16) можно записать в виде

$$\delta_{кв} = k_{пх}/\sqrt{3}N_{кв}, \quad (2.17)$$

где $k_{пх} = A_x/2x_{эфф}$ — *пик-фактор сообщения*, зависящий от статистики сообщения. Практика показывает, что для различных классов непрерывных сообщений $k_{пх} \approx 1,5-3,5$ и, следовательно,

$$\delta_{кв} \approx (1-2)/N_{кв}.$$

В инженерных приложениях среднеквадратическое значение ошибки квантования обычно оценивают величиной

$$\delta_{кв} = 1/N_{кв}, \quad (2.18)$$

которая соответствует пик-фактору $k_{пх} = \sqrt{3}$ (т. е. равновероятному распределению мгновенных значений сообщения).

Из (2.17) следует, что при равномерном квантовании сообщений ошибка квантования пропорциональна величине пик-фактора. Для ее уменьшения квантуемые сообщения можно предварительно подвергнуть нелинейному преобразованию. Примером такого преобразования является сжатие динамического диапазона передаваемых речевых сообщений (компрессия речи) с последующим расширением динамического диапазона принятых сообщений устройствами, амплитудные характеристики которых согласованы с амплитудными характеристиками сжимающих устройств. Заметим, что квантование непрерывных сообщений по уровню как отдельная самостоятельная операция практически не применяется; эта операция обычно является частью процесса преобразования непрерывных сообщений в цифровую форму и выполняется с помощью аналого-цифровых преобразователей.

2.5. ПРЕОБРАЗОВАНИЕ НЕПРЕРЫВНЫХ СООБЩЕНИЙ В ЦИФРОВУЮ ФОРМУ

В результате дискретизации по времени и квантования по уровню непрерывное сообщение заменяется последовательностью отсчетов, которые могут принимать только конечное число значений, равное числу уровней квантования $N_{кв}$. Каждое из этих значений (число) можно выразить в одной из систем счисления и передать по линии связи в виде кодовых комбинаций.

Как известно, запись числа N в позиционной системе счисления имеет вид

$$N = \sum_{i=m}^1 \gamma_i M^{i-1} = \gamma_m M^{m-1} + \dots + \gamma_1 M^0,$$

где M — основание системы счисления ($M \geq 2$); m — число разрядов; γ_i — весовой коэффициент разряда, принимающий одно из целых значений в интервале ($0 \leq \gamma_i \leq M-1$). Передача конкретного числа по линии связи сводится к передаче его весовых коэффициентов γ_i . Наиболее просто эта операция реализуется для двоичной системы счисления, когда γ_i принимают только два значения (0 и 1). В этом случае кодовые комбинации состоят из двоичных элементов (например, импульсов и пауз).

Преобразование непрерывных сообщений в цифровую форму связано с появлением ошибок за счет дискретизации по времени и квантования по уровню. В соответствии со свойством аддитивности критерия СКО средний квадрат ошибки цифрового преобразования δ_c^2 равен

$$\delta_c^2 = \delta_d^2 + \delta_{кв}^2, \quad (2.19)$$

где значения δ_d и $\delta_{кв}$ определяются выражениями (2.7) и (2.18).

При передаче непрерывных сообщений, преобразованных в

цифровую форму, наличие помех в канале связи приводит к тому, что некоторые элементы переданных кодовых комбинаций могут быть искажены и приняты неверно. Например, вместо элемента 0 принят элемент 1 или вместо 1 принят элемент 0. В результате, кроме указанных выше ошибок появляется дополнительная ошибка, средний квадрат которой равен [3]:

$$\delta_{\text{ш}}^2 = k P_{\text{ош}}, \quad (2.20)$$

где $P_{\text{ош}}$ — вероятность ошибки при приеме отдельного элемента цифровой последовательности; k — коэффициент, величина которого зависит от характеристик сообщений и помех ($k=1-4$), обычно принимают $k=4$.

При точностях передачи, представляющих практический интерес, ошибку $\delta_{\text{ш}}$ можно считать независимой от ошибок преобразования $\delta_{\text{д}}$ и $\delta_{\text{кв}}$. Поэтому общий средний квадрат ошибки передачи равен

$$\delta_{\text{общ}}^2 = \delta_{\text{д}}^2 + \delta_{\text{ш}}^2 = \delta_{\text{д}}^2 + \delta_{\text{кв}}^2 + \delta_{\text{ш}}^2. \quad (2.21)$$

Обычно параметры системы выбирают так, чтобы общая заданная ошибка передачи определялась в основном ошибками преобразования. В этом случае стремятся обеспечить такие условия работы системы, при которых выполняется условие

$$\delta_{\text{ш}}^2 \leq 0,1 \delta_{\text{д}}^2 \text{ или } \delta_{\text{общ}}^2 \leq 1,1 \delta_{\text{д}}^2. \quad (2.22)$$

Приняв во внимание (2.20) и эти условия, нетрудно установить допустимую вероятность ошибки в канале связи при общей заданной величине среднеквадратической ошибки передачи аналоговых сообщений цифровыми методами:

$$P_{\text{ош}} \leq 0,09 \delta_{\text{общ}}^2 / k \approx 2,5 \cdot 10^{-2} \delta_{\text{общ}}^2. \quad (2.23)$$

В зависимости от вида решаемых задач, типа канала связи и других факторов величина общей ошибки $\delta_{\text{общ}}$ может выбираться в диапазоне $10^{-1}-10^{-3}$. В этом случае допустимая вероятность ошибки при передаче непрерывных сообщений цифровыми методами лежит в интервале $2,5 \cdot (10^{-4}-10^{-8})$.

Зная допустимую величину $P_{\text{ош}}$ при выбранном способе передачи можно определить необходимое отношение сигнала к шуму на входе приемного устройства системы. Эти вопросы рассматриваются в гл. 7.

2.6. ОБЩЕЕ ОПИСАНИЕ МОДУЛИРОВАННЫХ СИГНАЛОВ

В качестве переносчика информации (несущей) в системах связи обычно используют гармонические колебания

$$s_0(t) = S_0 \cos(\omega_0 t + \varphi_0) = S_0 \cos \Phi_0(t). \quad (2.24)$$

Здесь S_0 — амплитуда, ω_0 — частота, φ_0 — начальная фаза, $\Phi_0(t)$ — полная фаза колебания. Если один из параметров несущей изменяется в соответствии с передаваемым сообщением, то переносчик

называется *сигналом*. Для дискретных сообщений указанные изменения параметра называются *манипуляцией*, а для непрерывных — *модуляцией*. В первом случае параметр несущей изменяется дискретно и может принимать только конечное число значений, а во втором — параметр изменяется непрерывно.

В общем случае модулированный сигнал конечной длительности T можно записать в виде

$$s(t) = S(t) \cos [\omega_0 t + \varphi(t)] = S(t) \cos \Phi(t), \quad (2.25)$$

где $S(t)$ и $\Phi(t)$ — огибающая и полная фаза сигнала; $\varphi(t)$ — модуляционная составляющая полной фазы сигнала.

При модуляции $S(t)$ или $\Phi(t)$ (а в общем случае оба эти параметра) становятся функциями передаваемого сообщения $x(t)$.

Выражение (2.25) иногда удобнее представить в форме

$$\begin{aligned} s(t) &= S(t) \cos [\varphi(t)] \cos \omega_0 t - S(t) \sin [\varphi(t)] \sin \omega_0 t = \\ &= S_c(t) \cos \omega_0 t - S_s(t) \sin \omega_0 t = s_c(t) - s_s(t), \end{aligned} \quad (2.26)$$

где $s_c(t)$ и $s_s(t)$ — квадратурные составляющие сигнала, а $S_c(t)$ и $S_s(t)$ — их огибающие.

Нетрудно видеть, что

$$S(t) = [S_c^2(t) + S_s^2(t)]^{1/2}; \quad \varphi(t) = \text{arctg} [S_s(t)/S_c(t)]. \quad (2.27)$$

Огибающие квадратурных составляющих являются медленно меняющимися функциями времени, следовательно, огибающая $S(t)$ и фаза сигнала $\varphi(t)$ также медленно меняющиеся функции. Это означает, что за время, равное периоду несущей $T_0 = 2\pi/\omega_0$, функции $S_c(t)$ и $S_s(t)$ (или $S(t)$ и $\varphi(t)$) изменяются незначительно и спектр сигнала занимает полосу частот $\Delta\omega_c = 2\pi\Delta f_c$, сосредоточенную вблизи несущей f_0 , т. е. $\Delta f_c/f_0 \ll 1$. Сигналы, удовлетворяющие такому условию, называются *узкополосными* или *полосовыми*. К ним практически относятся сигналы всех систем связи со свободным распространением.

Важной характеристикой модулированных сигналов является отношение ширины их спектра к ширине спектра модулирующего колебания:

$$\gamma_F = \Delta f_c / F_B \gg 1. \quad (2.28)$$

В соответствии с величиной γ_F различают *узкополосные* ($\gamma \approx 1 \div 2$) и *широкополосные* ($\gamma \gg 1$) виды модуляции:

Средняя удельная мощность узкополосного сигнала конечной длительности T равна (по определению)

$$P_c = \frac{1}{T} \int_0^T s^2(t) dt = \frac{1}{T} \int_0^T S^2(t) \cos^2 [\omega_0 t + \varphi(t)] dt. \quad (2.29)$$

Нетрудно видеть, что с большой точностью эту мощность можно принять равной величине

$$P_c = \frac{1}{2T} \int_0^T S^2(t) dt = \frac{P_{\text{огиб}}}{2}, \quad (2.30)$$

где $P_{\text{огиб}}$ — мощность огибающей модулированного сигнала.

Основываясь на этих соотношениях, можно найти мощность конкретного модулированного сигнала.

При решении задач передачи информации узкополосные сигналы часто удобно представлять в комплексной форме. В этом случае (2.25) принимает вид

$$s(t) = \operatorname{Re} [S(t) e^{j\varphi(t)} e^{j\omega_0 t}] = \operatorname{Re} [S(t) e^{j\omega_0 t}], \quad (2.31)$$

где $S(t)$ — комплексная огибающая сигнала; Re — реальная часть выражения в квадратных скобках. В (2.31) функцию $e^{j\omega_0 t}$ можно рассматривать как комплексную несущую (комплексную сигнальную функцию), а $S(t)$ — как комплексный модулируемый параметр (комплексную модулирующую функцию).

Развитием комплексного представления сигналов является аналитический сигнал

$$\tilde{s}(t) = s(t) + j\bar{s}(t),$$

представляющий собой комплексную функцию, действительная часть которой равна заданному сигналу, а мнимая — его преобразованию по Гильберту [4].

Функция $s(t)$ и сопряженная ей по Гильберту функция $\bar{s}(t)$ связаны парой преобразований:

$$\bar{s}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau} d\tau; \quad s(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\bar{s}(\tau)}{\tau-t} d\tau.$$

Нетрудно показать, что $\bar{s}(t)$ можно выразить через квадратурные составляющие сигнала:

$$\bar{s}(t) = S_c(t) \cos \omega_0 t + S_s(t) \sin \omega_0 t.$$

Представление сигналов в комплексной форме позволяет широко пользоваться при различных преобразованиях сигналов методом низкочастотных эквивалентов, разложениями огибающих сигнала по ортогональным системам низкочастотных базисных функций и геометрическими представлениями комплексных огибающих как векторов.

2.7. ПОМЕХИ В КАНАЛАХ СВЯЗИ

При передаче информации существенное влияние на показатели работы системы связи могут оказывать различные внешние помехи, попадающие на вход приемника вместе с полезным сигналом. Источниками таких помех могут быть:

другие системы, работающие в том же (или близком) диапазоне частот;

разнообразные физико-химические процессы, происходящие в среде распространения и вызывающие появление так называемых шумов среды;

различные промышленные установки, создающие возмущающие излучения.

Большинство внешних помех относится к классу аддитивных; они действуют на входе приемника независимо от того, есть ли там сигнал или он отсутствует. К числу аддитивных помех относятся и внутренние шумы приемника, которые обычно пересчитывают к его входу. По своей природе аддитивные помехи случайны и их можно описать только статистически.

Для удобства рассмотрения все аддитивные помехи обычно разбивают на три основные группы: *флуктуационные* (шумовые), *сосредоточенные* (квазисинусоидальные) и *импульсные*.

Как уже указывалось в гл. 1, значительная часть современных радиотехнических систем передачи информации работает в диапазонах дециметровых и сантиметровых волн, а характерной особенностью их развития является тенденция перехода к еще более коротким волнам. В связи с этим основным видом аддитивных помех в системах связи становятся флуктуационные помехи. Такие помехи имеют вид непрерывных случайных колебаний, а их спектр на входе приемника обычно значительно шире его полосы пропускания. К флуктуационным помехам относятся внутренние тепловые шумы приемника, шумы среды распространения, космические шумы. Большое внимание, которое уделяется в теории связи флуктуационным помехам и особенно внутренним тепловым шумам приемника, объясняется тем, что их влияние проявляется во всех диапазонах волн, включая оптические. Поэтому от них нельзя избавиться выбором диапазона или переходом от одного диапазона к другому.

Внутренние шумы, приведенные ко входу приемного устройства, представляют собой стационарный гауссовский процесс с нулевым средним и односторонней спектральной плотностью (энергетическим спектром)

$$N_0 = kT^\circ F_{\text{ш}} = kT_{\text{экв}}^\circ [\text{Вт/Гц}], \quad (2.32)$$

где $k = 1,38 \cdot 10^{-23}$ (Дж/К) — постоянная Больцмана; T° — температура окружающей среды по шкале Кельвина ($T^\circ = 273 + t^\circ \text{C}$); $F_{\text{ш}}$ — коэффициент шума приемника; $T_{\text{экв}}^\circ = T^\circ F_{\text{ш}}$ — эквивалентная шумовая температура входа.

Величина N_0 практически не зависит от частоты (до частот полосы пропускания приемника Δf_s , которая в свою очередь позволяет пользоваться при рассмотрении различных вопросов простой и удобной моделью *белого шума*). В реальных условиях белым считают шум с шириной спектра $\Delta f_{\text{ш}}$, значительно превышающей полосу пропускания приемника Δf_s , которая в свою очередь определяется полосой частот, занимаемой спектром сигнала Δf_c ($\Delta f_{\text{ш}} \gg \Delta f_s = \Delta f_c$). Практически при $\Delta f_{\text{ш}} \approx (5-10)\Delta f_s$ шум можно считать белым. Шум с равномерной спектральной плотностью в полосе Δf_s часто удобно представлять в виде квадратурных составляющих:

$$n(t) = N_c(t) \cos \omega_0 t + N_s(t) \sin \omega_0 t, \quad (2.33)$$

где $N_c(t)$ и $N_s(t)$ — низкочастотные квадратурные составляющие случайные нормальных процессов со спектральной плотностью $2N_0$ и шириной спектра от нуля до $\Delta f_0/2$.

Нетрудно убедиться, что средняя мощность белого шума в полосе Δf_0 равна

$$P_{\text{ш}} = N_0 \Delta f_0, \quad (2.34)$$

где эквивалентная шумовая полоса приемника Δf_0 определяется выражением

$$\Delta f_0 = \frac{1}{2\pi} \int_{-\infty}^{\infty} [K^2(\omega)/K_0^2] d\omega. \quad (2.35)$$

В (2.35) $K(\omega)$ — частотная характеристика избирательной части приемника, а K_0 — ее значение на центральной частоте.

Выражение (2.33) можно представить в виде

$$n(t) = N(t) \cos[\omega_0 t + \theta_{\text{ш}}(t)], \quad (2.36)$$

где $N(t)$ и $\theta_{\text{ш}}(t)$ — огибающая и фаза случайного шумового процесса, определяемые как

$$N(t) = [N_c^2(t) + N_s^2(t)]^{1/2}; \theta_{\text{ш}}(t) = \text{arctg}[N_s(t)/N_c(t)]. \quad (2.37)$$

Так как $N_c(t)$ и $N_s(t)$ — медленно меняющиеся колебания по сравнению с $\cos \omega_0 t$, то $N(t)$ и $\theta_{\text{ш}}(t)$ также медленно меняющиеся процессы.

Запись в форме (2.36) позволяет рассматривать шум как узкополосный случайный процесс, имеющий вид гармонического колебания со случайными медленно изменяющимися во времени амплитудой и фазой.

Выражения (2.33) или (2.36) можно записать в комплексном виде:

$$n(t) = \text{Re}[\dot{N}(t) e^{j\omega_0 t}], \quad (2.38)$$

где $\dot{N}(t)$ — комплексная огибающая шума;

$$\dot{N}(t) = N_c(t) + j N_s(t). \quad (2.39)$$

Рассмотренные виды представления полосовых случайных процессов широко используются в статистической теории связи при исследовании помехоустойчивости различных способов передачи и приема сигналов.

Широкое развитие цифровых методов передачи информации, усложнение систем и объединение их в более крупные комплексы и сети привело к необходимости осуществлять в процессе работы разнообразную коммутацию цепей, источников сообщений, каналов связи и т. п. Подобная коммутация связана с кратковременными прерываниями, что приводит к появлению переходных про-

цессов в различных точках устройств и систем. Возникающие при этом эффекты аналогичны действию импульсных помех на входе приемника. Иными словами, в рассматриваемом случае в качестве модели мешающих воздействий можно использовать импульсную помеху. При независимом возникновении импульсов импульсную помеху наиболее часто описывают распределением Пуассона

$$P(l) = (\nu T)^l e^{-\nu T} / l!, \quad (2.40)$$

где $P(l)$ — вероятность появления l импульсов за время T ; ν — среднее число импульсов в единицу времени (интенсивность потока импульсов).

Несмотря на кратковременность действия, импульсная помеха может существенно снизить точность передачи информации. Поэтому в современных системах связи (особенно в системах передачи данных, команд и т. п.), где требуется высокая достоверность передачи, мерам борьбы с такими помехами уделяется большое внимание.

2.8. ОРТОГОНАЛЬНЫЕ РАЗЛОЖЕНИЯ КОЛЕБАНИЙ

Сообщения, сигналы и помехи, с которыми приходится иметь дело в теории и технике связи, представляют собой колебания весьма сложной формы. Часто целесообразно представлять такие колебания $u(t)$ в виде линейной комбинации более простых колебаний. Это можно сделать путем подбора системы заранее известных функций $\varphi_1(t)$, $\varphi_2(t)$, ..., $\varphi_i(t)$, ... и некоторых числовых коэффициентов a_1 , a_2 , ..., a_i , ... таким образом, чтобы

$$u(t) = \sum_{i=1}^{\infty} a_i \varphi_i(t). \quad (2.41)$$

Представление (2.41) называется *разложением функции $u(t)$* по выбранной регулярной системе функций $\varphi_i(t)$. Эту систему называют *системой базисных функций*. Наиболее часто в разложениях используют системы функций, удовлетворяющих на интервале разложения $[0, T]$ условию ортогональности:

$$\frac{1}{T} \int_0^T \varphi_i(t) \varphi_j(t) dt = \begin{cases} Q & \text{при } i = j, \\ 0 & \text{при } i \neq j. \end{cases} \quad (2.42)$$

При $Q=1$ ортогональная система функций называется *ортонормированной* (ортонормальной). Нормировка ортогональной системы функций может осуществляться по мощности или по энергии. При этом выражение для ортонормированной функции имеет вид

$$\varphi_{i \text{ норм}}(t) = \varphi_i(t) / \|\varphi_i\|, \quad (2.43)$$

где $\|\varphi_i\|$ — норма функции $\varphi_i(t)$.

При нормировке по мощности квадрат нормы определяется соотношением

$$\|\varphi_i\|^2 = \frac{1}{T} \int_0^T \varphi_i^2(t) dt, \quad (2.44)$$

а при нормировке по энергии

$$\|\varphi_i\|^2 = \int_0^T \varphi_i^2(t) dt. \quad (2.45)$$

Выражения (2.44) и (2.45) определяют соответственно мощность и энергию функции $\varphi_i(t)$. Очевидно, что для нормированной функции независимо от вида нормировки должно выполняться условие $\|\varphi_i\|_{\text{норм}}^2 = 1$.

В теории связи и ее приложениях удобнее пользоваться нормировкой по мощности, так как она инвариантна к величине интервала ортогонального разложения T . Далее для определенности будем полагать, что система базисных функций $\varphi_i(t)$ ортонормальна.

При известной системе ортогональных функций колебание $u(t)$ полностью определяется совокупностью числовых коэффициентов a_i . Эта совокупность, представленная в порядке возрастания номера i , называется спектральной характеристикой (спектром) колебания $u(t)$ в базисе функций $\varphi_i(t)$. Если система функций $\varphi_i(t)$ ортонормальна, то коэффициенты a_i вычисляются по формуле

$$a_i = \int_0^T u(t) \varphi_i(t) dt. \quad (2.46)$$

Разложение (2.41), в котором коэффициенты a_i определяются формулой (2.46), называют *обобщенным рядом Фурье*. Если ограничить число членов ряда, то функция $u(t)$ будет представлена с некоторой погрешностью (ошибкой). В этом случае вместо истинного колебания $u(t)$ рассматривается его аппроксимация (приближение) $\hat{u}(t)$:

$$\hat{u}(t) = \sum_{i=1}^n a_i \varphi_i(t). \quad (2.47)$$

Ошибка такой аппроксимации равна

$$e(t) = u(t) - \hat{u}(t) = u(t) - \sum_{i=1}^n a_i \varphi_i(t). \quad (2.48)$$

С учетом (2.42) нетрудно показать, что относительная величина среднего квадрата ошибки аппроксимации равна

$$\delta_a^2 = \frac{\langle e^2(t) \rangle}{\langle u^2(t) \rangle} = \frac{P_e}{P_u} = 1 - \left(\sum_{i=1}^n a_i^2 \right) / P_u, \quad (2.49)$$

где P_e и P_u — средние удельные мощности, определяемые выражениями

$$P_e = \langle e^2(t) \rangle = \frac{1}{T} \int_0^T e^2(t) dt, \quad P_u = \langle u^2(t) \rangle = \frac{1}{T} \int_0^T u^2(t) dt. \quad (2.50)$$

В (2.49), (2.50) и далее ломаные скобки $\langle \cdot \rangle$ означают усреднение по времени.

Так как всегда $\delta^2_a \geq 0$, то из (2.49) следует

$$\sum_{i=1}^n a_i^2 \leq P_u. \quad (2.51)$$

Это условие, справедливое для любой ортонормированной базисной системы, называют неравенством Бесселя.

Из (2.46) видно, что коэффициенты a_i не зависят от числа членов ряда и, следовательно, добавление новых членов с целью улучшения точности аппроксимации не повлияет на вычисленные ранее значения a_i .

Если среднеквадратическую ошибку аппроксимации можно сделать сколь угодно малой за счет увеличения числа членов ряда, то базисную систему называют *полной*. При этом неравенство (2.51) переходит в равенство Парсеваля, показывающее, что мощность колебания равна сумме мощностей всех его спектральных составляющих [4].

2.9. НЕКОТОРЫЕ ВАЖНЫЕ ОРТОГОНАЛЬНЫЕ РАЗЛОЖЕНИЯ

В теории передачи информации и ее приложениях находят применение ортогональные разложения по различным системам базисных функций. Помимо свойства ортогональности такие системы должны удовлетворять требованиям полноты, удобства применения и простоты реализации. Одно из важнейших ортогональных разложений — ряд Котельникова, было рассмотрено в п. 2.3.

Рассмотрим кратко еще два важных разложения, широко используемых в теории передачи сообщений.

Разложение по системе тригонометрических функций (ряды Фурье). Такое разложение находит широкое применение по следующим причинам: гармонические колебания не искажаются при прохождении через линейные системы с постоянными коэффициентами; анализ с применением гармонических функций сравнительно прост и опирается на хорошо разработанный математический аппарат; гармонические функции легко генерируются.

Система образована из множества функций

$$\left\{ \cos i \frac{2\pi t}{T}, \sin i \frac{2\pi t}{T} \right\}, \quad (2.52)$$

ортogonalных на интервале $[0, T]$. В (2.52) i — номера гармоник ($i=1, 2, 3, \dots, \infty$). При нормировке (2.52) по мощности имеем

$$\{\varphi_i(t)\} = \{\sqrt{2} \cos i \Omega t, \sqrt{2} \sin i \Omega t\}, \quad (2.53)$$

где $\varphi_i(t)$ — ортонормированная функция; $\Omega = 2\pi/T$.

В этом случае разложение непрерывной функции времени $u(t)$ на интервале $[0, T]$ имеет вид

$$u(t) = a_0 + \sum_{i=1}^{\infty} a_i \varphi_i(t), \quad (2.54)$$

где

$$a_i \varphi_i(t) = \begin{cases} \sqrt{2} a_{2k-1} \sin k \Omega t & \text{для } i = 2k-1, \\ \sqrt{2} a_{2k} \cos k \Omega t & \text{для } i = 2k. \end{cases} \quad (2.55)$$

Коэффициенты a в этих выражениях представляют собой эффективные значения соответствующих гармоник. Они определяются по формуле (2.46):

$$\begin{aligned} a_0 &= \frac{1}{T} \int_0^T u(t) dt, \\ a_{2k-1} &= \frac{\sqrt{2}}{T} \int_0^T u(t) \sin k \Omega t dt, \\ a_{2k} &= \frac{\sqrt{2}}{T} \int_0^T u(t) \cos k \Omega t dt. \end{aligned} \quad (2.56)$$

В некоторых случаях ряд (2.54) удобно представить в виде

$$u(t) = a_0 + \sum_{i=1}^{\infty} A_i \cos(i \Omega t + \theta_i), \quad (2.57)$$

где A_i и θ_i — амплитуда и начальная фаза i -й гармоники

$$A_i = [2(a_{2k-1}^2 + a_{2k}^2)]^{1/2}; \quad \theta_i = -\operatorname{arctg}(a_{2k-1}/a_{2k}).$$

Представление функций рядом Фурье можно обобщить, используя базисную систему комплексных экспоненциальных функций

$$\{\exp(jk\Omega t)\}, \quad k=0, \pm 1, \pm 2, \dots \quad (2.58)$$

В этом случае имеем комплексное представление ряда Фурье

$$u(t) = \sum_{k=-\infty}^{\infty} \sqrt{2} F_k \exp(jk\Omega t), \quad (2.59)$$

коэффициенты этого ряда определяются выражением

$$F_k = \frac{1}{T} \int_0^T u(t) \exp(-jk\Omega t) dt. \quad (2.60)$$

При представлении непрерывных функций рядом Фурье практически достаточно учитывать конечное число членов ряда, так как любое реальное устройство обработки колебаний обладает конечной чувствительностью и может воспринимать их лишь в некоторой конечной полосе частот. Тогда ряд (2.54) примет вид

$$u(t) \approx \hat{u}(t) = \sum_{i=n_1}^{n_2} a_i \varphi_i(t). \quad (2.61)$$

Нетрудно видеть, что число составляющих такого усеченного ряда

$$N = 2(n_2 - n_1) = 2(f_B - f_H)T = 2\Delta f T, \quad (2.62)$$

где $f_B = \omega_B / 2\pi = n_2 1 / T$; $f_H = \omega_H / 2\pi = n_1 1 / T$ — верхняя и нижняя граничные частоты учитываемых гармоник ряда; $\Delta f = (f_B - f_H)$ — ширина спектра колебания.

Если $f_H = 0$, то

$$N = 2f_B T. \quad (2.63)$$

Выражение (2.62) определяет число степеней свободы колебания со спектром, расположенным в диапазоне частот от f_H до f_B (полосовое колебание), а (2.63) — число степеней свободы колебания со спектром от $f_H = 0$ до f_B (видеоколебание). Аналогичный результат был получен при рассмотрении

теоремы Котельникова. Различие только в том, что число N при частотном представлении колебания соответствует количеству учитываемых гармоник, а при временном — количеству отсчетов.

Строго говоря, представление непрерывных функций ограниченной длительности рядами с конечным числом составляющих математически некорректно, потому его называют *условно-аналитическим*.

По сути дела такое представление эквивалентно введению модели функции с ограниченным спектром.

Разложение по системе функций Уолша. Появление разнообразных цифровых фильтров и устройств, для реализации которых не нужны индуктивности, усилило интерес к исследованию базисных систем, более соответствующих специфике цифровой техники.

Большое внимание в последние годы уделяется ортогональным функциям Уолша. Особенность этих функций состоит в том, что при изменении во времени они могут принимать только два значения: $+1$ и -1 .

Функции Уолша можно получить из более простых ортогональных функций Радемахера [5]:

$$r_k(t) = \text{sign} \left[\sin \left(2^{k-1} \frac{2\pi t}{T} \right) \right] = \text{sign} \left[\sin \left(2^k \frac{\pi t}{T} \right) \right], \quad (2.64)$$

где $k = 1, 2, 3, \dots$; sign — сигнум-функция (знаковая функция), определяемая как

$$\text{sign}(x) = \begin{cases} +1 & \text{при } x > 0, \\ -1 & \text{при } x < 0. \end{cases} \quad (2.65)$$

Из (2.64) следует, что для функций Радемахера интервал ортогональности $[0, T]$ разбивается на 2^k равных подынтервалов ΔT , в каждом из которых функция $r_k(T)$ принимает попеременно значения $+1$ и -1 , т. е. имеет вид меандра с периодом $2\Delta T$ (см. рис. 2.5, на котором показаны функции $r_k(t)$ для $k=1; 2; 3$).

Рассмотрим процедуру получения функций Уолша разных порядков при $k=0, 1, \dots$

Функция Уолша нулевого порядка ($k=0$) вводится по определению, т. е.

$$\text{wal}_0(t) = 1.$$

Для получения функций других порядков (при $k=1, 2, \dots$) поступают следующим образом:

1. Записывают числа k в двоичной системе счисления (в прямом двоичном коде).

$$k = \sum_{i=m}^1 \gamma_i 2^{i-1} = \gamma_m 2^{m-1} + \dots + \gamma_1 2^0, \quad (2.66)$$

где m — число разрядов двоичного кода, необходимое для представления функций Уолша наиболее высокого порядка k ; γ_i — весовые коэффициенты, имеющие значение 1 или 0, в зависимости от того, учитывается или не учитывается данный i -й разряд при суммировании.

Запись числа k в двоичной системе представляет собой последовательность коэффициентов γ_i , расположенных в порядке убывания старшинства разрядов.

2. Числа k , представленные в прямом двоичном натуральном коде, перекодируют по правилу: кодовую комбинацию натурального кода складывают по модулю 2 с этой же комбинацией, сдвинутой на один разряд вправо, при этом младший разряд сдвинутой комбинации отбрасывается¹. Полученный код называется *кодом Грея*. Если число в прямом двоичном коде имеет вид $\gamma_m, \gamma_{m-1}, \dots, \gamma_1$, то в коде Грея оно записывается в виде b_m, b_{m-1}, \dots, b_1 , где в соответствии с изложенным правилом $b_m = \gamma_m$; $b_{m-1} = \gamma_m \oplus \gamma_{m-1}$; \dots ; $b_1 = \gamma_2 \oplus \gamma_1$.

3. Выражают функции Уолша через функции Радемахера по правилу

$$\text{wal}_k(t) = \prod_{i=1}^m [r_i(t)]^{b_i} = [r_1(t)]^{b_1} [r_2(t)]^{b_2} \dots [r_m(t)]^{b_m}. \quad (2.67)$$

Это правило показывает, что функции Уолша получаются перемножением функций Радемахера в определенных комбинациях, задаваемых сочетаниями коэффициентов b_i . Вид функций Уолша, полученных для $k=1, 2, 3$, в соответствии с изложенной процедурой, показан на рис. 2.6.

Отметим, что получение функций Уолша на основе функций Радемахера легко осуществляется с помощью простых логических устройств цифровой техники [5, 6].

Функции Уолша разных порядков, сведенные вместе по определенному правилу, образуют систему, называемую системой Уол-

¹ Сложение по модулю два выполняется по правилам:
 $1 \oplus 1 = 0$; $1 \oplus 0 = 1$; $0 \oplus 1 = 1$; $0 \oplus 0 = 0$.

ша $\{wal_h(t)\}$. Для этой системы характерно расположение функций в порядке возрастания числа перемен их знака (+ и -) на интервале $[0, T]$. В этой системе четные (относительно середины интервала) функции чередуются с нечетными. При этом число перемен знака на интервале $[0, T/2]$ для четных функций равно $m/2$, а нечетных $(m+1)/2$. Учитывая эти особенности, иногда используют модификацию системы $\{wal_h(t)\}$, сходную с тригонометрической и обозначаемую в виде $\{cal_h(t), sal_h(t)\}$.

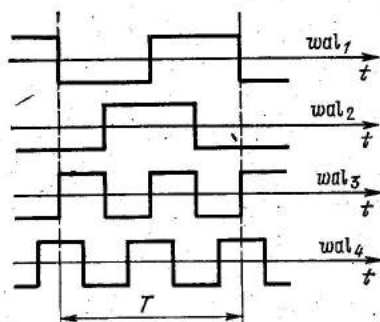


Рис. 2.6. Функции Уолша

Упорядочение (нумерация) функций Уолша может быть выполнена и иными способами. Известны системы Уолша — Адамара $\{had_h(t)\}$ и Уолша — Пейли $\{pal_h(t)\}$, в которых упорядочение функций в системе основано на применении матрицы Адамара и ее модификаций.

Каждый из способов упорядочения функций Уолша в системе имеет свои особенности, использование которых оказывается целесообразным для того или иного характера решаемых задач (стремление получить быструю сходимость ряда, удобство технической реализации и т. п.).

Теория и приложения функций Уолша находятся в стадии становления и многие вопросы ждут своего решения. Вместе с тем большое внимание к функциям Уолша и значительные достижения в этой области дают основания полагать, что теоретическое и прикладное значение таких функций будет непрерывно возрастать. В частности, в настоящее время интенсивно разрабатываются способы практического применения функций Уолша для уплотнения каналов в системах передачи цифровой информации.

Современное состояние теории и приложений функций Уолша рассмотрено в ряде работ, например [5, 6].

2.10. ГЕОМЕТРИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ КОЛЕБАНИЙ

Допустим, что некоторые математические объекты A_i (числа, функции, векторы и т. п.) являются элементами множества A . Символически это обозначается так: $A_i \in A$ (элемент A_i принадлежит множеству A).

Если над объектами A_i можно производить линейные операции (сложение и умножение на число), то множество A принадлежит к линейному пространству, а его элементы A_i — точки этого пространства. Это пространство может иметь любую размерность n . Если в таком пространстве определено расстояние между двумя точками A_i и A_j , то пространство называется метрическим. Расстояние между началом координат и какой-либо точкой A_i про-

пространства называется *нормой*, а пространство, где она введена, *нормированным*.

Линейное метрическое нормированное пространство, в котором введены следующие определения нормы и расстояния:

$$\|A_i\| = \left(\sum_{k=1}^n a_{ik}^2 \right)^{1/2}; \quad \|A_i - A_j\| = \left[\sum_{k=1}^n (a_{ik} - a_{jk})^2 \right]^{1/2} \quad (2.68)$$

называется *евклидовым*. При $n \rightarrow \infty$ (бесконечномерное пространство) евклидово пространство называется *гильбертовым*.

При геометрической интерпретации элементы множества A_i можно рассматривать как векторы, длина которых определяется величиной нормы каждого такого элемента. Тогда определения (2.68) равносильны введению в пространстве векторов A_i понятия скалярного произведения

$$A_i \cdot A_j = \sum_{k=1}^n a_{ik} a_{jk}. \quad (2.69)$$

Действительно, приняв $A_i = A_j$ из (2.69), получаем формулу для $\|A_i\|$; заменив в этой формуле A_i на $A_i - A_j$, получаем формулу для $\|A_i - A_j\|$.

Опираясь на эти краткие сведения, перейдем к рассмотрению геометрических представлений непрерывных колебаний.

Разложение непрерывных колебаний в ряд (2.47) допускает их удобную геометрическую интерпретацию в n -мерном линейном метрическом пространстве. Действительно, ортогональное разложение (2.47) по системе функций $\varphi_i(t)$ ставит в однозначное соответствие колебанию $u(t)$ набор чисел a_i (коэффициентов ряда). Эти числа можно рассматривать как координаты некоторой точки в пространстве или как проекции вектора, соединяющего начало координат с такой точкой.

Итак, колебанию $u(t)$ можно сопоставить точку или вектор \mathbf{u} в n -мерном пространстве, размерность которого равна числу степеней свободы колебания $u(t)$. При такой интерпретации между функциями и соответствующими им векторами возникают наглядные и полезные соотношения. Остановимся на некоторых из них.

Пусть колебания $u_a(t)$ и $u_b(t)$ разложены по ортогональной системе $\varphi_i(t)$, т. е. представлены в виде

$$u_a(t) = \sum_{i=0}^n a_i \varphi_i(t); \quad u_b(t) = \sum_{i=0}^n b_i \varphi_i(t). \quad (2.70)$$

В n -мерном векторном пространстве этим колебаниям соответствуют вектора \mathbf{u}_a и \mathbf{u}_b с координатами a_i и b_i . Длина (модуль) каждого из этих векторов соответственно равна:

$$|\mathbf{u}_a| = \left(\sum_{i=0}^n a_i^2 \right)^{1/2}; \quad |\mathbf{u}_b| = \left(\sum_{i=0}^n b_i^2 \right)^{1/2} \quad (2.71)$$

Приняв во внимание (2.42) и учитывая ортонормальность разложения ($Q=1$), нетрудно убедиться, что длина вектора и норма со-

ответствующего ему колебания равны. Так, для u_a имеем

$$|u_a| = \|u_a\| = \left(\frac{1}{T} \int_0^T \left[\sum_{i=0}^n a_i \varphi_i(t) \right]^2 dt \right)^{1/2} = \left(\sum_{i=0}^n a_i^2 \right)^{1/2} \quad (2.72)$$

или $|u_a| = \|u_a\| = u_a \text{эфф} = \sqrt{P_a}$.

Аналогично, для u_b имеем

$$|u_b| = \|u_b\| = \left(\frac{1}{T} \int_0^T \left[\sum_{i=0}^n b_i \varphi_i(t) \right]^2 dt \right)^{1/2} = \left(\sum_{i=0}^n b_i^2 \right)^{1/2} \quad (2.73)$$

или $|u_b| = \|u_b\| = u_b \text{эфф} = \sqrt{P_b}$, где P_a и P_b — средние удельные мощности колебаний $u_a(t)$ и $u_b(t)$.

Таким образом, длина вектора в n -мерном пространстве определяется эффективным значением соответствующего ему колебания.

Расстояние между векторами u_a и u_b равно

$$|u_a - u_b| = \left[\sum_{i=0}^n (a_i - b_i)^2 \right]^{1/2} \quad (2.74)$$

Оно характеризует степень близости векторов. Этому расстоянию соответствует величина

$$\|u_a - u_b\| = \left\{ \frac{1}{T} \int_0^T [u_a(t) - u_b(t)]^2 dt \right\}^{1/2}, \quad (2.75)$$

которая определяет степень различия между двумя колебаниями. Чем меньше величина $\|u_a - u_b\|$, тем меньше различаются колебания $u_a(t)$ и $u_b(t)$ друг от друга, и наоборот.

Скалярное произведение векторов определяется величиной

$$u_a u_b = |u_a| |u_b| \cos \theta_{ab} = \sum_{i=0}^n a_i b_i, \quad (2.76)$$

где θ_{ab} — угол между векторами.

Выражению (2.76) соответствует среднее значение произведения колебаний, определяемое как

$$\langle u_a u_b \rangle = \frac{1}{T} \int_0^T u_a(t) u_b(t) dt. \quad (2.77)$$

Это выражение характеризует эффект взаимодействия колебаний $u_a(t)$ и $u_b(t)$. Его численное значение можно рассматривать как взаимную мощность P_{ab} колебаний $u_a(t)$ и $u_b(t)$. Легко убедиться, что с учетом (2.70) и ортонормированности базиса $\{\varphi_i(t)\}$ выражение (2.77) полностью совпадает с (2.76).

Если колебания $u_a(t)$ и $u_b(t)$ ортогональны, то $\langle u_a u_b \rangle = 0$. В векторном представлении таким сигналам соответствуют взаимноперпендикулярные векторы.

Если $u_a = -u_b$ (противоположные сигналы), то $P_{ab} = -P_a = -P_b$ и соответствующие векторы одинаковой длины направлены в противоположные стороны.

Приведенные выше определения можно наглядно проиллюстрировать на примере двумерного векторного пространства (рис. 2.7).

Несмотря на то, что образы многомерного пространства графически изобразить нельзя, геометрические представления оказываются удобными и наглядными для иллюстрации многих сложных вопросов теории помехоустойчивости, разделения сигналов и др.

Рассмотренное геометрическое представление колебаний применимо в равной мере к сигналам, сообщениям и помехам. В частности, сообщения с ограниченным спектром можно интерпретировать как векторы в пространстве с числом координат, равным числу степеней свободы сообщения $N = 2F_b T$.

Так как сообщения и сигналы — некоторые функции времени, то их изменения во времени приводят к тому, что соответствующие им векторы в многомерных пространствах также меняют свое положение, описывая контурами некоторые линии, называемые *линиями сообщений и сигналов*.

Пространственное описание колебаний позволяет успешно решать различные задачи передачи сообщений с единых позиций. Впервые таким подходом очень эффективно воспользовался В. А. Котельников при исследовании проблемы помехоустойчивости [7].

В заключение отметим, что теоретической основой пространственного описания различных колебаний являются методы функционального анализа. В настоящее время эти методы широко используются при исследовании вопросов передачи информации.

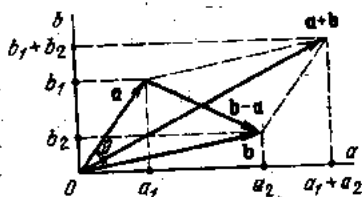


Рис. 2.7. Векторы в двумерном пространстве

КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАЧИ

1. Назовите основные виды сообщений и приведите примеры таких сообщений.
2. Каковы цели дискретизации непрерывных сообщений по времени и квантования по уровню?
3. Что такое предельная дискретизация по времени и каковы ее особенности?
4. Какие обстоятельства затрудняют реализацию предельной дискретизации по времени?
5. Что такое равномерное квантование по уровню? Как определяется ошибка при таком квантовании?
6. Энергетический спектр непрерывных сообщений определяется выражением $G_x(\Omega) = 1/[1 + (\Omega\tau)^2]$. Допустимая величина среднеквадратической ошибки преобразования этих сообщений в цифровую форму $\delta_n \leq 10^{-2}$. Определите число уровней квантования и минимально возможную частоту дискретизации по времени.

7. Какие сигналы называют модулированными?
8. Назовите основные виды помех и укажите их наиболее характерные особенности.
9. Укажите основные особенности разложений по ортогональным системам функций.
10. В чем отличие аппроксимирующей функции от исходной? Какова погрешность аппроксимации и как ее можно уменьшить?
11. Проведите нормировку следующих систем базисных функций по мощности и по энергии:
 - 1) $\{A \cos \omega_k t, A \sin \omega_k t\}, 0 \leq t \leq T;$
 - 2) $\{C \exp(j \omega_k t)\}, 0 \leq t \leq T;$
 - 3) $\left\{ \frac{\sin \Omega_B (t - k \Delta t)}{\Omega_B (t - k \Delta t)} \right\}, -\infty < t < \infty.$
12. Покажите, что при фиксированном числе членов разложения формула (2.46) обеспечивает наилучший выбор коэффициентов a_i в том смысле, что при любом другом способе их выбора СКО аппроксимации будет возрастать.
13. Получите выражения для функций Уолша первых восьми порядков и нарисуйте графики этих функций.
14. Что называется линейным метрическим пространством? В каком случае такое пространство называют евклидовым?
15. Покажите, что $d^2 = \|u_a - u_b\|^2 = P_a + P_b - 2P_{ab}$. При каком условии величина d^2 достигает максимума? Какой физический смысл имеет такое условие?
16. Определите величину различимости двух колебаний: $u_1(t) = U_1 \cos \omega_0 t$ и $u_2(t) = U_2 \cos 2\omega_0 t$.

ГЛАВА 3. ЭЛЕМЕНТЫ ОБЩЕЙ ТЕОРИИ СИСТЕМ ПЕРЕДАЧИ ИНФОРМАЦИИ

3.1. ВВЕДЕНИЕ

Теоретические основы систем передачи информации представляют собой результат усилий многих ученых и инженеров. Наиболее заметный вклад в разработку этих основ внесли в конце 40-х годов нашего столетия советский ученый В. А. Котельников и американский ученый К. Шеннон, которые широко использовали статистический подход к решению проблем передачи сообщений. Такой подход оказался очень эффективным и позволил решить ряд принципиально важных задач в различных областях радиотехники и ее приложений.

В настоящее время *статистическая теория передачи информации*, называемая также *общей теорией связи*, достигла весьма высокого уровня. Одна из наиболее важных задач этой теории — решение проблемы оптимизации систем передачи информации.

3.2. ПРОБЛЕМА ОПТИМИЗАЦИИ СИСТЕМ ПЕРЕДАЧИ ИНФОРМАЦИИ

В любой системе связи степень соответствия принятых сообщений переданным определяется тремя факторами: помехами, неидеальностью и нестабильностью характеристик передающей и приемной частей системы, неидеальностью характеристик среды распространения сигналов.

Если критерии оценки качества работы системы связи выбраны, то ее показатели будут тем выше, чем лучше выбраны способы формирования сигналов и их приема с учетом указанных факторов. Можно представить такую гипотетическую систему связи, в которой при прочих равных условиях соответствующие операции выполнены «наилучшим» (оптимальным) образом и, следовательно, показатели качества ее работы окажутся самыми высокими. С точки зрения выбранных критериев подобная система будет *оптимальной (идеальной)*.

При оптимизации систем связи желательно получить ответы на следующие вопросы:

какова структура оптимальной системы связи, т. е. какие конкретные операции необходимо выполнить, чтобы обеспечить оптимальную передачу информации в заданных условиях и при определенных наложенных ограничениях?

как оценить качество работы оптимальной системы и найти ее показатели?

как выбрать параметры оптимальной системы?

как реализовать оптимальную систему?

каковы методы приближения реальных систем к оптимальным?

К сожалению, получить исчерпывающие ответы на поставленные вопросы в настоящее время не удается из-за больших математических трудностей и отсутствия хорошо разработанных методов оптимизации сложных систем. Поэтому при решении задач оптимизации систем связи приходится существенно упрощать модели таких систем и вводить ряд дополнительных ограничений.

В настоящее время сложились два направления развития теории оптимизации систем связи: 1) оптимизация систем связи в целом; 2) оптимизация приемной части системы связи при условии, что способ формирования сигналов выбран или задан.

Рассмотрим суть указанных направлений.

Оптимизация системы связи в целом. Эта проблема впервые была поставлена и частично решена одним из основоположников теории информации К. Шенноном. Считая, что основная задача системы связи состоит в надежной (безошибочной) передаче информации с максимально возможной скоростью, Шеннон доказал ряд теорем, утверждающих существование оптимальных в указанном смысле систем. Ему удалось получить результаты, которые дают возможность оценивать качество работы оптимальных систем. Однако вопрос о структуре таких систем (синтез систем) остался открытым — теория Шеннона указала лишь общие пути его решения, суть которых состоит в отыскании наилучших методов пре-

образования сообщений — в сигналы на передающей стороне системы (*оптимальное кодирование*) и поступающей на вход приемного устройства системы смеси сигналов и помех в сообщения (*оптимальное декодирование*)¹.

Таким образом, в соответствии с подходом Шеннона *оптимальная система связи* — это такая система, в которой применены «наилучшие» методы кодирования и декодирования, обеспечивающие максимальную скорость передачи информации по каналу связи. Оптимизация в указанном смысле означает отыскание структуры системы связи, заключенной между сечениями A и A' на рис. 1.2.

При рассмотрении поставленных задач Шеннон использовал ряд существенных допущений и идеализаций. Так, он считал, что единственным видом помех является шум с нормальным законом распределения (гауссовский шум), а искажения в самой системе отсутствуют (характеристики системы идеальны). Однако, несмотря на это, конкретизировать суть оптимальных процедур кодирования и декодирования и найти способы их реализации не удалось.

Поэтому для решения задачи синтеза оптимальной системы связи приходится идти на дальнейшие упрощения, в частности, использовать *принцип декомпозиции (разбиения)* системы на отдельные части (подсистемы). В соответствии с этим принципом можно оптимизировать только ту часть системы, которая заключена между сечениями B и B' на рис. 1.2. Такую задачу называют оптимизацией системы «модулятор — демодулятор» (оптимизация «*модема*»). В результате решения этой задачи должен быть найден «наилучший» вид сигнала (оптимальный сигнал) и соответствующий оптимальный способ его приема. Подобная задача иногда называется задачей оптимизации типа «сигнал — фильтр».

Возможна и другая упрощенная постановка задачи оптимизации системы связи. Суть ее состоит в следующем. Часть системы, включая модем, считается «*черным ящиком*» и характеризуется стохастической матрицей трансформации входного множества символов $\{x_{hi}\}$ в выходное множество символов $\{\hat{x}_{hi}\}$ под действием помех. Такая матрица определяет все возможные вероятности переходов символов алфавита входного множества в символы алфавита выходного множества, т. е. элементами матрицы $\|P_{ij}\|$ являются переходные вероятности, характеризующие возможность перехода под действием помех i -го входного символа в j -й выходной символ. Значения переходных вероятностей P_{ij} определяются энергией и свойствами сигналов, характеристиками демодулятора, видом и уровнем помех. Задавая модем матрицей переходных вероятностей, можно формализовать задачу и отвлечься от его конкретных свойств. Такая формализация позволяет описать в общем виде статистические особенности системы и использовать это описание для оптимизации структуры подсистемы «*кодер — декодер*».

¹ Операциям кодирования и декодирования в подобной трактовке соответствуют рассмотренные в гл. 4 операции формирования сигналов и их приема.

(оптимизация «кодека») при заданной статистической характеристике остальной части системы, представленной «черным ящиком».

Таким образом, оптимизировать систему связи можно, оптимизируя ее отдельные части (подсистемы «модем» и «кодек»).

Необходимо отметить, что несмотря на многочисленные исследования [1, 2], проблема оптимизации даже в такой упрощенной постановке еще далека от решения. Кроме того, нужно иметь в виду, что хотя разбиение системы на части и позволяет в некоторых случаях решить задачу оптимизации до конца, такое решение не всегда эквивалентно оптимизации всей системы в целом.

Оптимизация приемной части системы связи. Для этого направления характерно допущение, что вся совокупность операций по формированию сигналов в передающей части системы и параметры сигналов-переносчиков заданы (выбраны) заранее и *точно известны* в пункте приема. Задача сводится к отысканию такого способа приема, который был бы оптимальным для заданного способа формирования сигналов и конкретных условий в пункте приема.

При решении задачи обычно налагается ряд ограничений на некоторые параметры системы (например, занимаемую полосу частот, величину пиковой или средней мощности излучаемого сигнала и др.).

Рассматриваемое направление оптимизации систем связи называется *теорией оптимальных методов приема сигналов*. Основы этой теории были разработаны В. А. Котельниковым в 1946—1947 гг. и изложены им в работе «Теория потенциальной помехоустойчивости» [3]. Эта работа положила начало интенсивным исследованиям различных вопросов оптимального приема.

Чем более полны и достоверны *априорные* (исходные) сведения о помехах и среде распространения сигналов, тем успешнее может быть решена задача оптимального приема. Наиболее полно к настоящему времени разработаны вопросы оптимального приема в каналах с аддитивными гауссовскими помехами и сигналами, точно известными в месте приема.

Несмотря на ряд существенных допущений и ограничений, при которых обычно решаются задачи оптимального приема, получаемые результаты позволяют ответить на многие принципиально важные вопросы. В частности, сравнивая показатели оптимального приема с показателями реального приема, можно судить о том, насколько целесообразно дальнейшее совершенствование последнего. Располагая показателями оптимального приема для разных способов формирования сигналов (разных видов модуляции) можно установить, какие из этих способов наиболее целесообразны в данных конкретных условиях. Важное достоинство теории оптимальных методов приема состоит в том, что во многих случаях она позволяет не только найти показатели качества приема, но и определить *структуру оптимального приемника*.

Теория оптимальных методов приема — эффективное средство при проектировании и разработке новых систем передачи инфор-

мации, позволяющее разобраться в том, какой из возможных систем и в каком случае необходимо отдать предпочтение; можно ли повысить качественные показатели системы, совершенствуя только способ приема; каковы методы реализации приема, близкого к оптимальному, и т. д.

Сочетание различных подходов к оптимизации систем связи позволяет более обоснованно и с меньшими затратами времени и средств проектировать и создавать такие системы. Вместе с тем важно помнить, что любая теория оптимизации всегда основывается на ряде упрощений и идеализаций. Поэтому нужно правильно оценивать возможности и результаты теории оптимальных систем, не преувеличивая и не преуменьшая их значения. Эта теория, как правило, указывает направление поиска решения задачи и позволяет определить предельные показатели качества работы систем, но не дает готовых рецептов реализации систем.

Ниже кратко излагаются некоторые важные положения и результаты общей теории связи.

3.3. ПРЕДЕЛЬНЫЕ ВОЗМОЖНОСТИ СИСТЕМ ПЕРЕДАЧИ ИНФОРМАЦИИ

Одна из основных концепций теории связи состоит в том, что информацию можно измерять количественно. Введя некоторую меру количественного измерения информации, можно использовать ее для оценки информационных свойств сообщений, сигналов и помех. Указанный подход впервые был детально разработан и применен к решению ряда важных задач теории связи К. Шенноном. Ключевые результаты его исследований связаны с двумя важнейшими понятиями теории информации — *производительностью источника сообщений* и *пропускной способностью канала связи*.

Любой источник создает сообщения с определенной скоростью. Чем выше эта скорость, тем больше сведений создается в единицу времени и тем выше производительность источника, которую можно рассматривать как скорость поступления информации от источника на вход канала связи. Чтобы передать нужную информацию к получателю, необходимо иметь канал с определенной пропускной способностью. Величины пропускной способности канала C и производительности источника H_i могут быть измерены количественно числом двоичных единиц (бит) в секунду и сопоставлены между собой. Рассматривая эти вопросы, Шеннон доказал следующую теорему.

Если канал связи с шумами обладает пропускной способностью C , а производительность источника определяется величиной H_i такой, что $H_i \leq C$, то возможно такое кодирование, которое обеспечивает передачу сообщений по этому каналу со сколь угодно малыми ошибками и со скоростью, сколь угодно близкой к величине C [4].

Важность теоремы состоит в том, что она утверждает принципиальную возможность безошибочной передачи сообщений по ка-

налам, в которых действуют аддитивные помехи. Это означает, что ошибки, вызванные действием таких помех в канале, могут быть устранены с помощью кодирования. Теорема Шеннона определяет условия наиболее эффективного использования канала связи и представляет собой одно из наиболее фундаментальных положений теории информации. Она служит основой многих исследований по теории кодирования. Однако следует иметь в виду, что указанная теорема представляет собой утверждение предельного типа. В действительности передавать информацию по каналу со скоростью C нельзя; к этой скорости можно только приближаться в пределе, используя все более и более сложные способы кодирования, что приводит к возрастанию задержки поступления информации к получателю [1, 4, 5].

Шеннон показал, что *пропускная способность идеального гауссовского канала связи* (канала, в котором единственной причиной искажений является белый гауссовский шум) с сигналами, у которых может быть ограниченной только средняя мощность P_c , определяется формулой [4].

$$C = \Delta f_c \log_2 (1 + P_c/P_{\text{ш}}) \text{ [бит/с]}, \quad (3.1)$$

где Δf_c — ширина полосы частот, занимаемой каналом (ширина спектра сигнала); $P_{\text{ш}} = N_0 \Delta f_c$ — средняя мощность шума в полосе частот Δf_c ; N_0 — спектральная плотность белого шума.

Формула Шеннона характеризует *предельные возможности* гауссовского канала, которые могут быть достигнуты только при «наилучших» способах формирования сигналов и их приема (оптимальном кодировании и декодировании). Исследования [1, 4, 5] показали, что такие способы физически нереализуемы и формула является асимптотическим соотношением, определяющим практически недостижимую границу скорости передачи информации по каналу связи. Это связано с тем, что по мере приближения к границе время, затрачиваемое на кодирование, стремится к бесконечности и, соответственно, увеличивается время задержки поступления информации к получателю.

Из выражения (3.1) следует, что одну и ту же пропускную способность можно получить при разных значениях полосы пропускания канала Δf_c и отношения сигнала к шуму в канале $q_c = P_c/P_{\text{ш}}$, т. е. возможен *обмен между полосой пропускания канала и отношением сигнала к шуму в нем*. Такая возможность означает, что одно и то же количество информации можно передавать по разным каналам с одинаковой скоростью. В каналах, физические свойства которых не позволяют использовать широкую полосу частот (т. е. при ограничениях на ширину спектра применяемых сигналов), приходится обеспечивать более высокое отношение сигнала к шуму по сравнению с каналами, в которых возможно применение сигналов с широким спектром.

Выясним, как изменяется пропускная способность канала при расширении его полосы пропускания. Для этого учтем, что мощность шумов в полосе пропускания канала равна $P_{\text{ш}} = N_0 \Delta f_c$ и за-

пишем (3.1) в виде

$$C = \Delta f_c \log_2(1 + P_c/N_0 \Delta f_c). \quad (3.2)$$

Нетрудно убедиться, что это выражение при беспредельном расширении полосы Δf_c стремится к пределу:

$$\lim_{\Delta f_c \rightarrow \infty} C = \frac{P_c}{N_0} \log_2 e \approx 1,44 \Delta f_0, \quad (3.3)$$

где $\Delta f_0 = P_c/N_0$ — полоса пропускания, при которой отношение сигнала к шуму в канале становится равным единице.

Разделив обе части (3.2) на Δf_c преобразуем его к виду

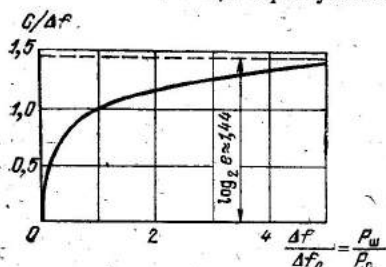
$$\frac{C}{\Delta f_0} = \frac{\Delta f_c}{\Delta f_0} \log_2 \left(1 + \frac{\Delta f_0}{\Delta f_c} \right). \quad (3.4)$$

Эта зависимость показана на рис. 3.1. Из нее следует, что с увеличением полосы Δf_c пропускная способность канала заметно возрастает только до тех пор, пока мощность сигнала в канале превышает мощность шума.

Исследования и практика показывают, что *скорость передачи информации во многих реальных системах связи значительно меньше величины, определяемой формулой Шеннона*, а попытки приблизиться к теоретическому пределу наталкиваются на ряд трудностей. В частности, в теории информации показано, что при ограниченной средней мощности сигналов пропускная способность канала определяется формулой (3.1) только тогда, когда сигналы являются шумовыми, т. е. представляют собой реализации нормального случайного процесса со средней мощностью P_c и равномерным спектром в полосе Δf_c . Такие сигналы называются гауссовскими. Несмотря на то, что гауссовские сигналы можно технически реализовать (с некоторой погрешностью), их применение с инженерной точки зрения нецелесообразно из-за неэффективного использования мощности передатчиков. Это объясняется тем, что шумовые сигналы должны иметь большой пикфактор, а при ограниченной пиковой (мгновенной) мощности реальных передатчиков такое требование приводит к существенному снижению средней мощности сигналов. По указанным причинам в большинстве прикладных задач ограничение налагается не на среднюю, а на пиковую (мгновенную) мощность излучаемого сигнала.

Рис. 3.1. Зависимость пропускной способности непрерывного канала связи от его полосы пропускания.

Определение пропускной способности канала с ограниченной пиковой мощностью сигнала наталкивается на значительные математические трудности. Они связаны с тем, что ограничение на пиковую мощность является более сильным, чем ограничение на среднюю мощность, а разнообразие возможных классов сигналов,



удовлетворяющих такому условию, весьма широко (синусоидальные, прямоугольные, треугольные и др.).

Опираясь на результаты теории информации, можно показать [4], что пропускная способность идеального гауссовского канала с ограниченной пиковой мощностью сигналов ниже пропускной способности, определяемой формулой (3.1), и может быть представлена в виде

$$C = \Delta f_c \log_2 (1 + \alpha_c P_c / P_{ш}), \quad (3.5)$$

где α_c — коэффициент, учитывающий ухудшение информационных свойств применяемого класса сигналов по сравнению с гауссовскими ($0 < \alpha_c \leq 1$); P_c — средняя мощность используемых сигналов.

Расчеты показывают, что для синусоидальных сигналов $\alpha_c \approx 0,3$, для треугольных $\alpha_c \approx 0,7$, а для прямоугольных (меандровых) $\alpha_c \approx 0,03$.

При $\alpha_c \approx 1$ (гауссовские сигналы) выражение (3.5) совпадает с (3.1), поэтому (3.5) можно рассматривать как обобщенную формулу Шеннона. Основываясь на этой формуле, найдем показатели оптимального гауссовского канала. Для этого в соответствии с критериями, введенными в п. 1.3, будем характеризовать удельные затраты энергии и полосы (затраты на передачу одного бита информации) величинами

$$\beta_E = E_0 / N_0 = P_c T_0 / N_0, \quad (3.6)$$

$$\beta_{\Delta f} = \Delta f_c / \mathcal{R}_{\max} = \Delta f_c / C, \quad (3.7)$$

где E_0 и T_0 — энергия сигнала и время, затрачиваемые на передачу одного бита информации в оптимальном канале; \mathcal{R}_{\max} — максимальная скорость передачи информации по каналу, равная его пропускной способности.

Приняв во внимание, что по определению

$$\mathcal{R}_{\max} T_0 = C T_0 = 1 \text{ [бит]} \quad (3.8)$$

и учитывая выражения (3.6) и (3.7), формулу (3.5) нетрудно преобразовать к виду

$$\frac{\Delta f_c}{C} \log_2 \left(1 + \alpha_c \frac{E_0}{N_0} \frac{C}{\Delta f_c} \right) = 1 \quad (3.9)$$

или

$$\log_2 \left(1 + \alpha_c \frac{\beta_E}{\beta_{\Delta f}} \right)^{\beta_{\Delta f}} = 1. \quad (3.10)$$

Из (3.10) следует, что

$$\beta_E = \frac{\beta_{\Delta f}}{\alpha_c} (2^{1/\beta_{\Delta f}} - 1). \quad (3.11)$$

Выражение (3.11) определяет связь между удельными затратами энергии и полосы в оптимальном гауссовском канале для разных классов сигналов. При увеличении удельных затрат полосы удельные затраты энергии уменьшаются, стремясь в пределе к

величине

$$\lim_{\beta_{\Delta f} \rightarrow \infty} \beta_E = \ln 2 / \alpha_c \approx 0,7 / \alpha_c. \quad (3.12)$$

Приняв во внимание, что

$$\frac{\beta_E}{\beta_{\Delta f}} = \frac{P_c}{P_{\text{ш}}} = \frac{P_c}{N_0 \Delta f c} \quad (3.13)$$

и учтя (3.11), имеем

$$q_c = \frac{P_c}{P_{\text{ш}}} = \frac{1}{\alpha_c} (2^{1/\beta_{\Delta f}} - 1). \quad (3.14)$$

Это выражение определяет необходимую величину отношения сигнала к шуму в оптимальном гауссовском канале связи в зависимости от удельных затрат полосы и вида сигнала.

Зависимость (3.11) для $\alpha_c = 0,3$ (синусоидальные сигналы) и $\alpha_c = 1$ (гауссовские сигналы) показана на рис. 3.2. Эта зависимость называется *границей Шеннона*. Она представляет собой геометрическое место точек; координаты каждой точки $(\beta_{\Delta f}, \beta_E)$ соответствуют вполне определенной оптимальной системе передачи информации. На этом же рисунке для $\alpha_c = 0,3$ и $\alpha_c = 1$ штриховыми кривыми нанесена зависимость (3.14).

Из проведенного рассмотрения следует, что принципиально в гауссовском канале связи можно реализовать *бесконечное множество* различных оптимальных систем, каждая из которых характеризуется соответствующими удельными затратами полосы и энергии и необходимым отношением сигнала к шуму.

Для оптимальных систем с малыми удельными затратами энергии характерны значительные удельные затраты полосы и небольшие ($q_c < 1$) отношения сигнала к шуму в канале. В системах с малыми удельными затратами полосы, наоборот, требуются значительные затраты энергии и большие отношения сигнала к шуму в канале.

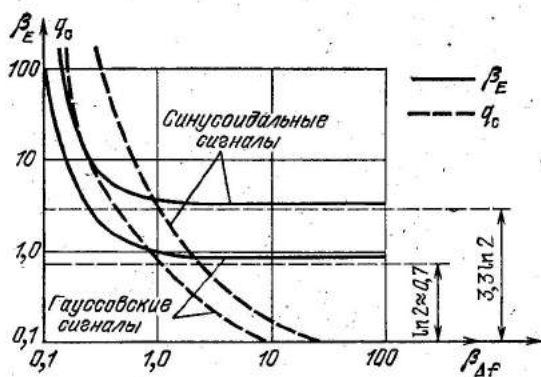


Рис. 3.2. Зависимость удельных затрат энергии от удельных затрат полосы в оптимальном гауссовском канале

Отмеченные особенности отражают общие закономерности, справедливые для любых систем связи: желание улучшить один из показателей работы системы приводит к неизбежному ухудшению другого показателя.

Так как показатели реальных систем ниже показателей оптимальных систем, то всем реальным системам связи соответствуют точки области, которая лежит выше границы Шеннона. Сравнивая координаты точки для реальной системы с границей Шеннона, можно установить, насколько далека реальная система от идеальной по соответствующим показателям.

Рассмотренные предельные соотношения теории информации и вытекающие из них выводы указывают пути решения ряда важных задач по проектированию и созданию систем связи. К числу таких задач относятся: выбор вида сигналов в зависимости от требований к системе; оценка предельных показателей работы систем и другие.

3.4. ВЫБОР СИГНАЛОВ В ОПТИМАЛЬНЫХ ГАУССОВСКИХ КАНАЛАХ

Важность задачи выбора вида сигналов определяется тем, что от этого существенно зависят показатели качества работы системы связи. Трудности решения этой задачи применительно к реальным системам обусловлены тем, что при выборе вида сигналов необходимо учитывать особенности передаваемых сообщений, специфику работы системы, технические ограничения, налагаемые на некоторые ее параметры и характеристики, а также ряд других факторов.

Указанные трудности можно уменьшить, если рассмотреть выбор сигналов в оптимальных гауссовских каналах связи. В этом случае отпадает необходимость учета ряда факторов, характерных для реальных каналов связи. Опираясь на результаты такого рассмотрения, можно выбрать сигналы, которые окажутся наиболее подходящими в конкретных условиях работы реальной системы связи. Такой подход представляется целесообразным, поскольку в реальных каналах возможное разнообразие сигналов не может быть шире, чем в оптимальных (идеальных).

Рассмотрение проведем для передачи как дискретных, так и непрерывных сообщений.

Передача дискретных сообщений. Из теории информации известно, что *максимальная средняя производительность стационарного источника дискретных сообщений* определяется величиной [4, 5]

$$(H_t)_{\max} = \lim_{T \rightarrow \infty} [\log_2 N(T)/T], \quad (3.15)$$

где $N(T)$ — число различных сообщений длительностью T , вырабатываемых дискретным источником.

В наиболее интересном для приложений частном случае, когда источник с алфавитом из M независимых равновероятных информационных символов одинаковой длительности t_0 каждый, выраба-

тывает сообщения, состоящие из m таких символов, имеем

$$(H_t)_{\max} = \lim_{m \rightarrow \infty} [\log_2 M^m / m \tau_0] = \frac{\log_2 M}{\tau_0}. \quad (3.16)$$

Это выражение показывает, что максимальная производительность источника дискретной информации в данном случае определяется только размером (основанием) алфавита M ($M \geq 2$) и длительностью информационного символа τ_0 .

В оптимальном (идеальном по Шеннону) канале связи источник информации с помощью кодирования согласован с каналом, т. е. в пределе его производительность равна пропускной способности канала

$$(H_t)_{\max} = \mathcal{R}_{\max} = C = 1/T_0. \quad (3.17)$$

Учитывая (3.8) и (3.16), имеем

$$T_0 = \tau_0 / \log_2 M. \quad (3.18)$$

Это выражение устанавливает связь между затратами времени на передачу одного бита информации T_0 и длительностью информационного символа τ_0 . Приняв во внимание (3.7) и (3.18), можно выразить удельные затраты полосы в оптимальной системе так:

$$\beta_{\Delta f} = \frac{\Delta f_c}{R_{\max}} = \Delta f_c T_0 = \frac{\Delta f_c \tau_0}{\log_2 M} = \frac{B_c}{\log_2 M}. \quad (3.19)$$

Величина

$$B_c = \Delta f_c \tau_0 \quad (3.20)$$

называется *базой сигнала*. Она характеризует соотношение между шириной спектра сигнала и длительностью его посылки, определяемой длительностью информационного символа. Из теории сигналов известно, что минимальное значение базы приблизительно равно единице.

Сигналы с базой, равной единице (или близкие к этому), называются *простыми*, а с базой, значительно превышающей единицу, — *сложными*. Рассмотрим два характерных случая.

1. Сигналы с малыми затратами полосы.

Такие сигналы должны обеспечивать выполнение условия $\beta_{\Delta f} < 1$. Из (3.19) следует, что для этого желательно применять сигналы с минимально возможной базой ($B_c \approx 1$) и увеличивать размер алфавита M ($M > 2$). Тогда для сигналов с $B_c = 1$

$$\beta_{\Delta f} = 1 / \log_2 M. \quad (3.21)$$

С учетом (3.11) и (3.21) нетрудно показать, что для таких сигналов затраты энергии определяются выражением

$$\beta_E = (M - 1) / \alpha_c \log_2 M. \quad (3.22)$$

Сигналы, удовлетворяющие условиям $B_c = 1$ и $M > 2$, называются *многоуровневыми* (многоосновными) *простыми сигналами*. Они представляют собой набор M различающихся посылок сигнала, занимающий ту же полосу частот, что и любая посылка, т. е.

полоса частот, занимаемая многоуровневым сигналом, не зависит от числа используемых посылок M . К таким сигналам относятся сигналы с многократной фазовой манипуляцией, многоуровневой амплитудной манипуляцией, а также комбинированные амплитудно-фазоманипулированные сигналы.

При выборе многоуровневых сигналов необходимо учитывать, что стремление уменьшать затраты полосы приводят к заметному увеличению затрат энергии даже в оптимальных системах (см. рис. 3.2). В реальных же системах связи увеличение затрат энергии происходит существенно более резко и поэтому в таких системах сигналы с $M \geq 8$ обычно не применяют¹.

Сигналы рассмотренного вида широко применяются в цифровых радиорелейных линиях, где несложно обеспечить большое отношение сигнала к шуму и очень важно экономно расходовать полосу.

Заметим, что получить малые затраты полосы с использованием сложных сигналов практически невозможно, так как при $B_c \gg 1$ требуется использовать огромные значения M ($M > 2^{B_c}$).

2. Сигналы с малыми затратами энергии.

Из выражения (3.12) следует, что стремление к минимально возможным затратам энергии в оптимальном канале связи приводит к очень значительным затратам полосы, т. е. при

$$\beta_E \rightarrow \beta_{E \min} = \ln 2 / \alpha_c, \quad \beta_{\Delta f} \gg 1.$$

Следовательно, в этом случае используемые сигналы должны удовлетворять условию $B_c \gg \log_2 M$, из которого следует, что $B_c \gg 1$. Это значит, что для получения малых затрат энергии нужно применять сложные сигналы с большой базой. Нетрудно убедиться также в том, что наиболее целесообразно при этом использовать двоичные сигналы ($M=2$). Из (3.19) следует, что в этом случае затраты полосы определяются величиной базы сигнала ($\beta_{\Delta f} = B_c$). С учетом (3.11) и (3.19) для таких двоичных сигналов имеем

$$\beta_E = \frac{B_c}{\alpha_c} (2^{1/B_c} - 1) \quad (\text{при } M=2). \quad (3.23)$$

Расчеты показывают, что уже при $B_c = 10$ величина β_E практически совпадает с минимальной величиной затрат $\beta_{E \min}$.

Естественно, возникает вопрос о целесообразности применения сигналов с базой $B_c > 10$, поскольку это не приводит к заметному снижению энергетических затрат. Оказывается, что с увеличением базы сигнала появляется ряд особенностей, одна из которых состоит в том, что прием сигналов с большой базой проводится в условиях очень малого отношения сигнала к шуму на входе прием-

¹ Кроме этого, в реальных системах приходится считаться с различными нестабильностями характеристик и другими неблагоприятными факторами, которые препятствуют значительному увеличению M .

ника. Действительно, из выражений (3.13) и (3.23) следует

$$q_c = \frac{P_c}{P_{\text{ш}}} = \frac{\beta_E}{\beta_{\Delta f}} = \frac{1}{\alpha_c} (2^{1/B_c} - 1). \quad (3.24)$$

При соответствующем выборе величины базы сигнала это отношение может быть сделано сколь угодно малым, т. е. окажется, что принимаемый сигнал будет «замаскирован» шумом. Отмеченная особенность сложных сигналов с большой базой открывает возможность построения скрытых систем связи, работу которых трудно обнаружить.

Известен также ряд других свойств сложных сигналов, позволяющих создавать системы передачи информации, обладающие весьма ценными качествами [6].

Передача непрерывных сообщений. При теоретических рассмотрениях непрерывные сообщения часто удобно заменять дискретными (с требуемой точностью). Предположим, что непрерывные сообщения рассматриваемого класса имеют спектр, ограниченный частотой F_B и задана среднеквадратическая ошибка их представления δ_a (см. гл. 2). Такие сообщения можно заменить последовательностью выборочных мгновенных значений (отсчетов) с частотой следования $F_{\text{п}} = 2F_B$, удовлетворяющей теореме Котельникова. Каждый отсчет этой последовательности с учетом заданной ошибки представления можно проквантовать на конечное число уровней. Указанная замена означает, что вместо источника непрерывных сообщений можно рассматривать источник дискретных сообщений. Нетрудно убедиться в том, что максимальная производительность такого источника определяется выражением

$$(H_t)_{\text{max}} = 2 F_B \log_2 \frac{1}{\delta_a} = F_B \log_2 \frac{1}{\delta_a^2}. \quad (3.25)$$

Действительно, число отсчетов (символов), создаваемых источником в одну секунду, определяется величиной $2F_B$, а число различных значений, которые может принимать каждый отсчет (размер алфавита) обратно пропорционально ошибке δ_a .

Как уже указывалось, в оптимальной шенноновской системе связи вся информация может быть передана от источника к получателю без потерь, т. е. без дополнительных ошибок, которые могут возникать в канале связи за счет помех.

Приравнявая (3.5) и (3.25), имеем

$$F_B \log_2 \frac{1}{\delta_a^2} = \Delta f_c \log_2 (1 + \alpha_c q_c). \quad (3.26)$$

Откуда

$$q_c = \frac{1}{\alpha_c} \left[\left(\frac{1}{\delta_a^2} \right)^{1/\gamma_F} - 1 \right]. \quad (3.27)$$

где $\gamma_F = \Delta f_c / F_B \geq 1$ — коэффициент, характеризующий широкополосность используемых сигналов (см. п. 2.6).

