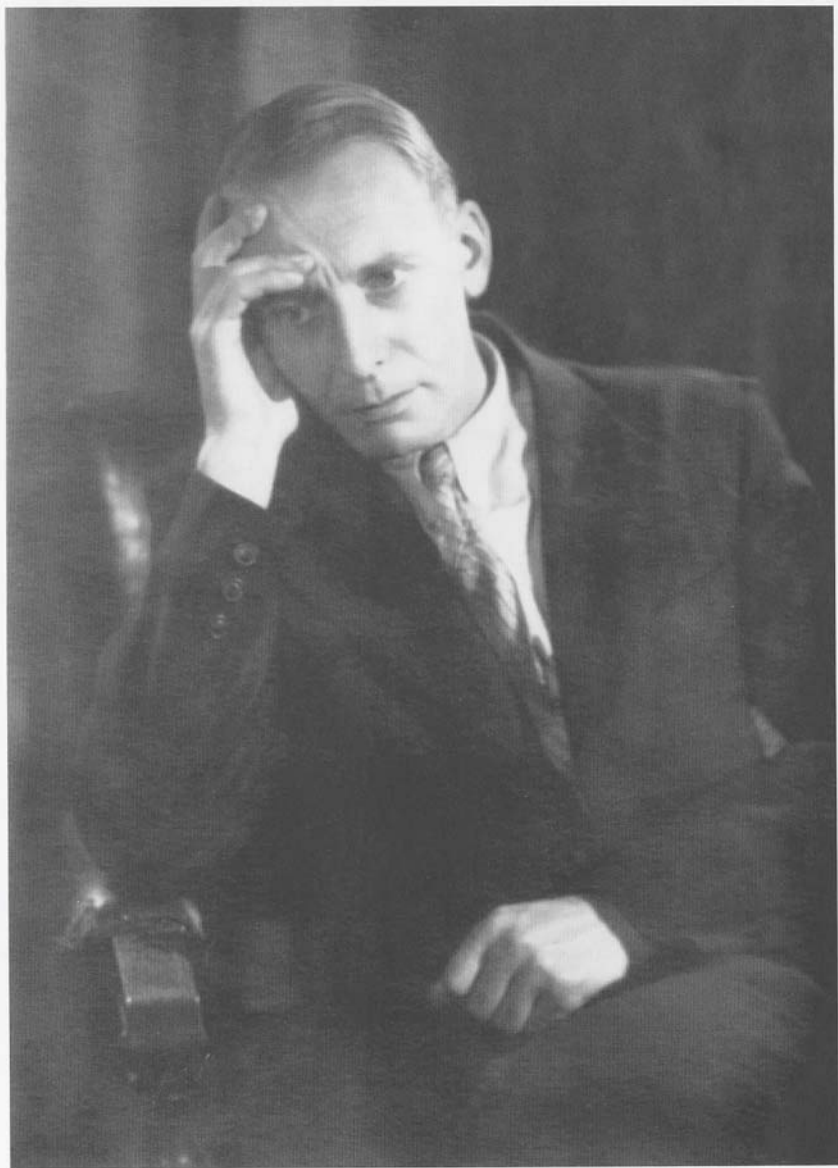


КЛАССИКИ ОБРАЗОВАНИЯ





Академик А.А. Харкевич (1904 – 1965)

РАДИОТЕХНИКА и ЭЛЕКТРОНИКА

А.А. ХАРКЕВИЧ

ОСНОВЫ РАДИОТЕХНИКИ

ИЗДАНИЕ ТРЕТЬЕ, СТЕРЕОТИПНОЕ



МОСКВА
ФИЗМАТЛИТ®
2007

УДК 621.37
ББК 32.84
Х 21

Харкевич А. А. **Основы радиотехники.** — 3-е изд., стер. — М.: ФИЗМАТЛИТ, 2007. — 512 с. — ISBN 978-5-9221-0790-7.

В курс теоретических основ радиотехники вошли: общие вопросы передачи и приема сигналов, исследование прохождения электрических сигналов через внутренние цепи аппаратуры и распространения сигналов по линиям и волноводам, исследование основных радиотехнических процессов.

Математический аппарат курса включает решение линейных дифференциальных уравнений с постоянными и переменными коэффициентами и решение нелинейных дифференциальных уравнений.

Настоящее издание полностью воспроизводит текст издания 1962 года, которое было допущено Министерством высшего и среднего специального образования СССР в качестве учебного пособия для высших учебных заведений СССР.

ОГЛАВЛЕНИЕ

Предисловие к третьему изданию	7
Введение	9

Часть I. Передача и прием сигналов

Глава 1. Сигнал и линия	12
§ 1. Основные понятия	12
§ 2. Понятие об излучении и распространении радиоволн	14
§ 3. Физические характеристики сигнала	19
Глава 2. Модуляция и коды	25
§ 4. Модуляция; общие понятия	25
§ 5. Спектры модулированных колебаний	30
§ 6. Векторные диаграммы модуляции	37
§ 7. Импульсная модуляция	39
§ 8. Код; общие понятия	42
§ 9. Теорема Котельникова	45
§ 10. Квантование	49
§ 11. Импульсно-кодовая модуляция	52
§ 12. Многоканальная связь	54
§ 13. Частотное и временное разделение	55
Глава 3. Помехи и помехоустойчивость	60
§ 14. Помехи; общее описание	60
§ 15. Флуктуационные помехи	62
§ 16. Шум и замирание	66
§ 17. Понятие помехоустойчивости	68
§ 18. Увеличение отношения сигнал/помеха	69
§ 19. Зависимость отношения сигнал/помеха от вида модуляции	73
§ 20. Некоторые свойства импульсно-кодовой модуляции	76
§ 21. Помехоустойчивость квантованного сигнала	79
§ 22. Корректирующие коды	81

Часть II. Лине́йные явления в радиотехнике

Глава 4. Задачи и методы линейной теории	86
§ 23. Лине́йные явления и устройства	86
§ 24. Принцип наложения	87
§ 25. Временной и спектральный подходы	89
§ 26. Интеграл Фурье	91
§ 27. Интеграл Дюамеля	95
§ 28. Частотные и временные характеристики	100
§ 29. Основы теории спектров	107
Глава 5. Системы с сосредоточенными параметрами	119
§ 30. Характеристика и назначение рассматриваемых систем	119
§ 31. Одиночный контур; частотные характеристики	124
§ 32. Одиночный контур; временные характеристики	136
§ 33. Связанные контуры	143
§ 34. Системы со многими степенями свободы	154
§ 35. Краткие сведения из теории четырехполосников	158
§ 36. Фильтры; общие понятия	165
§ 37. Нагрузка фильтра	175
§ 38. Фильтры типа M	180
§ 39. Линия задержки	184
§ 40. Некоторые схемы с обратной связью	191
§ 41. Дифференцирующие и интегрирующие схемы	196
§ 42. Электромеханические системы	204
Глава 6. Волновые системы	210
§ 43. Линии; общие понятия	210
§ 44. Волновые параметры	214
§ 45. Отражение; стоячие волны	217
§ 46. Коэффициенты отражения	219
§ 47. Входное сопротивление линии	226
§ 48. Методы согласования	230
§ 49. Резонансы в линии	233
§ 50. Линия с потерями	239
§ 51. Неустановившиеся процессы в линии	245
§ 52. Распространение импульсов по линии	254
§ 53. Сопоставление линии и фильтра	258
§ 54. Волноводы; общие понятия	261
§ 55. Элементарная теория прямоугольного волновода	265

§ 56. Картина поля в прямоугольном волноводе; типы волн	270
§ 57. Круглый волновод	278
§ 58. Возбуждение и фильтрация волн в волноводе	282
§ 59. Потери в волноводе	286
§ 60. Объемные резонаторы	291
§ 61. Сопоставление волновода и линии	295
§ 62. Разложение волн в волноводе на плоские волны	298

Часть III. Нелинейные и параметрические явления в радиотехнике

Глава 7. Введение	303
§ 63. Характеристика предмета	303
§ 64. Обзор методов нелинейной теории	305
Глава 8. Нелинейные цепи и основные нелинейные процессы . . .	309
§ 65. Нелинейная цепь	309
§ 66. Преобразование спектра в нелинейной цепи	314
§ 67. Умножение частоты	316
§ 68. Линеаризация и нелинейные искажения	318
§ 69. Выпрямление	322
§ 70. Детектирование	325
§ 71. Детектирование ЧМ и ИМ	335
§ 72. Инерционная нелинейность	338
§ 73. Ограничение и стабилизация	340
§ 74. Модуляция	345
§ 75. Преобразование частоты	349
§ 76. Нелинейный способ получения АМ	351
§ 77. Практические схемы модуляции	354
§ 78. Специальные электронные приборы	360
Глава 9. Генерирование колебаний	364
§ 79. Автоколебания	364
§ 80. Энергетика автоколебаний; динамическая устойчивость	366
§ 81. Условие самовозбуждения генератора	368
§ 82. Комплексная обратная связь	370
§ 83. Инженерные критерии устойчивости	372
§ 84. Установившийся режим генератора	378
§ 85. Установление колебаний в генераторе	384

§ 86. Генератор с инерционной нелинейностью	389
§ 87. Фазовая плоскость	390
§ 88. Фазовые портреты генераторов	394
§ 89. Графические методы	397
§ 90. Некоторые схемы генераторов.	401
§ 91. <i>RC</i> -генераторы синусоидальных колебаний.	406
§ 92. Релаксационные автоколебания.	409
§ 93. Релаксационные генераторы	412
Глава 10. Нелинейные системы под внешним воздействием	417
§ 94. Работа выпрямителя	417
§ 95. Резонанс в нелинейной цепи	420
§ 96. Триггеры.	423
§ 97. Автоколебательная система под внешним воздействием.	430
§ 98. Захватывание	435
§ 99. Регенеративный приемник	438
§ 100. Импульсная синхронизация	439
§ 101. Синхронизация релаксационного генератора	442
Глава 11. Параметрические явления.	446
§ 102. Синхронное детектирование	446
§ 103. Фазовое детектирование и разделение	450
§ 104. Параметрическое возбуждение	453
§ 105. Деление частоты.	457
§ 106. Регенеративное деление частоты	462
§ 107. Усиление как параметрический процесс	465
§ 108. Некоторые генераторы свч	477
Приложения	482
<i>Приложение 1. Доказательство теоремы Котельникова.</i>	<i>482</i>
<i>Приложение 2. Сравнение частотного и временного уплотнения по использованию полосы пропускания линии</i>	<i>484</i>
<i>Приложение 3. Обоснование критерия Найквиста</i>	<i>487</i>
<i>Приложение 4. Краткие сведения об уравнении Матъё.</i>	<i>493</i>
<i>Приложение 5. Основы параметрического усиления.</i>	<i>498</i>
<i>Приложение 6. Прохождение периодических колебаний сложной формы через линейные цепи</i>	<i>505</i>
Список рекомендуемой литературы	509

Предисловие к третьему изданию

Выдающийся отечественный ученый академик Александр Александрович Харкевич родился 3 февраля 1904 г. в г. Санкт-Петербурге.

В 1930 г. А. А. Харкевич заканчивает учебу в Петроградском электротехническом институте и получает диплом инженера. Во время учебы он работал в Центральной радиолaborатории (ЦРЛ) в Ленинграде, где им был разработан первый отечественный диффузорный динамический громкоговоритель.

В 1938 г. А. А. Харкевич защищает докторскую диссертацию и получает звание профессора в Военной электротехнической академии. Позже он организовал кафедру в ленинградском электротехническом институте связи. Во время войны А. А. Харкевич заведует Лабораторией в Физико-техническом институте АН СССР, которая занималась разработкой электроакустических преобразователей для гидроакустических устройств подводных лодок.

В 1944–1951 гг. А. А. Харкевич работает во Львове и Киеве, где он избирается членом-корреспондентом Украинской академии наук. В Киеве А. А. Харкевич успешно руководит разработками в новой области техники — магнитной записи сигналов и выполняет глубокие теоретические исследования в области волновых процессов, результаты которых изложены в его монографии «Неустановившиеся волновые явления» (1952 г.).

С 1952 г. А. А. Харкевич работает в Москве, где возглавляет кафедру теоретической радиотехники Московского электротехнического института связи (ныне — Московский технический университет связи и информатики).

В течение последующих десяти лет им создаются замечательные труды по радиотехнике, которые сыграли выдающуюся роль в подготовке специалистов в этой области. В 1952 г. издаются две книги «Спектры и анализ» и «Автоколебания», в 1956 г. — «Нелинейные и параметрические явления в радиотехнике», в следующем году — «Теоретические основы радиосвязи», а в 1962 г. — «Основы радиотехники». Эти талантливо написанные книги пользовались заслуженной популярностью; их переводили в Польше, Китае, США и Англии.

С 1954 г. А. А. Харкевич работает в Лаборатории по разработке проблем проводной связи АН СССР, сначала в должности старшего научного сотрудника, а позже в качестве заведующего лабораторией, которая преобразована в Лабораторию систем передачи информации.

На базе этой лаборатории в 1961 г. был создан Институт проблем передачи информации АН СССР (ИППИ АН СССР). Директором института был назначен академик А. А. Харкевич. Институт был создан для развития теории информации и ее приложений, разработки принципиальных вопросов построения системы передачи и распределения

информации при создании концепции единой сети связи страны, а также разработки теории распознавания образов и цифровой обработки изображений. В настоящее время Институт проблем передачи информации Российской академии наук (ИППИ РАН) носит имя академика А. А. Харкевича и является ведущим мировым центром в области теории информации и биоинформатики.

Следует особенно отметить выдающиеся способности А. А. Харкевича как педагога и популяризатора науки. Его живая творческая и научная мысль, его умение излагать высокие научные идеи простым и ясным языком и способность донести их до аудитории студентов, аспирантов и инженеров снискали ему заслуженную славу блестящего оратора и талантливого лектора и учителя. Именно поэтому созданные им в 50–60 гг. двадцатого века учебники для вузов были неоднократно переизданы и не потеряли актуальность и в наши дни.

В книге «Основы радиотехники» блестяще изложены основные принципы построения приемно-передающих устройств и физические основы теории распространения радиосигналов. Книга написана так, что физическая сущность явлений не теряется в дебрях математической техники. Конечно, за годы, прошедшие со времени издания книги, радиотехника существенно шагнула вперед в области методов обработки сигналов. С развитием вычислительной техники получила большое распространение цифровая обработка сигналов. Это означает, что вместо применявшихся ранее моделей сигналов в виде непрерывных спектров и дифференциальных уравнений используются методы описания сигналов с помощью дискретных частотных методов и разностных уравнений. Однако эти изменения касаются только формального математического аппарата, а физическая сущность явлений и исходные уравнения, описывающие эти явления, остались неизменными. Именно эта идея превосходно отражена в учебниках А. А. Харкевича.

Актуальность изучения студентами вузов основ радиотехники не вызывает сомнений именно в современных условиях, когда бурно развивается теория и практика применения беспроводной связи: сотовая телефония, сотовое телевидение, широкополосный доступ в Интернет, использование спутниковых систем в глобальных сетях, использование в региональных сетях — Wi-Max и в локальных сетях Wi-Fi технологий.

Академик *Н. А. Кузнецов*,
директор ИППИ РАН им. академика А. А. Харкевича
в 1990–2004 гг.

Введение

Радиотехника представляет собой обширную область науки и техники, продолжающую быстро развиваться. Каждый день ставит перед радиотехникой новые задачи. Поэтому теория становится все более сложной, техника все более изощренной. Объем радиотехнических знаний неуклонно возрастает. Меняется и наше понимание предмета радиотехники. Попытаемся с современной точки зрения ответить на вопрос: что такое радиотехника?

Все основные применения радиотехники связаны с передачей и приемом тех или иных сигналов. Так, к радиотехнике относятся все виды радиосвязи, начиная со старейшего — радиотелеграфа. Сюда относятся также радиотелефон, радиофототелеграф, телевидение, радиовещание. Но кроме радиосвязи, радиотехнике принадлежит также область радиотелеметрии, т. е. измерения на расстоянии с передачей результатов измерения при помощи радиосигналов. В радиолокации само измерение осуществляется при помощи радиосигналов, отражаемых от объекта наблюдения. Радионавигация и радиогеодезия применяют специальные радиосигналы, измеряя параметры которых можно определять направления, расстояния и местонахождение передатчика и приемника. Одно из новейших применений радиотехники — радиоастрономия — основано на приеме и изучении радиосигналов, поступающих от внеземных объектов — Солнца и других звезд, туманностей и межзвездного пространства. Радиолокация нашла космические применения; как известно, получены отраженные сигналы не только от Луны, но и от Венеры. В космических масштабах осуществляется телеметрическая и телевизионная связь со спутниками и космическими кораблями. Управление на расстоянии производится при помощи радиосигналов.

Под радиосигналами во всех перечисленных применениях понимаются сигналы, переносимые электромагнитными волнами. Для радиотехники характерно применение сравнительно коротких волн, т. е. высоких частот. Более того, для современного этапа развития радиотехники важную роль играет тенденция к дальнейшему укорочению волн и повышению частот, к освоению новых диапазонов. Диапазон дециметровых волн освоен совсем недавно. На очереди — сантиметровые волны, техника которых уже в значительной мере разработана. В ближайшем будущем предстоит освоение диапазона миллиметровых волн.

Еще не так давно радиотехника определялась как техника связи *без проводов* и противопоставлялась проводной связи. Это деление уже устарело. Во-первых, расширились применения радиотехники, которые охватывают не только связь в узком смысле слова. Во-вторых, утратило смысл выражение «без проводов». Известно, что независимо от наличия или отсутствия проводов электрический сигнал передается

посредством электромагнитной волны. Дело сводится лишь к наличию тех или иных *направляющих* систем. Термин «без проводов» относится по существу к передаче электромагнитными волнами, распространяющимися в неограниченном однородном пространстве. Такие условия в более или менее чистом виде имеются лишь в космическом пространстве. В земных условиях волны распространяются в шаровом слое между земной поверхностью и ионосферой. Кроме того, нужно учесть, что электромагнитные волны передаются не только по коаксиальным кабелям (где еще можно говорить о проводах), но и по волноводам, которые можно с равным правом относить как к «проводной», так и к «беспроводной» передаче. Более того, созданы диэлектрические, ленточные и другие волноводы, и, наконец, так сказать, «волновод наоборот» — одиночный провод, с успехом используемый в качестве направляющей для волн сверхвысокой частоты.

Учитывая все это, можно определить радиотехнику как *технику передачи сигналов при помощи электромагнитных волн высокой частоты*.

Надо пояснить, что техника высокой частоты не включается в радиотехнику. К технике высокой частоты относятся: электрический нагрев для разных технических целей (сушка, закалка, плавление и т. п.), высокочастотный электрический транспорт («ВЧТ»), а также биологические применения поля высокой частоты. Техника высокой частоты применяет элементы радиотехнической аппаратуры, как, например, генераторы и усилители. Но то же относится в наше время к *любой* отрасли науки и техники. Нельзя себе представить, например, современную исследовательскую лабораторию — к какой бы отрасли она ни относилась — без радиотехнической аппаратуры, т. е. без генераторов, усилителей и в особенности без радиоизмерительных приборов — осциллографов, вольтметров и т. п. В этой аппаратуре нуждаются физики и биологи, врачи и металлурги, химики и строители. Словом, радиотехническая аппаратура и методика широко используются во всех отраслях естествознания и техники, опирающихся на эксперимент.

Это отнюдь не означает, что все эти отрасли нужно причислить к радиотехнике, которая имеет свои собственные весьма обширные задачи. Но это значит, что квалифицированный радиоинженер найдет применение своим знаниям в очень многих отраслях, далеко за пределами основных задач собственно радиотехники.

Исходя из приведенного выше определения предмета радиотехники, можно составить себе представление о содержании курса теоретических основ радиотехники. Сюда должны войти прежде всего общие вопросы передачи и приема сигналов: строение системы передачи, построение сигналов, вопросы модуляции и кодов, помехи передаче и методы борьбы с ними. Все это составляет первую часть курса, озаглавленную «Передача и прием сигналов». Затем нужно исследовать прохождение электрических сигналов через внутренние цепи аппаратуры и распространение сигналов по линиям и волноводам. Эти вопросы

и общие методы их исследования рассматриваются во второй части курса «Линейные явления в радиотехнике». Такое название обусловлено тем, что с математической точки зрения исследование перечисленных вопросов сводится к составлению и решению линейных дифференциальных уравнений с постоянными коэффициентами. Наконец, третья часть курса, озаглавленная «Нелинейные и параметрические явления в радиотехнике», посвящена основным радиотехническим процессам, как-то: генерирование колебаний, модуляция, детектирование и т. п. Математическим аппаратом этой части являются нелинейные дифференциальные уравнения, а также линейные уравнения с переменными коэффициентами.

Курс «Основы радиотехники» базируется на физике, математике и теоретической электротехнике. Он читается параллельно с курсами электронных и ионных приборов и радиоизмерений и, в свою очередь, закладывает основу для последующих специальных радиотехнических курсов.

Некоторые теоретические вопросы затрагиваются в данном курсе лишь бегло, так как они подробно излагаются в специальных курсах. Таковы, например, вопросы теории антенн и распространений радиоволн.

Содержание курса подверглось, разумеется, некоторому обновлению по сравнению с предыдущим изданием (1956–1957 гг.). Однако требования к радиоинженеру быстро возрастают, и ощущается необходимость радикальной перестройки учебных планов, программ и учебников. Такая перестройка неизбежно произойдет в ближайшем будущем.

Часть I

ПЕРЕДАЧА И ПРИЕМ СИГНАЛОВ

Глава 1

СИГНАЛ И ЛИНИЯ

§ 1. Основные понятия

Посредством электрических сигналов передаются те или иные *сообщения*. Между сообщением и сигналом должно быть однозначное соответствие. Сигнал отображает сообщение в форме определенного электрического возмущения; по принятому сигналу можно восстановить переданное сообщение.

В случае радиотелеграфа сообщение — это телеграмма, т.е. некоторый словесный текст. На основе определенной телеграфной азбуки (телеграфного кода) он преобразуется в сигнал, например, в виде комбинаций коротких и длинных посылок («точек» и «тире», как в телеграфном коде Морзе). На приемной стороне принятые сигналы расшифровываются и восстанавливается сообщение, т.е. первоначальный текст телеграммы.

В этом примере отправителем и получателем сообщения является человек. Но во многих случаях функции отправителя и получателя сообщений выполняют те или иные автоматические устройства. Так, например, при автоматическом управлении на расстоянии принятые команды исполняются автоматическими устройствами. Другим примером служат автоматические измерительные устройства, вроде необслуживаемых метеорологических станций, автоматически передающих показания многочисленных измерительных приборов. Прием и предварительная обработка данных также могут быть автоматизированы.

В идеальном случае принятое сообщение тождественно переданному. В действительности же этого никогда не бывает; принятое сообщение в той или иной мере отличается от переданного. Степень соответствия принятого сообщения переданному называется вообще *верностью* передачи. На верность передачи влияют в основном два обстоятельства.

Во-первых, сигнал искажается в цепях и устройствах передатчика и приемника. Под искажениями понимаются такие изменения сигнала, которые обусловлены *известными* свойствами цепей и устройств, вы-

ражаемыми их *характеристиками*. Но если характеристики устройств известны, то обусловленные ими искажения сигнала могут быть (по крайней мере, в принципе) устранены либо путем совершенствования аппаратуры и методов обработки сигнала, либо путем исправления уже искаженного сигнала. Такое исправление носит название *коррекции*.

Во-вторых, сигнал подвергается воздействию различного рода *помех*. Под помехами понимаются сторонние возмущения, действующие в тракте передачи. Если помеха в точности известна, то ее нетрудно устранить путем компенсации — так обстоит, например, дело с помехой, называемой «фоном» переменного тока. Но в большинстве случаев помеха не только нами не контролируется, но имеет *случайный* характер. Ясно, что при таких условиях борьба с ней сильно затрудняется. А между тем проблема борьбы со случайными помехами становится все более актуальной, так как радиотехнике приходится решать все более трудные задачи: необходимо передавать сигналы на все большие расстояния при ограниченной мощности; при этом с каждым днем возрастают требования к надежности, т. е. к верности передачи. Поэтому методам борьбы с помехами в радиотехнике уделяется очень большое внимание.

Назовем *передатчиком* устройство, преобразующее сообщение в сигнал, а *приемником* устройство, производящее обратное преобразование сигнала в сообщение. Тогда можно изобразить *систему передачи* в целом, как показано схематически на рис. 1.

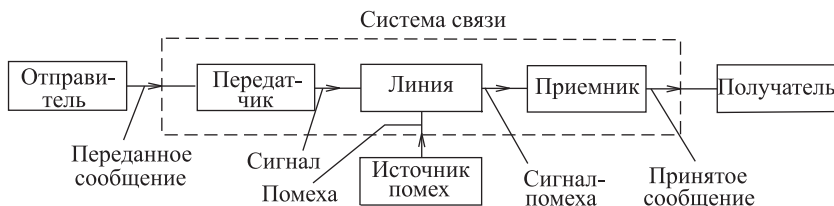


Рис. 1

Обведенную пунктирной рамкой часть системы называют *каналом*. Помехи могут действовать в различных звеньях системы; для удобства рассмотрения обычно объединяют все источники помех в один эквивалентный, как показано на схеме. Сигнал распространяется от передатчика к приемнику по некоторой *линии*. Линия есть тракт электромагнитной волны. В простейшем случае она представляется в виде пары проводов. Конструктивно линия выполняется либо в виде пары параллельных проводов, либо в виде *коаксиальной линии*, в которой второй провод имеет форму трубки, охватывающей первый провод. (Примерами могут служить обычные телевизионные кабели — симметричный и коаксиальный.) В качестве линии может применяться *волновод*. В простейшей форме волновод представляет собой трубу с проводящими стенками; электромагнитная волна распространяется

внутри трубы. Линии и волноводы являются направляющими системами, локализирующими поле электромагнитной волны. Но электромагнитная волна может распространяться и в открытом (неограниченном) пространстве. Раньше именно этот случай считался характерным для радиотехники; радиотехника определялась как техника передачи сигналов «без проводов». Строго говоря, в земных условиях пространство, в котором распространяются электромагнитные волны, ограничено. Волны распространяются в шаровом слое, заключенном между поверхностью Земли и ионосферой. Но при передаче направленным лучом в пределах прямой видимости это ограничение пространства не играет роли. В этом случае под линией нужно понимать телесный угол, опирающийся на антенну приемника. В общей формулировке линией следует называть зону пространства, по которой распространяется достигающая приемника волна.

Механизм распространения радиоволн может быть различным. Краткому обзору вопросов распространения радиоволн посвящен следующий параграф.

§ 2. Понятие об излучении и распространении радиоволн

Мы будем рассматривать здесь вопросы, относящиеся к передаче сигналов при помощи электромагнитных волн, распространяющихся без направляющих систем (т. е. без проводных линий или волноводов).

Следует сразу подчеркнуть различие между статическим электрическим (или магнитным) полем и полем электромагнитной волны. Дело в том, что напряженность статического электрического поля, создаваемого системой заряженных тел (или статического магнитного поля, создаваемого системой проводов, обтекаемых токами), при больших расстояниях убывает с третьей степенью расстояния или еще быстрее. В то же время напряженность как электрической, так и магнитной составляющих поля свободно распространяющейся электромагнитной волны убывает лишь с первой степенью расстояния. Этим и обусловлена возможность связи на больших расстояниях при помощи электромагнитных волн.

Процесс создания распространяющейся от источника электромагнитной волны называется *излучением*. Нужно прежде всего составить понятие об этом процессе. Теория излучения рассматривается со всеми необходимыми подробностями в курсе теории поля; мы ограничимся здесь только элементарными качественными представлениями.

Представим себе систему проводов, по которым протекает постоянный ток. В окружающем пространстве существует статическое магнитное поле. Если постепенно уменьшить ток до нуля, то поле, также постепенно убывая, исчезнет; содержащаяся в нем энергия будет возвращена источнику тока. Такая же картина наблюдается и при питании

системы переменным током низкой частоты: поле периодически возникает и исчезает, энергия в течение одной четверти периода (пока ток нарастает) запасается в поле, а в течение последующей четверти периода (пока ток убывает) возвращается источнику. Сопротивление системы будет при этих условиях чисто реактивным (если пренебречь активным сопротивлением проводов); система ведет себя как индуктивность. Образующееся в описанных условиях поле — это статическое магнитное поле.

Картина явлений совершенно изменится, если та же система проводов (назовем ее теперь антенной) будет питаться током высокой частоты. Качественное различие заключается в том, что при высокой частоте решающую роль играет конечная скорость распространения электромагнитного поля. Вышеописанное периодическое перемещение энергии из источника в окружающее поле и обратно происходит только в ограниченной области пространства, непосредственно примыкающей к антенне. Из более удаленных областей электромагнитное поле не успеет перенести энергию обратно в антенну за требуемое короткое время (около четверти периода). В этих, более удаленных, областях энергия перемещается поступательно по направлению от антенны в окружающее пространство. Существующее здесь электромагнитное поле и есть поле электромагнитной волны. В установившемся режиме источник все время отдает энергию и, следовательно, затрачивает мощность на поддержание поля; энергия разносится во все стороны расходящейся электромагнитной волной. В этих условиях сопротивление антенны получает активную составляющую, так называемое *сопротивление излучения*. Произведение из квадрата действующего значения тока на сопротивление излучения равно излучаемой мощности. Теория показывает, что сопротивление излучения пропорционально квадрату отношения высоты антенны h к длине волны λ . В принципе можно получить излучение при любой частоте (т. е. при любой длине волны); однако для получения эффективного излучения отношение h/λ должно быть порядка единицы, что приводит к совершенно фантастическим размерам антенных устройств для низких частот. Именно поэтому в радиотехнике применяются высокие частоты: они позволяют получить эффективное излучение при помощи антенн приемлемых размеров.

В радиотехнике применяются частоты примерно от 10 кГц до 100 ГГц. Этот огромный диапазон принято подразделять на части, присваивая отдельным поддиапазнам специальные наименования, приведенные в табл. 1.

Диапазон миллиметровых волн уже вплотную подходит к диапазону инфракрасных световых волн. Напомним, что диапазон видимых световых волн составляет 0,4–0,7 мкм (частоты порядка 10^{14} Гц).

С радиотехнической точки зрения важно отметить, что конструкция аппаратуры существенно изменяется при переходе к частотам, на порядок более высоким; меняются отчасти и применяемые принципы (например, принципы генерирования колебаний, усиления колебаний

Таблица 1

Наименование волн	Длины волн	Частоты
Сверхдлинные	$10 \div 100$ км	$30 \div 3$ кГц
Длинные	$1 \div 10$ км	$300 \div 30$ кГц
Средние	$100 \div 1000$ м	$3 \div 0,3$ МГц
Короткие	$10 \div 100$ м	$30 \div 3$ МГц
Метровые	$1 \div 10$ м	$300 \div 30$ МГц
Дециметровые	$1 \div 10$ дм	$3 \div 0,3$ ГГц
Сантиметровые	$1 \div 10$ см	$30 \div 3$ ГГц
Миллиметровые	$1 \div 10$ мм	$300 \div 30$ ГГц

и т. п.). С этим связаны специфические трудности «освоения» новых диапазонов радиоволн. С точки же зрения радиосвязи нужно в первую очередь указать на качественные различия в механизме распространения волн различной длины.

Условия распространения радиоволн между передатчиком и приемником, расположенными на поверхности земли (или вблизи от нее), характеризуются, во-первых, наличием земной (или водной) поверхности, обладающей конечной проводимостью, и, во-вторых, наличием ионосферы — верхних слоев атмосферы, разреженный газ которых ионизован действием солнечных и космических лучей. Радиоволны преломляются и поглощаются в ионосфере в тем большей мере, чем больше степень ионизации и чем длиннее волна. Этим обусловлены, с одной стороны, суточные и сезонные изменения условий распространения радиоволн, а с другой стороны, различия в условиях распространения волн различной длины.

Волна может достигать приемника как за счет огибания выпуклой земной поверхности (земная волна), так и за счет преломления в ионосфере (пространственная волна). Длинные и средние волны распространяются при помощи обоих механизмов, в зависимости от времени суток и расстояния. Короткие волны достигают места приема главным образом за счет преломления в ионосфере, т. е. за счет рефракции. Упрощенно можно трактовать явление как отражение от ионизованного слоя. Отражение происходит лишь при условии, что угол падения достаточно мал; другими словами, волны, распространяющиеся под большим углом к горизонту, проникают в ионосферу без отражения и назад не возвращаются. Поэтому вблизи передатчика образуется «зона молчания», в которой приема нет; прием делается возможным лишь за пределами зоны молчания. Волна может попасть в точку приема и после многократного отражения попеременно от ионосферы и от земной поверхности; наблюдался прием коротких волн, обогнувших кругом земной шар. Наконец, ультракороткие волны (укв) практически не отражаются от ионосферы. Поэтому надежная связь на укв возможна лишь в условиях прямой видимости, т. е. на расстоянии примерно до 50 км или несколько больше, в зависимости от высоты подъема

антенн над поверхностью земли. Заметим, что свойства укв позволяют применять их для связи с внеземными объектами.

Все сказанное иллюстрируется рис. 2. На рис. 2, *а* показано распространение земной волны. Сплошная линия — путь волны, огибающей земную поверхность, пунктирная — путь преломленной волны, интенсивность которой мала вследствие большого поглощения в ионосфере. Рис. 2, *б* дает картину распространения пространственной волны, пунктиром намечен путь волны с многократным отражением, а также пути волн, уходящих в ионосферу. Наконец, на рис. 2, *в* показано распространение укв.



Рис. 2

Новые возможности передачи укв на большом расстоянии выявились вследствие рассеяния укв на *неоднородностях тропосферы и ионосферы*. Дело заключается в том, что как тропосфера, так и ионосфера представляют собой неоднородные среды — их диэлектрическая проницаемость меняется от места к месту случайным образом. В тропосфере она зависит от давления, температуры и влажности воздуха, которые непостоянны и изменяются по мере движения воздушных масс. В ионосфере диэлектрическая проницаемость зависит от электронной плотности сильно ионизированного газа и также изменяется вследствие постоянного беспорядочного движения газа. В результате этих движений как в тропосфере, так и в ионосфере образуются неоднородности, т. е. области с отличным от среднего значением диэлектрической проницаемости.