Из рассмотрения (3.27) следует, что чем шире полоса, занимаемая спектром сигналов, применяемых для передачи непрерывных сообщений, тем меньше требуемое отношение сигнала к шуму в канале, т. е. тем выше помехоустойчивость системы. Нетрудно видеть, что это качество достигается за счет увеличения затрат полосы. В частности, отсюда следует, что сигналы с широкополосной частотной модуляцией (ЧМ) обеспечивают более высокую помехоустойчивость, чем сигналы с амплитудной модуляцией (АМ).

Величина удельных затрат полосы в оптимальных системах при передаче непрерывных сообщений определяется выражением

$$\beta_{\Delta f} = \frac{\Delta f_c}{(H_f)_{\max}} = \frac{\Delta f_c}{F_B} \frac{1}{\log_2(1/\delta_a^2)} \quad (3.28)$$

Используя (3.28) и (3.27), нетрудно получить выражение для соответствующих затрат энергии в таких системах:

$$\beta_E = q_c \beta_{\Delta f} = \frac{\gamma_F}{\alpha_c} \frac{[(1/\delta_a^2)^{1/\gamma_F} - 1]}{\log_2(1/\delta_a^2)} \quad (3.29)$$

Результаты проведенного рассмотрения возможных случаев передачи информации в каналах связи, оптимальных по Шеннону, можно использовать при решении задач, связанных с выбором сигналов, удовлетворяющих заданным требованиям. В частности, опираясь на эти результаты, можно более обоснованно подходить и к решению задач оптимального приема.

3.5. ЭЛЕМЕНТЫ ТЕОРИИ ОПТИМАЛЬНОГО ПРИЕМА СИГНАЛОВ

Прием сигналов — одна из наиболее сложных теоретических и инженерных задач передачи сообщений. Сложность состоит в том, что в пункте приема сообщения необходимо извлекать из модулированных сигналов-переносчиков, которые в процессе прохождения по линии связи не только ослабляются, но и подвергаются воздействию различных искажающих факторов и помех.

Весьма желательно располагать методами приема, которые были бы наилучшими (оптимальными) в данных конкретных условиях. Направление, связанное с отысканием таких методов, называется *теорией оптимального приема*. Эта теория занимается решением следующих задач:

синтез оптимального приемника, т. е. нахождение правила (алгоритма) работы оптимального приемника, соответствующего принятым критериям, заданным условиям работы и наложенным ограничениям, а также установление его структуры;

анализ оптимального приемника, т. е. определение показателей качества его работы, в первую очередь, помехоустойчивости;

анализ «устойчивости» алгоритма оптимального приема. Решение задачи синтеза оптимального приема, как правило, основывается на ряде допущений и идеализаций, которые не всегда могут быть в хорошем соответствии с действительностью. Неточное знание условий работы системы часто называют априорной недоста-

точностью (в реальных ситуациях это неизбежно). Оценка влияния априорной недостаточности на алгоритм работы оптимального приемника и составляет суть задачи. Ее цель — установление границ применимости найденного алгоритма;

отыскание способов инженерной реализации оптимального приемника. Структуру оптимального приемника, основанную на ряде идеализаций, часто не удается непосредственно реализовать на практике. Однако опираясь на теорию оптимального приема, можно построить реальные приемные устройства с показателями, близкими к оптимальным. Такие реальные устройства называются подоптимальными (субоптимальными).

Теоретической основой решения задач оптимального приема является *теорема Байеса*.

Некоторые понятия теории статистических решений. Пусть некоторая случайная физическая величина, которую назовем *причиной*, может принимать множество значений (исходов) Π с плотностью вероятностей $p(\Pi)$, которая считается *априорной* (заранее известной). Пусть причина вызывает появление другой случайной величины — *следствия* C , которое также может принимать множество значений. Плотность вероятностей этих значений зависит от конкретных исходов причины. Поэтому ситуация описывается *множеством условных плотностей вероятностей* $p(C|\Pi)$.

Статистическим решением называют процедуру, которая состоит в том, чтобы, наблюдая конкретное следствие $C_i (C_i \in C)$, указать вызвавшую его причину $\Pi_i (\Pi_i \in \Pi)$. Так как наблюдаемое следствие C_i может быть вызвано любым исходом причины Π , то можно определить плотность вероятностей всех возможных исходов, которые могли вызвать данное следствие, т. е. определить функцию $p(\Pi|C_i)$. Эта функция называется *апостериорной* (последопытной, установленной на основе имевшего место опыта или наблюдения) *плотностью вероятностей причин*.

Основой для принятия статистического решения является теорема Байеса [3, 7].

$$p(\Pi|C) = \frac{p(\Pi)p(C|\Pi)}{p(C)}, \quad (3.30)$$

где $p(C|\Pi)$ — условная плотность распределения следствий; $p(C)$ — безусловная плотность распределения следствий C , определяемая как

$$p(C) = \int_{\Gamma} p(\Pi)p(C|\Pi) d\Pi.$$

Значение этого интеграла не зависит от Π , поскольку интегрирование по этой переменной ведется по всей области ее существования Γ .

Из (3.30) следует, что апостериорная плотность вероятностей причины $p(\Pi|C)$ зависит от априорной плотности вероятностей причины $p(\Pi)$ и условной плотности вероятностей следствий $p(C|\Pi)$. Плотность $p(C|\Pi)$ является функцией Π , ее называют *функцией правдоподобия*.

В теории статистических решений показано [1, 5], что при принятии решения о конкретном значении действовавшей причины Π_i , вызвавшей наблюдаемое (или заданное) следствие C_i , наименьшую ошибку можно совершить, если выносить решение в пользу того значения причины, при которой условное распределение $p(\Pi | C_i)$ имеет наибольшее значение. Такое правило принятия решения называется *байесовским*.

Если априорная плотность $p(\Pi)$ неизвестна, то самое большее, что можно сделать — предположить равномерность ее распределения. Тогда решение будет выноситься в пользу того значения причины $\Pi_i \in \Pi$, при котором функция правдоподобия $p(C | \Pi)$ для наблюдаемого следствия $C_i \in C$ принимает наибольшее значение. Это означает, что такое значение причины считается наиболее правдоподобным среди других возможных значений. Подобная процедура принятия решения называется *правилом максимального правдоподобия*.

Применим изложенный подход к решению задачи оптимального приема сигналов.

Суть процедуры оптимального приема. В гл. 2 было установлено, что между колебаниями и векторами можно установить взаимно-однозначное соответствие. Поэтому вместо колебаний можно рассматривать соответствующие векторы. Исходя из этого, будем считать причиной Π случайный вектор x , соответствующий передаваемым сообщениям (или однозначно связанный с ним вектор сигналов s , переносящих эти сообщения), а следствием C — случайный вектор y , соответствующий смеси сигнала и шума на входе приемника. С учетом сказанного (3.30) можно записать либо в виде

$$p(x|y) = [p(x)p(y|x)]/p(y) \quad (3.31)$$

либо в эквивалентном выражению (3.31) виде

$$p(s|y) = [p(s)p(y|s)]/p(y), \quad (3.32)$$

где x , s , y — векторы в многомерных пространствах, соответствующие сообщениям $x(t)$, сигналам $s(t) = s[x(t), t]$ и входным реализациям $y(t) = s(t) + n(t)$.

При передаче дискретных сообщений множество сообщений $x(t)$ может принимать только конечное число дискретных значений $(x_1, x_2, \dots, x_j, \dots, x_M)$, которому однозначно соответствует конечное число различающихся сигналов $s_1(t), s_2(t), \dots, s_j(t), \dots, s_M(t)$.

Оптимальная процедура приема состоит в определении величин $p(s|y)$ для всех M значений $s_j(t)$, сравнении этих величин между собой и выборе наибольшей из них. Значение $s_j(t)$, которому соответствует максимальная величина $p(s_j|y)$, считается переданным сигналом и в соответствии с этим на выходе приемника воспроизводится сообщение x_j .

Основная трудность при решении такой задачи связана с нахождением апостериорного распределения $p(s|y)$. Наиболее детально задача решена для помехи типа гауссовского белого шума и набора сигналов, заранее известных в точке приема. Если при

этом все сообщения x_j равновероятны и независимы, то выражение для $p(s|y)$ можно привести к виду

$$p(s|y) = A \exp \left\{ -\frac{1}{N_0} (y-s)^2 \right\}. \quad (3.33)$$

где N_0 — односторонняя спектральная плотность мощности белого гауссовского шума; A — некоторая константа, см. [1—3, 7].

Нахождение сигнала $s_j(t)$, максимизирующего величину (3.33) при наблюдении на входе приемника некоторой реализации $y(t)$, эквивалентно минимизации показателя экспоненты. Следовательно, оптимальный приемник должен выносить решение о приеме того сигнала $s_j(t)$, при котором функция $p(s_j|y)$ достигает максимума, а величина

$$d_j^2 = (y - s_j)^2 \quad (3.34)$$

соответственно становится минимальной.

Учитывая свойства векторного представления функций времени, от выражения (3.34), можно перейти к эквивалентному ему выражению

$$\langle d_j^2(t) \rangle = \int_0^T \{y(t) - s_j(t)\}^2 dt. \quad (3.35)$$

Выражение (3.34) или (3.35) представляет собой алгоритм работы оптимального приемника дискретных сообщений. Работая по этому алгоритму, оптимальный приемник должен вычислить значения величины $\langle d_i^2(t) \rangle$ для всех M , используемых в системе сигналов $s_j(t)$ (где $j=1, 2, \dots, M$), сравнить их между собой, выбрать наименьшее значение и воспроизвести на выходе соответствующее ему дискретное сообщение.

Иными словами, оптимальный приемник всегда воспроизводит на выходе сообщение, переносимое тем сигналом, к которому наиболее близка входная реализация $y(t)$. В геометрической интерпретации это означает, что оптимальный приемник всегда относит вектор входной реализации y к ближайшему вектору сигнала.

Очевидно, что прием сигналов в присутствии шума может привести к ошибкам, поскольку вектор входной реализации случаен и с некоторой вероятностью может попасть в любую точку пространства. Допустим, что вектор y , образованный из переданного сигнала s_i и шума n , попал в точку, наиболее близко расположенную к вектору сигнала s_j . Если $i=j$, то приемник примет правильное решение, если же $i \neq j$, то решение приемника окажется ошибочным и вместо переданного сообщения x_i он ошибочно воспроизведет сообщение x_j .

Несмотря на то, что оптимальный приемник дискретных сообщений может допускать ошибочные решения, их вероятность у этого приемника минимальна по сравнению с любыми реальными приемниками таких сообщений.

Исследования показывают, что алгоритм (3.35) может быть представлен в более удобном для схемной реализации виде и позволяет получить структурные схемы оптимальных приемников и

выражения для расчета их помехоустойчивости. Эти вопросы рассмотрены в гл. 7.

Перейдем к рассмотрению особенностей оптимального приема при передаче непрерывных сообщений. В этом случае передаваемое сообщение $x(t)$ может иметь очень большое (практически бесконечное) число возможных реализаций, каждая из которых представляет собой непрерывную функцию времени. Поэтому в геометрической интерпретации сообщениям и сигналам соответствуют не отдельные точки (или векторы с фиксированной длиной) в многомерных пространствах (как это было при передаче дискретных сообщений), а континуум линий сообщений и сигналов, описываемых концами векторов x и s . Исследования [1, 7] показывают, что в этой ситуации оптимальный прием связан с формированием на приемной стороне такого сигнала $\hat{s}(t)$, который бы обеспечивал максимум максимум апостериорной плотности вероятности, определяемой выражением (3.32).

Применительно к каналу с гауссовским белым шумом и равновероятными сообщениями указанное условие сводится к минимизации величины

$$\langle d^2(t) \rangle = \int_0^T [y(t) - \hat{s}(t)]^2 dt. \quad (3.36)$$

Чтобы сформировать сигнал $\hat{s}(t)$, на приемной стороне нужно использовать принятое сообщение $\hat{x}(t)$, которое представляет собой результат обработки входной реализации $y(t)$ приемником. Сообщение $\hat{x}(t)$ называют *оценкой* переданного сообщения $x(t)$. Формирование сигнала $\hat{s}(t)$ представляет собой модуляцию несущей сигнала колебанием $\hat{x}(t)$ по тому же закону и с теми же параметрами, что и на передающей стороне.

Сформированный в приемнике сигнал $\hat{s}(t)$ используется при обработке входной реализации $y(t)$ и последующем формировании оценки сообщения $\hat{x}(t)$, которая, в свою очередь, необходима для создания сигнала $\hat{s}(t)$. Нетрудно понять, что указанная процедура может быть выполнена только в устройстве следящего типа, с использованием обратной связи по формируемой оценке сообщения $\hat{x}(t)$. Упрощенная функциональная схема, поясняющая принцип работы такого оптимального приемного устройства, показана на рис. 3.3.

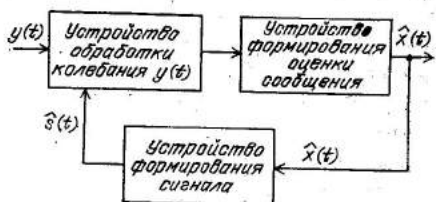


Рис. 3.3. Упрощенная функциональная схема оптимального приемника непрерывных сообщений

В геометрической интерпретации минимизация выражения (3.36) означает, что оптимальный приемник всегда относит входную текущую реализацию y к ближайшей линии сигналов и в соответствии с этим формирует на выходе оценку сообщения $\hat{x}(t)$. Из-за влияния шума оценка $\hat{x}(t)$ отличается от переданного со-

общения $x(t)$. Это отличие обычно характеризуют величиной *среднеквадратической ошибки* (см. п. 1.3). Оптимальный прием обеспечивает минимальное значение этой ошибки по сравнению с любым другим способом приема.

Теория оптимального приема непрерывных сообщений, часто называемая также *теорией оптимальной демодуляции* аналоговых видов модуляции, или *теорией нелинейной фильтрации*, представляет важный раздел общей теории связи, основы которой были заложены в работах А. Н. Колмогорова, В. А. Котельникова, Н. Винера, К. Шеннона и ряда других отечественных и зарубежных ученых.

Несмотря на интенсивное развитие и ряд достижений теории оптимальной демодуляции, задача синтеза оптимального физически реализуемого демодулятора в достаточно общем виде до сих пор еще не решена.

Развитие теории оптимальной демодуляции непрерывных модулированных сигналов в настоящее время идет по двум направлениям:

гауссовская теория оптимальной демодуляции, основанная на предположении, что сообщения и помехи являются гауссовскими процессами с известными статистическими характеристиками;

марковская теория оптимальной демодуляции, основанная на описании сообщений и помех марковскими случайными процессами.

Необходимо отметить, что описание сообщений и помех гауссовскими процессами в ряде случаев физически оправдано, так как достаточно хорошо соответствует реальной ситуации. В то же время в ряде приложений допустимо аппроксимировать реальные процессы марковскими. Это позволяет при решении задач оптимального приема пользоваться одновременно предположениями о марковском и гауссовском характере процессов, описывающих сообщения и помехи. В соответствии с этим при синтезе оптимальных демодуляторов часто можно пользоваться методами любого из указанных направлений. Однако ответить на вопрос о преимуществах того или иного подхода затруднительно, поскольку для одних задач удобнее гауссовское, а для других — марковское представление процессов. Оба направления детально разработаны лишь для достаточно больших отношений сигнала к шуму на входе приемника. При таком условии распределение (3.32) имеет только один явно выраженный максимум. Для более низких отношений сигнала к шуму это распределение может иметь несколько максимумов и тогда критерий максимума правдоподобия оказывается непригодным, так как существенно изменяется характер ошибки и основную роль начинают играть *аномальные ошибки* (см. п. 1.3).

Приложения теории оптимального приема к некоторым случаям передачи непрерывных сообщений рассмотрены в гл. 5.

Марковская теория нелинейной фильтрации и ее приложения детально рассмотрены в фундаментальной работе [8].

КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАЧИ

1. Перечислите основные задачи, решаемые теорией оптимизации систем связи, и поясните их сущность.
2. Дайте характеристику основных направлений развития теории оптимизации систем связи.
3. Какие идеализации и допущения характерны при решении задач оптимизации системы связи в целом и оптимизации только приемной части системы?
4. Какие вопросы рассматриваются в теории оптимальных методов приема?
5. Сформулируйте теорему Шеннона о безошибочной передаче сообщений по каналу связи с шумами. Поясните эту теорему с физической точки зрения.
6. Считая пропускную способность двух гауссовских каналов одинаковой, докажите справедливость соотношения

$$q_2 = (1 + q_1)^{\Delta f_1 / \Delta f_2} - 1,$$

- где q_1 , q_2 и Δf_1 , Δf_2 — соответствующие отношения сигнала к шуму и полосы пропускания каналов.
7. Дайте физическое объяснение, почему изменение полосы пропускания канала влияет на его пропускную способность более резко, чем изменение отношения сигнала к шуму?
 8. Покажите, что при увеличении отношения сигнала к шуму в канале влияние класса используемых сигналов на величину пропускной способности уменьшается, а при уменьшении — увеличивается. Дайте этому физическое объяснение.
 9. Определите удельные затраты энергии и необходимое отношение сигнала к шуму q_0 в оптимальном гауссовском канале, если $\beta_{\Delta f} = 0,1$; $\beta_{\Delta f} = 0,25$; $\beta_{\Delta f} = 20$; $\beta_{\Delta f} = 200$.
 10. Объясните, почему в системах связи с широкополосными сигналами требуются меньшие отношения сигнала к шуму, чем в системах с узкополосными сигналами?
 11. Что называют границей Шеннона?
 12. Какими задачами занимается теория оптимального приема?
 13. В чем сущность теоремы Байеса? Что такое статистическое решение?
 14. Нарисуйте структуру оптимального приемника дискретных сообщений.
 15. Почему оптимальный приемник дискретных сообщений всегда выносит решение о приеме того сигнала, к которому наиболее близка входная реализация $y(t)$?
 16. Как работает оптимальный приемник непрерывных сообщений? Какова его структура?

ГЛАВА 4. СИГНАЛЫ В СИСТЕМАХ ПЕРЕДАЧИ НЕПРЕРЫВНЫХ СООБЩЕНИЙ И МЕТОДЫ ИХ ФОРМИРОВАНИЯ

4.1. ВВЕДЕНИЕ

При передаче непрерывных сообщений необходимо выбирать сигналы, наиболее целесообразные в данных конкретных условиях. Для этого нужно знать основные особенности сигналов, формируемых в передающей части системы связи. Поскольку современные системы связи, как правило, многоканальные, важной операцией при таком формировании является уплотнение каналов. Использование различных методов уплотнения каналов в сочетании с разными видами модуляции сигналов позволяет получить ряд вариантов построения многоканальных систем передачи непрерывных сообщений.

В этой главе рассмотрены наиболее часто применяемые виды модуляции и уплотнения, а также методы формирования модулированных сигналов в указанных системах.

4.2. ОБЩИЕ СООТНОШЕНИЯ ПРИ МОДУЛЯЦИИ КОЛЕБАНИЙ

Пусть имеется некоторое колебание $u(t)$ с рядом характеризующих его постоянных параметров $\Pi_{01}, \Pi_{02}, \dots, \Pi_{0i}, \dots, \Pi_{0n}$, т. е.

$$u(t) = u(\Pi_{01}, \Pi_{02}, \dots, \Pi_{0n}, t). \quad (4.1)$$

Модуляция колебания $u(t)$ представляет собой операцию, связанную с изменением одного (или нескольких) его параметров в соответствии с модулирующим колебанием $u_x(t)$.

Наиболее часто для модуляции используют гармонические колебания и периодические импульсные последовательности. Такие колебания легко генерируются, сравнительно просто синхронизируются и удобны при обработке.

Модуляция может быть полезной, т. е. выполняться для получения нужного результата, и паразитной (вредной), возникающей под действием некоторых нежелательных факторов (случайных изменений условий работы и т. п.).

Ниже речь будет идти только о полезной модуляции, выполняемой специально предназначенным для этого устройством — модулятором.

Модулированный параметр можно записать так:

$$\Pi(t) = \Pi_0 + \Delta\Pi(t) = \Pi_0 + K_M u_x(t), \quad (4.2)$$

где K_M — чувствительность модулятора. Она показывает насколько изменяется модулируемый параметр при изменении сообщения. Зависимость приращения параметра $\Delta\Pi$ от модулирующего напряжения u_x называется *модуляционной характеристикой* (рис. 4.1). Чем больше крутизна этой характеристики, тем выше чувстви-

ность модулятора. Чтобы параметр Π изменялся пропорционально колебанию u_x , необходимо иметь $K_M = \text{const}$, т. е. использовать линейный участок модуляционной характеристики, где $K_M = \text{tg } \alpha$ (см. рис. 4.1).

Представим $u_x(t)$ в виде

$$u_x(t) = U_{x \max} \frac{u_x(t)}{U_{x \max}} = U_{x \max} x(t), \quad (4.3)$$

где $U_{x \max}$ — максимальное значение $u_x(t)$; $x(t)$ — нормированное сообщение, удовлетворяющее условию

$$-1 \leq x(t) \leq 1. \quad (4.4)$$

Далее везде будем считать, что рассматриваемые сообщения нормированные, их среднее значение равно нулю, а спектр ограничен частотой $F_a = \Omega_a/2\pi$.

Удельная средняя мощность сообщения $x(t)$ определяется выражением

$$P_x = \langle x^2(t) \rangle = \frac{1}{T} \int_0^T x^2(t) dt = x_{\text{эфф}}^2 = 1/k_{\text{пх}}^2, \quad (4.5)$$

где $k_{\text{пх}} = x_{\text{max}}/x_{\text{эфф}}$ — пикфактор сообщения. Так как $x_{\text{max}} = 1$ (по определению), то $k_{\text{пх}} = 1/x_{\text{эфф}}$.

С учетом (4.3)

$$\Delta \Pi(t) = K_M U_{x \max} x(t) = \Delta \Pi_m \frac{U_{x \max}}{U_{\text{лм}}} x(t),$$

где $\Delta \Pi_m$ — максимальное отклонение параметра, соответствующее линейной части характеристики с протяженностью $U_{\text{лм}}$ (см. рис. 4.1). Для полного использования характеристики необходимо выбрать $U_{x \max} = U_{\text{лм}}$. В этом случае

$$\Delta \Pi(t) = K_M U_{\text{лм}} x(t) = \Delta \Pi_m x(t). \quad (4.6)$$

Рассмотренные соотношения позволяют правильно согласовать передаваемые сообщения с параметрами модулятора.

4.3. ГАРМОНИЧЕСКИЕ МОДУЛИРОВАННЫЕ КОЛЕБАНИЯ

Гармонические колебания широко применяются как сигналы-переносчики информации в среде между передатчиком и приемником, а также как вспомогательные колебания при операциях уплотнения и разделения каналов. Наиболее распространены три основных вида модуляции таких колебаний: амплитудная (АМ), фазовая (ФМ) и частотная (ЧМ). Рассмотрим особенности этих видов модуляции.

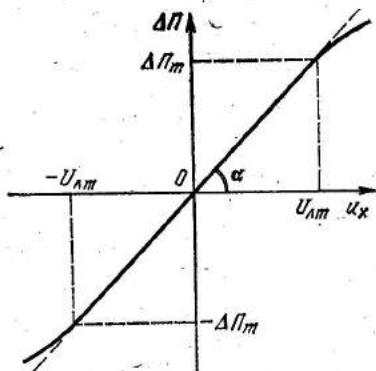


Рис. 4.1. Модуляционная характеристика

Амплитудная модуляция и ее разновидности. При амплитудной модуляции огибающая гармонического колебания становится линейной функцией сообщения, т. е.

$$S(t) = S_0 + \Delta S_m x(t) = S_0 [1 + (\Delta S_m/S_0) x(t)], \quad (4.7)$$

где $\Delta S_m = \Delta P_m$ — максимальное отклонение огибающей при модуляции.

С учетом (4.7) АМ сигнал принимает вид

$$s_{AM}(t) = S_0 [1 + M_a x(t)] \cos(\omega_0 t + \varphi_0), \quad (4.8)$$

где $M_a = \Delta S_m/S_0$ — коэффициент амплитудной модуляции.

Как известно, спектр АМ сигнала состоит из несущей и двух боковых полос (нижней и верхней). Он занимает диапазон частот от $(f_0 - F_B)$ до $(f_0 + F_B)$, т. е. ширина спектра АМ сигнала равна

$$\Delta f_{AM} = 2F_B, \quad (4.9)$$

где F_B — верхняя граничная частота спектра сообщения $x(t)$. Приняв во внимание выражения (2.30) и (4.5), нетрудно установить, что мощность АМ сигнала равна

$$P_{AM} = \frac{S_0^2}{2} [1 + M_a^2 \langle x^2(t) \rangle] = \frac{S_0^2}{2} \left(1 + \frac{M_a^2}{k_{пх}^2} \right). \quad (4.10)$$

Эта мощность складывается из мощности несущей $P_0 = S_0^2/2$ и мощности боковых полос $2P_6 = P_0 M_a^2/k_{пх}^2$.

Для большинства непрерывных сообщений $k_{пх} \approx 2-3$ и поэтому мощность боковых полос составляет не более 10—20% мощности несущей. Так как полезная информация заключена только в боковых полосах АМ сигнала, с информационной точки зрения при АМ мощность передатчика используется неэффективно. Для устранения этого недостатка иногда применяют АМ сигналы с *подавленной несущей* (АМ—ПН).

Нетрудно показать, что при использовании одного и того же передатчика мощность боковых полос сигнала АМ—ПН можно увеличить в четыре раза по сравнению с АМ сигналом (при $M_a = 1$).

Другой, еще более эффективной разновидностью АМ сигналов является *однополосная амплитудная модуляция с подавленной несущей* (ОАМ—ПН). Этот вид модуляции представляет собой такое преобразование, при котором спектр сигнала полностью совпадает со спектром сообщения, перенесенным по оси частот в нужную область. Однополосную модуляцию называют также *модуляцией с одной боковой полосой* (ОБП). При ОБП вся мощность передатчика тратится на создание информационного спектра. Полоса частот сигнала с ОБП равна ширине спектра сообщения, т. е. $\Delta f_{об} = F_B$.

В настоящее время известны два метода формирования сигналов с ОБП: *фильтрационный* (спектральный) и *фазовый* (корреляционный). Суть фильтрационного метода поясняется рис. 4.2. В этом методе высокие требования предъявляются к фильтру, ко-

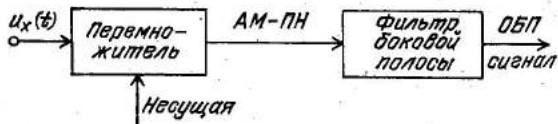


Рис. 4.2. Схема формирования ОБП сигнала фильтрационным методом

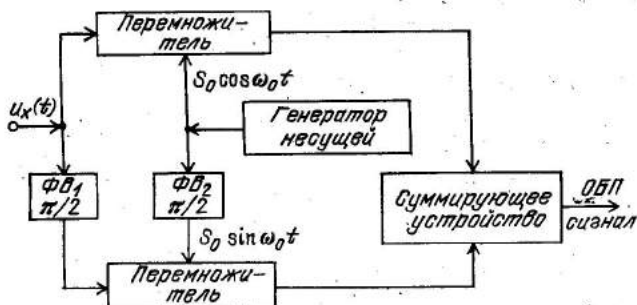


Рис. 4.3. Схема формирования ОБП сигнала фазовым методом

торый должен выделять без искажений нужную боковую полосу частот и подавлять все остальные составляющие спектра. Для этой цели обычно применяются кварцевые или электромеханические фильтры.

Фазовый метод поясняется рис. 4.3, где фазовращатель $\Phi В_1$ поворачивает фазу всех составляющих сообщения $x(t)$ на $\pi/2$, а $\Phi В_2$ — фазу несущей на $\pi/2$. Можно показать, что при суммировании выходных напряжений перемножителей образуется ОБП сигнал с нижней боковой, а при вычитании — ОБП сигнал с верхней боковой.

В качестве перемножителей в обоих методах обычно используют балансные модуляторы, которые позволяют уменьшить число побочных продуктов (гармоник), появляющихся при этой нелинейной операции. Необходимо отметить, что степень подавления «ненужной» боковой полосы в обоих методах существенно зависит от стабильности параметров схем и стабильности частоты используемых несущих колебаний. Особенно жесткие требования в фазовом методе предъявляются к стабильности амплитуд и фаз опорных напряжений, а также к точности балансировки схемы.

Трудности реализации сигналов с ОБП приводят к тому, что практически они могут применяться только для передачи сообщений, не содержащих очень низкочастотных составляющих (например, для телефонных сообщений со спектром от 300 Гц до 3400 Гц, телевизионных, факсимильных и т. п.). Для сообщений с узкими спектрами (например, телеметрических, спектр которых имеет ширину от нуля до нескольких герц или десятков герц) применение ОБП очень затруднительно или даже невозможно.

Наиболее широкое применение ОБП находит в многоканальных системах передачи телефонных сообщений.

Фазовая модуляция. При фазовой модуляции амплитуда сигнала остается неизменной ($S(t) = S_0$), а фаза изменяется в соответствии с сообщением, т. е.

$$\Phi(t) = \Phi_0(t) + \Delta\Phi(t) = \omega_0 t + \Delta\Phi_m x(t) + \Phi_0, \quad (4.11)$$

где $\Delta\Phi_m = \Delta\Pi_m$ — максимальное отклонение фазы, называемое *индексом фазовой модуляции*.

Для удобства далее будем полагать начальную фазу $\Phi_0 = 0$.

Выражение для сигнала ФМ с учетом (4.11) можно записать так:

$$s_{\text{ФМ}}(t) = S_0 \cos[\omega_0 t + \Delta\Phi_m x(t)]. \quad (4.12)$$

Как известно, спектр ФМ сигнала даже для сообщений с ограниченным спектром содержит бесконечное количество составляющих, образующих нижнюю и верхнюю боковые полосы. Поэтому недостаточное пропускание боковых полос спектра ФМ (ограничение спектра по частоте) может привести к появлению нежелательной паразитной амплитудной модуляции и искажению сообщения, содержащегося в ФМ сигнале.

Исследования и расчеты показывают, что при проведении инженерных расчетов ширину спектра ФМ сигнала можно определять по приближенной формуле [1]

$$\Delta f_{\text{ФМ}} \approx 2F_v(1 + \Delta\Phi_m), \quad (4.13)$$

где F_v — граничная частота спектра сообщения.

Средняя удельная мощность ФМ сигнала в соответствии с (2.30) равна

$$P_{\text{ФМ}} = S^2/2 = P_0. \quad (4.14)$$

Это означает, что мощность ФМ сигнала не зависит от параметров модуляции; в процессе модуляции она остается неизменной и равной мощности несущей. Поэтому передатчик ФМ может постоянно работать в режиме отдачи максимальной мощности. Модуляция вызывает только перераспределение мощности несущей между спектральными составляющими, которое зависит от выбранной величины индекса фазовой модуляции $\Delta\Phi_m$.

Частотная модуляция. При частотной модуляции амплитуда сигнала остается неизменной, а мгновенная частота изменяется в соответствии с сообщением

$$\omega(t) = \omega_0 + \Delta\omega(t) = \omega_0 + \Delta\omega_m x(t), \quad (4.15)$$

где $\Delta\omega_m = 2\pi\Delta f_m$ — максимальное отклонение частоты при модуляции (*девиация частоты*).

Учитывая, что полная фаза при ЧМ равна

$$\Phi_{\text{ЧМ}}(t) = \int_0^t \omega(t) dt = \omega_0 t + \Delta\omega_m \int_0^t x(t) dt, \quad (4.16)$$

выражение для ЧМ сигнала можно записать в виде

$$s_{\text{ЧМ}}(t) = S_0 \cos \left[\omega_0 t + \Delta\omega_m \int_0^t x(t) dt \right], \quad (4.17)$$

или

$$s_{\text{ЧМ}}(t) = S_0 \cos [\omega_0 t + \Delta\omega_m \theta(t)], \quad (4.18)$$

где

$$\theta(t) = \int_0^t x(t) dt \quad (4.19)$$

— интегральная функция сообщения.

Так как общая фаза ЧМ сигнала зависит от интегральной функции сообщения, ЧМ часто называют *интегральной модуляцией*. Выражения (4.12) и (4.18) имеют одинаковую структуру. Это объясняется тем, что ФМ и ЧМ представляют собой разновидности более общего вида модуляции — *угловой*. Однако то обстоятельство, что при ФМ изменения частоты — результат модуляции фазы, а при ЧМ изменения фазы — результат модуляции частоты приводит к существенному различию спектров этих видов модуляции.

Чтобы выяснить суть различия, рассмотрим выражение для мгновенной частоты при ФМ, которое с учетом (4.11) имеет вид

$$\omega(t) = d\Phi(t)/dt = \omega_0 + \Delta\Phi_m x'(t), \quad (4.20)$$

где $x'(t) = d[x(t)]/dt$ — производная сообщения. Максимальное отклонение мгновенной частоты при ФМ (девиация частоты при ФМ) определяются величиной

$$\Delta\omega_m = 2\pi \Delta f_m = \Delta\Phi_m \max |x'(t)|. \quad (4.21)$$

Приняв во внимание неравенство Бернштейна [2], в соответствии с которым

$$\max |x'(t)| \leq \Omega_B \max |x(t)|, \quad (4.22)$$

где $\Omega_B = 2\pi F_B$ — граничная частота спектра $x(t)$, и учтя, что $\max |x(t)| = 1$, имеем

$$\Delta\omega_m = \Delta\Phi_m \Omega_B \quad \text{или} \quad \Delta f_m = \Delta\Phi_m F_B. \quad (4.23)$$

Это выражение связывает параметры угловой модуляции с шириной спектра сообщения. С учетом (4.23) справедливы следующие соотношения между индексом модуляции и девиацией частоты:

$$\text{для ФМ } \Delta\Phi_m = \text{const}, \quad \Delta f_m = \Delta\Phi_m F_B, \quad (4.24)$$

$$\text{для ЧМ } \Delta f_m = \text{const}, \quad \Delta\Phi = \Delta f_m / F_B. \quad (4.25)$$

С учетом (4.25) формула (4.13) принимает вид

$$\Delta f_{\text{ЧМ}} = 2 F_B (1 + \Delta\Phi_m / F_m) = 2 \Delta f_m (1 + 1/\Delta\Phi_m). \quad (4.26)$$

Из формулы следует, что при $\Delta\Phi_m > 5$ ширина спектра ЧМ определяется в основном величиной девиации частоты. По этой при-

чине в инженерных расчетах часто пользуются приближенным выражением

$$\Delta f_{\text{ЧМ}} \approx 2,6 \Delta f_m. \quad (4.27)$$

Ширина спектра ФМ сигнала пропорциональна ширине спектра сообщения и, следовательно, для разных сообщений она различна. Поэтому на практике при передаче непрерывных сообщений предпочитают использовать ЧМ, поскольку ширина спектра ЧМ остается примерно постоянной для сообщений с разной шириной спектра.

Заметим, что свойства ФМ и ЧМ позволяют легко перейти от одного вида модуляции к другому. Действительно, если на вход модулятора ФМ передатчика включить частотно-зависимый линейный фильтр с коэффициентом передачи, обратно пропорциональным частоте (интегрирующее звено), то на выходе передатчика образуется ЧМ сигнал. При включении на входе модулятора ЧМ передатчика фильтра с коэффициентом передачи, пропорциональным частоте (дифференцирующее звено), образуется ФМ сигнал. Аналогично решается вопрос и в приемных устройствах, если возникает необходимость принимать на ЧМ приемник ФМ сигнал, или на ФМ приемник ЧМ сигнал. Соответствующая частотно-корректирующая цепь в этом случае включается на выходе частотного или фазового детектора.

Отмеченные особенности ЧМ и ФМ сигналов часто используют на практике для повышения качества передачи сообщений с помощью *гибридной модуляции*, когда для нижнего участка спектра модулирующего сообщения применяют ЧМ, а для верхнего ФМ. По сути дела гибридная модуляция представляет собой способ *аналогового кодирования сообщений*, основанный на использовании предсказаний таких сообщений.

4.4. МЕТОДЫ ФОРМИРОВАНИЯ СИГНАЛОВ С УГЛОВОЙ МОДУЛЯЦИЕЙ

Различают два основных метода получения сигналов с угловой модуляцией:

1) *модуляция на промежуточной частоте* с последующим умножением или преобразованием мгновенной частоты полученного модулированного колебания;

2) *модуляция на рабочей частоте* сигнала.

Первый метод широко применяется в ЧМ радиовещании, радиорелейных и космических системах связи. Второй метод часто используется в служебных связных системах малой мощности, для которых требования к характеристикам и показателям являются менее жесткими, и на первый план выдвигаются экономические соображения и простота реализации. Метод находит также применение в передатчиках, использующих такие СВЧ приборы, в которых получение угловой модуляции с большими индексами не вызывает трудностей (лавинно-пролетные диоды, диоды Ганна, отражательные клистроны, лампы бегущей волны и др.).

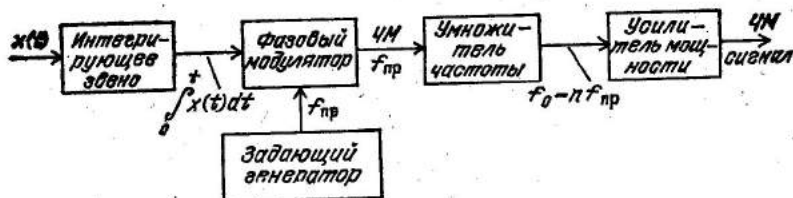


Рис. 4.4. Структурная схема получения ЧМ сигнала косвенным методом

При передаче непрерывных сообщений сигналы с фазовой модуляцией обычно не применяются по причинам, о которых говорилось выше¹. Наиболее широко при такой передаче применяются сигналы с ЧМ. Поэтому ниже основное внимание уделено рассмотрению методов получения сигналов с ЧМ.

В соответствии с рекомендациями МККР для многоканальных радиорелейных и космических линий связи с ЧМ используются промежуточные частоты $f_{пр} = 35$ МГц (для несущих $f_0 \leq 1000$ МГц) и $f_{пр} = 70$ МГц (для несущих $f_0 \geq 1000$ МГц).

Если ЧМ сигнал получают методом модуляции на промежуточной частоте с последующим умножением, то $f_0 = n f_{пр}$, $\Delta f_m = n \Delta f_{мпр}$, где n — коэффициент умножения, $\Delta f_{мпр}$ — девиация на промежуточной частоте. Если же ЧМ получают методом преобразования частоты от $f_{пр}$ до f_0 , т. е. сдвигом модулированного колебания с промежуточной частоты на частоту несущей, то девиация не изменяется.

Кроме прямых методов для получения ЧМ сигналов применяют также косвенные методы, суть которых состоит в том, что ЧМ сигнал формируется с помощью фазового модулятора, на входе которого включается интегрирующее звено. Косвенные методы позволяют использовать высокостабильные задающие генераторы и благодаря этому обеспечивают высокую стабильность несущей ЧМ сигнала. Структурная схема получения ЧМ сигнала косвенным методом на промежуточной частоте с последующим умножением приведена на рис. 4.4. Для повышения стабильности частоты сигнала в задающем генераторе применяется кварцевая стабилизация.

При умножении частоты ее относительная нестабильность не изменяется, поэтому при большом умножении абсолютная нестабильность может быть значительной. Например, если $f_{пр} = 70$ МГц, а $f_0 = 2100$ МГц, то при $n = 30$ и относительной нестабильности 10^{-5} абсолютная нестабильность составит более 20 кГц. С учетом случайного характера нестабильности это потребует расширения полосы пропускания системы по высокой частоте не менее чем на 40 кГц. Если же полосу не расширять, то нестабильность сигнала может привести к появлению рассогласования

¹ При передаче дискретных сообщений сигналы с фазовой модуляцией (вернее, с фазовой манипуляцией) обладают рядом ценных качеств и широко применяются (см. гл. 6—9).

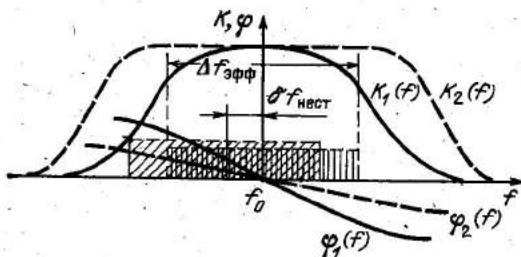


Рис. 4.5. Рассогласование между спектром сигнала и ФЧХ избирательной части системы связи

между положением спектра сигнала и частотной характеристики избирательной части общего тракта системы (рис. 4.5). Выход части спектра за пределы полосы пропускания системы и, соответственно, на нелинейный участок фазовой характеристики приводит к появлению паразитной амплитудной и фазовой модуляций. Необходимое расширение полосы пропускания для ослабления этого эффекта показано на рис. 4.5 пунктиром.

Остановимся на некоторых особенностях получения ЧМ и ФМ колебаний.

Современные системы связи характеризуются высоким качеством передачи информации. Поэтому к линейности модуляционных характеристик предъявляются весьма жесткие требования и допустимые значения коэффициента нелинейных искажений могут составлять доли процента (0,1—0,2% по 2-й и 3-й гармоникам гармонического испытательного напряжения).

При таких требованиях способы получения ЧМ, основанные на перестройке контура автогенератора за счет изменения реактивного сопротивления, вносимого в контур, не могут обеспечить больших девиаций частоты. Обычно девиация частоты в таких способах не превышает 1—1,5% от значения частоты автогенератора, а в схемах с кварцевыми генераторами составляет сотые доли процента.

Значительно более эффективными являются способы, основанные на изменении фазового сдвига в цепи обратной связи. Генераторы ЧМ колебаний, построенные по такому принципу, позволяют получать девиацию частоты, достигающую 10—30% от частоты автогенератора, при сохранении хорошей линейности модуляционной характеристики. Подобные генераторы ЧМ колебаний можно разделить на две группы:

автогенераторы с параллельными цепями обратной связи, в которых под действием модулирующего колебания (сообщения) происходит изменение фазового сдвига в цепи обратной связи;

автогенераторы RC, в которых управление фазовым сдвигом в цепи обратной связи осуществляется изменением сопротивления одного из резисторов, входящих в эту цепь.

В ЧМ автогенераторах с параллельными цепями обратной связи можно изменять частоту в очень широких пределах. Так,

на частоте $f=70$ МГц генератор подобного типа обеспечивает отклонение частоты ± 5 МГц при крутизне модуляционной характеристики 10 МГц/В и нелинейных искажениях, составляющих доли процента [3]. В [3] указывается, что имеются автогенераторы, которые при частоте 70 МГц обеспечивают отклонение частоты ± 26 МГц, а также автогенераторы, частота которых при модуляции может изменяться от 1 до 10 МГц. По сути дела здесь можно говорить уже не о модуляции, а об электронной перестройке частоты, поскольку на выходе автогенераторов формируются ЧМ колебания, которые уже нельзя отнести к полосовым сигналам.

Как правило, RC автогенераторы применяются для получения ЧМ колебаний малой мощности на частотах не выше $15-20$ кГц. Такие колебания используются как ЧМ поднесущие в некоторых многоканальных системах с частотным уплотнением каналов (например, в телеметрических системах).

Как уже отмечалось, в системах передачи непрерывных сообщений сигналы с ФМ обычно не применяются. Несмотря на это, методы получения ФМ колебаний представляют большой практический интерес и их разработке уделяется значительное внимание. Объясняется это тем, что формирование ФМ колебаний часто необходимо при выполнении ряда важных операций как в передающей, так в приемной частях системы связи (при получении ЧМ сигналов косвенным методом, автоматической подстройке частоты и фазы сигнала и т. д.).

Известно несколько способов получения ФМ колебаний:

- 1) пропускание гармонического колебания через контур с изменяющейся фазовой характеристикой;
- 2) использование фазовращательных цепей с переменным сдвигом фазы;
- 3) преобразование АМ колебаний в ФМ.

Рассмотрим кратко особенности этих способов.

При изменении реактивного сопротивления избирательного контура в соответствии с модулирующим колебанием его частот-



Рис. 4.6. Схема получения ФМ сигнала изменением реактивного импеданса усиленной ступени

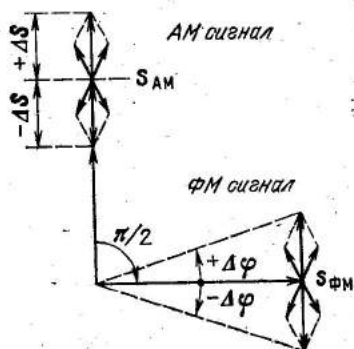


Рис. 4.7. Диаграмма, поясняющая получение ФМ сигнала из АМ

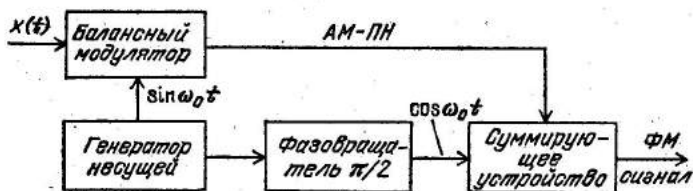


Рис. 4.8. Структурная схема модулятора с использованием преобразования АМ сигнала в ФМ

ная и фазовая характеристики также изменяются. Осуществляя указанные изменения в усилительном каскаде, расположенном после высокостабильного задающего генератора, можно получить ФМ сигнал с высокой стабильностью несущей частоты (рис. 4.6). В качестве управляемого сопротивления широко применяют варикапы.

Получение ФМ колебаний с помощью фазовращателей можно рассматривать как разновидность предыдущего способа. Отличие состоит в том, что изменения фазы достигаются изменением импеданса одного из элементов фазовращателя, через который пропускается гармоническое колебание.

Получение ФМ колебаний из АМ основано на том, что при повороте фазы несущего колебания на $\pi/2$ АМ колебание преобразуется в ФМ (рис. 4.7). Один из вариантов модулятора ФМ, основанного на преобразовании АМ в ФМ с помощью балансного модулятора и фазовращателя, показан на рис. 4.8.

Для всех указанных способов характерно появление сопутствующей паразитной модуляции и нелинейных искажений. Чтобы уменьшить эти нежелательные эффекты, приходится ограничиваться получением ФМ колебаний с малыми индексами ($\Delta f = 0,2-0,3$), а затем применять умножение частоты для повышения индекса до нужной величины.

Помимо рассмотренных способов, ФМ колебания можно получить, сформировав сначала колебания с помощью фазоимпульсной модуляции (ФИМ), а затем пропустив полученную последовательность импульсов через узкополосный резонансный усилитель, настроенный на частоту повторения импульсов F_n или на одну из ее гармоник. При таком способе можно получить довольно значительный индекс модуляции ($\Delta f \approx 1,5-2$), и, следовательно, не требуется столь большое умножение частоты, как в предыдущих способах. Этот способ находит применение для УКВ ЧМ радиовещания, а также для передачи звукового сопровождения в телевидении. Детальное изложение вопросов реализации сигналов с угловой модуляцией содержится в [3].

4.5. ИМПУЛЬСНЫЕ МОДУЛИРОВАННЫЕ КОЛЕБАНИЯ

Импульсные модулированные колебания получаются при модуляции периодической последовательности импульсов (импульсной поднесущей). Импульсная модуляция широко применяется при

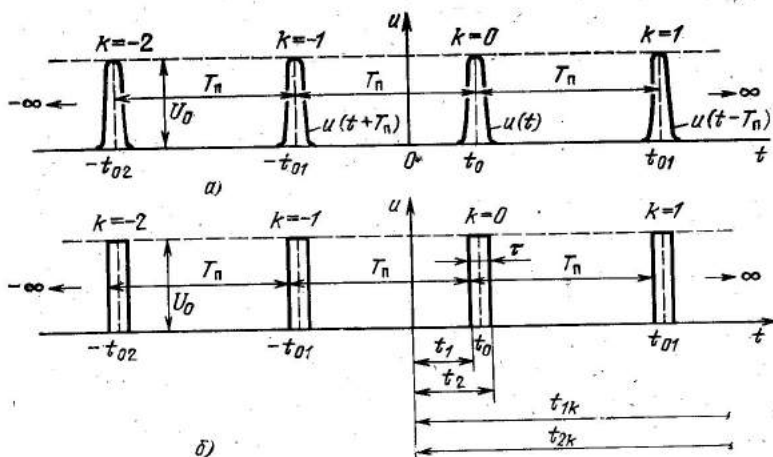


Рис. 4.9. Периодическая последовательность импульсов

временном уплотнении каналов, а также как промежуточная операция при различных преобразованиях сообщений.

4.5.1. Импульсная поднесущая. Любую периодическую последовательность импульсов (см. рис. 4.9,а) можно записать так:

$$u_0(t) = \sum_{k=-\infty}^{\infty} U(t-t_{0k}) = U_0 \sum_{k=-\infty}^{\infty} f(t-t_{0k}), \quad (4.28)$$

где U_0 — амплитуда импульсов; $t_{0k} = t_0 + kT_n$ — тактовые точки; T_n — период повторения импульсов; $f(t-t_{0k})$ — огибающая импульса с единичной амплитудой.

Выражение (4.28) можно разложить в ряд Фурье (см. 2.57), т. е. представить в виде

$$u_0(t) = A_0 + \sum_{k=1}^{\infty} A_k \cos k \Omega_n (t-t_0), \quad (4.29)$$

где A_0 — постоянная составляющая; $\Omega_n = 2\pi F_n = 2\pi/T_n$ — частота повторения импульсов; A_k — амплитуда гармоники с частотой $k\Omega_n$.

Подчеркнем, что выражение (4.28) и (4.29) полностью эквивалентны. Поэтому любые изменения параметров в одном из выражений соответствуют точно таким же изменениям этих параметров в другом выражении.

Значения A_0 и A_k зависят от конкретного вида огибающей $f(t)$. Наибольшее распространение в теории и приложениях получили периодические последовательности прямоугольных импульсов (рис. 4.9,б). Каждый такой импульс можно описать так:

$$f(t-t_{0k}) = \begin{cases} 1 & t_{0k} - \tau/2 \leq t \leq t_{0k} + \tau/2 \\ 0 & \text{для других } t. \end{cases} \quad (4.30)$$

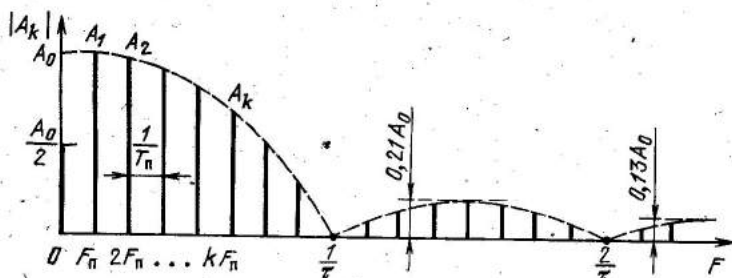


Рис. 4.10. Амплитудный спектр периодической последовательности прямоугольных импульсов

Спектральная функция такого импульса

$$g(\Omega) = g(j\Omega) = \frac{2}{\Omega} \sin\left(\frac{\Omega\tau}{2}\right) e^{-j\Omega t_0} \quad (4.31)$$

Значения A_0 и A_k с учетом (2.56), (2.57) и (4.31) после несложных преобразований принимают вид:

$$\begin{aligned} A_0 &= U_0 \tau / T_n = U_0 F_n \tau = U_0 / Q, \\ A_k &= \frac{2U_0}{k\pi} \sin\left(\frac{k\Omega_n \tau}{2}\right) = \frac{2U_0}{Q} \frac{\sin(k\Omega_n \tau/2)}{k\Omega_n \tau/2}, \end{aligned} \quad (4.32)$$

где $Q = T_n / \tau$ — скважность импульсов.

Выражение (4.29) с учетом (4.32) можно записать в виде

$$u_0(t) = \frac{U_0 \tau}{T_n} + \sum_{k=1}^{\infty} \frac{2U_0}{k\pi} \sin\left(\frac{k\Omega_n \tau}{2}\right) \cos(k\Omega_n(t-t_0)). \quad (4.33)$$

Амплитудный спектр периодической последовательности прямоугольных импульсов, определяемый выражением (4.32), показан на рис. 4.10. Нетрудно убедиться, что число гармоник в первом (основном) лепестке спектра равно скважности импульсов ($k_1 = Q$), а ширина спектра последовательности, определяемая этим лепестком, зависит только от длительности импульса.

$$\Delta f_{\text{имп}} = 1/\tau. \quad (4.34)$$

Основные виды импульсной модуляции. Возможны следующие виды модуляции периодической последовательности импульсов (см. рис. 4.11):

1. **Амплитудно-импульсная модуляция (АИМ).** При АИМ амплитуда импульсов изменяется в соответствии со значениями сообщения, а остальные параметры последовательности остаются неизменными, т. е.

$$U_0(t) = U_0 [1 + M_a x(t)]; \quad \tau = \text{const}; \quad T_n = \text{const}. \quad (4.35)$$

2. **Широтно-импульсная модуляция (ШИМ).** При ШИМ модулируется ширина (длительность) импульсов, а остальные параметры остаются неизменными. Различают одностороннюю и двустороннюю ШИМ. На практике наибольшее применение находит односторонняя ШИМ (иногда ее обозначают ОШИМ). При этом

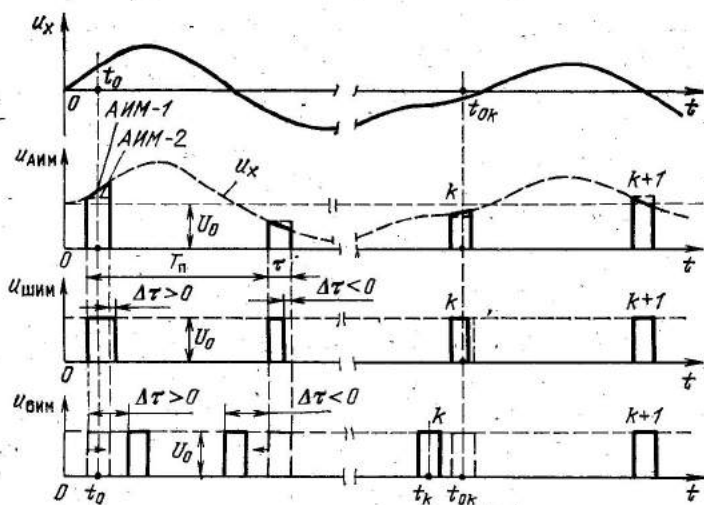


Рис. 4.11. Различные виды импульсной модуляции

положение одного из фронтов импульса (например, переднего) остается неизменным относительно тактовой точки, а положение другого фронта (заднего) меняется в соответствии с модулирующим сообщением.

Для односторонней ШИМ имеем:

$$\tau(t) = \tau + \Delta\tau_m x(t); \quad U = U_0 = \text{const}; \quad T_{\text{п}} = \text{const}. \quad (4.36)$$

Максимальный временной сдвиг модулируемого фронта $\Delta\tau_m$ относительно его начального значения называется *девиацией фронта импульса*. При ШИМ всегда должно выполняться условие $\Delta\tau_m/\tau < 1$, так как минимальная длительность импульса при модуляции должна быть конечной ($\tau_{\text{min}} = \tau - \Delta\tau_m > 0$).

3. *Время-импульсная модуляция (ВИМ)*. При ВИМ модулируется положение импульсов t_k относительно тактовых точек t_{0k} немодулированной последовательности. Длительность импульсов и их амплитуда при модуляции сохраняются неизменными. В соответствии с этим

$$t_k = t_{0k} + \Delta\tau_m x(t); \quad \tau = \text{const}; \quad U = U_0 = \text{const}. \quad (4.37)$$

Максимальный временной сдвиг импульса $\Delta\tau_m$ относительно тактовой точки называется *девиацией импульса*. При ВИМ (в отличие от ШИМ) возможен выбор таких значений $\Delta\tau_m$, при которых $\Delta\tau_m/\tau > 1$.

Изменение положений импульса при ВИМ относительно тактовой точки можно рассматривать как изменение «фазы» импульса. В этом легко убедиться, перейдя к безразмерному времени $\varphi = 2\pi t/T_{\text{п}}$. Тогда величина $\Delta\tau_m$ будет соответствовать следующему отклонению (девиация) фазы:

$$\Delta\varphi_m = 2\pi \Delta\tau_m/T_{\text{п}} = \Omega_{\text{п}} \Delta\tau_m.$$

С учетом этого выражения вместо (4.37) можно рассматривать эквивалентное ему выражение в безразмерном времени

$$\varphi_k = \varphi_{0k} + \Delta\varphi_m x(t),$$

где $\varphi_k = 2\pi t_k/T_n$; $\varphi_{0k} = 2\pi t_{0k}/T_n$.

Если величина $\Delta\tau_m$ (или $\Delta\varphi_m$) при ВИМ постоянна и не зависит от ширины спектра сообщения, то модуляцию называют фазо-импульсной (ФИМ). При наличии такой зависимости модуляцию называют частотно-импульсной (ЧИМ). Разделение ВИМ на ФИМ и ЧИМ аналогично разделению угловой модуляции на ФМ и ЧМ.

При анализе импульсных видов модуляции иногда вводят понятие о модуляции 1-го и 2-го рода. К 1-му роду относят те виды модуляции, у которых значение модулируемого параметра в рассматриваемый момент времени пропорционально мгновенному значению сообщения в этот же момент времени. Если значение модулируемого параметра пропорционально мгновенному значению сообщения в тактовых точках, то модуляцию относят ко 2-му роду.

Введение такой классификации объясняется тем, что при модуляции 2-го рода сообщение $x(t)$ на интервале длительности импульса принимается равным его значению в тактовой точке $x(t_{0k})$, т. е. не учитывается его изменение за этот промежуток времени. При реализации импульсных видов модуляции, в зависимости от конкретного способа, можно получить модуляцию либо 1-го, либо 2-го рода.

Соотношения (4.35—4.37) соответствуют модуляции 1-го рода. Для определенности дальнейшее рассмотрение в этом параграфе будем проводить применительно к этому роду модуляции. При большой скважности импульсов ($Q \gg 1$) различие между модуляцией 1-го и 2-го рода незначительно и его обычно не учитывают.

4.6. АНАЛИТИЧЕСКАЯ ЗАПИСЬ ИМПУЛЬСНЫХ ВИДОВ МОДУЛЯЦИИ И ИХ СПЕКТРАЛЬНЫЙ СОСТАВ

Амплитудно-импульсная модуляция. Аналитическую запись АИМ можно получить, подставив в (4.28) вместо U_0 значение $U(t)$ из (4.35):

$$u_{\text{АИМ}}(t) = \sum_{k=-\infty}^{\infty} U_0 [1 + M_a x(t)] f(t - t_{0k}). \quad (4.38)$$

Выражение в квадратных скобках не зависит от k , поэтому

$$u_{\text{АИМ}}(t) = U_0 [1 + M_a x(t)] \sum_{k=-\infty}^{\infty} f(t - t_{0k}). \quad (4.38a)$$

С учетом (4.28) и эквивалентного ему выражения (4.29) соотношение (4.38a) можно представить так:

$$u_{\text{АИМ}}(t) = A_0 + A_0 M_a x(t) + \sum_{k=1}^{\infty} A_k [1 + M_a x(t)] \cos k \Omega_n (t - t_0). \quad (4.39)$$

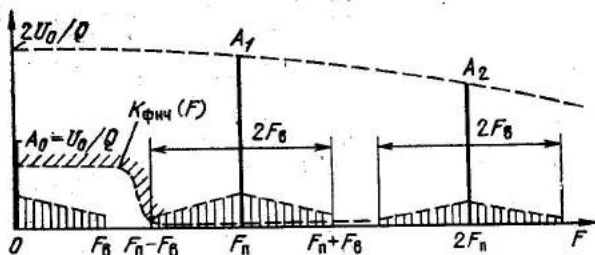


Рис. 4.12. Часть спектра АИМ

Это выражение определяет структуру АИМ, которая состоит из постоянной составляющей A_0 , составляющей $A_0 M_a x(t)$, пропорциональной сообщению $x(t)$, и бесконечного числа гармоник частоты повторения kF_n , каждая из которых модулирована по амплитуде сообщением $x(t)$.

Нетрудно видеть, что спектр АИМ имеет вид, показанный на рис. 4.12. Чтобы соседние участки спектра АИМ не перекрывались, необходимо выполнить условие $F_n \geq 2F_в$, предельное значение которого соответствует теореме Котельникова.

Заметим, что, применив фильтр нижних частот с характеристикой $K_{ФНЧ}(F)$ и полосой пропускания, определяемой шириной спектра сообщения, можно выделить это сообщение из АИМ последовательности. Этот вопрос более детально рассматривается в п. 5.6.

Широтно-импульсная модуляция. В соответствии с определением ШИМ ее аналитическая запись имеет вид

$$u_{\text{ШИМ}}(t) = \sum_{k=-\infty}^{\infty} U_0 f(t-t_k) = \sum_{k=-\infty}^{\infty} U_0 f_k(t), \quad (4.40)$$

где $f_k(t)$ — огибающая импульса с номером k ; t_k — момент времени, соответствующий середине этого импульса. Для прямоугольных импульсов

$$f_k(t) = \begin{cases} 1 & t_{1k} \leq t \leq t_{2k}, \\ 0 & \text{для других } t. \end{cases} \quad (4.41)$$

Здесь t_{1k} и t_{2k} — моменты времени, соответствующие началу и окончанию импульса с номером k . Значения моментов t_{1k} , t_{2k} и t_k зависят от конкретного вида ШИМ. Например, для односторонней ШИМ 1-го рода (ШИМ-1) с модуляцией заднего фронта импульса имеем:

$$\begin{aligned} t_{1k} &= t_{0k} - \tau/2 = t_0 + kT_n - \tau/2, \\ t_{2k} &= t_{0k} + \tau/2 + \Delta\tau_m x(t_{2k}), \\ t_k &= \frac{1}{2}(t_{1k} + t_{2k}) = t_{0k} + \frac{1}{2}\Delta\tau_m x(t_{2k}). \end{aligned} \quad (4.42)$$

Длительность импульса ШИМ-1 в момент времени $t = t_{2k}$ равна

$$\tau(t)|_{t=t_{2k}} = t_{2k} - t_{1k} = \tau + \Delta\tau_m x(t)|_{t=t_{2k}}.$$

Это соотношение справедливо для любого момента времени. Поэтому при односторонней ШИМ-1

$$\tau(t) = \tau + \Delta\tau_m x(t), t_k(t) = t_{0k} + 0,5 \Delta\tau_m x(t). \quad (4.43)$$

Нетрудно убедиться, что в отличие от этого случая при двусторонней симметричной ШИМ середина импульса всегда совпадает с положением тактовой точки $t_k(t) = t_{0k}$.

Спектральный состав односторонней ШИМ-1 можно определить, подставив в (4.33) вместо τ и t_0 соответствующие им значения $\tau(t)$ и $t_0(t)$ из (4.43). При этом $t_0(t) = t_k(t)$ для $k=0$, т. е. $t_0(t) = t_0 + 0,5 \Delta\tau_m x(t)$. После несложных преобразований получим

$$u_{\text{ШИМ}}(t) = \frac{U_0}{Q} + \frac{U_0}{Q} \frac{\Delta\tau_m}{\tau} x(t) + \sum_{k=1}^{\infty} B_k(t) \cos [k \Omega_{\Pi} (t - t_0) - 0,5 k \Delta\varphi_{\text{ШИМ}} x(t)], \quad (4.44)$$

где $B_k(t)$ — огибающая гармоники с номером k ;

$$B_k(t) = \frac{2 U_0}{\pi k} \sin \left\{ \frac{k \Omega_{\Pi} \tau}{2} \left[1 + \frac{\Delta\tau_m}{\tau} x(t) \right] \right\}, \quad (4.45)$$

$\Delta\varphi_{\text{ШИМ}} = \Delta\tau_m \Omega_{\Pi}$ — индекс модуляции при ШИМ.

Отличие спектральной структуры ШИМ от АИМ состоит в том, что гармоники частоты повторения модулированы более сложным образом: они изменяются одновременно и по амплитуде и по фазе. Так как $\Delta\varphi_{\text{ШИМ}} \ll 1$, то влияние фазовой модуляции сказывается незначительно по сравнению с амплитудной. (При двусторонней симметричной ШИМ фазовая модуляция гармоник отсутствует.) Нетрудно показать, что для всех гармоник с номерами $k \leq 0,1 Q$ спектральная структура АИМ и ШИМ практически одинакова.

Время-импульсная модуляция. Аналитическое выражение для ВИМ с прямоугольными импульсами имеет такой же вид, как и для ШИМ (см. (4.40), (4.41)). Различие состоит в том, что при ВИМ-1 значения t_{1k} и t_{2k} определяются так:

$$t_{1k} = t_{0k} - \tau/2 + \Delta\tau_m x(t_{1k}), t_{2k} = t_{0k} + \tau/2 + \Delta\tau_m x(t_{2k}).$$

Для любого произвольного момента времени эти соотношения принимают вид

$$t_{1k}(t) = t_{0k} - \tau/2 + \Delta\tau_m x(t), t_{2k}(t) = t_{0k} + \tau/2 + \Delta\tau_m x(t + \tau). \quad (4.46)$$

Заметим, что эти соотношения соответствуют случаю, когда ВИМ реализуется в виде фазо-импульсной модуляции ($\Delta\tau_m = \text{const}$).

Для определения спектрального состава ВИМ надо подставить в выражение (4.33) вместо τ и $t_0(t)$ соответствующие им значения $\tau(t)$ и $t_0(t)$, определяемые так:

$$\tau(t) = t_{2k}(t) - t_{1k}(t), t_0(t) = t_k(t)|_{k=0} = 0,5 [t_{2k}(t) + t_{1k}(t)],$$

и провести преобразования с учетом того, что при ВИМ длительность импульса и его амплитуда неизменны.

Приняв во внимание, что за время τ сообщение $x(t)$ изменяется незначительно, можно записать

$$\tau(t) = \tau + \Delta\tau_m \tau \left[\frac{x(t+\tau) - x(t)}{\tau} \right] \approx \tau [1 + \Delta\tau_m x'(t)], \quad (4.47)$$

$$t_0(t) \approx t_0 + 2 \Delta\tau_m \left[\frac{x(t+\tau) + x(t)}{2} \right] \approx t_0 + 2 \Delta\tau_m x(t + \tau/2).$$

Тогда (4.33) в случае ВИМ примет вид:

$$u_{\text{ВИМ}}(t) = \frac{U_0}{Q} + \frac{U_0}{Q} \Delta\tau_m x'(t) + \sum_{k=1}^{\infty} B_k(t) \cos \left[k \Omega_{\Pi} (t - t_0) - 2k \Delta\Phi_{\text{ВИМ}} x \left(t + \frac{\tau}{2} \right) \right], \quad (4.48)$$

где $B_k(t)$ — огибающая гармоники с номером k ;

$$B_k(t) = \frac{2U_0}{k\pi} \sin \left\{ \frac{k\Omega_{\Pi}\tau}{2} \left[1 + \Delta\tau_m x'(t) \right] \right\}, \quad (4.49)$$

$\Delta\Phi_{\text{ВИМ}} = \Delta\tau_m \Omega_{\Pi}$ — индекс модуляции при ВИМ.

Выражение (4.48) определяет спектральный состав ВИМ-1: постоянную составляющую, составляющую, пропорциональную производной сообщения и бесконечное число гармоник частоты повторения, модулированных одновременно по амплитуде и фазе. Можно показать, что при ВИМ-1 влияние амплитудной модуляции гармоник проявляется незначительно по сравнению с фазовой, т. е. информация в основном содержится в фазе гармоник. Можно также показать, что при ВИМ-2 спектр содержит постоянную составляющую и бесконечное число гармоник частоты повторения, модулированных сообщением только по фазе.

Характерной особенностью импульсных видов модуляции является то, что все они широкополосны — их спектр намного шире спектра сообщения и простирается от постоянной составляющей до частоты¹

$$F_{\text{в имп}} \approx 1/\tau. \quad (4.50)$$

4.7. МЕТОДЫ ФОРМИРОВАНИЯ ИМПУЛЬСНЫХ МОДУЛИРОВАННЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Последовательность АИМ можно рассматривать как результат дискретизации непрерывных сообщений по времени. Осуществить эту операцию можно многими способами. В настоящее время наиболее распространены способы, основанные на применении *ключевых устройств* (ключей). На один вход такого устройства (рис. 4.13) подается периодическая последовательность коротких

¹ Доля мощности, заключенная в составляющих с частотами выше $F_{\text{в}}$, настолько мала, что эти составляющие можно не учитывать.

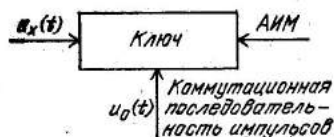


Рис. 4.13. Принцип ключевого формирования АИМ последовательностей

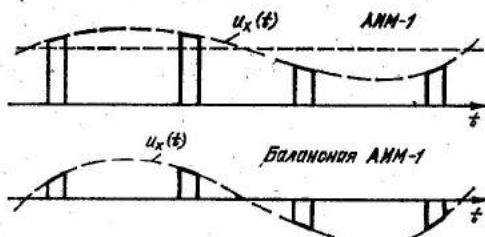


Рис. 4.14. Однополярная и балансная АИМ последовательности

прямоугольных импульсов (*коммутационная последовательность*) $u_0(t)$, а на другой — непрерывное сообщение (*модулирующая функция*) $u_x(t)$. Каждый коммутационный импульс открывает ключ на короткое время, в течение которого образуется выходной импульс с амплитудой, пропорциональной мгновенному значению сообщения. В зависимости от схемы ключа и режима его работы можно получить на выходе либо обычную однополярную последовательность АИМ, либо двухполярную последовательность, называемую *балансной АИМ* (рис. 4.14). Несложно показать, что в спектре балансной АИМ отсутствуют все гармоники частоты повторения, т. е. этот вид модуляции имеет аналогию с АМ—ПН.

Последовательность ШИМ можно получать различными способами. Наиболее часто для этой цели применяют импульсные устройства, имеющие одно или два устойчивых состояния равнове-

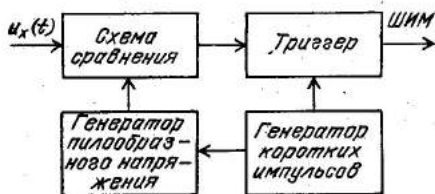


Рис. 4.15. Структурная схема модулятора ШИМ

схему сравнения модулирующее напряжение $u_x(t)$, и нарастающее пилообразное напряжение становятся равными, на выходе этой схемы формируется короткий импульс, возвращающий триггер в первоначальное состояние. В результате напряжение, снимаемое с нагрузки одного из плеч триггера, представляет собой последовательность импульсов с односторонней ШИМ.

Последовательность ФИМ обычно получают дифференцированием последовательности ШИМ с последующим формированием коротких импульсов, соответствующих положению модулируемого фронта импульсов ШИМ.

4.8. МЕТОДЫ УПЛОТНЕНИЯ КАНАЛОВ

Для организации многоканальной передачи по одной линии связи необходимы операция уплотнения каналов в передающей части системы связи и операция разделения каналов в ее приемной части. Ниже рассматриваются только вопросы уплотнения каналов, вопросы разделения каналов рассмотрены в гл. 5.

Принцип уплотнения каналов. Суть операции уплотнения каналов поясним с помощью рис. 4.16. Сообщения $x_i(t)$ от N независимых источников необходимо преобразовать в единое колебание — многоканальное (групповое) сообщение $u_{\text{мн}}(t)$. Для этого нужно иметь N вспомогательных колебаний $\varphi_i(t)$ $i=1, 2, \dots, N$. Обычно стремятся, чтобы эти колебания были ортогональны. Наиболее часто в качестве таких колебаний используют гармонические колебания или периодические импульсные последовательности (гармонические и импульсные поднесущие). Они генерируются в специальном устройстве, называемом генератором вспомогательных колебаний.

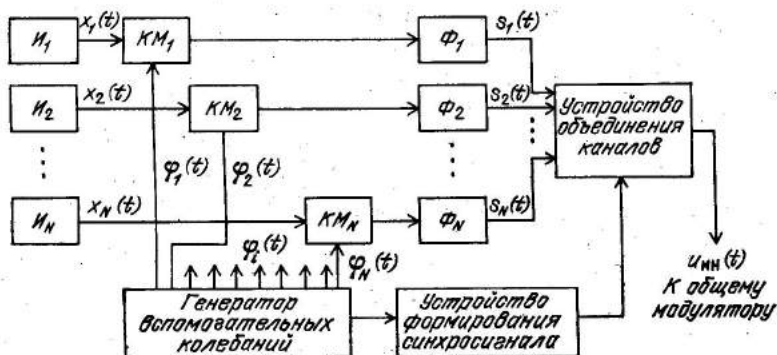


Рис. 4.16. Общая схема уплотнения каналов

Изменяя с помощью канальных модуляторов (КМ) параметры поднесущих в соответствии с передаваемыми сообщениями, можно получить N канальных сигналов $s_i(t)$. Поскольку такая операция нелинейна и приводит к появлению побочных спектральных составляющих, лежащих вне полосы, занимаемой спектром модулированной поднесущей, в схеме обычно содержатся фильтры Φ_i , пропускающие спектр соответствующей i -й модулированной поднесущей и устраняющие побочные продукты модуляции. В результате объединения канальных сигналов образуется групповой сигнал — многоканальное сообщение $u_{\text{мн}}(t)$, которое подается на вход общего модулятора многоканальной системы связи¹.

¹ В общем случае, когда в системе связи применяется кодирование, сообщение $u_{\text{мн}}(t)$ подается на вход кодирующего устройства, а на вход модулятора поступает уже результат кодирования.

Возможны два метода объединения: *линейное* и *нелинейное* уплотнение каналов. Линейное уплотнение находит наиболее широкое применение. Оно представляет собой простое суммирование канальных сигналов

$$u_{\text{мп}}(t) = \sum_{i=1}^N s_i(t), \quad (4.51)$$

выполняемое с помощью сумматора¹.

Вспомогательные колебания $\varphi_i(t)$, используемые при уплотнении каналов, должны обладать свойством разделимости, которое обеспечивает правильность разделения принятого многоканального сообщения на отдельные канальные сигналы с последующим извлечением из них информации для каждого получателя.

Из теории функционального анализа и основанной на ней теории разделения сигналов и сообщений известно, что необходимым и достаточным условием разделимости функций $\varphi_i(t)$ является условие их линейной независимости. Это означает, что ни одну из используемых функций нельзя получить с помощью линейных комбинаций других функций этого класса [1].

Чтобы функции $\varphi_i(t)$ ($i=1, 2, \dots, N$) были линейно-независимыми, необходимо и достаточно иметь определитель матрицы $\|a_{ij}\|$, элементы которой определены соотношением

$$a_{ij} = \int_0^T \varphi_i(t) \varphi_j(t) dt, \quad (4.52)$$

отличающимся от нуля. Это утверждение называется *теоремой Грама* [1].

Частным, но очень важным случаем линейно-независимых функций являются ортогональные функции. При уплотнении каналов в качестве вспомогательных, как правило, применяют ортогональные колебания, а параметры модуляции канальных сигналов выбирают так, чтобы эти сигналы оставались также ортогональными. В реальных условиях выполнить это требование удастся не всегда, т. е.

$$\frac{1}{T} \int_0^T s_k(t) s_i(t) dt = \begin{cases} P_s, & i=k \\ P_e \neq 0, & i \neq k, \end{cases} \quad (4.53)$$

где P_s — мощность поднесущей с номером i ; P_e — взаимная мощность k -0 и i -й поднесущих.

Для лучшей разделимости каналов в этом случае желательно обеспечивать выполнение условия $P_e \ll P_s$. Это условие имеет следующий физический смысл: доля мощности, проникающей на выход канала из другого канала, должна быть много меньше мощности поднесущей.

¹ *Нелинейное* (мажоритарное) уплотнение, основанное на использовании функций Уолша, иногда применяется при передаче цифровой информации (например, в телеметрии). Широкого распространения нелинейное уплотнение пока не получило [4, 5].

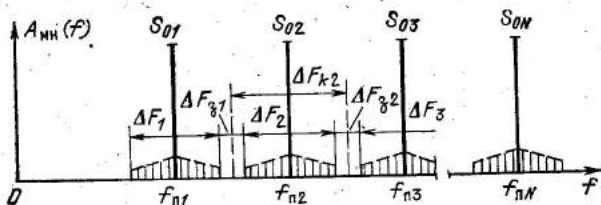


Рис. 4.17. Спектр многоканального сообщения при частотном уплотнении каналов

Полученное в результате уплотнения многоканальное сообщение $u_{mn}(t)$ модулирует сигнал-переносчик. Заметим, что в общем случае, при формировании многоканального сообщения в устройстве уплотнения предусматривается создание специального служебного колебания, предназначенного для синхронизации передающей и приемной частей системы связи. Синхронизация при некоторых видах уплотнения принципиально необходима.

При уплотнении каналов наиболее широко применяются ортогональные колебания следующих видов: гармонические, импульсные и кодовые последовательности. Их использование позволяет получить три основных вида уплотнения: *частотное, временное и кодовое* (уплотнение по форме).

Частотное уплотнение каналов. Такое уплотнение основано на принципе частотного преобразования спектров сообщений отдельных источников на передающей стороне системы связи. Для этого используется набор гармонических поднесущих $s_{oi}(t)$ с разными частотами $f_{n1}, f_{n2}, \dots, f_{nN}$. Модулируя (или манипулируя) поднесущие, можно получить N канальных сигналов $s_i(t)$, каждый из которых занимает полосу частот ΔF_i , зависящую от ширины спектра исходных сообщений $x_i(t)$ и вида модуляции. Чтобы уменьшить взаимное влияние соседних каналов и облегчить их разделение, между каналами вводят *защитные частотные промежутки* (полосы) ΔF_{zi} . Поэтому полная полоса частот, занимаемая каждым каналом

$$\Delta F_{ki} = \Delta F_i + \Delta F_{zi} = \Delta F_i (1 + \Delta F_{zi}/\Delta F_i) = \gamma_{zi} \Delta F_i, \quad (4.54)$$

где γ_{zi} — защитный коэффициент полосы (обычно выбирают $\gamma_{zi} = 1,2-1,3$).

Многоканальное сообщение при частотном уплотнении образуется линейным сложением канальных сигналов, а его спектр определяется суммой спектров этих сигналов (см. рис. 4.17). Из рассмотрения этого рисунка следует, что верхняя граничная частота многоканального сообщения равна

$$F_{в.мн} = f_{nN} + \Delta F_N / 2 \approx f_{nN}. \quad (4.55)$$

Если число каналов и полосы частот, занимаемые каждым из них, известны, то

$$F_{в.мн} \approx f_{nN} = f_{n1} + \sum_{i=1}^N \Delta F_{ki}. \quad (4.56)$$

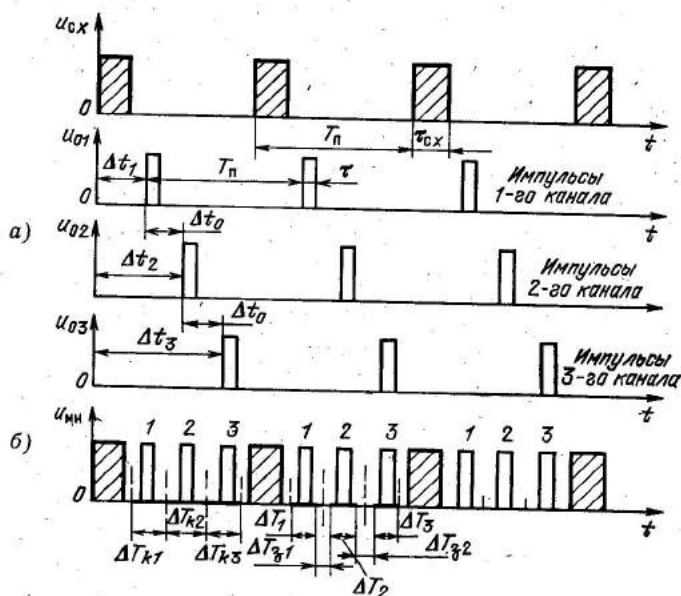


Рис. 4.18. Принципы формирования многоканального сообщения при временном уплотнении каналов

Значение частоты нижней поднесущей обычно выбирается не менее $(10-15)\Delta F_1$, где ΔF_1 — ширина спектра модулированной нижней поднесущей. При таком выборе частоты выделение и демодуляция нижней поднесущей в приемной части системы связи не встречает затруднений.

Временное уплотнение каналов. Временное уплотнение каналов основано на дискретизации непрерывных сообщений по времени. При таком уплотнении используется набор импульсных поднесущих, не перекрывающихся во времени (рис. 4.18,а). Каждая поднесущая модулируется своим непрерывным сообщением в соответствующем канальном модуляторе. Естественно, что частота повторения импульсов в этих поднесущих должна удовлетворять $F_n \geq 2F_{вi}$ для всех каналов ($i=1, 2, \dots$).

Многоканальное сообщение образуется в результате линейного объединения (суммирования) модулированных импульсов поднесущих (рис. 4.18,б)¹. Очевидно, что ширина спектра многоканального сообщения $F_{в.мн}$ однозначно определяется длительностью импульсов поднесущих и приблизительно равна величине $F_{в.мн} \approx 1/\tau$ (см. (4.50)).

Временное уплотнение осуществляется в синхронном режиме. Для этого в устройстве уплотнения формируется периодическая последовательность *синхроимпульсов* с периодом $T_n = 1/F_n$. (Для

¹ Для простоты на рис. 4.18 показано объединение в трехканальной системе.

наглядности эти импульсы на рис. 4.18 заштрихованы.) В каждом цикле (периоде T_n) сначала передается синхроимпульс длительностью $\tau_{сх}$, а затем поочередно импульсы всех N каналов. В зависимости от вида импульсной модуляции каждому каналу в интервале T_n отводится определенное время ΔT_{hi} , которое включает интервал ΔT_i , занимаемый каналом с учетом модуляции и *временной защитный интервал* ΔT_{zi} , вводимый для уменьшения взаимного влияния соседних каналов и облегчения разделения каналов в приемной части системы связи.

Итак,

$$\Delta T_{hi} = \Delta T_i + \Delta T_{zi} = \Delta T_i (1 + \Delta T_{zi}/\Delta T_i) = \gamma_{zi} \Delta T_i, \quad (4.57)$$

где γ_{zi} — защитный коэффициент канального промежутка. (Обычно γ_{zi} выбирают в пределах 1,5—2,0.) Опираясь на результаты п. 4.5, легко убедиться, что при АИМ $\Delta T_i = \tau$, при ШИМ $\Delta T_i = \tau + \Delta \tau_m$ и при ВИМ $\Delta T_i = \tau + 2\Delta \tau_m$. Если $\Delta T_{hi} = \Delta T_h = \text{const}$, то число каналов, которое может быть получено при временном уплотнении $N_{ВРК}$ равно

$$N_{ВРК} = (T_n - \tau_{сх})/\Delta T_h. \quad (4.58)$$

Нетрудно показать, что при временном уплотнении $N_{ВРК}$ пропорционально скважности импульсов Q . Так как выбор T_n определяется максимальной шириной спектра передаваемых сообщений $F_{в\max}$ и не может быть произвольным ($T_n \leq 1/2F_{в\max}$), то увеличить число каналов можно лишь укорачивая длительность импульсов, т. е. расширяя полосу частот, занимаемую многоканальным сообщением.

О других видах уплотнения. Частотное и временное уплотнение — основные виды уплотнения в современных системах связи. Эти виды уплотнения широко используются при многоканальной передаче как непрерывных, так и дискретных сообщений. В некоторых случаях (например, в телеметрии) применяют *комбинированное уплотнение*, при котором часть поднесущих, используемых для частотного уплотнения, в свою очередь уплотняют по времени. Это возможно потому, что в телеметрии ряд параметров изменяется очень медленно и ширина спектра многоканального сообщения при их временном уплотнении оказывается сравнительно узкой. Поэтому модуляция гармонической поднесущей таким спектром не вызывает затруднений.

Развитие цифровых методов передачи информации привело к появлению нового вида уплотнения, основанного на использовании функций Уолша, псевдослучайных импульсных последовательностей и других дискретных поднесущих, обладающих свойствами ортогональности. Такой вид уплотнения называется кодовым (уплотнение по форме). Характерной особенностью систем с кодовым уплотнением является то, что в них используются сложные сигналы. Известно, что такие сигналы позволяют получить ряд очень ценных свойств систем (высокую помехоустойчивость, скрытность работы, возможность одновременной работы многих

систем в одном и том же диапазоне частот и др.). Системы с кодовым уплотнением часто называют асинхронно-адресными, так как в них реализуется адресный способ передачи информации. Суть этого способа состоит в том, что каждому источнику (или абоненту) выделяется сигнал, форма которого является его отличительным признаком (адресом). Приемник, настроенный на сигнал определенной формы (на определенный адрес), не будет принимать сигналы других форм.

Кодовые методы уплотнения применяются в системах управления, телеметрии, в разветвленных и спутниковых системах связи и др.

Подробнее с упомянутыми видами уплотнения и их применением можно ознакомиться в работах [4, 5, 6].

4.9. СИГНАЛЫ В СИСТЕМАХ МНОГОКАНАЛЬНОЙ СВЯЗИ

При многоканальной связи сигналы формируются в два этапа: получение многоканального сообщения (уплотнение каналов); модуляция сигнала-переносчика многоканальным сообщением¹.

В результате образуются сигналы с *двуступенчатой модуляцией*. В сокращенном обозначении таких сигналов и соответствующих им систем сначала дается указание о виде модуляции в первой ступени, а затем — во второй. Например: ОБП—ЧМ, АМ—ЧМ, ЧМ—ЧМ — сигналы и системы с частотным уплотнением и частотной модуляцией; АИМ—ЧМ, ШИМ—ЧМ, ВИМ—ЧМ — сигналы и системы с временным уплотнением и частотной модуляцией. Во второй ступени возможна и амплитудная модуляция. Однако она применяется редко и в основном отдают предпочтение частотной модуляции, как более помехоустойчивой и менее критичной к неидеальности характеристик систем.

Остановимся кратко на некоторых особенностях сигналов в многоканальных системах связи.

Сигналы при частотном уплотнении. Нормированное многоканальное сообщение при частотном уплотнении можно записать так

$$x_{\text{мн}}(t) = u_{\text{мн}}(t)/U_{\text{max}} = \sum_{i=1}^N s_{\text{ки}}(t)/U_{\text{max}}, \quad (4.59)$$

где $u_{\text{мн}}(t)$ — случайное колебание, определяемое суммой модулированных поднесущих $s_{\text{ки}}(t)$; U_{max} — максимально допустимая величина колебания $u_{\text{мн}}(t)$.

При $N \gg 1$ и приблизительно одинаковых мощностях модулированных поднесущих выполняются условия центральной предельной теоремы и колебание $u_{\text{мн}}(t)$ можно считать стационар-

¹ В проводных, кабельных и гидроакустических многоканальных системах второй этап обычно отсутствует и многоканальное сообщение передается непосредственно по линии связи. Такой случай называется *передачей в полосе сообщения*. Наиболее часто в указанных системах применяют частотное уплотнение.

ным нормальным случайным процессом, одномерное распределение плотности вероятности которого имеет вид

$$p_1(u_{\text{мн}}) = \frac{1}{\sqrt{2\pi} u_{\text{эфф}}} \exp\left(-\frac{u_{\text{мн}}^2}{2u_{\text{эфф}}^2}\right), \quad (4.60)$$

где $u_{\text{эфф}}$ — эффективное значение $u_{\text{мн}}(t)$. По определению

$$u_{\text{эфф}}^2 = \frac{1}{T} \int_0^T u_{\text{мн}}^2(t) dt \approx \frac{S_{0i}^2 N}{2}. \quad (4.61)$$

При нахождении $u_{\text{эфф}}$ принято, что амплитуды поднесущих S_{0i} одинаковы, а мощность модулированных поднесущих такая же, как и немодулированных. Последнее верно для ЧМ и ФМ; для АМ имеющееся различие невелико и его можно не учитывать (см. формулу (4.10)).

Величина $u_{\text{мах}}$ выбирается так, чтобы вероятность события $|u_{\text{мн}}| > u_{\text{мах}}$ не превышала некоторой заранее установленной величины $P_{\text{доп}}$, т. е.

$$P(|u_{\text{мн}}| > u_{\text{мах}}) = 2 \int_{u_{\text{мах}}}^{\infty} p_1(u_{\text{мн}}) du_{\text{мн}} \leq P_{\text{доп}}. \quad (4.62)$$

Расчеты показывают, что при $P_{\text{доп}} = 10^{-3} - 10^{-6}$ пикфактор $k_{\text{п.мн}}$ многоканального сообщения $u_{\text{мн}}(t)$ принимает значения 3—4,5, т. е. (10—13) дБ¹.

В инженерных приложениях приведенными соотношениями можно пользоваться при $N \geq 10$. При этом нужно иметь в виду, что обычно заданной величиной является максимально допустимое значение многоканального сообщения, которое физически определяется линейным участком модуляционной характеристики общего модулятора системы. Связав между собой параметры многоканального сообщения и модулятора и наложив конкретные ограничения на них, можно решить ряд прикладных задач по выбору параметров частотного уплотнения каналов.

В заключение приведем пример записи сигнала. Пусть в первой ступени применена АМ, а во второй — ЧМ. В соответствии с (4.17) сигнал-переносчик, модулированный по частоте многоканальным сообщением, можно записать так:

$$s_{\text{АМ-ЧМ}}(t) = S_0 \cos \left[\omega_0 t + \Delta\omega_m \int_0^t x_{\text{мн}}(t) dt \right], \quad (4.63)$$

где

$$x_{\text{мн}}(t) = \frac{\sum_{i=1}^N S_{0i} [1 + M_i x_i(t)] \cos(\omega_{\text{ц}i} t + \varphi_{0i})}{k_{\text{п.мн}} u_{\text{эфф}}} \quad (4.64)$$

¹ Пикфактором многоканального сообщения $k_{\text{п.мн}}$ называют величину отношения максимально допустимого значения многоканального сообщения к его эффективному значению, т. е. $u_{\text{мах}}/u_{\text{эфф}}$.

Здесь $k_{п.мн}$ — пикфактор, а $u_{эфф}$ — эффективное значение многоканального сообщения, равное

$$u_{эфф}^2 = \frac{1}{2} \sum_{i=1}^N S_{0i}^2 [1 + M_i^2/k_{пi}^2].$$

Если принять одинаковыми амплитуды S_{0i} и коэффициенты M_i модуляции поднесущих, а также пикфакторы сообщений $k_{пi}$, то

$$u_{эфф} = 0,5 S_{0i}^2 N (1 + M_i^2/k_{пi}^2). \quad (4.65)$$

Аналогично можно получить выражения и для других видов сигналов в системах с частотным уплотнением каналов.

Сигналы при временном уплотнении. Многоканальное сообщение при временном уплотнении представляет собой последовательность синхрои импульсов и сумму N модулированных импульсных последовательностей, не перекрывающихся во времени:

$$u_{мн}(t) = \sum_{k=-\infty}^{\infty} U_{сх} f_{сх}(t - t_{0k}) + \sum_{i=1}^N \sum_{k=-\infty}^{\infty} U(t) f(t - t_k - i\Delta T_k). \quad (4.66)$$

Здесь первая сумма описывает последовательность синхрои импульсов с амплитудой $U_{сх}$ и единичной огибающей отдельного импульса $f_{сх}(t)$, а вторая двойная сумма — информационные последовательности импульсов N каналов, записанные в предположении, что интервал времени ΔT_k , занимаемый отдельным каналом, для всех каналов одинаков.

Для надежной и устойчивой работы канал синхронизации должен обладать значительно более высокой помехоустойчивостью по сравнению с информационными каналами. Это достигается тем, что энергия синхрои импульса выбирается больше энергии информационного импульса. Обычно энергию синхрои импульса увеличивают за счет увеличения его длительности. В современных системах связи синхрои импульсы чаще всего формируются в виде некоторой комбинации таких же импульсов, которые используются и в информационных каналах.

Канал синхронизации выполняет служебные функции, поэтому при записи многоканального сообщения его обычно не учитывают.

Чтобы записать выражение для сигнала в системе с временным уплотнением, необходимо представить многоканальное сообщение в нормированном виде. Нетрудно убедиться, что при использовании в первой ступени модуляции последовательностей ШИМ или ВИМ такое сообщение описывается выражением:

$$x_{мн}(t) = \frac{u_{мн}(t)}{U_0} = \sum_{i=1}^N \sum_{k=-\infty}^{\infty} f(t - t_k - i\Delta T_k), \quad (4.67)$$

где U_0 — амплитуда импульсов.

При использовании в первой ступени АИМ $u_{max} = 2U_0$ (при $M_a = 1$) и

$$x_{мн}(t) = \frac{u_{мн}(t)}{2U_0} = \frac{1}{2} \sum_{i=1}^N \sum_{k=-\infty}^{\infty} [1 + M_i x_i(t)] f(t - t_{0k} - i\Delta T_k). \quad (4.68)$$

Зная выражения для $x_{\text{мн}}(t)$, несложно записать сигнал для конкретного вида модуляции во второй ступени. В частности, если во второй ступени использована частотная модуляция, то сигнал имеет вид (4.63), где под девиацией частоты $\Delta\omega_m$ теперь нужно понимать половину максимального отклонения частоты при импульсной модуляции.

КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАЧИ

1. Какую операцию называют модуляцией колебаний?
2. Что такое модуляционная характеристика? Как получить заданную величину максимального отклонения параметра при модуляции?
3. Назовите основные разновидности амплитудной модуляции и запишите соответствующие им сигналы.
4. Считая максимальной (пиковую) мощность передатчика заданной, определите, во сколько раз средняя мощность сигнала с ОБП больше средней мощности одной боковой АМ сигнала. Решение найдите в общем виде, а также для случаев, когда пикфактор сообщения равен $\sqrt{2}$ и 3.
5. Запишите выражения для АМ, ФМ и ЧМ сигналов, если модулирующее напряжение имеет вид $u_x(t) = \sum_{i=1}^3 u_i \sin(\Omega_i t + \theta_i)$, где $u_1 = u_2 = 2\text{В}$; $u_3 = 3\text{В}$; $F_1 = 300\text{ Гц}$; $F_2 = 700\text{ Гц}$; $F_3 = 200\text{ Гц}$; фазы θ_i — независимы, случайны и равновероятны в интервале $[0, 2\pi]$. Чувствительность модулятора равна $K_{\text{АМ}} = 5$; $K_{\text{ФМ}} = 1\text{ рад/В}$; $K_{\text{ЧМ}} = 10^3\text{ Гц/В}$. Найдите удельную мощность и ширину спектра сигналов при условии, что в отсутствии модуляции амплитуда несущей во всех случаях одинакова.
6. Укажите основные особенности ФМ и ЧМ сигналов. В чем сходство и различие между этими сигналами?
7. Запишите выражения ФМ и ЧМ сигналов для произвольного непрерывного сообщения $u_x(t)$.
8. Как изменится ширина спектра ФМ и ЧМ сигналов, если эффективное напряжение сообщения на входе модулятора уменьшить вдвое?
9. Почему при передаче непрерывных сообщений предпочитают применять ЧМ, а не ФМ сигналы? Как преобразовать один вид сигнала в другой в передающей части системы связи? Что нужно сделать, чтобы на приемник ФМ сигнала принять ЧМ сигнал, и наоборот?
10. Назовите основные методы получения сигналов с угловой модуляцией и укажите их особенности.
11. Покажите, что при получении ФМ путем преобразования АМ сигнала, закон изменения фазы определяется выражением

$$\Delta\varphi(t) = \text{arctg} [M_a x(t)].$$

При каком значении M_a СКО, обусловленная нелинейными искажениями, не превышает 0,5%?

12. Какое колебание называют импульсной поднесущей? Запишите аналитическое выражение для такой поднесущей.
13. Получите выражения (4.31) и (4.32).
14. Покажите, что спектральные функции одиночного косинусоидального и треугольного импульсов определяются соответственно выражениями:

$$g(j\Omega) = \frac{2U_0\tau}{\pi} \frac{\cos(\Omega\tau/2)}{1 - [(2/\pi)\Omega\tau/2]^2}; \quad g(j\Omega) = \frac{U_0\tau}{2} \left[\frac{\sin(\Omega\tau/4)}{\Omega\tau/4} \right]^2,$$

где τ — длительность основания импульса.

15. Получите разложения в ряд Фурье для периодических последовательностей косинусоидальных и треугольных импульсов.
16. Каким требованиям должна удовлетворять периодическая последовательность импульсов, используемая для импульсной модуляции?
17. Назовите основные виды импульсной модуляции и дайте их краткую характеристику.
18. Рассчитайте необходимую чувствительность модулятора ШИМ, если $\max u_x = 1\text{В}$; полоса частот, занимаемая каналом, равна 200 кГц ; $\Delta\tau_m = 0,5\tau$.
19. В чем различие между импульсной модуляцией 1-го и 2-го рода? Почему при $Q \gg 1$ с этим различием можно не считаться?
20. Нарисуйте спектры АИМ, ШИМ и ВИМ для случая, когда сообщение: 1) гармоническое; 2) случайный процесс.
21. Покажите, что в спектре балансной АИМ отсутствуют гармоники частоты повторения F_n . Сравните спектральную структуру этого вида модуляции со структурой обычной АИМ 1-го рода.
22. Нарисуйте схему получения ШИМ с использованием триггера с одним устойчивым состоянием. Укажите недостатки такого способа по сравнению с использованием триггера, имеющего два устойчивых состояния. Нарисуйте эпюры формирования ШИМ, соответствующие этим двум способам.
23. Нарисуйте схему получения ФИМ из последовательности ШИМ и нарисуйте эпюры, соответствующие такому преобразованию.
24. Какую операцию называют уплотнением каналов? В чем суть этой операции?
25. Назовите основные виды уплотнения каналов и дайте их краткую характеристику.
26. Какие меры уменьшают взаимное влияние каналов, при уплотнении?
27. Запишите выражения для сигнала в следующих многоканальных системах: АМ—ЧМ; ЧМ—ЧМ; ШИМ—ЧМ; ВИМ—АМ; ВИМ—ЧМ. Получите приближенные выражения для ширины спектра таких сигналов, считая импульсы прямоугольными. Всеми необходимыми данными задаться в общем виде.
28. Рассчитайте максимальное число каналов в системе АИМ—ЧМ,

если частота повторения импульсов в каждом канале $F_n = 250$ Гц; полоса пропускания УПЧ приемного устройства $\Delta f_{\text{упч}} = 150$ кГц; девиация частоты сигнала $\Delta f_m = 50$ кГц. Иметь в виду, что в данном случае Δf_m — величина полного изменения частоты от f_0 в паузах до $f_0 + \Delta f_m$ в моменты, соответствующие максимальному значению импульсов АИМ, равному $2U_0$, где U_0 — амплитуда импульсов в отсутствии модуляции. При расчете принять $\tau_{\text{сх}} = 3\tau$; $\tau_3 = \tau$.

29. Рассчитайте максимальное число каналов в системах ВИМ—ЧМ (при $\Delta\tau_m = 2\tau$) и ШИМ—ЧМ (при $\Delta\tau_m = \tau/2$). Остальные данные те же, что в задаче 28.
30. Определите ширину спектра многоканального сообщения при частотном уплотнении каналов, если известно: $N = 12$; ширина спектра сообщений F_B во всех каналах одинакова ($F_B = 3400$ Гц). Задачу решить для двух видов модуляции поднесущих: АМ и ЧМ с индексом $\Delta\phi = 5$.

ГЛАВА 5. ПРИЕМ СИГНАЛОВ В СИСТЕМАХ ПЕРЕДАЧИ НЕПРЕРЫВНЫХ СООБЩЕНИЙ

5.1. ВВЕДЕНИЕ

Прием сигналов представляет собой одну из наиболее сложных теоретических и инженерных задач, возникающих при передаче сообщений по каналам связи. Сложность ее обусловлена тем, что переданные сообщения приходится извлекать из смеси ослабленного (и, как правило, искаженного при прохождении через среду) сигнала и помех. Это приводит к тому, что сообщения на выходе приемника удастся воспроизвести только приближенно, т. е. с некоторой ошибкой.

Хотя теория и указывает на принципиальные возможности снижения этой ошибки до сколь угодно малой величины (см. гл. 3), практические соображения заставляют ограничивать ошибку некоторым ее конечным допустимым значением.

В этой связи особую важность приобретают вопросы определения величины ошибок при передаче непрерывных сообщений и отыскание наиболее помехоустойчивых способов приема. Эти вопросы могут быть решены на основе теории оптимального приема, которая позволяет найти структуру оптимального приемника и определить его помехоустойчивость. Однако, как указывалось в гл. 3, теория оптимального приема непрерывных сообщений разработана только для достаточно больших отношений сигнала к шуму на входе приемника. Поэтому рассмотрению приложений этой теории предшествует материал п. 5.2, в котором дана геометрическая трактовка некоторых особенностей приема сигналов, модулированных непрерывными сообщениями. Это поз-

воляет выяснить, при каких условиях отношение сигнала к шуму на входе приемника можно считать достаточным, чтобы теория была верна.

При рассмотрении вопросов оптимального приема в этой главе полагается, что помехи представляют собой аддитивные гауссовские процессы с равномерным энергетическим спектром, а искажения в самой системе отсутствуют. Непрерывные сообщения считаются случайными процессами с ограниченным спектром.

В главе принят следующий порядок изложения материала. Сначала рассматривается *помехоустойчивость оптимального приема* (*потенциальная помехоустойчивость* по терминологии В. А. Котельникова), а затем структура некоторых оптимальных приемников. Такой «обратный» порядок изложения объясняется тем, что потенциальную помехоустойчивость приема сигналов с разными видами модуляции можно определить, не зная структуру приемных устройств. Это позволяет сравнивать разные виды сигналов и выбирать наиболее подходящие по помехоустойчивости, не затрагивая вопросы построения приемника. После этого рассматривается структура некоторых *оптимальных приемников (оптимальных демодуляторов)* и обсуждаются особенности их реализации.

Так как при передаче непрерывных сообщений широко используется принцип многоканальности, в главе уделено большое внимание вопросам разделения каналов на приемной стороне многоканальных систем связи и ошибкам, которые возникают при такой операции.

5.2. ГЕОМЕТРИЧЕСКАЯ ТРАКТОВКА ОСОБЕННОСТЕЙ ПРИЕМА СИГНАЛОВ В ПРИСУТСТВИИ ШУМА

Характерной особенностью АМ сигналов является то, что они представляют собой линейную функцию передаваемого сообщения $x(t)$. Поэтому АМ сигналы и их разновидности (АМ—ПН, ОБП) относятся к *линейным видам модуляции*.

Для сигналов с ФМ или ЧМ характерна нелинейная зависимость от передаваемых сообщений. Действительно, представив ФМ сигнал в квадратурных составляющих, получим

$$s_{\text{ФМ}}(t) = S_0 \cos[\omega_0 t + \Delta\varphi_m x(t)] = F_c[x(t)] \cos \omega_0 t + F_s[x(t)] \sin \omega_0 t \quad (5.1)$$

где $F_c[x(t)] = S_0 \cos[\Delta\varphi_m x(t)]$; $F_s[x(t)] = S_0 \sin[\Delta\varphi_m x(t)]$ — нелинейные функции сообщения $x(t)$. Аналогично можно представить и ЧМ сигнал.

Из-за отмеченной особенности ФМ и ЧМ относят к *нелинейным видам модуляции*, несмотря на то, что их модулируемые параметры (фаза или частота) являются линейными функциями сообщений.

Наиболее удобно и наглядно различия между линейными и нелинейными видами модуляции можно уяснить, пользуясь геометрическими представлениями модулированных сигналов.

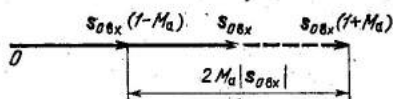


Рис. 5.1. Геометрическая трактовка АМ сигнала

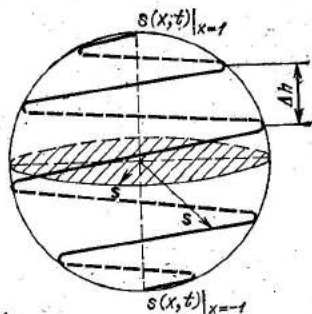


Рис. 5.2. Геометрическая трактовка ФМ и ЧМ сигналов

В геометрической трактовке АМ сигналу соответствует вектор s_{AM} переменной длины, всегда совпадающий по направлению с вектором несущей s_{0bx} . Линия, которую описывает конец вектора сигнала при изменении сообщения $x(t)$, называется *линией сигналов*. При АМ эта линия всегда расположена на прямой (рис. 5.1), а ее длина конечна и равна $M_a |s_{0bx}|$.

Сигналам ФМ и ЧМ в геометрической трактовке соответствуют векторы постоянной длины, концы которых расположены на поверхности гиперсферы с радиусом s . При изменении сообщения $x(t)$ конец вектора сигнала перемещается по линии сигналов, расположенной на поверхности гиперсферы. Для иллюстрации на рис. 5.2 в трехмерном пространстве показана поверхность шара с расположенной на ней линией сигналов, соответствующей ЧМ или ФМ. В отличие от линейной модуляции протяженность линии сигналов при ЧМ или ФМ может быть сделана сколь угодно большой без увеличения мощности (или энергии) сигналов. Однако при этом «витки» линии нужно укладывать на поверхности гиперсферы все плотнее и плотнее и расстояние Δh между соседними витками («шаг намотки») будет все меньше и меньше. Эта особенность приводит к существенному различию помехоустойчивости нелинейных видов модуляции по сравнению с линейными.

Рассмотрим небольшой участок поверхности гиперсферы, на котором расположены три соседних «витка» линии сигнала с номерами $(i-1)$, i и $(i+1)$ (рис. 5.3). Допустим, что в некоторый

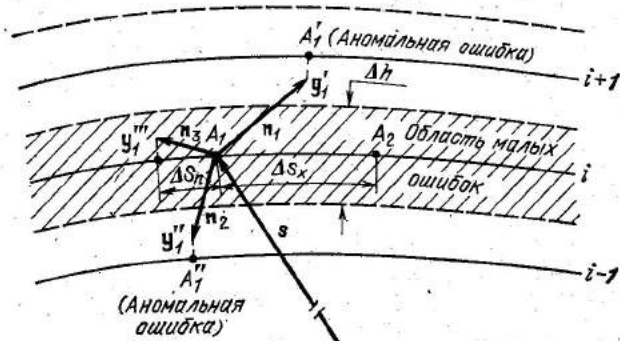


Рис. 5.3. Участок поверхности с тремя соседними витками линии ФМ и ЧМ сигналов

момент времени t_1 конец вектора сигнала s находился на линии сигнала в точке A_1 . При изменении непрерывного сообщения на величину Δx ($|\Delta x| \ll 1$) сигнал изменится на некоторую величину Δs_x , а конец вектора сигнала переместится по линии сигналов из точки A_1 в точку A_2 . Установим связь между Δx и Δs_x , взяв их отношение¹

$$K_s = \frac{\Delta s_x}{\Delta x} = \frac{s(t + \Delta t) - s(t)}{x(t + \Delta t) - x(t)} \approx \frac{\partial s(t)}{\partial x(t)}. \quad (5.2)$$

Представив (5.2) в форме

$$K_s = \frac{\partial s(t)}{\partial t} \Big/ \frac{\partial x(t)}{\partial t} = \frac{s'(t)}{x'(t)}, \quad (5.3)$$

и применив неравенство Бернштейна (см. (4.22)) к числителю и знаменателю выражения (5.3), получим

$$K_s \leq \Delta f_c \max |s(t)| / F_B \max |x(t)|,$$

где Δf_c — ширина спектра модулированного сигнала; F_B — верхняя граничная частота спектра сообщения.

Приняв во внимание, что $\max |x(t)| = 1$, а для ЧМ (или ФМ) $\max |s(t)| = S_{0 \text{ вх}}$, получим следующее выражение для K_s :

$$|K_s| \approx S_{0 \text{ вх}} \Delta \varphi_m. \quad (5.4)$$

Величина K_s показывает, во сколько раз приращение сигнала Δs больше соответствующего ему приращения сообщения Δx . Поэтому K_s можно назвать коэффициентом растяжения сигнала. Чем больше этот коэффициент, тем протяженнее линия сигналов. Из выражения (5.4) следует, что при ЧМ или ФМ можно увеличить длину линии сигналов за счет увеличения индекса модуляции $\Delta \varphi_m$, не изменяя при этом мощность (или энергию) сигнала.

Если кроме сигнала на входе приемника действует также аддитивная флуктуационная помеха $n(t)$, то она будет «сбивать» конец вектора с линии сигналов. Пока помеха «мала», результирующий вектор $y = s + n$ с большой вероятностью попадает в окрестность точки A_1 , находящейся на рассматриваемом витке линии сигналов.

Влияние «малой» помехи можно оценить величиной дополнительного («паразитного») смещения Δs_n точки A_1 по линии сигналов. Полное смещение, определяемое как сообщением, так и помехой, окажется равным

$$\Delta s = \Delta s_x + \Delta s_n = \Delta s_x (1 + \Delta s_n / \Delta s_x). \quad (5.5)$$

Случайную величину $\epsilon = \Delta s_n / \Delta s_x$ можно рассматривать как ошибку, вызванную помехой.

Чтобы влияние помехи было незначительным («малая» помеха), нужно обеспечить условие $|\epsilon| \ll 1$. Поскольку величина ϵ

¹ Здесь и ниже частная производная сигнала берется по сообщению, потому что модулированный сигнал является функцией двух аргументов — времени и сообщения.

случайна, то это условие носит вероятностный характер и должно выполняться с вероятностью, близкой к единице. Выполнить указанное условие (или эквивалентное ему условие $|\Delta s_n| \ll |\Delta s_x|$) можно по-разному: увеличением мощности сигнала на входе приемника; увеличением индекса модуляции; одновременным увеличением и мощности сигнала и его индекса модуляции.

Следует подчеркнуть, что стремление ослабить влияние помех только за счет увеличения индекса модуляции (т. е. за счет неэнергетического параметра), может привести к обратному результату и вместо уменьшения ошибки произойдет ее резкое возрастание. Это связано с тем, что при увеличении индекса модуляции расстояние между соседними витками линии сигнала уменьшается, а вероятность попадания конца вектора u в окрестность одного из соседних витков возрастает. «Перескоки» вектора u из окрестности «своего» витка в окрестность «чужого» витка связаны с появлением значительных ошибок, которые называются *аномальными*. Величина аномальных ошибок во много раз превосходит величину «нормальных» ошибок, которые характерны для «малой» (слабой) помехи.

Эффект, связанный с появлением аномальных ошибок, называется *пороговым*. Количественно пороговый эффект можно характеризовать величиной допустимой вероятности появления аномальных ошибок. До тех пор пока вероятность «перескоков» не превышает допустимой величины, можно считать, что имеет место «малая» помеха и оптимальный прием осуществляется в надпороговой области. Это означает, что отношение сигнала к шуму на входе превышает пороговое.

Для иллюстрации сказанного на рис. 5.4 приведена качественная зависимость ошибки приема δ от отношения средних мощностей шума и сигнала на входе приемника в случае нелинейной модуляции. Зависимость имеет резкий перегиб в окрестности порогового отношения шума к сигналу $(P_{ш}/P_c)_{пор}$. Область I, где $(P_{ш}/P_c)_{вх} \leq (P_{ш}/P_c)_{пор}$ — надпороговая. В этой области осуществляется прием, соответствующий случаю «малой» помехи (случай сильного сигнала). Область II — подпороговая. Она соответствует случаю «большой» помехи (случай слабого сигнала). Для этой области характерен резкий рост ошибки приема. В случае линейной модуляции зависимость $\delta = f[(P_{ш}/P_c)_{вх}]$ имела бы линейный характер (см. пунктирную линию на рис. 5.4).

Режим работы систем передачи непрерывных сообщений и параметры этих систем обычно стремятся выбрать так, чтобы прием происходил в надпороговой области.

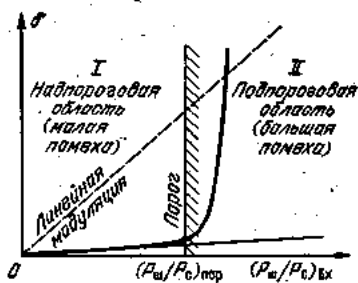


Рис. 5.4. Зависимость ошибки от отношения сигнала к шуму на входе приемника сигналов с нелинейной модуляцией

5.3. ПОМЕХОУСТОЯЧИВОСТЬ ОПТИМАЛЬНОГО ПРИЕМА НЕПРЕРЫВНЫХ СООБЩЕНИЙ

Задачей приемного устройства являются извлечение переданного сообщения $x(t)$ из входного колебания $y(t)$. Однако из-за помех и искажений эта процедура не может быть выполнена точно, и восстановить сообщение на выходе приемника можно только приближенно. Такое приближенное сообщение называют *оценкой* и обозначают $\hat{x}(t)$.

Критерием близости $\hat{x}(t)$ и $x(t)$ в теории и технике связи принята СКО (см. гл. 1), в соответствии с которой

$$\delta^2 = \langle [\hat{x}(t) - x(t)]^2 \rangle / \langle x^2(t) \rangle = k_{\text{пх}}^2 \langle [\hat{x}(t) - x(t)]^2 \rangle, \quad (5.6)$$

где скобки $\langle \cdot \rangle$ означают операцию усреднения реализации по времени¹.

Оптимальный приемник непрерывных сообщений обеспечивает наименьшую возможную в заданных условиях величину СКО. Определим эту ошибку.

Основываясь на теории ортогональных разложений передачу любого непрерывного сообщения можно заменить передачей совокупности числовых коэффициентов (параметров). Пусть непрерывное сообщение $x(t)$ представлено рядом (см. гл. 2)

$$x(t) = \sum_{k=1}^n a_k \varphi_k(t). \quad (5.7)$$

При известной системе базисных функций передача сообщений $x(t)$ эквивалентна передаче n значений коэффициентов a_k . Следовательно, передаваемый сигнал можно рассматривать как функцию времени и коэффициентов a_k , т. е.

$$s(t) = s[x(t), t] = s(a_1, a_2, \dots, a_k, \dots, a_n, t). \quad (5.8)$$

Влияние помех приведет к тому, что каждый коэффициент a_k будет принят с некоторой погрешностью Δa_k . В результате оценка сообщения примет вид

$$\hat{x}(t) = \sum_{k=1}^n (a_k + \Delta a_k) \varphi_k(t) = x(t) + \varepsilon(t), \quad (5.9)$$

где колебание $\varepsilon(t)$ нужно рассматривать как помеху на выходе приемника.

Если единственной причиной появления этой помехи является белый гауссовский шум на входе приемника, то нетрудно убедиться в том, что помеха $\varepsilon(t)$ имеет нормальное распределение. В. А. Котельников показал [1], что в режиме надпорогового опти-

¹ Для стационарных эргодических процессов такое усреднение, как известно, эквивалентно статистическому усреднению (усреднению по всем реализациям). В (5.6) учтено, что для нормированных сообщений квадрат пикфактора определяется величиной $1/\langle x^2(t) \rangle = k_{\text{пх}}^2$.

мального приема спектральная плотность такой помехи определяется выражением

$$N_{0s}(f) = N_0 / \frac{1}{T} \int_0^T \left[\frac{\partial s(x, t)}{\partial x(t)} \right]^2 dt = N_0 / \left\langle \left[\frac{\partial s(x, t)}{\partial x(t)} \right]^2 \right\rangle. \quad (5.10)$$

Знаменатель в (5.10) представляет собой усредненное значение квадрата коэффициента растяжения сигнала (см. выражение (5.2)):

$$\langle K_s^2 \rangle = \left\langle \left[\frac{\partial s(x, t)}{\partial x(t)} \right]^2 \right\rangle. \quad (5.11)$$

Чем больше эта величина, тем слабее проявляется действие входного шума и тем меньше спектральная плотность помехи на выходе.

Средний квадрат ошибки при оптимальном приеме непрерывных сообщений с учетом (5.6) можно найти по формуле

$$\delta^2 = k_{\text{пх}}^2 \langle s^2(t) \rangle = k_{\text{пх}}^2 \int_0^{F_B} N_{0s}(f) df. \quad (5.12)$$

Для выбранного (или заданного) вида модулированных сигналов помехоустойчивость оптимального приема будет наиболее высокой по сравнению с любым возможным реальным способом приема этих же сигналов. Поэтому такую помехоустойчивость часто называют *потенциальной* (предельно возможной для данного вида сигналов).

При анализе потенциальной помехоустойчивости полезно различать прямые виды модуляции, у которых передаваемое сообщение $x(t)$ непосредственно входит в выражение для сигнала, т. е. $s(t) = s[x(t), t]$, и интегральные, у которых сигнал — функция интеграла от передаваемого сообщения, т. е. $s(t) = s[\int x(t) dt, t]$.

Рассмотрим особенности расчета потенциальной помехоустойчивости для некоторых случаев.

Помехоустойчивость сигналов с амплитудной модуляцией. Пусть для передачи непрерывных сообщений используется АМ сигнал (см. (4.8) в гл. 4). В соответствии с (5.10) — (5.12)

$$\langle K_{s \text{ АМ}}^2 \rangle = \frac{1}{T} \int_0^T S_0^2 M_a^2 \cos^2(\omega_0 t + \varphi_0) dt = \frac{S_0^2 M_a^2}{2}, \quad (5.13)$$

$$N_{0s}(f) = N_0 / K_{s \text{ АМ}}^2 = 2 N_0 / S_0^2 M_a^2. \quad (5.14)$$

$$\delta_{\text{АМ}}^2 = 2 k_{\text{пх}}^2 N_0 F_B / S_0^2 M_a^2. \quad (5.15)$$

В (5.13) учтено, что $\cos^2 \alpha = 0,5(1 + \cos 2\alpha)$ и интеграл распадается на две составляющих, одна из которых (с частотой $2\omega_0$) близка к нулю и отброшена.

Из (5.14) следует, что при АМ сигнале спектральная плотность помехи на выходе оптимального приемника постоянна. Эта

особенность характерна не только для АМ, но и всех других сигналов с прямыми видами модуляции.

Приняв во внимание, что средние мощности сигнала и шума на входе приемника

$$P_c \approx S_0^2/2; \quad P_{ш} = 2N_0F_B = N_0\Delta f_{AM},$$

где $\Delta f_{AM} = 2F_B$ — ширина спектра АМ сигнала, определяющая полосу пропускания приемника, имеем

$$\delta_{AM}^2 = \frac{1}{2} \left(\frac{P_{ш}}{P_c} \right)_{вх} \frac{k_{пх}^2}{M_a^2}. \quad (5.16)$$

В соответствии с (5.16) потенциальная помехоустойчивость АМ сигналов в основном определяется отношением сигнала к шуму на входе приемника. Для получения малых значений ошибки это отношение должно быть весьма большим. Например, для получения $\delta_{AM} \leq 10^{-2}$ ($\leq 1\%$) при $k_{пх} = 3$ и $M_a = 1$ необходимо иметь на входе $(P_c/P_{ш})_{вх} \geq 4,5 \cdot 10^4$ (т. е. 46,5 дБ).

Из этого примера видно, что АМ сигналы обладают низкой потенциальной помехоустойчивостью.

Помехоустойчивость сигналов с угловой модуляцией. Пусть для передачи непрерывных сообщений используются сигналы с угловой модуляцией. Сначала рассмотрим случай фазовой модуляции, описываемой выражением (4.12). В соответствии с (5.10) и (5.11) имеем

$$\langle K_s^2 \Phi_M \rangle = \frac{1}{T} \int_0^T S_0^2 \Delta \Phi_m^2 \cos^2 [\omega_0 t + \Delta \Phi_m x(t)] dt = \frac{S_0^2 \Delta \Phi_m^2}{2}; \quad (5.17)$$

$$N_{ФМ} = 2N_0/S_0^2 \Delta \Phi_m^2. \quad (5.18)$$

При вычислении (5.17) приняты во внимание те же соображения, что и при вычислении (5.13).

Из (5.18) следует, что при ФМ сигнале, как и при АМ, спектральная плотность помехи на выходе постоянна, поскольку ФМ принадлежит к сигналам с прямой модуляцией. Приняв во внимание (5.12) и (5.18), имеем

$$\delta_{ФМ}^2 = \frac{2N_0 k_{пх}^2 F_B}{\Delta \Phi_m^2 S_0^2} = \frac{k_{пх}^2}{\Delta \Phi_m^2} \frac{F_B}{\Delta f_c} \left(\frac{P_{ш}}{P_c} \right)_{вх}, \quad (5.19)$$

где $P_{ш} = N_0 \Delta f_c$ и $P_c = S_0^2/2$ — средние мощности шума и сигнала на входе приемника; Δf_c — полоса частот, занимаемая спектром ФМ сигнала.

Проведем теперь рассмотрение для ЧМ сигнала, описываемого выражением (4.18). Так как ЧМ относится к интегральному виду

модуляции, то при определении величины $\langle K_s^2 \rangle$ частную производную сигнала по сообщению надо находить как производную сложной функции:

$$\langle K_s^2 \rangle_{\text{ЧМ}} = \left\langle \left[\frac{\partial s(\Theta, t)}{\partial \Theta(t)} \frac{\partial \Theta(t)}{\partial x(t)} \right]^2 \right\rangle = \left\langle \left[\frac{\partial s(\Theta, t)}{\partial \Theta(t)} \frac{\partial \Theta(t)}{\partial t} \frac{\partial t}{\partial x'(t)} \right]^2 \right\rangle = \\ = \left\langle \left[\frac{\partial s(\Theta, t)}{\partial \Theta(t)} \frac{x(t)}{x'(t)} \right]^2 \right\rangle.$$

Учитывая, что спектры сигнала и сообщения не перекрываются и среднее значение квадрата произведения равно произведению средних квадратов, для ЧМ сигнала имеем

$$\langle K_s^2 \rangle_{\text{ЧМ}} = \left\langle \left[\frac{\partial s(\Theta, t)}{\partial \Theta(t)} \right]^2 \right\rangle \left\langle \left[\frac{x(t)}{x'(t)} \right]^2 \right\rangle = \frac{S_0^2 \Delta \omega_m^2}{2} \left\langle \left[\frac{x(t)}{x'(t)} \right]^2 \right\rangle. \quad (5.20)$$

Приняв во внимание, что колебания $x(t)$ и $x'(t)$ ортогональны, можно показать, что

$$\langle K_s^2 \rangle_{\text{ЧМ}} = S_0^2 \Delta \omega_m^2 / 2 \Omega^2, \quad (5.21)$$

где Ω — текущая частота, принимающая значения в интервале $[0, \Omega_0]$. В соответствии с (5.10) спектральная плотность помехи на выходе оптимального приемника ЧМ сигналов равна

$$N_{0\text{в}}(f) = \frac{2 N_0 \Omega^2}{S_0^2 \Delta \omega_m^2} = \frac{2 N_0 F^2}{S_0^2 \Delta f_m^2}. \quad (5.22)$$

Эта формула показывает, что при ЧМ сигнале спектральная плотность помехи на выходе имеет квадратичную зависимость от частоты. Такая зависимость характерна для всех интегральных видов модуляции.

Средний квадрат ошибки при приеме ЧМ сигналов в соответствии с формулой (5.12) можно записать так:

$$\delta_{\text{ЧМ}}^2 = \frac{2 N_0 k_{\text{пх}}^2 F_{\text{в}}^3}{3 S_0^2 \Delta_m^2} = \frac{k_{\text{пх}}^2}{3 \Delta \varphi_{\text{ЧМ}}^2} \frac{F_{\text{в}}}{\Delta f_c} \left(\frac{P_{\text{ш}}}{P_{\text{с}}} \right)_{\text{вх}}, \quad (5.23)$$

где $\Delta \varphi_{\text{ЧМ}} = \Delta f_m / F_{\text{в}}$ — индекс частотной модуляции.

Проанализируем полученные результаты. Из (5.19) и (5.23) следует, что при ФМ и ЧМ помехоустойчивость приема можно повысить только за счет увеличения индекса модуляции $\Delta \varphi$ (не увеличивая при этом среднюю мощность сигнала $P_{\text{с}}$). Однако увеличение $\Delta \varphi$ приводит к расширению спектра ФМ и ЧМ сигналов и соответственно к необходимости использовать более широкую полосу частот. Это уменьшает отношение сигнала к шуму на входе приемника $q_{\text{с}} = (P_{\text{с}}/P_{\text{ш}})_{\text{вх}}$. При некотором значении индекса $\Delta \varphi = \Delta \varphi_{\text{кр}}$ величина $q_{\text{с}}$ снизится до пороговой величины $q_{\text{пор}}$, при которой условия надпорогового приема нарушаются и начинает резко возрастать вероятность аномальных ошибок $P_{\text{ан}}$. В этом случае формулами (5.19) и (5.23) пользоваться уже нельзя.

Исследования показывают, что для сигналов с угловой модуляцией величина $P_{ан}$ связана со значениями $\Delta\varphi_{кр}$ и $q_{пор}$ приближенным соотношением [2]

$$P_{ан} \approx \frac{1 + \Delta\varphi_{кр}}{2\sqrt{3}\pi} \frac{e^{-q_{пор}}}{\sqrt{q_{пор}}} \quad (5.24)$$

Задавая приемлемый уровень $P_{ан}$ (обычно принимают $P_{ан} \approx 10^{-5} - 10^{-6}$), можно установить связь между величинами $\Delta\varphi_{кр}$ и $q_{пор}$ и тем самым определить условия, при которых формулы (5.19) и (5.23) верны.

В соответствии с изложенной методикой можно найти величину СКО воспроизведения непрерывных сообщений на выходе оптимального приемника и для других видов модуляции сигналов. При решении таких задач в случае многоканальной передачи сообщений появляются некоторые особенности. Рассмотрим их.

5.4. ПОТЕНЦИАЛЬНАЯ ПОМЕХОУСТОЙЧИВОСТЬ ПРИ МНОГОКАНАЛЬНОЙ ПЕРЕДАЧЕ НЕПРЕРЫВНЫХ СООБЩЕНИЙ

Сигналы в многоканальных системах зависят от способов уплотнения каналов и видов модуляции в первой и второй ступенях. Иными словами, сигналы в таких системах — сложные функции времени и передаваемых сообщений, т. е.

$$s(t) = s[x_{мн}(t), t] = s\{D_{упл}[x_1(t), x_2(t) \dots x_N(t)], t\}, \quad (5.25)$$

где $x_{мн}(t)$ — многоканальное сообщение или групповой сигнал (см. п. 1.2); $D_{упл}$ — оператор уплотнения каналов.

Одновременную передачу непрерывных сообщений от N независимых источников можно заменить передачей соответствующих каждому сообщению коэффициентов разложения. Конечно, эта задача сложнее передачи сообщений по одному каналу, но принципиально она решается так же, как и в рассмотренных выше случаях.

Спектральную плотность помехи на выходе i -го канала можно определить с помощью выражения (5.10), в котором теперь нужно учитывать сложную зависимость сигнала (5.25) от сообщений $x_i(t)$, передаваемых по i -му каналу. В соответствии с этим имеем

$$N_{0e_i}(f) = N_0 \left\langle \left[\frac{\partial s(x_{мн}, t)}{\partial x_i(t)} \right]^2 \right\rangle = N_0 / \langle K_{si}^2 \rangle. \quad (5.26)$$

Сложная зависимость сигнала от передаваемых сообщений проявляется различно при разных видах уплотнения и модуляции. Опираясь на результаты, изложенные выше, можно найти развернутые выражения $\langle K_{si}^2 \rangle$ для многоканальных систем с разными сочетаниями видов модуляции в первой и второй ступенях. Эти выражения сведены в табл. 5.1.

Таблица 5.1

Вид модуляции в 1-й и 2-й ступенях	Выражение для коэффициента $\langle K_{si}^2 \rangle$	Тип многоканальной системы
прямая — прямая	$\left\langle \left[\frac{\partial s(x_{MH}, t)}{\partial x_{MH}(t)} \right]^2 \right\rangle \left\langle \left[\frac{\partial x_{MH}(t)}{\partial x_i(t)} \right]^2 \right\rangle$	АМ—АМ; АМ—ФМ; АИМ—АМ; ФИМ—АМ и т. п.
интегральная — прямая	$\left\langle \frac{1}{\Omega^2} \left[\frac{\partial s(x_{MH}, t)}{\partial x_{MH}(t)} \right]^2 \right\rangle \left\langle \left[\frac{\partial x_{MH}(t)}{\partial \theta_i(t)} \right]^2 \right\rangle$	ЧМ—АМ; ЧМ—ФМ; ЧИМ—АМ; ЧИМ—ФМ
прямая — интегральная	$\left\langle \frac{1}{\omega^2} \left[\frac{\partial s(x_{MH}, t)}{\partial \theta_{MH}(t)} \right]^2 \right\rangle \left\langle \left[\frac{\partial x_{MH}(t)}{\partial x_i(t)} \right]^2 \right\rangle$	АМ—ЧМ; ФМ—ЧМ; ФИМ—ЧМ; АИМ—ЧМ
интегральная — интегральная	$\left\langle \frac{1}{\omega^2} \left[\frac{\partial s(x_{MH}, t)}{\partial \theta_{MH}(t)} \right]^2 \right\rangle \left\langle \frac{1}{\Omega^2} \left[\frac{\partial x_{MH}(t)}{\partial x_i(t)} \right]^2 \right\rangle$	ЧМ—ЧМ; ЧИМ—ЧМ

В таблице Ω и ω — текущие частоты, которые принимают значения в пределах, занимаемых спектрами канального $x_i(t)$ и многоканального $x_{MH}(t)$ сообщений соответственно, т. е. $0 \leq \Omega \leq \leq \Omega_{вi}$; $0 \leq \omega \leq \leq \omega_{в.мн.}$ Функции $\theta_i(t)$ и $\theta_{MH}(t)$ соответственно равны

$$\theta_i(t) = \int_0^t x_i(t) dt; \quad \theta_{MH}(t) = \int_0^t x_{MH}(t) dt.$$

Зная выражение для $\langle K_{si}^2 \rangle$, с помощью (5.26) можно найти спектральную плотность помехи на выходе отдельного канала той или иной системы, а затем определить СКО воспроизведения принятого по этому каналу сообщения.

В общем случае для систем с частотным уплотнением каналов СКО на выходе i -го канала определяется выражением

$$\delta_i^2 = \frac{k_{nxi}^2}{\Delta F_i} \int_0^{F_{вi}} \int_{f_i}^{f_i} N_{0ei}(f, F) df dF, \quad (5.27)$$

где $f_i = f_{ni} - 0,5\Delta F_i$; $f_i = f_{ni} + 0,5\Delta F_i$; f_{ni} — частота поднесущей i -го канала; ΔF_i — полоса частот, занимаемая модулированной поднесущей; $F_{вi}$ — верхняя граничная частота спектра сообщения i -го канала.

Для систем с временным уплотнением каналов СКО на выходе i -го канала определяется выражением

$$\delta_i^2 = k_{nxi}^2 \int_0^{F_{вi}} N_{0ei}(f) df. \quad (5.28)$$

Приведем примеры расчета потенциальной помехоустойчивости некоторых многоканальных систем.

Система АМ—ЧМ. Сигнал в такой системе определяется выражением (4.63). Примем для простоты те же допущения, которые были сделаны в п. 4.9. ($S_{0i} = \text{const}$; $M_i = M$; $k_{\text{п.и.}} = k_{\text{п.к}}$). Тогда

$$\langle K_{st}^2 \rangle = S_0^2 \Delta \omega_m^2 k_{\text{п.к}}^2 M^2 / 2 \omega^2 N k_{\text{п.и.}}^2 (k_{\text{п.к}}^2 + M^2). \quad (5.29)$$

Приняв во внимание (5.26) и (5.27), получим

$$\begin{aligned} \delta_{i \text{ АМ-ЧМ}}^2 &= \frac{1}{\Delta F_i} \int_0^{F_{\text{вт}}} \int_{f_1}^{f_2} \frac{2 N k_{\text{п.и.}}^2 (k_{\text{п.к}}^2 + M^2) N_0 f^2}{M^2 S_0^2 \Delta f_m^2} df dF = \\ &= \frac{2 N k_{\text{п.к}}^2 (k_{\text{п.к}}^2 + M^2) N_0 F_{\text{вт}}}{3 M^2 S_0^2 \Delta f_m^2 \Delta F_i} (f_2^3 - f_1^3). \end{aligned}$$

После несложных преобразований это выражение можно привести к виду

$$\delta_{i \text{ АМ-ЧМ}}^2 = \frac{N k_{\text{п.к}}^2 (k_{\text{п.к}}^2 + M) F_{\text{вт}} f_{\text{пт}}^2}{M^2 \Delta f_m^2 \Delta f_c} \left(\frac{P_{\text{ш}}}{P_c} \right)_{\text{вх}}, \quad (5.30)$$

где $P_{\text{ш}} = \Delta f_c N_0$; $P_c = S_0^2 / 2$; N — число каналов; Δf_c — ширина спектра сигнала.

При выводе выражения (5.30) принято, что $1 + \frac{1}{12} \left(\frac{\Delta F_i}{f_{\text{пт}}} \right)^2 \approx 1$. Полученный результат показывает, что ошибка зависит от отношения сигнала к шуму на входе приемника, параметров системы и статистических свойств сообщений.

Система АИМ—ЧМ. В этой системе в первой ступени применен прямой вид модуляции, а во второй — интегральный. Высокочастотный сигнал в такой системе определяется выражением (4.63), в котором функция $x_{\text{и.к}}(t)$ имеет вид (4.68). В соответствии с табл. 5.1

$$\langle K_{st}^2 \rangle = \frac{S_0^2 \Delta \omega_m^2}{2 \omega^2} \left\langle \left[\sum_{k=-\infty}^{\infty} M_i f(t - t_{0k} - i \Delta T_k) \right]^2 \right\rangle.$$

Поскольку импульсы $f(t)$ для разных k не перекрываются во времени, усреднение квадрата суммы равно усреднению суммы квадратов и, следовательно,

$$\langle K_{st}^2 \rangle = \frac{M_i S_0^2 \Delta \omega_m^2}{2 \omega^2} \left\langle \sum_{k=-\infty}^{\infty} f^2(t - t_{0k} - i \Delta T_k) \right\rangle.$$

Можно доказать, что

$$\left\langle \sum_{k=-\infty}^{\infty} f^2(t - t_{0k} - i \Delta T_k) \right\rangle = \frac{P_c}{P_{\text{пик}}} = \frac{1}{Q}, \quad (5.31)$$

где P_c и $P_{\text{пик}}$ — средняя и пиковая мощности последовательности импульсов соответственно; Q — скважность импульсов.

Тогда с учетом (5.26) и (5.28), после некоторых преобразований получим

$$\delta_{i \text{ АИМ-ЧМ}}^2 = 2 k_{\text{п.к}}^2 Q N_0 F_{\text{вт}}^3 / 3 M_i^2 S_0^2 \Delta f_m^2. \quad (5.32)$$

Проводя подобные вычисления, можно определить потенциальную помехоустойчивость любой многоканальной системы передачи непрерывных сообщений.

Исследованию потенциальной помехоустойчивости многоканальных систем связи посвящен ряд работ. Однако теория этих вопросов разработана еще недостаточно, особенно в части установления более четких границ применимости полученных результатов и определения пороговых отношений сигнала к шуму на входе оптимальных приемников.

Каждому виду сигнала соответствует оптимальный способ приема и, следовательно, определенное значение потенциальной помехоустойчивости. Сравнивая ее с помехоустойчивостью реального способа приема такого сигнала, можно решить вопрос о целесообразности дальнейшего совершенствования этого реального способа.

Зная потенциальную помехоустойчивость приема различных сигналов, можно сравнивать их между собой и отбрасывать сигналы с низкой помехоустойчивостью. При этом можно совершенно не затрагивать вопросы реализации способов приема, поскольку реальные приемные устройства не могут обеспечить помехоустойчивость выше потенциальной.

Приведенные соображения показывают, что теория потенциальной помехоустойчивости, разработанная В. А. Котельниковым, является эффективным средством решения ряда важных задач передачи информации.

5.5. СТРУКТУРА ОПТИМАЛЬНЫХ ПРИЕМНИКОВ НЕПРЕРЫВНЫХ СООБЩЕНИЙ

Рассмотрим приложение результатов теории статистических решений к синтезу оптимальных приемников непрерывных сообщений для простого и хорошо изученного случая, когда помеха — белый гауссовский шум, сообщение — случайный процесс с ограниченным спектром, а искажения за счет неидеальности характеристик системы отсутствуют. Предполагается также, что параметры сигнала-переносчика $s_0(t)$ (частота и начальная фаза в отсутствие модуляции) априорно известны на приемной стороне. Такой случай часто называют приемом сигналов, известных точно [3].

Оптимальный прием АМ сигналов. В этом случае входная реализация имеет вид

$$y(t) = s_{AM}(t) + n(t) = S_0 [1 + M_a x(t)] \cos \omega_0 t + n(t).$$

Это выражение удобно записать так:

$$y(t) = s_0(t) + M_a x(t) s_0(t) + n(t). \quad (5.33)$$

Исследуя выражение (3.36), можно показать [2—4], что оптимальный приемник АМ сигналов обрабатывает входную реализацию $y(t)$ по правилу, которому соответствует структурная схема, показанная на рис. 5.5. В перемножителе реализация $y(t)$ умно-

жается на опорное напряжение $s_{оп}(t) = as_0(t)$, где a — постоянный коэффициент. Выходное колебание перемножителя содержит ряд составляющих, в том числе сообщение $x(t)$ и наложенный на него шум с равномерной спектральной плотностью. После прохождения этого колебания через идеальный фильтр нижних частот с полосой от нуля до F_n на выходе приемника появляется колебание $\hat{x}(t) = x(t) + \varepsilon(t)$, которое является оценкой переданного сообщения $x(t)$. Колебание $\varepsilon(t)$ — случайная помеха на выходе оптимального приемника.

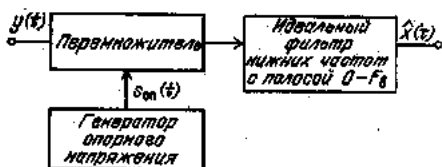


Рис. 5.5. Структурная схема оптимального приемника АМ сигналов, известных точно

Ее спектральная плотность определяется выражением (5.14), а величина СКО при оптимальном приеме АМ сигнала — выражением (5.15).

Если за счет неидеальности характеристик передающей части системы или среды распространения сигнал приобретает случайные медленные изменения начальной фазы, то

на приемной стороне необходимо применять обработку с использованием двух квадратурных опорных напряжений. В этом случае оптимальный приемник состоит из двух ветвей обработки с последующим объединением результатов. Такая двухканальная (квадратурная обработка) позволяет сделать приемное устройство инвариантным к начальной фазе сигнала.

Возможен и другой вариант оптимального приема, в котором производится оценка случайных изменений начальной фазы сигнала, в соответствии с которой фаза опорного колебания непрерывно подстраивается в соответствии с изменениями случайной фазы сигнала. В этом случае прием становится близким к приему сигнала, известного точно на приемной стороне.

Оптимальный прием сигналов с угловой модуляцией. Многочисленные исследования [1—4] показали, что задачу синтеза оптимального приемника сигналов с нелинейными видами модуляции удастся решить только при большом отношении сигнала к помехе на входе приемника, когда прием ведется в надпороговой области. Однако и в этом случае задача оказывается достаточно сложной. Опуская подробности ее решения, приведем лишь конечный результат и обсудим его.

В работе [4] показано, что оценка сообщения при оптимальном приеме представляет собой решение интегрального уравнения

$$\hat{x}(t) = \frac{2}{N_0} \int_{-\infty}^t R_x(t-\tau) \frac{\partial s[\tau, \hat{x}(\tau)]}{\partial \hat{x}(\tau)} \{y(\tau) - s[\tau, \hat{x}(\tau)]\} d\tau, \quad (5.34)$$

где $R_x(\tau)$ ($-\infty < \tau \leq t$) — функция корреляции стационарного гауссовского процесса, описывающего сообщение $x(t)$.

Применим это уравнение к угловой модуляции. В этом случае сигнал можно записать так (см. п. 4.3):

$$s[t, \theta(t)] = S_0 \cos[\omega_0 t + \beta\theta(t)], \quad (5.35)$$

где

$$\theta(t) = x(t), \beta = \Delta\varphi_m \text{ (для ФМ)} \quad (5.36)$$

$$\theta(t) = \int_0^t x(\tau) d\tau, \beta = \Delta\omega_m \text{ (для ЧМ)}.$$

Приняв во внимание (5.35) и учтя, что

$$\frac{\partial s(t)}{\partial \theta(t)} = S_0 \beta \sin[\omega_0 t + \beta\theta(t)], \quad (5.37)$$

уравнение (5.34), записанное для оценки $\hat{\theta}(t)$, после несложных преобразований можно привести к виду

$$\hat{\theta}(t) \approx \frac{2S_0\beta}{N_0} \int_{-\infty}^t R_x(t-\tau) \sin[\omega_0 \tau + \beta\hat{\theta}(\tau)] y(\tau) d\tau. \quad (5.38)$$

Приближенность этого выражения связана с тем, что при его получении отброшена составляющая, содержащая двойную частоту $2\omega_0$. Это допустимо, поскольку $\hat{\theta}(t)$ — низкочастотный процесс, спектр которого определяется спектром сообщения и, следовательно, не должен содержать составляющих с высокими частотами.

Уравнение (5.38) определяет алгоритм работы оптимального приемника сигналов с угловой модуляцией. Полученный алгоритм можно трактовать как свертку, описывающую линейную систему с импульсной реакцией $h(t) = R_x(t)$ и входным воздействием $u_{вх}(t) = \sin[\omega_0 t + \beta\hat{\theta}(t)] y(t)$. Соответствующая этому алгоритму структурная схема приемника показана на рис. 5.6. Особенность этой схемы в том, что она *следающая*. Оценка $\hat{\theta}(t)$ используется как управляющее воздействие для генератора, вырабатывающего модулированное опорное напряжение. В результате в схеме осуществляется непрерывное слежение за формируемой оценкой $\hat{\theta}(t)$, которая удовлетворяет условию максимума апостериорного распределения вероятностей (3.33). Усилитель с коэффициентом

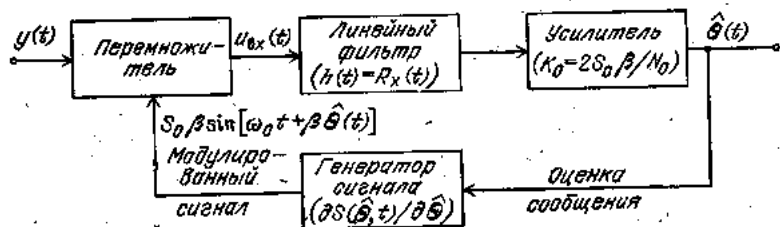


Рис. 5.6. Структурная схема оптимального приемника сигналов с угловой модуляцией

усиления $K_0 = 2S_0\beta/N_0$ выполняет роль масштабного множителя.

Схема, приведенная на рис. 5.6, соответствует оптимальному приему сигналов с ФМ, так как в этом случае оценка $\hat{\theta}(t)$ эквивалентна оценке сообщения $\hat{x}(t)$. При приеме сигналов с ЧМ для получения оценки $\hat{x}(t)$ эту схему следует дополнить дифференцирующим звеном, так как теперь $\hat{x}(t) = d\hat{\theta}(t)/dt$. Заметим, что возможность перехода от ФМ сигналов к ЧМ и обратно при помощи линейных фильтров в передающей и приемной частях системы позволяет получить ряд вариантов оптимальных приемников. Но принципиальная особенность всех их остается неизменной — все эти приемники являются следящими.

Развитие теории синтеза оптимальных приемников сигналов с угловой модуляцией обусловило появление многочисленных прикладных исследований и разработок инженерных схем, в которых используется принцип следящего приема. Некоторые особенности такого приема кратко рассмотрены в п. 5.6.

5.6. МЕТОДЫ ДЕМОДУЛЯЦИИ МОДУЛИРОВАННЫХ КОЛЕБАНИЙ

Выделение сообщений из модулированных колебаний называется *демодуляцией* и осуществляется специальным устройством — демодулятором. Демодулятор преобразует изменения модулируемого параметра $\Delta\Pi(t)$ в колебания $\hat{u}_x(t)$, соответствующие переданным сообщениям. Зависимость $\hat{u}_x(t) = f[\Delta\Pi(t), f]$ называется *демодуляционной характеристикой* (рис. 5.7). Ее можно записать в следующем виде:

$$\hat{u}_x(t) = K_d \Delta\Pi(t) = K_d \Delta\Pi_m \hat{x}(t). \quad (5.39)$$

Здесь $K_d = \text{tg } \alpha_d$ — чувствительность демодулятора (коэффициент передачи демодулятора). Остальные обозначения такие же, как и в гл. 4 (см. п. 4.2).

Чтобы в демодуляторе не возникали искажения, его демодуляционная характеристика должна быть линейной ($K_d = \text{const}$), а частотная характеристика — равномерной в пределах полосы частот, занимаемой спектром сообщения.

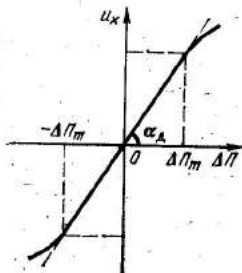


Рис. 5.7. Демодуляционная характеристика



Рис. 5.8. Структурная схема синхронного демодулятора АМ сигналов

Многие демодуляторы, применяемые на практике, появились в результате «естественного отбора» среди большого числа изобретений, инженерных находок и предложений. Теория оптимального приема позволила более четко разобраться с этим многообразием и указала пути создания демодуляторов, близких по своим показателям к оптимальным.

Опираясь на предыдущие результаты этой главы, рассмотрим некоторые прикладные методы демодуляции гармонических и импульсных модулированных колебаний.

Демодуляция АМ сигналов. В соответствии с теорией оптимального приема демодуляция АМ сигналов должна осуществляться в соответствии со схемой, приведенной на рис. 5.5. Наиболее сложным при реализации этой схемы является получение опорного напряжения. Обычно оно формируется с помощью местного генератора, подстраиваемого по частоте и начальной фазе к входящему сигналу с помощью системы автоподстройки. Демодулятор, в котором удается поддерживать условия $\omega_r \approx \omega_0$, $\varphi_r \approx \varphi_0$, называется *синхронным* или *когерентным* детектором. Его упрощенная схема приведена на рис. 5.8. В таком детекторе отсутствует нелинейное взаимодействие между сигналом и помехой, поэтому процесс детектирования входной смеси сигнала и помехи можно рассматривать отдельно (независимо) для сигнала и для помехи.

Это означает, что синхронный детектор ведет себя по отношению к огибающим входных колебаний как линейная система. Необходимо отметить, что отмеченное свойство синхронного детектора проявляется при условии, когда опорное напряжение существенно (в 10—15 раз) превышает максимальную величину входного сигнала.

Синхронное (когерентное) детектирование принципиально необходимо при демодуляции АМ сигналов с подавленной несущей, а также сигналов с ОБП. Для приема таких сигналов нужен высокостабильный местный генератор, с помощью которого формируется опорное напряжение. Несмотря на такое усложнение аппаратуры, сигналы с ОБП широко применяются в многоканальных радиорелейных и коротковолновых линиях телефонной связи.

При демодуляции обычных АМ сигналов, как правило, отказываются от учета случайных изменений начальной фазы сигнала (т. е. от синхронного детектирования) и применяют схемы демодуляторов, инвариантных к начальной фазе сигнала. Для простоты реализации оптимальную схему двухканального (квадратурного) демодулятора чаще всего заменяют подоптимальной, которая состоит из полосового фильтра с прямоугольной частотной характеристикой (с полосой $\Delta f_{\Phi} = 2F_{\text{в}}$ и центральной частотой, равной частоте сигнала f_0), детектора огибающей (амплитудного детектора) и усилителя низкой частоты с равномерной частотной характеристикой от нуля до $F_{\text{в}}$ (рис. 5.9).

При больших отношениях сигнала к шуму на входе приемника амплитудный (некогерентный) детектор почти не проигрывает

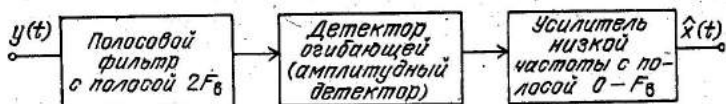


Рис. 5.9. Структурная схема подоптимального приемника АМ сигналов

синхронному (когерентному) в помехоустойчивости. Однако при уменьшении входного отношения сигнала к шуму проигрыш может стать заметным, так как в амплитудном детекторе начинает проявляться эффект подавления сигнала шумом. Исследования показывают, что этот эффект начинает заметно сказываться при $(P_c/P_{ш})_{вх} < 5$. Практически в системах с АМ минимальное отношение сигнала к шуму, при котором обеспечивается приемлемая точность передачи сообщений ($\delta_{общ} < 0,1$), значительно превышает указанную величину и с эффектом подавления можно не считаться.

Демодуляция ФМ и ЧМ сигналов. Теория оптимального приема показывает, что демодуляторы сигналов с угловой модуляцией представляют собой устройства, в которых осуществляется слежение за мгновенной частотой или фазой сигнала (см. п. 5.5).

Рассмотрим физическую сторону такой процедуры и особенности следящих демодуляторов. Изменения мгновенной частоты при угловой модуляции приводят к тому, что спектр мощности сигнала сосредоточен в области его мгновенной частоты (активный спектр) и занимает полосу, приблизительно равную удвоенному значению ширины спектра сообщения (рис. 5.10). В зависимости от величины мгновенного значения сообщения положение активного спектра сигнала на оси частот меняется в пределах общей полосы Δf_c , охватывающей все возможные ситуации. Таким образом, в оптимальном приемнике сигналов с угловой модуляцией полоса пропускания шириной $2F_B$ должна перестраиваться в соответствии с ожидаемым сообщением (оценкой $\hat{x}(t)$). Диапазон перестройки определяется максимальными отклонениями мгновенной частоты при модуляции. При ЧМ диапазон перестройки постоянен, а при ФМ он зависит от ширины спектра сообщения. Это обстоятельство является еще одной из причин, по которой при передаче непрерывных сообщений предпочитают применять ЧМ, а не ФМ сигналы.