Когда волна встречается на пути своего распространения с неоднородностью, то происходит ее рассеяние. Явление состоит в том, что создается излучение во все стороны, т. е. дело обстоит так, как если бы неоднородность сама стала источником излучения. При рассеянии интенсивность излучения в различных направлениях неодинакова. Наибольшая часть излучения направлена в сторону распространения исходной волны — это так называемое «рассеяние вперед». Но так как имеется, хотя и более слабое, излучение во всех направлениях, то вследствие рассеяния возможен прием укв и за пределами прямой видимости. Все это показано схематически на рис. 3, *а*, на котором заштрихованная область изображает отдельную неоднородность. Если бы использовалась одна индивидуальная неоднородность, то устойчивая связь не была бы возможна, так как неоднородности являются образо-

ваниями, все время случайным образом изменяющими свое положение, размеры и физические свойства. Но антенны передатчика и приемника обладают конечной остротой направленности, т. е. захватывают определенные телесные углы. Поэтому используется рассеяние от неоднородностей в пределах сравнительно большой области пространства (заштрихованной на рис. 3, б), так что эффект рассеяния от многих неоднородностей усредняется. Это позволяет получить достаточно надежную непрерывную связь.

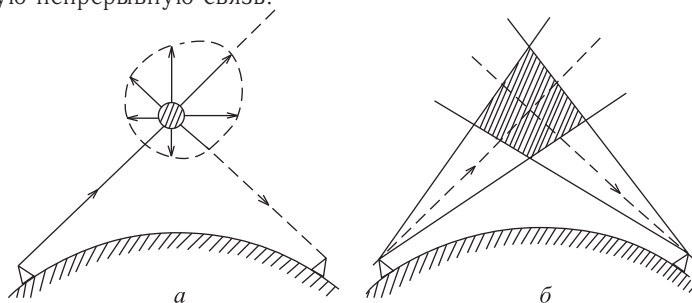


Рис. 3

Другая возможность состоит в использовании *метеорных следов*. Метеоры входят в плотные слои атмосферы с очень большими скоростями (порядка десятков километров в секунду). Разогреваясь вследствие трения о воздух до весьма высокой температуры, метеорное вещество превращается в сильно ионизированный газ, образующий облако, сильно растянутое вдоль траектории метеора. Это и есть метеорный след; он иногда легко наблюдается невооруженным глазом в виде светлой полоски. Метеорный след представляет собой неоднородность, рассеивающую падающую на него электромагнитную волну. При массе метеора, превышающей 10^{-5} г, ионизация настолько велика, что метеорный след отражает волну подобно проводящему телу. При благоприятном расположении следа относительно передатчика и приемника и при достаточной величине и степени ионизации следа можно осуществить передачу на большие расстояния (1000 км и больше).

Связь на метеорных следах имеет ту существенную особенность, что хотя общее количество метеоров (с массой более 10^{-5} г), попадающих в земную атмосферу, очень велико (порядка 10^{10} в сутки), но появление «подходящего» (по положению и степени отражения) следа на трассе данной радиолинии — событие, происходящее не так уж часто. Поэтому на метеорных линиях ведется *перемежающаяся связь*. Система выжидает, пока появившийся след создаст достаточный уровень принимаемого сигнала. Когда это произойдет, автоматически включается передача; также автоматически она прекращается, когда уровень принимаемого сигнала упадет ниже установленного порога: система ожидает возникновения следующего следа. Таким образом, пе-

редача ведется отдельными «вспышками». На протяжении «вспышки» передача идет в очень быстром темпе. Так, в одной из действующих телеграфных систем передается 1300 слов в минуту, а средняя скорость передачи составляет около 60 слов в минуту, так что на передачу затрачивается всего лишь 5 %, а на выжидание 95 % времени. Опыт эксплуатации показал высокую надежность этого нового вида связи.

Наконец, следует упомянуть о возможности связи при помощи отражения радиоволн от специальных искусственных спутников Земли. Проекты такой системы усиленно разрабатываются. Предложено много вариантов искусственных отражающих устройств. Ближайшее будущее покажет, какой из этих вариантов будет избран и осуществлен.

В заключение параграфа отметим характерную для современного этапа развития радиотехники тенденцию к переходу на все более короткие волны, т. е. на все более высокие частоты.

На это есть серьезные причины; некоторые из них мы вкратце укажем:

1. Применение очень высоких частот позволяет получить остронаправленное излучение, что имеет большое значение не только для радиолокации, но и для радиосвязи, в частности при построении радиорелейных линий.

2. В диапазоне коротких волн делаются неощутимыми атмосферные и многие виды промышленных помех.

3. Чем выше частота, тем меньше дает себя знать «теснота в эфире», т. е. тем большее число станций может работать без взаимных помех.

4. Большой «простор» позволяет применить помехоустойчивые широкополосные системы модуляции.

5. Чем больше скорость передачи, тем больше должна быть частота, на которой она ведется. Примером может служить телевидение: телевизионную передачу осуществляют только на укв.

Некоторые из этих соображений будут подробнее разъяснены в дальнейшем. Вопрос о направленности разбирается в курсе антенн.

§ 3. Физические характеристики сигнала

Всякий сигнал представляет собой изменяющуюся во времени электрическую величину (напряжение, ток, напряженность поля) и может быть, следовательно, выражен некоторой функцией времени.

Задание этой функции полностью определяет сигнал. Но нам желательно описывать свойства сигнала некоторым более экономным образом, выбрав для этого такие показатели, которые по возможности просто и в то же время достаточно полно характеризовали бы сигнал с точки зрения условий его передачи.

В качестве таких показателей употребляются следующие три величины: *длительность сигнала, динамический диапазон его и ширина спектра.*

Длительность сигнала — наиболее простая и в то же время практически важная его характеристика. Чем она больше, тем на большее время занимается линия.

Динамическим диапазоном называется отношение наибольшей мгновенной (так называемой пиковой) мощности сигнала к наименьшей мощности (называемой иногда пороговой). Обычно измеряют динамический диапазон логарифмической мерой и выражают его в децибелах. Ясно, что выбор наименьшей мощности сигнала связан с уровнем помех; для удовлетворительной передачи требуют обычно, чтобы наименьшая мощность сигнала несколько превышала мощность помех. Широко распространился также иной показатель, а именно отношение *средних* мощностей сигнала и помехи. Для краткости говорят просто «отношение сигнал/помеха», имея в виду именно отношение средних мощностей. Логарифм этого отношения называют превышением сигнала над помехой. Таким образом, динамический диапазон или превышение представляют собой показатели, характеризующие мощность сигнала, но не абсолютную, а отнесенную к мощности помехи.

Наконец, третьим показателем свойств сигнала является ширина его спектра, и на этом показателе нам придется остановиться подробнее.

Вначале мы будем предполагать, что сигнал представляет собой периодическую функцию времени (в дальнейшем мы откажемся от этого упрощающего предположения). В таком случае можно представить сигнал рядом Фурье, т.е. разложить его на синусоидальные составляющие.

Если сигнал представлен периодической функцией $x(t)$ с периодом T , то можно записать

$$x(t) = c_0 + \sum_{k=1}^{\infty} (a_k \cos k\omega_1 t + b_k \sin k\omega_1 t) \quad (1)$$

или

$$x(t) = c_0 + \sum_{k=1}^{\infty} c_k \cos(k\omega_1 t - \varphi_k), \quad (2)$$

где

$$c_k = \sqrt{a_k^2 + b_k^2}, \quad \operatorname{tg} \varphi_k = \frac{b_k}{a_k}, \quad \omega_1 = \frac{2\pi}{T},$$

$$a_k = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \cos k\omega_1 t \, dt,$$

$$b_k = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \sin k\omega_1 t \, dt.$$

Смысл формулы (2) состоит в том, что периодическая функция $x(t)$ может быть представлена суммой синусоидальных колебаний с ча-

стотами, кратными основной частоте ¹⁾ ω_1 , и с надлежащим образом подобранными амплитудами c_k и начальными фазами φ_k .

Отдельные слагаемые суммы (2) называются *гармониками*. Колебание основной частоты $\omega_1 = \frac{2\pi}{T}$ называется первой гармоникой, колебание с частотой $2\omega_1$ — второй гармоникой и так далее.

Постоянная составляющая

$$c_0 = \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt$$

представляет собой просто среднее значение функции $x(t)$; постоянная составляющая сигнала на практике чаще всего отсутствует, так как не проходит через такие элементы схемы, как трансформаторы или конденсаторы.

Совокупность величин c_k называется *спектром* амплитуд; совокупность величин φ_k — *спектром* фаз. Чаще всего интересуются только спектром амплитуд и называют его для краткости просто спектром; если речь идет о фазах, то это специально оговаривают.

Графически спектр представляют в координатах c_k, ω , как показано на рис. 4. Длины вертикальных отрезков представляют амплитуды соответствующих гармоник; эти отрезки называют *спектральными линиями*. Спектр вида рис. 4 называют *линейчатым*.

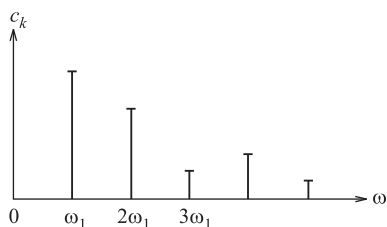


Рис. 4

В общем случае сумма (2) представляет бесконечный ряд. Но в действительности для всех сигналов число членов суммы, а следовательно, и число спектральных линий конечны, так как амплитуды гармоник, начиная с некоторого номера n , уже настолько малы, что ими можно пренебречь. Таким образом, сигналы практически представляются функциями с *ограниченным спектром*.

Интервал на шкале частот, в котором размещается ограниченный спектр, и называется *шириной спектра*.

Ограничение спектра производится в технике связи сознательно, так как мы заинтересованы в его сокращении. Это обусловлено тем, что как аппаратура, так и линия связи пропускают ограниченную полосу частот.

¹⁾ Мы будем на протяжении курса пользоваться, наряду с угловой частотой ω (выражаемой в радианах в секунду), также и частотой $f = \omega/2\pi$ (выражаемой в герцах), применяя термин *частота* как к той, так и к другой величине.

Ограничение спектра производят, исходя из допустимого искажения сигнала. Мы рассмотрим этот вопрос на нескольких примерах из области связи.

Начнем с телеграфного сигнала. Чтобы найти ширину спектра при телеграфировании, возьмем самый невыгодный случай, когда телеграфный сигнал претерпевает наиболее быстрые изменения. Это будет сигнал, представляющий собой чередование самых коротких посылок тока (точек) и самых коротких пауз, равных по длительности точкам. Периодический сигнал такого рода изображен на рис. 5, а. Основная частота сигнала есть $f_1 = 1/T$; в телеграфии ее называют частотой манипуляции. Спектр функции рис. 5, а безграничен. Но можно отбросить высшие гармоники. При этом, конечно, ухудшится форма принимаемого сигнала. Обычно полагают,

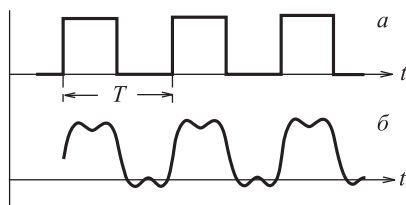


Рис. 5

что, сохранив гармоники до третьей включительно, можно получить еще удовлетворительную форму принятого сигнала. На рис. 5, б показана форма сигнала при устранении всех гармоник выше третьей.

Теперь можно подсчитать ширину спектра телеграфного сигнала. Спектр укладывается в интервал от нуля до частоты $3f_1$. Таким образом, $3f_1$ — это и есть ширина спектра. Частота манипуляции зависит от скорости передачи. Пусть передается 200 слов в минуту. Считая в среднем по шесть букв на слово, получим, что число букв в секунду составляет 20. При применении кода Морзе средняя длительность буквы, выраженная через длительность одной «точки», составляет около 10. В секунду будет передаваться в среднем 100 точек и пауз. Это и есть частота манипуляции.

Итак, ширина спектра телеграфного сигнала в нашем примере составляет

$$F = 3f_1 = 300 \text{ Гц.}$$

Обратимся к передаче звука — к телефонии и вещанию. Здесь спектр ограничивается, как правило, с учетом особенностей слухового восприятия. Человек, как правило, не слышит звуковых колебаний с частотами выше $12 \div 15$ кГц. Значит, более высокие частоты не нужно передавать, и при ограничении спектра частотой 15 кГц можно получить субъективно-идеальное воспроизведение звука. Однако передача такого широкого спектра связана с большими затратами и техническими трудностями. Многочисленными опытами установлена примерная оценка качества звучания в зависимости от верхней граничной частоты спектра, указанная в табл. 2.

К этому нужно добавить, что спектр можно ограничивать и снизу, используя слабую восприимчивость слуха к низким частотам. Это

Таблица 2

Верхняя граница спектра, кГц	Качество звучания
15	идеальное
10	высокое
8	хорошее
5	удовлетворительное

обстоятельство имеет особое значение при построении низкочастотной части радиоприемников.

Приведенные оценки относятся к передаче музыки и учитывают эстетическое впечатление, производимое на слушателей. При телефонии требования могут быть значительно снижены, так как телефонная связь должна удовлетворять только двум условиям:

- 1) чтобы речь была разборчива и
- 2) чтобы собеседники могли узнавать друг друга по голосу.

Для выполнения этих условий достаточно передать спектр от 200 до 3000 Гц.

В качестве третьего примера рассмотрим телевизионный сигнал. Его спектр ограничивается с учетом свойств зрительного восприятия. Глаз имеет ограниченную разрешающую способность. Этим определяется требуемая четкость изображения, зависящая от числа строк (в Советском Союзе принято 625 строк). Такая же четкость, естественно, должна обеспечиваться и вдоль строки. Следовательно, число различных элементов вдоль строки должно составлять $\frac{4}{3} 625 = 833$, где $\frac{4}{3}$ — отношение сторон кадра. Общее число различных элементов изображения будет $625 \cdot 833 \approx 500\,000$. Наивысшая частота при передаче телевизионного изображения получится в том крайнем случае, когда изображение будет представлять собой чередование черных и белых полей (наподобие шахматной доски), причем каждое поле имеет наименьший размер, т. е. элемента разложения.

При сканировании (развертывании) такого изображения получится сигнал точно такого вида, как на рис. 5, а: посылка соответствует черному полю, пауза — белому. На протяжении каждого кадра получится $2,5 \cdot 10^5$ периодов, а в секунду передается 25 кадров. Основная частота сигнала равна в рассматриваемом случае

$$f_1 = 25 \cdot 2,5 \cdot 10^5 \approx 6 \text{ МГц.}$$

Так как речь идет об очень мелких деталях изображения, то достаточно передавать только основную частоту f_1 ; она и является, следовательно, верхней граничной частотой спектра.

В состав телевизионного сигнала входят также синхронизирующие импульсы, следующие в принятой у нас системе с частотой 50 Гц. Таким образом, ограничение снизу невозможно, и мы видим, что спектр

телевизионного сигнала исключительно широк — он в тысячу раз шире спектра звукового сигнала. Это существенно затрудняет построение системы телевизионного вещания.

Ширина спектра сигнала является очень важным показателем. Для неискаженной передачи сигнала (точнее, для передачи с допустимым искажением) нужно, чтобы система обладала достаточной полосой пропускания. Если полоса пропускания уже, чем спектр, если спектр не вмещается в предоставленную полосу частот, то это влечет за собой недопустимое ухудшение качества передачи.

Глава 2

МОДУЛЯЦИЯ И КОДЫ

§ 4. Модуляция; общие понятия

Спектры сигналов всех видов простираются далеко в область низких частот, а между тем, как было показано в § 2, радиосвязь возможна только на высоких частотах. Отсюда следует, что для передачи посредством излучения электромагнитных волн нужно перенести, передвинуть спектр сигнала в область высоких радиочастот. В таком преобразовании низкочастотного сигнала и состоит сущность *модуляции*.

Возьмем синусоидальное колебание высокой частоты:

$$u = U_0 \cos(\omega_0 t + \varphi_0), \quad (3)$$

где U_0 , ω_0 и φ_0 — соответственно амплитуда, частота и начальная фаза. Если колебание синусоидально, то все три эти параметра постоянны.

Пусть теперь один из параметров, например амплитуда, изменяется во времени в соответствии с передаваемым низкочастотным сигналом, т. е.

$$U_m = U_0 + \Delta U x(t), \quad (4)$$

где ΔU — постоянная; $x(t)$ — низкочастотный сигнал.

В таком случае мы запишем вместо (3)

$$u_{\text{AM}} = U_0 \left[1 + \frac{\Delta U}{U_0} x(t) \right] \cos(\omega_0 t + \varphi_0). \quad (5)$$

Это есть выражение для амплитудно-модулированного сигнала. Колебание вида (5) уже не есть синусоида, так как синусоидальное колебание имеет постоянную амплитуду. Множитель (4) выражает так называемую *огibaющую* модулированного колебания. Функция $x(t)$ называется *модулирующей функцией*. Частота ω_0 немодулированного колебания (3) называется *несущей частотой*. Функции (3), (4) и (5) изображены на рис. 6. Заметим, что огибающая $U_m(t)$ воспроизводит кривую $x(t)$.

Как будет показано ниже, модулированное колебание есть высокочастотное колебание; низкочастотных составляющих оно не содержит.

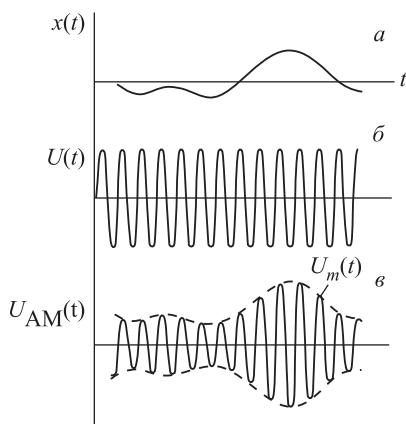


Рис. 6

Тем не менее модулированное колебание несет в себе низкочастотный сигнал, который заложен в изменениях амплитуды. Для того чтобы извлечь низкочастотный сигнал из модулированного колебания, на приемной стороне системы связи нельзя воспользоваться фильтрами, так как низкочастотный сигнал не входит в состав модулированного колебания в качестве *слагаемого*; низкочастотный сигнал входит в качестве *сомножителя* (см. (5)), поэтому для его извлечения потребуется особая операция, называемая *детектированием* (детектирование — обнаружение). В детек-

торе происходит операция, обратная модуляции; она состоит в том, что спектр низкочастотного сигнала снова возвращается на свое место, т. е. в область низких частот. Получающиеся при детектировании побочные продукты в виде составляющих высокой частоты отделяются фильтром.

Модуляция происходит в специальном устройстве, называемом модулятором. Назначение модулятора состоит в том, чтобы при подаче на один его вход напряжения несущей частоты (3), а на другой вход — низкочастотного сигнала $x(t)$, дать на выходе модулированное колебание (5), т. е. выполнить определяемую выражением (5) операцию перемножения.

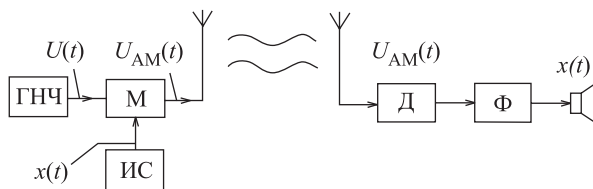


Рис. 7

Теперь мы можем изобразить блок-схему радиосвязи (рис. 7). Передающая часть системы состоит из ГНЧ — генератора несущей частоты, ИС — источника низкочастотного сигнала (телеграфный ключ, микрофон и т. п.), М — модулятора. Выход модулятора соединен непосредственно с передающей антенной. Приемная антенна связана с детектором Д, за которым включен фильтр Ф. На схеме опущены усилители,

не играющие принципиальной роли: мощная ступень передатчика и усилителя высокой и низкой частоты приемника.

Рассмотрим возможные виды модуляции для случая, когда немодулированное колебание (колебание несущей частоты) синусоидально и записывается в виде

$$u = U_0 \cos(\omega_0 t + \varphi_0).$$

Как уже говорилось, при чисто синусоидальном колебании все три параметра, определяющие колебание, т.е. амплитуда U_0 , частота ω_0 и начальная фаза φ_0 , — постоянны. Модуляция состоит в том, что один из параметров изменится в соответствии с изменениями передаваемого низкочастотного сигнала $x(t)$. Изменять можно любой из трех названных параметров. В соответствии с этим мы получим три вида модуляции: амплитудную модуляцию (АМ), частотную модуляцию (ЧМ) и фазовую модуляцию (ФМ). Во всех случаях модуляция состоит в том, что тот или иной параметр получает приращение, пропорциональное $x(t)$. Так, имеем

$$\text{АМ} \quad U_m = U_0 + \Delta U x(t), \quad (6)$$

$$\text{ЧМ} \quad \omega = \omega_0 + \Delta \omega x(t), \quad (7)$$

$$\text{ФМ} \quad \varphi = \varphi_0 + \Delta \varphi x(t). \quad (8)$$

Запишем аналитические выражения для модулированных колебаний при всех трех видах модуляции. Для АМ это уже было сделано: если подставить в (3) вместо U_0 выражение (6), то получим, отбросив несущественное в данном случае слагаемое φ_0 ,

$$u_{\text{АМ}} = U_0 \left[1 + \frac{\Delta U}{U_0} x(t) \right] \cos \omega_0 t. \quad (9)$$

При обычной амплитудной модуляции нужно позаботиться о том, чтобы амплитуда (6) не принимала отрицательных значений. В противном случае получатся специфические искажения — так называемая перемодуляция. Если модулирующая функция $x(t)$ определена так, что ее наибольшее абсолютное значение равно единице, то относительное изменение амплитуды

$$\frac{\Delta U}{U_0} = m$$

называют *коэффициентом модуляции*; вышеприведенное требование сводится к условию

$$m < 1.$$

Коэффициент амплитудной модуляции на практике не превосходит в среднем $0,4 \div 0,5$; хотя при такой неглубокой модуляции хуже используется мощность передатчика, но зато уменьшается вероятность перемодуляции при пиковых значениях модулирующей функции. Характер искажений при перемодуляции поясняется рис. 8. Здесь следует

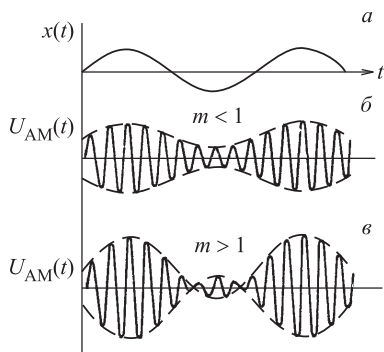


Рис. 8

отметить, что при перемодуляции (рис. 8, в) огибающая АМ колебания уже не повторяет формы кривой $x(t)$.

Обратимся к фазовой модуляции. Чтобы получить аналитическое выражение для ФМ колебания, достаточно подставить в (3) вместо φ_0 выражение (8). Это дает

$$u_{\text{ФМ}} = U_0 \cos [\omega_0 t + \Delta\varphi x(t)] \quad (10)$$

(постоянная φ_0 отброшена). Это выражение существенно отличается от (9); разница состоит в том, что модулирующая функция $x(t)$ входит не в множитель при $\cos \omega_0 t$, а под знак косинуса.

Наконец, в случае ЧМ мы имеем для частоты выражение (7). Однако это выражение нельзя непосредственно подставлять в (3) вместо ω_0 . Дело в том, что мы имеем в случае ФМ и ЧМ *переменную* частоту и должны опираться на общие определения.

Колебание с постоянной амплитудой (случай ФМ и ЧМ) можно представить в следующей общей форме:

$$u = U_0 \cos \vartheta(t), \quad (11)$$

где $\vartheta(t)$ — мгновенная фаза. Для синусоидального колебания $\vartheta(t) = \omega_0 t + \varphi_0$, т. е. фаза растет пропорционально времени. С точки зрения векторного представления это означает, что вектор вращается равномерно с постоянной угловой скоростью ω_0 . Если же частота непостоянна, то ее мгновенное значение определяется как

$$\omega = \frac{d\vartheta}{dt}. \quad (12)$$

Колебание с изменяющейся частотой представляется вектором, вращающимся с переменной угловой скоростью. В этом случае на основании определения (12) имеем для фазы

$$\vartheta = \int \omega dt \quad (13)$$

где ω — функция времени. Смысл соотношений (12) и (13) легко уяснить при помощи аналогии с теми соотношениями, которые связывают между собой скорость и пройденный путь. В нашем случае ϑ —

¹⁾ Знак $\int$ означает неопределенный интеграл без произвольной постоянной.

угол, на который повернулся вектор (аналогия пройденного пути), ω — переменная угловая скорость вектора (аналогия линейной скорости).

Таким образом, общее выражение (11) можно переписать в виде

$$u = U_0 \cos \left(\int_0^t \omega dt \right).$$

Теперь ясно, как нужно составить аналитическое выражение для ЧМ колебания. Мы должны взять выражение (7) для переменной частоты и проинтегрировать его:

$$\vartheta = \int_0^t [\omega_0 + \Delta\omega x(t)] dt = \omega_0 t + \Delta\omega \int_0^t x(t) dt.$$

Итак, получаем для ЧМ колебания

$$u_{\text{ЧМ}} = U_0 \cos [\omega_0 t + \Delta\omega X(t)], \quad (14)$$

где обозначено

$$X(t) = \int_0^t x(t) dt.$$

Сравнивая выражение (10) для ФМ и (14) для ЧМ, мы видим, что они различаются между собой только тем, что при ФМ в аргумент косинуса входит модулирующая функция $x(t)$, а при ЧМ — интеграл от этой функции. Что касается $\Delta\omega$ и $\Delta\varphi$, то это постоянные, определяющие глубину соответствующего вида модуляции и выбираемые по нашему усмотрению. Можно, следовательно, сказать, что ФМ и ЧМ находятся между собой в близком родстве, и более того, можно рассматривать ЧМ как разновидность ФМ или наоборот. Вместе с тем оба эти вида модуляции существенно отличаются от АМ.

Для пояснения различия трех видов модуляции на рис. 9 показаны формы модулированных колебаний. Модулирующая функция $x(t)$ изменяется по треугольному закону. Такой вид функции $x(t)$ выбран с целью подчеркнуть различие между ФМ и ЧМ. Если бы мы взяли для $x(t)$

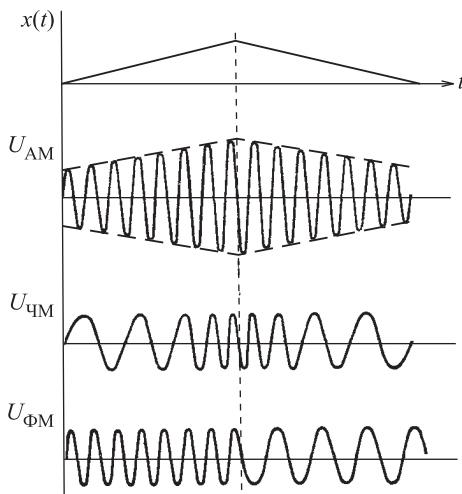


Рис. 9

синусоидальный закон изменения, то $X(t)$ изменялось бы также по синусоидальному закону и нельзя было бы обнаружить никакого различия в форме модулированных колебаний при ФМ и ЧМ.

§ 5. Спектры модулированных колебаний

Всякое модулированное колебание несинусоидально и имеет сложный спектр. Наша очередная задача состоит в нахождении спектров модулированных колебаний для всех трех рассмотренных выше видов модуляции — АМ, ЧМ и ФМ.

Начнем с АМ в том простейшем случае, когда модулирующая функция представляет собой синусоидальное колебание низкой частоты Ω

$$x(t) = \cos \Omega t.$$

АМ колебание записывается в виде

$$U_{\text{АМ}} = U_0 (1 + m \cos \Omega t) \cos \omega_0 t. \quad (15)$$

Раскрыв это выражение, получим

$$u_{\text{АМ}} = U_0 \left[\cos \omega_0 t + \frac{m}{2} \cos (\omega_0 + \Omega) t + \frac{m}{2} \cos (\omega_0 - \Omega) t \right]. \quad (16)$$

Таким образом, АМ колебание (15) состоит из трех составляющих: колебание несущей частоты ω_0 и два колебания с частотами $\omega_0 \pm \Omega$. Эти колебания носят название *спутников*; их частоты называются *боковыми частотами*. Состоящий из трех линий спектр АМ колебания изображен на рис. 10. Амплитуда спутников пропорциональна глубине модуляции. При отсутствии модуляции $m = 0$; при этом, очевидно, спутников нет. При наиболее глубокой модуляции, т. е. при $m = 1$, амплитуды спутников равны половине амплитуды колебания несущей частоты.

Мощность АМ колебания зависит от глубины модуляции. Мощность несущей частоты неизменна, она пропорциональна квадрату амплитуды, т. е. U_0^2 . Мощность каждого спутника пропорциональна квадрату его амплитуды, т. е. $\frac{m^2}{4} U_0^2$.

Таким образом, при наиболее глубокой модуляции ($m = 1$) мощность АМ колебания (равная сумме мощностей всех трех составляющих) в полтора раза превосходит мощность немодулированного колебания несущей частоты.

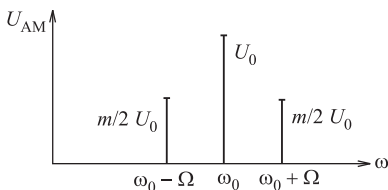


Рис. 10

Рассмотрим теперь более общий случай, когда модулирующая функция периодична и выражается рядом Фурье

$$x(t) = \sum_{k=1}^{\infty} c_k \cos(k\omega_1 t - \varphi_k)$$

(предполагается, что постоянная составляющая отсутствует). Тогда имеем для АМ колебания следующее выражение:

$$\begin{aligned} u_{\text{АМ}} &= U_0 \left[1 + m \sum_{k=1}^{\infty} c_k \cos(k\omega_1 t - \varphi_k) \right] \cos \omega_0 t = \\ &= U_0 \left\{ \cos \omega_0 t + \frac{m}{2} \sum_{k=1}^{\infty} c_k \cos[(\omega_0 + k\omega_1)t - \varphi_k] + \right. \\ &\quad \left. + \frac{m}{2} \sum_{k=1}^{\infty} c_k \cos[(\omega_0 - k\omega_1)t + \varphi_k] \right\}. \end{aligned}$$

Это выражение показывает, что спектр АМ колебания состоит из колебания несущей частоты ω_0 , колебаний с суммарными частотами $\omega_0 + k\omega_1$ и колебаний с разностными частотами $\omega_0 - k\omega_1$. Обе эти суммы носят название *боковых полос*.

Легко видеть, что спектр верхней боковой полосы ($\omega_0 + k\omega_1$) есть не что иное, как спектр модулирующей функции $x(t)$, сдвинутый на ω_0 в сторону высоких частот. Что же касается нижней боковой полосы ($\omega_0 - k\omega_1$), то ее спектр симметричен спектру верхней боковой полосы относительно несущей частоты ω_0 (т.е. он представляет собой как бы зеркальное отражение спектра верхней боковой полосы). Все эти соотношения показаны на рис. 11. На рис. 11, а изображен спектр модулирующей функции $x(t)$. Немодулированное колебание несущей частоты представлено на рис. 11, б. На рис. 11, в показан спектр модулированного колебания с двумя боковыми полосами.

Если спектр модулирующей функции ограничен сверху частотой F , то ширина спектра АМ колебания составляет, очевидно, $2F$. Для того чтобы одновременно работающие станции не мешали друг другу, нужно, чтобы их спектры не перекрывались. Для этого необходимо разнести несущие частоты не менее чем на $2F$. При обычном вещании ограничивают спектр звукового сигнала сверху частотой

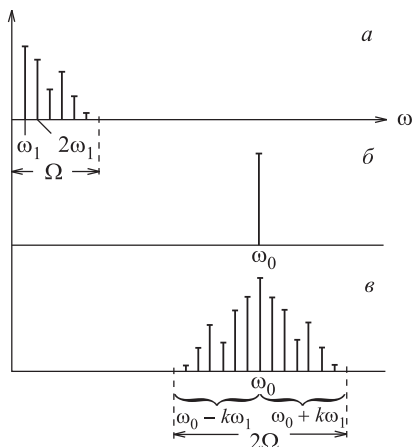


Рис. 11

5 кГц. Ширина спектра модулированного колебания составляет 10 кГц. На такой интервал и должны быть удалены друг от друга несущие частоты вещательных станций, если, конечно, станции территориально расположены настолько близко и мощности передатчиков настолько велики, что в месте приема возможны взаимные помехи. Так, например, в пределах Европы все мощные вещательные станции слышны. Поэтому в средневолновом диапазоне $200 \div 600$ м, что соответствует полосе частот $1,5 \div 0,5$ МГц, можно разместить только около 100 станций. Такое положение и есть то, что называется «теснотой в эфире». Однако легко видеть, что в коротковолновом диапазоне гораздо просторнее. Так, в диапазоне $20 \div 60$ м полоса частот в десять раз больше, а именно 10 МГц (от 15 до 5 МГц). Но нужно ведь разместить не только вещательные станции, но и огромное количество служебных радиосвязей на всевозможных волнах. Распределение волн регулируется специальными международными соглашениями.

По поводу АМ следует заметить, что, кроме обычной АМ с двумя боковыми полосами, для специальных связей применяется еще передача *одной боковой полосой* (ОБП); вторая боковая полоса подавляется. Эта система позволяет вдвое сократить ширину модуляционного спектра, но требует более сложной аппаратуры. Следует также упомянуть о передаче без несущей частоты, осуществляемой с применением *балансной модуляции*, сущность которой состоит в том, что модулирующая функция, перемножается с колебанием несущей частоты по формуле

$$u_{\text{БАМ}} = U_0 x(t) \cos \omega_0 t.$$

Так, например, при синусоидальной модуляции

$$x(t) = \cos \Omega t$$

имеем

$$u_{\text{БАМ}} = U_0 \cos \Omega t \cos \omega_0 t = \frac{1}{2} U_0 [\cos (\omega_0 + \Omega) t + \cos (\omega_0 - \Omega) t],$$

т. е. только два спутника — несущая отсутствует. Заметим, что с математической точки зрения балансная модуляция отличается от обычной отсутствием постоянной слагающей в скобках в формуле (15). При приеме балансно-модулированного колебания приходится добавлять несущую частоту от специального маломощного генератора, что, конечно, усложняет дело. Смысл балансной модуляции состоит главным образом в сокращении бесполезного расхода энергии на составляющую несущей частоты и, следовательно, в уменьшении средней мощности передатчика.

Перейдем к ЧМ и ограничимся рассмотрением синусоидальной модуляции. Для этого случая имеем (см. формулу (14)):

$$\begin{aligned} x(t) &= \cos \Omega t, & X(t) &= \frac{1}{\Omega} \sin \Omega t, \\ u_{\text{ЧМ}} &= U_0 \cos \left(\omega_0 t + \frac{\Delta \omega}{\Omega} \sin \Omega t \right). \end{aligned} \tag{17}$$

Величина $\Delta\omega$ называется *частотным отклонением* (или девиацией частоты); отношение

$$\beta = \Delta\omega/\Omega$$

носит название *индекса частотной модуляции*.

Чтобы найти спектр ЧМ колебания, нужно разложить (17) в тригонометрический ряд. Элементарными средствами это сделать нельзя, но известны следующие разложения:

$$\begin{aligned}\cos(x \sin \varphi) &= J_0(x) + 2 \sum_{n=1}^{\infty} J_{2n}(x) \cos 2n \varphi, \\ \sin(x \sin \varphi) &= 2 \sum_{n=0}^{\infty} J_{2n+1}(x) \sin (2n+1) \varphi,\end{aligned}\tag{18}$$

где $J_n(x)$ — бesselева функция порядка n аргумента x .

На основании (18) получаем из (17)

$$\begin{aligned}u_{\text{ЧМ}} &= U_0 \cos(\omega_0 t + \beta \sin \Omega t) = \\ &= U_0 [\cos \omega_0 t \cos(\beta \sin \Omega t) - \sin \omega_0 t \sin(\beta \sin \Omega t)] = \\ &= \frac{1}{2} U_0 \left\{ \cos \omega_0 t \left[J_0(\beta) + 2 \sum J_{2n}(\beta) \cos 2n \Omega t \right] - \right. \\ &\quad \left. - \sin \omega_0 t 2 \sum J_{2n+1}(\beta) \sin (2n+1) \Omega t \right\} = \\ &= \frac{1}{2} U_0 \left\{ J_0(\beta) \cos \omega_0 t + \sum J_{2n}(\beta) \cos (\omega_0 + 2n \Omega) t + \right. \\ &\quad + \sum J_{2n}(\beta) \cos (\omega_0 - 2n \Omega) t + \sum J_{2n+1}(\beta) \cos [\omega_0 + (2n+1) \Omega] t - \\ &\quad \left. - \sum J_{2n+1}(\beta) \cos [\omega_0 - (2n+1) \Omega] t \right\}.\end{aligned}$$

Первые две суммы берутся по всем четным числам $2n$, а вторые две суммы — по всем нечетным числам $2n+1$. Поэтому можно, объединить суммы попарно, учтя знаки, и получим окончательно

$$\begin{aligned}u_{\text{ЧМ}} &= U_0 \left\{ J_0(\beta) \cos \omega_0 t + \sum_{k=1}^{\infty} J_k(\beta) \cos (\omega_0 + k \Omega) t + \right. \\ &\quad \left. + \sum_{k=1}^{\infty} (-1)^k J_k(\beta) \cos (\omega_0 - k \Omega) t \right\}.\end{aligned}$$

Таким образом, оказывается, что при синусоидальной ЧМ получается теоретически безграничный спектр; амплитуда k -й составляющей пропорциональна $J_k(\beta)$. Однако практически ширина спектра ЧМ колебания ограничена. Это вытекает из свойств бesselевых функций. На рис. 12 дан график, на котором бesselевы функции $J_k(\beta)$ представлены в зависимости от k при различных значениях β .

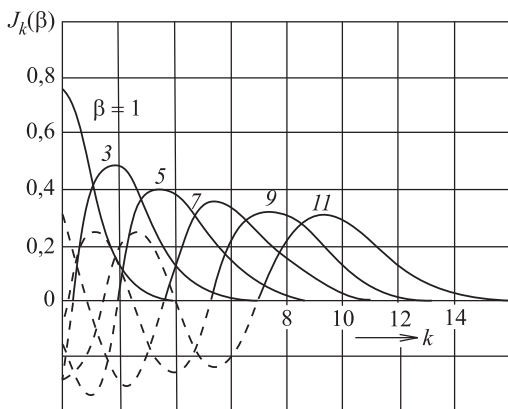


Рис. 12

Как видно из графика, функция $J_k(\beta)$ быстро убывает, когда k сравнивается по порядку величины с β ; в дальнейшем, при $k > \beta$ функция $J_k(\beta)$ имеет очень малые значения. Это значит, что амплитуды составляющих спектра ЧМ колебания становятся очень малыми для $k > \beta$ и ими можно пренебречь. Условливаются отбрасывать все составляющие, амплитуда которых не превышает 0,01. Этим определяется *действительная ширина* спектра ЧМ колебания. Пусть n — число спектральных линий в каждой боковой полосе. Интервалы между линиями на шкале частот равны Ω . Таким образом, действительная ширина каждой боковой полосы составляет $n\Omega$, а полная ширина спектра — $2n\Omega$. При $\beta \gg 1$ можно положить $n \approx \beta$; в этом случае, следовательно, полная ширина спектра равна

$$2n\Omega \approx 2\beta\Omega = 2\Delta\omega,$$

т.е. удвоенному частотному отклонению. Величина $2\Delta\omega$ называется также *полосой качания*, так как в процессе модуляции частота может принимать любое мгновенное значение внутри интервала $\omega_0 \pm \Delta\omega$.

Мы получили, что действительная ширина спектра ЧМ колебания равна в пределе $2\Delta\omega$. Но это лишь предельное соотношение, справедливое с достаточным приближением при больших значениях индекса модуляции β . Рассмотрим теперь соотношения при малом β . Вернемся для этого к выражению (17)

$$\begin{aligned} u_{\text{ЧМ}} &= U_0 \cos(\omega_0 t + \beta \sin \Omega t) = \\ &= U_0 [\cos \omega_0 t \cos(\beta \sin \Omega t) - \sin \omega_0 t \sin(\beta \sin \Omega t)]. \end{aligned}$$

Перейдем к пределу при $\beta \rightarrow 0$. Заменяя $\cos$ малого аргумента единицей, а $\sin$ малого аргумента — самим аргументом, получим

$$u_{\text{ЧМ}} \approx U_0 \cos \left[\cos \omega_0 t + \frac{\beta}{2} \cos(\omega_0 + \Omega)t - \frac{\beta}{2} \cos(\omega_0 - \Omega)t \right], \quad (19)$$

т.е. колебание, не отличающееся по спектральному составу от АМ колебания и состоящее из несущей частоты и двух спутников. Таким образом, при $\beta \rightarrow 0$ каждая боковая полоса состоит из одной спектральной линии, т.е. $n = 1$.

Если теперь при помощи графика рис. 12 или таблиц бесселевых функций найти n для любого β , то получится зависимость, показанная на рис. 13. Асимптотами кривой являются две прямые $n = 1$ (для $\beta \rightarrow 0$) и $n = \beta$ (для $\beta \rightarrow \infty$). Зависимость между n и β с достаточной для практики точностью может быть выражена простой формулой

$$n = 1 + \beta.$$

Заметим, что прямая $n = 1$ выражает соотношение для АМ (при синусоидальной модуляции). Таким образом, ширина спектра при ЧМ всегда больше, чем ширина спектра при АМ.

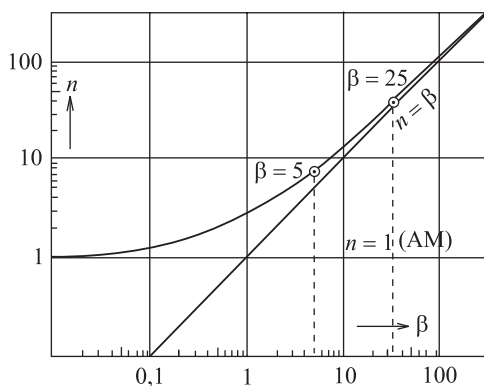


Рис. 13

Теперь мы можем отметить различие между спектрами АМ и ЧМ колебаний. При АМ ширина спектра модулированного колебания равна удвоенной ширине спектра модулирующей функции (точнее, удвоенной высшей частоте этого спектра). При ЧМ с большим индексом ширина спектра модулированного колебания не зависит от спектра модулирующей функции и определяется полосой качания. С другой стороны, при АМ ширина спектра модулированного колебания не зависит от интенсивности модулирующего низкочастотного сигнала, а при ЧМ ширина спектра прямо пропорциональна амплитуде модулирующего колебания. Иначе говоря, ширина спектра при АМ не зависит от глубины модуляции (изменяются только амплитуды спутников), а при ЧМ ширина спектра зависит от глубины модуляции. Мощность ЧМ колебания остается постоянной и не зависит от модуляции (так как остается постоянной амплитуда колебания).

На рис. 14 изображены спектры ЧМ колебания для двух разных значений модулирующей частоты Ω , т.е. для двух значений индекса

модуляции β . Эти значения отмечены на рис. 13 кружками. Рисунок показывает, насколько при данных значениях β действительная ширина спектра $2n\Omega$ отличается от полосы качания $2\Delta\omega$.

На практике применяются большие значения β . Так, для вещания принята величина частотного отклонения 75 кГц. Для $\Omega/2 = 5$ кГц получим $\beta = 15$. Ширина спектра 150 кГц, т. е. в 15 раз больше, чем при АМ. Поэтому ЧМ нельзя (из-за «тесноты в эфире») применять на средних и даже на коротких волнах; ЧМ-вещание ведется лишь на укв. Важное преимущество ЧМ перед АМ состоит в большей помехоустойчивости. Об этом будет речь позднее.

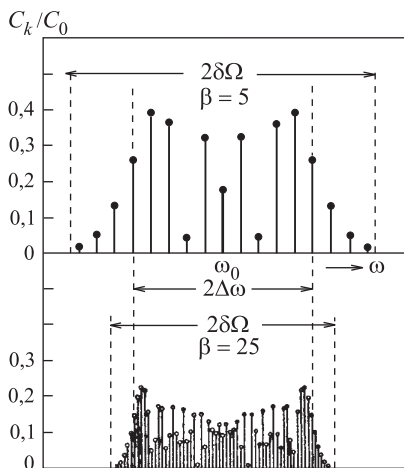


Рис. 14

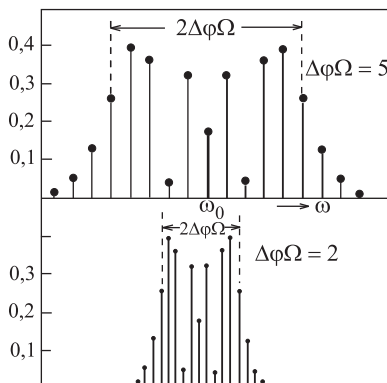


Рис. 15

Остается рассмотреть спектр ФМ колебания. Соотношения для ФМ и ЧМ различаются только тем, что в общее выражение для ФМ входит сама модулирующая функция, а для ЧМ — ее интеграл (см. (10) и (14)).

Таким образом, при синусоидальной модуляции $x(t) = \sin \Omega t$ будем иметь для ФМ

$$u_{\text{ФМ}} = U_0 \cos(\omega_0 t + \Delta\varphi \sin \Omega t). \quad (20)$$

Это выражение отличается от (17) только тем, что вместо $\Delta\omega/\Omega$ в (20) стоит постоянная величина $\Delta\varphi$ — индекс ФМ.

Ширина боковой полосы определится, как и раньше, произведением индекса на модулирующую частоту.

Следовательно, полная ширина спектра ФМ колебания равна $2\Delta\varphi\Omega$. Она зависит от модулирующей частоты, и в этом и состоит различие спектров ФМ и ЧМ. Число спектральных линий остается при изменении модулирующей частоты неизменным, но интервал между ними (равный Ω) изменяется; за счет этого и изменяется общая ширина

спектра. На рис. 15 показаны спектры ФМ колебания для двух разных значений модулирующей частоты Ω .

§ 6. Векторные диаграммы модуляции

Можно составить наглядное представление о модуляции путем построения векторных диаграмм. Возьмем выражение для АМ при синусоидальной модуляции (см. (16))

$$u_{\text{АМ}} = U_0 \left[\cos \omega_0 t + \frac{m}{2} \cos (\omega_0 + \Omega) t + \frac{m}{2} \cos (\omega_0 - \Omega) t \right]. \quad (21)$$

Модулированное колебание представляется суммой трех векторов. Вектор несущей частоты имеет длину U_0 и вращается с угловой скоростью ω_0 . Векторы спутников имеют длину $\frac{m}{2}U_0$ и вращаются один с угловой скоростью $\omega_0 + \Omega$, а другой со скоростью $\omega_0 - \Omega$. Следовательно, векторы спутников вращаются с угловой скоростью Ω в противоположные стороны *относительно вектора несущей частоты*. Все это показано на рис. 16. АМ колебание выражается суммой всех трех

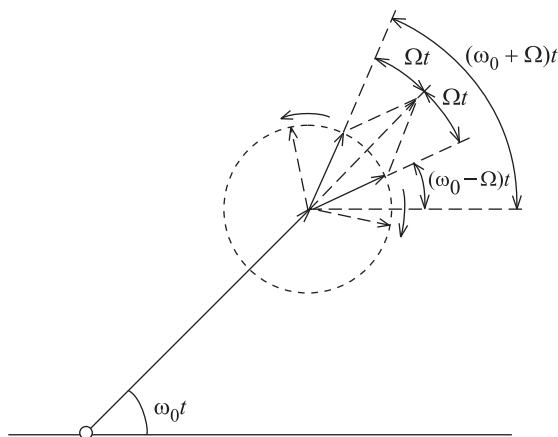


Рис. 16

векторов. Векторы спутников расположены симметрично относительно вектора несущей частоты; поэтому их сумма дает вектор, совпадающий по направлению с вектором несущей частоты. Добавление этой суммы изменяет, следовательно, длину, но не направление результирующего вектора; результирующий вектор вращается по-прежнему с неизменной угловой скоростью ω_0 , но длина его периодически изменяется по мере вращения векторов спутников. Это означает, что мы имеем колебание с неизменной частотой, но с изменяющейся амплитудой, а это и есть АМ.

Построение векторной диаграммы для ЧМ колебания затрудняется тем, что даже при простейшей синусоидальной модуляции получается богатый спектр, следовательно, нужно суммировать большое число (а именно $2\beta + 1$) векторов. Это затруднение мы обойдем, взяв вырожденный случай ЧМ при очень малом индексе. В этом случае ЧМ колебание (так же как и в случае АМ) имеет только три составляющие (см. (19))

$$u_{\text{ЧМ}} \approx U_0 \left[\cos \omega_0 t + \frac{\beta}{2} \cos (\omega_0 + \Omega) t - \frac{\beta}{2} \cos (\omega_0 - \Omega) t \right]. \quad (22)$$

Различие между (21) и (22) на первый взгляд невелико: оно состоит только в различии знаков перед последним членом. Но это различие имеет решающее значение.

Мы начинаем построение точно так же, как и раньше (рис. 17). Строим вектор несущей частоты длиной U_0 под углом $\omega_0 t$ к оси отсчета.

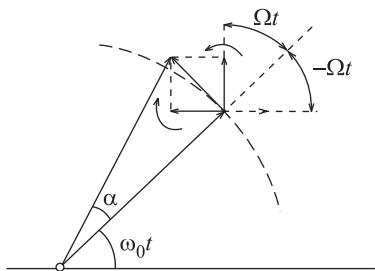


Рис. 17

Строим вектор первого спутника под углом Ωt к вектору несущей частоты. Что же касается второго спутника, то знак минус указывает на то, что он имеет обратную фазу, а вектор его — противоположное направление по отношению к направлению, которое он занимал в случае АМ (и которое намечено на рис. 17 пунктиром). Сумма векторов спутников образует теперь вектор, *перпендикулярный* к вектору основной частоты. Его длина изменяется по

мере вращения векторов спутников в противоположных направлениях (отмеченных на рисунке стрелками). Результирующий вектор при этом *качается* относительно вектора составляющей угловой скорости, т. е. кроме постоянной составляющей имеется еще и переменная составляющая угловой скорости. Это и есть частотная модуляция.

Длина результирующего вектора, представляющая амплитуду модулированного колебания, согласно нашему построению изменяется (так как она есть гипотенуза прямоугольного треугольника, в котором вектор несущей частоты является катетом). Но не следует забывать, что мы рассматриваем предельный случай очень малого индекса β . Поэтому угол α , наибольшее значение которого равно

$$\alpha = \arctg \beta \approx \beta,$$

также мал и длина результирующего вектора в этом приближении остается неизменной. Если же мы пожелаем исследовать тем же методом общий случай любого β , то придется складывать уже не три, а соответственно большее число векторов. При этом окажется, что конец результирующего вектора при его качании перемещается по дуге

окружности (намеченной пунктиром на рис. 17), т. е. длина результирующего вектора не меняется, как оно и должно быть при чистой ЧМ.

Векторная диаграмма ФМ не отличается от векторной диаграммы ЧМ. Нужно лишь иметь в виду, что ФМ определяется угловым отклонением результирующего вектора от положения вектора несущей частоты, а ЧМ скоростью этого отклонения, т. е. производной фазы по времени.

§ 7. Импульсная модуляция

В предыдущих параграфах мы рассматривали три вида модуляции: АМ, ЧМ и ФМ. Сущность этих видов модуляции состоит в том, что подвергается изменению один из трех параметров синусоидального колебания несущей частоты: амплитуда, частота или фаза.

Синусоидальное колебание высокой несущей частоты играет роль *переносчика*: передаваемый низкочастотный сигнал заложен в изменениях того или иного параметра переносчика. В телеграфии для этой цели применяется постоянный ток. Он характеризуется двумя параметрами, величиной и направлением. Изменяя либо то, либо другое, можно образовать телеграфный сигнал.

В качестве переносчика сигнала используют не только постоянный или синусоидальный ток, но и периодическую последовательность импульсов. Такого рода последовательность изображена на рис. 18, а.

Она характеризуется следующими параметрами (если отбросить параметры, определяющие форму импульсов):

1. Высота («амплитуда») импульсов h .
2. Длительность импульсов τ .
3. Частота следования импульсов $\omega_0 = 2\pi/T$.
4. Положение импульсов во времени относительно положения импульсов немодулированной последовательности — так называемая фаза импульсов.

Изменяя один из перечисленных параметров в соответствии с изменениями модулирующей функции, можно получить четыре основных вида *импульсной модуляции*, а именно:

1. Амплитудно-импульсная модуляция — АИМ.
2. Модуляция импульсов по длительности — ДИМ.
3. Частотно-импульсная модуляция — ЧИМ.
4. Фазо-импульсная модуляция — ФИМ.

На рис. 18, б изображен передаваемый сигнал — для примера взят телеграфный сигнал. На рис. 18, в, г, д и е показаны сигналы при различных видах импульсной модуляции. Вертикальными пунктирными линиями отмечены положения немодулированных импульсов. Соответствующие отметки на оси времени называются *тактовыми* точками. При ФИМ происходит смещение импульсов относительно тактовых точек. Предполагается, что при всех видах ИМ изменения того или иного параметра пропорциональны значениям модулирующей функции.

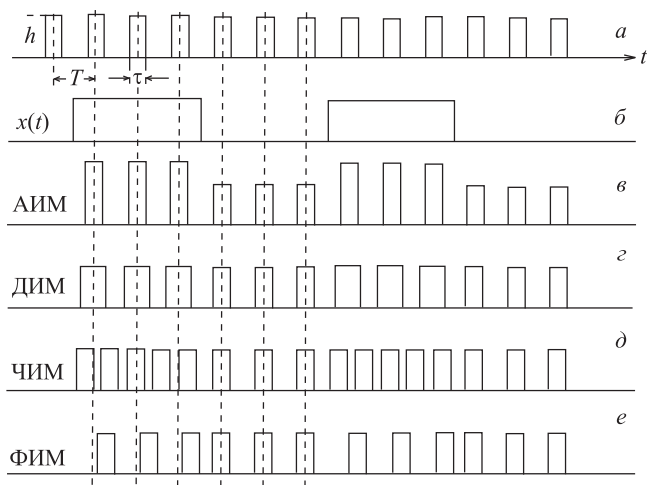


Рис. 18

Можно различать более мелкие разновидности ИМ. Так, например, на рис. 18, г изображен сигнал ДИМ, характеризующийся тем, что при изменяющейся длительности импульсы располагаются симметрично относительно тактовых точек. Но, видоизменив способ модуляции, можно получить изменение длительности за счет смещения только одного фронта импульса — переднего или заднего. Такого рода варианты не меняют сущности дела.

Спектр для всех видов ИМ определяется тем, что переносчик — последовательность импульсов — несинусоиден и имеет линейчатый спектр (рис. 19, а). При модуляции около каждой спектральной линии переносчика появляются боковые полосы (рис. 19, б). Строение полос зависит от вида модуляции; мы не будем вдаваться в подробности этого довольно специального вопроса.

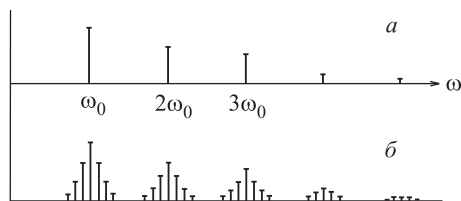


Рис. 19

Частота следования в современных импульсных системах не очень велика. Например, для телефонии применяется частота следования около 10 кГц (основания для такого выбора обсуждаются ниже, в § 9).

Поэтому для радиопередачи производится повторная модуляция: модулированными импульсами модулируется синусоидальное колеба-

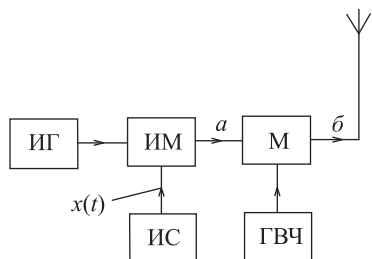


Рис. 20

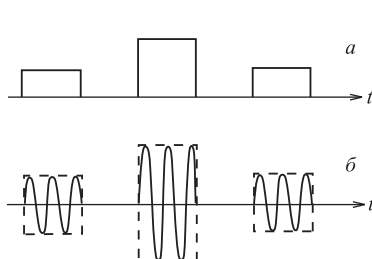


Рис. 21

ние высокой радиочастоты. Блок-схема передающей части системы импульсной радиосвязи (рис. 20) состоит из ИС — источника низкочастотного сигнала, ИГ — импульсного генератора, вырабатывающего периодическую последовательность импульсов, ИМ — импульсного модулятора, в котором осуществляется тот или иной вид импульсной модуляции, ГВЧ — генератора высокой несущей частоты, М — модулятора. Модулированное напряжение с выхода модулятора подводится к антенне. Выходная ступень передатчика на схеме рис. 20 не показана.

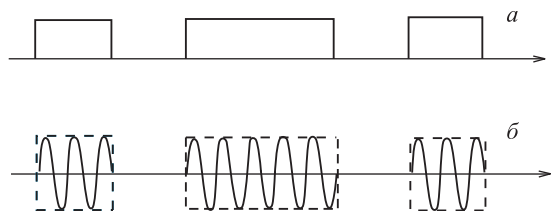


Рис. 22

В пояснение сказанного на рис. 21 изображены сигналы на различных этапах преобразования для АИМ, а на рис. 22 — то же самое для ДИМ. Буквой *а* обозначен сигнал на выходе импульсного модулятора, буквой *б* — сигнал на выходе высокочастотного модулятора. В литературе сигнал вида *а* иногда называют видеоимпульсами, а сигнал вида *б* — радиоимпульсами.

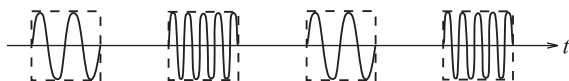


Рис. 23

При повторной модуляции возникает возможность еще одного вида импульсной модуляции, а именно модуляции высокой несущей частоты. Этот вид модуляции называется высокочастотной импульсной модуляцией — ВЧИМ. Он характеризуется тем, что ни высота, ни длительность, ни фаза, ни частота следования импульсов не меняются; таким

образом, огибающая высокочастотных импульсов остается неизменной. Меняется лишь частота заполнения, как показано на рис. 23.

§ 8. Код; общие понятия

В телеграфии для передачи сообщений применяются телеграфные коды. Суть дела состоит в том, что каждой букве или знаку соответствует определенная комбинация электрических импульсов. Совокупность этих комбинаций и образует код. Наиболее известны коды Морзе и Бодо.

Всякий телеграфный код строится из элементарных сигналов, настолько различающихся друг от друга, что на приемном конце системы связи их можно уверенно распознать. Различные элементарные сигналы мы будем называть *элементами кода*. В коде Морзе элементами являются: «точка» — короткая посылка тока, «тире» — вдвое более длинная посылка и пауза. В коде Бодо элементами являются посылка и отсутствие посылки или посылки одинаковой длительности и абсолютной величины, но разной полярности.

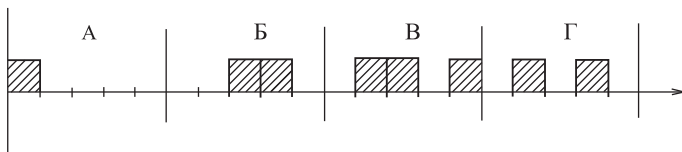


Рис. 24

Общее число элементов в кодовой комбинации называют *числом знаков*. Все комбинации кода Бодо имеют одинаковое число знаков, а именно пять. Такой код называется *равномерным*. Несколько комбинаций кода Бодо представлены на рис. 24; график показывает изменение тока во времени. Всего таких комбинаций можно составить $2^5 = 32$, что соответствует числу букв в алфавите. Но так как, кроме букв, нужно передавать еще цифры и знаки, то в аппаратах Бодо применяется второй регистр, содержащий такое же число комбинаций.

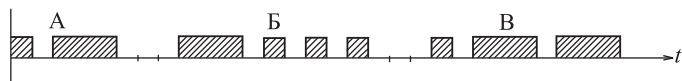


Рис. 25

Код Морзе является *неравномерным* — его комбинации содержат различное число знаков. Несколько комбинаций кода Морзе показаны на рис. 25. Очевидно, что для сокращения времени передачи нужно присвоить более короткие кодовые комбинации часто встречающимся буквам и наоборот. Именно по такому принципу и построен код Морзе: он основан на статистике букв в английском языке. Буква Е встречается чаще всего, и ей присвоено самое короткое кодовое обозначение —

одна точка. Следующая по частоте появления буква Т обозначается одним тире и так далее. Статистика букв в русском языке иная, поэтому можно было бы улучшить код Морзе, переставив обозначения букв в соответствии с частотой их появления. Однако расчеты показывают, что полученная от этого экономия была бы незначительна (около 8%).

Теперь мы поставим вопрос о кодах в несколько более общем виде. При телеграфной передаче передается, собственно говоря, не буква (сообщение), а отображающая букву кодовая комбинация (сигнал). Можно полагать, что кодовая комбинация изображает порядковый номер буквы или вообще некоторое условно приписанное данной букве число. При такой постановке вопроса кодовая комбинация отождествляется с некоторым числом, а элементы кода — с цифрами, посредством которых записывается число. Здесь придется рассмотреть различные системы записи чисел — системы счисления.

Десятичная система счисления располагает десятью цифрами — от 0 до 9. Одной цифрой можно записать любое число, не превосходящее 9. Число 10 будет уже двузначным. Запись 10 можно прочесть так: один десяток плюс 0 единиц. Такая система называется разрядной. Любое число представляется в следующем виде:

$$N = \alpha \cdot 10^0 + \beta \cdot 10^1 + \gamma \cdot 10^2 + \delta \cdot 10^3 + \dots,$$

где $\alpha, \beta, \dots$ — числа от 0 до 9. Таким образом, например, число 5379 содержит 5 тысяч, 3 сотни, 7 десятков и 9 единиц, т. е.

$$5379 = 9 \cdot 10^0 + 7 \cdot 10^1 + 3 \cdot 10^2 + 5 \cdot 10^3.$$

Можно сказать и так: число 5379 содержит 5 единиц в разряде тысяч, 3 единицы в разряде сотен, 7 единиц в разряде десятков и 9 единиц в разряде единиц.

Число в десятичной системе представляется суммой по степеням числа 10, которое называется *основанием* системы счисления. Но в качестве основания можно взять *любое* число m . Тогда для записи некоторого числа N в системе счисления с основанием m будем иметь

$$N = am^0 + bm^1 + cm^2 + dm^3 + \dots,$$

т. е. сумму по степеням числа m , где $a, b, \dots$ — числа, не превосходящие $m - 1$. Рассмотрим, к примеру, случай $m = 3$. Система счисления с основанием 3 называется троичной. Она располагает только тремя цифрами. Обозначения этих цифр произвольны; пусть это будет 0, 1 и 2. Заметим, что в этой системе число 3 будет уже двузначным: оно запишется единицей в разряде троек и нулем в разряде единиц. Запись чисел по троичной системе имеет следующий вид:

десятичная система	...	1	2	3	4	5	6	7	8	9	10	...
троичная система	...	1	2	10	11	12	20	21	22	100	101	...

Теперь нужно связать строение кода с системой счисления. Уже говорилось, что элементы кода отождествляются с цифрами. Число

элементов — число цифр — определяет основание системы счисления. Таким образом, кодовую комбинацию можно рассматривать как запись числа по соответствующей системе счисления. Вернемся к коду Бодо (рис. 24). Этот код строится из двух элементов — посылка и отсутствие посылки (пауза). Обозначим посылку цифрой 1, а паузу — цифрой 0. При таком обозначении можно записать комбинации кода Бодо так:

Буква	...	А	Б	В	Г	Д	Е	Ж	...
Кодовая комбинация	...	10000	00110	01101	01010	11110	01000	00011	...

Последняя строка представляет запись пятизначных чисел в системе с основанием два, так называемой *двоичной* системе. Число элементов кода будем называть по аналогии *основанием* кода. Код Бодо строится из двух элементов, имеет основание два и называется двоичным кодом.

Код Морзе также строится из двух элементов — точек и тире, но он не является двоичным. Дело в том, что код Морзе является неравномерным. Поэтому мы не можем знать, где кончается одна кодовая комбинация и начинается следующая, и возникает надобность в специальном *разделительном* знаке.

В качестве такого знака употребляется пауза между буквами. Она играет роль третьего элемента кода, который, таким образом, становится троичным.

Если обозначить точку — 1, тире — 2 и паузу — 0, то, например, слово «Москва» при записи по коду Морзе выглядит так:

22	0	222	0	111	0	212	0	122	0	12
М	—	О	—	С	—	К	—	В	—	А

Двоичный код имеет в технике особое значение. Это объясняется отчасти тем, что такие элементы кода, как отсутствие посылки и наличие посылки, наиболее уверенно различимы, отчасти же тем, что легко осуществить устройства, действие которых характеризуется тем, что они находятся в одном из *двух* возможных состояний (например, всякого рода реле, триггеры и т.п.). Поэтому двоичный код находит преимущественное применение не только в связи, но и в электронных вычислительных машинах, в автоматике и других отраслях техники, имеющих дело с передачей электрических сигналов.

Мы рассматривали до сих пор коды и соответствующие системы счисления с основанием, меньшим десяти, в частности, двоичные и троичные. Но с таким же успехом мы могли бы построить код с любым основанием больше десяти. Предположим, что мы желаем построить *однозначный* код, при помощи которого можно было бы передать любую из 32 букв. Для этого потребуется код с основанием 32, т. е. код, имеющий 32 элемента. В качестве элементов кода можно, например, взять импульсы различной высоты. Каждая буква будет передаваться только одним таким импульсом (это и есть однозначный код). Нужно, конечно, позаботиться о том, чтобы различие между импульсами было

достаточно велико для того, чтобы не произошло ошибки при приеме. Ясно, что это приведет к увеличению мощности сигнала (как пиковой, так и средней), зато, передавая для каждой буквы вместо кодовой комбинации только один импульс, можно соответственно ускорить передачу (или сократить ширину спектра — об этом говорится ниже).

Общее число комбинаций для равномерного кода можно подсчитать по формуле

$$N = m^n,$$

где m — основание кода; n — число знаков в комбинации.

Для неравномерного кода нужно просуммировать число однозначных, двузначных и т. д. комбинаций.

Для кода Бодо (двоичный, пятизначный код) имеем

$$N = 2^5 = 32.$$

Для однозначного кода с основанием 32, о котором только что говорилось

$$N = 32^1 = 32.$$

§ 9. Теорема Котельникова

При телеграфии передаваемое сообщение состоит из отдельных символов — букв, знаков или цифр. Такого рода сообщения называются *дискретными* и могут быть переданы при помощи соответствующего кода, в котором каждому символу сообщения соответствует определенная кодовая комбинация.

Но в технике приходится иметь дело и с *непрерывными* сообщениями, когда нужно передать значения некоторой непрерывно изменяющейся величины. Так обстоит дело, например, при передаче звука или изображения. При передаче звука мы получаем на выходе микрофона непрерывно изменяющееся напряжение, воспроизводящее изменения звукового давления. При передаче изображения (например, в случае фототелеграфа) оно разворачивается движущимся по изображению узким световым лучом, и на выходе фотоэлемента, воспроизводящего отраженный свет, также получается непрерывно изменяющееся напряжение. Аналогично обстоит дело и в телевидении.

Таким образом, во многих случаях нужно передать непрерывную функцию $x(t)$. На первый взгляд кажется, что передача непрерывного сообщения принципиально отличается от передачи дискретного сообщения, так что и техника передачи должна быть существенно иной; в частности, представляется, что передача непрерывного сообщения с помощью того или иного кода невозможна. Это, однако, не так. Дело в том, что все непрерывные функции, которые нужно передавать, представляют собой функции с ограниченным спектром, о чем уже говорилось в § 3. А для функций с ограниченным спектром справедлива следующая *теорема Котельникова*.

Функция с ограниченным спектром полностью определяется своими значениями, отсчитанными через интервалы $\Delta t = \frac{1}{2F}$, где F — ширина спектра функции.

В частности, если спектр функции отсчитывается от нулевой частоты, то F есть просто верхняя граничная частота спектра.

Смысл теоремы Котельникова состоит в том, что если требуется передать непрерывную функцию $x(t)$ с ограниченным спектром, то не нужно передавать все значения функции; достаточно передать отдельные мгновенные значения, отсчитанные через Δt . Так как функция полностью определяется этими значениями, то по ним непрерывная функция $x(t)$ может быть восстановлена на приемном конце системы связи.

Поясним смысл теоремы Котельникова на графике. На рис. 26 приведен график некоторой функции $x(t)$. Предполагается, что спектр ее ограничен сверху частотой ω_c . Произведем отсчеты мгновенных значений функции $x(t)$ в моменты $\Delta t = \pi/\omega_c$, $2\Delta t$, $3\Delta t$, ... и т. д.