В соответствии с изложенным оптимальный демодулятор ЧМ сигналов можно реализовать в виде устройства, структурная схе-

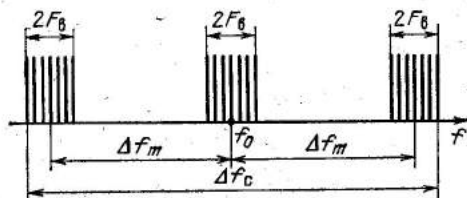


Рис. 5.10. Расположение активного спектра ЧМ сигнала на оси частот

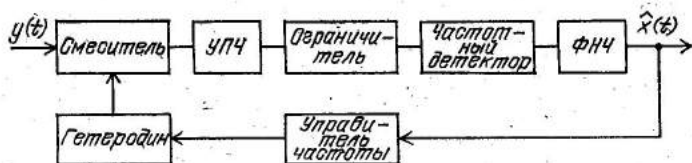


Рис. 5.11. Структурная схема следящего демодулятора ЧМ сигналов

ма которого показана на рис. 5.11. Полоса пропускания этой схемы, определяемая полосой УПЧ, равна $2F_v$ (или несколько шире).

Поскольку полезная информация при ЧМ заложена в изменениях частоты, все изменения других параметров необходимо устранить. В первую очередь это относится к огибающей ЧМ сигнала. Все ее изменения, обусловленные помехами и искажениями за счет неравномерности частотной характеристики УПЧ в пределах полосы пропускания, устраняются амплитудным ограничителем. Устранение амплитудных изменений ЧМ (или ФМ) сигналов является обязательным условием для демодуляции таких сигналов.

С выхода ограничителя ЧМ сигнал поступает на частотный детектор, который преобразует отклонения мгновенной частоты в пропорциональные им мгновенные значения напряжения. Так как это напряжение состоит не только из полезного сообщения, но и из шума, необходимо пропустить его через фильтр нижних частот (ФНЧ) с полосой пропускания, равной ширине спектра сообщения F_v . С выхода фильтра сообщение подается к получателю и на устройство, управляющее частотой местного генератора (гетеродина). По сути дела управитель частоты выполняет роль частотного модулятора и на выходе гетеродина формируется ЧМ сигнал, в котором роль сообщения выполняет его оценка $\hat{x}(t)$. Изменения мгновенной частоты гетеродина почти соответствуют изменениям мгновенной частоты входного ЧМ сигнала, т. е. частота гетеродина «следит» за частотой сигнала. Некоторая неточность такого слежения обусловлена действием помех и искажений, которое проявляется в различии между переданным сообщением $x(t)$ и полученной оценкой $\hat{x}(t)$.

Таким образом, в рассмотренном демодуляторе реализуется принцип слежения за мгновенной частотой ЧМ сигнала. Существует ряд разновидностей демодуляторов, использующих такой принцип: демодуляторы со следящим фильтром; демодуляторы с обратной связью по частоте; демодуляторы со следящим гетеродином и др. Все они относятся к устройствам, в которых важнейшей операцией является фазовая автоподстройка частоты (ФАП).

Следящие демодуляторы широко применяются в тропосферных, спутниковых, космических и некоторых других линиях связи, в которых трудно получить большие отношения сигнала к шуму и желательна понижать пороговое отношение ЧМ сигнала к шуму.

Реализация следящих демодуляторов приводит к заметному



Рис. 5.12. Структурная схема стандартного демодулятора ЧМ сигналов

усложнению приемников ЧМ сигналов. В тех случаях, когда в линии связи обеспечивается высокое отношение сигнала к шуму (например, в ЧМ радиовещании и т. п.) для упрощения приемников часто отказываются от следящего приема. Демодуляцию ЧМ сигналов в этих случаях выполняют демодулятором, схема которого приведена на рис. 5.12 и отличается от схемы на рис. 5.11 тем, что в ней отсутствует обратная связь по частоте, а полоса пропускания УПЧ выбирается в соответствии с формулой (4.26), т. е. так, чтобы охватить весь диапазон возможных значений мгновенной частоты (такой ЧМ демодулятор иногда называют стандартным). Помехоустойчивость ЧМ приемника со стандартным демодулятором ниже, чем со следящим. Это различие зависит от соотношения полос пропускания УПЧ сравниваемых приемников. Нетрудно показать, что при одинаковых значениях спектральной плотности шума N_0 и средней мощности сигнала на входах приемников, справедливо выражение

$$q_{\text{след}}/q_{\text{пр}} = \Delta f_{\text{пр}}/\Delta f_{\text{след}} \approx 1 + \Delta\Phi, \quad (5.40)$$

где $q_{\text{след}}$ и $q_{\text{пр}}$ — отношение средней мощности сигнала и средней мощности шума на выходе УПЧ для следящего и простого ЧМ приемников соответственно; $\Delta f_{\text{след}}$ и $\Delta f_{\text{пр}}$ — полосы пропускания УПЧ соответствующих приемников.

Из (5.40) следует, что для получения одного и того же отношения сигнала к шуму на выходе УПЧ в следящем приемнике требуется в $(1 + \Delta\Phi)$ раз меньшая мощность входного сигнала, чем в простом. На практике выигрыш оказывается меньше, так как в цепи обратной связи следящего приемника действует не сообщение $x(t)$, а его оценка $\hat{x}(t) = x(t) + \varepsilon(t)$ и слежение выполняется с некоторой погрешностью. Кроме того, для обеспечения устойчивости замкнутой следящей системы полосу пропускания УПЧ приходится выбирать несколько больше $2F_{\text{в}}$, что также снижает выигрыш. Опыт показывает, что в реальных приемниках со следящими ЧМ демодуляторами среднюю мощность входного сигнала удается снизить в 3—5 раз по сравнению с обычными ЧМ приемниками (при одинаковой точности передачи сообщений).

Детальное рассмотрение вопросов следящего приема ЧМ сигналов содержится в [5, 6].

В заключение отметим, что в некоторых многоканальных системах с частотным разделением каналов для демодуляции ЧМ поднесущих можно применять частотно-импульсные детекторы, работа которых основана на преобразовании ЧМ поднесущих в последовательности импульсов ЧИМ и их демодуляции фильтрами



Рис. 5.13. Структурная схема частотно-импульсного детектора

нижних частот. Структурная схема такого детектора показана на рис. 5.13. Поскольку скважность импульсов ЧИМ в данном случае невелика, их демодуляция с помощью фильтра нижних частот оказывается вполне эффективной. Важным достоинством подобных детекторов является то, что они не содержат индуктивностей и просты в настройке.

Демодуляция импульсных модулированных последовательностей. Рассмотрение спектральной структуры импульсных модулированных колебаний вида АИМ или ШИМ, проведенное в гл. 4, показало, что при таких видах модуляции появляется составляющая, пропорциональная модулирующему колебанию (сообщению $x(t)$). Эту составляющую можно выделить, подав модулированную импульсную последовательность на фильтр нижних частот с полосой пропускания $\Delta F_{\text{ФНЧ}} = F_{\text{в}}$, т. е. получить на выходе фильтра колебание, соответствующее исходному сообщению (см. рис. 4.12). Основываясь на результатах, изложенных в гл. 4, нетрудно показать, что при коэффициенте передачи фильтра $K_{\text{ф}} \approx 1$ в пределах полосы пропускания максимальное выходное напряжение составит:

$$(U_0/Q) M_a \text{ (для АИМ) и } (U_0/Q) (\Delta \tau_m/\tau) \text{ (для ШИМ)}. \quad (5.41)$$

Если $M_a \approx \Delta \tau_m/\tau$, то величина этих напряжений при АИМ и ШИМ одинакова. Поскольку при $Q \gg 1$ уровень сообщения на выходе фильтра получается малым (десятки милливольт), то фильтрацию обычно сочетают с усилением, чтобы получить необходимую величину выходного сообщения.

Таким образом, демодуляция импульсных модулированных последовательностей может быть непосредственно осуществлена с помощью линейной системы — фильтра нижних частот. Для АИМ и ШИМ такой метод демодуляции является основным.

Из-за неидеальности частотной характеристики фильтра нижних частот на его выход кроме переданного сообщения $u_x(t)$ попадают ослабленные составляющие $u_e(t)$ от других участков спектра. В результате выходное колебание $\hat{u}_x(t)$ отличается от переданного сообщения, т. е. появляются искажения, которые называются *демодуляционными*. Поскольку основной вклад в такие искажения вносит участок спектра, соответствующий области первой гармоники частоты повторения F_n , то при расчете демодуляционных искажений влиянием остальных участков спектра можно пренебречь. Более того, при оценке искажений можно ограничиться только учетом неотфильтрованной части амплитуды первой гармоники частоты повторения F_n , так как в ней заключена основная доля мощности рассматриваемого участка спектра. Не-

трудно показать, что эта мощность равна

$$P_{\text{искаж}} \approx 0,5 A_1^2 K_{\text{ФНЧ}}^2 (F_{\text{п}}) = 0,5 A_1^2 K_{\text{ФНЧ}}^2 (\mu F_{\text{в}}), \quad (5.42)$$

где $\mu = F_{\text{п}}/F_{\text{в}}$ — коэффициент следования импульсов.

С учетом изложенного средний квадрат ошибки демодуляционных искажений можно оценить по формуле

$$\delta_{\text{демод}}^2 = \langle u_e^2(t) \rangle / \langle u_x^2(t) \rangle \approx A_1^2 K_{\text{ФНЧ}}^2 (\mu F_{\text{в}}) / 2 \langle u_x^2(t) \rangle. \quad (5.43)$$

Для последовательностей АИМ и ШИМ с прямоугольными импульсами эта формула с учетом (4.32), (4.39), (4.44) принимает вид

$$\delta_{\text{демод}}^2 \approx (2 k_{\text{пх}}^2 / M^2) K_{\text{ФНЧ}}^2 (\mu F_{\text{в}}), \quad (5.44)$$

где $M = M_a$ для АИМ; $M = \Delta\tau_m / \tau$ для ШИМ; $k_{\text{пх}}$ — пикфактор сообщения.

Из (5.44) следует, что при заданном типе фильтра нижних частот существенно уменьшить демодуляционные искажения можно только за счет выбора более высокой частоты повторения $F_{\text{п}}$ (увеличения μ). Однако это влечет за собой ухудшение информационных показателей качества работы системы, так как при увеличении μ возрастает избыточность передаваемых сообщений и эффективность использования линии связи падает. Если увеличение избыточности нежелательно или недопустимо, то уменьшить величину демодуляционных искажений можно, применяя фильтр с частотной характеристикой, имеющей большой коэффициент прямоугольности. Однако, чем ближе характеристика фильтра к прямоугольной, тем сложнее и дороже фильтр. В реальных условиях при выборе демодулирующего фильтра приходится принимать компромиссное решение, которое учитывает обе отмеченные особенности.

Для последовательностей с ВИМ демодуляция с помощью фильтра нижних частот обычно не применяется. Это объясняется тем, что уровень полезного сообщения на выходе фильтра оказывается значительно меньше, чем при АИМ или ШИМ. Действительно, в соответствии с (4.48) при ВИМ напряжение на выходе фильтра нижних частот

$$u_x(t) = (U_0/Q) \Delta\tau_m x'(t). \quad (5.45)$$

С учетом (4.22) имеем

$$\max u_x \leq \frac{U_0}{Q} \Delta\tau_m \Omega_{\text{в}} = \frac{U_0}{Q} \frac{\Delta\varphi_{\text{ВИМ}}}{\mu}, \quad (5.46)$$

где $\Delta\varphi_{\text{ВИМ}} = \Delta\tau_m \Omega_{\text{п}}$; $\mu = \Omega_{\text{п}}/\Omega_{\text{в}}$.

Сравнение (5.46) и (5.41) дает

$$\max u_{x \text{ ВИМ}} / \max u_{x \text{ АИМ}} = \Delta\varphi_{\text{ВИМ}} / \mu M_a < 1. \quad (5.47)$$

Если учесть, что в большинстве случаев $\Delta\varphi_{\text{ВИМ}} \approx 0,1-0,2$; $\mu \geq 2-3$ и более, $M_a \approx 0,4-0,7$, то в соответствии с (5.47) уровни напряжений могут различаться примерно на порядок. Это озна-

чает, что для обеспечения необходимой величины выходного напряжения и заданной ошибки демодуляционных искажений при ВИМ потребовалось бы применение дополнительного усиления и более сложных фильтров, чем при АИМ или ШИМ. Кроме того, нужно иметь в виду, что для ФИМ (наиболее часто применяемой разновидности ВИМ) напряжение на выходе демодулирующего фильтра пропорционально производной исходного сообщения (см. (5.45)). Поэтому (чтобы сделать напряжение пропорциональным сообщению) демодулирующий фильтр необходимо дополнять интегрирующим звеном.

Недостатки непосредственной демодуляции ВИМ приводят к тому, что она практически не применяется. Демодуляцию ВИМ обычно осуществляют косвенным методом: сначала преобразуют последовательность ВИМ в последовательность ШИМ или АИМ, а затем уже проводят демодуляцию. Технически наиболее удобно и целесообразно преобразовывать ВИМ в ШИМ — эта операция проста и может быть выполнена одновременно с разделением каналов.

5.7. РАЗДЕЛЕНИЕ КАНАЛОВ ПРИ МНОГОКАНАЛЬНОЙ ПЕРЕДАЧЕ НЕПРЕРЫВНЫХ СООБЩЕНИЙ

Принцип разделения каналов. Ошибки при многоканальной передаче сообщений. Для получения сообщений отдельных каналов $\hat{x}_i(t)$ принятое многоканальное сообщение $\hat{x}_{\text{мн}}(t)$ нужно обработать так, чтобы выделить отдельные каналные сигналы $\hat{s}_i(t)$, а затем демодулировать их.

Схема устройства разделения каналов приведена на рис. 5.14. Многоканальное сообщение $\hat{x}_{\text{мн}}(t)$ с выхода общего демодулятора приемника поступает одновременно на N входов канальных трактов (каналов), в состав каждого из которых входит селектор канала (СК) и канальный демодулятор (КД). Селектор предназначен для выделения только «своего» канального сигнала

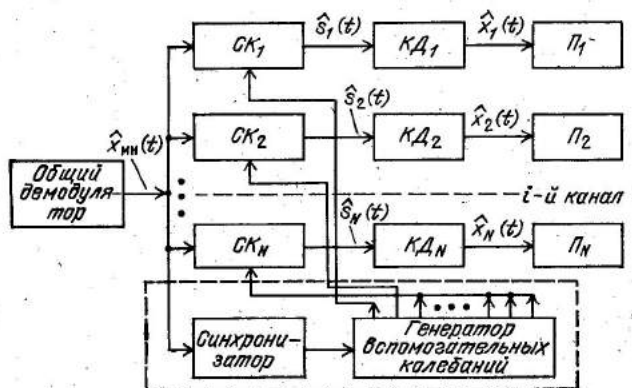


Рис. 5.14. Структурная схема устройства разделения каналов

и не реагирует на другие сигналы. Такие избирательные свойства можно получить, реализуя селектор в виде согласованного фильтра, или в виде коррелятора. В последнем случае необходимо иметь набор вспомогательных колебаний, каждое из которых выполняет роль опорного напряжения в соответствующем канальном корреляторе. Чтобы эти колебания не отличались от поднесущих (гармонических или импульсных — в зависимости от вида уплотнения каналов в передающей части), формирующий их генератор синхронизируется сигналами, специально передаваемыми по отдельному каналу синхронизации.

Выбор того или иного способа реализации канального селектора зависит от способа уплотнения каналов, вида модуляции поднесущих и несущей, а также от других факторов. На этих вопросах остановимся ниже при рассмотрении отдельных видов разделения каналов.

После разделения канальные сигналы $\hat{s}_i(t)$ демодулируются в канальных демодуляторах и сообщения $\hat{x}_i(t)$ поступают к получателям Π_i . Заметим, что разделение и демодуляция в некоторых случаях могут осуществляться одним устройством, совмещающим функции селектора и демодулятора.

Операции уплотнения и разделения каналов нельзя реализовать идеально, и они всегда сопровождаются некоторыми искажениями, обусловленными неидеальностью характеристик общего тракта многоканальной системы: нелинейностью модуляционных и демодуляционных характеристик, ограниченностью полосы пропускания общего тракта; неравномерностью частотных и нелинейностью фазовых характеристик тракта в пределах полосы пропускания¹. Эти причины приводят к искажению многоканальных сообщений, в результате чего на выходе каждого селектора помимо канального сигнала появляется дополнительное мешающее колебание, называемое *перекрестной помехой*. В отличие от аддитивных помех, которые не зависят от сигнала, перекрестные помехи появляются только при наличии сигнала. Влияние искажений, возникающих в общем тракте, а также действие аддитивных помех в канале приводят к тому, что принятое многоканальное сообщение $\hat{x}_{\text{пр}}(t)$ отличается от переданного $x_{\text{пр}}(t)$. В результате сообщения $\hat{x}_i(t)$ на выходе отдельных каналов также отличаются от переданных $x_i(t)$, т. е. принимаются с некоторой ошибкой. При анализе и расчетах удобно разделять эту ошибку на составляющую, обусловленную действием аддитивных помех и составляющую, обусловленную искажениями. При реальных требованиях к точности передачи эти составляющие можно считать независимыми и тогда средний квадрат ошибки передачи по i -му каналу равен (см. (1.6)).

$$\delta_i^2 = \delta_{i\text{ш}}^2 + \delta_{i\text{искаж}}^2,$$

¹ Напомним, что общим (или групповым) трактом многоканальной системы называют ту ее часть, которая заключена между входом общего модулятора и выходом общего демодулятора.

где $\delta_{i\text{ш}}$ и $\delta_{i\text{искаж}}$ — составляющие ошибки, обусловленные действием аддитивных помех и искажений соответственно.

При высоких требованиях к точности передачи сообщений необходимо стремиться уменьшать обе составляющие ошибки. Практика показывает, что при правильном выборе параметров многоканальной системы можно получить величину $\delta_{i\text{искаж}}$ (1—3)%. Дальнейшее снижение этой составляющей ошибки сопряжено с большими техническими трудностями и обычно оказывается нецелесообразным.

Для уменьшения шумовой составляющей ошибки необходимо увеличивать отношение сигнала к шуму на входе приемника. Можно принять, что это отношение достаточно, если выполняется условие $\delta_{i\text{ш}}^2 \leq 0,1\delta_{i\text{искаж}}^2$ ($\delta_{i\text{ш}} \leq 0,3\delta_{i\text{искаж}}$). При $\delta_{i\text{искаж}}$ (1—3)% этому условию соответствуют значения $\delta_{i\text{ш}} \approx (0,3—1) \cdot 10^{-2}$. Для расчета необходимого отношения сигнала к шуму на входе приемника реальной системы (т. е. определения ее помехоустойчивости) при заданной величине $\delta_{i\text{ш}}$ можно пользоваться формулами потенциальной помехоустойчивости, приведенными в п. 5.4, так как при высоких требованиях к точности передачи непрерывных сообщений помехоустойчивость реального приема практически не отличается от помехоустойчивости оптимального приема [1, 3, 7].

В заключение отметим, что в аналоговых многоканальных системах передачи непрерывных сообщений практически достижимая точность передачи (величина СКО) обычно находится в пределах $(1—3) \cdot 10^{-2}$, т. е. составляет (1—3)%.

При необходимости передавать непрерывные сообщения с более высокой точностью следует применять цифровые методы передачи.

Рассмотрим кратко особенности основных видов разделения каналов при передаче непрерывных сообщений.

Частотное разделение каналов (ЧРК). При частотном разделении каналные селекторы представляют собой полосовые фильтры, полоса каждого из которых рассчитана на пропускание спектра «своей» модулированной поднесущей. Для ослабления влияния других поднесущих (особенно соседних) фильтры должны иметь частотные характеристики с большим коэффициентом прямоугольности. Демодуляция поднесущих обычно выполняется некогерентными демодуляторами: при АМ — амплитудными детекторами, а при ЧМ — частотными дискриминаторами. Такие демодуляторы значительно проще и дешевле когерентных, а проигрыш в помехоустойчивости в ряде случаев невелик и с ним можно примириться.

Если в первой ступени рассматриваемых систем применяется ОБП (например, при передаче телефонных сообщений), то в канальных трактах принципиально необходимым является когерентное детектирование. В этом случае с помощью специального синхросигнала на приемной стороне формируются вспомогательные колебания с сеткой частот, аналогичной сетке частот поднесущих

в передающей части системы. Эти колебания используются как опорные напряжения в канальных демодуляторах ОБП.

Выходное напряжение каждого демодулятора подается на фильтр нижних частот с полосой пропускания, равной ширине спектра сообщения данного канала, а затем полученное сообщение поступает к получателю.

Анализ искажений в системах с ЧРК связан с большими математическими трудностями, поэтому обычно вводят следующие упрощающие задачу идеализации и допущения:

1) основная доля искажений обусловлена неидеальностью характеристик группового тракта. Искажения, возникающие в отдельных канальных трактах, значительно меньше и их можно не учитывать;

2) искажения в групповом тракте можно разделить на две независимые составляющие: искажения в высокочастотной части тракта (искажения модулированного сигнала) и искажения в видеотракте (искажения многоканального сообщения);

3) основной причиной искажений сигнала в высокочастотной части тракта является неравномерность частотной и нелинейность фазовой характеристик усилителя промежуточной частоты приемника в пределах полосы частот, занимаемой спектром модулированного сигнала¹;

4) основной причиной искажений в видеотракте является нелинейность его модуляционной и демодуляционной характеристик, которые при анализе можно рассматривать как единую модуляционно-демодуляционную характеристику.

Теоретические исследования показывают, что перечисленные выше допущения значительно упрощают анализ и обеспечивают приемлемую для инженерных приложений точность расчетов [7]. В частности, при правильно выбранной полосе пропускания высокочастотной части общего тракта искажения в системах с ЧРК в основном определяются нелинейностью модуляционно-демодуляционной характеристики системы, которая показывает зависимость мгновенных значений многоканального сообщения на выходе общего тракта от соответствующих значений такого сообщения на входе тракта. По сути дела эту зависимость можно рассматривать как амплитудную характеристику общего тракта². При расчетах часто ограничиваются оценкой искажений, обусловленных только этой характеристикой.

Для количественной оценки искажений модуляционно-демодуляционная характеристика аппроксимируется либо полиномом вида

$$u_{\text{вых}} = a_1 u_{\text{вх}} + a_2 u_{\text{вх}}^2 + a_3 u_{\text{вх}}^3 + \dots \quad (5.48)$$

¹ Как правило, во второй ступени модуляции многоканальных систем передачи непрерывных сообщений применяется частотная модуляция. Для ЧМ сигналов нелинейность амплитудной характеристики УПЧ практически не влияет на точность передачи.

² В соответствии с изложенными выше допущениями считается, что остальные причины искажений в общем тракте отсутствуют.

где a_1, a_2, a_3 — коэффициенты, величина и знак которых зависят от вольт-амперных характеристик электронных устройств тракта; либо линейно-ломанной зависимостью вида

$$u_{\text{вых}} = \begin{cases} a_1 u_{\text{вх}}, & |u_{\text{вх}}| \leq U_{\text{лм}}, \\ a_1 U_{\text{лм}}, & |u_{\text{вх}}| > U_{\text{лм}}. \end{cases} \quad (5.49)$$

Вид характеристик, соответствующих аппроксимациям (5.48) (кривая 1) и (5.49) (кривая 2) показан на рис. 5.15.

Приняв во внимание, что многоканальное сообщение на входе общего тракта представляет собой сумму модулированных поднесущих, сообщение на выходе этого тракта можно представить так:

$$u_{\text{МН}}(t) = a_1 \sum_{i=1}^N s_i(t) + e_{\text{иск}}(t), \quad (5.50)$$

где $e_{\text{иск}}(t)$ — перекрестная помеха, обусловленная нелинейностью характеристики тракта.

Основываясь на результатах приближенного анализа [7], можно показать, что при оценке искажений спектральная плотность перекрестной помехи в пределах полосы частот, занимаемой многоканальным сообщением, приблизительно постоянна и равна величине

$$N_0 \text{ иск} \approx 1,6 a_3^2 P_{\text{МН}}^2 N / f_{\text{пН}} \quad (5.51)$$

— для полиномиальной аппроксимации:

$$N_0 \text{ иск} \approx \frac{0,6 P_{\text{МН}}}{k_{\text{п.МН}}^2 f_{\text{пН}}} \exp\left(-\frac{k_{\text{п.МН}}^2}{2}\right) \quad (5.52)$$

— для линейно-ломанной аппроксимации.

В приведенных формулах $P_{\text{МН}}$ и $k_{\text{п.МН}}$ — средняя удельная мощность и пикфактор многоканального сообщения; $f_{\text{пН}}$ — верхняя поднесущая частота; N — число каналов.

При определении ошибки на выходе отдельного канала перекрестную помеху можно рассматривать как нормальный случайный процесс со спектральной плотностью, определяемой выражением (5.51) или (5.52), и пользоваться формулами п. 5.3.

Детальное исследование искажений в системах с ЧРК содержится в [7, 8].

Временное разделение каналов (ВРК). При временном разделении каналные селекторы (рис. 5.14) представляют собой устройства совпадения по времени (временные селекторы). На один вход такого устройства подается многоканальное сообщение

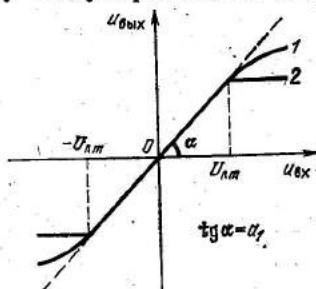


Рис. 5.15. Аппроксимации амплитудных характеристик группового тракта

(последовательность информационных импульсов всех каналов), а на другой — специально сформированная вспомогательная периодическая последовательность импульсов (стробов). Для каждого канала последовательность стробов формируется генератором вспомогательных колебаний, работа которого синхронизируется с работой генератора канальных импульсов, находящегося в передающей части системы. Для устойчивой синхронизации канал синхронизации должен обладать значительно более высокой помехоустойчивостью по сравнению с любым информационным каналом.

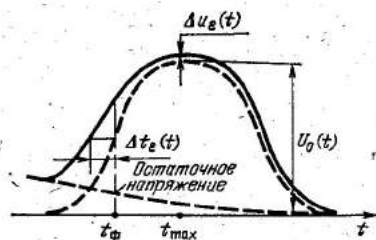


Рис. 5.16. Межканальные искажения в системах с временным уплотнением

Если информационный импульс совпадает по времени со стробом, то он проходит на выход данного временного селектора, если совпадения нет, то селектор оказывается закрытым. В результате такой операции многоканальное сообщение разделяется на импульсные последовательности, соответствующие отдельным каналам. Для правильного временного разделения помимо надежной синхронизации необходимо, чтобы длительность стробов была равна интервалу времени, отводимому на отдельный канал с учетом модуляционных изменений длительности импульса (при ШИМ) или его положения (при ВИМ) и канальных защитных интервалов (см. гл. 4).

В системах с временным разделением модулируемый параметр сигнала изменяется скачкообразно в соответствии с импульсным характером многоканального сообщения. При конечной полосе пропускания УПЧ и видеосуилителя это приводит к появлению заметных переходных процессов и наложению остаточных колебаний от предыдущих импульсов на последующие. В результате появляются *межканальные искажения*, которые после общего демодулятора проявляются в виде наложения остаточного напряжения переходных процессов на огибающую каждого импульса (см. рис. 5.16). Такое наложение приводит к некоторому сдвигу импульса во времени $\Delta t_\epsilon(t)$ и изменению величины его амплитуды $\Delta u_\epsilon(t)$.

Рассмотренные искажения обусловлены конечной полосой пропускания общего тракта системы в области верхних частот (в литературе эти искажения иногда называют искажениями 1-го рода). Кроме того, возможны искажения, связанные с «завалом» частотной характеристики общего тракта в области нижних частот (искажения 2-го рода). Современная элементная база позволяет строить схемы видеосуилителей без переходных емкостей, т. е. пропускать все составляющие спектра, включая постоянную составляющую. Поэтому искажения 2-го рода можно полностью исключить. Для систем с ВРК паразитные изменения $\Delta t_\epsilon(t)$ и $\Delta u_\epsilon(t)$ в основном определяются искажениями 1-го рода и по-

сравнению с ними другие виды искажений в общем тракте можно не учитывать.

Строгий анализ межканальных искажений весьма сложен, так как требуется учет конкретных частотных и фазовых характеристик УПЧ, вида модуляции и т. п.

Влияние переходных искажений в системах с ВРК обычно оценивается с помощью следующих соотношений:

$$\delta_i^2 = \overline{\Delta u_e^2} / \langle \Delta u_{xi}^2(t) \rangle \quad \text{при АИМ}, \quad (5.53)$$

$$\delta_i^2 = \overline{\Delta t_e^2} / \langle \Delta \tau_{xi}^2(t) \rangle \quad \text{при ВИМ или ШИМ}, \quad (5.54)$$

где $\Delta u_{xi}(t)$ и $\Delta \tau_{xi}(t)$ — полезные изменения амплитуды i -го импульса (при АИМ), его фронта (при ШИМ) или положения (при ВИМ); Δu_e и Δt_e — случайные изменения амплитуды и положения фронта i -го импульса под действием остаточного переходного напряжения от предыдущих импульсов. Усредненные значения $\overline{\Delta u_e^2}$ и $\overline{\Delta t_e^2}$ определяются в те моменты времени, когда влияние переходных искажений для данного вида модуляции оказывается наименьшим. Для АИМ этот момент соответствует максимальному значению i -го импульса (t_{\max} на рис. 5.16), а для ВИМ или ШИМ — максимальной крутизне фронта этого импульса (t_ϕ на рис. 5.16) ¹.

В инженерных приложениях часто ограничиваются еще более приближенной оценкой искажений, когда учитывается максимальное влияние только одного предыдущего импульса. Для получения такой оценки нужно определить значения Δu_e и Δt_e , соответствующие применяемому УПЧ, а также учесть, что

$$\langle \Delta u_{xi}^2(t) \rangle = U_0^2 M_{ai}^2 \langle x_i^2(t) \rangle = U_0^2 M_{ai}^2 / k_{\pi xi}^2, \quad (5.55)$$

$$\langle \Delta \tau_{xi}^2(t) \rangle = \Delta \tau_{mi}^2 \langle x_i^2(t) \rangle = \Delta \tau_{mi}^2 / k_{\pi xi}^2. \quad (5.56)$$

Приближенный анализ, соответствующий такому случаю, позволяет получить следующие выражения средних квадратов ошибок за счет переходных искажений

$$\delta_{i \text{ АИМ}} \approx (k_{\pi xi}^2 / M_{ai}^2) h_{\text{УПЧ}}^2(t_{\max}), \quad (5.57)$$

$$\delta_{i \text{ ВИМ ШИМ}}^2 \approx (k_{\pi xi}^2 / \Delta \tau_{mi}^2 \Delta f_{\text{экв}}^2) h_{\text{УПЧ}}^2(t_\phi), \quad (5.58)$$

где $h_{\text{УПЧ}}$ — переходная характеристика УПЧ; $t_{\max} \approx 2\tau + \tau_s$, $t_\phi \approx 1,5\tau + \tau_s$, $\Delta f_{\text{экв}}$ — эквивалентная полоса пропускания УПЧ; M_{ai} — коэффициент модуляции в i -м канале; $\Delta \tau_{mi}$ — максимальный полезный сдвиг фронта импульса (при ШИМ) или положения импульса (при ВИМ) в i -м канале. Отсчет времени в (5.57) и (5.58) ведется от начала предыдущего импульса (см. рис. 5.16).

¹ Напомним, что модулирующие колебания (сообщения) считаются эргодическими процессами, прямая черта означает усреднение по совокупности, а косые скобки — усреднение по времени.

Уменьшить влияние переходных искажений можно несколькими способами:

- 1) расширением полосы пропускания УПЧ (уменьшением длительности переходных процессов);
- 2) увеличением защитного канального промежутка между соседними каналами;
- 3) применением схем компенсации остаточных переходных напряжений.

Первые два способа наиболее просты, но связаны с ухудшением показателя использования полосы пропускания системы (растут удельные затраты полосы); при реализации третьего способа требуется заметное усложнение приемного устройства системы. Поэтому выбор того или иного способа (или их комбинации) зависит от особенностей конкретной решаемой задачи.

В заключение отметим, что в случае применения ВИМ временные селекторы одновременно выполняют и роль преобразователей ВИМ в ШИМ. Обычно такие селекторы выполняются в виде триггеров с двумя устойчивыми состояниями. Под действием импульса ВИМ, совпавшего по времени со стробом, триггер переходит в другое устойчивое состояние, а затем задним фронтом строба вновь возвращается в исходное состояние, которое сохраняется до появления следующего импульса ВИМ данного канала. Дополнительные сведения по вопросам разделения каналов в многоканальных системах можно найти в [7, 8].

5.8. ОСОБЕННОСТИ СХЕМНОЙ РЕАЛИЗАЦИИ НЕКОТОРЫХ УСТРОЙСТВ АНАЛОГОВЫХ СИСТЕМ ПЕРЕДАЧИ СООБЩЕНИЙ

Для многих современных систем передачи информации характерна тенденция постепенного превращения передающей и приемной частей в специализированные вычислительные устройства. Большую роль в этом играют быстро развивающиеся методы цифровой обработки информации. Вместе с тем говорить об отмирании классических способов построения передающих и приемных устройств систем передачи информации пока нет оснований. Особенно это касается аналоговых систем передачи непрерывных сообщений. Такие системы еще долгие годы будут широко применяться.

Аналоговые системы передачи информации состоят из разнообразных элементов и устройств, в том числе: усилителей, фильтров, интеграторов, ограничителей, перемножителей, ключей, детекторов, синтезаторов частоты, модуляторов и т. д.

Элементной базой современной аппаратуры таких систем являются *интегральные устройства*, выполненные в виде законченных функциональных узлов, разработанных специально для данного типа систем связи, а также устройства, представляющие отдельные многоцелевые интегральные блоки общего назначения.

Многоцелевые блоки можно рассматривать как универсальные

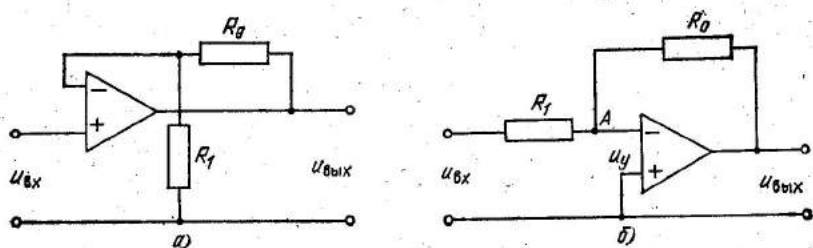


Рис. 5.17. Схемы неинвертирующего и инвертирующего усилителей

базовые устройства, на основе которых реализуются различные линейные и нелинейные операции: усиление колебаний, модуляция, перемножение колебаний и т. п. Типичными примерами таких базовых устройств являются операционные и дифференциальные усилители, аналоговые перемножители и балансные модуляторы.

Ниже кратко рассмотрены особенности универсальных базовых блоков, а также вопросы реализации некоторых устройств, входящих в состав систем передачи информации.

Операционные усилители и некоторые их применения. Идеальный операционный усилитель представляет собой схему усиления напряжения с бесконечно большим коэффициентом усиления, бесконечным входным и нулевым выходным сопротивлением. Реальные операционные усилители в интегральном исполнении имеют коэффициент усиления от нескольких сотен до нескольких десятков тысяч, входное сопротивление — сотни килоом, выходное — несколько ом и менее.

Операционные усилители широко применяются для реализации ряда математических операций (умножения, интегрирования, извлечения квадратного корня и т. д.), а также для построения схем различного функционального назначения (модуляторов, демодуляторов, избирательных усилителей и т. д.). Перечисленные операции реализуются путем введения обратных связей с помощью различных внешних цепей, подключаемых к операционному усилителю. Для существенного ослабления влияния дестабилизирующих факторов (изменений температуры, питающих напряжений) входной каскад операционного усилителя выполняют по дифференциальной схеме.

Некоторые примеры использования операционных усилителей показаны на рис. 5.17.

Схема, приведенная на рис. 5.17,а, обеспечивает хорошее разделение входа и выхода и с помощью цепи обратной связи позволяет усиливать входные напряжения без инвертирования. При замыкании резистора R_0 ($R_0=0$) схема выполняет функцию эмиттерного повторителя.

Для усиления с инвертированием входного напряжения принимают схему, представленную на рис. 5.17,б. При глубокой от-

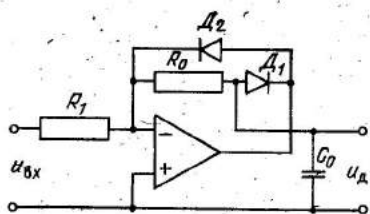


Рис. 5.18. Схема линейного амплитудного детектора с большим динамическим диапазоном входных сигналов

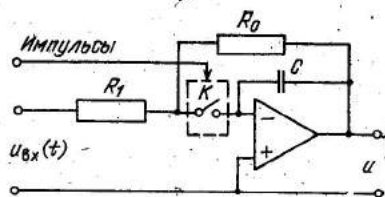


Рис. 5.19. Устройство запоминания отсчета

рицательной обратной связи управляющее напряжение u_y мало и потенциал точки A близок к потенциалу земли. Поэтому коэффициент усиления схемы близок к величине $k_u \approx -R_0/R_1$, а входное сопротивление близко к R_1 .

Заменяв в схеме рис. 5.17,б резистор R_1 конденсатором C_1 , можно реализовать операцию дифференцирования входного напряжения, а при замене резистора R_0 конденсатором C_1 — операцию интегрирования. Естественно, что для выполнения указанных операций следует правильно выбирать соотношение между постоянной времени схемы и скоростью изменения входного напряжения.

Как известно, диодные детекторы плохо детектируют сигналы низкого уровня. Однако развитие интегральных схем, миниатюризация аппаратуры и снижение мощностей преобразуемых сигналов приводят к необходимости детектирования сигналов низких уровней. На основе операционного усилителя можно создать линейный амплитудный детектор с большим динамическим диапазоном входных сигналов (от нескольких милливольт до нескольких вольт). Одна из схем такого детектора приведена на рис. 5.18. В этой схеме детектирование выполняется диодом D_1 с нагрузкой R_0 . Диод D_2 служит для замыкания цепи постоянного тока. Конденсатор C_0 совместно с нагрузкой R_0 образует фильтр нижних частот, который пропускает на выход только демодулированное сообщение и отфильтровывает все высокочастотные продукты детектирования.

При передаче непрерывных сообщений импульсными методами часто возникает необходимость запоминания мгновенных значений сообщения в точках отсчетов и т. п. Такая задача может быть решена с помощью схемы, показанной на рис. 5.19. При замыкании ключа K на короткое время ($\tau \ll 1/F_v$) конденсатор C заряжается до некоторого напряжения, величина которого пропорциональна входному напряжению. При размыкании этого ключа напряжение на конденсаторе C останется практически неизменным, так как разряд конденсатора через высокоомный вход незначителен. Низкое выходное сопротивление схемы позволяет подключить к ее выходу внешнюю нагрузку без опасения нару-

шить работу устройства. При некотором усложнении схемы ее можно использовать для последовательного запоминания отдельных отсчетов.

Частотно-избирательные схемы. При построении частотно-избирательных устройств в интегральном исполнении широко применяются методы *активной фильтрации*. Такие методы широко позволяют получить устройства с необходимыми избирательными свойствами без использования индуктивности. Основными элементами активных фильтров являются операционные усилители, резисторы и конденсаторы. Существуют два основных класса активных фильтров: с использованием линейной обратной связи и с использованием фазовой автоподстройки частоты (ФАП).

Применяя различные виды линейной обратной связи, можно получить стабильно работающие активные фильтры с высокой избирательностью. Однако следует иметь в виду некоторые недостатки, свойственные таким фильтрам:

- 1) в связи с трудностями обеспечения стабильной устойчивой работы применяемый диапазон частот активных фильтров обычно не превышает нескольких десятков килогерц (не более 100 кГц);
- 2) добротность фильтра и его центральная частота существенно зависят от стабильности коэффициента усиления усилителя и абсолютных величин элементов, образующих цепь обратной связи;

- 3) сравнительно высокая сложность и стоимость, связанные с высокими требованиями к точности абсолютных величин элементов активного фильтра и необходимостью применения гибридных интегральных схем с дополнительной подстройкой после изготовления.

Некоторые из перечисленных недостатков удается преодолеть в активных фильтрах с использованием ФАП.

Аналоговые перемножители и балансные модуляторы. Аналоговые перемножители так же, как и операционные усилители, являются универсальными базовыми устройствами, с помощью которых можно выполнять такие математические операции, как умножение, деление, возведение в квадрат, усреднение квадрата величины, извлечение квадратного корня и т. д. Перечисленные возможности обычно реализуются сочетанием перемножителей с операционными усилителями. Примеры выполнения некоторых математических операций показаны на рис. 5.20. В схемах, приведенных на рис. 5.20, *в* и *г*, для исключения возможности самовозбуждения надо обеспечивать положительный знак выходного напряжения ($y > 0$ для любого момента времени), поэтому все мгновенные значения входного напряжения должны быть отрицательными ($z < 0$ для любого момента времени).

Кроме реализации математических операций аналоговые перемножители применяются также и при построении различных модуляторов и демодуляторов. Наиболее широко в таких устройствах используются перемножители типа балансных модуляторов. В современной электронной аппаратуре находит применение ряд

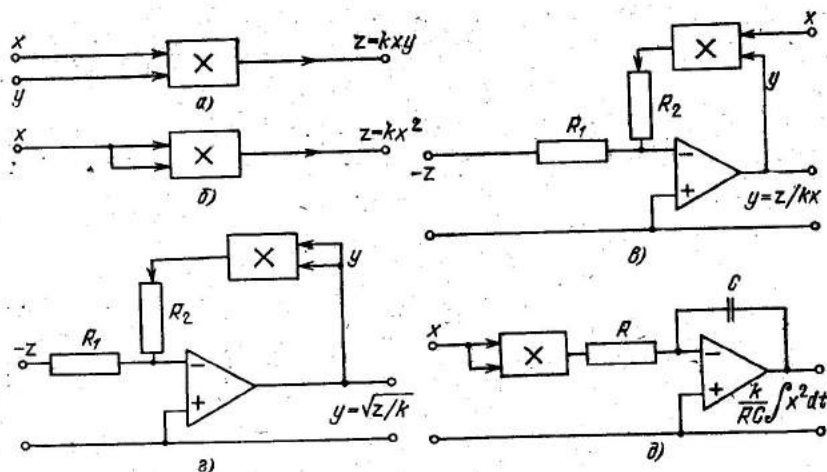


Рис. 5.20. Схемы выполнения математических операций

разновидностей балансных модуляторов в интегральном исполнении. Особенностью таких модуляторов является то, что при идеальной симметрии схемы выходное колебание не содержит составляющую несущей частоты. Поскольку в реальных схемах симметрия неидеальна и имеется некоторый разбаланс, полного подавления несущей не происходит. Для современных балансных модуляторов такое подавление составляет не менее 50 дБ.

На основе балансных модуляторов в системах аналоговой связи можно создавать различные устройства: модуляторы ОБП и АМ—ПН сигналов, умножители частоты, амплитудные, частотные и фазовые демодуляторы. Использование балансных модуляторов в качестве функциональных элементов интегральных схем позволило эффективно решить ряд сложных технических задач в цветном телевидении, стереофоническом радиовещании и других областях радиотехники.

Вопросы применения различных аналоговых интегральных схем в электронных устройствах изложены в работах [9—11].

КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАЧИ

1. От каких величин зависит линия сигналов при АМ и какова ее максимальная длина?
2. Что такое линия сигналов при ФМ и ЧМ и какими факторами определяется ее длина?
3. Поясните понятие «малая» помеха. При каких условиях помеху можно считать малой?
4. Что такое пороговый эффект при приеме сигналов? Каков механизм возникновения аномальных ошибок?
5. Нарисуйте примерный вид зависимости $\delta = f(\Delta\varphi_m)$ для ФМ (ЧМ) сигнала при различных фиксированных отношениях $N_0/P_{с.вх}$ и поясните их.

6. Покажите, что помеха на выходе оптимального приемника имеет нормальное распределение, если причиной ее появления является гауссовский шум на входе.
7. Что такое потенциальная помехоустойчивость и что она характеризует?
8. Покажите, что формулу (5.19) можно представить в виде

$$\delta_{\Phi M}^2 = \frac{k_{\text{пх}}^2}{2 \Delta \Phi_m^2 (1 + \Delta \Phi_m)} \left(\frac{P_{\text{п}}}{P_{\text{с}}} \right)_{\text{вх}}$$

9. Покажите, что колебания $x(t)$ и $x'(t)$ ортогональны и справедливо соотношение

$$\langle [x(t)/x'(t)]^2 \rangle = \langle x^2(t) \rangle \langle [1/x'(t)]^2 \rangle = 1/\Omega^2.$$
10. Рассчитайте и постройте график зависимости (5.24). Найдите значения $(P_{\text{с}}/P_{\text{ш}})_{\text{пор}}$ для следующих значений $\Delta \Phi_m$: 1, 2, 4, 5, 8, 10. Расчеты проведите для $P_{\text{ан}} = 10^{-4}$; 10^{-5} и сообщений с нормальным и равновероятными законами распределений.
11. Найдите потенциальную помехоустойчивость приема следующих видов сигналов: АМ; АИМ; АМ—АМ; АИМ—АМ; ВИМ—АМ; ВИМ—ЧМ.
12. Запишите выражения для колебаний на выходах перемножителя и фильтра нижних частот оптимального приемника АМ сигналов, рассмотренного в п. 5.5.
13. Нарисуйте схему оптимального приемника АМ сигналов, инвариантного к случайному значению начальной фазы сигнала.
14. Нарисуйте схему оптимального приемника АМ сигналов с оценкой случайных изменений начальной фазы сигнала. Объясните работу такого приемника.
15. Нарисуйте схемы оптимальных приемников сигналов с ЧМ и ФМ. Объясните работу этих приемников.
16. Как осуществляется демодуляция АМ и ЧМ сигналов в реальных приемниках?
17. Что такое следящие демодуляторы? Какие особенности характерны для них?
18. Как осуществляется демодуляция АИМ, ШИМ, ВИМ?
19. Назовите причины искажений при демодуляции импульсных модулированных последовательностей.
20. Рассчитайте ошибку демодуляционных искажений при демодуляции АИМ с помощью ФНЧ, если $k_{\text{пх}} = 3$; $M_a = 0,6$;

$$K_{\text{ФНЧ}}(\Omega) = 1/[1 + (\Omega \tau)^2]^2, \quad \mu = F_{\text{п}}/F_{\text{в}} = 5.$$
21. Решите предыдущую задачу для случая ШИМ с $\Delta \tau_{\text{п}}/\tau = 0,5$. Остальные данные прежние.
22. Что затрудняет увеличение числа каналов при частотном и временном разделении каналов? Укажите пути преодоления этих трудностей.
23. Покажите, что в каналах с АМ и ЧМ поднесущими за счет перекрестных помех со спектральной плотностью (5.52) воз-

никают ошибки, величина которых определяется выражениями

$$\delta_{i \text{ AM}}^2 \approx 0,6 N \frac{F_{\text{в}i} k_{\text{п}x}^2 t}{f_{\text{п}N} k_{\text{п.мн}}^2} \exp\left(-\frac{k_{\text{п.мн}}^2}{2}\right),$$

$$\delta_{i \text{ ЧМ}}^2 \approx \frac{0,2 N F_{\text{в}i} k_{\text{п}x}^2 t}{\Delta\varphi_{\text{м}i}^2 f_{\text{п}N} k_{\text{п.мн}}^2} \exp\left(-\frac{k_{\text{п.мн}}^2}{2}\right),$$

где $\Delta\varphi_{\text{м}i} = \Delta f_{\text{м}i} / F_{\text{в}i}$ — индекс модуляции i -й поднесущей; $F_{\text{в}i}$ — верхняя граничная частота спектра сообщения, передаваемого по i -му каналу.

24. Покажите, что при $k_{\text{п.мн}} = k_{\text{п}x}$ и $f_{\text{п}1} \ll f_{\text{п}N}$, а также при одинаковой ширине спектра сообщений во всех каналах ($F_{\text{в}i} = F_{\text{в}}$ для всех i) приведенные выше формулы принимают вид

$$\delta_{i \text{ AM}}^2 = \frac{0,3}{\gamma_{\text{зк}}} \exp\left(-\frac{k_{\text{п}x}^2}{2}\right), \quad \delta_{i \text{ ЧМ}}^2 = \frac{0,1 \exp(-k_{\text{п.мн}}/2)}{\gamma_{\text{зк}} \Delta\varphi_{\text{м}i} (1 + \Delta\varphi_{\text{м}i})},$$

где $\gamma_{\text{зк}}$ — защитный коэффициент полосы.

25. Назовите основные блоки аналоговых систем связи и приведите примеры их применения.
26. Нарисуйте схемы дифференцирования и интегрирования входного напряжения на операционных усилителях. Поясните их работу.
27. Нарисуйте схему линейного амплитудного детектора сигналов малого уровня. Поясните его работу.
28. Нарисуйте схемы ЧМ детектора в системе ЧМ—ЧМ.
29. Нарисуйте схемы активных фильтров: а) с линейной обратной связью, б) на основе ФАП.
30. Нарисуйте схемы устройств разделения систем: 1) АМ—ЧМ; 2) АИМ—ЧМ; 3) ВИМ—ЧМ с использованием современной элементной базы.

ГЛАВА 6. ПРИНЦИПЫ ФОРМИРОВАНИЯ СИГНАЛОВ В СИСТЕМАХ ПЕРЕДАЧИ ДИСКРЕТНЫХ СООБЩЕНИЙ. ИСКАЖАЮЩЕЕ ВЛИЯНИЕ ЛИНИИ СВЯЗИ

6.1. ПРИНЦИП ПЕРЕДАЧИ ДИСКРЕТНЫХ СООБЩЕНИЙ

В системах связи элементы x_i дискретных сообщений передаются от отправителя к получателю с помощью сигналов $s(t, x_i)$, однозначно сопоставляемых элементам x_i , где индекс i может принимать любое целое значение в интервале от 1 до M (напомним, что M — размер алфавита). Такое сопоставление по сути дела представляет собой кодирование в широком смысле (см. п. 1.2).

Сопоставление элементов x_i сигналам $s(t, x_i)$ может производиться различными способами в зависимости от назначения кодирования, требуемой простоты технической реализации системы или других обстоятельств.

Как уже указывалось в п. 1.2, кодирование в общем случае можно разделить на натуральное (первичное), эффективное и помехоустойчивое.

Остановимся подробнее на особенностях этих видов кодирования.

Натуральное кодирование. Основным требованием в этом случае является однозначность передачи элементов x_i . Кодирование состоит в непосредственном сопоставлении элементов сигналам. Число используемых сигналов определяет возможные разновидности кодирования. Крайними случаями являются M -ичное (M -арное) и двоичное (бинарное) кодирования.

При M -ичном кодировании каждому элементу x_i сопоставляется свой сигнал $s(t, x_i)$ некоторой длительностью T_c по элементарной логике следований: $x_i \rightarrow s(t, x_i)$. При этом число необходимых сигналов M_c равно размеру алфавита M .

При двоичном кодировании элементы x_i сначала нумеруются в двоичной системе счисления и представляются, таким образом, в виде последовательности цифр 1 и 0. Минимальное число двоичных знаков m , которое требуется для нумерации всех элементов x_i , называемое *позиционностью* или *разрядностью* двоичного кода, находится из соотношения

$$2^m \geq M.$$

Каждому элементу k_i двоичного кода (где $k_i=1$ или 0) сопоставляется свой сигнал $s(t, k_i)$. При этом число необходимых сигналов минимально ($M_c=2$), что является существенным преимуществом *бинарной системы передачи* при ее инженерной реализации.

Если увеличивать число элементов (цифр), используемых в коде (основание кода), то можно получить троичный код, четверичный и т. д. Такие коды называются *многоосновными*.

Эффективное кодирование. Основным требованием в этом случае является наиболее экономичное представление элементов x_i с точки зрения средней разрядности (длительности) закодированного элемента. (Требование однозначности при этом, конечно, сохраняется.) Разработанные эффективные коды также относятся к числу двоичных, так как используют два элемента k_0 и k_1 и для их реализации требуется только два сигнала.

Эффективное кодирование основано на том, что элементы x_i появляются в сообщении с различной вероятностью P_i . При эффективном кодировании разрядность двоичных кодовых комбинаций обратно пропорциональна вероятностям элементов x_i , т. е. более вероятным элементам x_i присваиваются более короткие кодовые комбинации. Разрядность кода становится при этом переменной величиной m_i , среднее значение которой \bar{m} можно сделать

меньше требуемой значности при натуральном бинарном кодировании. Это значит, что эффективное кодирование позволяет повысить скорость передачи дискретной информации.

Хотя число знаков кода при эффективном кодировании переменного, при передаче кодовых слов не требуются дополнительные разделительные знаки (или паузы) для однозначной дешифровки кодовых слов в элементы x_i . Для этого используется дополнительное требование кодирования: более короткие кодовые слова не должны быть началом более длинных кодовых слов.

Разработано несколько способов практического осуществления эффективного кодирования. Одни из них учитывают только различие вероятностей P_i элементов x_i , а другие наряду с этим принимают во внимание также и различные условные вероятности следования элементов x_i друг за другом, т. е. учитывают корреляцию элементов.

Например, для четырех элементов x_1, x_2, x_3 и x_4 , появляющихся в дискретных сообщениях с вероятностями соответственно 0,5; 0,25; 0,125; 0,125, эффективный код имеет вид

$$k_1 \rightarrow 1, \quad k_2 \rightarrow 01, \quad k_3 \rightarrow 001, \quad k_4 \rightarrow 000.$$

Этот пример наглядно иллюстрирует сформулированные выше требования эффективного кодирования.

В практике связи эффективное кодирование используется редко. Его основными недостатками являются усложнение кодирующей и декодирующей аппаратуры и пониженная помехоустойчивость, поскольку ошибочный прием одного элемента может привести к неверному декодированию ряда следующих за этим элементом кодовых слов. Вопросы эффективного кодирования подробно рассмотрены в [1].

Помехоустойчивое кодирование. Основным требованием в этом случае является возможность исправления кодовых комбинаций, в которых содержатся неверно принятые элементы k_i , или возможность обнаружения искаженных кодовых комбинаций. В первом случае код называют *исправляющим ошибки*, во втором — *обнаруживающим ошибки*.

Свойства помехоустойчивых кодов и особенности помехоустойчивого кодирования рассмотрены в гл. 8.

Сигналы $s(t, k_i)$ длительности T_c , соответствующие элементам кода k_i , подаются в линию связи, которая ослабляет их величину, искажает форму и добавляет помеху. В результате на входе приемного устройства на каждом интервале длительностью T_c имеет место колебание $y(t) = s_{вх}(t, k_i) + n(t)$, где $s_{вх}(t, k_i)$ — измененный линией сигнал $s(t, k_i)$, а $n(t)$ — колебание помехи. Основная задача приемного устройства состоит в том, чтобы определить, какой из переданных сигналов $s(t, k_i)$ присутствует в сумме $y(t)$, т. е. какой из элементов k_i был отправлен передающим устройством на рассматриваемом интервале времени длительностью T_c .

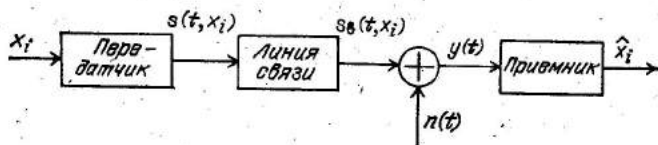


Рис. 6.1. Структурная схема системы передачи дискретных элементов

Следует подчеркнуть, что от приемного устройства не требуется восстанавливать форму сигнала $s(t, k_i)$, необходимо лишь определить его номер, т. е. решить, какой сигнал передан $s(t, k_0)$ или $s(t, k_1)$. Эта задача называется *задачей распознавания*.

Искажения формы сигналов в линии и наличие помех приводят к возможным ошибкам распознавания. Поэтому приемник не определяет элемент k_i , а выносит его оценку \hat{k}_i , которая может иногда и не совпадать с переданным элементом.

После приема (распознавания) элементов k_i приемное устройство осуществляет операцию декодирования, т. е. сопоставления принятых кодовых слов, состоящих из элементов \hat{k}_i , элементам x_i . Если элементы кода приняты верно, а схема декодирования работает правильно, то ошибок при приеме элементов x_i не происходит.

Структурная схема системы передачи дискретных элементов x_i приведена на рис. 6.1. Проведенное рассмотрение позволяет сформулировать следующие основные задачи теории и техники передачи дискретных сообщений:

- 1) выбор «наилучших» сигналов $s(t, x_i)$ и рассмотрение способов их формирования в передатчике;
- 2) определение и учет искажений, вносимых линией передачи в сигналы $s(t, x_i)$;
- 3) синтез оптимального приемника, т. е. определение алгоритма работы приемника, минимизирующего ошибки при распознавании сигналов;
- 4) анализ помехоустойчивости оптимального приемника, т. е. вычисление предельно достижимой минимальной вероятности ошибок;
- 5) изучение инженерных методов создания приемных устройств.

6.2. МЕТОДЫ СОЗДАНИЯ НАБОРА СИГНАЛОВ В СИСТЕМАХ ПЕРЕДАЧИ ДИСКРЕТНЫХ СООБЩЕНИЙ

При отсутствии помех и идеально стабильной аппаратуре связи в качестве M сигналов можно было бы выбрать M произвольных колебаний, сколь угодно мало отличающихся друг от друга, так как при указанных условиях они были бы безошибочно различимы в точке приема. Задача выбора оптимального (в определенном смысле) набора сигналов возникает лишь при наличии помех и нестабильностей аппаратуры. Для ее строгого решения необходимо ввести меру различимости сигналов между собою и

критерий оптимальности набора сигналов. В качестве последнего можно принять вероятность ошибок при различении сигналов выбранного набора в точке приема.

Предположим сначала, что линия вносит только случайные помехи. Рассмотрим следующий общий метод создания набора сигналов в этом случае. Выберем систему $N \leq M$ ортонормированных колебаний $\varphi_j(t)$, т. е. колебаний, обладающих свойством

$$\int_0^{T_c} \varphi_j(t) \varphi_k(t) dt = \begin{cases} 0, & j \neq k, \\ 1, & j = k, \end{cases} \quad (6.1)$$

и построим N генераторов, каждый из которых с периодом T_c генерирует последовательности колебаний из $\varphi_1(t), \varphi_2(t), \dots, \dots, \varphi_j(t), \dots, \varphi_N(t)$ соответственно. Способ выбора конкретного вида этого набора рассмотрим позднее.

Сопоставим каждому элементу x_i набор N чисел $a_{i1}, a_{i2}, \dots, \dots, a_{ij}, \dots, a_{iN}$ (i — номер элемента x_i , j — номер функции $\varphi_j(t)$).

После этого построим сигналы $s(t, x_i)$, соответствующие элементам x_i , по правилу

$$s(t, x_i) = \sum_{j=1}^N a_{ij} \varphi_j(t). \quad (6.2)$$

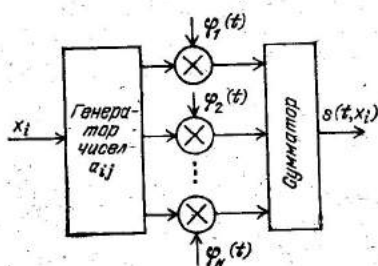


Рис. 6.2. Функциональная схема реализации набора дискретных сигналов

Функциональная схема реализации этого алгоритма показана на рис. 6.2. Ниже мы убедимся, что правило (6.2) является общим, не накладывающим принципиальных ограничений на вид сигналов.

Так как колебания $\varphi_i(t)$ заданы (однозначно выбраны), то каждый сигнал $s(t, x_i)$ однозначно определяется набором чисел a_{ij} . Эти числа можно рассматривать как координаты вектора сигнала s_i (см. п. 2.10), так что

$$s_i = (a_{i1}, a_{i2}, \dots, a_{ij}, \dots, a_{iN}). \quad (6.3)$$

Вектор (6.3) с координатами a_{ij} имеет N -мерную размерность.

Пример. Выберем систему из $N=2$ ортонормальных колебаний, определенных на интервале $0 < t < T_c$:

$$\varphi_j(t) = \begin{cases} \varphi_1(t) = \begin{cases} \sqrt{2/T_c} \sin \omega_0 t, & 0 < t < T_c, \\ 0 & \text{при других } t, \end{cases} \\ \varphi_2(t) = \begin{cases} \sqrt{2/T_c} \cos \omega_0 t, & 0 < t < T_c, \\ 0 & \text{при других } t. \end{cases} \end{cases}$$

Эти колебания показаны на рис. 6.3,а. Пусть необходимо получить $M=4$ сигнала. Выберем (пока без обоснования) следующие четыре набора чисел

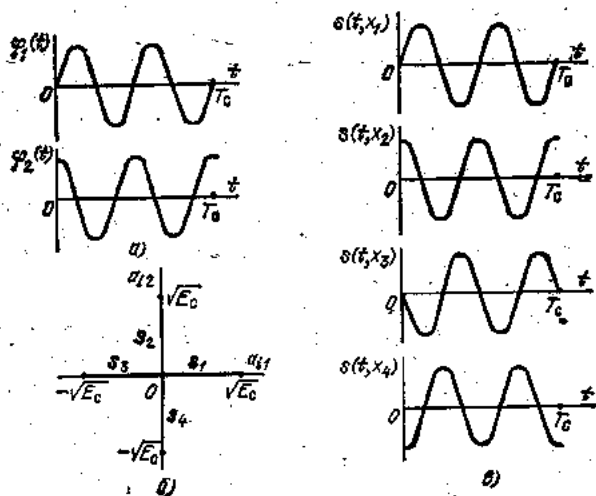


Рис. 6.3. К пояснению принципа создания четырех сигналов

a_{ij} (четыре вектора s_i):

$$s_1 = (a_{11}, a_{12}) = (\sqrt{E_c}, 0),$$

$$s_2 = (a_{21}, a_{22}) = (0, \sqrt{E_c}),$$

$$s_3 = (a_{31}, a_{32}) = (-\sqrt{E_c}, 0),$$

$$s_4 = (a_{41}, a_{42}) = (0, -\sqrt{E_c}),$$

где E_c — постоянная, имеющая, как будет видно ниже, смысл энергии сигнала. Векторы s_1, s_2, s_3, s_4 показаны на рис. 6.3,б. В соответствии с (6.2) запишем выражения сигналов $s(t, x_i)$ на интервале $0 < t < T_c$:

$$s(t, x_1) = \sum_{j=1}^2 a_{1j} \varphi_j(t) = a_{11} \varphi_1(t) + a_{12} \varphi_2(t) = \sqrt{\frac{2E_c}{T_c}} \sin \omega_0 t,$$

$$s(t, x_2) = \sum_{j=1}^2 a_{2j} \varphi_j(t) = a_{21} \varphi_1(t) + a_{22} \varphi_2(t) = \sqrt{\frac{2E_c}{T_c}} \cos \omega_0 t,$$

$$s(t, x_3) = \sum_{j=1}^2 a_{3j} \varphi_j(t) = a_{31} \varphi_1(t) + a_{32} \varphi_2(t) = -\sqrt{\frac{2E_c}{T_c}} \sin \omega_0 t,$$

$$s(t, x_4) = \sum_{j=1}^2 a_{4j} \varphi_j(t) = a_{41} \varphi_1(t) + a_{42} \varphi_2(t) = -\sqrt{\frac{2E_c}{T_c}} \cos \omega_0 t.$$

Эти сигналы показаны на рис. 6.3,в. Проверим, что величина E_c является удельной (выделяемой на единичном сопротивлении) энергией сигналов. По определению она равна

$$\int_0^{T_c} s^2(t, x_i) dt = \int_0^{T_c} \frac{E_c}{T_c} (\sin \omega_0 t)^2 dt = E_c$$

Следует обратить внимание на то, что концы векторов s_i , рассмотренных сигналов $s(t, x_i)$ в пространстве $N=2$ измерений парно предельно разнесены. Дальнейшее увеличение разноса можно получить только увеличив энергию сигналов E_c .

6.3. РАСПОЗНАВАНИЕ СИГНАЛОВ ПРИ ОТСУТСТВИИ ПОМЕХ

Убедимся, что по заданному сигналу $s(t, x_i)$ можно определить вектор этого сигнала s_i , т. е. определить его координаты a_{ij} . Действительно,

$$\begin{aligned} \int_0^T s(t, x_i) \varphi_l(t) dt &= \int_0^T \left[\sum_{j=1}^N a_{ij} \varphi_j(t) \right] \varphi_l(t) dt = \\ &= \sum_{j=1}^N \int_0^T a_{ij} \varphi_j(t) \varphi_l(t) dt = a_{il}. \end{aligned} \quad (6.4)$$

При последнем переходе в (6.4) использовано свойство ортогональности (6.1). Производя операцию (6.4), подставляя все колебания $\varphi_j(t)$ при $l=1, 2, \dots, N$, можно получить все составляющие вектора s_i .

Операцию (6.4) можно реализовать аппаратно в соответствии с функциональной схемой рис. 6.4, а. В точке приема необходимо иметь колебания $\varphi_j(t)$, аналогичные соответствующим колебаниям в передатчике (см. рис. 6.2). Указанные колебания должны вырабатываться синхронно и синфазно работающими генераторами в передатчике и приемнике.

Так как координаты a_{ij} однозначно определяют сигнал $s(t, x_i)$ и элемент x_i , то вычисление этих координат в приемнике эквивалентно определению переданного элемента x_i . Следует обратить внимание на то, что если линия не искажает форму сигналов и лишь ослабляет их так, что

$$s_{\text{ВЫХ}}(t, x_i) = \alpha s(t, x_i),$$

где $\alpha = \text{const}$, то вычисленные в точке приема координаты a_{ij} вектора s_i будут отличаться от соответствующих координат в передатчике только постоянным коэффициентом α . При этом геомет-

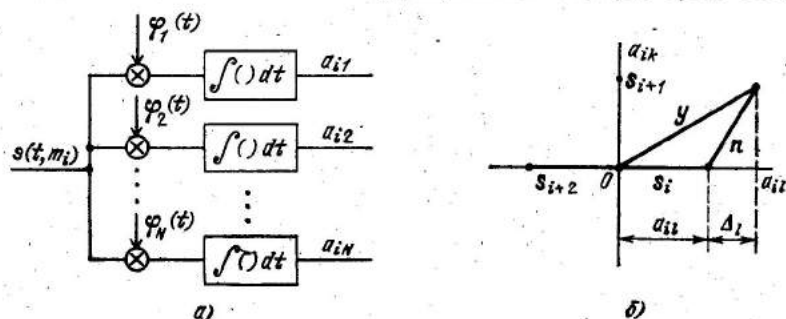


Рис. 6.4. К пояснению принципа восстановления вектора по форме сигнала (а) и распознавания сигналов при наличии помех (б)

рическая конфигурация векторов сигналов (взаимное расположение векторов s_i) не изменится и их безошибочное распознавание будет по-прежнему возможно. Если же выходные сигналы линии $s_{\text{вых}}(t, x_i)$ отличаются от входных $s(t, x_i)$ по форме, то восстановленный в точке приема вектор изменит свое направление, что может привести к ошибкам распознавания даже при отсутствии аддитивных помех.

6.4. РАСПОЗНАВАНИЕ СИГНАЛОВ ПРИ НАЛИЧИИ ПОМЕХ. КРИТЕРИЙ ОПТИМАЛЬНОСТИ НАБОРА СИГНАЛОВ

Рассмотрим случай, когда линия не искажает форму сигналов и только добавляет помеху, так что колебание в точке приема

$$y(t) = s_{\text{вых}}(t, x_i) + n(t), \quad (6.5)$$

где $s_{\text{вых}}(t, x_i)$ — ослабленный сигнал $s(t, x_i)$, а $n(t)$ — реализация помехи.

Подадим колебание (6.5) на схему рис. 6.4,а для восстановления вектора сигнала. На выходах схемы получим

$$y_i = \int_0^{T_c} y(t) \varphi_i(t) dt = \int_0^{T_c} [s_{\text{вых}}(t, x_i) + n(t)] \varphi_i(t) dt = a_{ii} + \Delta_i, \quad (6.6)$$

где

$$\Delta_i = \int_0^{T_c} n(t) \varphi_i(t) dt. \quad (6.7)$$

Величины Δ_i являются случайными вследствие случайности помехи $n(t)$.

Из (6.6) следует, что восстановленный вектор в общем случае не совпадает ни с одним из векторов ожидаемых сигналов. Эта ситуация показана на рис. 6.4, б. Вектор помехи n образован случайными координатами (6.7). Сумма вектора сигнала s_i и вектора помехи n образует вектор реализации $y(t)$:

$$y = s_i + n.$$

Из проведенного рассмотрения следует, что при наличии помех необходимо создать процедуру распознавания сигналов, т. е. соотнесения вектора y с одним из возможных векторов s_i . Такая процедура будет выведена в гл. 7. Из рассмотрения следует также, что можно предложить принцип выбора набора M векторов (сигналов), обеспечивающий понижение вероятности ошибок при распознавании сигналов $s(t, x_i)$. Этот принцип состоит в выборе набора векторов с наиболее удаленными концами. Так как координаты векторов пропорциональны энергии сигналов, то увеличить разнос концов векторов можно за счет увеличения энергии сигналов (тривиальный способ «грубой силы») или соответствующего взаимного расположения векторов сигналов, т. е. изменения их конфигурации.

6.5. ТИПОВЫЕ НАБОРЫ ПОМЕХОУСТОЙЧИВЫХ СИГНАЛОВ

Типовыми являются наборы сигналов, образованные в соответствии с общим правилом (6.2) и обладающие геометрической конфигурацией, обеспечивающей выполнение требования предельно возможного (в данных условиях) разнеса концов векторов сигналов. В зависимости от числа N ортонормальных функций $\varphi_j(t)$ и числа M образуемых сигналов принято выделять следующие конфигурации векторов сигналов [2].

Бинарные противоположные сигналы. В этом простейшем случае $N=1$ и образуется $M=2$ сигнала $s(t, x_1)$ и $s(t, x_2)$.

Векторы сигналов выбираются так:

$$s_1 = (a_{11}, a_{12}) = (\sqrt{E_c}, 0), \quad s_2 = (a_{21}, a_{22}) = (-\sqrt{E_c}, 0). \quad (6.8)$$

Векторы s_1 и s_2 показаны на рис. 6.5, а. Сигналы как функции времени в соответствии с (6.2) имеют вид

$$\begin{aligned} s(t, x_1) &= a_{11} \varphi_1(t) = \sqrt{E_c} \varphi_1(t), \\ s(t, x_2) &= -a_{11} \varphi_1(t) = -\sqrt{E_c} \varphi_1(t). \end{aligned} \quad (6.9)$$

Из (6.9) следует, что $s(t, x_1) = -s(t, x_2)$ независимо от выбора колебания $\varphi_1(t)$, которое с точки зрения теории может быть любым. Противоположными являются любые два сигнала, отличающиеся знаком.

Бинарные ортогональные сигналы. В этом случае используется $N=2$ ортонормальных колебаний $\varphi_1(t)$ и $\varphi_2(t)$; образуется $M=2$ сигнала, а векторы s_1 и s_2 выбираются так:

$$s_1 = (a_{11}, a_{12}) = (\sqrt{E_c}, 0), \quad s_2 = (a_{21}, a_{22}) = (0, \sqrt{E_c}). \quad (6.10)$$

Векторы s_1 и s_2 показаны на рис. 6.5, б. Временные графики сигналов зависят в соответствии с (6.2) от выбора ортонормаль-

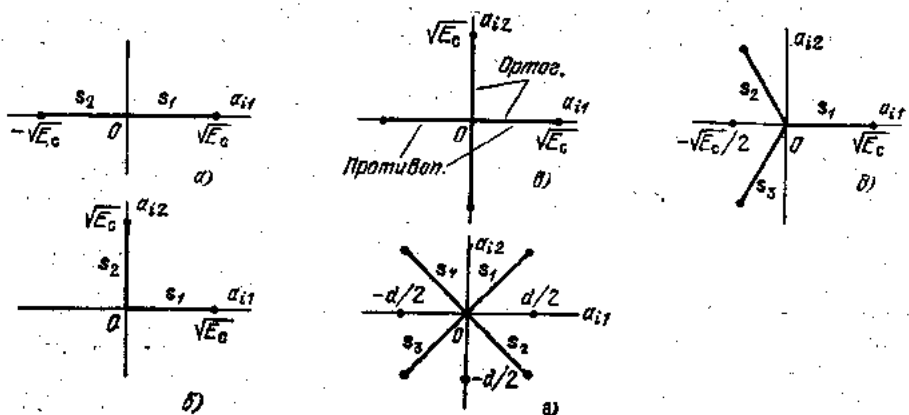


Рис. 6.5. Геометрические конфигурации типовых наборов сигналов: противоположных (а); ортогональных (б); биортогональных (в); прямоугольной конфигурации векторов (г); симплексных (д)

ных колебаний $\varphi_1(t)$ и $\varphi_2(t)$. Так, например, если принять те же колебания, что в примере п. 6.1, то в соответствии с (6.2)

$$\begin{aligned} s(t, x_1) &= a_{11} \varphi_1(t) + a_{12} \varphi_2(t) = \sqrt{E_c} \varphi_1(t), \\ s(t, x_2) &= a_{21} \varphi_1(t) + a_{22} \varphi_2(t) = \sqrt{E_c} \varphi_2(t). \end{aligned}$$

Легко проверить, что эти сигналы обладают свойством ортонормальности в смысле (6.1).

Из рассмотрения, проведенного в п. 6.4, следует, что при прочих равных условиях распознавание бинарных противоположных сигналов будет происходить с меньшей вероятностью ошибок, чем бинарных ортогональных, так как концы векторов первых больше разнесены. (Ниже мы увидим, что в некоторых реальных случаях целесообразнее применять ортогональные сигналы.)

М-арные ортогональные сигналы. Этот случай является обобщением предыдущего при образовании любого числа M сигналов и использования $N=M$ ортогональных функций $\varphi_j(t)$. При этом

$$\begin{aligned} s_1 &= (a_{11}, a_{12}, \dots, a_{1N}) = (\sqrt{E_c}, 0, \dots, 0); \\ s_2 &= (a_{21}, a_{22}, \dots, a_{2N}) = (0, \sqrt{E_c}, \dots, 0), \\ &\vdots \\ s_M &= (a_{M1}, a_{M2}, \dots, a_{MN}) = (0, 0, \dots, \sqrt{E_c}). \end{aligned} \quad (6.11)$$

Наглядное графическое изображение этих M -мерных векторов возможно лишь при $M=N=3$. Временные зависимости сигналов $s(t, x_i)$ определяются выбранной системой колебаний $\varphi_j(t)$. Так, если

$$\varphi_j(t) = \begin{cases} \sqrt{\frac{E_c}{T_c}} \sin \omega_j t, & 0 < t < T_c, \\ 0 & \text{при других } t, \end{cases} \quad (6.12)$$

то в соответствии с правилом (6.2) можно показать, что

$$s(t, x_i) = \sqrt{E_c} \varphi_i(t) = \sqrt{E_c T_c} \sin \omega_i t, \quad (0 < t < T_c), \quad (6.13)$$

т. е. сигналы являются отрезками гармонических колебаний разных частот $\omega_1, \omega_2, \dots, \omega_i, \dots, \omega_m$, удовлетворяющих условию ортогональности (6.1).