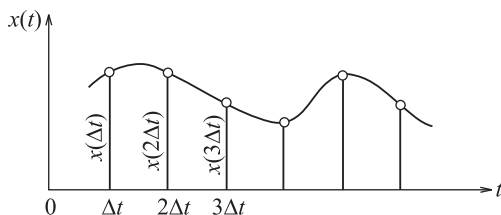


Рис. 26

Теорема утверждает, что задание отмеченных на рисунке точек полностью определяет кривую, т. е. что через эти точки кривая может быть проведена *единственным образом*. Это не должно казаться странным: ведь быстрые изменения функции на интервале между соседними точками возможны только в том случае, если в составе функции имеются достаточно высокие частоты. Все дело в том и заключается, что отсчеты берутся *достаточно часто* и тем чаще, чем шире спектр функции. То, что функция определяется полностью, если отсчеты следуют друг за другом через интервалы $\Delta t = \pi/\omega_c$, можно показать следующим рассуждением. Возьмем отрезок функции длительностью T и разложим функцию на этом интервале в ряд Фурье ¹⁾. Так как функция имеет ограниченный спектр, то разложение будет конечным и мы можем записать (постоянная составляющая отброшена)

$$x(t) = \sum_{k=1}^n c_k \cos(k\omega_1 t - \varphi_k), \quad (23)$$

¹⁾ Для придания строгости нижеследующему рассуждению нужно рассматривать предельный переход при $T \rightarrow \infty$.

где

$$\omega_1 = \frac{2\pi}{T}, \quad n = \frac{\omega_c}{\omega_1}$$

(n — наивысший номер гармоники в разложении (23)).

Таким образом, функция $x(t)$ полностью определяется n амплитудами c_k и n фазами φ_k гармоник, а всего $2n$ числами. Эти числа, если передавать их через равные интервалы, будут следовать друг за другом через

$$\Delta t = \frac{T}{2n} = \frac{2\pi}{\omega_1} \frac{\omega_1}{2\omega_c} = \frac{\pi}{\omega_c}.$$

Обычно теорема Котельникова доказывается путем разложения функции с ограниченным спектром в ряд следующего вида:

$$x(t) = \sum_{k=1}^{\infty} x(k\Delta t) \frac{\sin \omega_c(t - k\Delta t)}{\omega_c(t - k\Delta t)}. \quad (24)$$

Мы не будем делать выкладки, приводящие к этому разложению, но поясним его физический смысл.

Величины $x(k\Delta t)$, т.е. $x(\Delta t)$, $x(2\Delta t)$, ... — это и есть отсчеты функции $x(t)$ в моменты Δt , $2\Delta t$, ..., изображаемые соответствующими ординатами на рис. 25.

Функция $\frac{\sin z}{z}$ равна единице при $z = 0$ и убывает, колеблясь, при возрастании абсолютного значения z (функция, как легко видеть, четна). График функции $\frac{\sin z}{z}$ приведен на рис. 27. Эта функция представляет изменение во времени напряжения на выходе идеального фильтра, пропускающего все частоты от нуля до граничной частоты ω_c , когда на вход фильтра подается напряжение в форме очень короткого импульса.

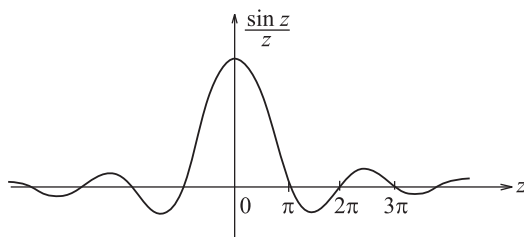


Рис. 27

На рис. 28 сверху изображена некоторая функция $x(t)$ с ограниченным спектром. В точках $k\Delta t$ взяты отсчеты $x(k\Delta t)$ — они представлены соответствующими ординатами. На последующих строках рис. 28 изображены отдельные слагаемые ряда (24). Они представляются функциями такого же вида, как на рис. 27, но сдвинутыми друг

относительно друга на Δt ; максимальные ординаты равны $x(k\Delta t)$. Если теперь сложить все эти составляющие, то получится снова исходная функция $x(t)$. Заметим, что в точке $t = k\Delta t$ k -я составляющая равна $x(k\Delta t)$, а все остальные составляющие равны нулю.

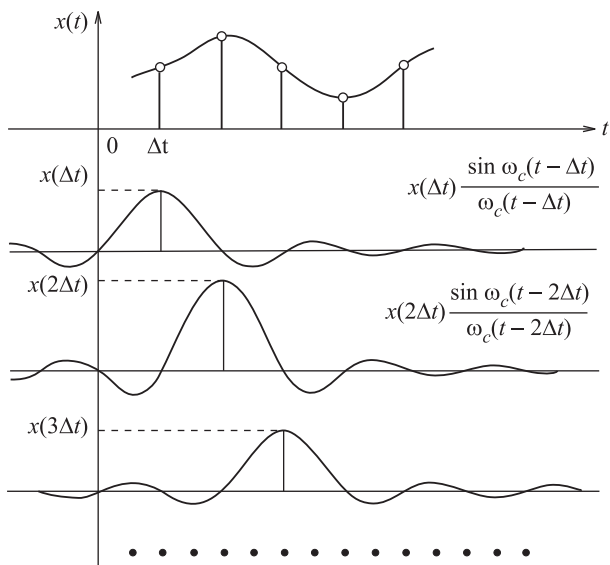


Рис. 28

Разложение (24) указывает технический способ передачи функции $x(t)$ с ограниченным спектром и восстановления ее на приемном конце. Способ этот состоит из следующих операций:

1. Берутся отсчеты $x(k\Delta t)$ функции $x(t)$ в моменты $k\Delta t$.
2. Полученные числа передаются любым способом по системе связи.
3. На приемной стороне вырабатываются короткие импульсы соответствующей высоты.
4. Эти импульсы подаются на вход идеального фильтра с верхней границей пропускания ω_c . На выходе фильтра получается исходная функция $x(t)$.

На практике спектр передаваемых функций ограничен нерезко. Кроме того, нужно учесть, что идеальный фильтр, не пропускающий вовсе частот выше граничной, физически неосуществим. Если принять все это во внимание, то окажется, что функция может быть восстановлена по своим отдельным отсчетам не точно, а приближенно. Это, конечно, усложняет соотношения, но не умалает значения теоремы Котельникова для теории и техники связи. На основе этой теоремы передача непрерывного сообщения сводится к точно такой же ситуации, как и передача дискретной последовательности некоторых чисел.

Таким образом, теорема Котельникова лежит в основе всей импульсной связи; она показывает, при каких условиях передача непрерывной функции может быть сведена к передаче отдельных импульсов или кодовых комбинаций.

Необходимая частота следования импульсов, называемая также *тактовой частотой*, определяется по формуле

$$f_0 = \frac{1}{\Delta t} = 2f_c,$$

где f_c — верхняя граница спектра. Так, например, для телефонной передачи хорошего качества, принимая $f_c = 4$ кГц, находим, что тактовая частота должна составлять 8 кГц, т. е. что должно посылаться 8 тысяч импульсов в секунду.

§ 10. Квантование

При передаче непрерывных функций импульсным методом нужно передавать мгновенные значения функции в моменты отсчетов. Но так как функция изменяется непрерывно, то ее мгновенные значения могут выражаться какими угодно числами. Передачу этих чисел легко осуществить: достаточно, например, передавать последовательность импульсов, высота которых пропорциональна отсчитанным мгновенным значениям передаваемой функции. Это будет обычная АИМ.

Но в импульсной радиосвязи прибегают к особому приему, состоящему в том, что передается только конечное число разрешенных значений, отстоящих друг от друга на конечный интервал. Если действительное мгновенное значение оказывается внутри этого интервала (т. е. принимает запрещенное значение), то оно заменяется ближайшим разрешенным. Эта операция называется *квантованием*, шкала разрешенных значений *шкалой квантования*, интервал между разрешенными значениями *шагом квантования*.

Нужно сразу пояснить, что с математической точки зрения квантование есть не что иное, как округление чисел, к которому мы постоянно прибегаем, не называя его, впрочем, квантованием. Никому не известны, да и не нужны *точные* значения иррациональных и трансцендентных чисел, как $\sqrt{2}$, π , e и т. д., так как точные значения этих чисел должны были бы быть записаны бесконечным числом знаков (по любой системе счисления). Когда мы пишем $\pi = 3,14$ или $\pi = 3,14159$, то в обоих случаях мы даем приближенное, округленное значение числа π . Приближение, разумеется, улучшается с увеличением числа знаков. Правило округления всем хорошо известно: оно состоит в том, что если отбрасываемые знаки составляют больше половины единицы последнего оставляемого знака, то в последнем знаке единица добавляется; в противном случае отбрасываемыми знаками просто пренебрегают.

Возьмем к примеру любую таблицу какой-нибудь трансцендентной функции, например четырехзначную таблицу логарифмов. Пользуясь вышеприведенными терминами, можно сказать, что таблица содержит

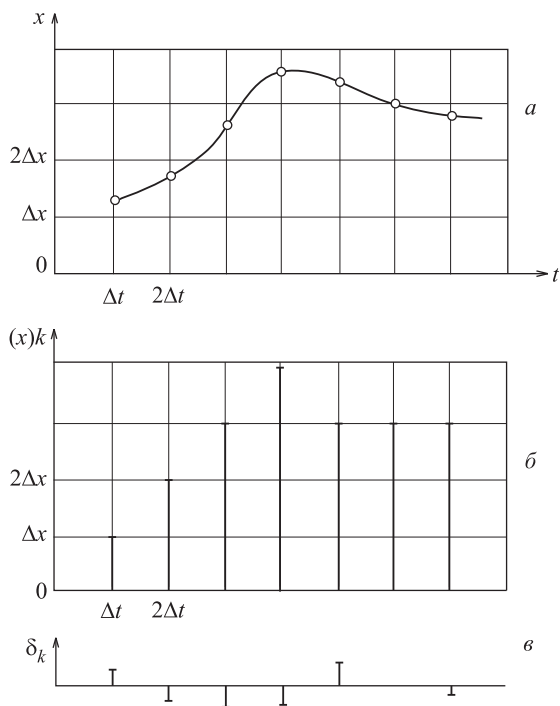


Рис. 29

квантованные значения логарифма; шаг квантования равен единице четвертого знака. В семизначной таблице шаг квантования меньше; семизначная таблица дает более точные значения логарифма.

Вернемся теперь к вопросам связи. На рис. 29, а изображена передаваемая функция $x(t)$ (здесь и в дальнейшем мы имеем в виду только функции с ограниченным спектром). График функции вписан в сетку с ячейками $\Delta x \Delta t$. Отметки на оси времен означают моменты отсчетов. Отметки на оси ординат — оси мгновенных значений функции $x(t)$ — образуют шкалу квантования, Δx — шаг квантования. В простейшем случае шкала квантования равномерна, т.е. имеет постоянный шаг. В этом случае разрешенные значения образуют последовательность $0, \Delta x, 2\Delta x, \dots, i\Delta x, \dots$ (шкала квантования может распространяться и в область отрицательных значений $x(t)$).

Истинные мгновенные значения $x_k = x(k\Delta t)$ (т.е. ординаты кривой $x(t)$ в точках $t = k\Delta t$) обозначены вертикальными отрезками на рис. 29, а. Процесс квантования состоит в том, что значения x_k заменяются ближайшими разрешенными значениями $(x)_k$, последовательность которых показана на рис. 29, б. Разность

$$x_k - (x)_k = \delta_k \quad (25)$$

можно назвать ошибкой квантования. Последовательность величин δ_k изображена на рис. 29, в. Перепишем (25) в виде

$$(x)_k = x_k - \delta_k.$$

Теперь мы можем сказать, что передача квантованных значений $(x)_k$ вместо истинных значений x_k равносильна наложению на истинные значения *помехи* δ_k . Это — удобная точка зрения. Последовательность δ_k так и называют: *помеха квантования* или *шум квантования*.

Применение квантования имеет двойное значение. Во-первых, оно открывает возможность использования импульсно-кодовой модуляции. Об этом будет рассказано в следующем параграфе. Во-вторых, квантование представляет собой мощное средство борьбы со случайными помехами, и на этой стороне дела мы сейчас остановимся.

При обсуждении проблем связи нельзя забывать, что в реальных условиях на сигнал всегда налагается помеха. Поэтому передача *истинных* значений x_k невозможна: на приемном конце мы будем получать не x_k , $y_k = x_k + \xi_k$, где ξ_k — мгновенное значение напряжения помехи в момент отсчета. Помеха имеет случайный характер, никак не контролируется, а потому восстановить истинные значения x_k по принятым значениям y_k нельзя.

Положим теперь, что мы знаем наибольшее значение помехи $\xi_{\text{макс}}$. В таком случае мы можем применить квантование, выбрав его шаг Δx так, чтобы выполнялось неравенство $\Delta x > 2\xi_{\text{макс}}$. При таком условии, получив на приемной стороне квантованные значения с наложенной на них помехой, можно снова проквантовать эти значения и очистить их таким образом от помехи. В самом деле, так как по условию помеха не превосходит половины шага квантования, то для искаженных помехой значений *ближайшим* уровнем остается тот, который передавался ¹⁾.

Таким образом, повторное квантование производит *восстановление* (регенерацию) поврежденного помехой квантованного сигнала. Эту операцию можно повторять любое число раз. Основываясь на этой идее, можно построить систему связи любой протяженности, разбив ее на участки и поместив на стыках участков ретрансляционные приемно-передающие станции. Назначение этих станций состоит в том, чтобы, приняв сигнал, восстановить его и переслать дальше. По этому принципу уже давно строились телеграфные восстанавливающие трансляции и по тому же принципу строятся радиорелейные линии связи.

Нужно пояснить, что квантование не освобождает сигнал от помехи, но заменяет одну помеху другой. Вместо случайной, неконтролируемой помехи мы накладываем на сигнал нами же созданную помеху — шум квантования. По интенсивности эта искусственная

¹⁾ Это рассуждение упрощено. В действительности максимальное значение помехи не ограничено. Поэтому можно говорить не о безошибочном восстановлении переданного сигнала, а лишь о восстановлении с заданной малой вероятностью ошибки.

помеха не меньше естественной, которую мы стремимся устранить. Но преимущество системы связи с квантованием состоит в том, что в ней может быть предотвращено *накопление* помех — это очень большое достоинство.

§ 11. Импульсно-кодовая модуляция

Под *импульсно-кодовой модуляцией* (ИКМ) понимается передача непрерывных функций при помощи двоичного кода. На рис. 30, *а* снова показана подлежащая передаче функция с ограниченным спектром, на рис. 30, *б* — квантованные отсчеты. Если бы мы передавали непосредственно последовательность изображенных на рис. 30, *б* импульсов различной высоты, то это была бы обычная амплитудно-импульсная модуляция. Но дело ведь заключается в том, чтобы любым способом передать числа, выражающие величину квантованных отсчетов функции $x(t)$. Для этого можно воспользоваться любым низшим кодом, в частности двоичным. Числа, подлежащие передаче, надо записать, по двоичной системе счисления — это и даст требуемые кодовые комбинации.

Напомним, что по двоичной системе число представляется в виде

$$N = a \cdot 2^0 + b \cdot 2^1 + c \cdot 2^2 + d \cdot 2^3 + \dots$$

Коэффициенты $a, b, c, \dots$ — однозначные числа, выражаемые одной цифрой. Цифр в двоичной системе только две — скажем, 0 и 1. Двоичная запись чисел приведена в табл. 3.

Таблица 3

Система	Запись чисел										
Десятичная	1	2	3	4	5	6	7	8	9	10	...
Двоичная	1	10	11	100	101	110	111	1000	1001	1010	...

Например, десятичное число $10 = 8 + 2 = 0 \cdot 2^0 + 1 \cdot 2^1 + 0 \cdot 2^2 + 1 \cdot 2^3$ записывается по двоичной системе единицей в разряде восьмерок, нулем в разряде четверок, единицей в разряде двоек и нулем в разряде единиц.

При помощи двузначных двоичных чисел можно записать десятичные числа от 0 до 3 — всего $4 = 2^2$ числа, при помощи трехзначных — числа от 0 до 7 — всего $8 = 2^3$ чисел, при помощи n -значных двоичных чисел можно представить всего 2^n чисел.

Обратимся к рис. 30. Подлежащая передаче последовательность чисел в десятичной записи выглядит так (см. рис. 30, *б*):

5, 7, 9, 5, 2, 3, 6...

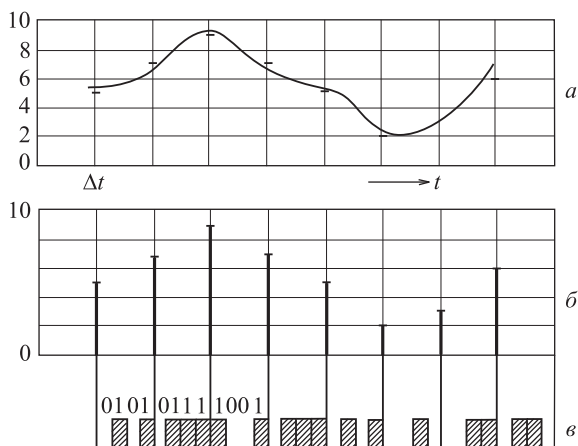


Рис. 30

Трехзначных чисел для записи этой последовательности по двоичной системе не хватит, придется взять четырехзначные числа. Получится последовательность (см. табл. 3)

0101, 0111, 1001, 0111, 0101, 0010, 0011, 0110.

Считая, что 1 означает посылку, а 0 — паузу (можно и наоборот — это совершенно безразлично), получим кодовые комбинации, изображенные на рис. 30, в. В таком виде сигнал и подается в линию или на высокочастотный модулятор.

Теперь становится ясным, какую роль играет во всем этом деле квантование. Благодаря квантованию мы сводим количество различных чисел, подлежащих передаче, до некоторой конечной величины N ; иначе говоря, N есть число разрешенных уровней шкалы квантования.

Если принять шаг квантования за единицу, то $N - 1$ будет означать наибольшее квантованное значение (наименьшее есть нуль). В таком случае можно определить необходимое число знаков в двоичной кодовой комбинации из соотношения:

$$N = 2^n, \quad n = \log_2 N.$$

Если n — не целое, то оно округляется до ближайшего большего целого числа.

Каковы же основания для выбора числа уровней N или числа ступеней квантования? Здесь надо учесть противоречивые соображения. С одной стороны, ясно, что чем меньше шаг квантования, т.е. чем больше число ступеней квантования, тем точнее передается данная функция (см. по этому поводу § 10). С другой стороны, увеличение числа ступеней требует удлинения кодовой комбинации, что, конечно, нежелательно. Компромисс находится опытным путем. Так, напри-

мер, для телефонной передачи установлено, что удовлетворительное качество передачи достигается при $N \approx 100$. Отсюда следует, что телефонный разговор можно передавать при помощи ИКМ, применяя семизначный двоичный код (так как $2^7 = 128$).

Мы не рассматриваем здесь аппаратуры ИМ и, в частности, ИКМ. Укажем лишь, что для получения различных видов ИМ разработаны и применяются специальные электронные устройства, очень остроумные по идее, компактные и надежные в работе. Некоторые из этих устройств описаны в части 3.

§ 12. Многоканальная связь

Всякая линия связи представляет собой дорогое сооружение, и потому совершенно естественно стремление наилучшим образом использовать его. Это достигается либо увеличением скорости передачи, либо одновременной передачей по одной линии нескольких независимых сообщений. Все меры, позволяющие сделать это, называют *уплотнением* линии связи.

Само собой разумеется, что передача нескольких независимых сообщений (например, телефонных разговоров) по одной и той же линии связи требует специальных устройств, позволяющих разделить сообщения на приемном конце и направить каждое из них соответствующему получателю. Если это обеспечено, то говорят, что каждое сообщение следует по своему *каналу* связи. Систему связи при наличии одной линии, но многих каналов называют *многоканальной*. Число каналов на одной линии может быть очень велико — оно достигает сотен и даже тысяч.

На рис. 31 изображена блок-схема многоканальной системы связи. На передающем конце n отправителей подают n сообщений. Эти сообщения поступают на передатчики, которые вырабатывают соответствующие сигналы. Сигналы отдельных каналов, которые мы назовем *канальными сигналами*, поступают на *групповую аппаратуру* или *аппаратуру уплотнения*, общую для всех каналов. Здесь происходит выработка *линейного сигнала*. Он представляет собой в простейшем случае смесь или сумму канальных сигналов. Линейный сигнал передается по линии и на приемной стороне воздействует одновременно на n *разделителей*, каждый из которых выделяет свой сигнал, т. е. сигнал данного канала. Выделенный канальный сигнал попадает на приемник и превращается в сообщение, поступающее к получателю.

Самой существенной особенностью многоканальной системы связи является наличие разделительных устройств, способных выделить и пропустить свой канальный сигнал и задержать чужие, т. е. принадлежащие другим каналам. На практике разделение не бывает совершенным: на сигнал данного канала накладываются в той или иной мере сигналы других каналов. Это создает специфические для многоканальной связи помехи — помехи от соседних каналов. Они могут быть

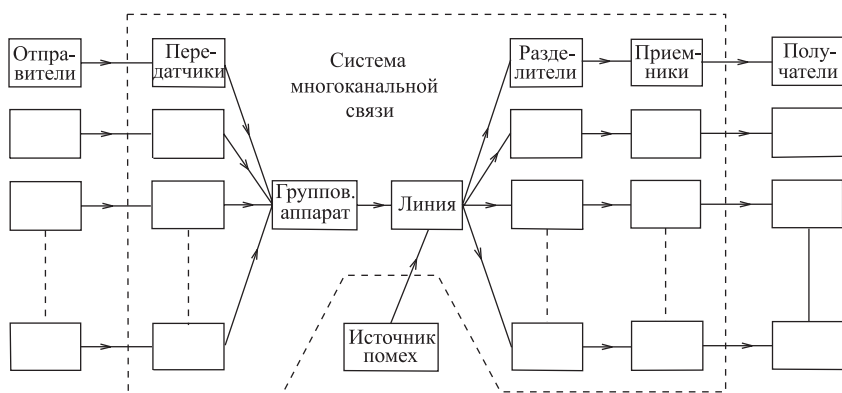


Рис. 31

уменьшены до допустимой величины улучшением системы разделения. Наряду с помехами от соседних каналов в многоканальной системе связи действуют, разумеется, и обычные помехи.

Для того чтобы сигналы могли быть разделены на приемной стороне, необходимо, очевидно, чтобы они различались между собой по некоторому признаку и чтобы разделители могли осуществить разделение на основе этого признака. Отличительный признак канальных сигналов может выбираться по-разному. В зависимости от этого выбора мы получим различные системы разделения сигналов. Преимущественное применение имеют частотный и временной способ разделения. Эти два главных способа рассматриваются в следующем параграфе. Теоретически возможны и другие способы разделения, однако в практике они еще применения не нашли.

§ 13. Частотное и временное разделение

Как уже говорилось выше, все сигналы связи имеют практически ограниченные спектры. Пусть имеется несколько сигналов, спектры которых не перекрываются, т. е. границы спектра одного сигнала лежат вне частотных полос, занимаемых спектрами других сигналов. В таком случае для выделения данного сигнала можно воспользоваться полосовым фильтром. *Идеальный* полосовой фильтр должен был бы пропускать колебания всех частот в пределах полосы пропускания и задерживать колебания всех частот вне полосы пропускания. При этом коэффициент передачи идеального фильтра должен быть постоянным в пределах полосы пропускания, в противном случае сигнал будет искажен.

Вопрос о том, насколько *реальные* фильтры, применяемые в радиосвязи, удовлетворяют этим требованиям, обсуждается подробно во второй части книги.

Для того чтобы разместить спектры нескольких низкочастотных сигналов (например, телефонных) в неперекрывающихся частотных полосах, нужно прибегнуть к модуляции. Несущие частоты должны быть разнесены на интервал, равный ширине спектра модулированного колебания. При обычной АМ интервал между несущими частотами должен равняться удвоенной ширине спектра низкочастотного сигнала.

На рис. 32, *а* показано расположение спектров колебаний, модулированных несколькими различными низкочастотными сигналами с одинаковой шириной спектра. На рис. 32, *б* изображена характеристика идеального фильтра, а на рис. 32, *в* — результат фильтрации, т. е. выделение спектра соответствующего сигнала. Как видим, модуляция является не только способом перемещения спектра сигнала в область высоких частот, но и средством размещения спектров сигналов в заданных частотных полосах в целях организации многоканальной связи.

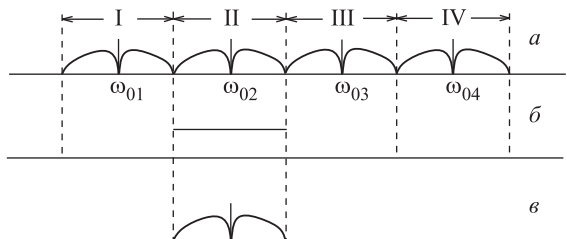


Рис. 32

Канальные сигналы в описанной системе различаются тем, что спектры их расположены в неперекрывающихся частотных полосах. Разделение сигналов осуществляется путем выделения соответствующей полосы при помощи фильтра. Этот способ разделения называется *частотным*.

Систему вещательных станций можно рассматривать как многоканальную систему с частотным разделением. Общей линией связи является «эфир». Разделение производится путем настройки приемника на несущую частоту той или иной станции; при этом приемник должен обеспечивать необходимую полосу пропускания, чтобы пропустить без заметных искажений весь модуляционный спектр данной станции.

При организации многоканальной связи по проводным линиям, а также на радиорелейных линиях пользуются этим же принципом, но часто прибегают к передаче одной боковой полосой. Это позволяет удвоить число каналов, т. е. в большей мере уплотнить линию. В профессиональных устройствах требуемое усложнение аппаратуры не представляет затруднений.

Следует заметить, что реальные фильтры не обладают прямоугольной характеристикой, показанной на рис. 32, *б*. Во-первых; характеристика реального фильтра имеет более пологие скаты, а, во-вторых, реальный фильтр в той или иной мере пропускает частоты, лежащие

вне полосы пропускания. Это влечет за собой возникновение взаимных помех между каналами. Для ослабления этих помех приходится увеличивать разнос несущих частот, оставляя между ними запасный интервал «на расфилтровку». Запасный интервал нужен тем меньший, чем совершеннее фильтры, но при проектировании системы нужно считаться и с тем, что улучшение фильтров связано с их удорожанием.

Итак, система многоканальной связи с частотным разделением строится по общей схеме рис. 31, причем каждый передатчик содержит в себе модулятор и генератор несущей частоты, а разделительные устройства на приемной стороне представляют собой полосовые фильтры.

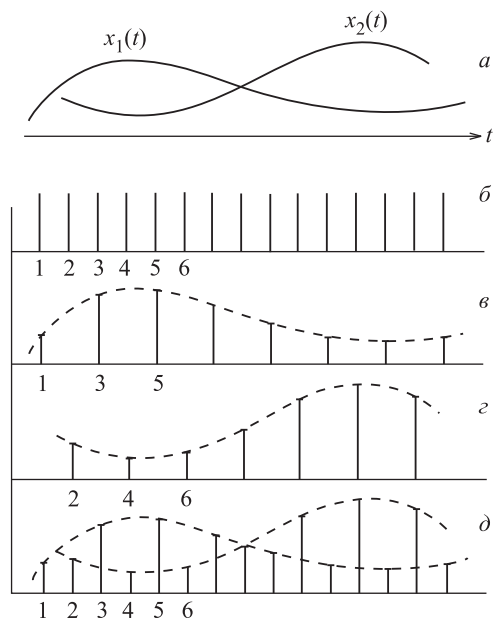


Рис. 33

Обратимся к *временному разделению*. Оно применимо только в импульсной радиосвязи. Дело в том, что при импульсной связи передача ведется при помощи импульсов, длительность которых может быть невелика по сравнению с периодом следования, или, как говорят, импульсы имеют большую *скважность* (скважность определяется отношением периода следования к длительности импульса; наименьшее значение скважности, очевидно, равно единице). При большой скважности между импульсами остается промежуток, на котором можно разместить импульсы других каналов. Эта идея поясняется рис. 33 на примере двух каналов АИМ. На рис. 33, а изображены два низкочастотных сигнала $x_1(t)$ и $x_2(t)$, подлежащие передаче по двум каналам.

На рис. 33,б представлена смодулированная последовательность импульсов; частота следования взята вдвое большей, чем это требуется для одного канала. (При n каналах надо взять частоту следования $f_0 = 2nf_c$, где f_c — верхняя граничная частота низкочастотного спектра.) Для первого канала используются нечетные номера импульсов, для второго — четные. Результат амплитудной модуляции импульсов обоих каналов показан на рис. 33,в и г. Обе модулированные последовательности наложены друг на друга, как показано на рис. 33,д; в таком виде импульсный сигнал направляется в линию или на высокочастотный модулятор. Для разделения канальных сигналов на приемной стороне достаточно применить синхронную коммутацию: некоторый переключатель должен замыкать цепь первого канала в моменты прихода нечетных импульсов и цепь второго канала — в моменты прихода четных импульсов.

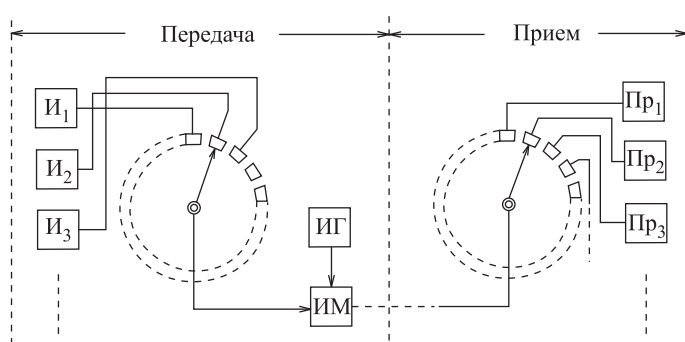


Рис. 34

Система многоканальной связи с временным разделением показана схематически на рис. 34. Основу системы составляют два синхронно работающих коммутатора — один на передающей стороне системы, второй — на приемной. На передающей стороне к ламелям коммутатора подключено несколько независимых источников низкочастотных сигналов И (например, несколько телефонных аппаратов). Вращающаяся щетка коммутатора подключает поочередно каждый из источников на вход импульсного модулятора ИМ, куда поступает также немодулированная последовательность импульсов от импульсного генератора ИГ. Модулированные сигналами всех каналов импульсы поступают на приемный коммутатор — разделитель. К его ламелям подключены приемники Пр (например, также телефонные аппараты), а щетка коммутатора подключает каждый приемник в момент, когда по линии поступает импульс данного канала. Так как все действие системы основано на синхронной работе коммутаторов, то обеспечению синхронности должно быть уделено самое серьезное внимание. Для синхронизации обычно передают вспомогательные синхронизирующие импульсы, для

которых занимают один или более каналов. В рассмотрение синхронизирующих устройств мы здесь вдаваться не будем.

В предыдущем описании коммутаторы были представлены как механические устройства с ламелями и щетками. Это сделано только для наглядности. В действительности применяются коммутаторы электронные. Роль щетки играет вращающийся или качающийся электронный луч. Механические устройства не смогли бы отчетливо работать на требуемых высоких скоростях. Так, например, при десяти телефонных каналах и $f_c = 4$ кГц период коммутации составил бы $\frac{1}{10 \cdot 8 \cdot 10^3} \text{ с} = 12,5 \text{ мкс}$.

Взаимные помехи между каналами при временном разделении могут быть обусловлены несовершенством коммутации, и в первую очередь нарушением синхронности. Но даже при идеальной коммутации взаимные помехи неизбежны вследствие того, что в процессе передачи форма импульсов искажается, импульсы расплываются и частично перекрывают друг друга. Для ослабления взаимных помех приходится вводить запасные (защитные) временные интервалы между импульсами, что ведет к снижению пропускной способности линии.

На одной и той же линии удастся разместить значительно большее число каналов путем применения частотного разделения, нежели при помощи временного разделения.

К достоинствам системы временного разделения следует отнести простоту и компактность системы в целом и разделяющих устройств в частности.

Частотное и временное разделения — основные методы разделения сигналов, применяемые в современной технике связи.

Известны, однако, и другие методы разделения, из которых упомянем здесь только один, а именно метод *фазового разделения*. Суть дела состоит в том, что, применяя в качестве переносчиков для двух каналов два синусоидальных колебания *одной и той же* несущей частоты, но сдвинутых по фазе на 90° , можно разделить каналные сигналы на приемной стороне. Для этого разделительное устройство должно реагировать на фазу несущей частоты. В качестве такого фазочувствительного приемного устройства может применяться синхронный детектор, действие которого описано в третьей части курса. Значение фазового разделения для техники ограничено тем, что на основе этого метода разделения может быть организовано не произвольное число каналов, а только два. Если же фазовое разделение по высокой частоте применяется в комбинации с каким-либо другим методом разделения (т.е. частотным или временным), то возможное число каналов удваивается. Аппаратура для фазового разделения относительно сложна и требует очень строгого соблюдения синхронности.

Глава 3

ПОМЕХИ И ПОМЕХОУСТОЙЧИВОСТЬ

§ 14. Помехи; общее описание

Помехи были выше определены как посторонние электрические возмущения, накладывающиеся на передаваемый сигнал и затрудняющие его прием. При большой интенсивности помех прием становится практически невозможным. Помехи, с которыми приходится иметь дело, весьма разнообразны как по своему происхождению, так и по физическим свойствам.

Говоря о помехах радиоприему, следует упомянуть прежде всего о *помехах от соседних* (по частоте) *станций*. Этот вид помех отличается тем, что его устранение полностью в нашей власти; причина этих помех — плохая техника и организация. Для устранения взаимных помех между станциями необходимо:

- 1) строго придерживаться установленного соответствующими соглашениями расписания волн;
- 2) фильтровать низкочастотный сигнал так, чтобы ширина спектра не превышала половины интервала между несущими частотами;
- 3) устранять гармоники несущей частоты.

Последнее очень важно, так как при несинусоидальной несущей модуляционный спектр принимает вид, как на рис. 19, б. Так, например, сильную помеху телевидению создают гармоники близко расположенных коротковолновых передатчиков.

Следующий вид помех — это помехи от всевозможных электрических устройств, как-то: электрический транспорт, коллекторные электрические двигатели, бытовые приборы (например, пылесосы), электромедицинские приборы (в особенности высокочастотные), промышленные установки высокой частоты (например, высокочастотные печи для плавки и закалки, установки для высокочастотной сушки и т. п.). Очень тяжелые помехи создает система зажигания автомобильных двигателей. Даже простой электрический звонок (с прерывателем) является источником сильной помехи, но, к счастью, кратковременной.

Все перечисленные помехи объединяются общим названием *промышленные помехи*. Сущность этих помех состоит в том, что при

работе перечисленных устройств создается мощное электромагнитное излучение. Либо возникает ударное возбуждение затухающих колебаний при искрообразовании, либо нормальный режим работы устройства состоит в генерировании незатухающих высокочастотных колебаний (как, например, в высокочастотных термических установках). Для борьбы с промышленными помехами нужны следующие меры: предотвращение искрообразования в тех устройствах, где это не является основой действия, или применение фильтров, замыкающих по возможности накоротко цепи для возникших высокочастотных колебаний; в установках высокой частоты — правильное конструирование, сводящее к минимуму излучение, а также экранирование установок. Промышленные помехи — также результат плохой техники и организации; но нужно учитывать особые организационные трудности в этом вопросе. Дело в том, что борьба с промышленными помехами требует широких междуведомственных мероприятий, проводимых на основе законодательства. Необходима и соответствующая инспекция. Все эти мероприятия осуществляются пока лишь частично.

Из числа помех природного происхождения отметим *атмосферные помехи*. Этим термином обозначаются помехи, обусловленные электрической деятельностью атмосферы. Электромагнитное излучение грозовых разрядов известно давно. Во время грозы при вспышке молнии в приемнике слышен сильный треск (в особенности на длинных волнах). Но и другие электрические атмосферные явления создают помехи. Различные формы тихого разряда также сопровождаются излучением. Сильную помеху создает, например, высаживание инея на антенну, так как каждая оседающая на антенну частица может нести на себе значительный электрический заряд. Аналогичный эффект могут вызвать пыльные бури, когда частицы пыли наэлектризовываются при трении о воздух, и т. п.

Мы не умеем пока устранять атмосферные помехи. Но энергия излучения грозовых разрядов сосредоточена главным образом в области низких радиочастот.

Поэтому от этого вида помех практически свободен коротковолновый диапазон, который широко используется для радиовещания и служебных радиосвязей.

В связи с освоением укв диапазона приобрели значение и до некоторой степени изучены помехи, обусловленные излучением внеземных объектов, в частности Солнца. Мощное излучение Солнца не только влияет на ионизацию верхних слоев атмосферы и изменяет таким образом условия связи на коротких волнах; большая доля энергии излучения лежит также в области укв (волн сантиметрового и дециметрового диапазонов) и непосредственно воспринимается в качестве помехи приемниками, работающими в этих диапазонах.

По своему характеру помехи разделяются на импульсные и непрерывные. Это — довольно грубое деление, так как отчетливой границы

между обоими видами помех нет; быстро и беспорядочно следующие друг за другом импульсы сливаются в непрерывный шум.

Особое место в радиосвязи (и технике связи вообще) занимают так называемые *флуктуационные помехи*.

§ 15. Флуктуационные помехи

Помехи могут проникать в систему связи не только извне; они могут также зарождаться внутри самой системы в различных ее звеньях. Такого рода помехи неизбежно сопутствуют основным явлениям, происходящим в аппаратуре. Помехи, обусловленные флуктуациями тех или иных физических величин, носят название флуктуационных.

Флуктуациями в физике называют вообще случайные колебания тех или иных величин около их средних значений. Флуктуации обусловлены статистической природой целого ряда физических величин. Поясним это понятие на примере.

Когда мы говорим о давлении газа на стенки заключающего газ сосуда, то имеем в виду эффект бесчисленных ударов движущихся молекул газа о стенки. Так как число молекул чрезвычайно велико, а массы их малы, то нельзя отметить результат каждого отдельного удара; можно наблюдать и измерять лишь некоторый средний, *статистический* эффект, проявляющийся в наличии действующей на стенку силы. Эту силу, отнесенную к единице площади, называют давлением газа. Если бы в некотором объеме была заключена всего одна (или небольшое число) молекула, то понятие давления потеряло бы смысл. Оно есть понятие статистическое и может применяться лишь к совокупности очень большого числа отдельных объектов.

При наблюдении средних значений величин всегда обнаруживаются случайные отклонения от них; эти отклонения и называются флуктуациями. Если бы мы применили для измерения давления газа прибор, способный регистрировать очень быстрые и небольшие изменения давления, то смогли бы непосредственно наблюдать флуктуации давления. В обычных условиях флуктуации были бы ничтожны: но если откачивать из сосуда газ, то относительная величина флуктуации будет возрастать, пока, наконец, явление естественным образом не распадется на отдельные толчки.

Сходным образом обстоит дело с электрическим током. Когда мы говорим о постоянном токе в 1 мА, то это значит, что через проводник в *среднем* проходит $6,3 \cdot 10^{15}$ электронов в секунду (заряд электрона равен $1,592 \cdot 10^{-19}$ кулона). В отдельные короткие интервалы число электронов в секунду может быть и больше и меньше вышеприведенного среднего значения; таким образом, будут существовать флуктуации постоянного тока, т.е. на постоянный в точном смысле ток будет наложена случайная переменная составляющая. Чем меньше ток, тем больше относительная величина флуктуации. Заметим, что понятие величина тока (так же, как и понятие давления) при известных усло-

виях теряет смысл. Так, например, при «токе» 10^{-20} А через проводник проходит в среднем меньше одного электрона в секунду; ток теряет непрерывный характер и представляет собой перемещение отдельных дискретных зарядов.

Дискретная природа электрического тока проявляется в электронных лампах в виде *дробового эффекта*. Флуктуации анодного тока первой лампы при значительном усилении обнаруживаются на выходе усилителя как характерный шум. Для количественной оценки дробового эффекта служат формулы, дающие значение среднего квадрата переменной составляющей анодного тока, обусловленной флуктуациями. Если весь ток эмиссии попадает на анод, то имеем

$$\overline{I_a^2} = 2e_0IF, \quad (26)$$

где $\overline{I_a^2}$ — средний квадрат переменной составляющей тока;

I — средний ток;

e_0 — заряд электрона;

F — полоса частот.

Формула (26) годится не только для электровакуумных приборов с горячим катодом, но и для фотоэлементов, так как механизм эмиссии (тепловая или фотоэмиссия) с точки зрения дробового эффекта не играет роли.

При наличии пространственного заряда имеем следующие формулы:

$$\begin{aligned} \text{для диода} \quad \overline{I_a^2} &= 0,6 \cdot 4kT_k gF, \\ \text{для триода} \quad \overline{I_a^2} &= a \cdot 4kT_k SF, \end{aligned} \quad (27)$$

где k — постоянная Больцмана, равная $1,37 \cdot 10^{-23}$ (Вт · с)/град;

T_k — абсолютная температура катода;

g — проводимость диода;

S — крутизна триода;

a — коэффициент, зависящий от конструкции триода и лежащий в пределах $0,5 \div 1,2$.

Так, например, для триода получим, взяв $a = 1$, $T_k = 1000$ °K, $S = 4,5$ мА/В, $F = 10^4$ Гц

$$\overline{I_a^2} = 4 \cdot 1,4 \cdot 10^{-23} \cdot 10^3 \cdot 4,5 \cdot 10^{-3} \cdot 10^4 \approx 25 \cdot 10^{-19} \text{ А}^2.$$

На нагрузке в 10^4 Ом в анодной цепи лампы эффективное напряжение шума равно

$$U = 10^4 \sqrt{25 \cdot 10^{-19}} \approx 1,6 \cdot 10^{-5} \text{ В} = 16 \text{ мкВ}.$$

Такая величина при неблагоприятных условиях может быть сравнима с полезным сигналом.

Еще большее значение имеют тепловые флуктуации. Мы считаем обычно *сопротивление* постоянным параметром и при том параметром пассивным, т.е. не являющимся электрическим генератором. Но в действительности на сопротивлении всегда имеется переменное на-

пряжение, обусловленное тепловым движением свободных электронов (электронов проводимости) в веществе сопротивления. Таким образом, всякое сопротивление является генератором шума.

Этот случай дает пример флуктуации около нуля: среднее значение напряжения на сопротивлении (при отсутствии постоянного тока) равно нулю, но переменная составляющая присутствует. Средний квадрат переменного флуктуационного напряжения дается формулой Найквиста

$$\overline{E^2} = 4kTRF, \quad (28)$$

где T — абсолютная температура, которую имеет сопротивление R ¹⁾. Формула (28) применима к любому комплексному сопротивлению; в этом случае под R понимается действительная составляющая комплексного сопротивления. Пусть $R = 10^4$ Ом, $T = 300$ °K, $F = 5 \cdot 10^3$ Гц, тогда

$$E = \sqrt{\overline{E^2}} \approx 0,9 \text{ мкВ.}$$

Если сопротивление R включено в цепь сетки первой лампы, то это означает, что уже на сетке первой лампы имеется довольно значительное шумовое напряжение, усиливается затем всеми последующими степенями.

Возвращаясь к дробовому эффекту, заметим, что для количественной его оценки часто вводят эквивалентное сопротивление, т. е. такое сопротивление, которое, будучи включено в цепь сетки, дало бы такой же шум, как и действительный дробовой эффект.

Вместо (28) можем записать

$$\overline{I_a^2} = 4kTRFS^2, \quad (29)$$

и, приравняв (29) и (27), найдем

$$R_{\text{экв}} = a \frac{T_{\text{к}}}{T} \frac{1}{S}.$$

Взяв, к примеру, $T_{\text{к}} = 1000$ °K, $T = 300$ °K, $S = 5$ мА/В, $a = 1$, найдем

$$R_{\text{экв}} = \frac{1000}{300} \frac{10^3}{5} \approx 670 \text{ Ом,}$$

т. е. величину, значительно меньшую, чем действительное значение сопротивления, обычно включаемого в цепь сетки. Такое прямое срав-

¹⁾ Предполагается, что R — постоянная величина. Если же активное сопротивление есть функция частоты $R(f)$, то вместо произведения RF в формулу (28) следует подставить

$$\int_F R(f) df.$$

Это замечание относится ко всем предыдущим формулам, в которые в качестве множителя входит полоса частот F .

нение показывает, что тепловой шум в обычных условиях является источником более интенсивных помех, нежели дробовой эффект.

Список явлений, порождающих флуктуационные помехи, не исчерпывается тепловыми флуктуациями и дробовым эффектом. Сюда следует отнести также флуктуации контактных разностей потенциалов, флуктуации магнитных параметров и другие явления, которые мы рассматривать не будем.

Характерной особенностью флуктуационных помех является то, что явления, порождающие эти помехи, принципиально не могут быть устранены. Они связаны с дискретным строением вещества, с дискретной природой электрического тока. Устранить тепловое движение невозможно; оно прекращается лишь при температуре абсолютного нуля.

Однако возможно ослабить флуктуационные помехи, прибегая к глубокому охлаждению.

По поводу флуктуационной помехи следует еще заметить, что происходящий процесс можно представить себе как беспорядочную последовательность очень малых и очень коротких импульсов, следующих друг за другом через случайные интервалы. Процесс такого рода, разумеется, непериодичен, так как импульсы следуют друг за другом беспорядочно. Тем не менее, можно говорить о спектре такого процесса. Под спектром понимается в данном случае распределение мощности по частотам. Вводят понятие спектральной плотности мощности, определяемой как отношение

$$G = \lim_{\Delta f \rightarrow 0} \frac{\Delta P}{\Delta f},$$

где ΔP — мощность, приходящаяся на полосу частот Δf . Для того чтобы найти мощность, заключенную в полосе частот $f_1 \div f_2$, нужно вычислить интеграл

$$P_{12} = \int_{f_1}^{f_2} G(f) df. \quad (30)$$

Флуктуационная помеха порождается явлениями молекулярных масштабов. Интервалы между импульсами имеют порядок времени пробега молекул, ионов и электронов, а длительность отдельных импульсов — и того меньше. При таких условиях оказывается, что для флуктуационных помех спектральная плотность мощности есть величина постоянная вплоть до очень высоких частот — порядка 10^{12} – 10^{13} Гц, т.е. до частот, намного превышающих самые высокие радиочастоты (частота 10^{12} Гц соответствует длине волны 0,3 мм). Если спектральная плотность в полосе F постоянна и равна G_0 , то для мощности в этой полосе получаем вместо (30)

$$P_F = G_0 F.$$

Этими соотношениями объясняется появление множителя F во всех вышеприведенных формулах для средних квадратов токов и напряжений.

Итак, спектральная плотность мощности флуктуационных помех постоянна во всем диапазоне вплоть до оптических частот ¹⁾. Из этого, между прочим, следует, что от флуктуационной помехи некуда уйти (как можно, например, уйти от грозových разрядов) и что флуктуационную помеху нельзя отфильтровать, не потеряв сигнала.

Вот почему флуктуационные помехи играют такую большую роль в современной технике связи, и вот почему способам борьбы с этими помехами уделяется так много внимания.

§ 16. Шум и замирание

До сих пор мы определяли помеху как стороннее электрическое возмущение, накладывающееся на сигнал. Математически можно представить результат воздействия помехи на сигнал следующим образом:

$$y(t) = x(t) + \xi(t), \quad (31)$$

где x — переданный сигнал;

y — принятый сигнал;

ξ — помеха, представляющая собой случайный процесс со средним значением нуль.

Таким образом, воздействие помехи проявляется в том, что она добавляется к сигналу, *складывается* с ним. Поэтому такого рода помеху называют *аддитивной*. В инженерной практике аддитивную помеху часто называют просто *шумом*.

Существует и другой вид помехи, при которой интенсивность принимаемого сигнала случайным образом изменяется с течением времени. Обычно эти изменения происходят сравнительно медленно. Явление носит название *замирания*. Вред этого типа помехи состоит в том, что при глубоком замирании интенсивность сигнала может стать ниже порога чувствительности приемника и в течение некоторого времени сигнал вовсе не будет приниматься.

Математически эффект замирания представляется соотношением

$$y(t) = \xi(t) x(t).$$

Обозначения здесь те же, что и в (31), но теперь $\xi(t)$ означает положительную, медленно изменяющуюся со временем случайную величину. Итак, при замирании происходит *умножение* сигнала на

¹⁾ Флуктуационную помеху часто называют «белым шумом». Подразумевается некоторая аналогия с белым светом, имеющим сплошной и более или менее равномерный спектр в определенной полосе частот.

некоторый случайный процесс. Поэтому помеху такого рода называют *мультипликативной* ¹⁾.

Всякую систему передачи можно характеризовать коэффициентом передачи, выражающим отношение принятого сигнала к переданному. В реальной системе коэффициент передачи подвержен случайным изменениям, т. е. флуктуациям, так что мультипликативная помеха наблюдается всегда. Однако флуктуации коэффициента передачи относительно малы, тогда как под замиранием понимают обычно глубокие изменения силы сигнала, вплоть до практически полного его пропадания. Здесь мы имеем дело с особым рода явлением интерференционного происхождения.

Дело заключается в том, что при обычной радиопередаче сигнал, посланный передатчиком, может достигать приемника по нескольким различным путям (например, на коротких волнах можно в одном и том же месте принимать сигналы, однократно и многократно отраженные от ионизированных слоев). Различные пути имеют разную длину, а потому колебания, одновременно достигающие приемника, имеют различные фазы. При сложении колебаний с несовпадающими фазами возникает интерференция, и амплитуда результирующего колебания может значительно уменьшиться. Это и есть замирание. Условия распространения не остаются постоянными; соотношение между сигналами, пришедшими по различным возможным путям, все время изменяется случайным образом; соответственно изменяется и интенсивность результирующего сигнала.

Существуют простые и эффективные методы борьбы с замиранием. Они основаны на том общем положении, что вероятность одновременного осуществления нескольких событий меньше, чем вероятность каждого из этих событий в отдельности. Один способ ослабления эффекта замирания состоит в том, что передача ведется не на одной несущей частоте, а на нескольких сразу. Выигрыш при применении этого способа получается вследствие того, что для разных частот фазовые сдвиги получаются различные (да и пути распространения могут различаться). Поэтому маловероятно, чтобы глубокое замирание возникло для всех несущих колебаний одновременно.

Второй способ основан на применении *разнесенных* приемных антенн. При достаточном расстоянии между антеннами маловероятно, чтобы замирание наблюдалось в них одновременно. Разнесенные антенны можно использовать по-разному. Можно просто складывать сигналы, получаемые от каждой из них. Но применяется и другой метод, состоящий в том, что приемники, подключенные к каждой из антенн, коммутируются так, что в каждый момент используется тот из них, который дает наибольший выходной сигнал.

¹⁾ Можно сказать, что замирание представляет собой медленную случайную амплитудную модуляцию.

Как видим, оба упомянутых способа борьбы с замиранием основаны на том, что организуется несколько, по возможности независимых каналов. При первом способе каналы разделяются по частоте, при втором — в пространстве. Независимость каналов обеспечивается достаточным разнесением соответственно по частоте и в пространстве. Мы не останавливаемся здесь на расчете необходимых разносов. Эти расчеты опираются на статистические данные о наблюдаемых в действительности замираниях.

Замирание представляет собой хотя и важное для радиосвязи, но довольно специфическое явление. В дальнейшем мы больше не будем им заниматься. Все последующие сведения и рассуждения о помехах и методах борьбы с ними будут относиться к аддитивной помехе.

§ 17. Понятие помехоустойчивости

Наличие помех затрудняет прием сигналов. При большой интенсивности помех распознавание сигнала может стать практически невозможным. Однако можно строить систему передачи и приема так, чтобы обеспечить достаточно надежную работу системы даже и в очень тяжелых условиях. Способность системы противостоять мешающему воздействию помехи носит общее название *помехоустойчивости*.

Для того чтобы дать качественному представлению о помехоустойчивости определенную количественную меру, рассмотрим положение подробнее.

При передаче всегда имеется некоторое множество возможных сообщений, из которого выбирается передаваемое. Каждому сообщению соответствует определенный сигнал, т. е. множеству сообщений отвечает определенное множество сигналов. Так, например, при передаче текста отдельные сообщения — это буквы; каждой букве соответствует сигнал в форме кодовой комбинации телеграфного кода.

В простейшей форме телеграфный сигнал состоит из посылок и пауз. При наличии помехи в паузе можно обнаружить напряжение, не равное нулю, и ошибочно принять эту паузу за посылку. Точно так же помеха, имеющая напряжение противоположного знака, может превратить при приеме посылку в паузу. В результате будет принят сигнал, соответствующий не той букве, которая была на самом деле передана, а какой-либо другой.

Замена фактически переданного сообщения каким-либо другим возможным является ошибкой при приеме сообщений. Вероятность *p* такого события может служить мерой ненадежности системы связи. Вероятность же противоположного события, т. е. вероятность правильного приема является мерой надежности.

В качестве меры помехоустойчивости можно взять надежность при заданной помехе.

Каковы же общие возможности повышения помехоустойчивости? Прежде всего отметим, что при данной интенсивности помех веро-

ятность правильного приема тем больше, чем сильнее различаются между собой сигналы, соответствующие различным сообщениям. На языке современной теории различие между сигналами называется *расстоянием*.

Таким образом, один из путей к повышению помехоустойчивости состоит в выборе системы сигналов, в которой любая пара возможных сигналов как можно далее отстоит друг от друга.

Затем нужно выбрать такой метод приема, который наилучшим образом реализует различие, существующее между сигналами. Приемник, дающий наибольшую возможную при данных условиях вероятность правильного приема, называется *идеальным*.

И, наконец, нужно указать одну возможность, основанную на использовании статистики сообщений. Суть дела мы поясним на примере. Пусть передается телеграмма, содержащая словесный текст, и пусть отдельные буквы заменены неверными, т.е. приняты с ошибкой. Но ошибки в отдельных *буквах* вовсе не означают невозможности правильно прочесть *слово*, которое эти буквы составляют. Обычно легко можно правильно прочесть телеграмму, несмотря на ошибки в отдельных буквах. Мы основываемся при этом на знании языка. Ошибка в одной букве может образовать такое сочетание букв, которое вообще не является словом в данном языке. Поэтому мы заменяем фактически принятую невозможную комбинацию букв ближайшей возможной, т.е. представляющей собой осмысленное слово. Если это не удастся сделать с полной уверенностью, то мы обращаемся к контексту, определяя данное слово по месту, которое оно занимает в предложении, или, наконец, по общему смыслу телеграммы. Все эти действия человек, восстанавливающий правильный текст телеграммы, принятой с ошибками, производит обычно бессознательно. Но именно на такого рода идеях основаны некоторые современные помехоустойчивые системы передачи сигналов.

Вероятность правильного приема зависит от интенсивности помехи по сравнению с интенсивностью сигнала. Интенсивности сигнала и помехи принято выражать их средними мощностями. Обстановка характеризуется отношением средней мощности сигнала к средней мощности помехи. Для краткости эта величина называется *отношением сигнал/помеха*. В случае флуктуационной помехи вероятность правильного приема тем выше, чем больше отношение сигнал/помеха.

Поэтому прежде чем изучать общие вопросы теории передачи сигналов, целесообразно рассмотреть некоторые простейшие возможности повышения отношения сигнал/помеха.

§ 18. Увеличение отношения сигнал/помеха

Самый простой и очевидный способ увеличения отношения сигнал/помеха состоит в увеличении мощности сигнала. К этому способу и прибегают во многих случаях. Однако для увеличения мощности сиг-

нала приходится соответственно увеличивать мощность источника питания, габариты и вес всей аппаратуры передатчика. Совершенно ясно, что по условиям эксплуатации, например, для всякого рода подвижной аппаратуры возникают строгие ограничения, не позволяющие идти по этому пути. Но, кроме увеличения мощности сигнала, существуют и другие возможности увеличения отношения сигнал/помеха, некоторые из которых мы здесь опишем в общих чертах.

Рассмотрим сначала *метод накопления*.

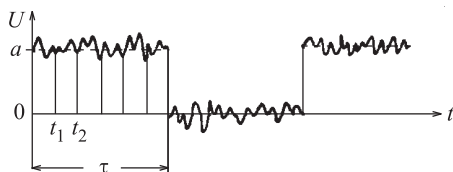


Рис. 35

Идею этого метода мы поясним на простейшем примере не из области радиотехники: пусть производится прием телеграфного сигнала, состоящего из чередующихся посылок и пауз. В посылке имеется постоянное напряжение a , в паузе напряжение равно нулю. На телеграфный сигнал наложена случайная помеха со средним значением, равным нулю (рис. 35). Если производить прием методом пробы, т. е. брать расчет принимаемого сигнала в некоторый момент на протяжении посылки, то получим

$$y = a + \xi,$$

где ξ — случайная величина, выражающая мгновенное значение помехи в момент отсчета.

Отношение сигнал/помеха можно в этом случае выразить как

$$\rho_0 = \frac{a^2}{\overline{\xi^2}},$$

где $\overline{\xi^2}$ — средний квадрат помехи.

Изменим теперь способ приема: возьмем на протяжении посылки не один, а несколько отсчетов в разные моменты времени. Получим:

$$y_1 = a + \xi_1, \quad y_2 = a + \xi_2, \dots, y_n = a + \xi_n,$$

Составим сумму этих отсчетов

$$y = \sum_{k=1}^n y_k = na + \sum_{k=1}^n \xi_k.$$

Первый член выражает полезный сигнал, второй — помеху. Беря средние квадраты обоих членов, составим отношение сигнал/помеха

$$\rho = \frac{n^2 a^2}{(\sum \xi_k)^2}.$$

Если случайные величины ξ_k независимы, то имеем просто

$$\rho = \frac{n^2 a^2}{n \overline{\xi^2}} = n \rho_0,$$

т. е. при n -кратном повторении отсчета отношение сигнал/помеха возрастает в n раз ¹⁾.

Отметим тут же, что вместо суммирования отдельных отсчетов можно выполнить интегрирование смеси сигнала с помехой. Так, если проинтегрировать телеграфный сигнал на протяжении одной посылки, то получим:

$$y = \int_0^{\tau} [a + \xi(t)] dt = a\tau + \int_0^{\tau} \xi(t) dt,$$

$$\rho = \frac{a^2 \tau^2}{\left(\int_0^{\tau} \xi(t) dt \right)^2}.$$

Для вычисления знаменателя нужно знать спектр помехи; однако всегда $\rho > \rho_0$, причем ρ растет пропорционально длительности посылки τ . Такой метод обработки телеграфного сигнала применяется в практике и известен под названием *интегрального приема*.

Метод накопления в различных формах и применениях известен издавна. Напомним, хотя бы, что надежность любого физического эксперимента, служащего для измерения какой-либо величины, повышается путем многократного повторения опыта с последующим усреднением результата. Этим ослабляется влияние случайных погрешностей.

Метод накопления в применении к приему сигналов (в том числе и радиосигналов) сводится по существу к тому, что суммируется несколько *экземпляров* одного и того же сигнала. Если помехи (с нулевым средним значением) в различных экземплярах сигнала независимы, то они суммируются по квадратичному закону (т. е. складываются их мощности или энергии), так что мощность суммы помех растет пропорционально n . Сигналы же (если они одинаковы) суммируются арифметически, так что мощность (или энергия) суммы растет пропорционально n^2 . Отсюда и получается выигрыш в отношении сигнал/помеха в n раз.

Из сказанного следует, что для реализации метода накопления нужно располагать несколькими экземплярами сигнала. При этом безразлично, каким образом они получены. Требуется лишь, чтобы помехи в каждом экземпляре были независимы (тогда выигрыш получается

¹⁾ Метод накопления дает увеличение отношения сигнал/помеха и когда ξ_k зависимы. Однако выигрыш при этом тем меньше, чем больше коэффициент корреляции помехи.

наибольший — в n раз). Иначе говоря, мы должны получить один и тот же сигнал по нескольким независимым (в отношении действия помех) каналам. В предыдущем примере независимые каналы были разделены во времени — мы образовали эти каналы, беря отсчеты в различные моменты времени. Но с таким же успехом можно использовать каналы с частотным разделением, передавая один и тот же сигнал в неперекрывающихся частотных полосах. Мы могли бы также прибегнуть к пространственному разделению каналов, ведя передачу несколькими направлениями лучами, или вообще по нескольким раздельным линиям. Таким образом, *возможно столько вариантов метода накопления, сколько существует способов разделения каналов.*

Перейдем к методу фильтрации.

Для увеличения отношения сигнал/помеха можно использовать различие в спектрах сигнала и помехи. Если бы спектры сигнала и помехи располагались в неперекрывающихся полосах, то сигнал мог бы быть полностью очищен от помехи. Для этого достаточно было бы пропустить смесь сигнала и помехи через полосовой фильтр с полосой пропускания, как раз покрывающей спектр сигнала. К сожалению, в действительности спектры сигнала и помехи всегда в той или иной мере перекрываются. И тем не менее, применение разумно подобранных фильтров может значительно увеличить отношение сигнал/помеха.

Рассмотрим для начала простой частный случай — прием радиотелеграфного сигнала при наличии помехи в виде белого шума с равномерным спектром, т. е. с постоянной спектральной плотностью G_0 . Ширина спектра сигнала F зависит от частоты манипуляции f_0 и может быть выражена соотношением

$$F = \kappa f_0,$$

где κ — коэффициент порядка нескольких единиц, зависящий от допустимого искажения формы сигнала.

Естественно, что искажения тем меньше, чем больше κ . Мощность сигнала в полосе F является заданной величиной; обозначим ее через P_c . Мощность помехи P_{Π} определяется шириной полосы пропускания; естественно взять эту полосу как можно более узкой, т. е. приравнять ее ширине спектра сигнала.

Тогда будем иметь

$$\rho = \frac{P_c}{P_{\Pi}} = \frac{P_c}{FG_0} = \frac{P_c}{\kappa f_0 G_0}.$$

Как видим, при заданных P_c , G_0 и κ можно увеличить ρ путем уменьшения частоты манипуляции f_0 . Уменьшая скорость телеграфирования, мы сужаем спектр телеграфного сигнала; это позволяет применить фильтр с более узкой полосой пропускания; мощность помехи на выходе фильтра уменьшается, так при постоянной спектральной плотности она просто пропорциональна ширине полосы.

Рассмотренный пример характерен тем, что в исходном положении мы имеем широкополосную помеху и узкополосный сигнал. Выбранное

решение — узкополосный *полосовой* фильтр, соответствующий спектру сигнала, — совершенно очевидно. Столь же очевидно решение и в противоположном случае, когда имеется сигнал с широким спектром и помеха с узким спектром (например, синусоидальная или близкая к синусоидальной). Здесь следует применить узкополосный *заграждающий* фильтр, устранивший часть спектра, содержащую помеху.

Обратимся к более общему случаю, когда исходные спектры сигнала и помехи имеют ширины одного порядка. Пусть оба спектра неоднородны, как показано на рис. 36, где изображены спектры *мощности* сигнала и помехи, обозначенные соответственно G_c и G_n . Для получения наибольшего отношения сигнал/помеха нужно найти то значение частоты f_1 , при котором имеет наибольшее отношение G_c/G_n , и применить полосовой фильтр, пропускающий узкую полосу частот около f_1 . Желательная характеристика фильтра изображена на рис. 36 в виде узкого заштрихованного прямоугольника.

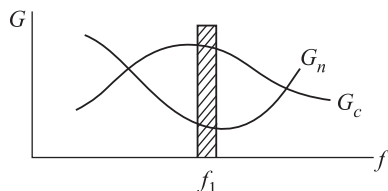


Рис. 36

При таком решении *форма* входного сигнала не сохраняется: на выходе узкополосного фильтра будет сигнал, близкий по форме к синусоиде с частотой f_1 . Можно лишь констатировать наличие сигнала на входе. Это называется задачей *обнаружения* сигнала.

В заключение этого параграфа приведем (без доказательства, которое выходит за рамки нашего курса) следующее важное общее положение: *нельзя получить увеличение отношения сигнал/помеха даром; за него всегда нужно чем-то расплатиться*. Оказывается, что *выигрыш может быть получен либо за счет увеличения мощности сигнала (что очевидно), либо за счет увеличения длительности сигнала, либо за счет расширения его спектра*. Примеры этого параграфа иллюстрируют вторую возможность. Что же касается расширения спектра, то примером использования этой возможности послужат широкополосные системы модуляции, рассматриваемые в последующих параграфах.

§ 19. Зависимость отношения сигнал/помеха от вида модуляции

Модуляция состоит в том, что тот или иной параметр переносчика изменяется во времени пропорционально модулирующей функции.

Представим переносчик функцией

$$f = f(a, b, c, \dots, t),$$

где $a, b, c, \dots$ — параметры, которые можно модулировать.

Если

$$a = a_0 + \delta_c a = a_0 + \Delta a x(t),$$

где $x(t)$ — модулирующая функция¹⁾, то имеем a — модуляцию. Если модулируется параметр b , т. е.

$$b = b_0 + \delta_c b = b_0 + \Delta b x(t),$$

то имеем b -модуляцию, и так далее. Заметим, что знаком δ здесь и ниже отмечены переменные приращения, а знаком Δ — постоянные. Индексами c и Π отмечены величины, относящиеся соответственно к сигналу и помехе.

Величины $\delta_c a$, $\delta_c b$, $\delta_c c$, ... определяют полезный сигнал на выходе приемника.

Если на немодулированный переносчик накладывается помеха, то появляется паразитная модуляция, т. е. параметры получают приращения $\delta_\Pi a$, $\delta_\Pi b$, $\delta_\Pi c$, ..., зависящие от помехи. Эти приращения проявляются в виде помехи на выходе приемника. Следовательно, отношение сигнал/помеха для a -модуляции будет

$$\rho_a = \frac{(\delta_c a)^2}{(\delta_\Pi a)^2} = \frac{(\Delta a)^2}{(\delta_\Pi a)^2} \overline{x^2(t)},$$

для b -модуляции

$$\rho_b = \frac{(\delta_c b)^2}{(\delta_\Pi b)^2} = \frac{(\Delta b)^2}{(\delta_\Pi b)^2} \overline{x^2(t)},$$

и так далее. Отношение сигнал/помеха для разных видов модуляции оказывается существенно различным.

Поясним сказанное примером. В качестве переносчика выберем синусоидальное колебание

$$f = a_0 \cos \omega_0 t.$$

Помеха для простоты пусть будет также синусоидальной, т. е.

$$\xi = \alpha \cos \omega_1 t.$$

Чтобы еще более упростить пример, будем рассматривать случай *малой* помехи, т. е. положим

$$\frac{\alpha}{a_0} \ll 1.$$

Найдем прежде всего паразитные приращения параметров, обусловленные действием помехи. Проще всего определить интересные нас

¹⁾ Предполагается, что модулирующая функция нормирована так, что $x(t) \leq 1$.

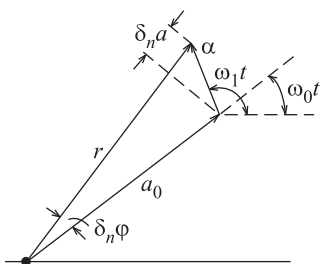


Рис. 37

величины, пользуясь векторной диаграммой рис. 37. Из этой диаграммы видно, что добавление к вектору переносчика вектора помехи вызывает как изменение длины результирующего вектора (обозначенного через r), так и изменение его углового положения. Это означает появление паразитной модуляции как амплитудной, так и угловой.

Приращение амплитуды равно

$$\delta_{\Pi} a = \alpha \cos (\omega_1 - \omega_0) t.$$

Приращение угла (фазы)

$$\delta_{\Pi} \varphi = \frac{\alpha}{a_0} \sin (\omega_1 - \omega_0) t,$$

и средние квадраты этих величин будут равны:

$$\overline{(\delta_{\Pi} a)^2} = \frac{1}{2} \alpha^2, \quad \overline{(\delta_{\Pi} \varphi)^2} = \frac{1}{2} \left(\frac{\alpha}{a_0} \right)^2.$$

Теперь обратимся к приращениям параметров, обусловленным полезной модуляцией. В случае амплитудной модуляции наибольшее возможное значение коэффициента модуляции

$$m = \Delta a / a_0$$

равно единице, так что

$$\Delta a = a_0.$$

Ограничимся случаем синусоидальной модуляции, т.е. положим $x(t) = \cos \Omega t$. Тогда:

$$\overline{x^2(t)} = \frac{1}{2}, \quad \overline{(\delta_c a)^2} = \frac{1}{2} a_0^2,$$

и мы находим отношение сигнал/помеха для АМ

$$\rho_a = a_0^2 / \alpha^2,$$

т.е. отношение сигнал/помеха на выходе АМ приемника равно тому же отношению на входе.

Обратимся к угловой модуляции. Полагая по-прежнему $x(t) = \cos \Omega t$, будем иметь для частотно-модулированного колебания

$$a_0 \cos \left(\omega_0 t + \frac{\Delta \omega}{\Omega} \sin \Omega t \right),$$

так что:

$$\delta_c \varphi = \frac{\Delta \omega}{\Omega} \sin \Omega t = \beta \sin \Omega t; \quad \overline{(\delta_c \varphi)^2} = \frac{1}{2} \beta^2$$

и отношение сигнал/помеха для ЧМ

$$\rho_\omega = \frac{\overline{(\delta_c \varphi)^2}}{\overline{(\delta_{\Pi} \varphi)^2}} = \beta^2 \frac{a_0^2}{\alpha^2} = \beta^2 \rho_0,$$

т. е. в β^2 раз больше, чем при тех же условиях для АМ. В этих формулах $\beta = \Delta\omega/\Omega$ — индекс частотной модуляции. Следует напомнить, что в практике применяются главным образом системы с большим индексом ($\beta \geq 10$), так что преимущество ЧМ перед АМ весьма значительно.