Биортогональные сигналы. В этом случае используется любое число N ортогональных колебаний $\varphi_j(t)$ и образуется $M=2N$ сигналов по следующему правилу: к каждому из построенных N ортогональных сигналов добавляется противоположный. Так, кроме сигнала с вектором $s_1 = (\sqrt{E_c}, 0)$ используется также сигнал с вектором $s_3 = (-\sqrt{E_c}, 0)$ а кроме ортогонального s_1 вектора $s_2 = (0, \sqrt{E_c})$ используется вектор $s_2 = (0, -\sqrt{E_c})$. Геометричес-

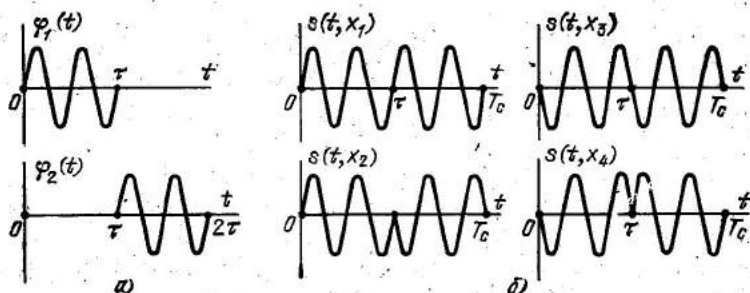


Рис. 6.6. Пояснения образования $M=4$ сигналов прямоугольной геометрической конфигурации векторов

кая конфигурация соответствующих векторов (при $N=2$) показана на рис. 6.5, в. Временные графики сигналов зависят от конкретно выбранного набора ортогональных колебаний $\varphi_j(t)$.

Сигналы с прямоугольной конфигурацией векторов. В этом случае используется любое число N колебаний $\varphi_j(t)$ и образуется $M=2^N$ сигналов. Геометрическая конфигурация векторов выбирается так, что их концы находятся в вершинах N -мерного куба. Так, при $N=2$ векторы имеют вид

$$\begin{aligned} \mathbf{s}_1 &= (a_{11}, a_{12}) = (d/2, d/2), & \mathbf{s}_3 &= (a_{31}, a_{32}) = (-d/2, -d/2), \\ \mathbf{s}_2 &= (a_{21}, a_{22}) = (d/2, -d/2), & \mathbf{s}_4 &= (a_{41}, a_{42}) = (-d/2, d/2), \end{aligned}$$

где d — постоянная величина. Эти векторы показаны на рис. 6.5 г. Их концы расположены в вершинах квадрата («двумерного куба»). Если в качестве $\varphi_1(t)$ и $\varphi_2(t)$ принять, например, смещенные на время τ синусоидальные посылки (см. рис. 6.6, а), то в соответствии с правилом (6.2) получим

$$\begin{aligned} s(t, x_1) &= \sum_{j=1}^2 a_{1j} \varphi_j(t) = \frac{d}{2} \varphi_1(t) + \frac{d}{2} \varphi_2(t), & s(t, x_2) &= \\ &= \sum_{j=1}^2 a_{2j} \varphi_j(t) = \frac{d}{2} \varphi_1(t) - \frac{d}{2} \varphi_2(t). \end{aligned} \quad (6.14)$$

Аналогично образуются остальные два сигнала. Полученные сигналы показаны на рис. 6.6, б. Они отличаются различным сочетанием начальных фаз составляющих их синусоидальных посылок длительностью τ . Аналогичным образом, принимая $N=3, 4, \dots$ и более, можно образовать $M=8, 16$ и более сигналов. При этом вовсе не обязательно использовать все сигналы. Некоторые могут оставаться лишь «потенциально возможными».

Симплексные сигналы. Это набор $M \geq 2$ сигналов, концы векторов которых равно и предельно удалены друг от друга. Так, при выборе двух ($N=2$) ортогональных колебаний $\varphi_j(t)$ можно построить три ($M=3$) симплексных сигнала, геометрическая конфигурация векторов которых показана на рис. 6.5, д. При $N=1$

можно построить два симплексных сигнала, которые совпадают с бинарными противоположными.

Симплексные сигналы можно получить из ортогональных. Действительно, будем характеризовать удаленность концов двух векторов i и j расстоянием d_{ij} (см. гл. 2). Нетрудно убедиться, что в этом случае

$$d_{ij} = \sqrt{\int_0^{T_c} [s(t, x_i) - s(t, x_j)]^2 dt} = \sqrt{E_i + E_j - 2\sqrt{E_i E_j} \kappa_{ij}} \quad (6.15)$$

где E_i, E_j — удельные (выделяемые на единичной нагрузке) энергии сигналов, а

$$\kappa_{ij} = \frac{1}{\sqrt{E_i E_j}} \int_0^{T_c} s(t, x_i) s(t, x_j) dt \quad (6.16)$$

— коэффициент взаимной корреляции рассматриваемых сигналов. Для ортогональных сигналов, например, $\kappa_{ij} = 0$ (по определению), для противоположных $\kappa_{ij} = -1$. Из (6.15) следует, что расстояние d_{ij} не изменится, если ко всем сигналам добавить одно и то же колебание $s(t)$. Суммарная энергия этих новых сигналов

$$E_{\Sigma} = \sum_{i=1}^M \int_0^{T_c} [s(t, x_i) + s(t)]^2 dt. \quad (6.17)$$

При нахождении минимума этой величины по $s(t)$ образуется (по определению) система симплексных сигналов. Анализ показывает, что

$$s_{\text{симпл}}(t, x_i) = s(t, x_i) - \frac{1}{M} \sum_{i=1}^M s(t, x_i), \quad (6.18)$$

где $s(t, x_i)$ — ортогональные сигналы.

Можно проверить, что векторы полученных симплексных сигналов $s_{\text{симпл}}(t, x_i)$ сдвинуты относительно друг друга на одинаковое расстояние ($d = \text{const}$), а энергия каждого из них равна $E_{\text{симпл}} = E(1 - 1/M)$, где E — энергия ортогональных сигналов (одинаковая для каждого из них). Отсюда следует, в частности, вывод, что при $M = 2$ применение симплексных сигналов при одинаковом расстоянии дает по сравнению с ортогональными выигрыш по энергии вдвое. Если $M \gg 2$, то симплексные сигналы практически эквивалентны ортогональным.

6.6. ВОПРОСЫ ИНЖЕНЕРНОЙ РЕАЛИЗАЦИИ СИГНАЛОВ

При выборе конкретного вида сигналов для системы связи, кроме его геометрической конфигурации, влияющей на помехоустойчивость системы, во внимание принимается ряд факторов, важнейшими из которых являются: свойства заданной линии передачи (полоса пропускания, постоянство параметров во времени); количество различных сообщений, которые одновременно необходимо передавать по линии (многоканальность); необходи-

мость совмещения работы нескольких систем передачи в одном частотном диапазоне с минимальными помехами (требование совместности); предпочтительность технической реализации генераторов на основе той или иной технологии.

При практическом формировании сигналов нет необходимости формально следовать алгоритму (6.2). Сигналы, обладающие требуемыми свойствами, часто можно получить непосредственным генерированием требуемых форм.

Широко используемым набором бинарных сигналов являются сигналы в виде отрезков гармонических колебаний с разными частотами или начальными фазами:

$$s(t) = S \cos(\omega t + \varphi), \quad 0 \leq t \leq T_c. \quad (6.19)$$

При сохранении неизменной частоты и изменении только начальной фазы по правилу $x_1 \rightarrow (\varphi=0)$, $x_2 \rightarrow (\varphi=\pi)$, получаются два противоположных сигнала (сигналы с фазовой манипуляцией). Если же изменить частоту по правилу $x_1 \rightarrow \omega_1$, $x_2 \rightarrow \omega_2$, причем

$$\omega_2 - \omega_1 \geq 2\pi/T_c, \quad (6.20)$$

то получается пара почти ортогональных сигналов, для которых

$$\int_0^{T_c} s(t, x_1) s(t, x_2) dt \approx 0$$

(сигналы с частотной манипуляцией).

Помехоустойчивость при почти ортогональных сигналах будет несколько ниже, чем при строго ортогональных сигналах вида (6.13), однако они располагаются более «компактно» по частоте. Внося сдвиг по частоте на величину $\omega\Delta \geq 2\pi/T_c$, можно образовать систему из $M > 2$ ортогональных (точнее, почти ортогональных) сигналов.

В бинарных системах иногда используется пара сигналов с амплитудной манипуляцией:

$$s(t, x_1) = S \cos(\omega_0 t + \varphi_0); \quad s(t, x_2) = 0, \quad 0 \leq t \leq T_c.$$

Этот случай называется в технике работой «с пассивной паузой» (при передаче символов $x_2=0$ сигнал не излучается), в отличие от случаев частотной или фазовой манипуляции, называемых «работой с активной паузой», когда энергия сигнала расходуется при передаче как символов $x_1=1$, так и символов $x_2=0$.

Полезной разновидностью сигналов являются сложные сигналы, у которых произведение ширины спектра Δf_c на длительность посылки T_c значительно больше единицы (см. п. 3.4). Применение таких сигналов позволяет успешно бороться с многолучевостью при приеме сигналов, улучшает электромагнитную совместимость различных систем, повышает помехоустойчивость приема при действии разного рода помех.

При решении вопроса о выборе вида сигналов для проектируемой системы связи необходимо учитывать искажающее действие среды распространения (физической линии связи).

6.7. ЛИНИЯ СВЯЗИ КАК ИСКАЖАЮЩИЙ ЧЕТЫРЕХПОЛЮСНИК

Подавляющее большинство реальных линий связи (сред распространения сигналов) можно рассматривать как линейные четырехполюсники, к которым, следовательно, применим принцип суперпозиции.

Из общего соотношения для входного и выходного колебаний линейного четырехполюсника

$$u_{\text{вых}}(t) = \int_{-\infty}^t u_{\text{вх}}(\tau) h(t-\tau) d\tau \quad (6.21)$$

следует, что такой четырехполюсник не вносит искажений в форму выходного колебания лишь тогда, когда его отклик $h(t)$ является δ -импульсом (это эквивалентно требованию бесконечно широкой полосы пропускания частот). В таком простейшем практически редко встречающемся случае импульсный отклик линии можно записать в виде

$$h_{\text{л}}(t) = a \delta(t - \tau_3),$$

где a — ослабление, а τ_3 — запаздывание.

Тогда сигнал на выходе

$$s_{\text{вых}}(t, x_i) = \int_{-\infty}^t s(\tau, x_i) a \delta(t - \tau - \tau_3) d\tau = a s(t - \tau_3, x_i),$$

т. е. будет ослаблен и задержан.

Большинство реальных сред относится к линейным четырехполюсникам с переменными параметрами. Для них импульсный отклик зависит не только от интервала τ , прошедшего с момента воздействия δ -импульса, но и от текущего времени:

$$h = h(\tau, t). \quad (6.22)$$

Зависимость импульсного отклика от времени носит, как правило, случайный характер, а изменения отклика $h(\tau)$ во времени обычно происходит медленно. Это означает, что за длительность посылки сигнала T_c отклик практически не изменяется.

Для линейных систем с переменными параметрами [2, 3]

$$u_{\text{вых}}(t) = \int_{-\infty}^t u_{\text{вх}}(\tau) h(t-\tau, t) d\tau. \quad (6.23)$$

Для большей наглядности анализа можно представить импульсный отклик как функцию τ в виде ряда Котельникова. Так как большинство радиотехнических сигналов относится к числу полосовых ($\Delta f_c \ll f_{\text{ср}}$, где $f_{\text{ср}}$ — средняя частота спектра полосового сигнала), то для описания отклика требуются выборки как из исследуемого отклика $h(\tau, t)$, так и из сопряженной (квадратурной) ему величины $\tilde{h}(\tau, t)$. Обозначим эти выборки через $h_k(t)$ и $\tilde{h}_k(t)$ соответственно. Эти выборки, как и сама функция $h(\tau, t)$, будут, очевидно, случайными медленно меняющимися функциями времени. Выборки необходимо брать через интервал $1/\Delta f_c$. Анализ

показывает [3], что при этом выходной сигнал с учетом (6.23) можно представить так:

$$s_{\text{вых}}(t, x_i) = \frac{2\pi}{\Delta\omega_c} \left\{ \sum_{k=0}^{\tau_{\text{л}} \Delta f_c} h_k(t) s\left(t - k \frac{2\pi}{\Delta\omega_c}, x_i\right) - \sum_{k=0}^{\tau_{\text{л}} \Delta f_c} \tilde{h}_k(t) \tilde{s}\left(t - k \frac{2\pi}{\Delta\omega_c}, x_i\right) \right\}, \quad (6.24)$$

где $\tau_{\text{л}}$ — длительность импульсного отклика линии в исследуемой полосе частот, а $\tilde{s}(t, x_i)$ — сопряженное (квадратурное) значение сигнала $s(t, x_i)$.

Из (6.24) видно, что влияние линии в общем случае сводится к появлению ряда запаздывающих на интервалы $k2\pi/\Delta\omega_c$ входных сигналов $s(t, x_i)$, а также сопряженных им колебаний $\tilde{s}(t, x_i)$, домноженных соответственно на выборки из отклика и сопряженного ему значения. Количество пар запаздывающих составляющих определяется произведением ширины полосы сигналов на длительность импульсного отклика ($\Delta f_c \tau_{\text{л}}$).

Причину появления в выходном колебании квадратурных составляющих можно понять, рассмотрев одно (произвольное) слагаемое суммы (6.24):

$$u_k(t) = \frac{2\pi}{\Delta\omega_c} h_k(t) s\left(t - k \frac{2\pi}{\Delta\omega_c}, x_i\right) - \frac{2\pi}{\Delta\omega_c} \tilde{h}_k(t) \tilde{s}\left(t - k \frac{2\pi}{\Delta\omega_c}, x_i\right) = \alpha_k(t) s(t - \tau_k) - \tilde{\alpha}_k \tilde{s}(t - \tau_k). \quad (6.25)$$

Здесь для упрощения обозначено: $\alpha_k(t) = (2\pi/\Delta\omega_c) h_k(t)$, $\tilde{\alpha}_k(t) = (2\pi/\Delta\omega_c) \tilde{h}_k(t)$, $k2\pi/\Delta\omega_c = \tau_k$ и опущено обозначение элемента x_i в выражении сигнала.

Пусть входной полосовой сигнал $s(t) = S(t) \cos \omega_0 t$. Сопряженное его значение $\tilde{s}(t) = S(t) \sin \omega_0 t$. Тогда k -е слагаемое выходного сигнала

$$u_k(t) = \alpha_k(t) S(t - \tau_k) \cos \omega_c(t - \tau_k) - \tilde{\alpha}_k(t) S(t - \tau_k) \sin \omega_c(t - \tau_k).$$

Путем элементарных преобразований оно приводится в виду (см. также п. 2.6).

$$u_k(t) = S_{\text{вых}}(t) \cos(\omega_0 t - \tau_k - \beta_k), \quad (6.26)$$

где

$$S_{\text{вых}}(t) = \sqrt{\alpha_k^2(t) + \tilde{\alpha}_k^2(t)} S(t - \tau_k), \quad \beta_k = \text{arctg}[\tilde{\alpha}_k(t)/\alpha_k(t)]. \quad (6.27)$$

Из приведенных выражений видно, что наличие прямой и квадратурной составляющих в (6.25) эквивалентно изменению как амплитуды, так и начальной фазы каждой составляющей выражения (6.26). Эти изменения, как видно из (6.27), определяются значениями выборок $h_k(t)$ и $\tilde{h}_k(t)$.

Физическую причину появления ряда составляющих в выходном колебании линии можно пояснить на примере линии связи, использующей отражения радиоволн от ионосферы. Последняя состоит из совокупности неоднородных областей с повышенной и пониженной концентрацией («электронные облака»). Колебание в точке приема складывается из многих составляющих, имеющих различные амплитуды и запаздывания. Эти составляющие часто называют *лучами*. В выражении (6.24), являющемся математической моделью (приближенным математическим описанием) явления, слагаемые, запаздывающие на интервалы $k2\pi/\Delta\omega_c$, можно рассматривать как отдельные лучи.

6.8. ОСНОВНЫЕ ВИДЫ ИСКАЖЕНИЙ, ВНОСИМЫХ ЛИНИЕЙ

Из (6.24) следует, что по отношению к прохождению через заданную линию связи все сигналы целесообразно разделить на две группы.

1. Сигналы, для которых $\Delta f_{стл} < 1$ назовем *узкополосными* относительно линии (при этом произведение $\Delta f_c T_c$ может быть $\gg 1$, т. е. сигналы могут быть широкополосными в спектральном смысле). Для этих сигналов в ряде (6.26) остается единственное слагаемое при $k=0$. Поэтому произвольный полосовой сигнал на входе линии

$$s(t, x_i) = S(t) \cos [\omega_0 t + \varphi(t)]$$

преобразуется линией в колебание, аналогичное (6.26), т. е.

$$s_{ввых}(t, x_i) = K(t) S(t) \cos [\omega_0 t + \varphi(t) - \beta(t)], \quad (6.28)$$

где

$$K(t) = \frac{2\pi}{\Delta\omega_c} \sqrt{h_0^2(t) + \tilde{h}_0^2(t)}, \quad \beta(t) = \arctg \frac{\tilde{h}_0(t)}{h_0(t)}. \quad (6.29)$$

Узкополосный относительно линии сигнал претерпевает лишь изменение амплитуды (появляется множитель $K(t)$) и фазы (член $\beta(t)$ в аргументе косинуса).

В зависимости от характера изменения выборок $h_0(t)$ и $\tilde{h}_0(t)$ из импульсного отклика линии при этом возможны различные ситуации:

а) параметры K и β постоянны. В этом случае их можно с любой точностью измерить и скомпенсировать в точке приема или передачи. Такой случай, называемый случаем *точно известных сигналов*, на практике встречается крайне редко;

б) коэффициент передачи $K(t) = K$ — постоянная величина, а сдвиг фазы $\beta(t)$ — медленная функция времени, практически постоянная на интервалах, равных длительности сигналов T_c . Это случай сигналов со случайной (или неизвестной) начальной фазой. Он обычно имеет место в радиорелейных, проводных и кабельных линиях;

в) оба параметра K и β — медленные случайные функции, практически постоянные на интервалах длительностью T_c . Это случай сигналов с неизвестными (или случайными) амплитудой и начальной фазой. Он является наиболее типичным для радиолиний;

г) оба параметра K и β изменяются во времени настолько быстро, что их нельзя считать постоянными на интервалах длительностью T_c . Из (6.28) видно, что при этом в процессе прохождения сигналов через линию они подвергаются случайной амплитудной и фазовой модуляции.

При построении оптимальных приемников имеет значение не только факт постоянства или переменности величин K и β (или h_0 и \tilde{h}_0), но и статистические законы их изменения. В инженерной практике принято различать два основных случая: если h_0 и \tilde{h}_0 изменяются независимо и распределены по нормальному закону с нулевым средним значением, то, как можно убедиться, величина K будет при этом распределена по закону Рэлея, а начальная фаза — равномерно в пределах от $-\pi$ до π . Такая линия (или канал) называется *однолучевой рэлевской*. Если же h_0 и \tilde{h}_0 нормальны, но с ненулевым средним, то величина K распределена

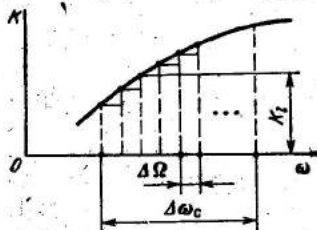


Рис. 6.7. Пояснение частотно-разнесенного представления канала передачи сигналов

при этом по закону Райса (закон распределения β при этом имеет сложное выражение). Такая линия называется *однолучевой райсовской*.

2. Сигналы, для которых $\Delta f_{стл} \gg 1$, назовем *широкополосными* относительно линии связи. При этом сами сигналы могут быть узкополосными в спектральном смысле, т. е. $\Delta f_c T_c \approx 1$ (все зависит от величин τ_l и Δf_c , точнее, от их произведения). В этом случае сигналы на выходе могут существенно отличаться по форме от входных, что, как увидим ниже, значительно затрудняет их безошибочно распознавание. Линии, для которых $\Delta f_{стл} \gg 1$, называются еще *многолучевыми*, так как в выходном сигнале (6.24) появляется ряд запаздывающих слагаемых или лучей.

Многолучевую линию (канал) можно описать, пользуясь понятием частотно-разнесенного представления. Рассмотрим модель коэффициента передачи $K(\omega)$ линии в исследуемом диапазоне частот (рис. 6.7) и произведем его «ступенчатую» аппроксимацию в интервалах частот $\Delta\Omega \ll \Delta\omega_c$. Сдвиг фазы $\beta(\omega)$ аппроксимируем кусочно-ломаной прямой. При этом выходной сигнал линии можно представить в виде суммы частотно-смещенных (разнесенных) составляющих в отдельных участках спектра шириною $\Delta\Omega$:

$$s(t, x_l) = \sum_{l=1}^{\Delta\omega_c/\Delta\Omega} s_l(t, x_l), \quad (6.30)$$

где через $s_l(t, x_l)$ обозначена составляющая колебания входного сигнала в l -й полосе шириною $\Delta\Omega$.

При прохождении через линию каждое слагаемое (6.32) приобретает некоторые задержку τ_l , ослабление K_l и сдвиг фазы ψ_l . В результате

$$s_{\text{вых}}(t, x_l) = \sum_{l=1}^{\Delta\omega_c/\Delta\Omega} K_l s_l(t - \tau_l, \psi_l). \quad (6.31)$$

При таком подходе приходится учитывать так называемые частотно-селективные замирания (фединг) в линии, так как ослабление частотно-разнесенных составляющих может происходить не одновременно (не «дружно») и по-разному. Частотно-селективная линия с выходным сигналом вида (6.31) на основании проведенного выше рассмотрения является многолучевой (ср. с. (6.24)).

В зависимости от характера случайных изменений параметров в (6.24) и (6.31), в случае многолучевой линии целесообразно рассматривать следующие ситуации:

а) параметры K_l , τ_l и ψ_l (или h_k и \bar{h}_k) в выражении (6.24) постоянны и, следовательно, могут быть изменены с любой точностью. Этот случай по существу, не отличается от ранее рассмотренного случая точно известных сигналов;

б) параметры K_l , τ_l , ψ_l — случайные медленно изменяющиеся функции. В этом случае, в зависимости от статических свойств параметров составляющих в (6.24) или (6.31), выделяют случаи многолучевых рэлеевских линий с независимыми замираниями компонент, а также многолучевых райсовских линий или с зависимыми или независимыми замираниями компонент;

в) параметры K_l , τ_l , ψ_l изменяются настолько быстро, что не сохраняются постоянными на интервалах длительностью T_c . Этот случай крайне неблагоприятен для распознавания.

КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАЧИ

1. Поясните, почему передача любых дискретных элементов эквивалентна передаче цифр.
2. Дайте определение первичного кодирования и укажите его разновидности.
3. Назовите основные цели эффективного кодирования и помехоустойчивого кодирования.
4. Дайте определение задачи распознавания сигналов.
5. Перечислите основные задачи теории и техники передачи дискретных сообщений.
6. Укажите общее правило синтеза набора M сигналов с помощью N ортонормальных функций.
7. Дайте определение вектора s_i , эквивалентного сигналу $s(t, x_i)$.
8. Сформулируйте правило вычисления вектора сигнала по заданной его форме.

9. Дайте определение противоположных, ортогональных, биортогональных, симплексных сигналов и сигналов прямоугольной конфигурации векторов.
10. Перечислите основные применяемые в технике типы сигналов и их особенности.
11. Создайте набор сигналов с $M=4$ и $M=8$, в которых бы использовалась амплитудно-фазовая манипуляция.
12. Объясните характер искажений, вносимых линией в соответствии с выражением (6.24).
13. Дайте определение однолучевой и многолучевой линий.
14. Перечислите основные случаи искажений, вносимых однолучевой и многолучевой линиями.
15. Объясните причины частотно-селективных замираний в линии.
16. Определите значность (разрядность) n требуемого двоичного кода при натуральном кодировании, если алфавит сообщения состоит из всех букв и знаков русского языка, латинского языка и десятичных цифр.
17. Допустим ли нижеследующий эффективный код для 4-элементного алфавита:

$$\begin{array}{ll} x_1 \rightarrow 1 & x_3 \rightarrow 10 \\ x_2 \rightarrow 01 & x_4 \rightarrow 001 \end{array}$$

Если нет, то почему?

18. Постройте графики трех симплексных сигналов при $N=2$ ортонормальных колебаниях примера п. 6.1.
19. Постройте график $M=4$ ортогональных сигналов при заданных $N=4$ ортонормальных колебаниях $\varphi_j(t)$. Вид колебаний $\varphi_j(t)$ выбрать из соображений простоты генерирования.
20. Сколько ортогональных колебаний $\varphi_j(t)$ потребуется для построения $M=100$ сигналов прямоугольной конфигурации?
21. Какой минимальный диапазон частот потребуется для размещения $M=32$ почти ортогональных сигналов длительностью $T_c = 1$ мс?
22. Будет ли однолучевой линией с длительностью отклика $\tau_n = 1$ мс по отношению к сигналам в виде гармонических посылок с прямоугольной огибающей длительностью $T_c = 10$ мкс?
23. Определите ослабление и сдвиг фазы, вносимые однолучевой линией с $h_0 = 0,1$, и $h_0 = 0,2$.

ГЛАВА 7. ОПТИМАЛЬНЫЙ ПРИЕМ СИГНАЛОВ В СИСТЕМАХ ПЕРЕДАЧИ ДИСКРЕТНЫХ СООБЩЕНИЙ

7.1. ПОСТАНОВКА ЗАДАЧИ СИНТЕЗА ОПТИМАЛЬНОГО ПРИЕМНИКА ДИСКРЕТНЫХ СООБЩЕНИЙ

Предположим сначала, что линия связи не вносит искажений в форму сигнала, но лишь ослабляет сигналы и добавляет помеху. Тогда на выходе линии (на входе приемника) на интервалах

времени $kT_c < t < (k+1)T_c$, где k — целые числа, будут иметь место реализации

$$y(t) = s(t, x_i) + n(t). \quad (7.1)$$

Здесь $s(t, x_i)$ — один из возможных сигналов из известного в точке приема множества, а $n(t)$ — реализация помехи. Задачей приемного устройства является определение номера i сигнала или, что то же, переданного элемента x_i длительностью T_c .

Распознавание сигналов можно, в принципе, производить различными способами с помощью различных приемников. Среди множества мыслимых приемников существует приемник, обеспечивающий прием с минимальной вероятностью ошибок. Такой приемник называется *оптимальным*.

Задача теории состоит в определении алгоритма работы и структуры оптимального приемника, а также в вычислении достижимой вероятности ошибок.

Из-за наличия помех безошибочное однозначное распознавание, очевидно, невозможно. Поэтому наибольшее, что может сделать приемник на первом шаге — это указать (вычислить) вероятности присутствия в заданной реализации $y(t)$ каждого из возможных сигналов передатчика. На следующем шаге необходимо установить правило вынесения однозначного решения. Здесь возможны различные подходы (критерии). В дальнейшем будем предполагать (как это чаще всего делаем), что присутствует тот сигнал, вероятность которого наибольшая.

В основе синтеза оптимального приемника лежит теорема Байеса, связывающая вероятности причин и следствий (см. п. 3.4). В рассматриваемом случае распознавания дискретных сигналов «причинами» являются случайно возникающие элементы сообщения x_i и соответствующие им сигналы $s(t, x_i)$. Обозначим это дискретное множество причин через $\Pi_1, \Pi_2, \dots, \Pi_i, \dots, \Pi_M$, а их априорные вероятности через $P(\Pi_1), P(\Pi_2), \dots, P(\Pi_M)$. «Следствиями» являются получаемые на входе приемника реализации $y(t)$. Множество этих следствий континуально, поэтому можно говорить лишь о плотности вероятности такого следствия (реализации $y(t)$). Однако сначала для наглядности рассмотрения предположим, что следствия также дискретны и их можно представить как $C_1, C_2, \dots, C_j, \dots, C_L$.

Ясно, что из-за случайности помех каждая причина (сигнал $s(t, x_i)$) может вызвать появление каждого следствия (любой реализации $y(t)$). Эта статистическая ситуация описывается матрицей условных вероятностей

$$\left\| \begin{array}{cccc} P(C_1|\Pi_1), & P(C_2|\Pi_1), & \dots, & P(C_L|\Pi_1) \\ P(C_1|\Pi_2), & P(C_2|\Pi_2), & \dots, & P(C_L|\Pi_2) \\ \vdots & \vdots & \ddots & \vdots \\ P(C_1|\Pi_M), & P(C_2|\Pi_M), & \dots, & P(C_L|\Pi_M) \end{array} \right\| \quad (7.2)$$

Следует подчеркнуть, что в рассматриваемой ситуации приемник, в принципе, не может сделать больше, чем вычислить по-

полученному следствию — реализации $y(t)$ апостериорные вероятности причин — сигналов $s(t, x_i)$ и, следовательно, соответствующих им элементов сообщения x_i . Как было пояснено в п. 3.4, апостериорные вероятности причин при заданном следствии C_j вычисляются в соответствии с теоремой Байеса:

$$P(\Pi_i|C_j) = \frac{P(\Pi_i)P(C_j|\Pi_i)}{P(C_j)}, \quad i = 1, 2, \dots, M. \quad (7.3)$$

Здесь $P(C_j)$ — безусловная вероятность j -го следствия, которая определяется по теореме о полной вероятности [1]:

$$P(C_j) = \sum_{i=1}^M P(\Pi_i)P(C_j|\Pi_i). \quad (7.4)$$

Получив в соответствии с (7.3) набор M апостериорных вероятностей $P(\Pi_i|C_j)$, получатель должен выбрать критерий, по которому он будет принимать окончательные решения о том, какой именно из сигналов передавался на рассматриваемом интервале. В нашем подходе это будет сигнал с наибольшей апостериорной вероятностью.

Приняв во внимание физическую природу причин и следствий рассматриваемой задачи, в соответствии с (7.3) и (7.4) получим

$$P[x_i|y(t)] + \frac{P(x_i)p[y(t)|x_i]}{p[y(t)]}.$$

Здесь и далее P — вероятности, а p — плотности вероятностей. Напомним, что вероятности $P(C_j|\Pi_i)$ или плотности вероятностей $p(y(t)|x_i)$ при заданных результатах наблюдений называются *правдоподобиями причин*.

7.2. СИНТЕЗ ОПТИМАЛЬНОГО ПРИЕМНИКА ДЛЯ РАСПОЗНАВАНИЯ ИЗВЕСТНЫХ СИГНАЛОВ

В задаче распознавания сигналов, не содержащих случайных параметров (т. е. точно известных), «причинами» являются поступающие на вход сигналы $s(t, x_i)$, вероятности которых равны, очевидно, вероятностям появления соответствующих элементов x_i . «Следствиями» являются реализации суммы сигнала и помехи (7.1).

Количественно описание ситуации удобно производить с помощью рассмотрения векторов соответствующих колебаний. Вместо сигналов $s(t, x_i)$ будем оперировать однозначно соответствующими им векторами $s_i = (a_{i1}, a_{i2}, \dots, a_{iN})$, а вместо реализаций $y(t)$ — векторами $y = (y_1, y_2, \dots, y_N)$, координаты которых определяются выражением (6.6.), которое в нашем случае запишем так:

$$y_j = \int_{kT_c}^{(k+1)T_c} y(t) \Phi_j(t) dt = a_{ij} + \Delta_j, \quad (7.5)$$

где Δ_j определяется соотношением (6.7).

В соответствии с теоремой Байеса

$$P(s_i|y) = P(s_i) p(y|s_i) / p(y) \quad (7.6)$$

Как было отмечено, решение обычно выносится в пользу сигнала, имеющего наибольшую апостериорную вероятность. Так как знаменатель (7.6) не зависит от номера i , то решающее правило (алгоритм решения) определяется так:

$$R = \max_i P(s_i) p(y|s_i) = \max_i P(x_i) p(y|s_i) \quad (7.7)$$

Следует обратить внимание на то, что в этих выражениях $p(y|s_i)$ — плотности вероятностей, так как компоненты вектора y , как видно из (7.5), являются непрерывными случайными величинами.

В выражении (7.7) априорные вероятности $P(x_i)$ передачи элементов x_i должны быть заданы. Следовательно, необходимо определить только правдоподобия $p(y|s_i)$. Это можно сделать исходя из того, что помеха аддитивна. Так как

$$y = s_i + n,$$

то плотность вероятности некоторого значения вектора $y = y_1$ равна плотности вероятности, что вектор помехи n примет значение $n_1 = y_1 - s_i$. Отсюда следует, что если $p_{\text{пом}}(n)$ — известная нам плотность вероятности вектора помехи, то

$$p(y|s_i) = p_{\text{пом}}((y_1 - s_i)|s_i) = p_{\text{пом}}(y_1 - s_i) \quad (7.8)$$

Последний переход справедлив потому, что сигнал и помехи — независимые процессы.

Для дальнейшей конкретизации алгоритма необходимо задать определенный вид помехи. В большинстве случаев имеют место нормальные (гауссовские) или близкие к ним помехи. Вычисления в этом случае оказываются наиболее простыми. При гауссовских помехах каждая компонента вектора $n = (n_1, n_2, \dots, n_j, \dots, n_N)$ распределена по нормальному закону

$$p_{\text{пом}}(n_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{n_j^2}{2\sigma^2}\right) \quad (7.9)$$

В ряде случаев, в частности, при равномерном распределении энергии помехи по полосе рассматриваемых частот, компоненты вектора n_j являются независимыми случайными величинами. Тогда, как известно, [1]

$$\begin{aligned} p_{\text{пом}}(n_1, n_2, \dots, n_N) &= \prod_{j=1}^N p_{\text{пом}}(n_j) = \\ &= \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{j=1}^N n_j^2\right\}. \end{aligned} \quad (7.10)$$

При зависимых компонентах n_j выражение для $p_{\text{пом}}(n_1, n_2, \dots, n_N)$ существенно усложняется и этот случай здесь рассматривать не будем.

Отметим, что $\sum n_i^2 = |n|^2$, т. е. является квадратом длины (нормы) вектора помехи. Следовательно, с учетом (7.8) и (7.10)

$$p(y|s_i) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left\{-\frac{|y-s_i|^2}{2\sigma^2}\right\}. \quad (7.11)$$

Отбросив множители, не зависящие от номера сигнала i , решающее правило (7.7) можно представить в виде

$$R_B = \max_i P(x_i) \exp\{-|y-s_i|^2/2\sigma^2\}. \quad (7.12)$$

Приемник, работающий по алгоритму (7.12), называется *байесовским* или *приемником максимальной апостериорной вероятности*. Если априорные вероятности элементов $P(x_i)$ одинаковы, то решающее правило упрощается:

$$R_{\text{мп}} = \max_i \exp\{-|y-s_i|^2/2\sigma^2\}. \quad (7.13)$$

Соответствующий приемник называется *приемником максимального правдоподобия*. Правило (7.13) раскрывает механизм работы оптимального приемника. Напомним суть его работы (см. п. 3.4).

Получив вектор y , с помощью обработки реализации $y(t)$ необходимо вычислить расстояние от его конца до концов векторов всех возможных сигналов s_i и вынести решение в пользу того сигнала, для которого величина $|y-s_i|$ будет минимальной, так как именно в этом случае функция (7.13) достигнет максимума. Коротко можно сказать, что оптимальный приемник выносит решение в пользу сигнала «ближайшего» к $y(t)$. Описанное правило наглядно представлено на рис. 7.1, где показаны векторы двух возможных сигналов s_1 и s_2 и вектор некоторой конкретной реализации $y=y_k$. Так как $y_k-s_1 < y_k-s_2$, то в данном случае приемник вынесет решение (возможно ошибочное), что в реализации $y(t)$ присутствует сигнал $s(t, x_1)$.

Выражение (7.13) достигает максимума при минимуме показателя экспоненты. Следовательно, правило (7.13) можно записать в ином виде:

$$R_{\text{мп}} = \min_i |y-s_i|^2$$

или, учитывая векторное представление y и s_i ,

$$R_{\text{мп}} = \min_i \left(\sum_j y_j^2 - 2 \sum_j y_j a_{ij} + \sum_j a_{ij}^2 \right). \quad (7.14)$$

Здесь первый член в скобках не зависит от номера i . Последний член — есть энергия i -го сигнала. Если энергии всех сигналов одинаковы, что обычно имеет место, то этот член также не зависит от номера i . Таким образом, решающее правило можно записать так:

$$R_{\text{мп}} = \max_i \sum_{j=1}^N y_j a_{ij}. \quad (7.15)$$

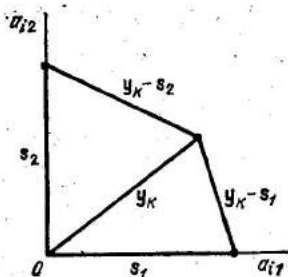


Рис. 7.1. К пояснению алгоритма работы оптимального приемника

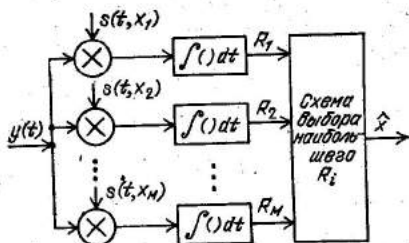


Рис. 7.2. Структурная схема оптимального корреляционного приемника M равновероятных сигналов

Справедливость такого перехода обусловлена тем, что второй член в (7.14) имеет знак минус и выражение (7.14) минимизируется, если этот член достигает максимума. Выражение (7.15) уже позволяет определить структуру оптимального приемника. Однако удобнее это выражение представить в другом виде. Действительно, учтем, что

$$\sum_{j=1}^N y_j a_{ij} = \int_{kT_c}^{(k+1)T_c} y(t) s(t, x_i) dt. \quad (7.16)$$

Тогда окончательно получим

$$R_{\text{мп}} = \max_i \int_{kT_c}^{(k+1)T_c} y(t) s(t, x_i) dt. \quad (7.17)$$

Структурная схема, реализующая последовательность операций, соответствующих правилу (7.17), представлена на рис. 7.2: эта структура называется *оптимальным корреляционным приемником*, так как основная операция, лежащая в его основе, — операция корреляции $y(t)$ со всеми возможными сигналами $s(t, x_i)$.

Из проведенного рассмотрения следует, что в состав оптимального приемника должны входить генераторы, вырабатывающие образцы сигналов $s(t, x_i)$, тождественные тем, которые используются на передатчике. Кроме того, между работой генераторов передатчика и приемника должна соблюдаться синхронность и синфазность, т. е. обеспечиваться идеальная синхронизация.

7.3. ФИЗИКА РАБОТЫ ОПТИМАЛЬНОГО КОРРЕЛЯЦИОННОГО ПРИЕМНИКА

Полагая, что линия связи идеальна, т. е. не вносит искажений, рассмотрим физику работы приемника на примере бинарной системы связи, использующей ортогональные сигналы в виде отрезков гармонических колебаний на интервалах $kT_c \leq t \leq (k+1)T_c$:

$$s(t, x_1) = S \sin \omega_1 t, \quad s(t, x_2) = S \sin \omega_2 t.$$

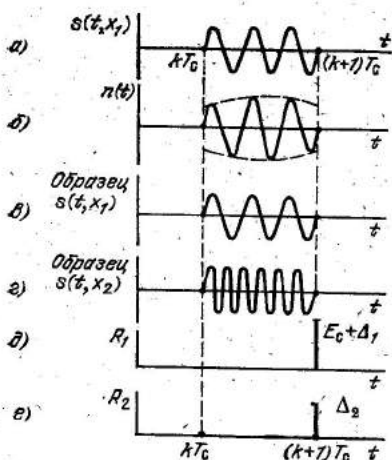


Рис. 7.3. К пояснению физики работы оптимального корреляционного приемника при двух сигналах

Предполагается, что разность частот $\Delta\omega = |\omega_2 - \omega_1|$ выбран так, чтобы выполнялось условие $\Delta\omega \geq 2\pi/T_c$, т. е. чтобы сигналы были практически ортогональными.

Пусть на рассматриваемом интервале $kT_c \leq t \leq (k+1)T_c$ передан сигнал $s(t, x_1)$. Тогда в точке приема действует реализация из смеси сигнала и шума

$$y(t) = S \sin \omega_1 t + n(t).$$

(Ослабление сигналов при распространении по неискажающей линии не приводит к изменению их формы и поэтому в данном рассмотрении не учитывается.) Сигнал и помехи на рассматриваемом интервале показаны на рис. 7.3, а, б. Корреляционный приемник вычисляет две величины:

$$R_1 = \int_{kT_c}^{(k+1)T_c} [S \sin \omega_1 t + n(t)] S \sin \omega_1 t dt,$$

$$R_2 = \int_{kT_c}^{(k+1)T_c} [S \sin \omega_1 t + n(t)] S \sin \omega_2 t dt.$$

Каждый интеграл состоит из двух частей: одна обусловлена интегрированием произведения первого сигнала, входящего в $y(t)$, на образец первого (или второго) сигнала, другая — интегрированием произведения помехи, входящей в $y(t)$ на образец первого (bR_1) или второго (bR_2) сигнала. (Образцы сигналов показаны на рис. 7.3, в, г.) В рассматриваемом случае первая часть интеграла в R_1 равна энергии сигнала, так как

$$\int_{kT_c}^{(k+1)T_c} (S \sin \omega_1 t)^2 dt = E_c,$$

а в R_2 равна нулю, так как сигналы практически ортогональны, т. е.

$$\int_{kT_c}^{(k+1)T_c} S \sin \omega_1 t S \sin \omega_2 t dt \approx 0.$$

Что касается вторых частей обоих интегралов, то они образуют случайные величины

$$\Delta_1 = \int_{kT_c}^{(k+1)T_c} n(t) S \sin \omega_1 t dt, \quad \Delta_2 = \int_{kT_c}^{(k+1)T_c} n(t) S \sin \omega_2 t dt.$$

Таким образом,

$$R_1 = E_c + \Delta_1, \quad R_2 = \Delta_2.$$

Эти величины показаны на рис. 7.3, *д*, *е*. При многократном повторении описанной процедуры вычислений чаще всего окажется, что при передаче сигнала $s(t, x_1)$ $E_c + \Delta_1 > \Delta_2$ и, таким образом, в соответствии с правилом (7.17) приемник почти всегда будет выносить правильное решение. Однако иногда, несмотря на то, что E_c — постоянная положительная величина, может оказаться, что $E_c + \Delta_1 < \Delta_2$. Тогда в соответствии с (7.17) приемник вынесет неправильное решение, т. е. произойдет ошибка.

Продолжим рассмотрение особенностей работы оптимального приемника. Предположим, что при тех же исходных условиях теперь в системе связи используются противоположные сигналы, например,

$$s(t, x_1) = S \sin \omega_0 t, \quad s(t, x_2) = -s(t, x_1) = -S \sin \omega_0 t.$$

В этом случае, как легко проверить, величины R_1 и R_2 принимают следующие значения:

$$R_1 = E_c + \Delta_1, \quad R_2 = -E_c + \Delta_2.$$

Ясно, что из-за наличия постоянного отрицательного слагаемого в $R_2 (-E_c)$ ошибочные решения ($R_2 > R_1$) при прочих равных условиях будут приниматься реже, чем в предыдущем случае и вероятность ошибок снизится. Рассмотренные примеры показывают, что применение оптимального приемника позволяет учесть не только энергию сигналов, но и их форму.

7.4. ПРИЕМНИК НА СОГЛАСОВАННЫХ ФИЛЬТРАХ

Решающее правило (7.17) непосредственно определяет структуру оптимального приемника в корреляционном варианте (рис. 7.2). Однако возможен и другой вариант реализации — с помощью согласованных фильтров.

Фильтр называется *согласованным* с сигналом, если его импульсный отклик

$$h_i(t) = as(T_c - t, x_i), \quad (7.18)$$

где a — коэффициент, а T_c — длительность сигнала $s(t, x_i)$. Операция (7.18), которую можно назвать обращением, показана на рис. 7.4, *а*. Покажем, что колебание на выходе согласованного фильтра в момент $t = T_c$ с точностью до константы равно корреляции с образцом сигнала. Действительно,

$$u_{\text{вых}}(t) = \int_0^t u_{\text{вх}}(\tau) h(t - \tau) d\tau = \int_0^t y(\tau) as(T_c - \tau, x_i) d\tau.$$

В момент $t = T_c$

$$u_{\text{вых}}(T_c) = a \int_0^{T_c} y(\tau) s(\tau, x_i) d\tau, \quad (7.19)$$

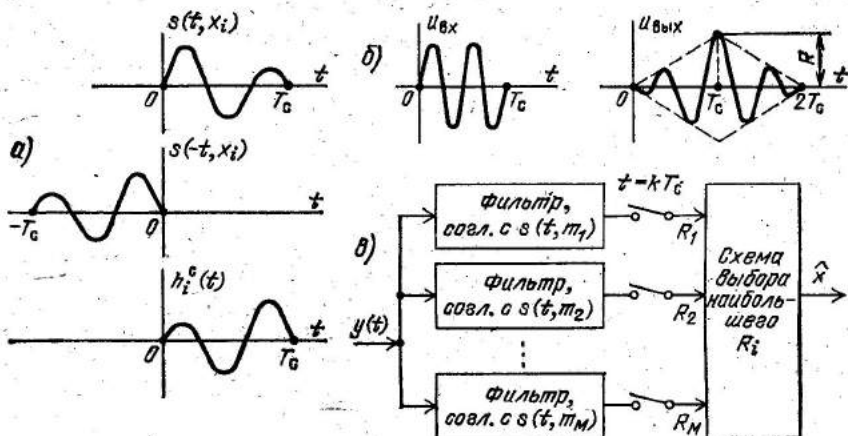


Рис. 7.4. Пояснение идеи построения оптимального приемника на согласованных фильтрах

что с точностью до константы a совпадает с результатом корреляции реализации $y(t)$ с сигналом $s(t, x_i)$. Пример прохождения сигнала в виде отрезка гармонического колебания через согласованный фильтр показан на рис. 7.4, б. Из рисунка видно, что в момент $t = T_c$ напряжение сигнала на выходе достигает максимума. Для помехи такого «пика» не получается. Поэтому согласованный фильтр создает при $t = T_c$ наибольшие возможные отношения «сигнал/помеха». Он полностью заменяет перемножитель и интегратор корреляционного приемника.

Схема оптимального приемника на согласованных фильтрах показана на рис. 7.4, в. Ключи позволяют произвести отсчет величин $u_{\text{вых}}(T_c)$, которые затем поступают на схему выбора наибольшего значения.

Рассмотрим физику работы согласованного фильтра. Если $G_s(\omega)$ — спектральная функция заданного сигнала, то

$$s(t, x_i) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{G}_s(\omega) e^{j\omega t} d\omega. \quad (7.20)$$

С учетом (7.18) можно записать выражение для коэффициента передачи согласованного фильтра в виде

$$K_i^c(\omega) = \int_{-\infty}^{\infty} h_i(t) e^{-j\omega t} dt = e^{-j\omega T_c} \dot{G}_s^*(\omega). \quad (7.21)$$

Из (7.21) следует, что этот коэффициент с точностью до множителя $e^{-j\omega T_c}$ комплексно сопряжен спектральной функции сигнала.

Выделим теперь бесконечно малую спектральную составляющую сигнала (7.20) на произвольной частоте ω_k в полосе $d\omega$. Ее комплексная амплитуда

$$\dot{U}_{\text{вх}} = \frac{1}{2\pi} \dot{G}_s(\omega_k) d\omega. \quad (7.22)$$

Комплексная амплитуда этой составляющей на выходе согласованного фильтра равна

$$\begin{aligned} \dot{U}_{\text{ВЫХ}} &= \dot{U}_{\text{ВХ}} K_i^c = \frac{1}{2\pi} G_s(\omega_h) d\omega e^{-j\omega_h T_c} G_s^*(\omega) = \\ &= \frac{1}{2\pi} G_s^2(\omega_h) e^{-j\omega_h T_c} d\omega. \end{aligned}$$

Мгновенное значение этой составляющей определяется выражением

$$u_{\text{ВЫХ}}(t) = \frac{1}{2\pi} G_s^2(\omega_h) d\omega \cos(\omega_h t - \omega_h T_c).$$

В момент окончания передачи сигнала ($t = T_c$) эта составляющая примет значение

$$u_{\text{ВЫХ}}(t)|_{t=T_c} = \frac{1}{2\pi} G_s^2 d\omega \cos(\omega_h T_c - \omega_h T_c) = \frac{1}{2\pi} G_s^2 d\omega.$$

Отсюда видно, что все гармонические составляющие сигнала, согласованного с фильтром, складываются на выходе фильтра в момент $t = T_c$ «согласованно», создавая пик, показанный на рис. 7.4, б.

7.5. РЕАЛИЗАЦИЯ СОГЛАСОВАННОГО ФИЛЬТРА. ФИЛЬТР, СОГЛАСОВАННЫЙ ПО ПОЛОСЕ

Построение фильтра, обладающего заданным импульсным откликом, является задачей теории синтеза линейных цепей. Решение ее не всегда возможно и, как правило, затруднено. Однако для весьма распространенных сигналов виде отрезков гармонического косинусоидального колебания согласованный фильтр можно реализовать сравнительно просто.

Пусть $s(t, x_i) = S \cos \omega_i t$, причем за время T_c укладывается целое число периодов частоты ω_i (рис. 7.5, а). Тогда, как легко проверить,

$$h_i(t) = a s(t, x_i), \quad (7.23)$$

т. е. отклик согласованного фильтра повторяет форму сигнала.

Рассмотрим колебательный контур LC , возбуждаемый δ -импульсом тока (рис. 7.5, б). Напряжение на его выходе

$$u_{\text{ВЫХ}}(t) = \frac{1}{C} e^{-\alpha t} \cos \omega_p t, \quad (7.24)$$

где $\alpha = r/2L$, $\omega_p = 1/\sqrt{LC}$. Если α весьма мало, а $\omega_p = \omega_i$, то на интервале $0 < t < T_c$ колебание (7.24) практически совпадает с импульсным откликом (7.23). Для того чтобы выходное колебание контура совпало с (7.23) на всем интервале, достаточно погасить колебание контура в момент $t = T_c$, для чего можно использовать короткозамыкающий ключ. Таким образом, схема, приведенная на рис. 7.5, в, имеет практически такой же отклик, как согласо-

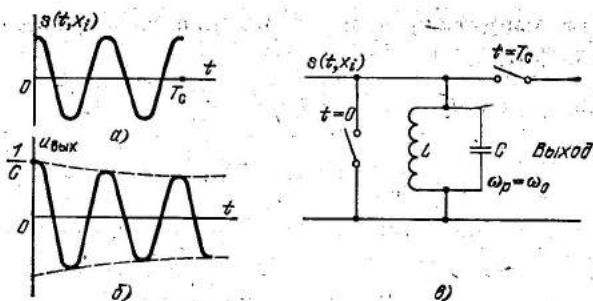


Рис. 7.5. Реализация согласованного фильтра с помощью колебательного контура и ключей

важный фильтр (7.23), и поэтому она является одной из реализаций согласованного фильтра для сигнала в форме отрезка косинусоидального колебания.

В большинстве реальных систем передачи дискретной информации согласованные фильтры даже для простых сигналов в виде отрезков гармонического колебания в настоящее время не используются в основном из-за относительной сложности системы и необходимости строгой синхронизации передатчика и приемника. В большинстве случаев используются колебательные контуры, согласованные с сигналами только по полосе. Если сигнал представляет собой отрезок гармонического колебания длительностью T_c с амплитудой S , то при воздействии его на идеальный полосовой П-образный фильтр с полосой пропускания $\Delta\omega$, амплитуда напряжения на выходе (без учета запаздывания) изменяется по закону [1]

$$U = \frac{1}{\pi} S \left[\text{Si} \left(\frac{\Delta\omega}{2} t \right) - \text{Si} \frac{\Delta\omega}{2} (t - T_c) \right],$$

где $\text{Si}(x)$ — интегральный синус.

Можно показать, что максимум амплитуды сигнала равен

$$U_{\max} = (2/\pi) S \text{Si}(\Delta\omega T_c/4).$$

Эффективное значение шума пропорционально корню из полосы фильтра

$$U_{\text{ш}} = k \sqrt{\Delta\omega}.$$

Отношение сигнала к шуму на выходе П-образного фильтра определяется выражением

$$\frac{U_{\max}}{U_{\text{ш}}} = k 2 S \text{Si} \left(\frac{\Delta\omega T_c}{2} \right) / \sqrt{\Delta\omega},$$

которое достигает максимума, как нетрудно определить, при по-

$$\Delta\omega_{\text{опт}} = 8,6/T_c \quad \text{или} \quad \Delta f_{\text{опт}} = 1,37/T_c \approx 1,4/T_c.$$

Эта полоса называется *оптимальной*. Для реальных избирательных фильтров это соотношение остается приблизительно справедливым.

Оптимальное значение полосы не очень критично. Как показывает анализ при расширении или сужении полосы фильтра в два раза отношение сигнал/помеха ухудшается приблизительно на 15%. Поэтому в инженерной практике для определения оптимальной полосы часто пользуются выражением

$$\Delta f_{\text{опт}} = (1 \dots 1,5) / T_c.$$

7.6. РАСЧЕТ ВЕРОЯТНОСТИ ОШИБОК ПРИ РАСПОЗНАВАНИИ ДИСКРЕТНЫХ СИГНАЛОВ

Оптимальный приемник, получив реализацию $y(t)$, выносит решение в пользу того сигнала, к которому вектор y окажется ближе всего. Физическим объяснением этого является большая вероятность реализации слабой помехи (и соответственно малых норм векторов n).

Правило (7.13), таким образом, делит пространство сигналов на области сигналов J_i , состоящие из точек, ближайших к вектору соответствующего сигнала (см. п. 3.4). Проиллюстрируем такое разбиение на области сигналов примером.

Пример. Рассмотрим четыре биортогональных сигнала с векторами

$$s_1 = (2, 0), s_2 = (0, 2), s_3 = (-2, 0), s_4 = (0, -2).$$

Соответствующие векторы показаны на рис. 7.6.а. Соединим концы векторов и проведем из начала координат пунктирные прямые, проходящие через середины полученных отрезков. Легко убедиться, что, например, все точки заштрихованной области J_1 находятся ближе всего к концу вектора сигнала s_1 . Поэтому они составляют область J_1 сигнала $s(t, x_1)$. Аналогично определяются области остальных трех сигналов.

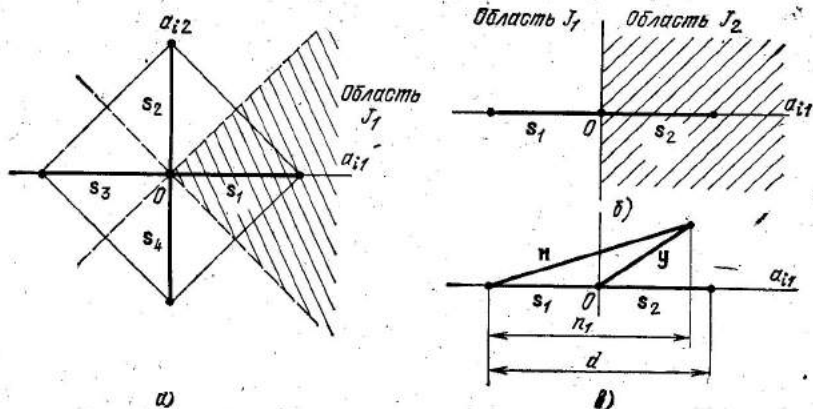


Рис. 7.6. Области сигналов при вычислении вероятности ошибок распознавания

Из сказанного следует, что условная вероятность правильного распознавания элемента x_i равна вероятности попадания конца вектора y_i , образованного сигналом s_i и шумом в области J_i :

$$P_{\text{прав } x_i} = P(y_i \in J_i),$$

$$y_i = s_i + n.$$

Следовательно, вероятность ошибок с учетом передачи всех элементов

$$P_{\text{ош}} = 1 - \sum_{i=1}^M P(x_i) P(y_i \in J_i). \quad (7.25)$$

Расчет $P_{\text{ош}}$ для произвольных сигналов при произвольном M оказывается весьма сложным, громоздким.

Для пояснения метода расчета найдем $P_{\text{ош}}$ при передаче информации бинарными противоположными сигналами. Векторы и области сигналов J_1 и J_2 для этого случая показаны на рис. 7.6, б. Предположим, что многократно передается первый сигнал. Тогда вероятность ошибочного приема такого сигнала

$$P_{\text{ош} | x_1} = P(y_1 \in J_2).$$

Случай ошибки при приеме сигнала s_1 показан на рис. 7.6, в. Из этого рисунка видно, что вектор y попадает в область J_2 , если координата n_1 конца вектора помехи n превысит величину $s_1 = \sqrt{E_c}$.

Следовательно, вероятность ошибочного приема сигнала s_1 (т. е. символа x_1) определяется вероятностью того, что случайная величина n_1 превысит величину $d/2 = \sqrt{E_c}$:

$$P_{\text{ош} | x_1} = P(n_1 > d/2 = \sqrt{E_c}).$$

Координата n_1 помехи — нормальная случайная величина с нулевым средним и дисперсией $\sigma_1^2 = N_0/2$, следовательно, одномерная плотность ее вероятности определяется выражением [2]:

$$p_1(n_1) = (1/\sqrt{2\pi\sigma_1^2}) \exp\{-n_1^2/2\sigma_1^2\} = (1/\sqrt{\pi N_0}) \exp\{-n_1^2/N_0\}.$$

С учетом этого выражения после некоторых преобразований

$$P_{\text{ош} | x_1} = P\left(n_1 > \frac{d}{2}\right) = \int_{d/2}^{\infty} p_1(n_1) dn_1 = \frac{1}{\sqrt{2\pi\alpha}} \int_{\alpha}^{\infty} e^{-x^2/2} dx = V(\alpha), \quad (7.26)$$

где $\alpha = 0,5 d/\sqrt{N_0/2} = d/\sqrt{2N_0}$.

Функция $V(\alpha)$ называется *интегралом ошибок*. Она табулирована и приводится в ряде книг. Приняв во внимание, что для рассматриваемого случая $d = 2\sqrt{E_c}$, имеем

$$\alpha = \sqrt{2E_c/N_0}; \quad P_{\text{ош} | x_1} = V(\sqrt{2E_c/N_0}). \quad (7.27)$$

Из-за симметрии картины (см. рис. 7.6) вероятность ошибок при передаче элемента x_2 выразится такой же формулой, т. е.

$$P_{\text{ош} | x_2} = V(\sqrt{2E_c/N_0}).$$

Кроме функции $V(\alpha)$ в приложениях часто пользуются функцией

$$\Phi(\alpha) = 0,5 - V(\alpha) = \frac{1}{\sqrt{2\pi}} \int_0^{\alpha} e^{-x^2/2} dx, \quad (7.28)$$

которая называется интегралом вероятностей или функцией Лапласа.

Вероятности $P_{\text{ош}}|x_1$ и $P_{\text{ош}}|x_2$ являются условными. Первая из них определяет ошибку при передаче символа x_1 , т. е. прием x_2 вместо переданного x_1 , а вторая — ошибочный прием символа x_1 вместо переданного x_2 .

Если априорные вероятности передачи символов $P(x_1)$ и $P(x_2)$ известны, то полная вероятность ошибочного приема переданного символа определяется выражением

$$P_{\text{ош}} = P(x_1)P_{\text{ош}}|x_1 + P(x_2)P_{\text{ош}}|x_2.$$

Если $P(x_1) = P(x_2)$, то

$$P_{\text{ош}} = \frac{1}{2} [P_{\text{ош}}|x_1 + P_{\text{ош}}|x_2].$$

Для бинарных сигналов $P_{\text{ош}}|x_1 = P_{\text{ош}}|x_2$, поэтому

$$P_{\text{ош}} = V(\sqrt{2E_c/N_0}) = 0,5 - \Phi(\sqrt{2E_c/N_0}). \quad (7.29)$$

Если передача информации производится бинарными ортогональными сигналами, то расстояние между концами соответствующих им векторов будет равно $d = \sqrt{2E_c}$ (см. рис. 6,5, б). Нетрудно убедиться, что при $P(x_1) = P(x_2)$ вероятность ошибки при оптимальном приеме таких сигналов определяется выражением

$$P_{\text{ош}} = V(\sqrt{E_c/N_0}) = 0,5 - \Phi(\sqrt{E_c/N_0}). \quad (7.30)$$

Графики, соответствующие выражениям (7.29) и (7.30), приведены на рис. 7.7. Исследование этих графиков показывает, что:

а) величина ошибок при оптимальном приеме зависит от отношения энергии сигналов к спектральной плотности помехи и от геометрической конфигурации сигналов;

б) для получения одной и той же вероятности ошибок при приеме противоположных сигналов требуется в два раза меньшая энергия сигналов, чем при приеме ортогональных;

в) при обоих типах сигналов наблюдается «пороговый эффект»: вероятность ошибок резко снижается после того, как величина E_c/N_0 превышает несколько единиц.

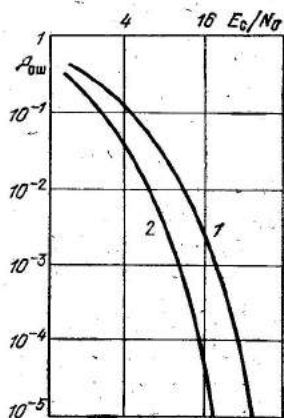


Рис. 7.7. Кривые вероятности ошибок оптимального распознавания бинарных сигналов для ортогональных (кривая 1) и противоположных сигналов (кривая 2)

Поступая в соответствии с изложенной методикой, можно найти вероятность ошибок оптимального приема и других видов сигналов [3].

В заключение отметим, что при инженерных расчетах часто удобнее пользоваться не табулированными значениями функций $V(\alpha)$ или $\Phi(\alpha)$, а аппроксимациями этих функций. Весьма удобной аппроксимацией функции $V(\alpha)$ является следующая:

$$V(\alpha) \approx W_1(\alpha) / \sqrt{1 + \alpha^2},$$

где

$$W_1(\alpha) = \frac{1}{\sqrt{2\pi}} e^{-\alpha^2/2}.$$

При $\alpha > 2$ такая аппроксимация обеспечивает точность не хуже 5%. Более грубой, но еще более удобной аппроксимацией будет

$$V(\alpha) \approx \sqrt{\frac{\pi}{2}} W_1(\alpha) = \frac{1}{2} e^{-\alpha^2/2}.$$

При $\alpha > 4$ такая аппроксимация во многих случаях вполне приемлема и значительно облегчает расчеты.

7.7. ПРИЕМ СИГНАЛОВ В ЛИНИЯХ, ВНОСЯЩИХ СЛУЧАЙНОЕ ОСЛАБЛЕНИЕ И СДВИГ ФАЗЫ

В п. 6.7 было показано, что если импульсный отклик линии представляет собой δ -функцию, то такая линия только ослабляет передаваемый сигнал, не изменяя его формы. Пусть ослабление сигнала a — медленно изменяющаяся случайная величина, практически постоянная на интервалах длительностью T_c . Если бы a была постоянной и известной величиной, то осуществлялся бы прием точно известных сигналов с решающим правилом

$$R = \max P(x_i) \cdot p(y|s_i, a). \quad (7.31)$$

При случайном значении a следует усреднить результат по закону распределения $p(a)$; тогда при равновероятностных сигналах решающее правило примет вид

$$R = \max_i \int_0^\infty p(y|s_i, a) p(a) da. \quad (7.32)$$

Из соотношения (7.32) следует, что при таком подходе структура оптимального приемника останется прежней (инвариантной к случайным значениям a). Вероятность же ошибок (при прочих равных условиях) возрастает. Действительно, при известном значении a и бинарных сигналах выражение для $P_{\text{ош}}$ в зависимости от вида сигналов определяется соотношениями (7.29) и (7.30). При случайном значении a эти выражения необходимо усреднить по $p(a)$. В частности, для противоположных сигналов усредненное значение вероятности ошибки $P_{\text{ош}}$ должно определяться в соответствии с выражением

$$\overline{P_{\text{ош}}} = \bar{V} \left(a \sqrt{\frac{2E_c}{N_0}} \right) = \int_0^{\infty} V \left(a \sqrt{\frac{2E_c}{N_0}} \right) p(a) da. \quad (7.33)$$

Для распределения $p(a)$, подчиняющегося закону Рэлея можно показать, что [3]

$$\overline{P_{\text{ош}}} = \frac{1}{2} \left[1 - \frac{a}{\sqrt{2+a^2}} \right], \quad (7.34)$$

где $a^2 = 2E_c/N_0$.

Нетрудно видеть, что при одинаковых значениях a вероятность ошибок, рассчитанная по формуле (7.34), значительно превышает вероятность ошибок, найденную по формуле (7.29). Физическая причина увеличения вероятности ошибок ясна: возрастание a (и, следовательно, энергии сигналов) приводит к некоторому уменьшению вероятности ошибок, однако падение a приводит к более значительному возрастанию этой вероятности вследствие отмеченного выше «порогового эффекта» (нелинейной зависимости $P_{\text{ош}}$ от E_c/N_0).

Рассмотрим далее случай, когда линия вносит в сигналы только случайный сдвиг начальной фазы, имеющий место в подавляющем большинстве реальных ситуаций. При этом, если

$$s(t, x_i) = S(t) \cos \omega_i t \quad (0 \leq t \leq T_c),$$

то сигналы на выходе линии (входе приемника)

$$s_{\text{вых}}(t, x_i) = S(t) \cos(\omega_i t - \theta). \quad (7.35)$$

Выходные сигналы (7.35) можно представить в виде двух составляющих со случайными амплитудами, но постоянными фазами:

$$s_{\text{вых}}(t, x_i) = S(t) \cos \theta \cos \omega_i t - S(t) \sin \theta \sin \omega_i t = a S(t) \cos \omega_i t - b S(t) \sin \omega_i t, \quad (7.36)$$

где a и b могут, в отличие от предыдущего случая, принимать и положительные и отрицательные значения.

Из (7.36) видно, что действие линии можно свести к появлению в точке приема двух составляющих сигнала: косинусоидальной и синусоидальной. Анализ этого случая, связанный с выполнением усреднения по обоим случайным параметрам a и b , довольно громоздок. Отсылая интересующихся к работам [2, 4, 5], приведем конечное выражение для решающего правила:

$$\hat{R} = \max_t \left[\left(\int_{kT_c}^{(k+1)T_c} y(t) S(t) \cos \omega_i t dt \right)^2 + \left(\int_{kT_c}^{(k+1)T_c} y(t) S(t) \sin \omega_i t dt \right)^2 \right]. \quad (7.37)$$

Из него следует, что оптимальный приемник производит корреляцию принятой реализации $y(t)$ с образцами обоих слагаемых сигнала (7.36). Возведение результатов в квадраты перед сложением и выбором максимума вызвано тем, что величины a и b могут быть как положительными, так и отрицательными.

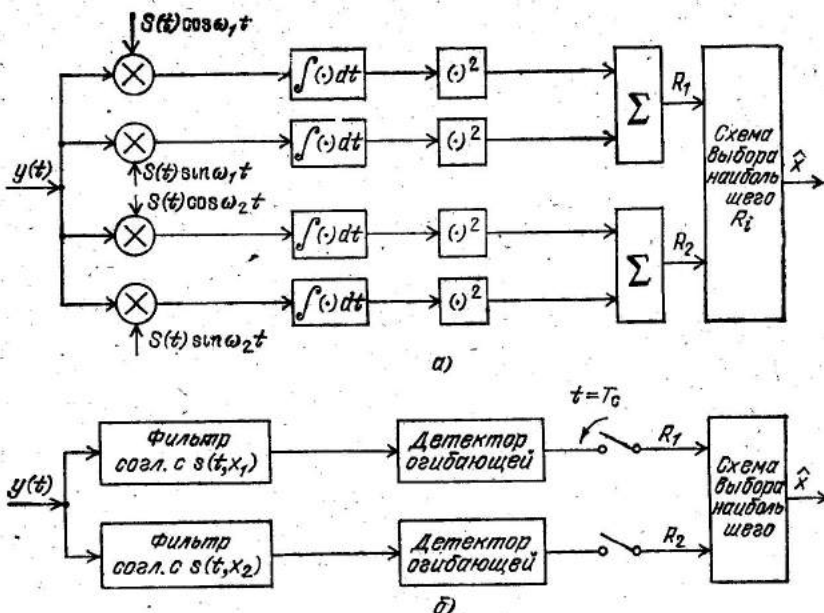


Рис. 7.8. Структурные схемы оптимальных приемников при неизвестной случайной фазе сигналов, реализованных на основе корреляционной обработки (а) и на основе согласованных фильтров (б)

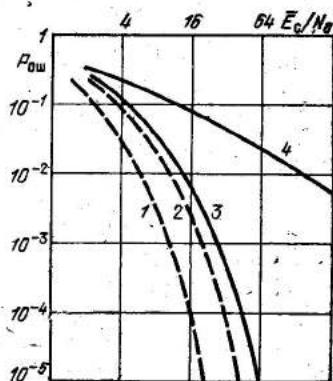
Структурная схема, соответствующая алгоритму (7.37) при $M=2$, приведена на рис. 7.8, а. Этот алгоритм можно реализовать и с помощью согласованных фильтров (рис. 7.8, б). В отличие от схемы, представленной на рис. 7.4, в, здесь содержатся детекторы огибающих выходных колебаний согласованных фильтров, после которых и производится отсчет. Физика процессов также ясна: если на вход согласованного с сигналом $s(t, x_1)$ фильтра подать сдвинутый по фазе сигнал, то в силу линейности фильтра произойдет запаздывание колебания и на выходе фильтра. Поэтому отсчет в момент $t=T_c$ (рис. 7.4, б) не совпадет с максимумом напряжения. В силу случайности этого сдвига наилучшей стратегией оказывается отсчет огибающей, а не мгновенного значения колебания.

Сравним случай приема сигналов при отсутствии случайной фазы (т. е. точно известных по форме сигналов) и при наличии случайной фазы. Первый случай принято называть *когерентным*, а второй — *некогерентным приемом* (именно этот случай чаще всего имеет место на практике).

При когерентном приеме (и ортогональных сигналах передатчика) $P_{\text{ош}}$ определяется соотношением (7.30). При некогерентном приеме вероятность ошибки равна [3]

$$P_{\text{ош}} = 1/2 \exp \{-E_c/2 N_0\}. \quad (7.38)$$

Рис. 7.9. Кривые вероятностей ошибок при приеме бинарных сигналов точно известных и противоположных (1), ортогональных (2), со случайной фазой (3) и случайными фазой и амплитудой (4)



Сравнивая выражения (7.30) и (7.38) при одинаковом значении вероятности ошибки, можно установить, какой энергетический проигрыш дает применение некогерентного приема по сравнению с когерентным. Расчеты показывают, что для обеспечения $P_{ош} = 10^{-3} - 10^{-6}$ при некогерентном приеме требуется увеличение энергии сигнала на 15—30% по сравнению с когерентным [3], т. е. проигрыш невелик.

В более общем случае неидеальность линии обуславливает случайные изменения амплитуды и фазы. Вероятность ошибок от этого увеличивается, так как независимо действуют оба рассмотренных фактора. Можно показать [3], что в этом случае вероятность ошибок при распознавании бинарных ортогональных сигналов равна

$$P_{ош|a, \varphi} = 1/(2 + \bar{E}_c/N_0), \quad (7.39)$$

где \bar{E}_c — среднее значение энергии принимаемых сигналов.

Кривые $P_{ош}$, соответствующие рассмотренным случаям, приведены на рис. 7.9.

7.8. ВЫБОР И РАСЧЕТ НЕКОТОРЫХ ПАРАМЕТРОВ СИСТЕМ ПЕРЕДАЧИ ДИСКРЕТНЫХ СООБЩЕНИЙ

В настоящее время существует несколько различных методов расчета основных энергетических параметров линии связи. Такие методики обычно направлены на обеспечение необходимого отношения сигнал/помеха в точке приема, при котором гарантируется заданная точность передачи (вероятность ошибок).

Следует отметить невозможность однозначного расчета даже основных параметров линии связи, так как переменных оказывается больше, чем связывающих их уравнений. Поэтому часть показателей и параметров приходится выбирать исходя из практических соображений, опираясь на предшествующий опыт. Необходимо также отметить невозможность прямого расчета системы, т. е. по заданным конечным показателям однозначно определить параметры системы. Поэтому на практике приходится рассчитывать несколько вариантов и из них выбирать «наилучший» по некоторому комплексу показателей. Наиболее часто обычно задают следующие показатели и параметры:

1. Размер первичного алфавита дискретных сообщений M . Примерами таких алфавитов являются: число букв, цифр и слу-

жебных знаков в различных кодовых словарях, используемых при передаче буквенных и числовых текстов (см. [3]); число уровней квантования непрерывных сообщений при их преобразовании в дискретные и т. п.

2. Требуемая техническая скорость передачи элементов дискретных сообщений \mathcal{R}_T [Бод].

3. Допустимая вероятность ошибки приема на элемент сообщения $P_{ош}$.

Кроме того, должен быть известен (или задан) вид и уровень помех. Обычно помеху считают гауссовским случайным процессом с равномерной спектральной плотностью N_0 . Возможен следующий порядок выбора и расчета параметров.

1. Выбирается вид первичного кодирования. В настоящее время при первичном кодировании наиболее широко применяется натуральный двоичный код. Число разрядов такого кода m определяется из соотношения: $2^m \geq M$. Как правило, при первичном кодировании используются стандартные коды с $m=7$ (для передачи буквенно-цифровых текстов) или с $m=4$ (для передачи только цифровых данных) [3].

2. Определяется допустимое время на передачу одного элемента (символа) сообщения, представленного первичным кодом $\tau_0 = 1/\mathcal{R}_T$ [с].

3. Исходя из свойств линии связи выбирается вид сигнала. В линиях с постоянными или незначительно изменяющимися параметрами предпочтительнее использовать противоположные сигналы (как наиболее помехоустойчивые). В каналах с заметными случайными изменениями амплитуды и фазы часто применяют ортогональные сигналы с частотной манипуляцией.

4. Для выбранного вида сигналов по заданной величине $P_{ош}$ с учетом характера линии находится (по графикам или расчетам) минимально допустимое значение $(E_c/N_0)_{доп}$.

5. По найденной величине $(E_c/N_0)_{доп}$ и заданному значению N_0 определяется $E_{с,доп} = N_0 (E_c/N_0)_{доп}$.

6. По величине $E_{с,доп}$ и длительности τ_0 определяется минимально допустимая средняя мощность сигналов в точке приема

$$P_{с.вх} = E_{с,доп} / \tau_0 = \mathcal{R}_T E_{с,доп}.$$

Зная ослабление мощности, которое вносит линия при распространении сигналов, можно определить требуемую мощность передатчика. Эта мощность оказывается, как правило, заниженной вследствие предположения об оптимальности приемника, полном согласовании элементов тракта и отсутствии потерь энергии в них. В реальных условиях эти предположения не выполняются. Кроме того, на качество приема влияют искажения в системе, неидеальность синхронизации, нестабильность работы отдельных узлов и устройств системы. Все эти факторы приводят к тому, что величину требуемой мощности сигнала на входе приемника реальной системы приходится увеличивать в несколько раз по сравнению с найденным выше значением.