Напомним также, что ширина спектра ЧМ в β раз больше (при большом индексе), чем ширина спектра АМ. Таким образом, увеличение отношения сигнал/помеха при переходе от АМ к ЧМ в β раз по напряжению (в β^2 раз по мощности) покупается ценой расширения занимаемой сигналом полосы в β раз.

Здесь происходит, таким образом, как бы обмен мощности на полосу, однако соотношения при этом обмене не наивыгоднейшие, как будет видно из дальнейшего.

Мы рассматривали простейший пример, в котором помеха и модулирующая функция были синусоидальны. Можно обобщить полученный результат для случая произвольной помехи и сигнала. Это потребует интегрирования по спектру. При АМ это делается очень просто, и результат остается неизменным: отношения сигнал/помеха на входе и выходе АМ приемника равны. Что же касается ЧМ, то вычисления довольно сложны и мы их здесь не приводим. Укажем лишь, что преимущество ЧМ перед АМ сохраняется при всех условиях — пока речь идет о *малой* помехе.

§ 20. Некоторые свойства импульсно-кодовой модуляции

Начнем с рассмотрения квантованной амплитудно-импульсной модуляции. Ошибка при приеме квантованных импульсов состоит в том, что вследствие наложения помехи может быть принят не тот уровень, который фактически передавался, а больший или меньший — в зависимости от знака помехи.

Как уже говорилось в § 11, ошибки не произойдет, если мгновенное значение помехи не превосходит по абсолютной величине половины шага квантования Δx . Но помеха есть случайная величина, и возможны *любые* мгновенные значения. Однако при любых свойствах помехи вероятность ошибки тем меньше, чем больше шаг квантования.

Поэтому, когда речь идет о квантованном сигнале, разумно характеризовать интенсивность сигнала не отношением средних мощностей сигнала и помехи, как обычно, а отношением половины шага квантования к среднеквадратичному (действующему) значению помехи

$$\delta = \sqrt{\xi^2} = \sqrt{P_{\Pi}}.$$

Квадрат этого отношения, т. е.

$$\rho = \frac{\left(\frac{1}{2}\Delta x\right)^2}{P_{\Pi}}, \quad (32)$$

заменяет в рассматриваемом случае обычное отношение сигнал/помеха.

Рассмотрим теперь кодовый метод передачи. Пусть число уровней квантования есть N . Будем передавать каждое из N значений при помощи n -значной кодовой комбинации, составленной из импульсов, квантованных на m уровней (код с основанием m). Общее число возможных комбинаций равно m^n , так как каждому из N уровней должна соответствовать своя кодовая комбинация, то, очевидно,

$$N = m^n \text{ } ^1).$$

Подсчитаем среднюю мощность такого кодированного сигнала. Пусть применяются как положительные, так и отрицательные уровни, т.е. пусть шкала уровней симметрична относительно нуля, так что разрешенными являются следующие уровни: $-\frac{m-1}{2}\Delta x, \dots, -i\Delta x, \dots, -2\Delta x, -\Delta x, 0, \Delta x, 2\Delta x, \dots, i\Delta x, \dots, \frac{m-1}{2}\Delta x$, а всего m уровней. Тогда, считая все уровни равновероятными, получим для мощности выражение

$$P_c = \frac{\Delta x^2}{m} \sum_{i=-\frac{m-1}{2}}^{\frac{m-1}{2}} i^2 = \frac{1}{12} \Delta x^2 (m^2 - 1).$$

Найдем отсюда шаг квантования

$$\Delta x = \sqrt{\frac{12P_c}{m^2 - 1}}$$

и подставим это значение в (32)

$$\rho = \frac{3}{m^2 - 1} \frac{P_c}{P_{\Pi}} = \frac{3}{m^2 - 1} \rho_0. \quad (33)$$

Таким образом, при неизменных мощностях сигнала и помехи выгодно уменьшать основание кода. Наименьшее значение m равно двум (двоичный код). Кодовый метод с применением двоичного кода — это и есть обычная ИКМ. Для нее из (33) получаем

$$\rho_2 = \frac{P_c}{P_{\Pi}} = \rho_0. \quad (34)$$

т.е. введенная нами специальная величина (32) совпадает, в случае двоичного кода, с обычным определением отношения сигнал/помеха. Наихудший случай будет при $N = m$, $n = 1$. Но это не что иное, как

¹⁾ N , m и n — целые числа; если это равенство не выполняется в точности, то нужно взять $N < m^n$. Это значит, что часть кодовых комбинаций останется неиспользованной.

обычная квантованная АИМ. Так как обычно $N = m \gg 1$, то из (33) получаем для АИМ

$$\rho_N = \frac{3}{N^2} \rho_0. \quad (35)$$

Сравнивая (34) и (35), мы видим, что переход от АИМ к ИКМ дает выигрыш в отношении сигнал/помеха в

$$\frac{\rho_2}{\rho_N} = \frac{1}{3} N^2 = \frac{1}{3} 2^{2n} \quad (36)$$

раз.

Посмотрим теперь, какой ценой покупается этот выигрыш. Если при АИМ за каждый тактовый интервал передается один импульс, то при ИКМ за тот же интервал должны быть переданы n импульсов. При неизменной скважности каждый из этих n импульсов в n раз короче, а следовательно, ширина спектра сигнала ИКМ ровно в n раз больше, чем ширина спектра сигнала АИМ. Таким образом, за увеличение отношения сигнал/помеха мы расплачиваемся расширением полосы.

Теперь нужно уточнить результат (36). Он справедлив, если мощность помехи остается неизменной. Но при помехе с равномерным спектром мощность помехи пропорциональна ширине полосы. Если учесть это обстоятельство, то чистый выигрыш будет не в $\frac{1}{3} 2^{2n}$ раз, а в $\frac{1}{3n} 2^{2n}$ раз.

При переходе от АМ к ЧМ также получается выигрыш в отношении сигнал/помеха за счет расширения полосы (см. § 19).

Однако очень важно заметить, что переход от АИМ к ИКМ приводит к значительно более выгодным соотношениям, что видно из сопоставления, сделанного в табл. 4.

Таблица 4

Переход	Относительное расширение полосы	Увеличение отношения сигнал/помеха пропорционально
От АМ к ЧМ	β	β^2
От АИМ к ИКМ	n	$\frac{1}{n} 2^{2n}$

При расширении полосы в одинаковое число раз отношение сигнал/помеха растет при переходе АМ–ЧМ по степенному закону, а при переходе АИМ–ИКМ — по показательному закону, т.е. значительно быстрее (множитель $\frac{1}{n}$ не играет заметной роли). Здесь происходит своего рода «обмен мощности на полосу», причем для ИКМ этот обмен совершается на более выгодных условиях.

Можно показать, что из всех известных и применяемых на практике систем ИКМ наиболее приближается к теоретически идеальной системе передачи.

§ 21. Помехоустойчивость квантованного сигнала

До сих пор о помехоустойчивости говорилось в плане общих определений и рассуждений. В этом параграфе мы покажем на простом примере конкретное вычисление вероятности ошибки.

Будем рассматривать квантованный сигнал с шагом квантования Δx . На сигнал наложен белый шум ξ мощностью P_{Π} , чему соответствует среднеквадратичное значение

$$\sigma = \sqrt{P_{\Pi}}.$$

При передаче на значение данного уровня может накладиться помеха, превосходящая половину шага квантования. В результате уровень будет определен неверно. Вероятность такого события, т. е. вероятность ошибки, запишется как

$$p = p \left\{ |\xi| > \frac{1}{2} \Delta x \right\}. \quad (37)$$

Для вычисления этой вероятности нужно знать плотность распределения вероятностей помехи. Белый шум имеет нормальное распределение, т. е. плотность распределения вероятностей выражается функцией

$$\varphi(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{x^2}{2\sigma^2}}. \quad (38)$$

Вероятность того, что значение случайной величины ξ окажется в интервале a, b , выражается общей формулой

$$p = \{a < \xi < b\} = \int_a^b \varphi(x) dx. \quad (39)$$

Рассмотрим вместо вероятности ошибки (37) вероятность противоположного события — вероятность правильного приема

$$q = 1 - p = p \left\{ |\xi| < \frac{1}{2} \Delta x \right\}.$$

Применяя (39), получим

$$q = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \varphi(x) dx = \frac{2}{\sqrt{2\pi} \sigma} \int_0^{\frac{1}{2}\Delta x} \varphi(x) dx$$

или, вводя $\varphi(x)$ из (38),

$$q = \frac{2}{\sqrt{2\pi}\sigma} \int_0^{\frac{1}{2}x} e^{-\frac{x^2}{2\sigma^2}} dx = \frac{2}{\sqrt{\pi}} \int_0^{\frac{x/2}{\sqrt{2}\sigma}} e^{-z^2} dz.$$

Интеграл в правой части есть интеграл вероятностей, или функция Лапласа. Он является функцией верхнего предела и обозначается так:

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz.$$

Введем обозначение

$$\mu = \frac{\frac{1}{2}\Delta x}{\sigma}.$$

Тогда вероятность правильного приема

$$q = \Phi\left(\frac{\mu}{\sqrt{2}}\right),$$

а вероятность ошибки

$$p = 1 - q = 1 - \Phi\left(\frac{\mu}{\sqrt{2}}\right). \quad (40)$$

Функция $\Phi(x)$ табулирована; пользуясь таблицами, можно построить график $p(\mu)$. Так как эта функция очень быстро убывает, то целесообразно выбрать логарифмический масштаб по оси ординат. График зависимости (40) дан на рис. 38. Пользуясь этим графиком, можно найти необходимую относительную величину шага квантования, если задана

вероятность ошибки. Так, если вероятность ошибки не должна превышать 10^{-6} (т.е. в среднем одна ошибка на миллион отсчетов, что может считаться хорошей надежностью в большинстве практических условий передачи), то достаточно взять:

$$\mu = \frac{\frac{1}{2}\Delta x}{\sigma} = 5, \quad \Delta x = 10\sigma.$$

Такое значение и принимают обычно при расчетах.

Выше говорилось (§ 18), что помехоустойчивость возрастает с увеличением различия (или расстояния) между сигналами. В рассмотренном случае различие определяется шагом квантования. Говорилось также, что помехоустойчивость возрастает с увеличением

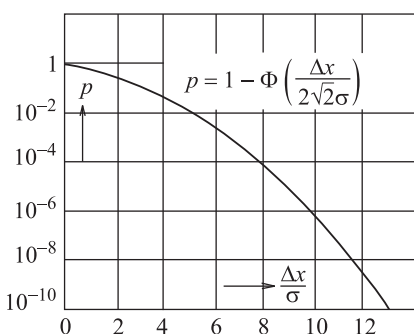


Рис. 38

мощности сигнала. В нашем случае мощность сигнала растет пропорционально квадрату шага квантования.

Таким образом, рассмотренный пример дает простые иллюстрации общим положениям.

§ 22. Корректирующие коды

Одним из новых средств борьбы с помехами являются *корректирующие коды*, т. е. коды, позволяющие обнаружить и исправить ошибки при приеме.

Рассмотрим n -значный двоичный код, кодовые комбинации которого представляются n -значными двоичными числами. Всего можно составить

$$N = 2^n$$

различных комбинаций. Ошибка при приеме кодовой комбинации состоит в том, что вследствие действия помехи нуль заменяется единицей или, наоборот, единица нулем. Если в кодовой комбинации один знак заменяется ошибочным, то такую ошибку будем называть *одиночной*, если два — *двойной* и т. д.

Если при передаче используются все возможные кодовые комбинации, то ошибка любой кратности остается незамеченной. Действительно: ошибки состоят в замене нулей единицами и наоборот. Но любое число таких замен превращает данную кодовую комбинацию в некоторую другую. А так как любая кодовая комбинация может встретиться в передаче, то у нас нет никаких оснований сомневаться в правильности принятой комбинации (если рассматривать ее отдельно, не используя статистику, контекст и т. п.).

Пусть, например, четыре различных сообщения A , B , C и D закодированы двузначным кодом ($n = 2$, $N = 4$, табл. 5).

Если передается сообщение B , то одиночная ошибка в первом знаке кодовой комбинации 01 дает 11 и мы ошибочно принимаем сообщение D . Ошибка во втором знаке дает сообщение A . Сообщение C может получиться только в результате двойной ошибки.

Таблица 5

Сообщение	A	B	C	D
Кодовая комбинация	00	01	10	11

Таблица 6

Сообщение	A	B	C	D
Кодовая комбинация	000	011	101	110

Различие между кодовыми комбинациями определяется числом различающихся знаков. Мы видим, что комбинации 00 и 01 различаются в одном знаке; то же относится к парам 00 и 10; 10 и 11; 01 и 11. Комбинации 00 и 11; 01 и 10 различаются между собой в двух знаках.

Для того чтобы можно было обнаружить одиночную ошибку, достаточно взять такие кодовые комбинации, которые различались бы между собой *не менее чем в двух знаках*. Тогда одиночная ошибка

даст кодовую комбинацию, которая отличается от истинной в одном знаке, но и от любой другой комбинации, использованной в данном коде, отличается *не менее* чем в одном знаке. Иначе говоря: принцип построения кода, обнаруживающего одиночную ошибку, состоит в том, что используются не все возможные комбинации, а лишь половина. Вторая половина образует запрещенные (т. е. не применяемые в данном коде) комбинации. Одиночная ошибка превращает разрешенную комбинацию в запрещенную; тем самым и обнаруживается произошедшая ошибка.

Для примера возьмем те же четыре сообщения и построим для них код, обнаруживающий одиночную ошибку (табл. 6).

Для того чтобы удовлетворить условию различия не менее чем в двух знаках, пришлось взять трехзначные комбинации. Но всего возможно $2^3 = 8$ таких комбинаций. Мы использовали половину; комбинации 001, 010, 100 и 111 — запрещены. Возьмем, к примеру, сообщение *B*. Ошибка в любом знаке комбинации 011 превращает эту комбинацию в одну из запрещенных, благодаря чему ошибка и обнаруживается.

Перейдем к коду, исправляющему одиночную ошибку. Такой код строится из кодовых комбинаций, различающихся *не менее чем в трех знаках*. Одиночная ошибка даст кодовую комбинацию, отличающуюся от истинной в одном знаке, но от ближайшей другой разрешенной комбинации не менее чем в двух знаках. Таким образом, ошибочно принятая комбинация всегда будет *ближе* к истинной, чем к любой другой разрешенной комбинации данного кода, и истинная комбинация может быть на этом основании восстановлена, т. е. ошибка исправлена. В табл. 7 дан пример исправляющего одиночную ошибку кода.

Таблица 7

Сообщение	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
Кодовая комбинация	00000	01101	10110	11011

Здесь пришлось применить уже пятизначные комбинации для того, чтобы получить требуемое различие любой пары комбинаций не менее чем в трех знаках. Пусть принята комбинация 01001. Такой комбинации в кодовой таблице нет. Следовательно, произошла ошибка. Но какое же сообщение в действительности передавалось? Сличим принятую комбинацию с каждой из разрешенных комбинаций данного кода. Принятая комбинация 01001 отличается от 00000 в двух знаках, от 01101 — в одном знаке, от 10110 — в пяти знаках, от 11011 — в двух знаках. Ближайшая кодовая комбинация есть 01101; стало быть, передавалось сообщение *B*.

Аналогичным образом можно построить коды, обнаруживающие или обнаруживающие и исправляющие ошибки любой кратности. Чем

больше исправляющая способность кода, т.е. чем выше кратность исправляемых ошибок, тем больше требуемое число знаков.

Способ приема при применении исправляющего кода состоит в том, что принятая комбинация сличается поочередно со всеми разрешенными комбинациями данного кода. Если она совпадает с какой-либо из них, то считается, что ошибки нет ¹⁾. Если же обнаружена ошибка, то истинной считается та комбинация, от которой принятая наименее отличается. При большом числе комбинаций такой способ приема приводит к громоздким техническим решениям.

Усиленно разрабатываются специальные коды, строение которых позволяет осуществить декодирование более простыми способами. Таких кодов имеется очень много и число их продолжает возрастать. Одновременно развивается и математическая теория исправляющих кодов. В рамках нашего курса нет никакой возможности изучать эту обширную и сравнительно трудную область. Мы ограничимся лишь самыми краткими сведениями о так называемых *систематических* кодах.

Систематическими называются такие коды, у которых общее число n знаков делится на κ : *информационных* знаков и r *контрольных* знаков, так что

$$n = \kappa + r.$$

Число информационных знаков определяется числом личных сообщений, которые нужно передать, а именно:

$$N = 2^{\kappa}.$$

Из r контрольных знаков можно образовать 2^r двоичных комбинаций. Число r легко определить, если учесть, что для исправления одиночной ошибки нужно: 1) указать, есть ошибка или нет, 2) если есть, то указать, в какой из n позиций (для исправления ошибки в двоичном коде достаточно указать *место*; исправление сводится к замене в указанном месте нуля единицей или единицы нулем).

Таблица 8

r	1	2	3	4	5...
κ	0	1	4	11	26...
n	1	3	7	15	31...

Итак, нужна одна кодовая комбинация для ответа «да» или «нет» на вопрос о наличии ошибки и n кодовых комбинаций для указания номера ошибочной позиции. Следовательно, должно быть

$$2^r \geq n + 1$$

или

$$\frac{2^n}{n+1} \geq N.$$

¹⁾ Хотя, строго говоря, это возможно при ошибке более высокой кратности, чем исправляемая данным кодом. Впрочем, такое событие обычно очень маловероятно.

Наивыгоднейший случай — когда имеется равенство; тогда относительное число контрольных знаков (т. е. r/n) получается наименьшим. Это будет при значениях κ , r и n , указанных в табл. 8.

В ранее рассмотренном примере:

$$N = 2^\kappa = 4; \quad \kappa = 2.$$

Наименьшее значение n равно пяти (так как $\frac{2^5}{5+1} \approx 5,3$, а $\frac{2^4}{4+1} = 3,2$). Таким образом, $r = n - \kappa = 3$ ¹⁾. Приведенный в табл. 7 код можно рассматривать как систематический: первые два знака можно считать информационными, последние три — контрольными.

Займемся теперь процедурой приема. Сначала вернемся к обнаруживающему коду табл. 6. Кодовые комбинации этого кода различаются между собой в двух знаках. Это значит, что *число единиц* в кодовой комбинации либо остается неизменным (когда одна единица заменяется нулем и один ноль единицей), либо увеличивается на два (когда два нуля заменяются двумя единицами), либо, наконец, уменьшается на два (когда две единицы заменяются нулями). Короче говоря, *четность*

Таблица 9

Позиция про- верки	1	2	3	4	5
	Знаки				
	Информа- ционные		Контроль- ные		
I	×	×	×		
II	×			×	
III		×			×

числа единиц во всех кодовых комбинациях обнаруживающего кода одинакова. Как видим (см. табл. 6), разрешенные кодовые комбинации имеют четное число единиц, а запрещенные — нечетное. Отсюда, кстати говоря, вытекает и способ построения систематического обнаруживающего кода. Способ состоит в том, что к исходным кодовым комбинациям (табл. 5) приписывают в качестве добавочного (контроль-

ного) знака нуль или единицу с таким расчетом, чтобы получить всегда четное число единиц (или всегда нечетное — это безразлично).

Итак, обнаружение ошибки производится путем *проверки на четность*. Это сравнительно простая операция, легко осуществимая технически.

Принцип проверки на четность сохраняется и в применении к исправляющим кодам. Здесь, однако, производятся несколько проверок на четность по особой схеме, на способе составления которой мы не будем останавливаться. Приведем лишь для примера схему проверок для исправляющего кода табл. 7. Эта схема дана в табл. 9.

Проверка состоит в том, что составляется сумма числа единиц на позициях, отмеченных крестиками. Контрольные знаки подобраны так, чтобы при отсутствии ошибки эта сумма для всех проверок была

¹⁾ При наивыгоднейших соотношениях трех контрольных знаков хватило бы для исправления одиночной ошибки в семизначном коде (см. табл. 8).

четной. Нечетность при хотя бы одной проверке указывает на наличие ошибки.

Местоположение одиночной ошибки определяется следующим образом: если нечет дает только одна из трех проверок, то ошибка имеется в контрольном знаке, входящем в данную проверку. Если нечет дают проверки I и II, то ошибка в первом информационном знаке; если нечет дают проверки I и III, то ошибка во втором информационном знаке. Нечет в проверках II и III может получиться только при двойной ошибке ¹⁾.

Покажем применение этой процедуры на примере. Пусть принята комбинация 01001. Первая проверка дает $0 + 1 + 0 = 1$ — нечет. Вторая — $0 + 0 = 0$ — чет. Третья — $1 + 1 = 2$ — чет. Итак, ошибка в первом контрольном знаке, стоящем на третьей позиции. Исправленная комбинация будет 01101, что соответствует сообщению *B* (см. табл. 7).

В заключение отметим, что повышение помехоустойчивости, достигаемое с помощью корректирующих кодов, связано с увеличением значности кода. Это означает либо увеличение длительности сигнала, либо (если кодовая комбинация корректирующего кода должна передаваться за то же время, что и кодовая комбинация обычного кода) соответствующее расширение спектра сигнала. Мы получаем, таким образом, еще одну иллюстрацию высказанного выше общего положения о цене помехоустойчивости.

¹⁾ Заметим, что под кодом, исправляющим одиночную ошибку, понимается код, исправляющий все возможные одиночные ошибки. Однако, применяя соответствующую схему проверок, можно исправить и некоторую часть ошибок более высокой кратности. Так, пятизначный, код при $k = 2$ исправляет все 5 одиночных ошибок и 2 из 10 двойных.

Часть II

ЛИНЕЙНЫЕ ЯВЛЕНИЯ В РАДИОТЕХНИКЕ

Глава 4

ЗАДАЧИ И МЕТОДЫ ЛИНЕЙНОЙ ТЕОРИИ

§ 23. Лине́йные явления и устройства

Всякое радиотехническое устройство представляет собой сложную цепь, отдельные звенья которой выполняют определённые функции. К числу таких функций относятся: генерирование колебаний, их усиление, модуляция, преобразование частоты, умножение частоты, деление частоты, фильтрация, детектирование модулированных колебаний и т. п.

Классификацию радиотехнических устройств можно строить по разным признакам. Так, можно различать их по техническому назначению, например делить на передающие устройства, приемные устройства, усилители и т. п. Именно этот принцип применяется при построении цикла специальных дисциплин в высшем учебном заведении радиотехнического профиля.

Но с точки зрения *теоретической* радиотехники имеет преимущество другой способ различения, а именно по математическому аппарату, применяемому при изучении свойств и действия тех или иных устройств. С этой точки зрения классификация строится на основе различий в математических уравнениях, описывающих действие того или иного звена сложного радиотехнического устройства.

Все уравнения, которые приходится составлять и решать при исследовании различных радиотехнических проблем, делятся на три вида:

1. Лине́йные уравнения с постоянными коэффициентами.
2. Лине́йные уравнения с переменными коэффициентами.
3. Нелинейные уравнения.

Мы сохраняем традиционное деление курса, состоящее в том, что проблемы, сводящиеся к линейным уравнениям с переменными коэффициентами и к нелинейным уравнениям, объединены в один раздел, называемый обычно «нелинейной радиотехникой». Этот раздел составляет содержание третьей части курса. В настоящей же части мы будем заниматься только вопросами, разрешимыми с помощью уравнений с постоянными коэффициентами.

Эти вопросы сводятся к исследованию тех звеньев радиотехнических устройств, которые представляют собой линейные системы с постоянными параметрами. Назначение этих звеньев состоит в передаче электрических возмущений от предшествующего звена к последующему. Передача может сопровождаться тем или иным преобразованием, о чем подробнее говорится в § 30. Таким образом, всякое такого рода звено можно представить в виде четырехполюсника, на вход которого электрическое возмущение подается, а с выхода — снимается.

Обозначим электрическое возмущение на входе четырехполюсника через $x(t)$, а соответствующее возмущение на выходе — через $y(t)$. Будем называть $x(t)$ *воздействием*, а $y(t)$ — *откликом*, подчеркивая этим то, что $x(t)$ и $y(t)$ находятся между собой в причинной взаимосвязи. Исследуемое устройство будем называть *системой*. Все задачи линейной теории сводятся к следующим трем:

1. По заданному воздействию $x(t)$, зная свойства системы, найти отклик $y(t)$.
2. По наблюдаемому отклику, зная свойства системы, найти неизвестное воздействие.
3. По заданным воздействию и отклику определить требуемые свойства системы.

Первая задача — главнейшая задача линейной теории. Вторая задача встречается реже; впрочем, она решается теми же средствами, что и первая. Третья задача — это, по существу, задача *синтеза* системы с заданными свойствами. Эта задача совершенно специфична.

Линейные уравнения с постоянными коэффициентами делятся, в свою очередь, на обыкновенные уравнения (т. е. уравнения в полных производных) и уравнения в частных производных. Системы, описываемые этими уравнениями, называются соответственно системами с сосредоточенными параметрами и системами с распределенными параметрами. К первым относятся контуры (любой сложности), фильтры; ко вторым — линии, волноводы, объемные резонаторы. Обыкновенные уравнения проще уравнений в частных производных. Поэтому, располагая материалом, как обычно, в порядке возрастающей трудности, мы будем рассматривать сначала системы с сосредоточенными параметрами.

Линейная теория располагает мощными общими методами исследования. Эти методы основываются на весьма универсальном *принципе наложения*.

§ 24. Принцип наложения

Сущность принципа наложения (или принципа суперпозиции) поясним сначала на примере.

Пусть в некоторой произвольной линейной электрической цепи действует электродвижущая сила e_1 . При этом в некотором плече схемы протекает ток i_1 . Пусть теперь источник эдс e_1 устранен, и взамен него на любом участке цепи действует эдс e_2 , вызывающая в том же плече схемы ток i_2 . Если затем ввести в действие на прежних местах

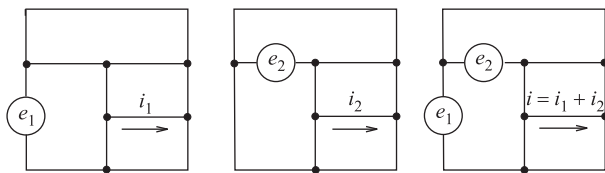


Рис. 39

оба источника эдс, то в цепи потечет ток $i = i_1 + i_2$ (рис. 39). То же самое относится к любому числу электродвижущих сил, т. е. $i = \sum_{k=1}^n i_k$, где i_k — ток, порождаемый отдельно действующей электродвижущей силой e_k .

Таким образом, действие каждой эдс независимо: создаваемая ею составляющая тока остается неизменной, вне зависимости от того, действуют в системе другие эдс или нет. Поэтому принцип наложения называют также принципом независимости действия.

Принцип наложения применим не только к электрическим цепям (хотя это применение интересует нас в первую очередь); он является общезначимым принципом. Чтобы подчеркнуть общность принципа наложения, можно выразить его следующей достаточно широкой формулой: *в линейной системе действие суммы причин равно сумме действий, вызываемых каждой причиной, отдельно взятой*. В рассмотренном примере эдс трактуются как причины, а токи — как действия.

Простейшим следствием принципа наложения является пропорциональность между причиной и действием. Если эдс возрастает в a раз, то и ток, порождаемый этой эдс, возрастает в a раз. Пропорциональность между током и напряжением, выражаемая законом Ома, вытекает из принципа наложения и имеет место только в линейной цепи.

Применимость принципа наложения может служить физическим определением линейной системы. Таким образом, с математической точки зрения линейной системой называется система, описываемая линейным уравнением, с физической же точки зрения линейной называется система, для которой справедлив принцип наложения. Оба определения, разумеется, не противоречат друг другу; одно может быть выведено из другого.

На принципе наложения основаны сильные общие методы решения линейных задач. Рассмотрим одну из задач, сформулированных в предыдущем параграфе: найти отклик линейной системы на заданное воздействие при известных свойствах системы.

В этой задаче воздействие $x(t)$ рассматривается как причина, а отклик $y(t)$ — как действие, применение принципа наложения состоит в следующем: мы разбиваем входное воздействие на произвольные слагаемые, т. е. полагаем

$$x = \sum_{k=1}^n x_k. \quad (41)$$

При этом мы выбираем способ разбиения (т. е. вид слагаемых x_k) так, чтобы отклик системы на каждое слагаемое x_k мог быть легко найден. Определив отклик y_k на каждое слагаемое x_k воздействия, основываясь на принципе наложения, находим полный отклик y на воздействие x путем суммирования y_k , т. е.

$$y = \sum_{k=1}^n y_k. \quad (42)$$

что и решает поставленную задачу.

Не обязательно разбивать x на *конечные* слагаемые. Слагаемые могут быть и бесконечно малыми. В таком случае функция $x(t)$ будет представлена не суммой (41), а соответствующим интегралом. При вычислении полного отклика нужно также выполнить интегрирование, так что решение задачи будет в этом случае выражаться не суммой (42), а некоторым интегралом.

Остается нерешенным вопрос о том, как именно следует выбирать элементарные слагаемые, на которые разбивается при применении этого метода входное воздействие $x(t)$. На этот вопрос нельзя дать однозначный ответ, так как функцию $x(t)$ можно разбить на любые слагаемые бесчисленными способами. Но можно выбрать наиболее подходящие способы разбиения, принимая во внимание наличие уже подготовленного математического аппарата, или опираясь на какие-либо хорошо усвоенные физические представления.

Из всех возможных способов разбиения применяются преимущественно только два, использующие интегралы Фурье и Дюамеля. Методы интеграла Фурье и интеграла Дюамеля описаны в последующих параграфах. Предварительно нужно сказать несколько слов о двух различных подходах к описанию явлений и свойств систем — о временном подходе и о спектральном подходе, так как выбор того или иного подхода определяет и выбор соответствующего математического аппарата.

§ 25. Временной и спектральный подходы

Всякий сигнал может быть представлен некоторой функцией $x(t)$, выражающей изменение во времени той или иной электрической величины (например, тока или напряжения). Это относится, впрочем, не только к сигналам, но и к самым разнообразным явлениям и процессам; отличие сводится к различию в наблюдаемых величинах, т. е. в физическом смысле величины x .

Предположим для начала, что функция $x(t)$ — периодическая. В таком случае она может быть представлена рядом Фурье, который мы на этот раз запишем в комплексной форме

$$x(t) = \sum_{k=-\infty}^{\infty} C_k e^{ik\omega_1 t}. \quad (43)$$

Эта запись равносильна формулам (1) или (2). C означают комплексные амплитуды гармоник и связаны с величинами § 3 следующим образом:

$$2C_k = a_k - ib_k = c_k e^{-i\varphi_k}, \quad C_0 = c_0.$$

Комплексные амплитуды C_k определяются по формуле

$$C_k = \frac{1}{T} \int_{-T/2}^{T/2} x(t) e^{-ik\omega_1 t} dt, \quad (44)$$

где $T = 2\pi/\omega_1$ — период функции $x(t)$.

Теперь заметим, что если амплитуды C_k определены по формуле (44), то равенство (43) есть *тождество*. Это значит, что задание совокупности C_k , т. е. комплексного спектра, *полностью* определяет периодическую функцию $x(t)$. Поэтому можно определять процесс *либо* заданием описывающей его функцией *времени*, *либо* заданием комплексного спектра, являющегося функцией *частоты*. При этом важно заметить, что оба представления совершенно равносильны.

Дело сводится к тому, что мы по-разному представляем один и тот же процесс; мы говорим об одном и том же явлении на двух различных языках — на языке *временном* и на языке *спектральном (частотном)*.

В радиотехнике применяются оба языка; радиоспециалист должен свободно изъясняться на любом из них и быть в состоянии в случае надобности легко делать перевод с одного языка на другой, т. е. переходить при описании того или иного явления от временных представлений к спектральным и обратно.

Следует пояснить, что в принципе можно было бы обойтись *только* временными или *только* спектральными представлениями. Но дело в том, что назначение различных радиотехнических устройств различно; в одних случаях естественно и удобно пользоваться спектральными представлениями, в других — временными. Так, например, назначение всякого фильтра состоит в том, чтобы пропустить колебания одних частот и задержать колебания других частот; иначе говоря, фильтр видоизменяет спектр воздействия. Поэтому действие и свойства фильтра естественно описывать на спектральном языке. В качестве другого примера возьмем какую-нибудь схему для передачи импульсов (хотя бы телеграфных сигналов). Если при этом нас интересует искажение *формы* импульсов, то их естественно представлять функциями времени. К этому нужно еще добавить, что *одна и та же* система может использоваться по-разному и в зависимости от этого рассматриваться с различных точек зрения. Так, например, схема обычного *фильтра нижних частот* (т. е. фильтра, пропускающего низкие частоты и задерживающего высокие частоты) может применяться в качестве *линии задержки*, т. е. устройства, служащего для получения запаздывания некоторого импульса на определенное время. Ясно, что в этом случае удобнее описывать работу системы на временном языке.

Этот пример особенно отчетливо показывает, что выбор того или другого способа описания данной системы зависит не столько от ее устройства, сколько от ее назначения и использования. Нужно понимать, что свойства данной физической системы представляют собой некоторую объективную сущность, которая не меняется от способа ее описания. Меняется только наша точка зрения на предмет, но не самый предмет.

История развития частотных и временных представлений дает много поучительного. Спектральные представления возникли сравнительно недавно и долгое время были мало распространены. Затем они бурно развились и получили широкое распространение, завоевав, в частности в радиотехнике, монопольное положение. Это принесло известный вред, так как безоговорочное применение спектрального метода *во всех* случаях нецелесообразно и может даже тормозить научно-технический прогресс. В новейшее время обозначилась тенденция к тому, чтобы, не отдавая безусловного предпочтения одному какому-либо методу исследования, применять каждый из них там, где он уместен. Следуя этой разумной тенденции, мы будем в дальнейшем излагать все вопросы линейной теории параллельно как на спектральном, так и на временном языках.

§ 26. Интеграл Фурье

До сих пор мы применяли разложение на синусоидальные составляющие только к периодическим функциям, представляя их рядами Фурье. Теперь введем такого же рода разложение применительно к *непериодическим* функциям. Понятие об этом важном обобщении можно составить путем нижеописанного предельного перехода.