7.9. ОТНОСИТЕЛЬНОЕ КОДИРОВАНИЕ КАК СПОСОБ БОРЬБЫ СО СЛУЧАЙНОЙ НАЧАЛЬНОЙ ФАЗОЙ СИГНАЛОВ

Рассмотренный до сих пор способ кодирования бинарных элементов в сигналы по логике $x_i \rightarrow s(t, x_i)$, где $i=1$ или 2 , можно назвать *прямым кодированием*. Однако возможен другой способ, при котором выполняется относительное кодирование [4].

Практическое распространение получил способ, называемый *относительной фазовой телеграфией* (ОФТ) или *относительной фазовой манипуляцией* (ОФМ). При этом используются два противоположных сигнала: $s(t, x_1)$ и $s(t, x_2) = -s(t, x_1)$ в виде посылок гармонического колебания с противоположными фазами, а передача выполняется по следующему правилу. Если на i -м интервале передается элемент x_1 , то фаза соответствующего ему сигнала сохраняется такой же, как и предыдущего сигнала на $(i-1)$ -м интервале: если же на i -м интервале передается элемент x_2 , то фаза сигнала изменяется на 180° по отношению к фазе сигнала на $(i-1)$ -м интервале.

В начале каждого сеанса связи необходимо передавать вспомогательную посылку, относительно которой начнется отсчет фазы первой посылки сигнала.

Сущность приема сигналов ОФТ состоит в определении того, сохранилась или изменилась начальная фаза в двух смежных сигналах. Для этого достаточно произвести операцию вычисления величины

$$R = \int_{T_c}^{2T_c} y_2(t) y_1(t - T_c) dt,$$

где $y_1(t)$ и $y_2(t)$ — реализации суммы сигналов и помехи на интервалах $0-T_c$ и T_c-2T_c соответственно, и сравнить R с нулем. При передаче элемента x_1 величина R (чаще всего) будет больше нуля, при передаче x_2 — меньше нуля (рис. 7.10, а).

Анализ вероятности ошибок при приеме сигналов с ОФТ проведен в [5]. По результатам этой работы рассчитаны кривые зависимости $P_{\text{ош}}$ от E_c/N_0 в канале, не вносящем случайных параметров при разных величинах базы сигналов $\Delta f_c T_c$ (см. рис. 7.10, б). На рис. 7.10, в приведены кривые $P_{\text{ош}}$ в однолучевом канале с медленно изменяющимися амплитудой и начальной фазой. Кривая 1 относится к случаю, когда параметры изменяются очень медленно (в двух смежных интервалах T_c они одинаковы), кривая 2 относится к случаю, когда эта скорость конечная, так что $\Delta f_{\text{кан}} T_c = 0,1$ ($\Delta f_{\text{кан}}$ — ширина энергетического спектра выборок $h_0(t)$ и $\bar{h}_0(t)$).

Расширять полосу частот при ОФТ в неискажающем канале нецелесообразно (рис. 7.10, а). При повышении скорости изменения параметров величина $P_{\text{ош}}$ возрастает и остается конечной даже при $E_c/N_0 \rightarrow \infty$ (рис. 7.10, в). Из сравнения рисунков видно, что появление случайных изменений параметров существенно повышает вероятность ошибок [4, 5].

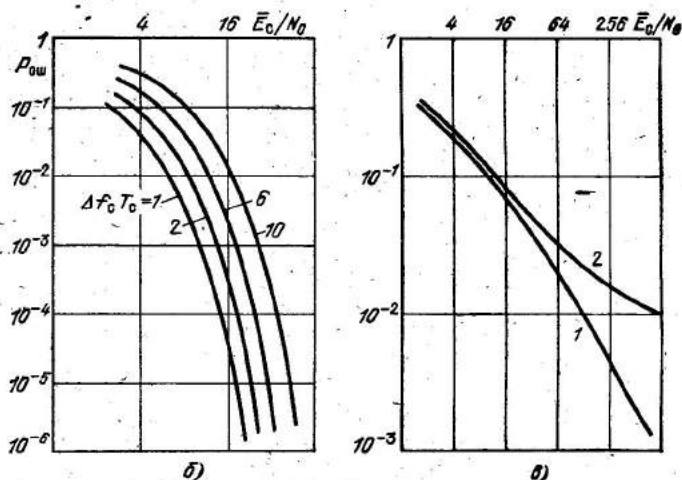
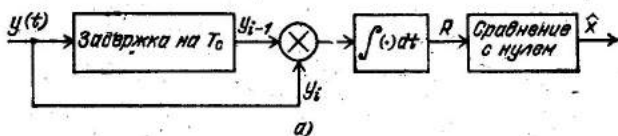


Рис. 7.10. Функциональная схема приемника сигналов с ОФТ (а) и кривые вероятностей ошибок для сигналов с указанной величиной базы (б), сигналов со случайно изменяющейся амплитудой и начальной фазой (в)

7.10. ПРИЕМ СИГНАЛОВ В КАНАЛАХ С БЫСТРО ИЗМЕНЯЮЩИМИСЯ ПАРАМЕТРАМИ. ПРИЕМ ПРИ НЕРАВНОМЕРНЫХ ПО СПЕКТРУ ПОМЕХАХ

До сих пор предполагалось, что в интервалах T_c параметры линии h_0 и \tilde{h}_0 постоянны. Если этого постоянства нет (велика длительность сигналов или быстро флуктуируют параметры), то выражение (6.26) для узкополосных относительно линии сигналов можно представить в виде

$$y(t) = \frac{2\pi}{\Delta\omega_c} h_0(t) s(t, x_i) - \frac{2\pi}{\Delta\omega_c} \tilde{h}_0(t) \tilde{s}(t, x_i) + n(t), \quad (7.40)$$

где $h_0(t)$ и $\tilde{h}_0(t)$ — случайные функции времени, а не случайные величины.

Если эти функции разложить в ортогональные ряды со случайными коэффициентами, по которым необходимо произвести усреднение, то в предположении почти ортогональных сигналов можно получить следующий алгоритм работы оптимального приемника:

$$R = \max_t \int_{kT_c}^{(k+1)T_c} \left[\int_0^t y(\tau) h_\phi^t(t-\tau) d\tau \right]^2 dt. \quad (7.41)$$

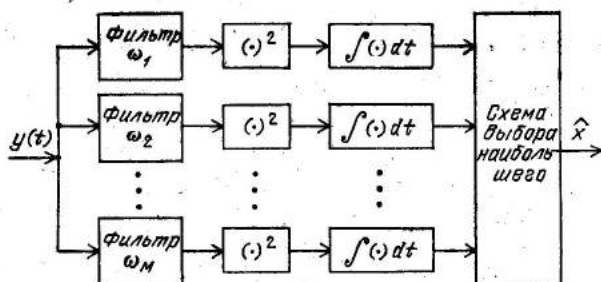


Рис. 7.11. Функциональная схема энергетического приемника

Внутренний интеграл описывает прохождение колебания $y(t)$ через фильтры с откликами $h_{i\phi}$, настроенные на средние частоты сигналов (полоса их определяется скоростью изменения параметров линии), а внешний — есть вычисление энергии реализации. В целом такой способ приема можно назвать *энергетическим*. Его схема приведена на рис. 7.11.

До сих пор предполагалось, что энергетический спектр помех $N(\omega)$ постоянен в полосе сигналов. Для узкополосных в спектральном смысле сигналов ($\Delta f_c T_c \approx 1$) это практически всегда справедливо. При широкополосных сигналах ($\Delta f_c T_c \gg 1$) помеху часто нельзя считать равномерной. В этом случае можно построить близкий к оптимальному приемник, разделив полосу сигнала Δf_c на малые участки, в пределах которых $N(\omega)$ считается постоянной [5].

7.11. ПРИЕМ СИГНАЛОВ С ОЦЕНКОЙ СЛУЧАЙНЫХ ПАРАМЕТРОВ, ВНОСИМЫХ ЛИНИЕЙ. АДАПТИВНЫЙ ПРИЕМ

До сих пор рассмотрение велось в предположении, что случайные параметры h_0 и \tilde{h}_0 , вносимые линией в узкополосные сигналы, не могут быть уточнены и максимум, что можно сделать в рассматриваемой ситуации, — произвести усреднение по их априорному распределению. Однако если параметры изменяются медленно, то на интервале времени $-nT_c, 0$, ($n \gg 1$), предшествующем интервалу $0, T_c$ анализа и вынесения решения, можно произвести оценку (измерение) случайных параметров h_0 и \tilde{h}_0 и использовать ее для повышения качества работы системы.

Нетрудно убедиться, что, обрабатывая последовательность реализаций вида (7.40), т. е. $Y(t)$, можно найти оценки параметров h_0 и \tilde{h}_0 :

$$\hat{h}_0 = \frac{\Delta\omega_c}{2\pi} \int_{-nT_c}^0 Y(t) S(t) dt, \quad \hat{\tilde{h}}_0 = \frac{\Delta\omega_c}{2\pi} \int_{-nT_c}^0 Y(t) \tilde{S}(t) dt, \quad (7.42)$$

где $S(t)$ — последовательность сигналов, содержащихся в реализации на интервале $-nT_c, 0$, $\tilde{S}(t)$ — последовательность сопряженных сигналов.

Параметры h_0 и \tilde{h}_0 в силу ортогональности сигналов $s(t, x_i)$ и $\tilde{s}(t, x_i)$ разделены. Так как интегрирование производится на протяженных интервалах $nT_c \gg T_c$, то точность оценок может быть весьма высокой.

Зная оценки \hat{h}_0 и $\hat{\tilde{h}}_0$, можно построить в точке приема «образцы» искаженных сигналов по правилу

$$s_B(t, x_i) = \frac{2\pi}{\Delta\omega_c} \hat{h}_0 s(t, x_i) - \frac{2\pi}{\Delta\omega_c} \hat{\tilde{h}}_0 \tilde{s}(t, x_i), \quad (7.43)$$

а затем произвести корреляцию с ними принятой реализации на интервале $0, T_c$. С течением времени все операции постепенно смещаются на последующие интервалы.

Анализ рассматриваемого случая позволяет построить следующий алгоритм работы приемника [5]:

$$R = \max_i \int_{kT_c}^{(k+1)T_c} y(t) \frac{2\pi}{\Delta\omega_c} [\hat{h}_0 s(t, x_i) - \hat{\tilde{h}}_0 \tilde{s}(t, x_i)] dt. \quad (7.44)$$

Функциональная схема, соответствующая описанным операциям, приведена на рис. 7.12. Так как последовательность бинарных сигналов в реализации $Y(t)$, т. е. $S(t)$, в точке приема неизвестна, в схеме оценки параметров на умножители на каждом интервале T_c приходится подавать сумму обоих сигналов: $s(t, x_1) + s(t, x_2)$ и $\tilde{s}(t, x_1) + \tilde{s}(t, x_2)$.

Описанный приемник называется *когерентным* приемником с *оценкой параметров* или *адаптивным* приемником.

Аналогично можно построить оптимальный приемник для широкополосных относительно линии сигналов (см. п. 6.8). Подробнее описание адаптивных приемников и выражения для вероятностей ошибок можно найти в [5].

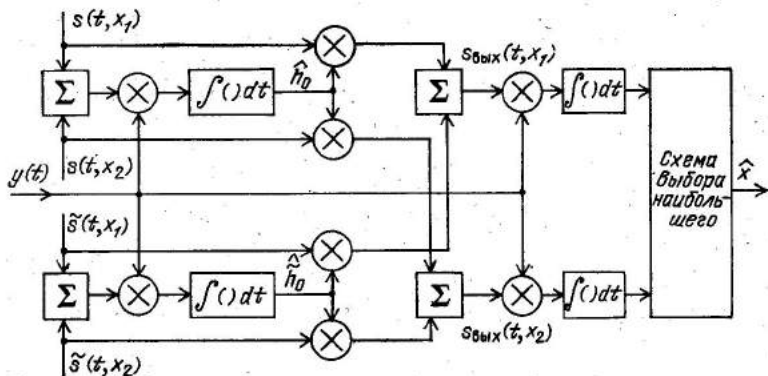


Рис. 7.12. Функциональная схема адаптивного приемника узкополосных сигналов

7.12. О СИСТЕМАХ ПЕРЕДАЧИ ИНФОРМАЦИИ С ОБРАТНЫМ КАНАЛОМ

Пусть имеется линия (канал) передачи от точки *A* к точке *B*, а также линия от *B* к *A*, которая используется для сообщения в пункт отправления *A* информации о «качестве» сигналов, получаемых в пункте *B*. В этом случае канал от *B* к *A* называется *обратным*.

Иногда статистические свойства линий (например, замирания) в обоих направлениях одинаковы (происходят одновременно). О таких линиях говорят, что они обладают свойством *взаимности*.

Обратный канал может быть использован по-разному. Одним из простейших вариантов использования этого канала является посылка по нему периодических испытательных сигналов, с помощью которых в точке *A* определяются интервалы наилучшего состояния прямой линии (например, интервалы наименьшего ослабления или минимальных помех). Передача из точки *A* в точку *B* осуществляется в течение этих интервалов. Такой метод передачи называется *прерывистой связью*.

Другим вариантом использования обратного канала является работа с переспросом и повторением. При этом искаженные (бесмысленные) кодовые комбинации, полученные в точке приема, «переспрашиваются» у передатчика, который повторяет правильные комбинации.

В системах с *информационной обратной связью* проверка правильности кодовых комбинаций, принятых приемником, осуществляется на передающей стороне. Для этого принятая комбинация (или условный сигнал) по обратному каналу возвращается на передающую сторону, где сверяется с ранее отправленной, хранящейся в оперативной памяти. При положительном результате сверки по прямому каналу передается следующая комбинация, при отрицательном «сигнал отказа» и предыдущая комбинация повторяется.

В системах с *решающей обратной связью* проверка правильности кодовых комбинаций, принятых приемником, осуществляется на приемной стороне по установленному правилу. По обратному каналу отправляется или подтверждение или требование на повторение комбинации. Системы с решающей обратной связью могут быть построены по принципу «блокировки» или «адресного переспроса». В первом варианте при приеме ошибочной комбинации приемник прекращает прием, посылая переспрос. Передатчик, получив сигнал переспроса, повторяет последовательности с момента появления ошибочной комбинации до сигнала переспроса. Во втором варианте (имеющем свои разновидности) по обратному каналу отправляются адреса (идентификаторы) искаженных кодовых комбинаций.

До сих пор предполагалось, что помехи в линии являются флуктуационными и обладают нормальным законом распределения мгновенных значений. Этот случай относится к одному из предельных и часто встречающихся. Вторым предельным случаем являются импульсные помехи, т. е. последовательность случайных по форме, величине и времени возникновения импульсов, длительность которых в среднем мала по сравнению с интервалами между ними. Импульсные помехи часто являются искусственными по происхождению (например, искрения и разряды в инженерных устройствах). Это позволяет бороться с ними, применяя их экранировку в точках возникновения. Для предотвращения распространения помех по проводам, питающим искрящее устройство, включают фильтры нижних частот, ослабляющие энергию высоко-частотной части спектра помехи.

Единой теории борьбы с импульсными помехами пока не создано вследствие их большого разнообразия, а также трудностей нахождения многомерного закона распределения помехи, необходимого для синтеза оптимального приемника (см. п. 7.2). Для различных моделей импульсных помех можно найти одномерные законы распределения позволяющие определять отношение сигнал/помеха для отдельных методов борьбы с импульсными помехами и таким образом сопоставлять их эффективность.

Для ослабления воздействия импульсных помех на приемное устройство используются несколько различных методов.

Структурная схема, используемая при методе ШОУ, приведена на рис. 7.13,а. Она состоит из широкополосного фильтра (Ш), ограничителя (О) и узкополосного фильтра (У). Полоса $\Delta f_{ш}$ выбирается так, чтобы выполнялось условие

$$\Delta f_{ш} \gg 1/\tau_{и},$$

где $\tau_{и}$ — предполагаемая средняя длительность импульсов помех.

Этим обеспечивается незначительное «размытие» импульсов помехи, которое имело бы место при непосредственном воздей-

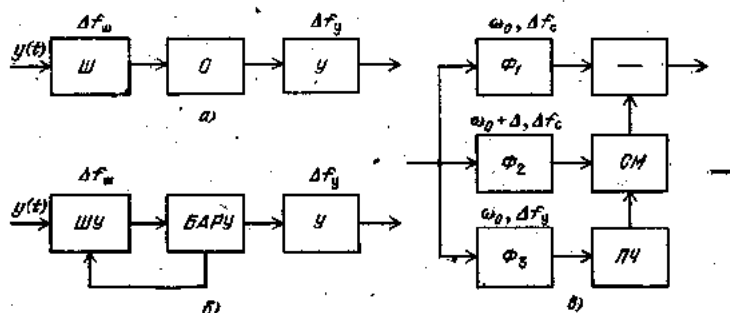


Рис. 7.13. Структурная схема приема по методу ШОУ (а), прерывания (б) и компенсации (в)

вии импульсов на узкополосный фильтр, согласованный по полосе с сигналами. Ограничитель O «обрезает» выбросы, обусловленные импульсными помехами, способствуя этим увеличению отношения сигнал/помеха. Полоса пропускания узкополосного фильтра $У$ согласована с полосой сигналов. Этим обеспечивается ослабление влияния флуктуационной составляющей помех.

На рис. 7.13,б приведена структурная схема реализации метода прерывания. Она состоит из широкополосного усилителя $Ш$, схемы быстродействующей автоматической регулировки усиления (БАРУ) и узкополосного фильтра $У$. Принцип ослабления влияния импульсов помех основан на закрывании приемника с помощью БАРУ на время воздействия импульсов.

При компенсационном методе каким-либо способом создаются синхронные реализации импульсных помех, которые затем вычитаются из суммы сигналов и помехи $y(t)$. Один из возможных вариантов структурной схемы компенсации приведен на рис. 7.13,в. Фильтр Φ_1 , предназначенный для ослабления воздействия флуктуационной составляющей помех, настроен на среднюю частоту сигналов ω_0 и согласован с ними по полосе. Средняя частота ω_2 фильтра Φ_2 выбрана так, чтобы в него практически не попадали сигналы, т. е. $\omega_2 = \omega_0 + \Delta$, где $\Delta > \Delta f_c / 2$. Полоса этого фильтра равна полосе фильтра Φ_1 . Короткий импульс помехи вызывает в фильтрах Φ_1 и Φ_2 приблизительно одинаковые медленно затухающие колебания, отличающиеся только средними частотами. С помощью узкополосного фильтра Φ_3 с полосой $\Delta f_y \ll \Delta f_c$, настроенного на среднюю частоту сигналов ω_0 , выделяется колебание с частотой ω_0 , из которого в схеме преобразования частоты (ПЧ) создается колебание, играющее роль местного генератора для смесителя (СМ). С помощью смесителя сформированные фильтром Φ_2 колебания переносятся на частоту сигналов ω_0 и вычитаются из суммы сигнала и помехи на выходе фильтра Φ_1 . Однако добиться полной компенсации практически невозможно.

КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАЧИ

1. Дайте определение оптимального приемника дискретных сигналов.
2. Сформулируйте и разъясните теорему Байеса о связи вероятностей причин и следствий.
3. Определите входящие в теорему Байеса вероятности.
4. Сформулируйте задачу распознавания сигналов в терминах теоремы Байеса.
5. Укажите решающее правило при распознавании сигналов, не искаженных каналом.
6. Выразите правдоподобия сигналов через плотность вероятности помех.
7. Дайте определения байесовского оптимального приемника и приемника максимального правдоподобия.

8. Приведите схему оптимального корреляционного приемника точно известных сигналов.
9. Поясните физику работы корреляционного приемника.
10. Дайте определение согласованного фильтра.
11. Докажите эквивалентность операций согласованной фильтрации и корреляции.
12. Приведите схему реализации согласованного фильтра для гармонических посылок.
13. Дайте определение оптимальной полосы пропускания приемника.
14. Определите области сигналов в многомерном пространстве векторов.
15. Укажите порядок определения вероятностей ошибок при распознавании бинарных точно известных сигналов.
16. Приведите структурную схему оптимального приемника в канале со случайными амплитудой и фазой.
17. Дайте определение когерентного и некогерентного приема.
18. Поясните принцип и назначение относительного кодирования.
19. Дайте определение энергетического приемника и укажите области его применения.
20. Поясните принцип приема с оценкой случайных параметров канала.
21. Укажите три основных метода борьбы с импульсными помехами.
22. При распознавании двух сигналов с априорными вероятностями $P(k_1)=0,3$ и $P(k_2)=0,7$ в конкретной реализации правдоподобия равны $p(y|k_1)=0,6$, $p(y|k_2)=0,4$. Какой сигнал был передан?
23. Вычислите вероятность ошибок при распознавании бинарных сигналов оптимальным приемником, если средняя мощность в точке приема $P_c=10^{-3}$ Вт, скорость передачи 100 Бод, спектральная плотность помехи $N_0=10^{-5}$ Вт/Гц.
24. Проведите ориентировочный расчет линии передачи, если $M=64$, $\mathcal{A}=10$ Бод, $P_{\text{ош}}=10^{-5}$, $N_0=10^{-5}$.
25. Постройте относительный бинарный код для последовательности элементов $k_1 k_1 k_2 k_1 k_2 k_2$.

ГЛАВА 8. ПРИНЦИПЫ ПОМЕХОУСТОЙЧИВОГО КОДИРОВАНИЯ

8.1. ИДЕЯ ПОСТРОЕНИЯ ПОМЕХОУСТОЙЧИВЫХ КОДОВ

При безыбыточном двоичном кодировании некоторого множества M дискретных элементов (знаков) каждому элементу такого множества необходимо поставить в соответствие определенную

двоичную кодовую комбинацию с числом разрядов k , удовлетворяющим условию $2^k \geq M$.

Очевидно, что изменение даже одного элемента двоичной кодовой комбинации в процессе ее передачи (замены 1 на 0, или наоборот) приведет к неверному приему знака x_i , т. е. безызбыточный код обладает большой чувствительностью к помехам.

Идея построения помехоустойчивого кода состоит в увеличении числа разрядов (позиций) в кодовых комбинациях с k до n ($n > k$). Из увеличенного числа позиций n только k позиций заполняются 1 и 0 в соответствии с правилом первичного кодирования знаков x_i . Эти k позиции называют *информационными*. Остальные $r = n - k$ позиций заполняются символами 0 и 1 по определенным правилам помехоустойчивого кодирования. Эти r позиций называют *проверочными*.

При использовании $n = k + r$ позиций возможно создание 2^{k+r} различных кодовых комбинаций. Из них только 2^k комбинаций используются для кодирования знаков x_i и называются *разрешенными*, а остальные $2^n - 2^k$ неиспользуемых комбинаций — *запрещенными*. Увеличивая n (при заданном значении k), можно построить помехоустойчивые коды двух типов: коды, обнаруживающие ошибки, и коды, исправляющие ошибки. В первом случае в точке приема можно лишь обнаружить искаженные кодовые комбинации. Во втором — ошибки на позициях кодовой комбинации можно исправить и, таким образом, принять ее правильно, несмотря на возможные ошибки в передаче двоичных сигналов $s(t, 1)$ и $s(t, 0)$, соответствующих элементам 1 и 0 двоичного кода.

Поясним сказанное простейшим примером. Пусть $k=2$. При безызбыточном двоичном кодировании в этом случае можно образовать четыре кодовых комбинации ($M=2^k=4$): $k_1=00$; $k_2=01$; $k_3=10$; $k_4=11$. Любая одиночная ошибка приведет к тому, что одна из комбинаций будет принята за другую (например, при ошибке в первой позиции k_1 получим 10, т. е. комбинацию k_3 , а во второй — 01, т. е. комбинацию k_2). Введем один дополнительный разряд, т. е. $r=1$ и $n=k+r=3$. Общее количество комбинаций теперь станет 8:

000; 001; 010; 011; 100; 101; 110; 111.

Используем как разрешенные для передачи только те четыре комбинации, число единиц в которых четно (они подчеркнуты), а остальные четыре (с нечетным числом единиц) будем считать запрещенными. Теперь одиночная ошибка на любой позиции разрешенной комбинации может быть обнаружена. Действительно, одиночная ошибка в разрешенной комбинации делает общее число единиц в этой комбинации нечетным, т. е. переводит разрешенную комбинацию в запрещенную, что и является признаком ошибки. Проверая число единиц в каждой принятой комбинации, можно установить, какие из этих комбинаций содержат ошибку.

Описанный код, обнаруживающий одиночные ошибки, называется *кодом с проверкой на четность* и широко используется на практике. Для его построения необходимо, очевидно, ко всем 2^k кодовым комбинациям безызбыточного кода добавить по одной позиции и заполнить ее символом 0 или 1 по правилу четности числа единиц.

Нетрудно видеть, что добавление дополнительного разряда увеличивает общее число возможных комбинаций вдвое по сравнению с числом комбинаций исходного первичного кода, а условие четности делит все комбинации на разрешенные и неразрешенные.

Увеличивая число дополнительных разрядов и формируя по определенным правилам проверочные символы 0 или 1, соответствующие этим разрядам, можно усилить корректирующие свойства кода так, чтобы код позволял не только обнаруживать, но и исправлять как одиночные, так и ошибки более высокой кратности. Для обнаружения ошибки достаточно установить факт, что в данной кодовой комбинации произошла ошибка (одиночная, двойная и т. п.). Для исправления ошибки необходимо не только обнаружить ошибку, но и указать ту позицию в кодовой комбинации, где произошла эта ошибка. Естественно, что такая задача значительно сложнее, чем обнаружение ошибки, и требует применения более сложных кодов.

Помехоустойчивое кодирование представляет собой обширную область теоретических и прикладных исследований. К числу основных задач, возникающих при этом, относятся отыскание кодов, эффективно исправляющих ошибки требуемого вида, нахождение методов кодирования и декодирования, способов их реализации.

8.2. ОСНОВНЫЕ ВИДЫ КОРРЕКТИРУЮЩИХ КОДОВ

Коды, исправляющие ошибки, часто называют также корректирующими или помехоустойчивыми. Их исследованию посвящено огромное количество работ, начиная от сугубо математических и кончая отдельными частными работами прикладного характера. Поэтому в этой области, более чем в любой другой, существует множество различных терминов, часто выражающих одно и то же понятие. Известно большое число кодов, названных по фамилиям авторов, впервые предложивших эти коды. Все это затрудняет пользование литературой по вопросам кодирования, особенно для читателей, начинающих впервые знакомиться с этой обширной областью.

В настоящее время чаще всего применяются двоичные равномерные корректирующие коды, так как они обладают хорошими исправляющими свойствами и их реализация сравнительно проста. Одна из особенностей таких кодов состоит в том, что все комбинации имеют одинаковое общее число разрядов n (одинаковую длительность). Двоичные равномерные коды делятся на блочные и сверточные. Остановимся кратко на рассмотрении некоторых особенностей этих классов кодов.

Блочные коды. При использовании блочных кодов цифровая информация передается в виде отдельных кодовых комбинаций (блоков) равной длины. Кодирование и декодирование каждого блока осуществляется независимо друг от друга.

Почти все блочные коды относятся к *разделимым кодам*, т. е. таким, комбинации которых состоят из двух различающихся час-

тей: информационной и проверочной. Информационные и проверочные разряды во всех кодовых комбинациях делимого кода всегда занимают одни и те же позиции. Делимые коды обычно условно обозначают в виде (n, k) , где n указывает значность кода (число позиций в блоке), k — число информационных позиций. Таким образом, число проверочных символов в делимых кодах равно $r = n - k$.

Среди делимых кодов различают систематические и не систематические коды.

Систематические, или линейные, коды представляют собой наиболее обширную группу делимых кодов. В таких кодах проверочные символы образуются различными линейными комбинациями информационных символов. Теоретической основой получения таких комбинаций является аппарат линейной алгебры, который позволяет формировать проверочные символы по определенным правилам (по определенной системе — отсюда и происхождение термина «систематические» коды). Использование аппарата линейной алгебры, в которой важное значение имеет понятие «группа», привело к тому, что в ряде работ рассматриваемые коды называют также «алгебраическими» или «групповыми» [1, 2].

Систематические коды хорошо изучены. Наиболее известны среди них *циклические коды*. Основное свойство циклических кодов состоит в том, что при циклическом сдвиге любой разрешенной кодовой комбинации получается также разрешенная комбинация. Циклические коды обладают хорошими корректирующими свойствами, а реализация кодирующих и декодирующих устройств для таких кодов оказывается проще, чем для других систематических кодов.

К циклическим относятся коды Хэмминга, которые исторически появились раньше многих других кодов и сыграли большую роль в развитии теории корректирующих кодов, а также коды Боуза — Чоудхури.

Сверточные коды. Формирование проверочных символов в таких кодах осуществляется по рекуррентным правилам, поэтому сверточные коды часто называют также *рекуррентными*, или *цепными*. Особенностью сверточных кодов является то, что они формируются непрерывно и в них проверочные символы перемежаются с информационными по всей длине кодовой последовательности, подчиняясь одному и тому же рекуррентному соотношению.

Принцип формирования простейшего сверточного кода можно пояснить с помощью рис. 8.1, где $\{x_{i,1}\}$ — последовательность информационных символов, а $\{x_{i,2}\}$ — проверочных. Устройство задержки τ создает запаздывание (сдвиг) входной последовательности $\{x_{i,1}\}$ на один разряд, а проверочные символы формируются по рекуррентному правилу

$$x_{i,2} = x_{i,1} \oplus x_{i-1,1}, \quad (8.1)$$

где \oplus означает суммирование по модулю 2 (т. е. $0 \oplus 1 = 1$; $1 \oplus 1 = 0$). Ключ K переключается поочередно из положения 1 в положение 2, и обратно. В ре-

результате на выходе формируется последовательность вида

$$\dots x_{i,1}, x_{i,2}, x_{i+1,1}, x_{i+1,2}, x_{i+2,1}, x_{i+2,2}, \dots$$

которая состоит из элементарных блоков (подчеркнутых снизу). Элементарным блоком сверточного кода называется последовательность символов с одинаковым первым номером, например $x_{1,1}$, $x_{1,2}$ или $x_{2,1}$, $x_{2,2}$ и т. д.

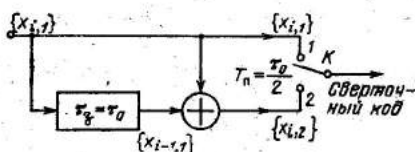


Рис. 8.1. Принцип формирования простейшего сверточного кода

При приеме правило (8.1) позволяет установить, какой из символов элементарного блока искажен, информационный или проверочный. Действительно, пусть информационный символ $x_{i,1}$. Так как он участвует в проверке дважды:

$$x_{i,2} = x_{i,1} \oplus x_{i-1,1}, x_{i+1,2} = x_{i+1,1} \oplus x_{i,1},$$

то его искажение приведет к невыполнению правила (8.1) в i -м и $(i+1)$ -м тактах. Если же искажен проверочный символ $x_{i,2}$, то уравнение (8.1) не будет выполняться только в i -м такте.

Рассмотренный простейший сверточный код позволяет исправлять одиночные ошибки. При искажении $x_{i,1}$ и $x_{i,2}$, т. е. всего элементарного блока, для обнаружения ошибки следует формировать более сложный сверточный код.

Теории и приложениям сверточных кодов в последние годы уделяется большое внимание, а область их применений непрерывно расширяется. Это объясняется тем, что для многих применений схемы кодеров и декодеров таких кодов получаются значительно проще, чем для блочных кодов с такой же корректирующей способностью [1, 2, 3].

Вопросы помехоустойчивого кодирования наиболее полно разработаны применительно к систематическим кодам. Такие коды широко применяются в вычислительной технике и различных автоматизированных цифровых устройствах. Некоторые из предложенных кодов оказались удобными и для цифровых систем передачи информации.

Ниже кратко изложены особенности систематических кодов, методы их кодирования и декодирования, а также рассмотрены возможности применения таких кодов для повышения точности передачи цифровой информации.

8.3. ОСНОВНЫЕ ХАРАКТЕРИСТИКИ КОРРЕКТИРУЮЩИХ КОДОВ

К основным характеристикам корректирующих кодов относят: избыточность кода, кодовое расстояние, число обнаруживаемых или исправляемых ошибок, корректирующие возможности кодов.

Избыточность корректирующего кода. Избыточностью корректирующего кода называют величину

$$\omega = r/n = (n - k)/n = 1 - k/n. \quad (8.2)$$

Из (8.2) следует

$$k/n = 1 - w. \quad (8.3)$$

Эта величина показывает, какую часть от общего числа символов кодовой комбинации составляют информационные символы. В теории кодирования величину k/n принято называть *скоростью передачи* [1]. Необходимо иметь в виду, что величина k/n характеризует относительную скорость передачи информации. Если производительность источника информации равна H бит/с, то скорость передачи после кодирования этой информации окажется равной

$$R = H k/n, \quad (8.4)$$

поскольку в закодированной последовательности из каждых n символов только k являются информационными.

Для обнаружения или исправления значительного числа ошибок, необходимо иметь код с большим числом проверочных символов. При этом существенно возрастает длительность кодовых блоков, что приводит к задержке информации при передаче и приеме. Чем сложнее кодирование, тем длительнее временная задержка информации.

Кодовое расстояние. Расстоянием d между двумя кодовыми комбинациями называется число позиций, в которых эти комбинации имеют разные символы. Например, расстояние между комбинациями 0001101 и 1001010 равно четырем ($d=4$). Расстояние между различными комбинациями некоторого конкретного кода может быть различным. Так, в частности, в безыбыточном натуральном (первичном) коде это расстояние для различных комбинаций может различаться от единицы до величины, равной значности кода. Особенно важной характеристикой корректирующих свойств кода является минимальное расстояние d_{\min} между кодовыми комбинациями. Это расстояние называется кодовым, или хэмминговым.

В безыбыточном коде все комбинации являются разрешенными и, следовательно, его кодовое расстояние равно единице. Поэтому при искажении хотя бы одного символа вместо переданной комбинации будет принята другая разрешенная комбинация. Чтобы код обладал корректирующими свойствами, необходимо ввести в него некоторую избыточность, которая обеспечивала бы минимальное расстояние между любыми двумя разрешенными комбинациями $d_{\min} \geq 2$.

Число обнаруживаемых или исправляемых ошибок. Кодовое расстояние является основной характеристикой корректирующей способности данного кода. Если код используется только для обнаружения ошибок кратности a , то необходимо и достаточно, чтобы минимальное расстояние удовлетворяло условию¹

$$d_{\min} \geq a + 1. \quad (8.5)$$

¹ Кратностью ошибки a называют число позиций кодовой комбинации, в которых под действием помехи одни символы оказались замененными другими. Например, 0 на 1 или 1 на 0.

В этом случае никакая комбинация из a ошибок не может перевести одну разрешающую кодовую комбинацию в другую разрешенную. Таким образом, условие обнаружения всех ошибок кратностью a можно записать в виде

$$a_{\text{обн}} \leq d_{\text{min}} - 1. \quad (8.6)$$

Для исправления всех ошибок кратностью a и менее необходимо, чтобы минимальное расстояние удовлетворяло условию

$$d_{\text{min}} \geq 2a + 1. \quad (8.7)$$

В этом случае любая кодовая комбинация с кратностью ошибок a отличается от каждой разрешенной комбинации не менее чем в $a+1$ позициях. Если условие (8.7) не выполнено, возможен случай, когда ошибка кратности a исказит переданную комбинацию так, что последняя станет ближе к одной из разрешенных комбинаций, чем к переданной, или даже перейдет в другую разрешенную комбинацию. Итак, условие исправления всех ошибок кратностью не более a можно записать в виде

$$a_{\text{испр}} \leq (d_{\text{min}} - 1)/2. \quad (8.8)$$

Корректирующие коды можно одновременно использовать и для обнаружения и для исправления ошибок. В [1] указывается, что минимальное расстояние, при котором можно исправить все ошибки кратностью a или меньше и одновременно обнаружить все ошибки кратностью b или меньше, определяется условием

$$d_{\text{min}} \geq a + b + 1, \quad (8.9)$$

где, в свою очередь, всегда должно выполняться условие $b > a$.

Корректирующие возможности кодов. Вопрос о минимально необходимой избыточности, при которой код обладает нужными корректирующими свойствами, является одним из важнейших в теории кодирования. Этот вопрос до сих пор не получил окончательного решения. В настоящее время известен ряд верхних и нижних оценок (границ), которые устанавливают связь между максимально возможным минимальным расстоянием корректирующего кода и его избыточностью [1, 2]. Чтобы дать представление о количественной стороне этих оценок приведем краткие сведения о некоторых оценках.

Нижняя граница Варшавова — Гильберта. Для больших значений n указанная граница определяется асимптотическим соотношением [1]

$$\frac{r}{n} = 1 - \frac{k}{n} \geq H\left(\frac{d_{\text{min}} - 2}{n - 1}\right), \quad (8.10)$$

где $H(x) = -x \log_2 x - (1-x) \log_2 (1-x)$.

Из выражения (8.10) следует

$$r \geq n H\left(\frac{d_{\text{min}} - 2}{n - 1}\right), \quad (8.11)$$

$$\frac{k}{n} \leq 1 - H\left(\frac{d_{\text{min}} - 2}{n - 1}\right). \quad (8.12)$$

Условия (8.11) и (8.12) позволяют оценить необходимое количество проверочных символов r и относительную скорость передачи k/n при заданных (или выбранных) значениях n и d_{\min} .

Верхняя граница Хэмминга. Основываясь на результатах, изложенных в [1], эту границу для двоичного корректирующего кода можно записать в виде

$$\frac{r}{n} \geq H\left(\frac{d_{\min}-1}{2n}\right). \quad (8.13)$$

Отсюда следуют выражения для оценки числа проверочных символов и относительной скорости передачи:

$$r \geq n H\left(\frac{d_{\min}-1}{2n}\right), \quad (8.14)$$

$$\frac{k}{n} \leq 1 - H\left(\frac{d_{\min}-1}{2n}\right). \quad (8.15)$$

Соотношение (8.15) известно в литературе как верхняя граница Хэмминга. Кроме рассмотренной оценки, известна также оценка, называется *верхней границей Плоткина* [1, 4]. Для двоичного корректирующего кода эта граница определяется выражением

$$r \geq 2(d_{\min}-1) - \log_2 d_{\min}, \quad (8.16)$$

которое справедливо при $n \geq 2d_{\min}-1$. Для значений $d_{\min}/n \leq 0,3$ разница между границей Хэмминга и границей Плоткина сравнительно невелика.

Заметим, что для некоторых частных случаев Хэмминг указал более простые соотношения, позволяющие определить необходимое число проверочных символов:

$$r \geq \log_2(n+1), \text{ если } d_{\min}=3, \quad (8.17)$$

$$r \geq \log_2 2n, \text{ если } d_{\min}=4. \quad (8.18)$$

Систематические коды с $d_{\min}=3; 4$ в литературе обычно называют *кодами Хэмминга*.

8.4. ВОЗМОЖНОСТИ ПОВЫШЕНИЯ ПОМЕХОУСТОЙЧИВОСТИ ПРИЕМА ЦИФРОВОЙ ИНФОРМАЦИИ С ПОМОЩЬЮ КОРРЕКТИРУЮЩИХ КОДОВ

В теории помехоустойчивого кодирования принято различать два основных вида ошибок: статистически независимые, т. е. некоррелированные ошибки; статистически зависимые, т. е. коррелированные ошибки (ошибки типа пачек или пакетов).

Известно большое число математических моделей, позволяющих описать закономерности возникновения ошибок в различных каналах передачи цифровой информации [5]. Рассмотрение вопросов применения корректирующих кодов для повышения достоверности передачи цифровой информации с учетом различных моделей ошибок представляет сложную задачу, особенно, если ошибки возникают «пачками». При этом для совокупности определенных условий требуется определенный код; например, если код

предназначен для исправления только одиночных или двойных ошибок, то применять его в каналах с замираниями бесполезно. Исследование подобных вопросов представляет собой большую самостоятельную проблему.

Основная цель дальнейшего рассмотрения состоит в том, чтобы проиллюстрировать на простом примере возможность повышения помехоустойчивости с помощью корректирующего кода и выяснить, при каких условиях применение такого кода целесообразно. При рассмотрении будем полагать, что канал является бинарным симметричным, а ошибки в нем независимы. Такие условия характерны для гауссовских каналов, в которых применяются двоичные посылки сигналов с одинаковыми энергиями¹.

При статистически независимых ошибках симметричный канал можно полностью характеризовать одним параметром — вероятностью ошибочного приема символа. Все соотношения, характеризующие работу канала, при этом оказываются наиболее простыми.

Пусть по гауссовскому каналу с двоичными противофазными сигналами передается цифровая информация, закодированная избыточным кодом, содержащим k информационных разрядов с длительностью символов, равной τ_0 . Вероятность ошибочного приема символа в этом случае определяется выражением (см. гл. 7)

$$p_1 = 0,5 - \Phi(\sqrt{2}h_0), \quad (8.19)$$

где $h_0^2 = E_0/N_0 = P_c\tau_0/N_0$.

Применение корректирующего кода означает увеличение избыточности в используемых комбинациях. Допустим, что каждая кодовая комбинация такого кода состоит из n символов. Здесь возможны два случая: 1) увеличение n производится при сохранении прежней длительности кодовой комбинации ($\tau_k = k\tau_0 = \text{const}$); 2) увеличение n производится при сохранении прежней длительности одного символа ($\tau_0 = \text{const}$).

Нетрудно видеть, что в первом случае

$$h_{0к}^2 = E_{0к}/N_0 = (k/n)h_0^2, \quad (8.20)$$

$$p_{1к} = 0,5 - \Phi(\sqrt{2k/n}h_0), \quad (8.21)$$

где $p_{1к}$ — вероятность ошибочного приема символа в кодированной последовательности с учетом увеличения избыточности ($n > k$).

Во втором случае

$$h_{0к} = h_0, p_{1к} = p_1. \quad (8.22)$$

Чтобы оценить возможность повышения помехоустойчивости с помощью корректирующего кода, найдем вероятность неправильного приема отдельного символа с учетом корректирующих свойств кода. Поскольку такой код позволяет обнаружить или исправить часть ошибок, то, очевидно, не все ошибки при приеме

¹ Бинарным симметричным называется такой канал, в котором ошибки, связанные с переходом символов 1 в 0 и 0 в 1, одинаковы.

отдельных символов имеют значение. Пусть кратность обнаруживаемой или исправляемой ошибки равна a . Тогда вероятность того, что в кодовой комбинации, состоящей из n символов, каждый из которых может быть принят с вероятностью ошибки $p_{1к}$, произойдет ошибка на a позициях кода, равна

$$p_{a \text{ ош}} = p_{1к}^a. \quad (8.23)$$

Вероятность того, что на остальных $(n-a)$ позициях ошибки не будет, равна

$$p_{(n-a) \text{ прав}} = (1 - p_{1к})^{n-a}. \quad (8.24)$$

Вероятность совместного события, состоящего в том, что в данной кодовой комбинации на a позициях произошла ошибка, а на $n-a$ позициях ошибки нет, равна

$$p = p_{1к}^a (1 - p_{1к})^{n-a}. \quad (8.25)$$

В кодовой комбинации, состоящей из n символов, ошибки кратности a могут произойти при различных сочетаниях позиций. Число таких сочетаний определяется выражением

$$C_n^a = n! / a!(n-a)! \quad (8.26)$$

Следовательно, вероятность того, что в кодовой комбинации из n символов появится ошибка кратности a (при любом сочетании позиций), определяется выражением

$$p_a = C_n^a p_{1к}^a (1 - p_{1к})^{n-a}. \quad (8.27)$$

Это выражение можно рассматривать как вероятность обнаруживаемой или исправляемой ошибки, так как по условию было принято, что все ошибки кратности a и менее могут быть обнаружены или исправлены.

Если при передаче произойдут ошибки кратности $a+1$ и выше, то такие ошибки уже нельзя обнаружить или исправить¹. Нетрудно видеть, что полная вероятность необнаруживаемых или неисправляемых ошибок определяется выражением

$$p_{\text{ош}} = \sum_{j=a+1}^n C_n^j p_{1к}^j (1 - p_{1к})^{n-j}, \quad (8.28)$$

где $j = a+1, a+2, \dots, n$.

При малой вероятности ошибки выражение (8.28) можно приближенно заменить первым членом суммы:

$$p_{\text{ош}} \approx C_n^{a+1} p_{1к}^{a+1} (1 - p_{1к})^{n-a-1}. \quad (8.29)$$

Это выражение можно рассматривать как вероятность ошибочного приема кодовой комбинации с учетом корректирующих свойств кода.

¹ Вообще говоря, корректирующий код способен исправлять некоторые ошибки и более высокой кратности, чем a . Однако «вес» таких ошибок среди других возможных ошибок невелик, и это обстоятельство можно не принимать во внимание.

Для оценки эффективности корректирующего кода удобно пользоваться эквивалентной вероятностью ошибки приема одного символа, т. е.

$$p_{1 \text{ экв}} = \frac{p_{\text{ош}}}{n} = \frac{1}{n} C_n^{a+1} p_{1к}^{a+1} (1 - p_{1к})^{n-a-1}. \quad (8.30)$$

Эта величина определяет вероятность необнаруживаемой ошибки приема символа при использовании корректирующего кода.

Сравнивая эту величину с вероятностью ошибки приема символа для безыбыточного кода, определяемой выражением (8.19), можно установить, при каких условиях применение корректирующего кода позволяет повысить помехоустойчивость приема. Естественно, что такое сопоставление должно проводиться при одинаковых условиях (например, при одинаковой полосе занимаемых частот или одинаковой длительности кодовой комбинации и т. п.).

В качестве примера рассмотрим эффективность кода Хэмминга с кодовым расстоянием $d_{\text{мин}}=3$. Такой код позволяет обнаружить двойную ошибку или исправить одиночную.

Из соотношения (8.17) следует, что

$$n = 2^r - 1, \quad (8.31)$$

где r — число проверочных символов.

Возьмем $r=3$. Так как $k=n-r$, то с учетом (8.31) нетрудно установить, что код имеет в данном случае вид $(n, k) = (7, 4)$.

Путь передача безыбыточным 4-разрядным кодом ($k=4$) в гауссовском канале с оптимальным приемником характеризуется вероятностью ошибки приема символа, равной $p_1=10^{-2}$. В соответствии с (8.19) находим, что вероятность $p_1=10^{-2}$ достигается при величине $\sqrt{2 h_0} \approx 2,33$.

Будем считать, что избыточность вводится за счет сокращения длительности символа первичного кода (расширения используемой полосы частот). Тогда в соответствии с (8.20) и (8.21)

$$h_{0к}^2 = 4/7 h_0^2 \text{ и } p_{1к} = 0,5 - \Phi(\sqrt{4/7} 2,33) \approx 3,75 \cdot 10^{-2}.$$

Пусть код Хэмминга используется для обнаружения ошибок. Тогда $a_{\text{ош}} \leq 2$ и эквивалентная вероятность необнаруживаемой ошибки определяется выражением (8.30), которое в данном случае принимает вид $p_{1 \text{ экв}} = (1/7) C_7^3 p_{1к}^3 (1 - p_{1к})^4$. Подставив сюда найденное значение вероятности $p_{1к}$, получим $p_{1 \text{ экв}} \approx 2,3 \cdot 10^{-4} < p_1$. Этот результат означает, что в данных условиях код Хэмминга, используемый для обнаружения двойных ошибок, оказывается эффективным. Если этот же код применить для исправления одиночных ошибок, то расчеты показывают, что при тех же условиях его эффективность ниже (в этом случае $p_{1 \text{ экв}} = 3,5 \cdot 10^{-3}$).

Такая особенность характерна и для других корректирующих кодов. При заданной избыточности эффективность корректирующего кода при обнаружении ошибок всегда значительно выше эффективности этого же кода, используемого для исправления ошибок.

При протяженных кодах ($n \gg 1$) расчеты по оценке эффективности кодов становятся весьма громоздкими. Для ориентировоч-

ных оценок можно использовать соотношение, полученное в [6]:

$$k/n > 1/d_{\min}. \quad (8.32)$$

Если это условие не выполняется, то применение выбранного корректирующего кода приведет к уменьшению помехоустойчивости.

8.5. МЕТОДЫ КОДИРОВАНИЯ И ДЕКОДИРОВАНИЯ СИСТЕМАТИЧЕСКИХ КОДОВ

Кодирование систематического кода сводится к образованию необходимого числа проверочных (избыточных) символов, которые обеспечивают обнаружение или исправление ошибок заданного вида. Любую кодовую комбинацию безыбыточного двоичного кода, имеющего k информационных разрядов, можно представить в виде

$$x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{ik}, \quad (8.33)$$

где индекс i принимает одно из двух значений ($i=1, 2$), указывая тем самым, какой вид имеет данный символ (например $x_{1j}=1, x_{2j}=0$).

Введем в код избыточность, добавив к каждой кодовой комбинации (8.33) по r проверочных символов. Задача состоит в том, чтобы сформировать проверочные символы наиболее простым способом и обеспечить при этом обнаружение или исправление ожидаемого вида ошибок. Это достигается тем, что совокупность проверочных символов

$$z_{i1}, z_{i2}, \dots, z_{im}, \dots, z_{ir} \quad (8.34)$$

формируется так, чтобы каждый из них представлял собой линейную комбинацию информационных символов. Для этого нужно образовать систему r линейных уравнений, каждое из которых имеет вид

$$z_{im} = \sum_{j=1}^k c_{ij} x_{ij}. \quad (8.35)$$

Здесь c_{ij} — весовые коэффициенты, принимающие значение 1 или 0 в зависимости от того, участвует или нет данный информационный символ x_{ij} в образовании проверочного символа z_{im} с порядковым номером m . Суммирование в (8.35) ведется по модулю 2.

Правило формирования кода обычно выбирают так, чтобы при декодировании имелась возможность выполнять ряд проверок на четность для некоторых определенным образом выбранных подмножеств информационных и проверочных символов каждой кодовой комбинации. Анализируя результаты проверок, можно обнаружить или исправить ошибку ожидаемого вида.

Правила построения систематических кодов довольно сложны, и мы на них останавливаться не будем. Подробное рассмотрение задач, связанных с построением корректирующих кодов, можно найти в [1, 2].

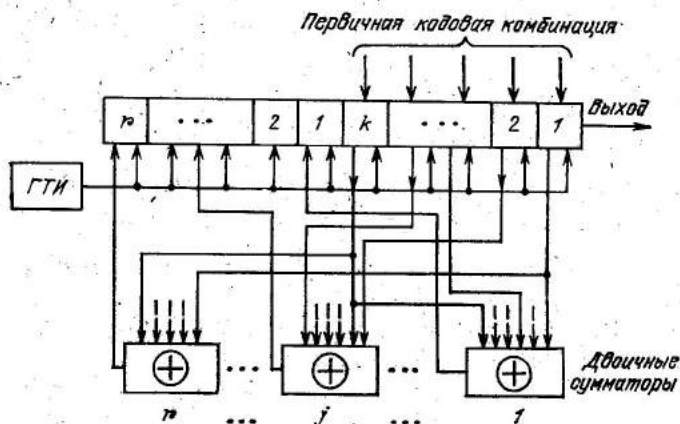


Рис. 8.2. Упрощенная структурная схема кодирующего устройства для (n, k) кода

Как отмечалось выше, достоинством систематических кодов является сравнительная простота их технической реализации. Кодирующее устройство для получения (n, k) кода состоит из n -разрядного регистра сдвига и g сумматоров по модулю 2 (рис. 8.2). Регистр содержит две части: информационную (k ячеек) и проверочную (g ячеек). Каждый сумматор служит для формирования проверочного символа, находящегося на определенной позиции. Подключение информационных ячеек регистра к соответствующим сумматорам выполняется по правилам построения данного корректирующего кода.

Устройство работает следующим образом. Цифровая информация в виде кодовой комбинации первичного кода записывается параллельно (одновременно) в k информационных ячеек регистра. Одновременно с этим во всех g двоичных сумматорах происходит формирование проверочных символов, которые записываются в g проверочных ячейках регистра. Следующим этапом является выведение полученной кодовой комбинации из регистра. Это достигается подачей тактовых импульсов (импульсов сдвига) от генератора тактовых импульсов (ГТИ). После n тактовых импульсов кодовая комбинация окажется выведенной из регистра. При этом генератор ГТИ должен быть отключен от регистра для подготовки последнего к записи и формированию следующей кодовой комбинации. Кодовая комбинация с выхода регистра поступает к модулятору передающей части системы связи.

Формирование корректирующих кодов с помощью регистров сдвига получило очень широкое распространение. Этот метод прост в реализации, позволяет создавать высоконадежные схемы и обеспечивает возможность изменения параметров кода в процессе работы. Действительно, если изменить частоту следования тактовых импульсов, то изменится длительность каждого символа

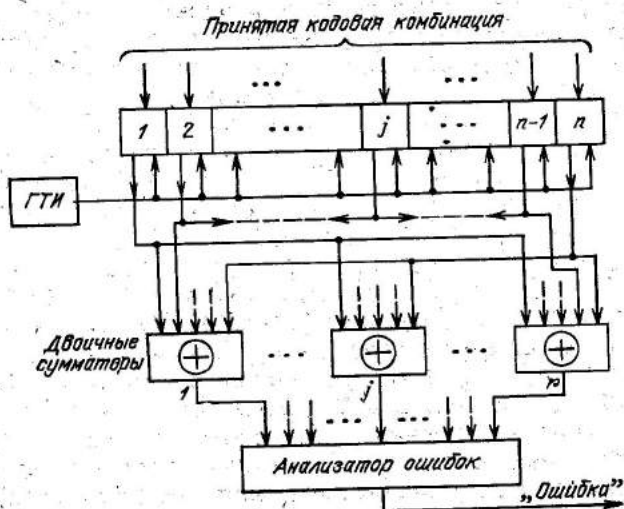


Рис. 8.3. Упрощенная структурная схема декодирующего устройства для (n, k) кода

ла, а следовательно, и энергия кодовой комбинации, скорость передачи. Если при этом согласовано (по команде) изменить соответствующие параметры в приемном устройстве, можно получить адаптивную (в некотором смысле) систему связи.

Рассмотрим принцип работы декодирующего устройства, упрощенная структурная схема которого показана на рис. 8.3. Устройство состоит из n триггерных ячеек, r сумматоров по модулю 2 и анализатора ошибок. Последовательность символов принятой кодовой комбинации (часть этих символов может быть принята неверно из-за действия помех) записывается в n ячеек регистра и с помощью сумматоров проводится проверка на четность. Для r подмножеств информационных и проверочных символов. Как уже говорилось выше, эти подмножества формируются при кодировании по определенным для данного кода правилам. Если в принятой кодовой комбинации ошибок нет, то на выходе всех сумматоров формируются символы, соответствующие нулю. Если в принятой комбинации имеются ошибочные символы, то на выходе некоторых сумматоров появятся символы, соответствующие единице, поскольку условия проверки на четность в этих сумматорах окажутся не выполненными. В этом случае анализатор ошибок выдает сигнал «ошибка».

При увеличении избыточности кода и усложнении устройства проверки и анализа ошибок декодирующее устройство сможет не только обнаружить ошибки в приеме отдельных символов, но и исправить часть этих ошибок. Для этого схема декодирующего устройства должна быть существенно усложнена по сравнению с

указанной на рис. 8.3. Схема должна содержать ряд дополнительных устройств, таких, как устройство памяти для запоминания исправляемой кодовой комбинации, логические устройства для исправления символов на тех позициях кодовой комбинации, где они оказались принятыми ошибочно и т. п.

Многочисленные исследования показали, что для кодов, обладающих большими корректирующими возможностями (а следовательно, представляющих собой последовательности с большим числом символов n), число операций при оптимальном декодировании растет экспоненциально с увеличением числа избыточных символов¹. Это обстоятельство существенно усложняет техническую реализацию декодирующих устройств. Поэтому в последние годы большое внимание уделяется вопросам разработки таких методов декодирования, для которых зависимость числа необходимых операций от длины кодовой последовательности является не экспоненциальной, а степенной или даже линейной функцией [1, 2]. В настоящее время разработаны два подхода к решению задачи упрощения процедуры декодирования: вероятностный и алгебраический.

При *вероятностном* или *последовательном декодировании* вся совокупность возможных кодовых комбинаций (включая разрешенные и запрещенные) разбивается на две группы: высоковероятные и маловероятные. Если принятая комбинация ближе к высоковероятной группе, она декодируется сразу же без проверки. Если же принятая комбинация ближе к маловероятным комбинациям, при декодировании выполняется проверка и исправление ошибок. Иными словами, сильно искаженные комбинации исправляются, а слабо искаженные декодируются без исправления. Такое декодирование несколько уступает в помехоустойчивости оптимальному, но заметно проще в реализации [1, 2].

При *алгебраическом* декодировании исправляется только определенная часть ошибок, т. е. корректирующие возможности кода используются лишь частично. Это означает, что применяется неоптимальный алгоритм декодирования, допускающий более простую схемную реализацию. Особенно удобны с точки зрения простоты декодирования циклические и сверточные коды. Для таких кодов разработаны эффективные методы алгебраического декодирования, которые получили название *пороговых*, или *мажоритарных* [1, 2].

Стремление к повышению скорости передачи информации по линиям связи и возрастающие требования к точности ее доставки потребителю стимулируют развитие и применение корректирующих кодов. Особое внимание в последние годы уделяется разработке теории и приложений сверточных кодов, поскольку их кодирование и декодирование оказывается не очень сложным и успешно реализуется на основе современной элементной базы.

¹ Оптимальным называют такое декодирование, при котором корректирующие возможности кода используются полностью.

КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАЧИ

1. Оцените корректирующие способности следующих систематических кодов: (15, 5); (31, 16); (63, 7); (7, 4). Установите, какие ошибки и какой кратности можно обнаруживать, а какие — исправлять с помощью указанных кодов.
2. Считая, что $n \gg 1$, $d_{\min} = 5$, определите параметры кода для верхней и нижней границ оценок.
3. Покажите, что для кодов Хэмминга справедливы следующие соотношения:

$$\frac{k}{n} \leq 1 - \frac{1}{n} \log_2(n-1), n \leq 2^r - 1, d_{\min} = 3;$$

$$\frac{k}{n} \leq 1 - \frac{1}{n} \log_2 2n, n \leq 2^{r-1}, d_{\min} = 4.$$

4. Покажите, что код Хэмминга при $d_{\min} = 3$ может иметь ряд вариантов: (7, 4); (15, 11); (31, 26); (63, 57); (127, 120) и т. д. При каких условиях целесообразно применять те или иные варианты кода, имея в виду, что корректирующие возможности этих вариантов одинаковы?
5. Докажите, что при сохранении энергии кодовой комбинации неизменной и увеличении избыточности кода вероятность ошибки при приеме имеет минимум для некоторого значения $r = r_{\text{opt}}$.
6. В чем суть методики, позволяющей оценить эффективность применения корректирующего кода?
7. Объясните принцип работы кодирующего устройства.
8. Объясните принцип работы декодирующего устройства.
9. В чем суть вероятностного и алгебраического методов декодирования?

ГЛАВА 9. ПРИНЦИПЫ РЕАЛИЗАЦИИ СИСТЕМ ПЕРЕДАЧИ ДИСКРЕТНЫХ СООБЩЕНИЙ

9.1. ОСНОВНАЯ ТЕРМИНОЛОГИЯ

При построении оптимальных систем передачи используются количественные критерии оптимальности, которые, хотя и являются строгими, но не учитывают всех реальных условий. Практика вносит свои коррективы в теорию, обусловленные сложностью реализации некоторых оптимальных систем и экономическими соображениями.

В настоящей главе рассмотрены практические методы реализации отдельных частей системы передачи сообщений и существующие регламентации.

Согласно ГОСТ 17657—72 и рекомендациям международного консультативного комитета по телефонии и телеграфии (МККТТ)

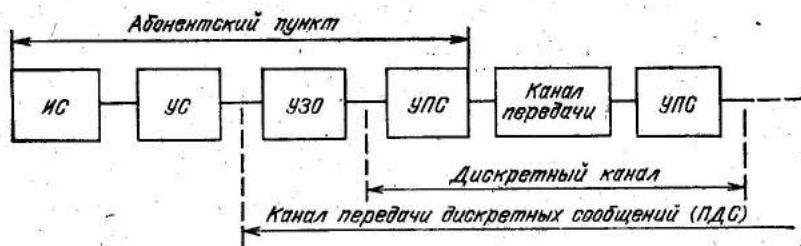


Рис. 9.1. Структурная схема системы передачи дискретных сообщений

под передачей дискретных сообщений понимается два вида связи: «традиционная» телеграфная связь — *телеграфия* и сравнительно новая область — *передача данных*.

Телеграфия предназначена для передачи любого знакового текста, составленного в любой форме из заданного алфавита. Под передачей данных понимают передачу дискретных сообщений в строго формализованном виде, пригодном для автоматической обработки ее (переприема, адресации), в особенности с помощью ЭВМ. Для обоих видов связи часто используются одни и те же физические линии, одинаковые (обычно бинарные) сигналы и сходная аппаратура. Важнейшие отличия состоят в большем быстродействии систем передачи данных, более высоких требованиях к вероятности ошибок при приеме и высокой степени автоматизации.

Структурная схема системы передачи дискретной информации приведена на рис. 9.1. Дискретные сообщения вырабатываются источником сообщений (ИС), в качестве которых могут выступать любые автоматические или ручные датчики: телеграфные аппараты, трансмиттеры, аналого-цифровые преобразователи, ЭВМ. Устройство сопряжения (УС) согласует источник сообщений с последующими частями системы по электрическим параметрам, скорости передачи, типу кодирования. Устройство защиты от ошибок (УЗО) предназначено для помехоустойчивого кодирования. В устройстве преобразования сигналов (УПС) кодовым комбинациям сопоставляются бинарные сигналы, пригодные для распространения по заданным каналам передачи (КП). На противоположном конце линии перечисленные устройства включены в обратном порядке, обеспечивая, таким образом, двустороннюю передачу сообщений.

Тракт от источника до абонента (получателя информации) называется *каналом передачи сообщений*. Любой канал можно характеризовать либо занимаемой полосой частот, либо занимаемым интервалом времени. В частности, аналоговый стандартный телефонный канал занимает полосу частот от 300 до 3400 Гц. В зависимости от вида передаваемых сообщений один и тот же канал может быть использован как аналоговый или дискретный. Сово-

купность дискретного канала и устройства защиты от ошибок называется *каналом передачи дискретных сообщений* (канал ПДС).

Совокупность ИС, УС, УЗО и УПС образует *абонентский пункт*.

Из отдельных каналов строятся *сети передачи*, позволяющие вести обмен сообщениями не между двумя, а многими абонентами. В этом случае в состав сети обязательно входят *устройства (узлы) коммутации* — автоматические или ручные.

Системы ПДС в зависимости от скорости передачи сообщений принято делить на низкоскоростные (до 200 бинарных посылок в секунду или Бод), среднескоростные (до 4800 Бод) и высокоскоростные (более 4800 Бод).

9.2. ИСТОЧНИКИ ДИСКРЕТНОЙ ИНФОРМАЦИИ. СИНХРОННЫЙ И СТАРТСТОПНЫЙ РЕЖИМЫ РАБОТЫ

Первичным источником дискретного сообщения (последовательности элементов x_i) могут быть непосредственно человек и автоматические устройства. Преобразование же элементов x_i в последовательность бинарных дискретных сигналов, образующих позиционный код, осуществляется техническими устройствами, которые называются *передающими оконечными устройствами (ПОУ)*. В большинстве случаев, как было отмечено ранее, используются равномерные коды, упрощающие инженерную реализацию автоматической передачи, обработки и приема (распознавания) элементов. При этом применяются два основных режима отправления сообщений: синхронный и стартстопный.

При *синхронном режиме* работы дискретные элементы x_i должны создаваться через строго определенные интервалы времени T и, следовательно, посылки (сигналы), составляющие кодовую комбинацию и последовательность их, т. е. сообщение, передаются через строго определенные интервалы времени τ_0 (рис. 9.2, а, б). Передающее и приемное устройства работают синхронно и синфазно, поэтому в точке приема априорно известны интервалы времени, на которых появляются сигналы.

При *стартстопном режиме* работы элементы x_i могут создаваться в любые моменты времени t_i , определяемые источником (рис. 9.2, в). Для того чтобы в точке приема были известны интервалы времени, на которых появляются сигналы, перед началом каждой комбинации посылается *стартовый импульс*, запускающий приемную систему, а в конце кодовой комбинации — *стоповый импульс* с длительностью, превышающей длительность переходных процессов в приемном устройстве (рис. 9.2, г). В современных устройствах передачи сообщений стартовый импульс сходен с посылкой «0» и по длительности равен τ_0 , а стоповый импульс — по знаку совпадает с посылкой «1», а по длительности равен $1,5\tau_0$, что создает условия однозначного (при отсутствии помех) определения начала и окончания передачи очередного элемента сообщения x_i .

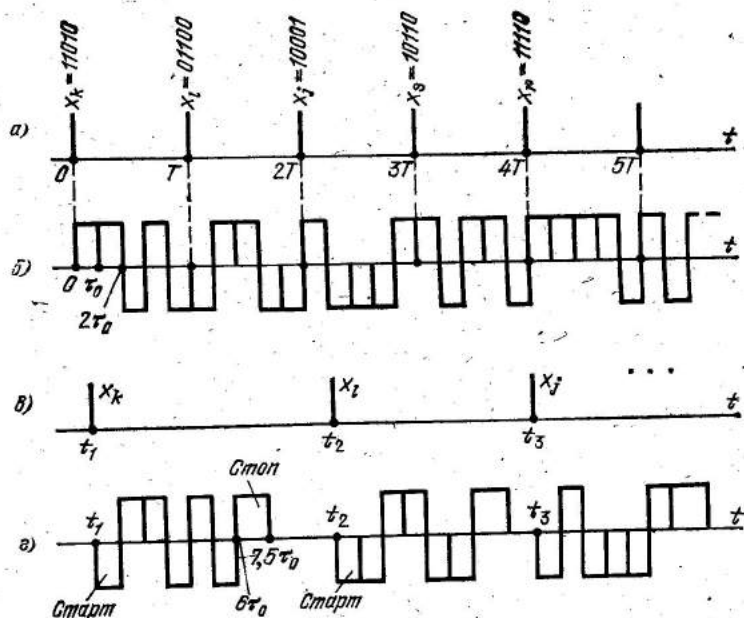


Рис. 9.2. Пояснение синхронного и стартстопного режимов передачи элементов сообщения

Из рис. 9.2 видно, что при стартстопном режиме скорость отправления при прочих равных условиях ниже, чем при синхронном. Добавление стартового и стопного импульсов в кодовую комбинацию осуществляется в оконечном передающем устройстве. Оно в явном или неявном (совмещенном) виде содержит:

кодирующее устройство, создающее на n выходах n кодовых элементов, соответствующих поступающему элементу информации x_i ;

накопительное (наборное) устройство, запоминающее кодовую комбинацию на время, необходимое для ее преобразования во временную последовательность (преобразования параллельного кода в последовательный);

распределительное устройство, преобразующее кодовую комбинацию, хранящуюся в наборном устройстве, в последовательность электрических импульсов;

генератор служебных разрядов (позиций), формирующий синхронизирующие импульсы, в частности стартовый и стопный;

управляющее устройство, определяющее способ и скорость создания кодовых комбинаций.

Ввод информации в ПОУ осуществляется вручную или автоматически. В первом случае каждому элементу x_i соответствует нажатие клавиши на клавиатуре типа стандартной пишущей ма-

шинки. При этом возможен практически только стартстопный режим отправления при скорости, не превышающей 400 элементов x_i в минуту, чему при $k=5$ соответствует скорость отправления

$$v = N(k + 2,5)/60 = 400(5 + 2,5)/60 = 50 \text{ Бод.}$$

Автоматический ввод может осуществляться непосредственно от источника (ЭВМ, аналого-цифрового преобразователя) или с промежуточного носителя. В обоих случаях сообщение передается в кодированном виде, так что кодирующее устройство в ПОУ не требуется. В качестве промежуточного носителя используются бумажная перфолента, перфокарты, магнитные ленты и магнитные карты, киноленты. Во всех случаях необходимы соответствующие устройства записи и считывания информации (перфораторы, магнитофоны, трансмиттеры, фотосчитыватели).

9.3. ПРИНЦИПЫ ОБРАЗОВАНИЯ КАНАЛОВ ПЕРЕДАЧИ

В современных системах непосредственная подача в линию импульсов кодовых комбинаций, созданных оконечными передающими устройствами, производится только на низших уровнях связи. Уже после первого узла коммутации абонентов (в районе, городе) дальнейшая передача сообщений (между районами, городами и странами) производится с помощью уплотнения линии или создания многоканальности, т. е. передачи по одной линии одновременно нескольких сообщений. Уплотнение линии оказывается возможным потому, что полоса частот, занимаемых сигналами оконечных передающих (и приемных) устройств значительно уже полос пропускания линий.

Как и при передаче непрерывных сообщений (гл. 4), в инженерной практике наиболее широко используются два основных способа уплотнения — частотный и временной (существует аппаратура, использующая оба принципа, т. е. *смешанное уплотнение*). При частотном уплотнении полоса линии $\Delta F_{\text{л}}$ разделяется набором фильтров на смежные практически не перекрывающиеся участки шириною Δf_c , в каждом из которых передаются сигналы, соответствующие элементам разных сообщений. Каждый такой отдельный участок образует *канал передачи*¹. При временном уплотнении линия поочередно коммутируется на все источники сообщений (сигналов) так, что позволяет за время между поступлением двух очередных сигналов каждого источника осуществлять передачу сигналов всех остальных источников.

Так как исторически сначала были построены инженерные сооружения для передачи непрерывных сообщений, в частности телефонных, то на первом этапе уплотнение дискретных сообщений производится в одном из двух диапазонов: 300—3400 Гц (каналы *тонального телеграфирования* (ТТ)) и 3400—10 000 Гц (каналы

¹ В технике под каналом понимают часто всю совокупность образующих его приборов (устройств).

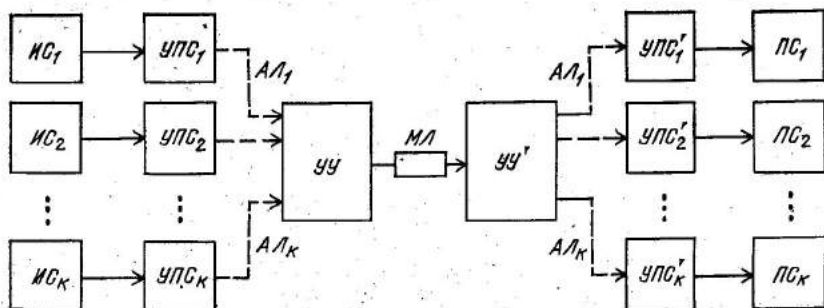


Рис. 9.3. Структурная схема многоканальной системы передачи

надтонового телеграфирования (НТТ)). Система передачи с точки зрения двустороннего обмена может использовать как *симплексный*, так и *дуплексный* режимы. В первом случае линии поочередно используется для отправления сигналов от точки А к точке В, а затем в обратном направлении. Во втором случае в каждом канале сообщения могут отправляться одновременно в обоих направлениях. Тривиальным способом создания дуплексного режима является использование двух линий (так называемая «четырепроводная система»). При наличии одной линии дуплексность создается частотным разделением сигналов разных направлений.

Таким образом, система передачи имеет структуру, представленную на рис. 9.3. Здесь K источников дискретных сообщений ИС (оконечных передающих устройств) управляют работой K устройств преобразования сигналов (УПС) (см. п. 9.1). Сигналы поступают через абонентские линии (АЛ) на устройство уплотнения (УУ), а уплотненные сигналы передаются в магистральную (междугородную или иную) линию (МЛ), на выходе которой производится разделение сигналов в приемных УПС для направления их к соответствующим получателям сигналов (ПС). Так как линии, как правило, дуплексные, то фактически на обоих концах системы находится аппаратура отправления и получения сообщений.

9.4. ОСНОВНЫЕ РЕГЛАМЕНТАЦИИ ДЛЯ СИСТЕМ ПЕРЕДАЧИ РАЗЛИЧНОГО БЫСТРОДЕЙСТВИЯ

Низкоскоростные системы (до 200 имп/с или Бод), или телеграфные. Исторически эти системы появились ранее других и не были специально предназначены для обмена информации между автоматическими или вычислительными устройствами. В соответствии с рекомендациями МККТТ низкоскоростные системы строятся по принципу тонального телеграфирования (частотного уплотнения) в стандартном тональном (300—3400 Гц) или надтоновом (3400—10 000 Гц) диапазонах. Общепринятые скорости отправления 50, 100 и 200 Бод при разности средних частот 120,

240 и 480 Гц; при этом в стандартном телефонном канале размещается 24, 12 или 6 каналов. При скорости 50 Бод используются сигналы как с амплитудной (посылки и паузы), так и с частотной манипуляцией (пара почти ортогональных сигналов). В отечественной аппаратуре широко применяется также скорость 75 Бод (при сигналах с частотной манипуляцией).

При использовании надтонального диапазона частот в нем размещают от 4 до 20 каналов при скоростях 50—75 Бод.

Низкоскоростные системы могут быть построены на основе двухпроводных или четырехпроводных линий (кабелей). Для реализации дуплексной передачи по двухпроводной линии применяются разделительные фильтры. При этом каналы, например с 1-го по 6-й, работают в одном направлении, а каналы с 7-го по 12-й — в другом, а для их разделения используются специальные фильтры прямого и обратного направлений.

При четырехпроводной системе передачи в прямом и обратном направлениях ведутся по разным парам проводов. Число каналов при прочих равных условиях при этом, очевидно, удваивается. При использовании надтонального диапазона участок спектра от 300 до 2700 Гц обычно отводится для телефонного сообщения, а частоты свыше 3400 Гц — для организации каналов передачи дискретной информации.

В низкоскоростных системах чаще всего используются коммутируемые линии (каналы), которые с помощью ручной или автоматической коммутации создаются из физически различных участков линии. Необходимость прохождения сигналов через устройства коммутации и участки линии с различными параметрами приводит к тому, что характеристики коммутируемых линий (затухание, полоса, уровень помехи) в среднем оказываются хуже, чем у некоммутируемых линий, используемых данными абонентами.

Среднескоростные системы передачи дискретных сообщений. Это относительно новые системы, предназначенные для передачи данных, т. е. формализованной бинарной информации, пригодной для ввода в ЭВМ со скоростями до 4800 Бод. В этом случае при наличии большого потока информации значительно чаще используются некоммутируемые (выделенные) каналы передачи. При этом в стандартном тональном диапазоне (300—3400 Гц) в соответствии с рекомендациями МККТТ скорость отправления может составлять 1200, 2400 и 4800 Бод (в зависимости от расстояний и допустимых вероятностей ошибок). Для двусторонней передачи используются четырехпроводные линии.

Для передачи данных используются также двухпроводные коммутируемые каналы. При этом передача может быть организована как в дуплексном, так и в симплексном режимах.

Дуплексная работа осуществляется с помощью разделения стандартного тонального диапазона на два равных по ширине поддиапазона, используемых для передачи в прямом и обратном направлениях (каналы 1-й и 2-й). Согласно рекомендациям МККТТ используются бинарные сигналы с частотной манипуляци-

ей со средними частотами $f_1=980$ Гц и $f'_1=1180$ Гц для 1-го канала и $f_2=1650$ Гц и $f'_2=1850$ Гц для 2-го канала. При этом нормальная скорость отправления составляет 200 Бод.

Предусмотрена возможность так называемого *полудуплексного режима работы*. При этом создается обратный канал служебного назначения с пониженной пропускной способностью. Поэтому стандартный тональный диапазон разделяется на две существенно неравные части: для служебного (обратного) и прямого каналов. Согласно рекомендациям МККТТ в обратном канале используются сигналы частотной манипуляции $f_{обр}=390$ Гц и $f'_{обр}=450$ Гц и скорости 75 Бод. В прямом канале также применяются сигналы с частотной манипуляцией при скоростях 600 Бод ($f_{пр}=1300$ Гц и $f'_{пр}=1700$ Гц) и 1200 Бод ($f_{пр}=1300$ и $f'_{пр}=2100$ Гц).

Абонентские пункты обычно строятся так, что канал может быть поочередно использован для телефонной связи или для передачи данных. С этой целью в системе имеется вызывное устройство и переключатель вида работ.

Высокоскоростные системы передачи дискретных сообщений. В стандартном тональном диапазоне частот скорость отправления практически не может превышать 4800 Бод. Для повышения быстродействия (при приемлемых ошибках распознавания) необходимо расширение полосы частот используемых каналов, обеспечивающих пропускание сигналов с малыми искажениями формы и приемлемым затуханием. Эта задача может быть решена за счет построения новой аппаратуры и использования широкополосных линий либо путем использования уже существующих систем передачи. Второй путь является, очевидно, экономически более выгодным, если только существующие линии и аппаратура уплотнения каналов не полностью загружены передачей информации. Для повышения скорости передачи сообщений несколько выделенных телефонных каналов объединяют в один широкополосный канал.

9.5. СТРУКТУРНЫЕ СХЕМЫ НИЗКОСКОРОСТНЫХ МНОГОКАНАЛЬНЫХ СИСТЕМ С ЧАСТОТНЫМ УПЛОТНЕНИЕМ

В этом случае используются сигналы или с амплитудной манипуляцией, АМн (посылка сигнала и пауза), или с частотной, ЧМн (два сигнала со смещением средних частот): В подавляющем большинстве случаев в аппаратуре используются сигналы в виде отрезков гармонического колебания длительностью τ_0 и, следовательно, полосой частот $\Delta f_c \approx 1/\tau_0$.

На рис. 9.4 приведена структурная схема многоканальной системы при использовании сигналов с АМн, на рис. 9.5 — временные диаграммы, поясняющие работу этой схемы. При этом для упрощения изображен лишь комплект устройств для отправления сообщений слева направо. Кривые, приведенные на рис. 9.5, относятся к выходам соответствующих блоков в произвольном i -м кана-

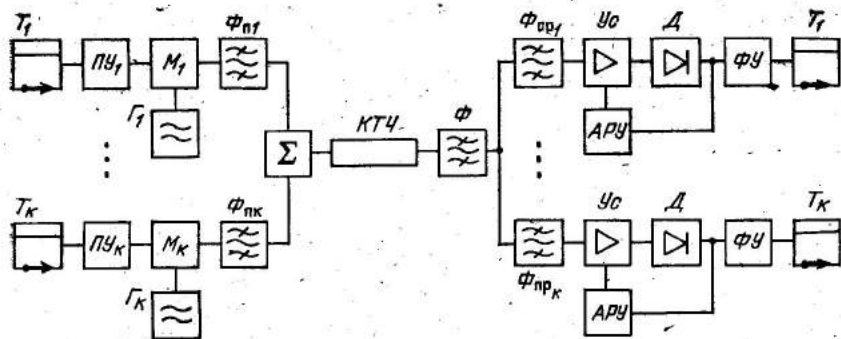


Рис. 9.4. Структурная схема многоканальной системы с частотным уплотнением бинарных сигналов с АМ

ле уплотнения (возможные искажения форм импульсов не указаны).

Импульсы кода от телеграфного аппарата T_i ($i=1, 2, \dots, K$) поступают на модулятор M_i через переходное устройство $ПУ_i$, преобразующее паузы сигналов телеграфного аппарата в отрицательные импульсы, что упрощает построение модулятора. Последний в случае сигналов с АМ, по существу, является ключом, пропускающим или не пропускающим гармонические колебания генератора Γ_i (начальная фаза сигналов при этом имеет случайные значения, однако сохраняется непрерывность фазы в последовательности их). Фильтры $\Phi_{пi}$ уменьшают взаимное влияние каналов передатчика. Частоты генераторов $\Gamma_1, \Gamma_2, \dots, \Gamma_K$ смещены на некоторую величину Δf , определяемую в основном полосой сигналов и качеством фильтров. Сумма канальных сигналов образует *линейный сигнал*, поступающий на вход канала тональной частоты (КТЧ) телефонного канала.

В точке приема имеется фильтр Φ , пропускающий суммарную полосу сигналов, предотвращающий попадание в приемник излишних помех. Приемные фильтры $\Phi_{прi}$ распределяют сигналы по различным каналам схемы. После усиления сигналы детектируются обычным амплитудным детектором (Д). Предусматривается схема автоматической регулировки усиления (АРУ). Выходные импульсы детектора поступают на формирующее устройство (ФУ), где производится распознавание импульсов по одному из принятых алгоритмов, а также превращение их в посылки и паузы, пригодные для срабатывания телеграфного приемного аппарата.

При использовании сигналов с частотной манипуляцией структурная схема рис. 9.4 должна быть изменена: в передатчике амплитудный манипулятор и генератор заменяются частотным манипулятором, представляющим собой управляемый реактивный элемент (емкость или индуктивность), присоединенный параллельно

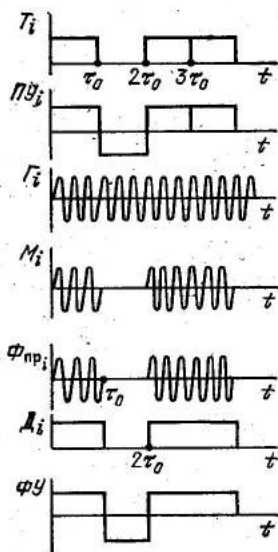


Рис. 9.5. Временные диаграммы, поясняющие работу схемы рис. 9.4

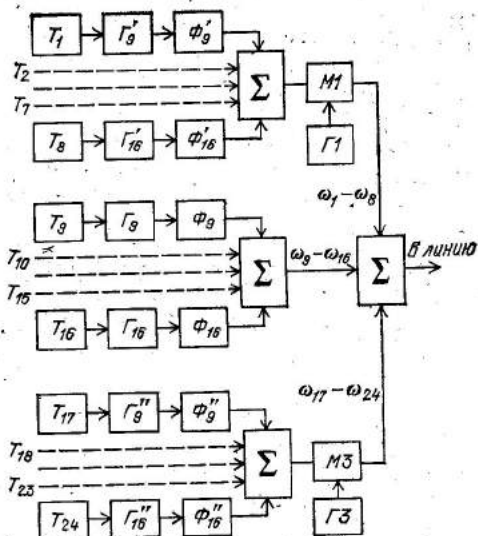


Рис. 9.6. К пояснению принципа группообразования при построении многоканальной системы с частотным уплотнением

контур автогенератора; в приемнике амплитудный детектор заменяется на частотный.

В рассмотренных структурных схемах число различных по параметрам фильтров, генераторов и некоторых других устройств равно числу каналов. Для увеличения однородности структуры используется принцип группообразования. При этом все каналы разделяются на несколько групп, например, с 1-го по 8-й, с 9-го по 16-й и с 17-го по 24-й (обычно групп три) (рис. 9.6). Каналы второй группы образуются вышеописанным способом с помощью фильтров $\Phi_9, \Phi_{10}, \dots, \Phi_{16}$ и генераторов $\Gamma_9, \Gamma_{10}, \dots, \Gamma_{16}$. Каналы же первой и третьей групп образуются с помощью фильтров и генераторов $\Phi'_9, \Phi'_{10}, \dots, \Phi'_{16}$ и $\Phi''_9, \Phi''_{10}, \dots, \Phi''_{16}$, имеющих те же параметры, что устройства второй группы. Затем сигналы этих двух групп переносятся в необходимые диапазоны частот с помощью процесса гетеродинирования. Для этого в структурную схему, показанную на рис. 9.6 упрощенно, вводятся генераторы несущих частот $G1$ и $G3$, модуляторы $M1$ и $M3$.

9.6. СТРУКТУРНЫЕ СХЕМЫ СРЕДНЕ- И ВЫСОКОСКОРОСТНЫХ СИСТЕМ ПЕРЕДАЧИ ДИСКРЕТНЫХ СООБЩЕНИЙ

Среднескоростные системы создаются в стандартном тональном диапазоне на дискретные скорости передачи 600, 1200, 2400, 4800 Бод. Передача сигналов в полосе 300—3400 Гц с приемлемой вероятностью ошибок при скорости отправления свыше 4800 Бод

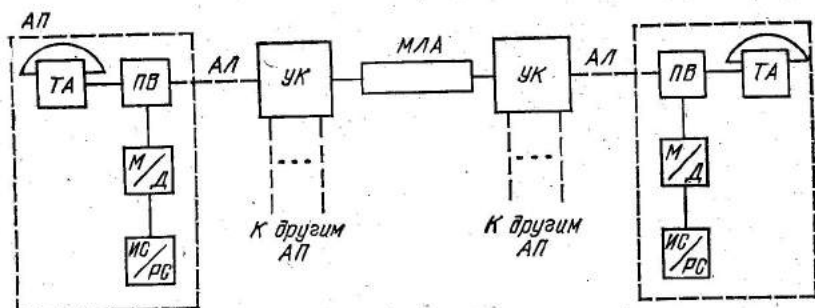


Рис. 9.7. Структурная схема коммутируемой сети передачи:

ТА — телефонные аппараты, ИС/РС — источники и регистраторы данных, М/Д — модемы (модулятор и демодулятор), УК — узлы коммутации, ПВ — переключатели вида работы, МЛ — магистральная линия с аппаратурой передачи, АЛ — абонентская линия

практически невозможна. В стандартном тональном диапазоне необходимо организовать дуплексный режим работы, поэтому среднескоростные системы являются одноканальными (один канал в прямом направлении, один — в обратном). По этой причине абонентский пункт среднескоростной системы строится так, чтобы канал поочередно (по автоматическому вызову) можно было предоставлять для передачи или телефонных или дискретных сообщений. С этой целью предусматривается *переключатель вида сообщений*.

Аппаратура строится в расчете на коммутируемые (проходящие через ручные и автоматические узлы коммутации) и некоммутируемые каналы. Некоммутируемые каналы обладают более высокими показателями.

Структурная схема коммутируемого канала передачи данных приведена на рис. 9.7.

При построении модемов, разделении стандартного канала на прямой и обратный и создании сигналов с частотной манипуляцией используются рекомендации МККТТ по номиналам частот (см. п. 9.4).

В среднескоростных коммутируемых системах предусматривается полудуплексный режим работы.

При дуплексном режиме скорость отправления (в каждом канале) равна 200 Бод, при полудуплексном пара сигналов с частотами 1,3 и 1,7 кГц обеспечивает отправление со скоростью 600 Бод, а пара с частотами 1,3 и 2,1 кГц — со скоростью 1200 Бод. В служебном канале скорость отправления равна 75 Бод.

Для высокоскоростных систем передачи данных (скорость >4800 Бод) требуется расширить полосу используемых частот по сравнению со стандартным тональным диапазоном. Этого можно достигнуть построением специальной аппаратуры или усложнением существующей многоканальной аппаратуры передачи телефонных сообщений. Во-втором, случае передачу дискретных сооб-

щений можно производить только за счет уменьшения числа телефонных каналов в сети. Используя стандартную аппаратуру каналообразования К-12, К-60 и др. (с частотным уплотнением) и телефонную систему с импульсно-кодовой модуляцией (с временным уплотнением), можно образовать широкополосные (по сравнению со стандартным тональным диапазоном) каналы передачи дискретных сообщений. При этом пользуются такими методами, как:

объединение нескольких телефонных каналов аппаратуры К-12; использование одной (или нескольких) многоканальных групп аппаратуры К-60;

применение манипуляции различных импульсных несущих в аппаратуре с ИКМ.

В аппаратуре К-12 производится частотное уплотнение 12 телефонных каналов, на каждый из которых с учетом защитных промежутков отведено по $\Delta f_{\tau} = 4$ кГц. При использовании амплитудной модуляции ряда генераторов несущих частот и выделении верхних боковых полос в линии применяется диапазон от 60 до 108 кГц. Образованное групповое колебание (12-канальная группа) путем гетеродинирования при несущей частоте $f_{ог} = 120$ кГц и выделения нижней боковой полосы переносится в диапазон от 12 до 60 кГц, который является выходным диапазоном для аппаратуры К-12. Часть из 12 фильтров в диапазоне 60—108 кГц может быть объединена для высокоскоростной передачи данных, поступающих от соответствующей аппаратуры.

На рис. 9.8,а показаны спектральные диаграммы описанных выше процессов: на верхнем рисунке — результат первого гетеродинирования (указана лишь несущая 12-го телефонного канала), на среднем — результат второго гетеродинирования и коэффициент передачи фильтра K_{ϕ} , выделяющего нижнюю боковую полосу частот, на нижнем — коэффициент передачи фильтра, объединяющего три канала (9-й, 10-й и 11-й).

На рис. 9.8,б приведена упрощенная структурная схема аппаратуры К-12 с добавлением цепей для передачи дискретных сообщений в тройной полосе (после объединения 9-го, 10-го и 11-го каналов в один). Подготовка сигналов, согласованных с объединенным каналом, производится в устройстве подготовки сигналов (УПС). Г — генератор несущего колебания на частоте $f_{ог} = 120$ кГц. Остальные блоки ясны из их условных обозначений. Нижняя часть схемы относится к приемной части.

Аппаратура К-60 позволяет объединить пять групп сигналов, созданных аппаратурой К-12 в одну 60-канальную группу. Принцип объединения остается тот же: полосы шириной 48 кГц (от 12 до 60 кГц) гетеродинируются так, что все пять групп занимают диапазон от 312 до 552 кГц. Канал любой 12-канальной группы может быть приспособлен для передачи дискретного сообщения (за счет ликвидации 12 телефонных сигналов).

Указанные варианты позволяют достигнуть при допустимой вероятности ошибок скорости соответственно 12 и 60 кБод.

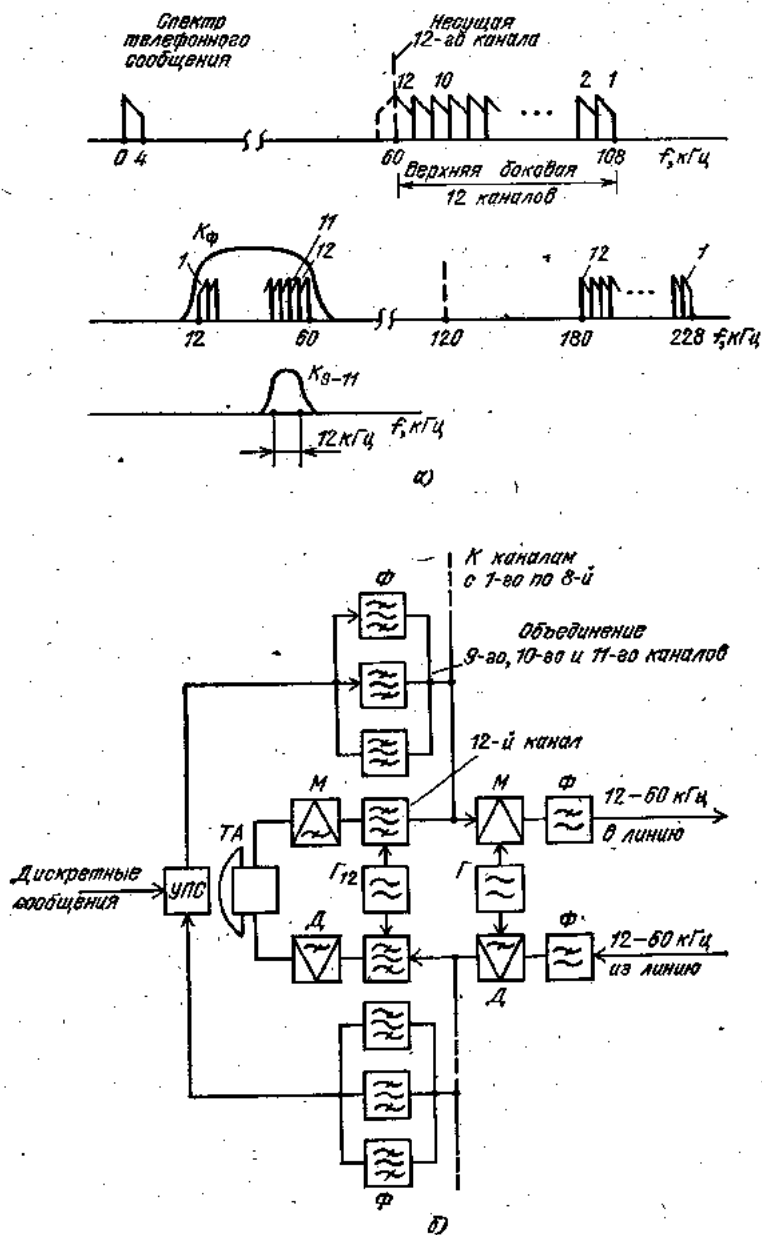


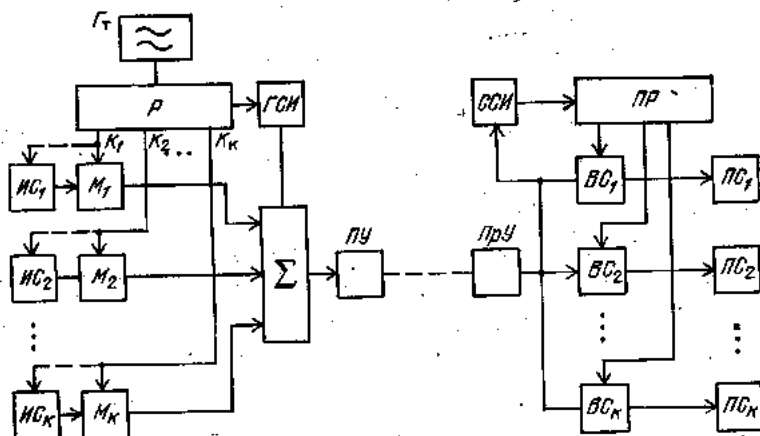
Рис. 9.8. К пояснению принципа действия высокоскоростной системы передачи данных с помощью объединения стандартных телефонных каналов (а) и упрощенная структура аппаратуры К-12 (б)

9.7. ПРИНЦИП ОБРАЗОВАНИЯ МНОГОКАНАЛЬНОСТИ С ПОМОЩЬЮ ВРЕМЕННОГО УПЛОТНЕНИЯ

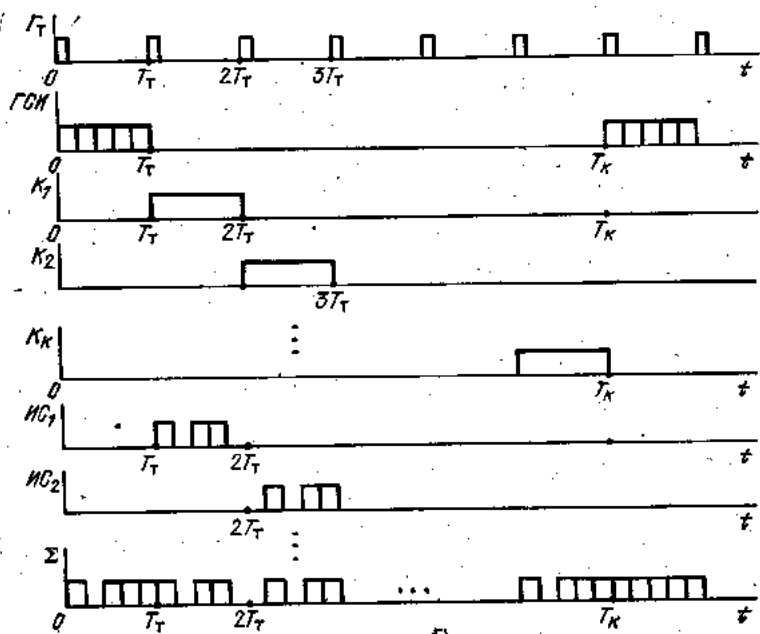
При временном уплотнении для передачи сигналов каждого источника отводится отдельный участок времени T_k . Физическое выделение участков производится с помощью синхронных на передающем и приемном устройстве коммутаторов или распределителей. Один из участков (временных каналов) отводится для передачи синхронизирующих сигналов, относительно которых производится однозначный отсчет участков времени отдельных каналов в обоих пунктах системы. При временном уплотнении, кроме полосовых сигналов, используемых при частотном уплотнении, могут применяться видеоимпульсы, минимальная длительность которых определяется, очевидно, предоставленной для системы передачи полосой частот и числом каналов.

Структурная схема синхронной системы с временным уплотнением приведена на рис. 9.9,а, а на рис. 9.9,б — даны временные диаграммы, поясняющие основные физические процессы (искажения сигналов в линии и аддитивные помехи не учтены). Ритм работы системы определяется генератором тактовой частоты Γ_T , вырабатывающим последовательность импульсов с периодом T_T , который в $(K+1)$ раз меньше периода T_k передачи канальных сигналов (рис. 9.9,б). Распределитель P за каждый период T_k вырабатывает $K+1$ канальных импульсов, сдвинутых на интервалы $T_T = T_k / (K+1)$, где K — число рабочих каналов. Выработанные периодические последовательности поступают соответственно на канальные модуляторы M_1, M_2, \dots, M_K и генератор синхронизирующихся импульсов (ГСИ). Последний вырабатывает синхронизирующую кодовую комбинацию. На вторые входы модуляторов поступают кодовые последовательности сообщений от соответствующих источников $ИС_1, ИС_2, \dots, ИС_K$. В рассматриваемом случае кодовые комбинации должны, очевидно, поступать (считываться с носителя) синхронно и синфазно с канальными импульсами. Для этого канальные импульсы от распределителя P должны быть поданы к источникам дискретных сообщений (штриховые линии на рис. 9.9,а). Сумма канальных и синхронизирующего импульсов подается на передающее устройство (ПУ), в котором при необходимости может производиться модуляция второй ступени, т. е. перенос видеоимпульсов на некоторую несущую частоту.

На приемной стороне сигналы поступают в приемное устройство (ПрУ), где производится предварительная фильтрация и усиление, а также детектирование, если в передающем устройстве имела место двухступенчатая модуляция. Импульсы синхронизации, выделенные с помощью селектора синхроимпульсов (ССИ), управляют работой приемного распределителя (ПР). Последний вырабатывает последовательности канальных импульсов, открывающих соответствующие временные селекторы $ВС_1, ВС_2, \dots, ВС_K$, представляющие собой схемы И. Таким образом, временные селекторы пропускают последовательности бинарных сигналов каж-



а)



б)

Рис. 9.9. К пояснению принципа действия синхронной системы передачи с временным уплотнением (а) и временные диаграммы ее работы (б).

дого канала соответствующим получателям сообщений ПС₁, ПС₂, ..., ПС_к (декодирующие устройства включены в блоки ПС).

В рассмотренной схеме временного уплотнения необходима подача импульсов от распределителя (Р) к источникам сообщений (ИС), которые могут быть удалены от устройства уплотнения и, следовательно, при этом необходимы специальные линии (штриховые линии на рис. 9.9,а).

9.8. СПОСОБЫ ФАЗИРОВАНИЯ (СИНХРОНИЗАЦИИ) ПОСЫЛОК

Процесс фазирования (синхронизации) состоит в обеспечении синфазного начала обработки бинарных посылок в точке приема и однозначного распределения принятых посылок по ячейкам накопителя (*i*-я посылка в *i*-ю ячейку). В первом случае говорят о процессе *фазирования по посылкам*, а во втором — о *фазировании по циклам*. Необходимость фазирования возникает вследствие недостаточной стабильности тактовых частот приемника и передатчика. Действительно, если относительная нестабильность каждого генератора тактовой частоты равна $k = \Delta F_{\max} / F_0$, где ΔF_{\max} — наибольшее отклонение во время работы, а $F_0 = 1/\tau_0$ — частота посылок, то расхождение генераторов на полный период произойдет за время

$$\Delta t = 1/2 \Delta F_{\max} = 1/2 k F_0 = \tau_0 / 2 k$$

(предполагается худший случай расхождения частот в разные стороны). Однако допустимое время расхождения $\Delta t_{\text{доп}}$, определяющее время сохранения синфазности $T_{\text{синф}}$, должно составлять лишь малую часть $\epsilon_{\text{доп}}$ величины Δt . Следовательно, время сохранения синфазности

$$T_{\text{синф}} = \epsilon_{\text{доп}} \Delta t = \epsilon_{\text{доп}} \tau_0 / 2 k.$$

Даже при минимальной скорости работы (50 Бод), кварцевых стабильных генераторах с нестабильностью порядка 10^{-5} и допустимой несинфазности порядка 0,1 время сохранения синфазности составит

$$T_{\text{синф}} = 0,1/2 \cdot 10^{-5} \cdot 50 = 100 \text{ с.}$$

Это время уменьшается пропорционально скорости передачи и для высокоскоростных систем составляет доли секунды.

В стартстопных системах фазирование по посылкам еще можно обеспечить за счет имеющейся стабильности генераторов (имеет место *автономное фазирование*). Однако в синхронных системах, где генераторы работают непрерывно, необходимо применять устройства *автоматического фазирования* по посылкам.

Фазирование по циклам (кодovým комбинациям) в стартстопных системах обеспечивается самым принципом передачи, а в синхронных — передачей специальных кодových комбинаций перед началом работы.

Фазирование по посылкам. Автоматическое фазирование по посылкам в современных системах осуществляется с помощью информационных (рабочих) сигналов без применения особых пилот-сигналов. Наибольшее распространение получили способ автоподстройки фазы генератора приемника (непосредственное воздействие на генератор) и способ деления частоты следования импульсов (или способ без непосредственного воздействия на генератор). Существуют различные варианты схемной реализации этих способов.

На рис. 9.10,а приведена упрощенная структурная схема автоподстройки с непосредственным воздействием на генератор. На

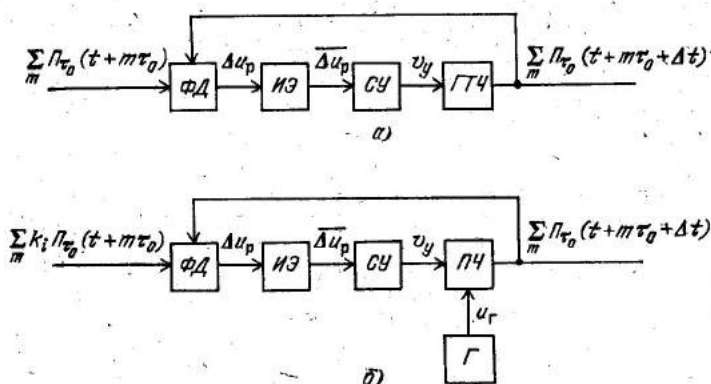


Рис. 9.10. Структурные схемы устройств фазирования по посылкам с непосредственным воздействием (а) и без непосредственного воздействия (б) на автогенератор

фазовый дискриминатор (ФД) воздействует последовательность принятых посылок

$$\sum_m k_i \Pi_{\tau_0}(t+mT)$$

и последовательность импульсов генератора тактовой частоты (ГТЧ)

$$\sum_m \Pi_{\tau_0}(t+m\tau_0 + \Delta t),$$

имеющих расхождение по времени на Δt . Выходное напряжение фазового детектора Δu_p , пропорциональное расхождению Δt , усредняется интегрирующим элементом (ИЭ) для ослабления случайных флуктуаций, вызванных помехами. Схема управления (СУ) вырабатывает непосредственное управляющее воздействие на генератор u_y . Управляющее воздействие может быть дискретным (релейным), при котором подстройка производится квазипериодически, и непрерывным (плавным), при котором автоподстройка частоты (например, с помощью управляемого реактивного элемента) осуществляется непрерывно во времени.

На рис. 9.10,б приведена упрощенная структурная схема фазирования без непосредственного воздействия на генератор приемника. В этой схеме генератор (Г) приемника вырабатывает импульсные колебания с частотой $f_r = kF_T$, где $k \gg 1$. Колебания генератора u_r воздействуют на вход преобразователя частоты (ПЧ), на другой вход которого подается принятая последовательность импульсов. В преобразователе частоты происходит не только деление частоты f_r в k раз, но и добавление (или исключение) отдельных импульсов в зависимости от расхождения фаз. В результате на выходе ПЧ образуется последовательность импульсов тактовой частоты, имеющая предельно достижимое в заданных условиях расхождение с последовательностью входных импульсов (случайная манипуляция входных импульсов передаваемой информацией устраняется особыми схемами).

Наибольшее распространение получили фазирющие устройства без непосредственного воздействия на генератор приемника. При этом не происходит снижение стабильности приемного генератора, что имеет место в схеме непосредственного воздействия из-за присоединения к автогенератору дополнительных дестабилизирующих цепей (это, в свою очередь, приводит к увеличению времени установления фазы).

Фазирование по циклам. Фазирование по циклам осуществляется с помощью передачи специальных фазирющих кодовых комбинаций, структура которых в точке приема априорно известна. Если эти комбинации передаются во время пауз в передаче сообщений, то способ фазирования называется *безмаркерным*. Если же элементы фазирющих комбинаций чередуются с элементами информационных посылок, то способ фазирования называется *маркерным*.

На рис. 9.11,а приведена упрощенная схема, поясняющая безмаркерный способ фазирования по циклам. В интервалах между передачей сообщений (но не реже чем с интервалом, на котором сохраняется синфазность) в передающее накопительное устройство (ПНУ) поступает фазирющая комбинация («маркер») от генератора фазирющих комбинаций (ГФК), которая отправляется в канал передачи. В точке приема фазирющая комбинация поступает на дешифратор (ДШ), который создает управляющий импульс u_r только в том случае, если распределители передатчика и приемника (РП и РПр) работают в фазе. При отсутствии этого импульса (и, следовательно, при расхождении фаз) управляющее устройство (УУ) изменяет фазу работы распределителя приемника РПр (управляющее воздействие указано штриховой стрелкой).

На рис. 9.11,б приведена упрощенная схема реализации маркерного способа фазирования по циклам. Генератор фазирющей комбинации (ГФК) («маркера») связан с одной из ячеек распределителя передатчика (РП) и, таким образом, фазирющая комбинация передается за несколько (l) циклов работы (длительностей кодовых информационных комбинаций). В распределителе

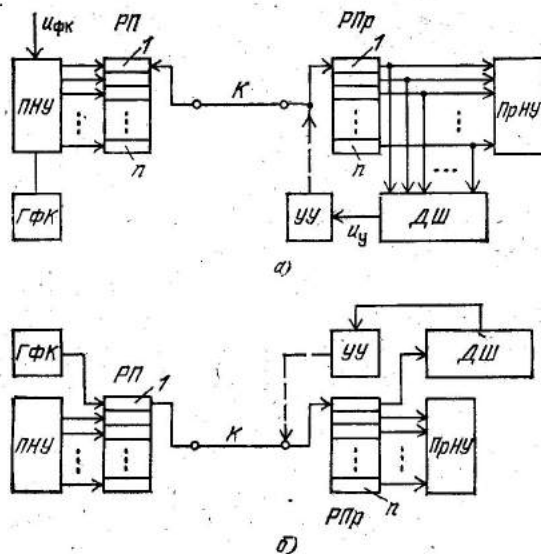


Рис. 9.11. Структурные схемы устройств безмаркерного (а) и маркерного (б) фазирования по циклам

приемного устройства (РПр) - к соответствующей ячейке присоединен дешифратор фазировочной комбинации (ДШ). При отсутствии синфазности распределителей приемника и передатчика дешифратор за l циклов работы не выдает выходного импульса на управляющее устройство и оно смещает фазу работы РПр на один элемент за каждый цикл работы. При вхождении в фазу дешифратор вырабатывает импульс, останавливающий работу управляющего устройства. Пока сохраняется синфазность, дешифратор через каждые l циклов вырабатывает импульс, поддерживающий «блокировку» управляющего устройства. При нарушении синфазности процесс подстройки начинается снова. Длина фазировочной комбинации выбирается так, чтобы вероятность ее появления на другой ячейке распределителя (что в принципе возможно) была весьма мала.

Маркерный способ фазирования по циклам имеет некоторые преимущества перед безмаркерным: при маркерном способе осуществляется автоматический контроль за синфазностью и нет необходимости периодически приостанавливать передачу сообщений, а при безмаркерном способе расхождение фаз обнаруживается лишь по резкому возрастанию числа ошибок в принятом сообщении. При наличии дуплексной (двусторонней) системы передачи сообщений возможна организация «автоматического» фазирования за счет посылки команд на передачу фазировочных комбинаций по обратному каналу. При этом фазирование можно организовать в обоих направлениях. Однако расхождение в одном кана-

ле также «автоматически» выводит из строя другой. При маркерном способе этого не происходит.

Рассмотренное выше маркерное фазирование называют иногда *синхронным фазированием*. Кроме того, используется также *стартстопное фазирование*. В этом случае до поступления сообщений распределители передатчика и приемника не работают. При поступлении сообщения отправляется стартовая кодовая комбинация и запускается распределитель передатчика. По окончании передачи сообщения отправляется *стоповая* кодовая комбинация. Обе комбинации дешифруются в приемном устройстве и соответственно запускают и останавливают работу распределителя приемного устройства. При стартстопном способе фазирования удается обеспечить более быстрое достижение синфазности, а также произвольность моментов начала передачи. Однако вероятность ошибок при этом в среднем может оказаться выше вследствие того, что ошибки при приеме стартстопных сигналов приводят к неверному приему всей последовательности сообщений. С последним недостатком можно бороться, увеличивая длину стартстопных фазирующих комбинаций, однако при этом неизбежно снижается средняя пропускная способность системы передачи.

Дополнительные сведения об особенностях аппаратуры систем передачи дискретной информации можно найти в [1—3].

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Дайте определения источника сообщений, устройства сопряжения, устройства защиты от помех, устройства преобразования сигналов, канала передачи, дискретного канала.
2. Дайте определения синхронного и стартстопного способов передачи сообщений.
3. Поясните принципы частотного и временного уплотнения, симплексного и дуплексного режимов передачи.
4. Перечислите основные регламентации для систем передачи различного быстродействия.
5. Приведите структурные схемы многоканальных систем передачи с частотным уплотнением при сигналах с АМн и ЧМн.
6. Поясните принцип группобразования при построении многоканальной системы.
7. Приведите структурные схемы среднескоростных систем передачи.
8. На каких принципах строятся высокоскоростные системы передачи?
9. Поясните принцип создания многоканальных систем с помощью временного уплотнения.
10. Объясните назначение фазирования по посылкам по циклам.

ГЛАВА 10. ПРИНЦИПЫ ПОСТРОЕНИЯ СЕТЕЙ СВЯЗИ

10.1. ВВЕДЕНИЕ

Связь является важнейшей составной частью инфраструктуры народного хозяйства страны. Быстрый и надежный обмен информацией необходим во всех сферах деятельности общества. Поэтому вопросам развития и совершенствования средств связи в нашей стране уделяется очень большое внимание. Создание надежной связи — общегосударственная задача. Особо важное значение в решении этой задачи придается созданию сетей связи и ускоренному развитию Единой автоматизированной сети связи страны (ЕАСС).

Сеть связи можно рассматривать как сложную систему, которая позволяет с помощью различных линий (каналов) и узлов коммутации объединить в требуемых сочетаниях большое число пространственно рассредоточенных источников и потребителей информации. В состав сети могут входить как стационарные (кабельные, радиорелейные), так и нестационарные линии и пункты связи (спутниковые линии связи, линии связи с подвижными объектами и др.).

Объединение линий связи различного вида в единую систему представляет очень сложную проблему, решение которой связано с решением большого числа научных, технических, экономических и организационных задач. Указанная проблема является одной из важнейших в нашей стране и необходимость ее решения привела к разработке общегосударственной Единой автоматизированной сети связи (ЕАСС), которая предусматривает эффективное совместное использование различных линий для обеспечения народного хозяйства и населения страны всеми необходимыми видами связи.

ЕАСС организационно и технически объединит в единое целое кабельные, радиорелейные, спутниковые и другие виды линий связи для передачи по ним телефонных, телеграфных, телевизионных, телеметрических и прочих видов сообщений. Широко используется ЕАСС для скоростной факсимильной передачи центральных газет в областные центры, автоматического управления технологическими процессами и объектами и т. п.

Наличие в составе ЕАСС различных видов линий связи и передача по ним разнообразной информации приводит к необходимости широкой унификации и стандартизации требований как к техническим параметрам и характеристикам применяемых систем, так и к показателям качества передачи информации.

Развитие и укрепление содружества социалистических стран и расширение международных связей сделало актуальной разработку единых норм и рекомендаций на качественные показатели и параметры систем связи в международном масштабе. Нормы,

принятые для ЕАСС, учитывают рекомендации МККР и МККТТ, в соответствии с которыми качественные показатели передачи по радиорелейным и спутниковым линиям не должны отличаться от аналогичных показателей передачи по кабельным линиям. Для стандартизации проектирования и эксплуатации разработан ряд рекомендаций на технические характеристики различных систем связи (полосы используемых частот, несущие частоты, девиация частоты, допустимое отношение сигнала к шуму в канале и т. д.).

ЕАСС представляет собой сложную развивающуюся систему, в которой стремятся учесть не только современные потребности в связи, но и существенное их увеличение в будущем. В настоящее время часть информационных потоков передается по линиям и сетям связи в аналоговом виде (за исключением передачи данных и некоторых других видов информации). Однако возрастающие потребности в гибкой, высококачественной и надежной связи приводят к быстрому развитию цифровых методов передачи, которые постепенно вытесняют аналоговые методы из различных областей, связанных с передачей информации. Эти тенденции учитываются при планировании перспектив развития и совершенствования ЕАСС.

В этой главе рассмотрены общие вопросы построения линий и сетей связи. Особенности построения ЕАСС рассмотрены в гл. 11.

10.2. ОСНОВНЫЕ ТИПЫ СЕТЕЙ, ОСОБЕННОСТИ ИХ ОПИСАНИЯ И РАБОТЫ

Сеть связи с геометрической точки зрения можно интерпретировать в виде графа, состоящего из узлов (узлов коммутации) и ребер (линий и устройств связи). Такой подход облегчает систематизацию и анализ сетей связи различного вида.

При построении сети возникает вопрос о выборе структуры соединения потребителей. Ниже рассмотрены основные типы сетей, используемые в зависимости от интенсивности обмена, территориального расположения абонентов и ряда других соображений (рис. 10.1).

В полносвязной сети (рис. 10.1,а) имеет место соединение каждого узла с каждым, чем обеспечивается, очевидно, предельная маневренность за счет дороговизны. При $У$ узлах число ребер $R_p = 0,5У(У-1)$. Нетрудно убедиться, что при изъятии $У-2$ ребер

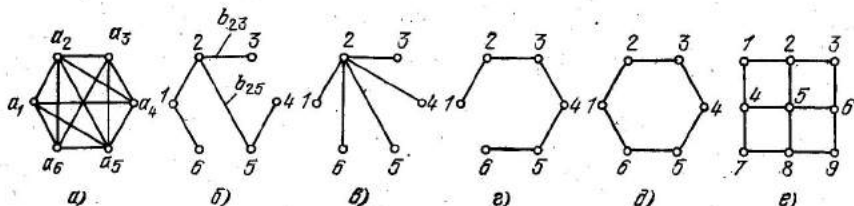


Рис. 10.1. Основные типы структур сетей

связность сети (возможность соединения любой пары абонентов) не нарушается.

В древовидной сети (рис. 10.1,б) никакая пара узлов не соединена более, чем одним ребром, а часть узлов не имеет прямых соединений друг с другом. Число ребер в такой сети $R_p = Y - 1$. Ее разновидностями являются узловая и линейная сети (рис. 10.1, в, г).

В матричной сети каждый узел соединен ребрами с частью узлов. Здесь возможно большое разнообразие вариантов, среди которых выделяют кольцевую сеть (рис. 10.1,д) и регулярные сети. Последние, в свою очередь, могут иметь разнообразные структуры, например, матричную (рис. 10.1,е), сотовую и другие в зависимости от числа сходящихся в узлах ребер.

Числа сходящихся в углах ребер r_i , называемые *рангами узлов*, на границе сети будут различными для граничных и внутренних узлов. Для регулярной сети

$$R_p = \frac{1}{2} \sum_{i=1}^{N-1} r_i Y_i,$$

где Y_i — число узлов ранга r_i .

Для количественного описания сети используется аппарат теории графов.

Взаимосвязь узлов и ребер описывается матрицей связности A , элементы которой a_{ij} определяются следующим образом:

$$a_{ij} = \begin{cases} 1, & \text{если между узлами } a_i \text{ и } a_j \text{ есть ребро,} \\ 0, & \text{если между узлами } a_i \text{ и } a_j \text{ нет ребра.} \end{cases}$$

Упорядоченный (каскадный) набор ребер μ_{ij} , соединяющих узлы a_i и a_j , не проходя дважды через один и тот же узел, называют *путем* или *маршрутом* из узла a_i в узел a_j .

Если путь начинается и заканчивается в одном и том же узле (не проходя дважды через какой-либо узел), то он является *контуром*. Так, древовидная сеть контуров не содержит, а кольцевая сеть — является контуром.

Если ребрам присваиваются положительные (или отрицательные) направления, то путь называют *направленным*. Число ребер r , входящих в путь μ_{ij} , называют *рангом пути*. Нетрудно доказать, что $1 < r < Y - 1$. Путь между узлами a_i и a_j по выбранному пути μ_{ij} называют *расстоянием* l .

Между двумя заданными узлами a_i и a_j существует, в зависимости от типа сети, несколько разных путей μ^k_{ij} .

Каждый k -й путь μ^k_{ij} от a_i к a_j можно записать упорядоченным перечнем узлов или дуг. Так, на рис. 10.1,б путь между a_1 и a_4 можно записать в виде

$$\mu_{14} = a_1 a_2 a_6 a_4 = b_{12} b_{26} b_{64}.$$

Для практического применения важно отыскать пути, обладающие минимальными рангами r , расстоянием l или другими тех-

нико-экономическими показателями. Решение этих задач возможно методами линейного программирования.

Число различных путей между избранной парой узлов a_i, a_j в сети называют связностью узлов h_{ij} . Минимальное из h_{ij} число по всей сети называют *связностью сети*. Так, для древовидной сети, приведенной на рис. 10.1,б, $h_{ij}=h=1$. Для полностью связной сети (рис. 10.1,а) $h=U-1=5$. Если есть пара узлов, для которой $h_{ij}=0$, то сеть называется *несвязной*.

Минимальный набор ребер, при котором сеть распадается на две не связанных (хотя бы одним ребром) называют *рассекающим набором* или сечением σ . Оно задается перечнем входящих в него ребер. Так, для сети, приведенной на рис. 10.1,б, рассекающим набором является единственное ребро: $\sigma_1 = \{b_{12}\}$ или $\sigma_2 = \{b_{25}\}$.

Иногда определяют сечение по отношению к конкретным узлам. Если сечение произведено так, что каждая из образовавшихся несвязных сетей является связной, то сечение ребер называют его рангом $r(\sigma_i)$.

Ребро графа является топологической моделью линии передачи, которая содержит как собственно линию (кабель, волновод или иную), так и аппаратуру обработки сигналов. Поэтому каждое ребро можно количественно описать несколькими показателями, существенными при оценке качества функционирования сети. Каждый из этих показателей в зависимости от задачи можно считать «весом» (значимостью) ребра. Знание весов позволяет определить основные количественные характеристики сети.

В матрице длин ребер L элементами являются геометрические расстояния l_{ij} между узлами a_i и a_j :

$$L = \|l_{ij}\|,$$

при этом, очевидно, $l_{ii}=0$, а при отсутствии ребра $l_{ij} \rightarrow \infty$.

В матрице пропускных способностей элементами являются суммы пропускных способностей c_{ij} всех каналов данного ребра:

$$C = \|c_{ij}\|.$$

Пропускная способность конкретного пути в силу каскадного соединения ребер определяется, очевидно, пропускной способностью наихудшего ребра (ребра с наименьшей пропускной способностью).

Аналогично строится матрица *емкостей* ребер, элементами которой является число стандартных каналов в соответствующих ветвях:

$$V = \|v_{ij}\|.$$

При этом, очевидно, величины v_{ii} неопределены, а при отсутствии ребра $v_{ij}=0$.

Аналогичным образом могут быть построены другие матрицы: стоимости, надежности, затуханий, вероятностей ошибок и другие. Так, матрица стоимостей ребер $\Pi = \|\pi_{ij}\|$, где π_{ij} — стоимость

линии между узлами (c_{ii} — стоимость узла). Матрица надежности $P = \|p_{ij}\|$, где p_{ij} — показатель надежности ребра b_{ij} .

На основе рассмотренных характеристик могут быть построены обобщенные показатели. Так, вводится *мощность сети*, которая определяется как сумма почленных произведений матрицы L и C .

Теория сетей позволила решить ряд специфических задач: определение оптимального пути (по заданному показателю), построение узловой сети с минимальным общим весом линий, определение и максимизация надежности путей в линии [1].

При функционировании сети возникает важная задача организации тракта передачи между удаленными корреспондентами сети. В настоящее время используются организация сквозного тракта и последовательная передачи с запоминанием в промежуточных пунктах. Первый способ называют *способом коммутации каналов*, а второй — *способом коммутации сообщений*.

Прямой (некоммутируемый) канал создается кроссированием (соединением) в узлах связи на определенное (обычно большое) время. По такому каналу не надо передавать адрес абонента. Параметры канала, как правило, высоки за счет исключения коммутационных устройств типа реле и возможности относительно просто производить корректировку электрических параметров канала с учетом реальных характеристик постоянно работающих устройств формирования этого канала.

Коммутируемый канал устанавливается каждый раз при вызове абонента с помощью коммутационных устройств (вручную или автоматически). При автоматической коммутации путь создания сквозного канала каждый раз может быть другим в зависимости от степени загрузки каналов. В каждой точке коммутации вносятся дополнительные помехи, поэтому сквозная характеристика канала становится случайной и ее коррекция без применения сложных адаптивных методов невозможна.

При коммутации сообщений абоненты имеют постоянную связь с ближайшим узлом. Сообщение вместе с адресом передается в узел коммутации сообщений (УКС), где оно записывается в запоминающем устройстве (ЗУ), становится на очередь и передается при освобождении канала на один из следующих узлов. Этот процесс записи и отправления продолжается до прихода в узел назначения.

На каждом узле производится проверка правильности сообщения. При ошибке в сообщении на передающий пункт передается запрос о повторении сообщения. В узлах коммутации сообщений может быть организовано длительное хранение сообщений для последующего контроля их содержания и правильности. В программу работы таких узлов можно заложить разнообразные формы обслуживания абонентов (приоритет, срочность, многоадресность и другие). Особенно полезна возможность связи оконечных устройств между собой в разное время, с разными скоростями, разными кодами.

Такое разнообразие функций сетей с коммутацией сообщений возможно осуществить только при использовании вычислительной техники и соответствующего математического обеспечения. Эти сложные коммутационные устройства в литературе часто называются *центрами коммутации сообщений (ЦКС)*.

Разновидностью коммутации сообщений является *коммутация пакетов*. Исходное сообщение делится в первом ЦКС на одинаковые по объему информации пакеты (в битах, байтах или машинных словах). Этими пакетами «плотно» упаковывается канал связи к одному из узлов. На оконечный узел пакеты одного сообщения могут прийти по разным каналам и в разное время. Из этих пакетов формируется первоначальное сообщение и передается абоненту. Возможна заготовка и объединение пакетов непосредственно на абонентских пунктах, которые должны быть оборудованы соответствующими вычислительными комплексами.

Следует подчеркнуть, что коммутация сообщений и пакетов в основном предназначена для передачи данных в дискретном виде. Однако в последнее время в связи с развитием проблемы прямого общения человека с ЭВМ на естественном языке появились методы анализа и синтеза речи с помощью дискретной техники и возможность организовывать передачу речи пакетным методом.

10.3. ИЕРАРХИЯ СЕТЕЙ И ЛИНИЙ

Передача сообщений происходит между исторически сложившимися сообществами потребителей информации, образующих иерархию. Укрупненными пользователями сетей являются: производства, населенные пункты, районы, области, республики. Каждый из перечисленных пользователей нуждается в передаче сообщений как на своем уровне, так и с пользователями всех других уровней.

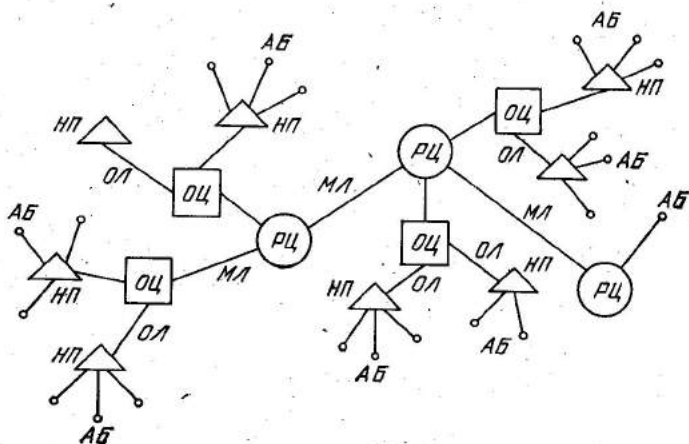


Рис. 10.2. Иерархия потребителей сообщений

Указанная ситуация упрощенно показана на рис. 10.2. Отдельные абоненты (АБ) соединены с узлами в населенных пунктах (НП) — городах, районных центрах — в *местную сеть*. Районные центры соединены областскими линиями (ОЛ) с областскими центрами (ОЦ), образуя при этом *областную сеть*. Областные центры магистральными линиями (МЛ) соединены с республиканскими (РЦ), образуя *магистральную сеть*.

Ситуация, указанная на рис. 10.2, очень упрощена. В действительности необходимо передавать различные виды сообщений (телефонные, телеграфные, телевизионные; дискретные данные, звуковое вещание). При этом к системам передачи предъявляются различные требования. На территории нашей страны передаются разнообразные потоки сообщений различного вида, что обусловлено необходимостью разделения всех сетей на *первичные* и *вторичные*.

Первичная сеть (рис. 10.3) образуется системой относительно широкополосных линий передачи и узлов коммутации двух видов — сетевых узлов (СУ) и сетевых станций (СС), обеспечивающих возможность передачи любого вида сообщений (в этом смысле первичная сеть «обезличена»). Сетевым узлом обеспечивается транзитную передачу типовых сигналов и их групп (групповых трактов), принадлежащих различным линиям передачи, а также производит передачу необходимого числа каналов вторичным сетям. Сетевая станция является, по существу, окончательным сетевым узлом. На ней производятся те же операции, кроме транзита типовых сигналов и групповых трактов.

Вторичные сети подключаются к каналам и групповым трактам первичной сети через сетевые станции. Вторичные сети разделяются по видам передаваемых по ним сообщений (телефонные, телеграфные, передача данных и др.).

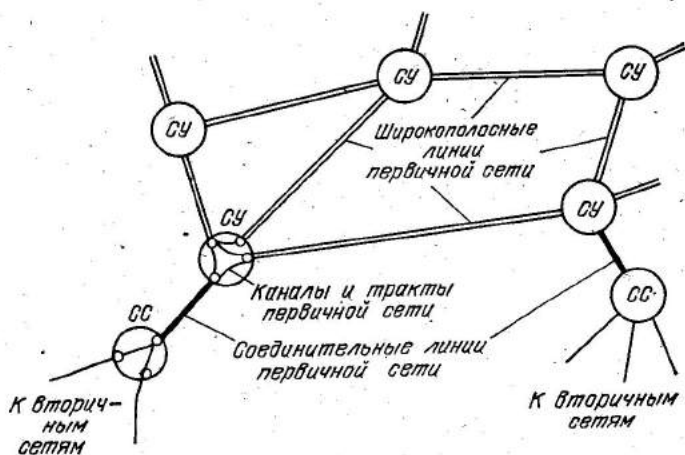


Рис. 10.3. Структура первичной сети и ее присоединение ко вторичным сетям

Общегосударственные вторичные сети (или вторичные сети общего пользования), как правило, независимы друг от друга. Так, телефонная и телеграфная сети имеют свои местные линии, абонентские установки и коммутационное оборудование. Обособленно существуют также вторичные сети звукового вещания, телевидения и др.

Указанные обстоятельства позволяют относительно независимо рассматривать первичные и вторичные сети, а также различные виды физических линий. С особенностями построения вторичных сетей связи и входящей в их состав аппаратурой можно познакомиться в [1].

Первичные сети связи и особенности аппаратуры таких сетей рассматриваются в гл. 11, посвященной общим принципам построения ЕАСС.

В сетях применяются различные по своей природе линии связи. Очень часто информация, поступающая от одного абонента к другому по сети связи может передаваться по составным каналам, которые на отдельных участках могут включать проводные, кабельные и другие линии связи, в том числе радиорелейные и спутниковые. Рассмотрим основные особенности различных линий передачи, используемых при создании сетей.

10.4. ОСОБЕННОСТИ РАДИОЛИНИЙ, ПРИМЕНЯЕМЫХ В СЕТЯХ

Передача сигналов в радиоканалах связи осуществляется с помощью электромагнитного поля излучения, создаваемого передающей антенной, и преобразования этого поля в сигнал с помощью приемной антенны. Антенны, выполняющие роль устройств, согласующих передатчик и приемник со средой распространения, являются важной инженерной частью радиолинии и в значительной степени определяют величину поля в различных точках пространства при заданной мощности генератора.

Свойства антенн концентрировать мощность излучаемых колебаний в определенном направлении (направленность излучения) и особенности распространения радиоволн (дифракция, поглощение, отражение) существенно зависят от частоты колебаний. Это приводит к тому, что свойства радиолиний в различных диапазонах волн существенно различаются [2, 3].

Линии на сверхдлинных (100 000—10 000 м) и длинных (10 000—1000 м) волнах. В этих диапазонах любую реальную антенну можно рассматривать как элементарный излучатель. Конструктивно антенны выполняются в виде вертикальных мачт и башен. На расстояниях порядка тысячи километров поле создается за счет дифракции электромагнитных волн вокруг земного шара при сравнительно слабом их поглощении в земле. Роль преломления волн ионосферой несущественна. Для определения напряженности поля обычно используются графики МККР [2].

На расстояниях, превышающих 1000 км, необходимо учитывать отражение волн от ионосферы и образование сферического волновода между поверхностью земного шара и слоем *E* ионосферы.

Характерными особенностями радиолинии в рассматриваемом диапазоне являются: низкий КПД антенн (отношение излученной мощности к подводенной), слабо выраженные суточные изменения напряженности поля со слабым увеличением ночью, практическая независимость поля от времени года, узкая полоса пропускания частот (от нескольких герц до нескольких килогерц).

Линия на средних (1000—100 м) волнах. В этом диапазоне на формирование поля в точке приема существенно влияют дифракция волн вокруг земного шара и преломление в ионосфере. Свойства линии значительно изменяются в течение суток. В дневные часы слой *D* ионосферы практически поглощает всю энергию волн, идущих под углом к горизонту, и поле в удаленных точках создается за счет дифракции. Расчет напряженности поля производится по графикам МККР, скорректированным с учетом дифракции. При реально применяемых мощностях излучения предельная длина линии передачи составляет около 500 км по суше и 1000 км над морем. В ночные часы сказывается преломление волн в слое *E* (слой *D* исчезает из-за рекомбинации электронов и ионов). Вследствие изменения преломляющих свойств слоя и сложения преломленного поля с полем, обусловленным дифракцией волн, поле в точке приема значительно флуктуирует. Появляются замирания длительностью от нескольких секунд до нескольких минут. Для борьбы с замираниями используются антифединговые антенны с пониженным уровнем излучения в направлениях, отличных от горизонта (для ослабления преломленного поля).

Характерными особенностями радиолиний в этом диапазоне являются: более высокий КПД антенн, ярко выраженные суточные изменения напряженности поля, наличие быстрых и медленных замираний.

Линия на коротких (100—10 м) волнах. В коротковолновом диапазоне явление дифракции не имеет существенного значения, и поле в точке приема создается в основном за счет преломления и отражения в слоях F_1 и F_2 ионосферы (слои *D* и *E* определяют в основном поглощение энергии волн).

Вследствие нестабильности и случайности характеристик отражающих слоев, возможен приход в точку приема нескольких лучей, отраженных от ионосферы и от земли или от разных слоев ионосферы. В результате этого, а также рассеяния на неоднородностях ионосферы, напряженность поля в точке приема изменяется (флуктуирует) в значительных пределах.

Работа коротковолновых линий связи также существенно зависит от состояния солнечной активности и изменений магнитного поля земли.

Основными особенностями линии на коротких волнах являются: сильно выраженная зависимость напряженности поля от времени года и суток, наличие глубоких и быстрых замираний поля (Фединга). Для борьбы с ошибками, вызываемыми замираниями, используется разнесенный прием. В стационарных радиоприемах связи обычно применяют два пространственно разнесенных пункта, в каждом из которых прием одних и тех же сообщений ведется на двух разных частотах. Организация такого приема сообщений соответствует 4-кратному разнесению. Иногда для упрощения линии связи организуют только пространственное или только частотное разнесение.

Помимо разнесенного приема эффективным средством борьбы с замираниями является адаптивный прием широкополосных сигналов (см. гл. 7).

Линия на ультракоротких (10—1 м) волнах (УКВ). Волны этого диапазона практически не дифрагируют и не отражаются от ионизированных слоев ионосферы. Поэтому устойчивая радиосвязь образуется лишь в пределах прямой видимости между передатчиком и приемником. Дальность прямой видимости определяется формулой

$$r \approx 3,6(\sqrt{h_1} + \sqrt{h_2}),$$

где h_1 и h_2 — высоты расположения передающей и приемной антенн над поверхностью в метрах, а r — дальность в километрах.

При благоприятных условиях УКВ могут распространяться на значительные расстояния за счет отражения от ионосферы. Такие условия возникают сравнительно редко, поэтому практически использовать дальнейшее распространение УКВ затруднительно.

Линии на сантиметровых, миллиметровых и более коротких волнах. Для сантиметровых волн (10—1 см) дифракция и отражение от ионосферы не имеют практического значения. Применение остронаправленных антенн позволяет избавиться от отражения от Земли. Поэтому связь возможна только в пределах прямой видимости. Однако уровень напряженности поля на входе приемника существенно зависит от поглощения энергии в осадках, молекулах воды и кислорода.

Заметное молекулярное поглощение происходит на волнах короче 1,5 см и имеет явно выраженный частотно-зависимый характер. В «максимумах» на волнах 1,35; 0,5; 0,25 см; 1,5; 0,75 мм поглощение достигает 1—10 дБ/км. Наиболее слабое поглощение в «окне» 0,5—1,35 см имеет порядок 0,06 дБ/км.

Линия на миллиметровых и субмиллиметровых (0,1—1 мм) волнах характеризуется резким возрастанием затухания в дожде и тумане (до 40 дБ/км для $\lambda=1$ мм при среднем дожде, до 100 дБ/км для $\lambda=0,1$ мм в среднем тумане). Молекулярное поглощение имеет ярко выраженный частотно-зависимый характер с затуханием в «пиках» до 20—30 дБ/км. Организация линии связи возможна практически лишь в «окнах прозрачности». В миллиметровом диапазоне их приблизительно десять, в субмиллимет-

ровом — десятки (наиболее широкое окно в окрестности $\lambda = 0,87$ мм).

Линия на оптических (световых) волнах. Оптические линии в открытом пространстве делятся на линии видимого (0,4—0,75 мкм) и инфракрасного света (0,75—100 мкм). На этих частотах существенно возрастают затухание в тумане (до 100—1000 дБ/км) и дожде (40 дБ/км), поэтому оптические линии нельзя использовать в этих условиях. Молекулярное поглощение образует большое количество пиков и окон прозрачности.

Самое прозрачное «окно» образуется в диапазоне 0,4—0,85 мкм (совпадает с областью чувствительности глаза). В инфракрасном диапазоне насчитывают семь-восемь окон прозрачности (2—2,5; 3,2—4,2; 8—13,5 мкм и другие).

На устойчивость линии связи на приземных трассах оказывают влияние неоднородности атмосферы, вызывающие медленные качания луча, изменение его поперечного сечения, нарушение когерентности и поляризации. В результате даже в пределах прямой видимости принятые сигналы подвержены флуктуациям.

10.5. ОСОБЕННОСТИ «ЗАКРЫТЫХ» ЛИНИЙ СВЯЗИ

Для связи между подвижными объектами может быть использована только «открытая» линия. Однако в других случаях при выборе типа линий для построения сети связи решающее значение в зависимости от ситуации могут иметь различные показатели. Чаще всего это стоимость эксплуатации в единицу времени стандартного канала или (для новых линий) стоимость построения линии в расчете на один канал. При более общем подходе следует учитывать комбинацию этих показателей. В особых ситуациях решающее значение может иметь предельная широкополосность линии (например, при передаче телевизионного изображения), ее надежность, неуязвимость и другие показатели.

Кроме радиолиний в сетях широко используются воздушные (проводные), городские телефонные кабели, коаксиальные и симметричные кабели, радиочастотные коаксиальные кабели, волноводы и световоды.

Воздушные линии применяются для организации местных связей. Их максимальная полоса пропускания составляет около сотен килогерц.

Городские телефонные кабели изготавливаются в нескольких вариантах в зависимости от вида изоляции и внешней защитной оболочки (типы Т, ТПК и другие). В городских телефонных кабелях предусмотрено «пространственное разнесение». Число пар проводов может быть весьма различным: от 5 до 1200.

Выпускаемые промышленностью в настоящее время *коаксиальные кабели* можно разделить на пять основных групп: микрогабаритные, малогабаритные, средние, большие и подводные (являющиеся разновидностью больших кабелей). Кабели этих групп отличаются диаметрами проводников, полосой пропускания и на-

значением (местная, зонавая, магистральная связь или связь под водой).

Радиочастотные кабели применяются в широком диапазоне волн (от длинных до ультракоротких волн). Согласно государственному стандарту выпускаются кабели трех типов: РК — радиочастотные коаксиальные, РС — радиочастотные со спиральными проводниками, РД — радиочастотные симметричные двужильные или из двух коаксиальных пар. Установлены стандартные ряды номиналов по волновому сопротивлению (например, для типа РК: 50, 75, 100, 150 и 200 Ом).

В сантиметровом диапазоне волн в качестве закрытых направлений систем применяются прямоугольные и круглые полые волноводы с металлическими токонесущими стенками. Возможны также диэлектрические и металло-диэлектрические волноводы. Однако затухание в таких волноводах значительно выше.

Согласно стандарту волновод представляет собой стальную трубку с внутренним диаметром 6 см, толщиной стенок 3 мм, с нанесением на внутреннюю поверхность слоя меди толщиной 20 мкм и лаковой пленки. Наружная защита имеет вид полиэтиленовой оболочки.

Одной из особенностей волновода является наличие критической длины волны, т. е. по волноводу могут распространяться только волны меньше критической (волновод действует как фильтр верхних частот).

В волноводах могут распространяться поперечные электрические волны типа ТЕ или Е волны и поперечные магнитные волны типа ТМ или Н. Наибольший интерес для дальней связи представляет поперечная электрическая волна типа TE_{01} , называемая иначе H_{01} . Частотная зависимость затухания этой волны, в отличие от других волн, имеет непрерывно падающий характер (чем выше частота, тем меньше потери и тем меньше затухание в волноводе). Однако волна типа H_{01} сравнительно неустойчива и при неизбежных изгибах и деформациях волновода подвержена перерождению в другие типы волн. Чтобы исключить эти нежелательные явления при передаче волны типа H_{01} используют волноводные фильтры (периодические структуры), обладающие свойствами пропускать волны типа H_{01} и задерживать волны всех остальных типов. Такие фильтры размещаются через каждые 1—2 км волновой линии.

В реальной волноводной линии затухание определяется наличием потерь нескольких видов. Основные потери — это тепловые потери в проводящих стенках волновода, зависящие от длины волны и диаметра волновода. Для уменьшения затухания следует уменьшить длину волны и увеличить диаметр волновода.

Дополнительные потери складываются из потерь на преобразование в паразитные типы волн, потерь в диэлектрической пленке (наносимой на внутреннюю поверхность цельнометаллических волноводов) и потерь в газе.

Экспериментальные исследования показывают, что в реальной волноводной линии диаметром 60 мм общее затухание имеет порядок 2—4 дБ/км в диапазоне 30—100 ГГц.

Технико-экономическое сравнение кабельных и волноводных линий показывает, что последние эффективны лишь при необходимости создавать большое число каналов (более 30 тыс.).

Световодные линии можно разделить на два класса: с дискретной фокусировкой (линзовые и зеркальные) и с непрерывной фокусировкой (зеркальные трубы и волоконные нити).

Линзовый световод представляет собой непрозрачную трубу диаметром 5—10 см с периодически расположенными линзами или зеркалами через каждые 100—200 м. Теоретически при этом возможно весьма малое затухание (менее 1 дБ/км), однако оно реализуется при весьма высокой прямолинейности оси (отклонения меньше 1 мм), большом радиусе кривизны при поворотах (не менее 100 м) или же требуется применение дорогой системы юстировки. Световоды в форме зеркальных труб (аналогичные обычным волноводам) дороги и экономически нецелесообразны.

Волоконный световод является в настоящее время одной из наиболее перспективных линий передачи сообщений. Наиболее распространенный волоконный световод представляет собой двухслойное стекловолокно. Распространение энергии в нем происходит за счет полного внутреннего отражения. В реальных конструкциях стекловолокна собираются в пучки, а из них монтируется кабель с наружной защитной оболочкой.

В одномодовых системах используется когерентное излучение лазера при одномодовом распространении световых волн по волокну толщиной в несколько микрон (толщина соответствует критической длине волны).

Важной характеристикой световодов являются потери пропускания, обусловленные поглощением и рассеянием энергии светового сигнала в материале световода. Количественно потери оценивают величиной затухания сигнала в дБ/км. Потери в современных световодах приближаются к теоретически возможным. Как известно [4, 5], нижний теоретический предел потерь для световодов из кварца составляет менее 1 дБ/км.

При увеличении диаметра волокна до 50—100 мкм устанавливается многомодовый режим. При этом в качестве источника излучения можно использовать обычные некогерентные источники света (светодиоды и др.). В многомодовом режиме возрастает затухание и появляются существенные искажения из-за дисперсии различных типов волн одной частоты.

В одномодовом оптическом кабеле теоретически можно образовать до 1 млн. телефонных каналов. Неидеальная когерентность источников и дисперсия волн в реальных волокнах снижают эту величину до нескольких десятков тысяч.

Оптический кабель имеет много других преимуществ: малогабаритность, гибкость, легкость прокладки с поворотами, защищен-

Таблица 10.1

Тип линии	Используемый диапазон, Гц	Порядок длины волны	Порядок числа каналов	Реальная система уплотнения
Воздушная (проводная)	10^3-10^5	км	10	K-12
Симметричный кабель	10^3-10^6	сотни	100	K-60, K-120
Коаксиальный кабель	10^3-10^8	м	10000	K-1920 K-3600
Волновод	$10^{10}-10^{11}$	мм	100000	K-10800
Световод (волоконный)	$10^{14}-10^{15}$	мкм	1000000	

ность от помех, легкость изготовления большими длинами, влагостойкость.

Для того чтобы дать представление о возможностях различных видов «закрытых» линий (не излучающих электромагнитное поле в открытое пространство) в табл. 10.1 приведены частотные свойства основных линий и реально достижимое число аналоговых стандартных телефонных каналов.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Перечислите основные типы топологических структур сетей.
2. Дайте определение матрицы связности сети.
3. Что такое маршрут, ранг пути, расстояние и связность сети?
4. Какие основные задачи можно решить на сети?
5. Перечислите основные особенности способов коммутации каналов и коммутации сообщений.
6. Дайте определение первичной и вторичных сетей.
7. Перечислите основные типы вторичных сетей.
8. Укажите основные используемые типы физических линий передачи сигналов.
9. Определите возможное число каналов в стандартном волноводе диаметром 6 см.
10. Чем отличаются одномодовый и многомодовый режим в световоде?
11. Укажите особенности радиолинии в диапазонах длинных, средних, коротких и ультразвуковых волн.

11.1. ВВЕДЕНИЕ

Единая автоматизированная сеть связи (ЕАСС) в нашей стране предназначена для сбора, обработки и распределения всех видов сообщений: телефонных, видеотелефонных, телеграфных, фототелеграфных, телевизионных, звукового вещания, дискретных данных для ЭВМ и автоматизированных систем управления (АСУ). Эти виды сообщений существенно отличаются друг от друга по спектрам, требованиям к качеству передачи, необходимой скорости доставки к потребителю и допустимой задержке. Кроме того, необходимо учитывать сложившиеся различия в территориальном распределении различных источников и потребителей. В СССР ЕАСС строится с учетом этой исторически сложившейся ситуации, с использованием оптимальных методов образования сетей и построения аппаратуры, максимально возможной унификации методов передачи сигналов, целесообразной автоматизации контроля и управления сетью.

Исторически сети электросвязи начали развиваться как сети передачи телефонных сообщений. Эксперименты показали, что хорошая разборчивость и даже достаточная натуральность воспроизведения речи обеспечивается при передаче полосы частот от 300 до 3400 Гц. Именно эту полосу рекомендует МККТТ в качестве стандартного канала тональной частоты («канала ТЧ»). Позднее была организована передача по такому каналу дискретной информации со скоростью 50 Бод, а затем и выше (до 4800 Бод). При этом использовался принцип частотного уплотнения. Типовые устройства такого уплотнения рассмотрены в гл. 9.

Таким образом, телефонный канал стал универсальным каналом в сети связи. Он коммутируется с помощью коммутационных устройств сети. Создание других, более широкополосных каналов производится путем кратного объединения стандартных каналов ТЧ.

Канал передачи ЕАСС — это совокупность технических средств и среды распространения сигналов, обеспечивающих передачу сообщений в определенной полосе и с определенной скоростью. С точки зрения способа передачи и технических средств каналы разделяются на *аналоговые* (непрерывные) и *цифровые* (дискретные). По цифровым каналам могут передаваться не только дискретные сообщения, но и аналоговые, представленные своими выборочными значениями (гл. 2, 4).

Основной характеристикой аналогового канала является полоса пропускания (передачи), а цифровой — скорость передачи в битах в секунду при заданной вероятности ошибок.

При передаче информации по сети отдельные каналы ТЧ объединяются в группы или пучки. Это целесообразно с точки зрения

не только эффективности использования систем передачи, но и оптимального распределения потоков сообщений по всей сети через коммутационные устройства и линии. Последнее особенно важно для тех пунктов сети, между которыми традиционно осуществляется обмен информацией. Формирование групп каналов увеличивает также оперативность управления сетью как в режимах нормального функционирования, так и при повреждениях, ремонтно-профилактических работах, изменении нагрузки.

Типовые групповые тракты ЕАСС — это совокупность технических средств, включая среды распространения сигналов, обеспечивающих передачу сообщений нормализованной группы каналов. Для каналов ТЧ образуются четыре такие группы:

первичный групповой тракт из 12 каналов ТЧ;

вторичный групповой тракт из 60 каналов;

третичный и четвертичный групповые тракты соответственно из 300 и 900 каналов ТЧ.

Кроме стандартного канала ТЧ и групповых трактов на его основе в ЕАСС формируются каналы передачи телевизионных изображений в полосе 50 Гц—6 МГц, звукового сопровождения телевидения высшего класса (30—15 000 Гц) и первого класса (50—10 000 Гц), а также каналы передачи звукового вещания в полосах 30—15 000, 50—10 000 и 100—6300 Гц.

Цифровые методы передачи непрерывных сообщений начали развиваться также применительно к телефонным разговорам. В ЕАСС по ряду соображений основным видом передачи принята импульсно-кодовая модуляция (ИКМ). При этом скорость передачи в одном телефонном канале составляет 64 кбит/с. Эта величина соответствует рекомендациям МККТТ и принята фактически всеми странами. Исследуются и применяются также другие виды передачи, например дельта-модуляция [1].

Типовой цифровой канал со скоростью передачи 64 кбит/с также может быть вторично уплотнен для передачи по нему до нескольких сот телеграфных сообщений со скоростью 50 Бод.

11.2. ПЕРВИЧНЫЕ И ВТОРИЧНЫЕ СЕТИ ЕАСС

Технико-экономические исследования принципов построения ЕАСС показали целесообразность разделения ее на два основных вида сетей — первичные и вторичные (см. гл. 10).

Первичная сеть состоит из совокупности линий передачи, сетевых узлов и сетевых станций, образующих сеть типовых каналов и типовых групповых трактов, обеспечивающих передачу любых видов информации и распределение их между вторичными сетями.

Вторичные сети, подключаемые к типовым каналам и групповым трактам первичной сети, разделяются по виду передаваемых по ним сообщений. Они подразделяются на телефонные, телеграфные, сети передачи данных, звукового и телевизионного вещания, сеть передачи газет фотоспособом и другие (рис. 11.1).

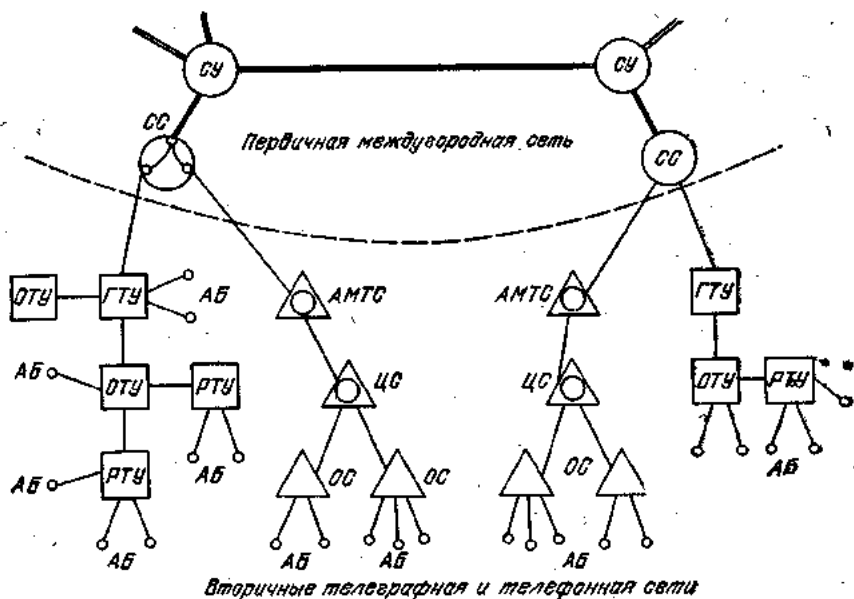


Рис. 11.1. К пояснению структуры вторичных сетей:

РТУ, ОТУ, ГТУ — районные, областные, главные телеграфные узлы коммутации; ОС, ЦС, АМТС — оконечные, центральные, автоматические междугородные телефонные станции; АБ — абоненты

Сетевые узлы первичной сети обеспечивают переключение и транзитную передачу типовых каналов и групп (трактов), принадлежащих различным физическим линиям передачи. В сетевом узле производится передача вторичным сетям необходимого числа каналов. Непосредственное подключение вторичной сети осуществляется через сетевые станции, которые, по существу, являются оконечными узлами сети. На них, в отличие от сетевых узлов, не производится сквозной транзит каналов.