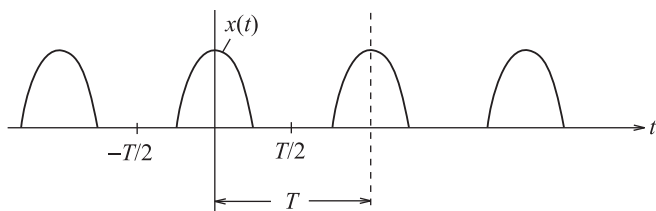


Рис. 40

На рис. 40 представлен график периодической функции в виде последовательности одинаковых импульсов произвольной формы, выражаемой функцией

$$x(t) = x(t - iT),$$

где i — любое целое число. Последовательность может быть представлена рядом Фурье

$$x(t) = \sum_{k=-\infty}^{\infty} C_k e^{ik\omega_1 t},$$

где

$$C_k = \frac{1}{T} \int_{-T/2}^{T/2} x(t) e^{-ik\omega_1 t} dt. \quad (45)$$

Будем рассматривать вместо амплитуды C_k произведение

$$s_k = C_k T = \int_{-T/2}^{T/2} x(t) e^{-ik\omega_1 t} dt. \quad (46)$$

Спектр амплитуд функции $x(t)$ — линейчатый; его примерный вид показан на рис. 41, а; спектр построен в масштабе $|s_k|$ (а не $|C_k|$) по оси ординат. Расстояния между спектральными линиями по оси абсцисс равны $\omega_1 = 2\pi/T$.

Будем теперь увеличивать T . При этом интервал между спектральными линиями соответственно уменьшится. Однако для тех же

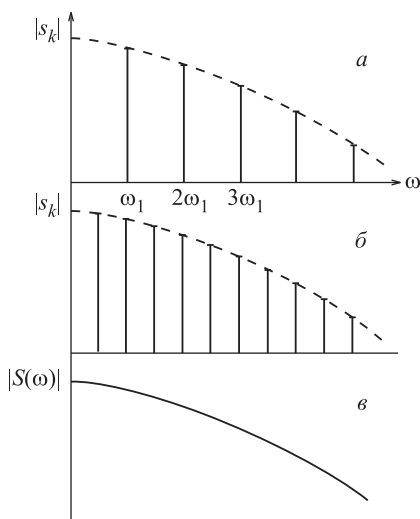


Рис. 41

значений $\omega = k\omega_1$, что и прежде, значение s_k останется неизменным, как это видно из формулы (46). Поэтому, если, например, период увеличился вдвое, то число линий удваивается, причем первоначальные линии спектра сохраняют свою величину и положение, а между ними появляются новые, как показано на рис. 41, б. Так как $s_k = C_k T$ для некоторой частоты $\omega = k\omega_1$ остается неизменным, то это значит, что амплитуда C_k убывает с увеличением периода T . Если продолжать увеличивать период, то линейчатый спектр будет становиться все более густым. При переходе к пределу при $T \rightarrow \infty$ интервал между линиями будет стремиться к нулю и совокуп-

ность дискретных точек с координатами $(s_k, k\omega_1)$ превратится в непрерывную последовательность точек, т.е. в линию, представляющую некоторую непрерывную функцию частоты, которую мы обозначим $S(\omega)$ (рис. 41, в). Но при неограниченном увеличении периода периодическая последовательность рис. 40 вырождается и переходит в один единственный импульс, представляемый функцией $x(t)$. Таким образом, устремляя T к бесконечности, мы переходим в пределе от периодической функции к непериодической.

Выразим амплитуду C_k как

$$C_k = \frac{s_k}{T} = \frac{1}{2\pi} s_k \omega_1.$$

Но при $T \rightarrow \infty$ величина ω_1 , представляющая собой интервал между линиями, превращается в $d\omega$, а s_k переходит в непрерывную функцию $S(\omega)$. Заменяя в (45) сумму интегралом, получаем

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{i\omega t} d\omega. \quad (47)$$

Это и есть *интеграл Фурье*. Он позволяет любую (с некоторыми ограничениями) непериодическую функцию представить в виде суммы бесконечного числа синусоидальных колебаний с бесконечно малыми амплитудами $S(\omega) d\omega$ и с бесконечно малым интервалом $d\omega$ по частоте. Говорят, что в составе непериодической функции имеются «все частоты». Функция $S(\omega)$ выражает спектральную плотность комплексной амплитуды ¹⁾. Она определяется на основании формулы (46), в которую вместо дискретных значений частоты $k\omega_1$ нужно ввести *непрерывно изменяющуюся* (текущую) частоту ω

$$S(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt. \quad (48)$$

Формула (48) позволяет найти $S(\omega)$ по заданному $x(t)$, а формула (47) — $x(t)$ по заданному $S(\omega)$. Формулы (47) и (48) имеют симметричное строение и называются *парой преобразований Фурье*.

Функция $S(\omega)$ — комплексная функция; она содержит сведения как о спектре амплитуд, так и о спектре фаз. Мы будем в дальнейшем называть функцию $S(\omega)$ для краткости *комплексным спектром*. Ее модуль

$$\Phi(\omega) = |S(\omega)|$$

дает сведения только о частотном распределении амплитуды. Во многих случаях только это нас и интересует. Мы будем называть функцию $\Phi(\omega)$ просто спектром, если это не потребует оговорок.

Спектр непериодической функции, выражаемый непрерывной функцией частоты, называется *сплошным* (в отличие от линейчатого спектра периодической функции).

¹⁾ Спектральная плотность *амплитуд* и спектральная плотность *фаз* — понятия, применяемые только к определенным функциям времени. Эти понятия не применимы к случайным процессам. Для случайного процесса существует только статистический спектр, физический смысл которого состоит в том, что он выражает спектральную плотность *мощности*.

Прodelанное рассуждение, приведшее нас к основным формулам (47) и (48), не является математически строгим. Однако оно дало нам попутно весьма полезный результат, а именно: мы установили связь между спектром одиночного импульса и спектром периодической последовательности, получаемой периодическим повторением импульса. Эта связь может быть выражена следующим правилом: *линейчатый спектр амплитуд периодической последовательности импульсов (построенный в масштабе $|s_k| = |C_k|T$) вписывается в кривую $\Phi(\omega) = |S(\omega)|$ сплошного спектра одиночного импульса*. Аналогичное правило может быть сформулировано и для спектра фаз.

Теперь займемся вопросом о применении математического аппарата преобразований Фурье к решению основной задачи линейной теории. Задача эта состоит в нахождении отклика линейной системы на произвольное воздействие $x(t)$. Основная идея состоит в том, что легко можно найти отклик системы на синусоидальное воздействие. Для этого достаточно знать *коэффициент передачи* системы, т. е. комплексное отношение

$$K(\omega) = \frac{\dot{U}_2}{\dot{U}_1},$$

где $\dot{U}_1$ и $\dot{U}_2$ — комплексные амплитуды соответственно входного и выходного напряжений. Зная коэффициент передачи, можно определить выходное напряжение по заданному входному

$$\dot{U}_2 = K(\omega) \dot{U}_1. \quad (49)$$

При таких обстоятельствах для решения задачи об отклике на произвольное воздействие $x(t)$ удобно разложить воздействие на синусоидальные составляющие. Именно такое разложение и представляет собой интеграл Фурье (47)

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) e^{i\omega t} d\omega. \quad (50)$$

(Мы снабдили спектр индексом x , чтобы подчеркнуть, что речь идет о спектре функции $x(t)$.)

Отдельное бесконечно малое слагаемое воздействия $x(t)$ имеет вид

$$dx(t) = \frac{1}{2\pi} S_x(\omega) e^{i\omega t} d\omega = d\dot{U}_1 e^{i\omega t}, \quad (51)$$

где

$$d\dot{U}_1 = \frac{1}{2\pi} S_x(\omega) d\omega$$

— бесконечно малая амплитуда. Отклик на синусоидальное колебание (51), поданное на вход, будет также синусоидальным колебанием

$$dy(t) = d\dot{U}_2 e^{i\omega t},$$

причем комплексная бесконечно малая амплитуда $d\dot{U}_2$ на выходе определится по формуле (49)

$$d\dot{U}_2 = K(\omega) d\dot{U}_1.$$

Таким образом, отклик на воздействие $dx(t)$ будет

$$dy(t) = \frac{1}{2\pi} K(\omega) S_x(\omega) e^{i\omega t} d\omega. \quad (52)$$

Чтобы получить полный отклик $y(t)$, нужно просуммировать (52) по всем частотам, т. е. составить (на основании принципа наложения) интеграл

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K(\omega) S_x(\omega) e^{i\omega t} d\omega. \quad (53)$$

Это и есть формула для нахождения отклика линейной системы на произвольное воздействие методом интеграла Фурье.

Заметим, что функция $y(t)$ может быть вообще представлена интегралом Фурье следующего вида (см. формулу (50)):

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_y(\omega) e^{i\omega t} d\omega. \quad (54)$$

Сопоставляя (53) и (54), получаем простое, но важное соотношение

$$S_y(\omega) = K(\omega) S_x(\omega), \quad (55)$$

т. е. спектр отклика равен спектру воздействия, умноженному на коэффициент передачи. Во многих случаях нас интересует не отклик как функция времени $y(t)$, а именно *спектр* отклика $S_y(\omega)$. В этих случаях формула (55) дает сразу нужный нам результат.

Касаясь практической стороны дела, нужно заметить, что вычисление интегралов вида (53) обычными методами затруднительно. Наиболее эффективно применение для этой цели метода, основанного на *теореме о вычетах* (из теории функции комплексного переменного). С этой точки зрения проще в применении метод интеграла Дюамеля.

§ 27. Интеграл Дюамеля

Прежде всего введем понятия *единичной функции* и *единичного импульса*, необходимые для дальнейшего. Единичная функция определяется следующим образом:

$$\sigma(t) = \begin{cases} 0 & \text{при } t < 0, \\ 1 & \text{при } t \geq 0. \end{cases} \quad (56)$$

График единичной функции показан на рис. 42, а.

Физически подобного рода функция реализуется, например, при включении в момент $t = 0$ источника постоянного напряжения. Мож-

но также рассматривать функцию $\sigma(t)$ как изменение проводимости цепи с активным сопротивлением при замыкании ключа, введенного последовательно в цепи. На этом основании единичную функцию $\sigma(t)$ называют иногда «функцией включения».

В более общей форме можно записать определение (56) с запаздывающим аргументом в виде:

$$\sigma(t - \tau) = \begin{cases} 0 & \text{при } t < \tau, \\ 1 & \text{при } t \geq \tau. \end{cases}$$

График этой функции (рис. 42, б) поясняет и физический смысл параметра τ .

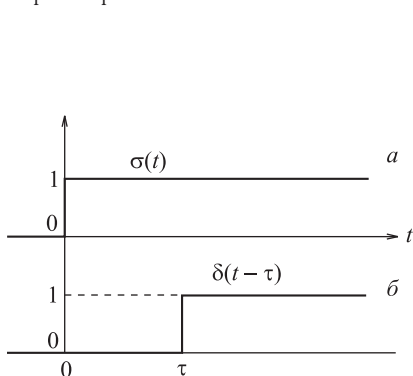


Рис. 42

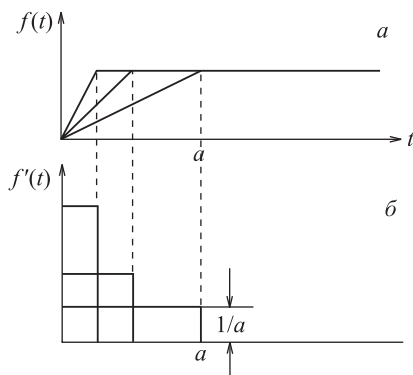


Рис. 43

Единичный импульс $\delta(t)$ определяется как импульс бесконечно малой длительности, но конечной площади, равной единице. Единичный импульс можно выразить следующим интегральным соотношением:

$$\sigma(t) = \int_{-\infty}^t \delta(t) dt. \quad (57)$$

Понятие об единичном импульсе можно составить путем предельного перехода. Рассмотрим вспомогательную функцию:

$$f(t) = \begin{cases} 0 & [t < 0], \\ t/a & [0 < t < a], \\ 1 & [t > a]. \end{cases}$$

Графики функции $f(t)$ для различных значений a показаны на рис. 43, а. Найдем производную функции $f(t)$:

$$f'(t) = \begin{cases} 0 & [t < 0], \\ 1/a & [0 < t < a], \\ 0 & [t > a]. \end{cases}$$

Таким образом, функция $f'(t)$ представляется в виде прямоугольного импульса длительностью a и высотой $1/a$ (рис. 43, б). Площадь такого импульса равна единице независимо от значения a . Переходя к пределу при $a \rightarrow 0$, получим:

$$f(t) \xrightarrow{a \rightarrow 0} \sigma(t), \quad f'(t) \xrightarrow{a \rightarrow 0} \delta(t).$$

В смысле указанного предельного перехода можно записать

$$\delta(t) = \frac{d}{dt}[\sigma(t)]. \quad (58)$$

Формально это соотношение может быть получено непосредственно из (57) путем дифференцирования по верхнему пределу.

Функция $\delta(t)$ введена в теоретическую физику Дираком; она называется дельта-функцией, или функцией Дирака.

Единичный импульс, имеющий бесконечную малую длительность и бесконечно большую высоту, является, конечно, математической абстракцией. Физически можно реализовать только *короткий импульс*, т. е. импульс очень малой длительности, причем смысл терминов «короткий» и «малая длительность» должен быть уточнен, т. е. должно быть указано с чем *сравнивается* длительность импульса.

Отклик системы на единичную функцию называется *переходной функцией* и обозначается $h(t)$. Отклик системы на единичный импульс называется *импульсной реакцией* и обозначается $g(t)$. В силу (58) между переходной функцией и импульсной реакцией существует простая взаимозависимость

$$g(t) = \frac{d}{dt}[h(t)].$$

Если воздействие запаздывает на время τ , то на такое же время запаздывает, очевидно, и отклик. Если воздействие увеличивается в a раз, то в силу линейности системы во столько же раз возрастает и отклик. Эти соотношения можно представить в виде табл. 10.

Таблица 10

Воздействие	$\sigma(t)$	$\sigma(t - \tau)$	$a\sigma(t - \tau)$	$\delta(t)$	$\delta(t - \tau)$	$a\delta(t - \tau)$
Отклик	$h(t)$	$h(t - \tau)$	$ah(t - \tau)$	$g(t)$	$g(t - \tau)$	$ag(t - \tau)$

Если известна переходная функция или импульсная реакция системы, то можно для нахождения отклика системы на произвольное воздействие разложить последнее по единичным функциям или по единичным импульсам. Это делается следующим образом. Пусть дана произвольная функция $x(t)$, равная нулю при $t = 0$; график такой функции показан на рис. 44. Представим себе, что функция $x(t)$ построена из запаздывающих бесконечно малых ступенек, намеченных на том же

рисунке. Каждая ступенька имеет высоту

$$dx = x' d\tau$$

и запаздывает на τ (эти соотношения показаны для одной ступеньки в увеличенном виде на рис. 45). Следовательно, аналитически каждая ступенька выражается как

$$dx(t) = x' \sigma(t - \tau) d\tau. \quad (59)$$

Для того чтобы получить значение функции $x(t)$ в момент времени t , нужно просуммировать все элементарные ступеньки на интервале от нуля до t .

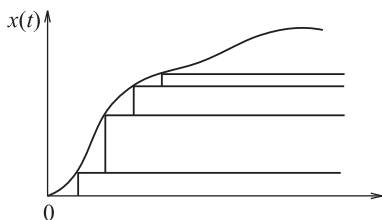


Рис. 44

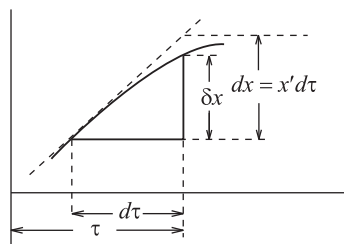


Рис. 45

Это даст для $x(t)$ следующее интегральное выражение:

$$x(t) = \int_0^t x'(\tau) \sigma(t - \tau) d\tau. \quad (60)$$

Отклик системы на элементарную ступеньку (59) будет выражен через переходную функцию

$$dy(t) = x' h(t - \tau) d\tau. \quad (61)$$

Для получения полного отклика нужно суммировать все элементарные отклики, т. е. проинтегрировать (61) от 0 до t

$$y(t) = \int_0^t x'(\tau) h(t - \tau) d\tau. \quad (62)$$

Это и есть формула, дающая решение задачи и известная под названием *интеграла Дюамеля*.

Приведем некоторые видоизменения этой формулы. Прежде всего распространим ее на тот случай, когда функция $x(t)$ имеет скачок при $t = 0$, т. е. когда $x(0) \neq 0$ (рис. 46). В этом случае к интегралу (60) нужно добавить еще ступень конечной высоты $x(0)$, так что вместо (60) запишем

$$x(t) = x(0) \sigma(t) + \int_0^t x'(\tau) \sigma(t - \tau) d\tau.$$

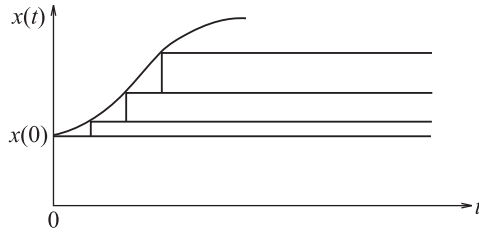


Рис. 46

Это дополнительное слагаемое воздействия вызовет соответствующий отклик, так что вместо (62) получим

$$y(t) = x(0) h(t) + \int_0^t x'(\tau) h(t - \tau) d\tau. \quad (63)$$

Заменой переменной ($t_1 = t - \tau$) и интегрированием по частям можно получить еще три формы, эквивалентные (63):

$$y(t) = x(0) h(t) + \int_0^t x'(t - \tau) h(\tau) d\tau, \quad (64)$$

$$y(t) = h(0) x(t) + \int_0^t h'(\tau) x(t - \tau) d\tau, \quad (65)$$

$$y(t) = h(0) x(t) + \int_0^t h'(t - \tau) x(\tau) d\tau. \quad (66)$$

Одна из формул (63)–(66) может оказаться более удобной для вычислений, в зависимости от вида функций $x(t)$ и $h(t)$. Так, например, если $h(t)$ выражается через экспоненциальные функции, то следует предпочесть формулу (66), так как, во-первых, экспоненциальные функции просто дифференцируются, а, во-вторых, подстановка разностного аргумента приводит к появлению не зависящего от τ множителя, который выносится за знак интеграла. В случае надобности вышеприведенные формулы легко распространяются на случай, когда функция $x(t)$ имеет разрыв не только при $t = 0$, но при $t = t_1, t_2, \dots, t_n$, причем величины скачков рав-

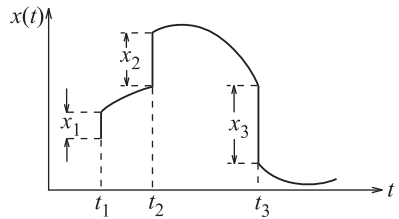


Рис. 47

ны соответственно $x_1, x_2, \dots, x_n$ (рис. 47). В этом случае будем иметь, например, вместо (63)

$$y(t) = \sum_{k=1}^n x_k h(t - t_k) + \int_0^t x'(\tau) h(t - \tau) d\tau \quad (67)$$

Сходным образом может быть сделано и разложение воздействия по единичным импульсам. Это разложение приведет нас к тем же самым формулам.

Заметим, что мы рассматривали специальный вид функции $x(t)$: мы полагали, что $x(t) \equiv 0$ при $t < 0$. Это означает, что интеграл Дюамеля представляет собой аппарат, особенно удобный для исследования всякого рода устанавливающихся процессов, т. е. явлений, разыгрывающихся в системе непосредственно вслед за началом воздействия на систему. (Начало воздействия отнесено к моменту $t = 0$.) Если же мы не желаем воспользоваться этой возможностью и хотим рассматривать отклик на воздействие, начало которого относится к любому моменту t_0 , хотя бы и в отдаленном прошлом, то в формуле (62) нужно взять нижний предел t_0 вместо 0. Соответственно изменятся и все последующие формулы.

§ 28. Частотные и временные характеристики

Отклик линейной системы зависит, во-первых, от оказываемого на систему воздействия, а во-вторых, от свойств самой системы. Из этого следует, что всякая формула, позволяющая найти отклик системы, должна содержать две функции, из которых одна выражает воздействие, а вторая характеризует свойства системы. Рассмотрим с этой точки зрения формулы интеграла Фурье и интеграла Дюамеля. Мы имеем (см. (53) и (62)):

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K(\omega) S_x(\omega) e^{i\omega t} d\omega,$$

$$y(t) = \int_0^t x'(\tau) h(t - \tau) d\tau.$$

¹⁾ Формула (67) может быть записана в компактной общей форме так называемого интеграла Стильтьеса:

$$y(t) = \int_0^t h(t - \tau) d[x(\tau)].$$

Формула (67) представляет собой не что иное, как развернутую запись этого интеграла.

Заметим, что обе формулы представляют собой интегральные выражения. Это обусловлено тем, что они получены на основе принципа наложения путем суммирования элементарных откликов, определяемых подынтегральными выражениями. В интеграле Фурье воздействие представлено его спектром $S_x(\omega)$, в интеграле Дюамеля — самой функцией $x(t)$. В интеграле Фурье свойства системы выражены через коэффициент передачи $K(\omega)$, в интеграле Дюамеля — через переходную функцию $h(t)$. Мы называем функции, служащие для описания свойств системы, вообще характеристиками. Как видим, свойства системы можно описывать по-разному: можно пользоваться различными характеристиками. Коэффициент передачи, являющийся функцией частоты, называется *частотной характеристикой*; переходная функция как функция времени называется *временной характеристикой*.

Понятие и определение той или иной характеристики возникает в известном смысле автоматически, когда мы, намереваясь применить принцип наложения, выбираем некоторый определенный способ разбиения воздействия на слагаемые, т. е. представляем функцию $x(t)$ той или иной суммой или интегралом. Если пересмотреть еще раз рассуждения § 26 и § 27, то станет ясно, что характеристика системы есть не что иное, как функция, выражающая отклик системы на воздействие, выбранное при данном способе решения задачи в качестве элементарного.

В методе интеграла Фурье мы разлагаем воздействие на элементарные синусоидальные составляющие; частотная характеристика — коэффициент передачи $K(\omega)$ — непосредственно выражает отклик системы на синусоидальное воздействие через отношение комплексных амплитуд на входе и на выходе. В методе интеграла Дюамеля мы разлагаем воздействие по единичным функциям, представляя его суммой элементарных запаздывающих ступенек; временная характеристика — переходная функция $h(t)$ — непосредственно выражает отклик системы на воздействие в виде единичной функции.

Если бы мы воспользовались еще каким-либо новым способом разложения воздействия (в принципе разложение можно осуществить бесчисленным множеством способов), то сейчас же возникла бы и соответствующая новая характеристика.

Всякая характеристика выражает свойства системы с определенной, удобной и естественной при данных обстоятельствах точки зрения. Но система остается одной и той же. Различные характеристики описывают по-разному одни и те же свойства, а отсюда следует, что одни характеристики могут быть получены из других, что между различными характеристиками должны существовать вполне определенные связи¹⁾. Установлением этих связей мы сейчас и займемся.

¹⁾ При условии, что различные характеристики отражают свойства системы с одинаковой полнотой, т. е. что различные характеристики содержат одинаковое количество сведений о свойствах системы. Для частотных и временных характеристик дело обстоит именно так.

Пусть на вход системы подается воздействие в виде единичного импульса $\delta(t)$. Для нахождения отклика воспользуемся формулой интеграла Фурье (53). Для этого нужно найти спектр единичного импульса. Мы имеем общую формулу (48)

$$S_x(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt.$$

Но функция $\delta(t)$ равна нулю при всех значениях t , кроме $t = 0$. При этом значении t экспоненциальный множитель обращается в единицу, и мы получаем для единичного импульса

$$S_\delta(\omega) = \int_{-\infty}^{\infty} \delta(t) dt = 1$$

(по определению единичного импульса). Отклик системы на единичный импульс есть импульсная реакция $g(t)$. Таким образом,

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K(\omega) e^{i\omega t} d\omega. \quad (68)$$

Эта формула и устанавливает связь между временной и частотной характеристиками. Как видим, коэффициент передачи $K(\omega)$ является спектром импульсной реакции $g(t)$. Существует и другое соотношение, вытекающее из (68),

$$K(\omega) = \int_{-\infty}^{\infty} g(t) e^{-i\omega t} dt. \quad (69)$$

Таким образом, $K(\omega)$ и $g(t)$ связаны между собой парой преобразований Фурье. Если одна из характеристик известна, то другая может быть найдена с помощью формул (68) или (69).

Рассмотрим некоторые подробности, относящиеся к частотным характеристикам. Коэффициент передачи

$$K(\omega) = \dot{U}_2 / \dot{U}_1$$

есть комплексная величина, которую можно записать в виде

$$K(\omega) = A e^{-i\varphi}.$$

Величина $A = |K|$ выражает отношение (действительных) амплитуд на выходе и на входе; функция $A(\omega)$ называется *амплитудно-частотной характеристикой*. Величина φ выражает фазовый сдвиг между синусоидальными напряжениями на выходе и на входе; функция $\varphi(\omega)$ называется *фазо-частотной характеристикой*. Можно также пользоваться зависимостью A от φ ; эта характеристика называется *амплитудно-фазовой*. График амплитудно-фазовой характеристики строится

в полярных координатах. A есть радиус-вектор, φ — соответствующий полярный угол.

По поводу амплитудных и фазовых частотных характеристик нужно заметить, что они не независимы, а представляют собой модуль и аргумент одной и той же комплексной функции частоты. Основываясь на теории функций комплексного переменного, можно вывести связи общего характера между амплитудно-частотной и фазо-частотной характеристиками. Эти связи довольно сложны и здесь разбираться не будут. Важно лишь иметь в виду, что какое-либо изменение параметров системы, влияющее на амплитудно-частотную характеристику, в общем случае оказывает соответствующее влияние и на фазо-частотную характеристику. Из наличия связи между характеристиками следует также, что нельзя конструировать систему, вполне произвольно задавая в отдельности амплитудно-частотную и фазо-частотную характеристики, так как система с требуемыми характеристиками может оказаться физически неосуществимой.

Нам остается выяснить, как определяются характеристики данной системы. Разумеется, для реальной системы любые характеристики могут быть получены экспериментально. Техника подобных экспериментов обсуждается в курсе радиоизмерений. Мы будем рассматривать вопрос теоретически, основываясь на уравнении системы.

Уравнение, связывающее напряжение на входе и на выходе линейной системы, может быть записано в следующем достаточно общем виде:

$$a_n \frac{d^n u_2}{dt^n} + a_{n-1} \frac{d^{n-1} u_2}{dt^{n-1}} + \dots + a_1 \frac{du_2}{dt} + a_0 u_2 = b_m \frac{d^m u_1}{dt^m} + \dots + b_0 u_1. \quad (70)$$

При установившемся синусоидальном режиме имеем:

$$u_1 = \dot{U}_1 e^{i\omega t}, \quad u_2 = \dot{U}_2 e^{i\omega t}$$

и, следовательно, для синусоидального режима уравнение (70) может быть переписано в виде

$$[a_n(i\omega)^n + a_{n-1}(i\omega)^{n-1} + \dots + a_1 i\omega + a_0] \dot{U}_2 = [b_m(i\omega)^m + \dots + b_0] \dot{U}_1,$$

или, сокращенно,

$$H_2(i\omega) \dot{U}_2 = H_1(i\omega) \dot{U}_1,$$

откуда

$$K(\omega) = \frac{\dot{U}_2}{\dot{U}_1} = \frac{H_1(i\omega)}{H_2(i\omega)} \quad ^1).$$

¹⁾ Мы рассматриваем здесь так называемые системы с сосредоточенными параметрами, описываемые дифференциальными уравнениями в полных производных типа (70). Для этих систем функции H_1 и H_2 представляют собой некоторые многочлены. Для систем же с распределенными параметрами

Итак, для нахождения комплексного коэффициента передачи, являющегося комплексной частотной характеристикой системы, нужно:

- 1) записать уравнение системы, связывающее величины на входе и на выходе;
- 2) заменить $\frac{d^k}{dt^k}$ на $(i\omega)^k$;
- 3) выразить искомый коэффициент передачи отношением полученных многочленов от $i\omega$ в правой и левой частях. (Заметим, что, как видно из этой процедуры, свойства системы отражены частотной характеристикой с той степенью полноты, с какой они представлены уравнением системы. Смысл этого замечания состоит в том, что любое уравнение отображает действительность лишь приближенно; см. § 30.)

Для нахождения переходной характеристики поступим так: запишем прежде всего уравнение (70) в операционной форме

$$(a_n p^n + a_{n-1} p^{n-1} + \dots + a_1 p + a_0) \bar{u}_2 = (b_m p^m + \dots + b_0) \bar{u}_1,$$

или, сокращенно,

$$H_2(p) \bar{u}_2 = H_1(p) \bar{u}_1,$$

где H_1 и H_2 — те же многочлены, что и выше, но аргументом является не $i\omega$, а p ; $\bar{u}_1$ и $\bar{u}_2$ — *изображения* напряжений на входе и на выходе. Мы ищем переходную функцию, т.е. отклик системы на единичную функцию $\sigma(t)$. Изображение этой функции есть единица. Таким образом, переходная функция определится как оригинал для следующего изображения:

$$h(t) \leftrightarrow \frac{H_1(p)}{H_2(p)} = K(p).$$

Достаточно универсальным средством решения этого операционного уравнения является формула Хевисайда

$$h(t) = \frac{H_1(0)}{H_2(0)} + \sum_{k=1}^n \frac{H_1(p_k)}{p_k H_2'(p_k)} e^{p_k t}. \quad (71)$$

Здесь $H'(p_k) = \left. \frac{dH}{dp} \right|_{p=p_k}$, p_k — корень характеристического уравнения

$$H_2(p) = 0$$

функции H_1 и H_2 являются трансцендентными. Системы с сосредоточенными параметрами рассматриваются подробно в гл. 5, а системы с распределенными параметрами — в гл. 6.

(всего имеется n корней). Несколько более простое выражение получается для импульсной реакции $g(t)$

$$g(t) = \frac{dh}{dt} = \sum_{k=1}^n \frac{H_1(p_k)}{H_2'(p_k)} e^{p_k t}. \quad (72)$$

Итак, процедура нахождения временной характеристики состоит в следующем:

1. Уравнение системы записывается в операторной форме путем замены d^k/dt^k на p^k .
2. Составляется выражение для операторного коэффициента передачи

$$K(p) = \frac{H_1(p)}{H_2(p)}.$$

3. Переходная функция находится по формуле (71) (или импульсная реакция по формуле (72)). В некоторых случаях можно воспользоваться специализированными приемами; это будет сделано, например, при анализе неустойчившегося режима линий.

Для составления дифференциальных уравнений удобно пользоваться методом сопротивлений. Метод состоит в том, что элементам схемы приписываются операторные выражения сопротивлений, т. е. R , pL и $1/(pC)$, после чего к схеме применяются законы Кирхгофа.

Покажем применение метода на примере. Пусть дана система, схема которой представлена на рис. 48. Требуется составить дифференциальное уравнение системы. Вводя вспомогательное напряжение u' , можно записать:

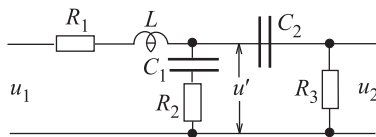


Рис. 48

$$\frac{\bar{u}_2}{\bar{u}_1} = \frac{R_3}{R_3 + \frac{1}{pC_2}}, \quad \frac{\bar{u}'}{\bar{u}_1} = \frac{R_2 + \frac{1}{pC_1}}{R_1 + R_2 + \frac{1}{pC_1} + pL}.$$

Перемножая эти два отношения, исключаем u' :

$$\frac{\bar{u}_2}{\bar{u}_1} = \frac{R_3 \left(R_2 + \frac{1}{pC_1} \right)}{\left(R_3 + \frac{1}{pC_2} \right) \left(R_1 + R_2 + \frac{1}{pC_1} + pL \right)}.$$

Раскрывая скобки и избавляясь от отрицательных степеней p , получаем

$$\frac{\bar{u}_2}{\bar{u}_1} = \frac{H_1(p)}{H_2(p)} = \frac{b_2 p^2 + b_1 p}{a_3 p^3 + a_2 p^2 + a_1 p + a_0},$$

где

$$\begin{aligned}a_0 &= 1, \quad a_1 = R_3 C_2 + (R_1 + R_2)C, \\a_2 &= R_3(R_1 + R_2)C_1 C_2 + LC_1, \\a_3 &= R_3 LC_1 C_2, \\b_1 &= R_3 C_2, \quad b_2 = R_2 R_3 C_1 C_2.\end{aligned}$$

В обычной форме уравнение (70) будет иметь вид

$$a_3 \frac{d^3 u_2}{dt^3} + a_2 \frac{d^2 u_2}{dt^2} + a_1 \frac{du_2}{dt} + a_0 u_2 = b_2 \frac{d^2 u_1}{dt^2} + b_1 \frac{du_1}{dt}.$$

В заключение параграфа рассмотрим некоторые полезные предельные соотношения между частотными и временными характеристиками. Эти соотношения можно вывести следующим образом. Представим переходную функцию $h(t)$ степенным рядом

$$h(t) = \alpha_0 + \alpha_1 \frac{t}{1!} + \alpha_2 \frac{t^2}{2!} + \dots \quad (t > 0).$$

Изображение этой функции будет

$$\bar{h}(p) = \alpha_0 + \alpha_1 \frac{1}{p} + \alpha_2 \frac{1}{p^2} + \dots$$

Перейдем к пределу при $t \rightarrow 0$. Это даст

$$\lim_{t \rightarrow 0} h(t) = \lim_{p \rightarrow \infty} \bar{h}(p) \quad (= \alpha_0).$$

Но изображение переходной функции есть

$$\bar{h}(p) = \frac{H_1(p)}{H_2(p)} = K(p).$$

При предельном переходе безразлично, какая переменная стоит в аргументе функции. Поэтому вместо $K(p)$ можем рассматривать для предельного перехода $K(\omega)$. Можно следовательно, записать

$$\lim_{t \rightarrow 0} h(t) = \lim_{\omega \rightarrow \infty} K(\omega). \quad (73)$$

Другое соотношение получают непосредственно из (71), переходя к пределу при $t \rightarrow \infty$. При этом сумма стремится к нулю, так как для всякой реальной системы корни p_k характеристического уравнения имеют отрицательные вещественные части и, значит, члены суммы убывают по экспоненциальному закону.

Таким образом,

$$\lim_{t \rightarrow \infty} h(t) = \frac{H_1(0)}{H_2(0)}$$

или

$$\lim_{t \rightarrow \infty} h(t) = \lim_{\omega \rightarrow 0} K(\omega). \quad (74)$$

Формулы (73) и (74) позволяют сразу, не решая задачи полностью, выяснить поведение системы при предельных значениях частоты, если известно поведение системы в начальный момент ($t \rightarrow 0$) и в установившемся состоянии ($t \rightarrow \infty$) и наоборот. Во многих случаях это представляет интерес.