В сетевых узлах и сетевых станциях производится, как правило, долговременное соединение каналов и трактов. Оно осуществляется вручную на специальных коммутационных панелях, называемых кроссами. В последние годы начинают также применяться автоматические устройства переключения.

Таким образом, первичная сеть «обезличена» по отношению к видам сообщений. Вторичные сети, наоборот, «специализированы».

Вторичные сети можно классифицировать по ведомственной принадлежности. Так, существует вторичная сеть Министерства путей сообщения. Она тесно «привязана» к структуре железных дорог и выполняет функции передачи телефонных сообщений, контроля и управления движением на железной дороге. Существуют вторичные ведомственные сети связи для газопроводов и нефтепроводов и других отраслей народного хозяйства.

В нашей стране координация строительства и развития всех ведомственных средств электросвязи осуществляется Межведомственным координационным советом (МВКС) по созданию Единой автоматизированной сети связи. Координация ведется с учетом потребностей ведомств, обеспечения общегосударственных интересов в развитии связи с учетом взаимного резервирования ведомственных и общегосударственных средств связи на основе максимальной унификации и стандартизации технических средств.

Первичная сеть ЕАСС, исходя из особенностей эксплуатации и сложившейся структуры управления, разделяется по территориальному признаку на магистральные, внутриобластные и местные сети.

Магистральная первичная сеть соединяет между собой все областные и республиканские центры, *областная первичная сеть* предназначена для связи всех районных центров друг с другом и с областным центром, *местные сети* подразделяются на городские и сельские. Они объединяют узлы и станции сельской местности между собой и с районным центром. Городская сеть объединяет городские станции и узлы между собой и также с районным центром, если город или поселок не является райцентром.

В последнее время введено понятие *зоновой* первичной сети, используемое применительно к принятой системе нумерации абонентов автоматической коммутируемой телефонной сети СССР. В ряде случаев в зону могут входить несколько областей или, наоборот, одна большая область разделена на две-три зоны.

Общегосударственные вторичные сети (вторичные сети общего пользования), как правило, независимы друг от друга. Например, телефонная и телеграфные сети общего пользования имеют отдельные коммутационные станции, свои местные абонентские линии и абонентские установки. Эксплуатация осуществляется разными службами: телефонной и телеграфной. Обособленно существуют также вторичные сети звукового и телевизионного вещания.

Описанная ситуация представлена на рис. 11.1, где изображены фрагменты двух вторичных сетей — телефонной и телеграфной.

Ведомственные вторичные сети представляют собой набор различных вторичных сетей, причем для реализации этих сетей может создаваться своя первичная сеть каналов и трактов (при условии максимального использования трактов и каналов первичной сети ЕАСС).

Развитие цифровых методов передачи и коммутации различных видов информации приводит к необходимости объединения практически всех вторичных сетей в единую (интегральную) сеть.

Потребителей в основном интересуют принципы построения и основные технические характеристики вторичных сетей общего пользования. В идеальном случае потребитель может ничего не знать о структуре и основных технических характеристиках пер-

вичных сетей — это входит в компетенцию Министерства связи или другого ведомства, ответственного за создание ведомственной первичной сети. Для широкого круга потребителей наибольший интерес представляют три типа вторичных сетей: телефонная, телеграфная и сеть передачи данных.

11.3. ТЕЛЕФОННАЯ СЕТЬ ОБЩЕГО ПОЛЬЗОВАНИЯ

Международный опыт и научное прогнозирование потоков различной информации показывают, что в обозримом будущем основным источником информации в сетях связи останутся телефонные переговоры. Поэтому наиболее разветвленной вторичной сетью является общегосударственная автоматическая коммутируемая телефонная сеть, определяющая как структуру, так и пропускную способность первичной сети.

Вся страна разбивается на зоны, границы которых определяются количеством абонентов на территории зоны. Размеры зоны выбираются такими, чтобы в установившемся состоянии в ней было от 1 до 8 млн. абонентов. Границы зоны по возможности привязываются к территориальному делению (области, края, республики). В этом случае кроме удобства административного управления телефонной зоной большая часть телефонного потока (трафика) замыкается внутри зоны и уменьшается необходимость в междугородных каналах. Каждой зоне присваивается семизначная нумерация. При установлении соединения внутри зоны абонент должен набрать семизначный номер, в котором первые две цифры являются внутризональным кодом сельских и городских сетей зоны с емкостью до 100 000 абонентов. Для междугородной связи каждой зоне присваивается трехзначный междугородный (междузоновый) код.

На рис. 11.2 приведена структурная схема автоматической коммутируемой телефонной сети общего пользования (ТФ—ОП).

Внутри зоны сеть построена по радиально-узловому принципу. Абоненты подключаются к оконечным станциям (ОС) (сельским и городским АТС). Оконечные станции соединены между собой через центральные станции (ЦС). Центральные станции каждой зоны соединяются между собой через зонный узел коммутации (ЗУ). В развитой сети при сильном «тяготении» между определенными группами ЦС они могут соединяться между собой через дополнительные узловые станции (УС). В свою очередь УС могут иметь соединения между собой, минуя ЗУ.

Для выхода в другие зоны зонный узел соединяется с автоматической междугородной станцией (АМТС), которая имеет каналы связи с любой другой АМТС страны. Эти каналы передаются из первичной междугородной сети ЕАСС от сетевых узлов (СУ) или сетевых станций (СС). Переключение осуществляется узлами автоматической коммутации (УАК), включаемых между АМТС и ОС или СС, причем УАК территориально могут располагаться или на АМТС или на ОС и СС.

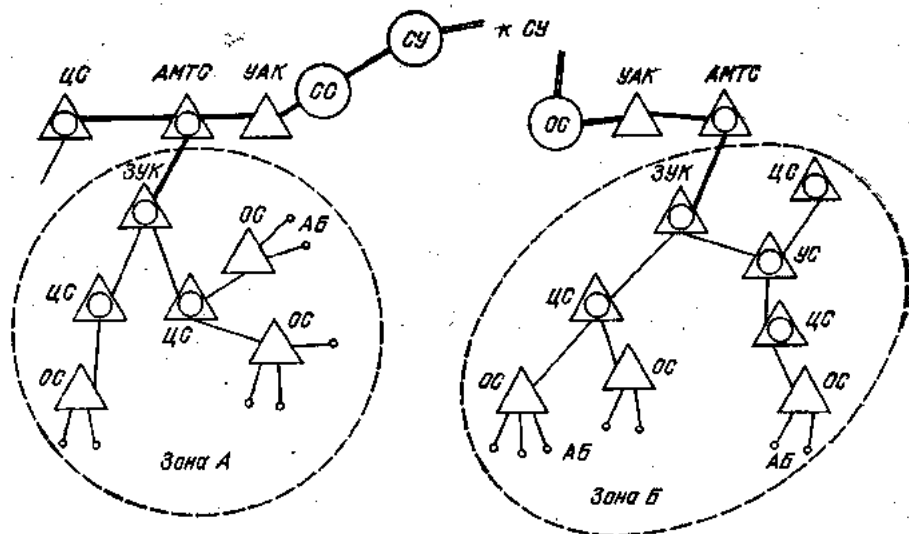


Рис. 11.2. Структурная схема автоматической коммутуруемой телефонной сети:
ЗУ — зонный узел коммутации

Общегосударственная телефонная сеть общего пользования должна удовлетворять потребности всех основных абонентов страны в телефонной связи. Однако на крупных заводах, производственных комбинатах, совхозах и т. д. телефонная связь для производственных нужд часто концентрируется внутри них. Для таких производственных и внутриведомственных телефонных сетей разработаны специальные учрежденческо-производственные автоматические телефонные станции (УПАТС). Эти станции обеспечивают связь между абонентами станции, имеют специальное оборудование для подключения удаленных до нескольких километров абонентов, выхода на общегосударственную сеть и содержат всего несколько соединительных линий, число которых в десятки и сотни раз меньше, чем число включенных абонентов.

Сеть ТФ—ОП может обеспечивать телефонную связь с подвижными объектами с помощью специального радиотелефонного и коммутационного оборудования. Сеть связи с подвижными объектами (автомобилями, поездами, самолетами, судами) состоит из абонентских радиостанций (АС) и центральных радиостанций (ЦРС) на несколько десятков или сотен абонентов; ЦРС подключаются к АТС телефонной сети общего пользования на правах учрежденческой или ведомственной сети.

Общегосударственная коммутуруемая телефонная сеть при установке у абонента соответствующего оконечного оборудования позволяет передавать также факсимильные сообщения и дискретные данные со скоростью от 50 до 1200 бит/с (в пределе до 1800 бит/с).

Передача данных по сети ТФ—ОП осуществляется по телефонному каналу связи в полосе 0,3—3,4 кГц при затухании от абонента до абонента, не превышающем 28,7 дБ на частоте 800 Гц и 40 дБ на частоте 1700 Гц следующими способами:

последовательным со скоростью до 200 бит/с в дуплексном или полудуплексном режиме;

последовательным со скоростью 600 или 1200 бит/с в полудуплексном режиме;

параллельным (многочастотным) со скоростью передачи до 40 знаков/с в полудуплексном режиме.

В отдельных случаях допускается включение аппаратуры передачи данных (АПД) непосредственно в автоматические междугородные телефонные станции по четырехпроводной схеме. При этом может быть осуществлен дуплексный режим работы со скоростью отправления до 1200 бит/с.

При организации первичных и вторичных сетей на их отдельных участках устанавливаются многоканальные системы связи с соответствующей аппаратурой уплотнения, обеспечивающие передачу определенных групп каналов ТЧ. Для получения необходимого качества передачи сигналов в системах с частотным уплотнением по каждому каналу ТЧ следует передавать сигналы средней мощности в одном канале ТЧ не более 32 мкВт (допускается увеличение мощности до 50 мкВт при условии ограничения количества таких каналов в одной системе передачи). Поэтому на входе неуплотненной абонентской линии, представляющей собой, как правило, физическую пару городских телефонных кабелей, мощность сигналов в аппаратуре передачи данных не должна превышать мощность сигнала от обычного телефонного аппарата, т. е. 1 мВт (условный нулевой уровень мощности в технике электросвязи, т. е. 0 дБ). В этом случае на входе ближайшей системы передачи мощность сигнала не превысит 32 мкВт или 15 дБ относительно точки с нулевым относительным уровнем. Выполнение этих требований является обязательным условием при включении АПД в ТФ—ОП.

Нагрузка телефонной сети зависит от количества, времени возникновения и продолжительности телефонных и других сообщений. Сеть ТФ—ОП проектируется с учетом статистики телефонных разговоров различных абонентов.

Для численной оценки нагрузки в телефонной сети пользуются понятием интенсивности. Под *интенсивностью нагрузки* понимается математическое ожидание поступающей нагрузки, отнесенное к единице времени (в телефонии за единицу времени принят час). Поэтому единицей измерения интенсивности нагрузки является Эрланг (1 часо-занятие за 1 ч)¹.

Интенсивность нагрузки изменяется по часам, дням, неделям, сезонам года. Для правильного проектирования сети надо знать

¹ Эта единица названа в честь крупного шведского специалиста по телефонии Эрланга.

кривые распределения нагрузки по этим параметрам. Одним из главных параметров является распределение нагрузки в течение суток. Час с наибольшей интенсивностью нагрузки называется *часом наибольшей нагрузки* (ЧНН).

Каждый абонент дает в среднем нагрузку от 0,06 до 0,15 Эрл (последнее значение характерно для учреждений телефонов). По этим величинам рассчитывается телефонная сеть, в первую очередь ее коммутационные устройства. Современная городская сеть обеспечивает качество обслуживания с потерями не более 2—5 вызовов на каждую точку (при междугородной связи не более 10). Поэтому непереносимое условие при передаче данных по сети ТФ—ОП—ограничение продолжительности сеанса, оно не должно превышать 10 минут в час ЧНН. Часы ЧНН сообщаются абонентам представителями Министерства связи СССР. В среднем основной пик нагрузки приходится на 9—12 часов по местному времени (ЧНН около 11 часов), второй пик наступает от 16 до 19 часов.

11.4. ТЕЛЕГРАФНАЯ СЕТЬ

Телеграфная сеть в СССР построена по трехзвенной радиально-узловой структуре главных, областных и районных телеграфных узлов (см. рис. 11.1). Главные телеграфные узлы (ГТУ) располагаются в Москве, в некоторых столицах союзных республик и в областных центрах, где пересекаются значительные потоки телеграфных сообщений. Эти узлы соединены между собой по принципу «каждый с каждым» и с областными узлами своих зон. Областные телефонные узлы соединяются с ГТУ и со своими районными узлами. При построении телеграфной сети предусматривается возможность включения абонентов непосредственно в ОТУ и ГТУ.

Телеграфная сеть состоит из четырех сетей:

телеграфная сеть общего пользования (ТГ—ОП), по которой передаются телеграммы, принятые в городских отделениях связи и доставляемые адресатам почтовой службой в виде отпечатанного текста на телеграфных бланках;

сеть абонентского телеграфирования (АТ), по которой передаются телеграммы или организуются телеграфные переговоры (диалог) по телеграфному аппарату с помощью телеграфистки по методу «запрос—ответ». Поскольку в настоящее время в этой сети повсеместно применяются телеграфные аппараты со скоростью телеграфирования 50 Бод, эта сеть носит название АТ-50;

международная сеть абонентского телеграфирования «Телекс» для связи с зарубежными странами;

международная сеть телеграфирования «Гентекс», обеспечивающая связь между странами социалистического содружества. Международные сети отделены от внутрисоюзной телеграфной сети, причем каждая из них имеет свои узлы коммутации и свои выделенные международные каналы связи.

Внутрисоюзная телеграфная сеть состоит из двух сетей — ТГ—ОП и АТ-50. Обе эти сети рассчитаны на скорость 50 Бод и для их построения применяются одни и те же телеграфные аппараты и одна и та же аппаратура уплотнения. Раздельный принцип построения этих сетей пока исключает возможность абоненту ТГ—ОП устанавливать связь с абонентом АТ-50, и наоборот.

В настоящее время создается еще одна подсеть телеграфного типа со скоростью отправления до 200 бит/с — ПД-200, предназначенная в основном для передачи данных. Особенностью этой подсети является тесная организационная и техническая увязка с другими телеграфными сетями. Например, коммутационные станции трех сетей АТ-50, ТГ—ОП и ПД—200 объединены конструктивно в единую станцию типа АТ—ПС—ПД.

Телеграфная сеть общего пользования структурно и организационно имеет несколько вариантов.

Система с неавтоматическим транзитом. Сообщение фиксируется в виде буквенного текста, который механически транспортируется к телеграфному аппарату, соединенному с требуемым пунктом. (Сообщение отправляется ручным набором на клавиатуре).

Система реперфораторного переприема с отрывом перфоленты. Сообщение фиксируется на перфоленте и в форме буквенного текста. Отрезок перфоленты с законченным текстом сообщения через узел сортировки транспортируется к аппарату, соединенному с требуемым узлом, к которому сообщение отправляется через транзиттерную приставку. Таким образом, эта система является полуавтоматической с коммутацией сообщений. (Она еще называется системой АТОЛ — автоматической транспортировки оторванной ленты.)

Система реперфораторного переприема без отрыва перфоленты. Сообщение строится по установленному «формату» с указанием в начале его адреса телеграммы. Распределение телеграмм по маршрутам производится автоматически с помощью электромеханических устройств. Это автоматическая система с коммутацией сообщений.

11.5. ОБЩЕГОСУДАРСТВЕННАЯ СИСТЕМА ПЕРЕДАЧИ ДАННЫХ (ОГСПД)

Общегосударственная система передачи данных (ОГСПД) представляет собой организационно-техническую совокупность средств связи, входящих в ЕАСС и предназначенных для передачи и обмена данными с широким кругом потребителей. Эта система должна обеспечивать связь между собой вычислительных центров (ВЦ), вычислительных центров с абонентскими пунктами (АП), а также АП между собой. В ОГСПД абонентам ВЦ и АП обеспечивается одноадресная, многоадресная и циркулярная передача сообщений, обмен данными в режиме «запрос—ответ» и в режиме «диалог», различные режимы передачи (дуплекс, полудуплекс, симплекс), работа с различными скоростями, кодами, методами защиты от ошибок и др.

Эта ОГСПД создается на базе технических средств ЕАСС и является ее составной частью. Создание ведомственных сетей передачи данных допускается только в тех случаях, когда эксплуатационно-технические характеристики ОГСПД не удовлетворяют специфическим требованиям этих ведомств, например, при передаче данных в АСУ—ТП таких сложных производственных комплексов, как нефтепроводы, нефтехимические производства, непрерывные металлургические процессы, системы управления движением самолетов и др.

В ОГСПД используются все принципы доставки информации по сети электросвязи: некоммутируемые и коммутируемые каналы, центры коммутации сообщений и пакетов. Сеть строится по территориальному радиально-узловому принципу, отражающему территориальное деление страны и соответствующему структуре первичной и вторичных сетей ЕАСС.

В состав ОГСПД входят:

специализированная сеть передачи данных телеграфного типа с коммутацией каналов (ПД—КК) при скоростях от 50 до 2400 бит/с;

специализированная сеть передачи данных с коммутацией сообщений (ПД—КС) при скоростях от 50 до 4800 бит/с;

коммутируемая телефонная сеть общего пользования (ТФ—ОП) для передачи данных;

коммутируемая сеть абонентского телеграфирования АТ-50 для передачи данных со скоростью 50 бит/с 5-элементным кодом МТК-2 при 7,5 элементах в знаке и при стартопном методе передачи и нагрузке не более 0,2 Эрл (12 мин в ЧНН);

выделенная интегральная сеть телефонной связи и передачи данных аналого-цифрового типа ВИС—АЦ с коммутацией каналов в крупных городах для обеспечения потребностей в передаче данных со скоростями до 4800 бит/с и для оперативной телефонной связи. Эти сети будут создаваться на базе электронных АТС и линий передачи с импульсно-кодовой модуляцией. Основной особенностью этой сети является низкое значение вероятности ошибок (10^{-6} даже без устройств защиты от ошибок);

сеть некоммутируемых каналов и трактов «телеграфного типа» для скоростей 50, 100, 200, 600, 1200 и 2400 бит/с;

сеть некоммутируемых каналов ТЧ с полосой 300—3400 Гц для скоростей 600, 1200, 2400, 4800 и 9600 бит/с;

сеть групповых трактов (широкополосных каналов) на полосе 60,6—107,7 кГц и скорости 24, 28, 96 кбит/с и на полосы 312,3—551,4 кГц и скорости 120, 240 и 480 кбит/с.

В некоммутируемых сетях коэффициент ошибки по битам на расстоянии до 5 тыс. км без устройств защиты от ошибок не будет превышать 10^{-3} , а при наличии этого устройства — 10^{-6} .

При проектировании сетей предусматривается сопряжение сети ПД—КК с сетью ПД—КС, АТ—50 и ВИС—АЦ. Сеть ПД—КС сопрягается с сетью АТ-50, ТФ—ОП и ПД—КК.

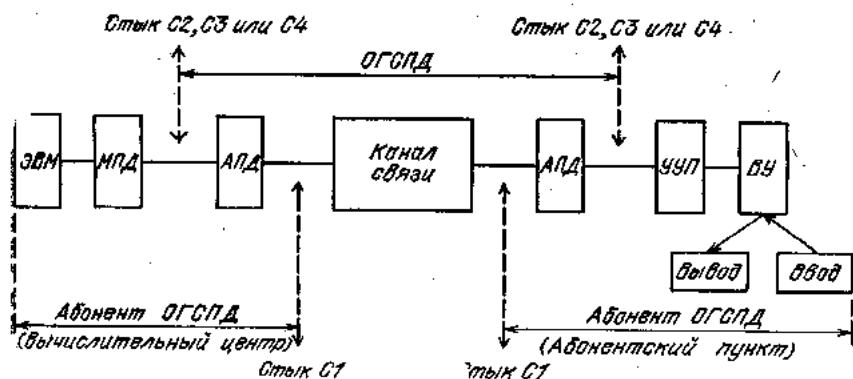


Рис. 11.3. Структурная схема ОГСРД

Обобщенная структурная схема ОГСРД приведена на рис. 11.3. В ее состав входят аппаратура передачи данных АПД (номенклатура АПД должна обеспечивать возможность работы в составе любой сети, входящей в ОГСРД); каналы связи, включая абонентские линии и коммутационное оборудование; контрольно-измерительное оборудование и устройства для управления.

Аппаратура передачи данных согласно ГОСТ 17657—72 может содержать устройство преобразования сигнала, обеспечивающее прямое и обратное преобразование сигнала в вид, пригодный для передачи по каналу связи, устройство защиты от ошибок (УЗО) и вызывные устройства (ВУ).

В состав ОГСРД не входят средства технологической обработки данных (оконечное оборудование данных): ЭВМ, оборудование сопряжения ЭВМ с каналами передачи данных, мультиплексоры передачи данных (МПД), устройства ввода — вывода данных на промежуточные носители, устройства печати данных.

Границей между ОГСРД и средствами обработки данных является стык (С) аппаратуры передачи данных с устройством сопряжения абонентского пункта. Сопряжение средств ОГСРД с оконечным оборудованием данных ВЦ и АП может быть непосредственно электрическим или через промежуточный носитель. При электрическом стыке С2, С3, С4 его параметры должны соответствовать ГОСТ 18145—72 и 18146—72. Сопряжение АПД с каналами связи осуществляется также по стандартному стыку С1: с телеграфными каналами до 200 бит/с (С1—ТГ), с каналом ТЧ (С1—ТЧ) и с неуплотненной физической линией (С1—ФЛ).

11.6. ЛИНИИ ПЕРЕДАЧИ И ТИПОВАЯ СТРУКТУРА ТРАКТА ЕАСС

В зависимости от исторически сложившейся ситуации и технико-экономической целесообразности передачи сигналов на различных участках сети электросвязи может осуществляться по про-

водам, кабелям, волоконным линиям, волноводам, с помощью ре-трансляции (в том числе через ИСЗ) или путем излучения в пространство.

Рабочий диапазон частот системы передачи определяется как свойствами используемых линий, так и достигнутыми параметрами приемопередающей аппаратуры. Высшая частота диапазона определяется, в частности, затуханием в линии, предельной мощностью передатчика и требуемым отношением сигнал/помеха в точке приема. Отношения излученной (переданной в линию) и принятой мощностей лежат обычно в пределах 50—70 дБ. Зная величину затухания в различных линиях, можно определить предельное расстояние связи. Как правило, оно оказывается значительно меньше требуемого. Поэтому линия разделяется на усилительные участки, в точках соединения которых устанавливаются усилители или регенераторы (восстановители) сигналов.

Так, для воздушной пары проводов в рабочем диапазоне от 10 до 150 кГц длина усилительного участка составляет соответственно от 300 до 50 км. Длина усилительных участков в некоторых других линиях приведена в табл. 11.1. Приведенные данные показывают, что длинные линии связи должны содержать до нескольких сотен усилительных станций.

Кроме собственно приемопередающего оборудования и усилительных станций в состав линии входят устройства телеуправления и электропитания аппаратуры.

Таблица 11.1

Вид линии	Диапазон частот	Длина усилительного участка, км	Примечание
Симметричный ВЧ кабель	12—500 кГц	30—10	
Коаксиальный кабель	6—60 МГц	6—1,5	
Волновод диаметром 50—70 мм	5—10 ГГц	50—20	при волне типа Н ₀₁
Радiorелейная наземная линия	2—8 ГГц	60—40	в зависимости от рельефа местности
Световод	световой	20—5	при затухании 2—5 дБ/км

Наряду с задачей передачи групп сигналов между оконечными пунктами линии, в сети возникает задача выделения и ввода «пучков» сигналов. Так, областная магистральная линия должна обеспечить выделение и ввод сигналов в каждом районном центре и других узлах связи, через которые она проходит. Аппаратура выделения совмещается с одной из усилительных станций и располагается в специальных помещениях, где производится ее

периодическое обслуживание. Промежуточные усилительные станции в кабельной линии связи, как правило, выполняются необслуживаемыми и располагаются под землей.

При использовании радионезлучения и ретрансляции сигналов через ИСЗ линии радиосвязи, работающие обычно в диапазоне от 10 до 100 м, не имеют, как правило, промежуточных станций до расстояний в несколько тысяч километров. При этом применяются передатчики мощностью до 100 кВт и сложные антенные системы. На одной несущей частоте из-за помех и селективных замираний удается передать с приемлемыми ошибками не более четырех стандартных телефонных каналов, по каждому из которых можно передать около шести телеграфных сообщений со скоростью до 200 Бод (напомним, что по кабельной линии в стандартном телефонном канале передается до нескольких десятков телеграфных сообщений со скоростью 200 Бод).

Технически (а чаще всего и экономически) наиболее эффективными оказываются системы с использованием ИСЗ и ретранслятор с *многостанционным доступом* наземных станций. Такой ретранслятор на ИСЗ позволяет образовать коммутируемую сеть связи, обслуживающую десятки узлов связи или оконечных станций. В ближайшем будущем один ИСЗ будет содержать до 24 ретрансляторов, каждый с пропускной способностью до 1000 стандартных каналов.

Весьма перспективными являются системы с использованием ИСЗ для передачи телевизионных изображений. Так, отечественная система «Экран» с помощью стационарного спутника обеспечивает передачи телевизионных сигналов в районы Западной и Восточной Сибири. Прием сигналов производится сравнительно небольшими неподвижными антеннами.

В соответствии с изложенными особенностями типовая структурная схема системы передачи показана на рис. 11.4,а. Ее укрупненными частями являются *аппаратура уплотнения* и *аппаратура линейного тракта*.

В ЕАСС принят групповой иерархический принцип уплотнения стандартных телефонных каналов (каналов ТЧ). *Аппаратура индивидуального преобразования* (АИП) объединяет каналы ТЧ в первичные группы. В аналоговых системах для этого применяются 12-канальные, а в цифровых — 30-канальные (иногда разделенные на две 15-канальные подгруппы) группы.

Аппаратура группового преобразования (АГП) объединяет первичные группы в группы высших порядков в соответствии со строго установленной иерархией. В аналоговых системах установлены: 12-канальная (первичная) группа, 60-канальная (вторичная, объединяющая, следовательно, пять первичных групп), 300-канальная (третичная) группа, 900-канальная (четверичная) и 1800-канальная (пятиричная) группа. В цифровых системах имеются: субпервичная (15-канальная), первичная (30-канальная), вторичная (120-канальная), третичная (480-канальная) и четверичная (1920-канальная).

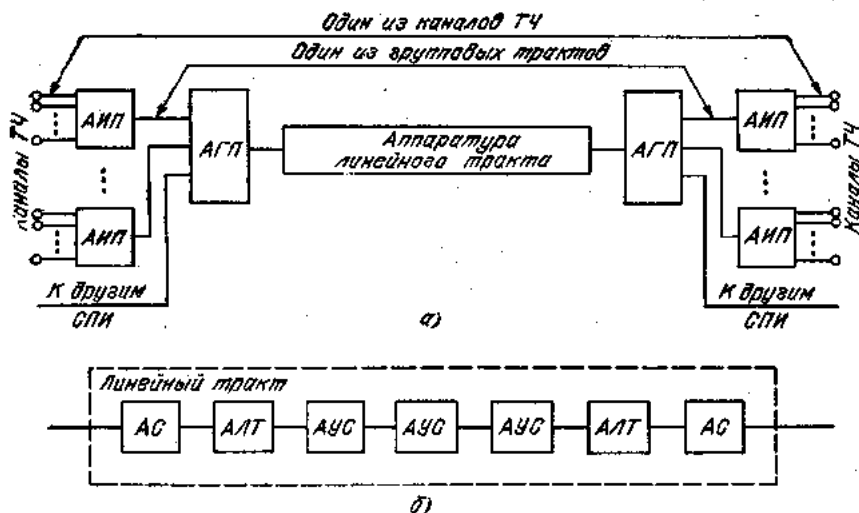


Рис. 11.4. Обобщенные структурные системы передачи в ЕАСС (а) и линейного тракта (б)

Аппаратура уплотнения АИП и АГП унифицирована и может быть использована для образования групповых трактов на первичных и вторичных сетях при использовании любых линий связи.

Рассмотрим далее структуру линейного тракта (рис. 11.4,б). В его состав входят: аппаратура сопряжения (АС), оконечная аппаратура линейного тракта (АЛТ), аппаратура усилительных станций (АУС) и собственно линии (направляющие системы), не указанные на рисунке.

Аппаратура сопряжения производит согласование выходных сигналов АГП с входными сигналами АЛТ по спектру, уровням и, возможно, другим параметрам.

Аппаратура линейного тракта (АЛТ) содержит приемо-передающие устройства, устройства телеуправления и сигнализации для всего оборудования линейного тракта (включая электропитание усилительных станций).

Аппаратура усилительных станций (АУС) содержит необслуживаемые усилители, устройства сигнализации и электропитания.

По условиям технико-экономической целесообразности длина линейных трактов в первичных и зональных сетях колеблется от 100 до 300 км. На магистральных линиях она может достигать 1500 км. Для передачи сигналов на большие расстояния строится цепочка из линейных трактов. Только системы с использованием ИСЗ позволяют перекрывать большие расстояния с помощью одного линейного тракта.

11.7. ПОСТРОЕНИЕ ГРУПП УПЛОТНЕНИЯ И ЛИНЕЙНЫХ ТРАКТОВ ПРИ АНАЛОГОВЫХ (НЕПРЕРЫВНЫХ) СИГНАЛАХ

При объединении стандартных телефонных каналов (300—3400 Гц) в типовые группы применяется метод однополосной модуляции с полностью подавленной несущей каждого канала (ОПМ—ПН). При этом более экономично используется полоса пропускания линейного тракта системы передачи, а подавление несущего колебания облегчает задачу обеспечения работы усилительного оборудования с допустимыми искажениями, в частности ослабляет нелинейные эффекты взаимодействия между отдельными сигналами в сложном многоканальном сигнале. Одновременно, однако, усложняется процесс демодуляции и разделения каналов с допустимыми искажениями формы сообщений в отдельных каналах. Опыт показывает, что при передаче речи между людьми допустимое расхождение частот несущих передатчика и местного гетеродина, восстанавливающего несущие колебания, могут доходить до 20 Гц без существенной потери разборчивости речи. При передаче же телеграфных сигналов (и сигналов передачи данных) эти расхождения не должны превышать 1—2 Гц. Для поддержания в сети ЕАСС требуемой стабильности колебаний местных генераторов применяется система передачи эталонных частот.

На рис. 11.5,а—г схематически показаны спектральные диаграммы первичной (ПГ), вторичной (ВГ), третичной (ТГ) и четверичной (ЧГ) групп уплотнения. Так, из рис. 11.5,а видно, что первичная группа из 12 стандартных каналов занимает диапазон частот от 60 до 108 кГц (см. также п. 9.6). При этом несущие частоты стандартных каналов (подавляемые в оконечных каскадах передатчика) устанавливаются равными

$$f_n^I = 108 - 4(n - 1),$$

где n — номер канала ($n = 1, 2, \dots, 12$). При изображении спектров стандартных каналов принято, что плотность мощности линейно возрастает с ростом частот до 3400 Гц. На каждый канал отведено 4 кГц с учетом защитных интервалов.

При образовании спектра вторичной группы из спектров пяти первичных групп (т. е. для 60 стандартных каналов) использован диапазон от 312 до 552 кГц (рис. 11.5,б). Несущие частоты колебаний вторичной группы приняты равными

$$f_n^{II} = 420 + 48(n - 1),$$

причем используют нижние боковые полосы. (Возможен также инверсный способ образования группы, когда $f_n^{II} = 252 + 48(n - 1)$ кГц с использованием верхних боковых полос.)

На рис. 11.5,в и г показано образование спектров третичной и четверичной группы.

В аналоговых системах передачи необходимо весьма точно поддерживать уровни сигналов, так как при их возрастании про-

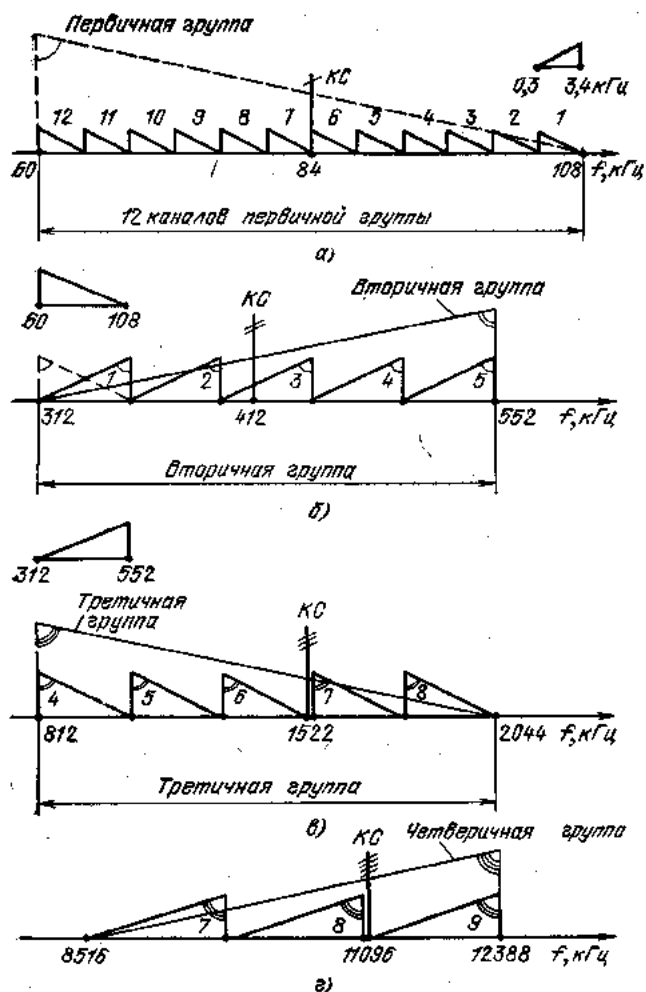


Рис. 11.5. Спектральные диаграммы различных групп уплотнения сигналов в ЕАСС

исходит перегрузка усилителей и недопустимые искажения формы сигналов. Для контроля уровня в каждую группу вводится контрольный сигнал (КС) — узкополосное гармоническое колебание. Частоты КС указаны на рис. 11.5. Контрольные сигналы используются для автоматического поддержания сигналов многоканальной группы.

Кроме образования соответствующих многоканальных сигналов из первичных сигналов, в каждой группе можно образовать широкополосные каналы с полосой от 60 до 108 кГц, от 312 до 552 кГц и т. д. Их можно использовать для передачи широкополосных сообщений (видеотелефонных, телевизионных, высокосо-

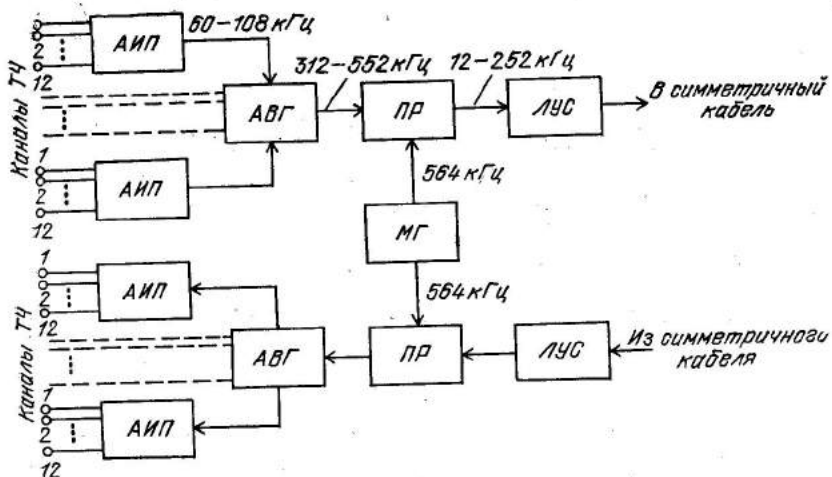


Рис. 11.6. Обобщенная структурная схема системы передачи К-60

ростной передачи данных). Для образования широкополосных каналов используется отдельное каналообразующее оборудование. Абонентские пункты потребителей должны, очевидно, быть оснащены соответствующим оконечным оборудованием.

Образованные аппаратурой уплотнения первичную и последующие группы необходимо перенести в рабочий диапазон линейного тракта (рис. 11.4,а) системы передачи с помощью аппаратуры сопряжения (рис. 11.4,б).

Так, в симметричных высокочастотных кабелях целесообразно с точки зрения технико-экономических показателей использовать диапазон частот от 12 до 252 кГц, в котором укладывается полоса вторичной 60-канальной группы. Она переносится в диапазон 12—252 кГц с помощью специального преобразовательного оборудования. Структурная схема соответствующей аппаратуры К-60 представлена на рис. 11.6. С помощью аппаратуры индивидуального преобразования (АИП) (ср. с рис. 11.4) образуются первичные группы в диапазоне 60—108 кГц. С помощью аппаратуры образования вторичной группы (АВГ) образуется вторичная группа в диапазоне 312÷552 кГц, которая с помощью преобразователя (ПР) и местного генератора (МГ) переносится в диапазон 12—252 кГц, после чего через линейный усилитель (ЛУС) поступает в симметричный кабель (линейный тракт). На приемной стороне преобразование спектров производится в обратном порядке.

Система К-60 позволяет передавать сообщения с требуемым качеством на расстояния до 12 000 км (при длине усилительного участка около 16 км). Аналогичным образом строятся более «емкие» системы передачи. Так, на рис. 11.7 показаны структуры двух типов линейных спектров, которые можно образовать с по-

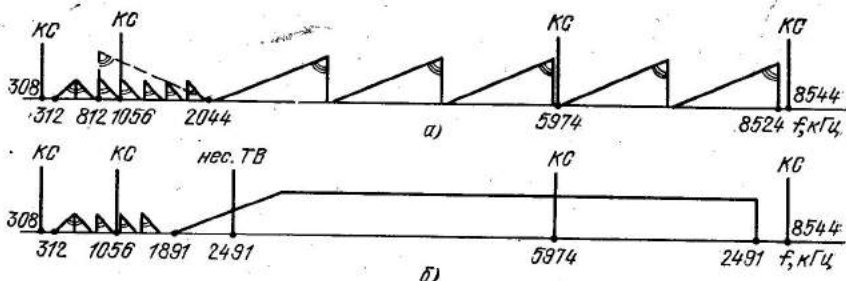


Рис. 11.7. Линейный спектр аппаратуры уплотнения К-1920

мощью аппаратуры К-1920. При формировании спектра первого типа (рис. 11.7,а) в диапазоне от 312 до 8524 кГц в одном коаксиальном кабеле формируется 1920 стандартных телефонных каналов, а при создании спектра второго типа (рис. 11.7,б) — 300 телефонных каналов и один канал телевидения. Кроме сигналов, несущих сообщения, в линейный спектр вводятся контрольные сигналы (КС) для автоматической регулировки усиления в линии передачи. При использовании широко распространенных 4- и 8-парных кабелей система К-1920 позволяет образовать соответственно 3840 и 7680 стандартных телефонных каналов по 4-проводной схеме.

Следующими по «емкости» систем передачи являются устройства К-3600 и К-10800. С помощью последней системы удастся обеспечить предельную в настоящее время пропускную способность 10 000 стандартных телефонных каналов при использовании одного 20-парного коаксиального кабеля. Ясно, что при расширении рабочего диапазона частот уменьшается длина усилительного участка (так, для аппаратуры К-60 она составляет уже только 6 км).

При построении сети ЕАСС широко используются радиорелейные линии (наземные и действующие через ретрансляторы ИСЗ). При этом линейный сигнал, например в аппаратуре К-1920, подается на модулятор передающего устройства соответствующей ретрансляционной системы и обычно при этом используется частотная модуляция. На оконечных приемных станциях производится демодуляция ЧМ сигналов и передача их (в случае необходимости) в кабельную линию.

В сети ЕАСС применяется единая иерархия построения групп каналов, линейных спектров, а также аппаратуры контроля и резервирования.

11.8. О ПОСТРОЕНИИ ЦИФРОВЫХ СИСТЕМ ПЕРЕДАЧИ ИНФОРМАЦИИ В ЕАСС

В настоящее время в цифровых системах, входящих в состав ЕАСС, используется импульсно-кодовая модуляция (ИКМ). Эти системы применяются для передачи телефонных сообщений в ре-

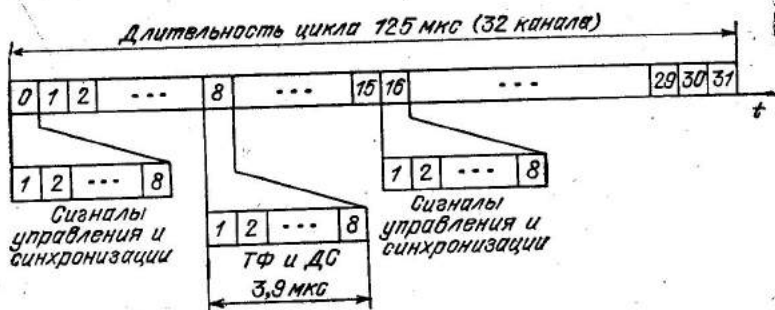


Рис. 11.8. Структура интервала временного уплотнения аппаратуры ИКМ-30

жиме временного уплотнения каналов. По ряду технико-экономических соображений и в соответствии с теоремой Котельникова частота дискретизации по времени в каждом канале выбрана равной 8 кГц. Полученные выборочные значения подвергаются операциям компрессирования (уменьшения динамического диапазона с последующим восстановлением в точке приема), квантования по уровням и бинарного кодирования (получения собственно сигналов ИКМ).

В сети ЕАСС принято квантование компрессированных выборок на 256 уровней, что соответствует 8-значному бинарному коду. Таким образом, на выходе кодирующего устройства образуется последовательность бинарных посылок, следующих со скоростью 64 кбит/с группами по 8 посылки, следующих в соответствии с выборками с частотой 8 кГц. Эта последовательность образует основной цифровой канал, параметры которого регламентированы в международном масштабе.

Структура интервала временного уплотнения цифровых каналов приведена на рис. 11.8 применительно к аппаратуре ИКМ-30. Аналогично строятся временные структуры в аппаратуре ИКМ-120, ИКМ-480, ИКМ-1920.

Временной интервал длительностью 125 мкс (соответствующий частоте 8 кГц) с помощью аппаратуры ИКМ-30 разделяется на 32 канальных интервала длительностью около 3,9 мкс, 30 из которых предназначены для передачи кодированных выборок 30 независимых сообщений (в двоичном восьмизначном коде), а два (0 и 16) — для передачи служебных сообщений (сигналов управления, взаимодействия и синхронизации). Длительность каждой посылки внутри канальных интервалов около 0,4 мкс (тактовая частота 2048 кГц).

При кодировании выборочных значений независимых сообщений используются два метода. При первом методе выборки каждого сообщения кодируются (и декодируются в точке приема) индивидуальными устройствами кодирования и декодирования (кодеками). Устройства кодирования и декодирования располагаются при этом у абонентов. При втором методе сигналы объеди-

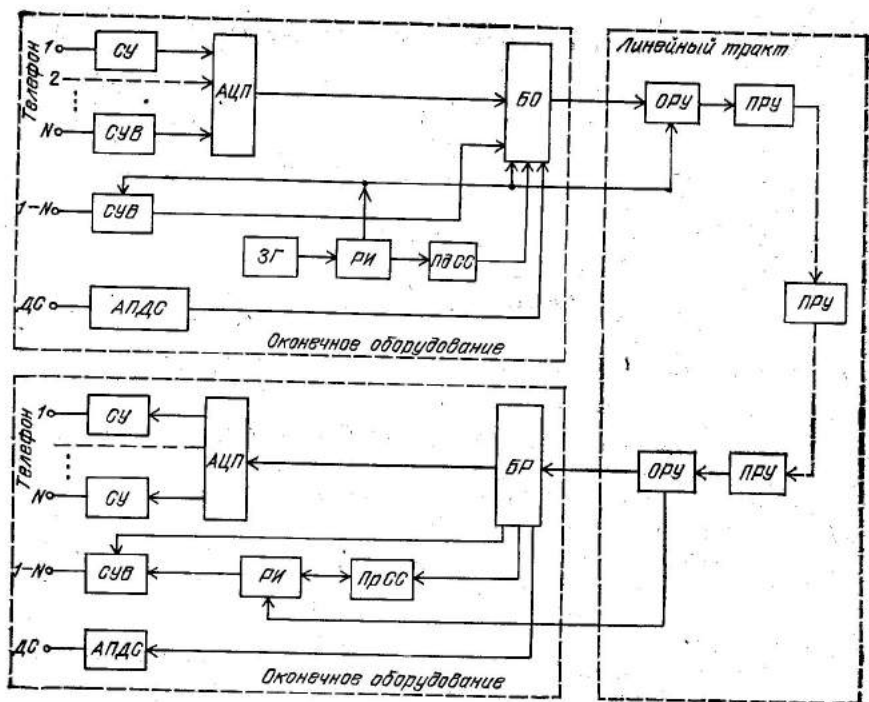


Рис. 11.9. Структурная схема цифровой системы передачи с ИКМ

няются в единый поток выборок из всех сообщений, который затем кодируется одним устройством. В сети ЕАСС применяется второй метод как экономически более выгодный.

Структурная схема цифровой системы передачи с ИКМ на N каналов показана на рис. 11.9. На ней предусмотрена передача непрерывных (телефонных) сообщений из 1, 2, ..., N -го каналов, передача данных (дискретных сообщений) через аппаратуру передачи АПДС (при этом заняты соответствующие каналы из N), передача сигналов управления и взаимодействия через аппаратуру СУВ (к последним относятся сигналы типа «отбой», «занято» и др.) и сигналов синхронизации.

Непрерывные сообщения 1, 2, ..., N по абонентским линиям поступают через согласующие устройства (СУ) на входы аналого-цифрового преобразователя (АЦП), где подвергаются указанным ранее операциям (дискретизации, компандирования и др.). В блоке объединения (БО) кодированные сигналы всех N каналов объединяются по времени с сигналами управления и взаимодействия (СУВ), вырабатываемыми в блоке СУВ и сигналами синхронизации, вырабатываемыми в блоке ПДСС.

Режим временного уплотнения реализуется с помощью высокостабильного задающего генератора (ЗГ) и распределителя импульсов (РИ), выполняющего роль делителя частоты. С выхода

БО последовательность бинарных посылок через оконечный регенерационный усилитель (ОРУ) поступает в линейный тракт. (В ОРУ могут производиться некоторые дополнительные операции, например преобразование униполярной последовательности в биполярную.)

На приемной стороне системы передачи импульсная последовательность поступает на приемный ОРУ, а затем на блок разделения сигналов (БР), блок распределения импульсов (РИ) и приемник синхросигналов (ПрСС). С помощью этих блоков производится разделение сигналов на основные сигналы, несущие переданные сообщения, и сигналы управления и взаимодействия. Сигналы передачи данных выводятся через аппаратуру АПДС.

Дополнительные сведения об особенностях построения современных систем и сетей можно найти в [2, 3].

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Назовите типовые каналы ЕАСС и дайте их характеристику.
2. Что понимают под первичными и вторичными сетями ЕАСС?
3. Перечислите основные особенности первичных и вторичных сетей ЕАСС.
4. Укажите основные особенности телефонной сети общего пользования.
5. Как осуществляется передача данных в телефонной сети общего пользования?
6. Укажите основные особенности телеграфной сети.
7. Перечислите основные особенности общегосударственной системы передачи данных.
8. Назовите состав типового тракта ЕАСС и объясните назначение входящих в него блоков.
9. Как осуществляется уплотнение каналов и построение линейных трактов в аналоговых системах ЕАСС?
10. Как осуществляется цифровая передача по сети ЕАСС?

ЗАКЛЮЧЕНИЕ

Теория и техника передачи информации достигли в настоящее время весьма высокого уровня и характеризуются следующими особенностями: существенным развитием математических методов исследований, широким применением вычислительной техники, использованием новых физических эффектов и достижений современной технологии.

Проектирование, разработка и изготовление радиотехнических систем передачи информации становятся очень сложными и дорогостоящими, поэтому особое значение приобретают вопросы технико-экономического обоснования целесообразности тех или иных вариантов создаваемых систем.

Системы передачи информации, имевшие ранее относительно самостоятельное значение, все чаще начинают применяться в составе различных сетей связи, включая и единую сеть связи страны. Эта особенность заставляет уделять большое внимание вопросам системного подхода к проектированию и созданию современных систем передачи информации¹.

Остановимся кратко на некоторых направлениях развития систем передачи информации и входящих в них устройств.

Основными тенденциями развития радиопередающих устройств систем являются: повышение генерируемой мощности и стабильности частоты, увеличение КПД устройств, применение малогабаритных и эффективных электронных приборов, использование новых видов сигналов, предварительное преобразование передаваемых сообщений с целью сокращения их избыточности. В связи с возрастанием требований к пропускной способности линий связи важное значение приобретают вопросы разработки методов и способов генерирования высокостабильных колебаний в миллиметровом и световом диапазонах волн.

Основными тенденциями развития радиоприемных устройств систем являются: использование алгоритмов работы, близких к оптимальным, повышение достоверности приема сообщений, применение различных методов ослабления влияния неидеальности характеристик среды распространения сигналов и аппаратуры систем, снижение уровня внутренних шумов приемников.

В последние годы благодаря достижениям микроэлектроники резко возросла роль современной элементной базы. Микроэлектронные устройства высокого уровня интеграции (БИС, СБИС, микропроцессоры) позволяют создавать компактную надежную аппаратуру с очень высокими показателями качества работы.

Применение современной элементной базы с ее широкими возможностями реализации самых сложных алгоритмов обусловило

¹ См. напр., книгу Л. С. Гуткина «Современная радиоэлектроника и ее проблемы». Изд. 2-е. — М.: Сов. радио, 1983. — 193 с.

быстрое развитие цифровых методов передачи разнообразной информации, включая речь, музыку, телевизионные изображения и др. Такие методы позволяют не только существенно повысить показатели работы систем, но и значительно облегчают построение сетей связи.

Сеть связи является очень сложной системой с большим количеством входящих в ее состав устройств, обладающих определенной надежностью работы. Необходимость быстрого и безошибочного определения места повреждения в сложных системах привела к появлению и развитию нового важного направления — автоматической диагностики сложных систем.

Коммунистическая партия и Советское правительство поставили перед учеными и инженерами ряд важных и ответственных задач в области электросвязи. Многие из этих задач долговременные, и успех их решения зависит от качества подготовки молодых специалистов. Авторы надеются, что данное учебное пособие будет полезным в такой подготовке.

СПИСОК ЛИТЕРАТУРЫ

К г л. 1

1. Немировский А. С., Рыжков Е. В. Системы связи и радиорелейные линии. — М.: Связь, 1980. — 432 с.
2. Радиоэлектроника и связь в народном хозяйстве/Под ред. Г. М. Крылова. — М.: Связь, 1980. — 120 с.
3. Тимищенко М. Г. Радиорелейные системы передачи прямой видимости. — М.: Радио и связь, 1982.

К г л. 2

1. Хургин Я. И., Яковлев В. П. Финитные функции в физике и технике. — М.: Наука, 1971. — 408 с.
2. Возенкрафт Дж., Джекобс И. Теоретические основы техники связи: Пер. с англ./Под ред. Р. Л. Добрушина. — М.: Мир, 1969. — 640 с.
3. Кэтермоул К. В. Принципы импульсно-кодовой модуляции: Пер. с англ./Под ред. В. В. Маркова. — М.: Связь, 1974. — 408 с.
4. Гоноровский И. С. Радиотехнические цепи и сигналы. — М.: Сов. радио, 1978. — 608 с.
5. Хармут Х. Теория секвентного анализа: Пер. с англ. — М.: Мир, 1980. — 574 с.
6. Дядюнов Е. Г., Сенин А. И. Ортогональные и квазиортогональные сигналы. — М.: Связь, 1977. — 224 с.
7. Котельников В. А. Теория потенциальной помехоустойчивости. — М.: Госэнергоиздат, 1966. — 152 с.

К г л. 3

1. Возенкрафт Дж., Джекобс И. Теоретические основы техники связи: Пер. с англ./Под ред. Р. Л. Добрушина. — М.: Мир, 1969. — 640 с.
2. Гуткин Л. С. Оптимизация радиоэлектронных устройств. — М.: Сов. радио, 1975. — 366 с.
3. Котельников В. А. Теория потенциальной помехоустойчивости. — М.: Госэнергоиздат, 1956. — 152 с.
4. Шеннон К. Работы по теории информации и кибернетике: Пер. с англ./Под ред. Р. Л. Добрушина и О. Б. Лупанова. — М.: ИЛ, 1963. — 829 с.
5. Галлагер Р. Теория информации и надежная связь: Пер. с англ./Под ред. М. С. Пинскера и Б. С. Цыбакова. — М.: Сов. радио, 1974. — 720 с.
6. Пенин П. И. Системы передачи цифровой информации. — М.: Сов. радио, 1976. — 364 с.
7. Ван Трис Г. Теория обнаружения, оценок и модуляции: Пер. с англ./Под ред. В. Т. Горяинова. — М.: Сов. радио, 1975. Т. 2. — 343 с.
8. Тихонов В. И., Кузьман Н. К. Нелинейная фильтрация и квазикогерентный прием сигналов. — М.: Сов. радио, 1975. — 704 с.

К г л. 4

1. Борисов Ю. П., Пенин П. И. Основы многоканальной передачи информации. — М.: Связь, 1967. — 436 с.
2. Беккенбах Э., Беллман Р. Неравенства: Пер. с англ./Под ред. В. И. Левина. — М.: Мир, 1965. — 273 с.
3. Верещагин Е. М., Никитенко Ю. Г. Частотная и фазовая модуляция в технике связи. — М.: Связь, 1974. — 224 с.
4. Тепляков И. М., Калашников И. Д., Рошин Б. В. Радиолинии космических систем передачи информации. — М.: Сов. радио, 1975. — 400 с.
5. Дядюнов Н. Г., Сенин А. И. Ортогональные и квазиортогональные сигналы. — М.: Связь, 1977. — 222 с.

К г л. 5

1. Котельников В. А. Теория потенциальной помехоустойчивости. — М.: Госэнергоиздат, 1956. — 152 с.
2. Возенкрафт Дж., Джекобс И. Теоретические основы техники связи: Пер. с англ./Под ред. Р. Л. Добрушина. — М.: Мир, 1969. — 640 с.
3. Гуткин Л. С. Теория оптимальных методов приема при флуктуационных помехах. — М.: Сов. радио, 1972. — 488 с.
4. Ван Трис Г. Теория обнаружения, оценок и модуляции: Пер. с англ./Под ред. В. Т. Горяинова. — М.: Сов. радио, 1975. Т. 2. — 343 с.
5. Кантор Л. Я., Дорофеев В. М. Помехоустойчивость приема ЧМ сигналов. — М.: Связь, 1977. — 336 с.
6. Клэппер Дж., Фрэнкл Дж. Системы фазовой и частотной автоподстройки частоты: Пер. с англ./Под ред. А. Ф. Фомина. — М.: Энергия, 1977. — 440 с.
7. Борисов Ю. П., Пенин П. И. Основы многоканальной передачи информации. — М.: Связь, 1967. — 435 с.
8. Немировский А. С., Рыжков Е. В. Системы связи и радиорелейные линии. — М.: Связь, 1980. — 432 с.
9. Современные линейные интегральные микросхемы и их применение: Пер. с англ./Под ред. М. В. Гальперина. — М.: Энергия, 1980. — 273 с.
10. Алексенко А. Г., Коломбет Е. А., Стародуб Г. И. Применение прецизионных аналоговых ИС. — М.: Радио и связь, 1981. — 224 с.
11. Справочник по интегральным микросхемам. — 2-е изд., перераб. и доп. — М.: Энергия, 1980. — 816 с.
12. Аналоговые интегральные микросхемы: Справочник. — М.: Радио и связь, 1981. — 160 с.

К г л. 6

1. Справочник по теоретическим основам радиоэлектроники. — М.: Энергия, 1977, т. 1 — 471 с.
2. Возенкрафт Дж., Джекобс И. Теоретические основы техники связи: Пер. с англ./Под ред. Р. Л. Добрушина. — М.: Мир, 1969. — 640 с.
3. Филиппов Л. И. Теория передачи дискретных сигналов. — М.: Высшая школа, 1981. — 175 с.

К г л. 7

1. Гоноровский И. С. Радиотехнические цепи и сигналы. — М.: Сов. радио, 1977. — 606 с.
2. Возенкрафт Дж., Джекобс И. Теоретические основы техники связи: М.: Мир, 1969. — 640 с.
3. Пенин П. И. Системы передачи цифровой информации. — М.: Сов. радио, 1976. — 364 с.
4. Петрович Н. Т. Передача дискретной информации в каналах с фазовой манипуляцией. — М.: Сов. радио, 1965. — 263 с.
5. Филиппов Л. И. Теория передачи дискретных сигналов. — М.: Высшая школа, 1981. — 175 с.

К г л. 8

1. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки: Пер. с англ./Под ред. Р. Л. Добрушина и С. И. Самойленко. — М.: Мир, 1976. — 594 с.
2. Касами Т., Токура Н., Ивадари Е., Инагаки Я. Теория кодирования: Пер. с японского/Под ред. Б. С. Цыбакова и С. И. Гельфанда. — М.: Мир, 1978. — 576 с.
3. Нейфах А. Э. Сверточные коды для передачи дискретной информации. — М.: Наука, 1979. — 221 с.
4. Харкевич А. А. Борьба с помехами. — М.: Наука, 1965. — 274 с.
5. Шварцман В. О., Емельянов Г. А. Теория передачи дискретной информации. — М.: Связь, 1979. — 424 с.
6. Финк Л. М. Теория передачи дискретных сообщений. — М.: Сов. радио, 1970. — 726 с.

К г л. 9

1. Шляпоберский В. И. Основы техники передачи дискретных сообщений. — М.: Связь, 1973. — 480 с.
2. Емельянов Г. А., Шварцман В. О. Передача дискретной информации. — М.: Радио и связь, 1981. — 240 с.
3. Прагер Э., Шимек Б., Дмитриев В. П. Цифровая техника в связи. — М.: Радио и связь, 1981. — 280 с.

К г л. 10

1. Теория сетей связи/Под ред. В. Н. Рогинского. — М.: Радио и связь, 1981. — 192 с.
2. Долуханов М. П. Распространение радиоволн. — М.: Связь, 1972. — 336 с.
3. Калинин А. И., Черенкова Е. Л. Распространение радиоволн и работа радиолиний. — М.: Связь, 1971. — 440 с.
4. Носов Ю. Р. Оптоэлектроника. — М.: Сов. радио, 1977. — 232 с.
5. Мурадян А. Г., Гинзбург С. А. Системы передачи информации по оптическому кабелю. — М.: Связь, 1980. — 158 с.

К г л. 11

1. Стил Р. Принципы дельта-модуляции: Пер. с англ./Под ред. В. В. Маркова. — М.: Связь, 1979. — 368 с.
2. Немировский А. С., Рыжков Е. В. Системы связи и радиорелейные линии. — М.: Связь, 1980. — 432 с.
3. Цифровые системы передачи: Пер. с польского/Под ред. В. Д. Романова. — М.: Связь, 1979. — 262 с.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Абонентский пункт 193
Адаптивная дискретизация 26
Адаптивный прием 171
Алгоритм оптимального приема 61
— — — дискретных сообщений 64, 155
— — — непрерывных сообщений 65, 66, 111
Алфавит сообщений 25, 60, 132, 167
Амплитудно-модулированные сигналы 70, 71, 98, 103, 113
Аппаратура:
временного уплотнения 243
группового преобразования 237
линейного тракта 238
передачи данных 235
сопряжения 238, 241
Асинхронно-адресные системы 92
База сигнала 58, 59, 60, 144
Байеса формула 62, 151
— правило решения 63, 153, 154
Бит 12, 13, 14, 181, 231
Бод 13, 193, 195, 198, 200, 201, 226
Векторное представление колебаний 44, 47, 136
Вероятность ошибки 110, 33, 64
— — аномальной 105, 106
— — при помехоустойчивом кодировании 184, 185
— — распознавания сигналов 161, 163—166, 169, 184
Восстановление функции по отсчетам 29, 117—119
Гауссовские каналы 50, 51, 53, 54, 55—57
Групповой сигнал 9
— тракт (общий) 9, 120, 122
Двоичные коды 133
— — натуральные (первичные) 133
— — помехоустойчивые 134, 176—190
— — эффективные 133, 134
Девияция частоты 72, 73
— фронта импульса 81
Декодер 8, 50, 180, 189, 190
Декомпозиция (разбиение) системы 50
Демодулятор 8, 50, 112
— оптимальный ЧМ-сигналов 114, 115
— синхронный (когерентный) 113
— стандартный ЧМ-сигналов 116
—, чувствительность 112
Дискретизация сообщений по времени:
адаптивная 26
избыточность 29
интервал дискретизации 26—29
ошибка 29, 118
предельная 27—29
равномерная 26
Дискретные сообщения (см. Сообщения)
Единая автоматизированная сеть связи (ЕАСС) 5, 211, 212, 225—245
— — — первичные и вторичные сети 226
— — — — типовая структура тракта 235
— — — — уплотнение 237—239
— — — — цифровая передача 242
Информация (см. Сообщения)
Искажения:
демодуляционные 117—119
в линиях связи 147—149
перекрестные 120
в системах с ЧРК 122, 123
в системах с ВРК 124—126
частотно-селективные 149
Каналы связи 8, 192, 225
— — дискретные 8, 193, 225
— — коммутируемые 215
— — некоммутируемые 215
— — непрерывные 8, 225
— — с коммутацией сообщений 215
— — с направляемыми сигналами 15, 221—224
— — с постоянными параметрами 16
— — со свободным распространением 15, 218—221
— — со случайными параметрами 16, 17, 21, 145
Квантование по уровню:
неравномерное 30
ошибка, 31, 32
равномерное 30
шаг квантования 30
Код двоичный 133
— исправляющий 134
— многоосновной 133
— обнаруживающий 134
— с проверкой на четность 177, 187
Кодек 50
Кодер 7, 50

- Кодирование натуральное 7, 133
 — эффективное 7, 133
 — помехоустойчивое 7, 134
 Корректирующие коды 176—190
 — — блочные 178
 — — избыточность 180, 181
 — — кодовое расстояние 181
 — — корректирующие возможности 181—183
 — — разделимые 178, 179
 — — сверточные 179, 180
 — — систематические 179, 187
 — — циклические 179
 Котельникова теорема 27—29
 — теория помехоустойчивости 48, 51, 102, 103
 Линия связи 6, 135, 145—149, 218—224
 Манипуляция сигнала 34
 — — амплитудная 144
 — — относительная фазовая 169
 — — фазовая 144
 — — частотная 144
 Многоканальная система 9, 196, 199
 — — передача 87, 92
 Многостанционный доступ 19, 20, 237
 Модель канала 8, 9, 50
 — коэффициента передачи линии 148
 — сообщений с ограниченным спектром 27
 Модем 50, 201
 Модулятор 7, 75
 Модуляция 34
 — амплитудная 69, 70
 — амплитудно-импульсная 80, 82, 83
 — время-импульсная 81, 84, 85
 — интегральная 103
 — линейная 98
 — нелинейная 98
 — прямая 103
 — фазовая 69, 72
 — частотная 69, 72, 73
 — широко-импульсная 81, 84, 85
 Оптимальная полоса 160
 Оптимальный прием сигналов
 — — задачи 61, 135
 — — — теоретические основы 62—66, 150—155
 — — — помехоустойчивость 101—109, 161—164
 Оптимальный приемник:
 дискретных сообщений 155
 корреляционный 155, 156
 непрерывных сообщений 64, 65, 109—112
 фильтровый 156, 157
 Оптимизация систем:
 в целом 49—51
 выбор сигналов 57—61
 приемной части 51, 52
 предельные возможности 52—57
 Ортогональные разложения 38—44, 136
 Относительная фазовая манипуляция 169, 170
 Отсчеты 26, 29
 Оценка сообщений 65, 102, 172
 Ошибка аномальная 11, 101, 105, 106
 — дискретизации по времени 29
 — кратность 181
 — необнаруживаемая 185
 — преобразования в цифровую форму 32, 33
 — среднеквадратическая 11, 29, 31—33, 103—108
 Пик-фактор сообщения 31, 69, 118
 Показатели качества систем связи:
 помехоустойчивость 11
 скорость передачи 12
 точность передачи 10
 удельные затраты 55—60
 эффективность 13
 Помеха аддитивная 8, 36
 — «большая» 101
 — «малая» 101
 — импульсная 37, 38, 174
 — шумовая 36, 37
 Помехоустойчивость оптимального приема 98, 102, 103—109, 162—168
 Пороговый эффект 101, 165
 Промежуток защитный временной 91
 — — частотный 89
 Пропускная способность канала связи 13, 52—55
 Разделение каналов 9, 119—126
 — — временное 123—126, 204—206
 — — кодовое (по форме) 91, 92
 — — частотное 121—126, 199—202
 Разрядность кода 32, 133
 Распознавание сигналов 138, 139
 Режим работы синхронный 193
 — — стартовый 193
 Ретранслятор 18, 19, 21, 237
 Сети связи 193, 211—218
 — — вторичные 217
 — — иерархия 216—218
 — — описание 212
 — — первичные 217
 — — структура 212, 213
 Сигналы дискретные:
 инженерная реализация 143, 144
 ортогональные 140, 141
 прямоугольной конфигурации 142
 симплексные 142
 со случайными параметрами 148
 точно известные 147
 узкополосные 147

- широкополосные 148
- Сигналы непрерывные:**
 - аналитические 35
 - квадратурные составляющие 34
 - комплексная форма 35
 - модулированные 33, 68—74, 78
 - огibaющая 34
 - средняя удельная мощность 34
- Синхронизация (фазирование)** 90, 94, 204—210
 - по посылкам 206—208
 - по циклам 208—210
- Системы связи:**
 - аналоговые (непрерывные) 15, 68, 97
 - дискретные (цифровые) 15, 191
 - коротковолновые 22
 - космические 21, 237
 - оптические 23, 223
 - радиорелейные 18, 236
 - с обратным каналом 173
 - спутниковые 18, 237
 - тропосферные 21
- Скорость передачи информации** 12, 13, 52, 54, 225, 226, 234
- Сообщения** 6, 24
 - дискретные 6, 25, 57, 132
 - многоканальные 9, 87—90 (см. Групповой сигнал)
 - непрерывные 6, 25, 60, 69
 - нормированные 69
 - пик-фактор 69
 - удельная мощность 69
- Телеграфная сеть** 232
- Телефонная сеть** 229
- Типовая структура тракта ЕАСС** 235
- Типы линий связи** 218—234
- Уплотнение каналов временное** 90, 91, 204—206, 243
 - — по форме (кодвое) 91, 92
 - — принцип уплотнения 87—89
 - — частотное 89, 90, 198—203
- Устройства аналоговых систем** 126—129
 - декодирования 189
 - кодирования 188
 - дискретных систем 192—210
- Фазирование (см. Синхронизация)**
- Фазоманипулированные сигналы** 144
- Фазовая модуляция** 72
- Фильтр согласованный** 159, 160
- Цифровые системы передачи информации** 19, 21, 191, 192, 242—245
- Частотная модуляция** 72—74
- Частотно-манипулированные сигналы** 144
- Шеннона граница** 56
 - теорема ρ канале с шумом 52
 - формула 53—55

ОГЛАВЛЕНИЕ

Предисловие	3
Глава 1. Общие сведения о системах передачи информации	5
1.1. Введение	5
1.2. Система связи. Основные понятия и определения	6
1.3. Показатели качества систем связи	10
1.4. Классификация систем и линий передачи информации	15
1.5. Основные виды радиотехнических систем передачи информации	17
Контрольные вопросы	23
Глава 2. Сообщения, сигналы и помехи	24
2.1. Введение	24
2.2. Основные виды сообщений и их краткая характеристика	24
2.3. Дискретизация непрерывных сообщений по времени	26
2.4. Квантование непрерывных сообщений по уровню	30
2.5. Преобразование непрерывных сообщений в цифровую форму	32
2.6. Общее описание модулированных сигналов	33
2.7. Помехи в каналах связи	35
2.8. Ортогональные разложения колебаний	38
2.9. Некоторые важные ортогональные разложения	40
2.10. Геометрическое представление колебаний	44
Контрольные вопросы и задачи	47
Глава 3. Элементы общей теории систем передачи информации	48
3.1. Введение	48
3.2. Проблема оптимизации систем передачи информации	49
3.3. Предельные возможности систем передачи информации	52
3.4. Выбор сигналов в оптимальных гауссовских каналах	57
3.5. Элементы теории оптимального приема сигналов	61
Контрольные вопросы и задачи	67
Глава 4. Сигналы в системах передачи непрерывных сообщений и методы их формирования	68
4.1. Введение	68
4.2. Общие соотношения при модуляции колебаний	68
4.3. Гармонические модулированные колебания	69
4.4. Методы формирования сигналов с угловой модуляцией	74
4.5. Импульсные модулированные колебания	78
4.6. Аналитическая запись импульсных видов модуляции и их спектральный состав	82
4.7. Методы формирования импульсных модулированных последовательностей	85
4.8. Методы уплотнения каналов	87
4.9. Сигналы в системах многоканальной связи	92
Контрольные вопросы и задачи	95
Глава 5. Прием сигналов в системах передачи непрерывных сообщений	97
5.1. Введение	97
5.2. Геометрическая трактовка особенностей приема сигналов в присутствии шума	98
5.3. Помехоустойчивость оптимального приема непрерывных сообщений	102
5.4. Потенциальная помехоустойчивость при многоканальной передаче непрерывных сообщений	106
5.5. Структура оптимальных приемников непрерывных сообщений	109
5.6. Методы демодуляции модулированных колебаний	112

5.7. Разделение каналов при многоканальной передаче непрерывных сообщений	119
5.8. Особенности схемной реализации некоторых устройств аналоговых систем передачи сообщений	126
Контрольные вопросы и задачи	130
Глава 6. Принципы формирования сигналов в системах передачи дискретных сообщений. Искажающее влияние линий связи	132
6.1. Принцип передачи дискретных сообщений	132
6.2. Методы создания набора сигналов в системах передачи дискретных сообщений	135
6.3. Распознавание сигналов при отсутствии помех	138
6.4. Распознавание сигналов при наличии помех. Критерий оптимальности набора сигналов	139
6.5. Типовые наборы помехоустойчивых сигналов	140
6.6. Вопросы инженерной реализации сигналов	143
6.7. Линия как искажающий четырехполюсник	145
6.8. Основные виды искажений, вносимых линией	147
Контрольные вопросы и задачи	149
Глава 7. Оптимальный прием сигналов в системах передачи дискретных сообщений	150
7.1. Постановка задачи синтеза оптимального приемника дискретных сообщений	150
7.2. Синтез оптимального приемника для распознавания известных сигналов	152
7.3. Физика работы оптимального корреляционного приемника	155
7.4. Приемник на согласованных фильтрах	157
7.5. Реализация согласованного фильтра. Фильтр, согласованный по полюсе	159
7.6. Расчет вероятности ошибок при распознавании дискретных сигналов	161
7.7. Прием сигналов в линиях, вносящих случайные ослабление и сдвиг фазы	164
7.8. Выбор и расчет некоторых параметров систем передачи дискретных сообщений	167
7.9. Относительное кодирование как способ борьбы со случайной фазой сигналов	169
7.10. Прием сигналов в каналах с быстро изменяющимися параметрами. Прием при неравномерных по спектру помехах	170
7.11. Прием сигналов с оценкой случайных параметров, вносимых линией. Адаптивный прием	171
7.12. О системах передачи информации с обратным каналом	173
7.13. О борьбе с импульсными помехами	174
Контрольные вопросы и задачи	175
Глава 8. Принципы помехоустойчивого кодирования	176
8.1. Идея построения помехоустойчивых кодов	176
8.2. Основные виды корректирующих кодов	178
8.3. Основные характеристики корректирующих кодов	180
8.4. Возможности повышения помехоустойчивости приема цифровой информации с помощью корректирующих кодов	183
8.5. Методы кодирования и декодирования систематических кодов	187
Контрольные вопросы и задачи	191
Глава 9. Принципы реализации систем передачи дискретных сообщений	191
9.1. Основная терминология	191
9.2. Источники дискретной информации. Синхронный и стартстопный режимы работы	193
9.3. Принципы образования каналов передачи	195
9.4. Основные регламентации для систем передачи различного быстродействия	196

9.5. Структурные схемы низкоскоростных многоканальных систем с частотным уплотнением	198
9.6. Структурные схемы средне- и высокоскоростных систем передачи дискретных сообщений	200
9.7. Принцип образования многоканальности с помощью временного уплотнения	204
9.8. Способы фазирования (синхронизации) посылок	206
Контрольные вопросы и задачи	210
Глава 10. Принципы построения сетей связи	211
10.1. Введение	211
10.2. Основные типы сетей и особенности их описания и работы	212
10.3. Иерархия сетей и линий	216
10.4. Особенности радиолиний, применяемых в сетях	218
10.5. Особенности «закрытых» линий связи	221
Контрольные вопросы и задачи	224
Глава 11. Основы построения Единой автоматизированной сети связи	225
11.1. Введение	225
11.2. Первичные и вторичные сети ЕАСС	226
11.3. Телефонная сеть общего пользования	229
11.4. Телеграфная сеть	232
11.5. Общегосударственная система передачи данных (ОГСПД)	233
11.6. Линии передачи и типовая структура тракта ЕАСС	235
11.7. Построение групп уплотнения и линейных трактов при аналоговых (непрерывных) сигналах	239
11.8. О построении цифровых систем информации в ЕАСС	242
Контрольные вопросы	245
Заключение	246
Список литературы	248
Предметный указатель	251

ПЕТР ИЛЬИЧ ПЕНИН. ЛЕОНИД ИВАНОВИЧ ФИЛИППОВ

РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Редактор А. И. Давыдовский, Редактор издательства Ю. И. Суханов
Художник К. М. Прасолов, Художественный редактор Л. Н. Сильянов
Технический редактор Т. Н. Зыкина, Корректор Т. Г. Захарова

ИБ № 806

Сдано в набор 2.08.83 г. Подписано в печать 6.12.83 г.
Т-20096 Формат 60×90^{1/16} Бумага тип. № 3 Гарнитура литературная Печать высокая
Усл. печ. л. 16,0 Усл. кр.-отт. 16,0 Уч.-изд. л. 16,94 Тираж 12 000 экз. Изд. № 19694
Зак. № 88 Цена 90 к.

Издательство «Радио и связь». 101000 Москва, Почтамт, а/я 693

Московская типография № 5 ВГО «Союзучиздат»
101000 Москва, ул. Кирова, д. 40