§ 29. Основы теории спектров

Для пользования спектральными представлениями вообще и методом интеграла Фурье в частности необходимо владеть основами теории спектров, которые и даны в кратком изложении в этом параграфе.

Мы приведем несколько важнейших теорем о спектрах, т. е. теорем, выражающих свойства преобразований Фурье.

Основой послужат формулы пары преобразований Фурье:

$$S(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt,$$

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{i\omega t} d\omega.$$

Теорема сложения. Пусть

$$x(t) = \sum x_k(t).$$

Тогда

$$S(\omega) = \int_{-\infty}^{\infty} \sum x_k(t) e^{-i\omega t} dt = \sum \int_{-\infty}^{\infty} x_k(t) e^{-i\omega t} dt = \sum S_k(\omega),$$

т. е. спектр суммы функций равен сумме спектров слагаемых.

Итак, преобразование Фурье линейно.

Теорема об изменении масштаба. Пусть

$$x_a(t) = x(at).$$

Тогда

$$S_a(\omega) = \int_{-\infty}^{\infty} x_a(t) e^{-i\omega t} dt = \int_{-\infty}^{\infty} x(at) e^{-i\omega t} dt = \frac{1}{a} \int_{-\infty}^{\infty} x(t_1) e^{-i\frac{\omega}{a} t_1} dt_1,$$

т. е.

$$S_a(\omega) = \frac{1}{a} S\left(\frac{\omega}{a}\right).$$

Таким образом, при изменении масштаба времени в a раз масштаб частот для спектра меняется в $1/a$ раз.

Эта теорема имеет важный физический смысл. Она показывает, что единственный способ *сжатия* спектра (т. е. сокращения ширины спектра) без изменения его характера состоит в том, чтобы растянуть

явление во времени. То же относится и к расширению спектра; для расширения спектра нужно ускорить ход явления. Технически такого рода преобразование можно выполнить так: записать функцию $x(t)$ (например, на магнитную ленту) и затем воспроизвести запись с иной скоростью (быстрее или медленнее). Такое преобразование обратимо: проделав его дважды, можно восстановить функцию в ее первоначальном временном масштабе. Любой способ сжатия спектра, не затрагивающий масштаба времени, должен повлечь за собой невосстановимое искажение спектра.

Теорема запаздывания. Пусть

$$x_\tau(t) = x(t - \tau).$$

Тогда

$$S_\tau(\omega) = \int_{-\infty}^{\infty} x_\tau(t) e^{-i\omega t} dt = \int_{-\infty}^{\infty} x(t - \tau) e^{-i\omega t} dt.$$

Заменяя переменную по формуле $t_1 = t - \tau$, найдем

$$S_\tau(\omega) = e^{-i\omega\tau} \int_{-\infty}^{\infty} x(t_1) e^{-i\omega t_1} dt_1,$$

т. е.

$$S_\tau(\omega) = e^{-i\omega\tau} S(\omega).$$

Таким образом, чтобы получить спектр функции $x(t - \tau)$, изображающей процесс $x(t)$, запаздывающий на время τ , достаточно умножить спектр $S(\omega)$ процесса $x(t)$ на $e^{-i\omega\tau}$. Заметим, что умножение на $e^{-i\omega\tau}$ изменяет только спектральную плотность фаз; спектральная плотность амплитуд остается при смещении функции во времени неизменной, что видно из следующего соотношения:

$$\Phi_\tau(\omega) = |S_\tau(\omega)| = |e^{-i\omega\tau} S(\omega)| = |S(\omega)| = \Phi(\omega).$$

Спектр производной и интеграла. Найдем спектр функции

$$x'(t) = \frac{dx}{dt}.$$

Интегрируя по частям выражение

$$S_{(1)}(\omega) = \int_{-\infty}^{\infty} x'(t) e^{-i\omega t} dt,$$

получим

$$S_{(1)}(\omega) = x(t) e^{-i\omega t} \Big|_{-\infty}^{\infty} + i\omega \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt.$$

Если

$$\lim_{t \rightarrow \pm\infty} x(t) = 0,$$

то

$$S_{(1)}(\omega) = i\omega S(\omega).$$

Поступая аналогичным образом, получим для n -й производной

$$S_{(n)}(\omega) = \int_{-\infty}^{\infty} \frac{d^n x}{dt^n} e^{-i\omega t} dt = (i\omega)^n S(\omega),$$

если для всех производных до $(n-1)$ -го порядка включительно выполняется условие

$$\lim_{t \rightarrow \pm\infty} \frac{d^k x}{dt^k} = 0.$$

Применяя тот же прием, найдем, что спектр $S_{(-1)}(\omega)$ интеграла от функции $x(t)$, взятого в пределах от $-\infty$ до t , выражается соотношением

$$S_{(-1)}(\omega) = \frac{1}{i\omega} S(\omega)$$

при условии, что

$$\int_{-\infty}^{\infty} x(t) dt = 0.$$

Это условие выполняется, например, для всякой нечетной функции.

Теорема о спектре произведения (дается без доказательства). Пусть

$$x(t) = x_1(t) x_2(t).$$

Тогда спектр функции $x(t)$ выражается через спектры сомножителей следующей формулой:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_1(\nu) S_2(\omega - \nu) d\nu.$$

Эта формула в особенности удобна для вычисления модуляционных спектров.

Теорема о спектре свертки (дается без доказательства). Сверткой (или складкой) двух функций называется интеграл

$$x(t) = \int_{-\infty}^{\infty} x_1(\tau) x_2(t - \tau) d\tau.$$

Спектр свертки выражается произведением спектров свертываемых функций, т. е.

$$S(\omega) = S_1(\omega) S_2(\omega).$$

Теорема энергий. Возьмем выражение интеграла Фурье.

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{i\omega t} d\omega,$$

умножим обе части на $x(t)$ и проинтегрируем в бесконечных пределах. Получим

$$\int_{-\infty}^{\infty} x^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) d\omega \int_{-\infty}^{\infty} x(t) e^{i\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) S(-\omega) d\omega.$$

Но $S(-\omega)$ и $S(\omega)$ — комплексно-сопряженные величины (т. е. величины, различающиеся только знаком мнимой части). Поэтому

$$S(\omega) S(-\omega) = |S(\omega)|^2 = \Phi^2(\omega).$$

Итак,

$$\int_{-\infty}^{\infty} x^2 dt = \frac{1}{\pi} \int_0^{\infty} \Phi^2(\omega) d\omega$$

(соотношение, аналогичное равенству Парсеваля из теории рядов Фурье).

Смысл этого соотношения состоит в том, что энергию некоторого процесса можно вычислять двояким, способом: либо интегрируя квадрат функций времени (выражающий мгновенную мощность), либо интегрируя квадрат амплитудного спектра.

Соотношение между Δf и Δt . Рассмотрим одно из важнейших общих положений теории спектров. Если Δf означает ширину спектра некоторого процесса, а Δt — его длительность, то справедливо следующее соотношение:

$$\Delta f \Delta t = \mu, \quad (75)$$

где μ — постоянная, порядка единицы. Смысл этого соотношения состоит в том, что ширина спектра обратно пропорциональна длительности процесса, и более того: ширина спектра по *порядку величины* определяется как $1/\Delta t$. Общий смысл этого утверждения подсказывается уже теоремой об изменении масштаба. Более определенных утверждений относительно связи между Δf и Δt нельзя сделать прежде всего потому, что произведение $\Delta f \Delta t$ зависит от вида функции. Это произведение имеет *порядок* единицы, но может изменяться в довольно широких пределах. Кроме того, нужно учесть, что для применения соотношения (75) необходимо так или иначе условиться о том, что именно мы называем шириной спектра Δf и длительностью Δt . Дело в том, что в общем случае как функция времени $x(t)$, так и ее спектр $\Phi(\omega)$ могут простирались как угодно далеко как во времени, так и по частоте. Выбор критерия для оценки Δf и Δt есть в значительной мере дело нашего произвола. Между тем ясно, что от выбора этого

критерия зависит величина как Δf и Δt , так и их произведения. Мы не будем исследовать этот вопрос в общем виде, но в дальнейшем поясним смысл и способ применения соотношения (75) на примерах.

Спектр короткого импульса. Рассмотрим короткий импульс, начинающийся при $t = 0$ и заканчивающийся при $t = \tau$. Вне интервала $0 < t < \tau$ функция, выражающая импульс, тождественно равна нулю; в интервале $0 < t < \tau$ она положительна и обозначена через $x(t)$ (рис. 49). Спектр такого импульса выражается формулой (обратите внимание на пределы интегрирования)

$$S(\omega) = \int_0^{\tau} x(t) e^{-i\omega t} dt.$$

Положим, что длительность импульса очень мала, так что экспоненциальный множитель в подынтегральном выражении, принимающий значения от 1 до $e^{-i\omega t}$, можно приближенно считать равным единице. Тогда

$$S(\omega) \approx \int_0^{\tau} x(t) dt = q,$$

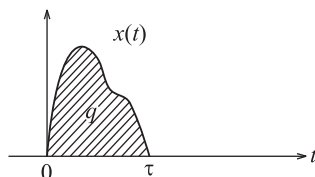


Рис. 49

где через q обозначена *площадь* импульса. Получается, что короткий импульс имеет равномерный спектр и притом не зависящий от вида функции $x(t)$. Этот результат нам уже знаком: он получен в § 28 для бесконечно короткого единичного импульса. Но там результат был точным; нужно выяснить смысл приближения, которое делается в рассматриваемом теперь случае импульса хотя и малой, но конечной длительности. Мы положили при выводе

$$e^{-i\omega\tau} \approx 1.$$

Но это приближение справедливо при

$$\omega\tau \ll 1.$$

Следовательно, спектральная плотность $S(\omega)$ остается постоянной и равной q вплоть до частот, сравнимых по порядку величины с $1/\tau$. При этих частотах она начинает убывать, так что график спектра имеет вид, как на рис. 50. Таким образом, величина $1/\tau$ *по порядку величины* определяет ширину спектра импульса. Но τ — это не что иное, как длительность импульса. Мы получили, следовательно, иллюстрацию общего положения о соотношении между длительностью и шириной спектра. Одно из практических применений этого соотношения состоит в том, что можно сразу грубо оценить ширину спектра одиночного импульса или последовательности импульсов, зная только длительность отдельного импульса. Так, если длительность импульса (независимо от формы) составляет одну микросекунду, то ширина спектра имеет

порядок 1 МГц. Заметим, что переход к пределу при $\tau \rightarrow 0$ и при $q = \text{const} = 1$ даст определенный в § 28 единичный импульс.

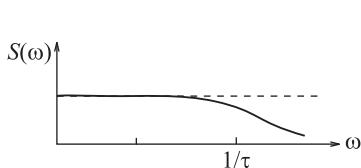


Рис. 50

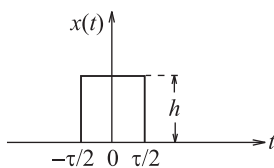


Рис. 51

Спектр прямоугольного импульса. Вычислим спектр прямоугольного импульса, записанного аналитически следующим образом:

$$x(t) = \begin{cases} 0 & \left[t < -\frac{\tau}{2} \right], \\ h & \left[-\frac{\tau}{2} < t < \frac{\tau}{2} \right], \\ 0 & \left[t > \frac{\tau}{2} \right], \end{cases}$$

где h — высота импульса; τ — длительность (рис. 51).

По общей формуле имеем для спектра

$$S(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt = h \int_{-\tau/2}^{\tau/2} e^{-i\omega t} dt = \frac{h}{-i\omega} e^{-i\omega t} \Big|_{-\tau/2}^{\tau/2}.$$

Подставляя пределы и вводя $q = h\tau$, получим

$$S(\omega) = q \frac{\sin \omega \frac{\tau}{2}}{\omega \frac{\tau}{2}}. \quad (76)$$

График модуля спектра $\Phi(\omega)$ изображен на рис. 52. Спектральная плотность обращается в нуль при

$$\omega \frac{\tau}{2} = n\pi \quad (n = 1, 2, 3, \dots).$$

При $\omega = 0$ $S(\omega) = q$.

Рассмотрим произведение $\Delta f \Delta t$. Длительность Δt в нашем случае вполне определена: она равна τ . Что же касается спектра, то он безграничен, хотя спектральная плотность и убывает с частотой. Будем считать (совершенно произвольно), что верхней границей спектра, определяющей его ширину, является частота

$$\omega_1 = \frac{2\pi}{\tau},$$

при которой спектральная плотность первый раз обращается в нуль.

Тогда $\Delta f = \frac{1}{2\pi} \omega_1$ и

$$\Delta f \Delta t = \frac{1}{2\pi} \omega_1 \tau = 1.$$

Делая в формуле (76) предельный переход при $\tau \rightarrow 0$ (и увеличивая соответственно h так, чтобы $q = h\tau$ оставалось неизменным), получим равномерный спектр с постоянной на всех частотах спектральной плотностью, равной площади импульса q .

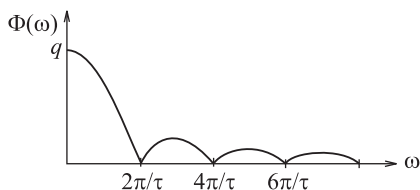


Рис. 52

Спектр «колокольного» импульса. Пусть импульс описывается функцией

$$x(t) = e^{-\beta^2 t^2}.$$

График этой функции имеет колоколообразную форму, откуда и название этого рода импульса. Найдем его спектр

$$S(\omega) = \int_{-\infty}^{\infty} e^{-\beta^2 t^2} e^{-i\omega t} dt.$$

Этот интеграл вычисляется путем приведения показателя к полному квадрату

$$S(\omega) = e^{-\frac{\omega^2}{4\beta^2}} \int_{-\infty}^{\infty} e^{-(\beta t + i\frac{\omega}{2\beta})^2} dt = \frac{2}{\beta} e^{-\frac{\omega^2}{4\beta^2}} \int_0^{\infty} e^{-x^2} dx.$$

Последний интеграл известен; он равен $\frac{1}{2}\sqrt{\pi}$. Итак,

$$S(\omega) = \frac{\sqrt{\pi}}{\beta} e^{-\frac{\omega^2}{4\beta^2}}.$$

Рассмотренный пример интересен тем, что спектр выражается такой же функцией *частоты*, как и исходная функция времени. Другой интересной особенностью колокольного импульса является то, что из всех известных форм импульсов для него получается наименьшее произведение $\Delta f \Delta t$.

Спектр модулированного колебания. Найдем спектр колебания, получаемого при балансной амплитудной модуляции синусоидального

колебания несущей частоты ω_0 произвольной функцией $f(t)$. Модулированное колебание как функция времени запишется в виде

$$x(t) = f(t) \cos \omega_0 t.$$

Спектр этой функции

$$\begin{aligned} S_x(\omega) &= \int_{-\infty}^{\infty} f(t) \cos \omega_0 t e^{-i\omega t} dt = \\ &= \frac{1}{2} \int_{-\infty}^{\infty} f(t) e^{-i(\omega - \omega_0)t} dt + \frac{1}{2} \int_{-\infty}^{\infty} f(t) e^{-i(\omega_0 + \omega)t} dt = \\ &= \frac{1}{2} S_f(\omega - \omega_0) + \frac{1}{2} S_f(\omega_0 + \omega), \end{aligned}$$

или (так как $|S(\omega)| = |S(-\omega)| = \Phi(\omega)$)

$$\Phi_x(\omega) = \frac{1}{2} \Phi_f(\omega_0 - \omega) + \frac{1}{2} \Phi_f(\omega_0 + \omega).$$

Итак, если спектр модулирующей функции $f(t)$ сплошной, то и спектр модулированного колебания сплошной; он состоит из двух боковых полос, повторяющих спектр Φ_f , но расположенных симметрично относительно несущей частоты (рис. 53). Таким изображением модуляционного спектра мы уже пользовались в § 13.

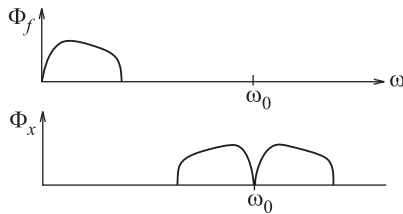


Рис. 53

Спектр отрезка синусоиды. Рассмотрим функцию, представляющую собой отрезок синусоиды, содержащий целое число периодов n , т. е.

$$x(t) = \begin{cases} 0 & \left[t < -\frac{1}{2}nT \right], \\ \cos \omega_0 t & \left[-\frac{1}{2}nT < t < \frac{1}{2}nT \right], \\ 0 & \left[t > \frac{1}{2}nT \right], \end{cases}$$

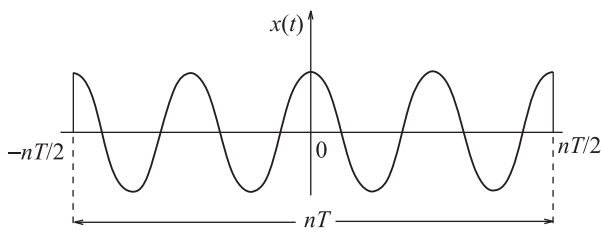


Рис. 54

(рис. 54). Спектр этой функции

$$S(\omega) = \int_{-\frac{1}{2}nT}^{\frac{1}{2}nT} \cos \omega_0 t e^{-i\omega t} dt = \frac{e^{-i\omega t}}{\omega_0^2 - \omega^2} (-i\omega \cos \omega_0 t + \omega_0 \sin \omega_0 t) \Big|_{-\frac{1}{2}nT}^{\frac{1}{2}nT}.$$

Так как $T = \frac{2\pi}{\omega_0}$ и $\frac{1}{2}n\omega_0 T = n\pi$, то

$$S(\omega) = (-1)^n \frac{2\omega}{\omega_0^2 - \omega^2} \sin n\pi \frac{\omega}{\omega_0},$$

откуда

$$\Phi(\omega) = \frac{2}{\omega_0} = \frac{\omega/\omega_0}{\left|1 - \left(\frac{\omega}{\omega_0}\right)^2\right|} \left| \sin n\pi \frac{\omega}{\omega_0} \right|. \quad (77)$$

При $\omega = \omega_0$ получается максимум; раскрыв неопределенность по обычным правилам, находим

$$\Phi(\omega_0) = \frac{n\pi}{\omega_0}. \quad (78)$$

Разбираемый пример имеет очень большое значение и заслуживает подробного обсуждения.

Синусоидальная функция есть периодическая функция. Всякая периодическая функция удовлетворяет определению

$$x(t) = x(t + nT),$$

где T — период, а n — любое целое число.

Следовательно, периодическая функция определена для всех значений аргумента. Физически это означает, что процесс, представляемый периодической функцией времени, существует *вечно*. Ясно, что в действительности таких процессов быть не может и, стало быть, периодическая функция есть математическая абстракция, не соответствующая никакому реальному физическому процессу. Реальный процесс всегда имеет начало и конец, т. е. конечную длительность. Если на протяжении своей длительности процесс выражается периодической функцией,

то это значит, что процесс в целом выражается *отрезком* периодической функции. Таким образом, рассматриваемый отрезок синусоиды — эта та *реальная* форма синусоидального колебания, с которой можно встретиться в действительности. Весь вопрос в том, какова длительность отрезка. Длительность отрезка синусоиды естественно измерять числом периодов. Так как отрезок синусоиды — непериодический процесс, то он должен обладать сплошным спектром. Сплошной спектр отрезка синусоиды выражается формулой (77). Но если неограниченно увеличивать длительность отрезка, устремляя число периодов n к бесконечности, то в пределе мы получим синусоиду в точном смысле математического определения периодической функции. Спектр синусоиды представляется одной-единственной спектральной линией на соответствующей частоте.

Итак, при увеличении длительности отрезка синусоиды его сплошной спектр должен вырождаться в пределе в одну спектральную линию. Этот предельный переход мы и проследим. Для этого предварительно перейдем к относительной спектральной плотности, разделив (77) на (78),

$$\Psi(\omega) = \frac{\Phi(\omega)}{\Phi(\omega_0)} = \frac{2}{n\pi} \frac{\omega/\omega_0}{\left|1 - \left(\frac{\omega}{\omega_0}\right)^2\right|} \left|\sin n\pi \frac{\omega}{\omega_0}\right|.$$

Это делается для того, чтобы привести максимальную ординату кривой спектральной плотности к постоянной (не зависящей от n) величине, равной единице (без такого приведения, как видно из (78), максимальная ордината растет пропорционально n , что неудобно с точки зрения построения графиков).

На рис. 55 построены графики относительной спектральной плотности для различных значений n . При $n = 1$ спектр получается очень широкий и расплывчатый. Чем больше n , тем лучше выявляется периодический характер явления, тем уже спектр. Для реального процесса, характеризующегося отрезком синусоиды, спектр представляется не бесконечно тонкой спектральной *линией*, а узкой *полоской*. Ширина спектральной полоски связана с длительностью отрезка все той же универсальной зависимостью

$$\Delta f \Delta t = \mu.$$

В данном случае легко найти значение произведения $\Delta f \Delta t$, если условиться принять за ширину спектра ширину главного максимума, т. е. интервал между двумя нулями. Соответствующие значения частоты найдем из уравнения

$$\sin n\pi \frac{\omega_0 \pm \Delta\omega}{\omega_0} = 0$$

(рис. 56) или

$$n\pi \left(1 \pm \frac{\Delta\omega}{\omega_0}\right) = k\pi.$$

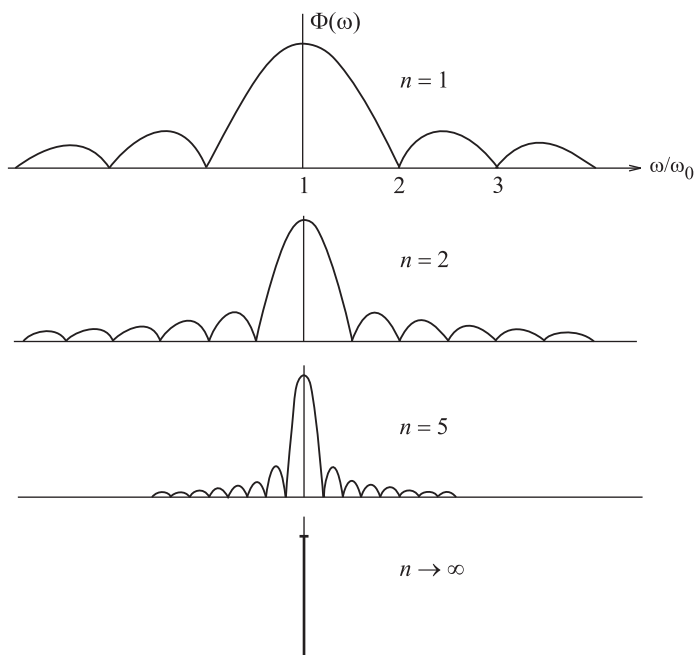


Рис. 55

Дадим k значения $n \pm 1$ — это соответствует ближайшим к максимуму нулям спектральной плотности. Получим

$$\Delta\omega = \omega_0/n,$$

откуда ширина главного максимума

$$\Delta f = \frac{2\Delta\omega}{2\pi} = \frac{\omega_0}{n\pi} = \frac{2f_0}{n}.$$

Длительность отрезка

$$\Delta t = nT = \frac{2\pi n}{\omega_0} = \frac{n}{f_0}.$$

Итак, в нашем случае

$$\Delta f \Delta t = 2.$$

Ширина спектральной полосы равна

$$\Delta f = 2/\Delta t,$$

т.е. зависит только от длительности отрезка. При $\Delta t = 1$ с ширина спектральной полосы составляет 2 Гц. Для телеграфной точки длительностью 1 мс ширина спектра составляет уже 2 кГц. Иногда удобнее

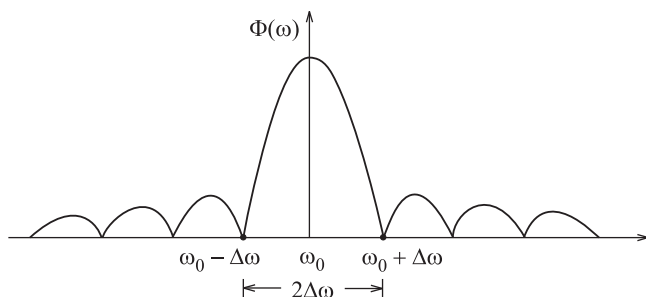


Рис. 56

обращаться с относительной шириной спектра, которая равна

$$\frac{\Delta f}{f_0} = \frac{2\Delta\omega}{\omega_0} = \frac{2}{n}.$$

Так, при $f_0 = 1$ МГц на протяжении телеграфной точки длительностью 1 мс уложится 1000 периодов; относительная ширина спектра будет $2 \cdot 10^{-3}$.

Глава 5

СИСТЕМЫ С СОСРЕДОТОЧЕННЫМИ ПАРАМЕТРАМИ

§ 30. Характеристика и назначение рассматриваемых систем

К числу параметров линейной электрической цепи, т. е. физических констант, определяющих ее свойства, относятся сопротивление, индуктивность и емкость. Эти параметры в общем случае распределены по цепи. Так, например, каждый элемент длины dl одиночного проводника обладает сопротивлением dR , индуктивностью dL , емкостью (относительно земли или других проводников) dC . Проводник длиной l , взятый в целом, имеет, конечно, некоторые суммарные параметры R , L и C , но эти параметры *распределены* по длине проводника равномерно, а может быть и неравномерно, т. е. в некоторых местах, например, емкость, отнесенная к единице длины, может быть больше или меньше, чем в других местах.

В целом ряде случаев можно отвлечься от рассмотрения распределенных параметров и полагать, что суммарные параметры цепи сосредоточены в определенных местах, т. е., например, что все сопротивление провода сосредоточено в его начале или в конце, или что емкость подключена только в одной точке провода и т. п. Такое представление заведомо отличается от действительности, но может служить удовлетворительным приближением к ней. Мы охотно пользуемся такого рода приближением, т. е. заменяем распределенные параметры сосредоточенными, потому что это упрощает как физическую картину явлений, так и их математическое описание. В самом деле, считая параметры распределенными, мы вводим для токов и напряжений зависимость от координат, а это значит, что система с распределенными параметрами должна описываться уравнением в частных производных (по времени и по координатам). В то же время для системы с сосредоточенными параметрами зависимость от координат выпадает, токи и напряжения зависят только от времени, и такая система описывается значительно более простым обыкновенным уравнением (т. е. уравнением в полных

производных, взятых по времени). Именно поэтому мы и рассматриваем в первую очередь системы с сосредоточенными параметрами как более простые.

Представление физической системы, как системы с сосредоточенными параметрами, всегда подразумевает некоторое уклонение от действительности. Мы не можем пока оценить степень этого уклонения и установить критерий, на основе которого можно было бы судить о допустимости приближенного представления реальной системы в виде системы с сосредоточенными параметрами. Для этого нужно рассмотреть те или иные схемы с волновой точки зрения. К этому мы будем подготовлены лишь на основе материала шестой главы. Пока что возьмем на веру, что приближенное рассмотрение действительных систем в виде систем с сосредоточенными параметрами вполне допустимо во многих случаях и что интересующие нас черты явлений при этом приближении не теряются.

Итак, будем представлять действительные устройства в виде схем, составленных из *элементов*, соединенных проводниками с нулевым сопротивлением. Под элементами понимаются двухполюсники, каждый из которых обладает одним и только одним свойством: сопротивлением, индуктивностью или же емкостью. Элементы называются так же, как и свойства, которыми они обладают: мы говорим, что схема составлена из сопротивлений, индуктивностей и емкостей. Свойства элементов определяются уравнениями, связывающими ток, протекающий через элемент, с напряжением на его концах. Эти уравнения таковы:

сопротивление

$$u = Ri,$$

индуктивность

$$u = L \frac{di}{dt},$$

емкость

$$u = \frac{1}{C} \int i dt = \frac{q}{C}.$$

Условное изображение названных элементов на схемах известно, подчеркнем лишь одно важное обстоятельство, непонимание которого может повести к недоразумениям. Дело в том, что мы изображаем на схеме только те элементы, которые учитываем в своих расчетах; схема не содержит никаких *подразумеваемых элементов*. И в этом смысле всякая схема представляет собой лишь приближенное отображение свойств действительного устройства. Необходимая степень приближения зависит от характера задач, которые мы себе ставим. Можно улучшать приближение; при этом схема по своим свойствам приближается к свойствам реального устройства, так сказать, асимптотически. Поясним это на примере. На рис. 57, *a* схема из источника эдс E индуктивности L и емкости C должна представлять реальное устройство из генератора, катушки и конденсатора. Но реальный

генератор обладает внутренним сопротивлением; если оно должно быть учтено, то его нужно показать на схеме (рис. 57, б). Реальная катушка обладает не только индуктивностью, но и активным сопротивлением R_L — оно показано на схеме рис. 57, в. Конденсатор обладает утечкой; в хорошо выполненном воздушном конденсаторе утечкой можно пренебречь в подавляющем большинстве случаев. Но в случае твердого диэлектрика его активную проводимость приходится иногда учитывать. Соответственно усовершенствованная схема показана на рис. 57, г. Далее, можно, если это требуется, учесть емкость обмотки и также показать ее на схеме и т. д.

Заметим, что в технике принимают специальные меры к тому, чтобы различные детали, из которых собираются схемы, обладали по возможности свойствами чистых элементов. Так, например, сопротивление стараются сделать безындукционным; для этого применяют непроволочные сопротивления малых размеров, а проволочные сопротивления изготавливают с бифилярной намоткой. Стремятся уменьшить активное сопротивление высокочастотных катушек, изготавливая их из многожильного провода и т. п.

Возникает естественный вопрос: насколько нужно усложнять изображаемую схему, что именно нужно учитывать с целью приближения к действительности?

В общем виде на этот вопрос можно ответить только так: схема должна быть настолько подробной, чтобы результаты ее теоретического анализа не расходились существенно с результатами физического эксперимента над действительным устройством. Итак, и в данном случае критерием является опыт. Опираясь на опыт, мы умеем в общем правильно составлять схемы. Но при освоении новых областей старый опыт может оказаться недостаточным. Так, например, многие хорошо изученные схемы ведут себя совершенно неожиданно, когда мы заставляем их работать на ультракоротких волнах, а тем более на сверхвысоких частотах. Это объясняется именно тем, что неучитываемые в обычных условиях так называемые паразитные параметры (например, какая-нибудь емкостная проводимость между проводами, ничтожная на средних радиочастотах) начинают оказывать решающее влияние на свч. Чтобы привести расчет в соответствие с опытом, приходит-

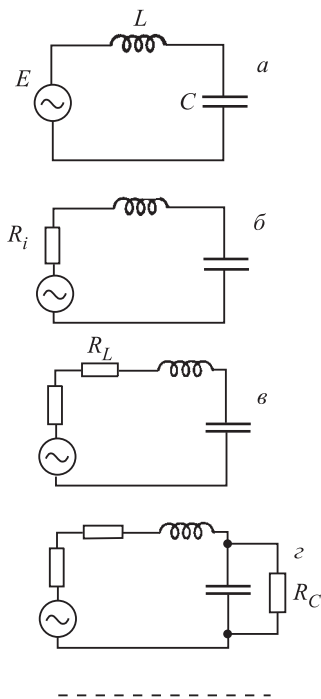


Рис. 57

ся, с одной стороны, совершенствовать принципиальные схемы, вводя в них дополнительные параметры, а с другой стороны, принимать меры к ослаблению влияния новых параметров путем изменения способов монтажа, конструкции деталей, ламп и т.п. В результате устройства свч даже по внешнему виду разительно отличаются от обычных радиотехнических устройств.

Обычный порядок исследования того или иного радиотехнического устройства или явления состоит в том, что сначала составляется принципиальная схема, а затем на основании этой схемы составляется дифференциальное уравнение, те или иные решения которого и дают ответ на интересующие нас вопросы. Из этого следует, что исследуемые уравнения описывают свойства реальной системы с совершенно такой же степенью полноты, с точно таким же приближением, как и принципиальная схема, на основании которой уравнение составлено.

То же самое относится и к любым характеристикам системы, так как всякая характеристика представляет собой не что иное, как некоторую специализированную форму решения уравнения системы.

Обратимся теперь к вопросу о назначении различных линейных систем с сосредоточенными постоянными параметрами. Выше уже говорилось (§ 23), что общим назначением таких систем является *передача* электрических возмущений от предшествующего звена сложного радиотехнического устройства к последующему. Однако на условия этой передачи могут быть наложены самые различные требования. Перечислим здесь наиболее часто встречающиеся.

В некоторых случаях требуется, чтобы отклик системы по возможности точно повторял воздействие, т.е. чтобы система не вносила

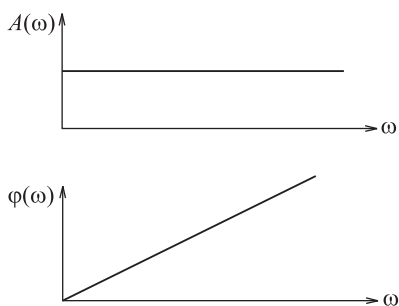


Рис. 58

искажений. Такое требование ставится, в частности, в отношении усилителей. Если достаточно сохранения спектра амплитуд, то для этого нужно лишь обеспечить постоянную (выражаемую горизонтальной прямой) амплитудно-частотную характеристику. Если же требуется сохранение *формы* воздействия, то нужно обеспечить постоянную амплитудно-частотную и линейную фазо-частотную характеристику (рис. 58). Выражая это же требование на временном

языке, нужно потребовать, чтобы переходная характеристика наименьшим образом отличалась от единичной функции.

В других случаях требуется, чтобы отклик некоторым определенным образом отличался от воздействия. Так, например, часто требуется устранить из спектра воздействия определенные полосы частот, другие же полосы сохранить. Эту задачу решают электрические фильтры. Их

действие описывается частотными характеристиками, которым может быть придан самый различный вид. В радиосвязи, в частности, большое значение имеют полосовые фильтры, пропускающие полосу частот, в которой заключен весь спектр принимаемого сигнала, и не пропускающие всех остальных частот. Такой фильтр играет роль частотно-избирательного устройства, идеальная частотная характеристика которого должна была бы иметь вид, показанный на рис. 59.

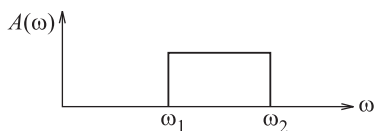


Рис. 59

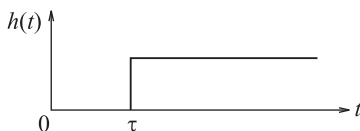


Рис. 60

Иногда требуется, чтобы отклик системы повторял воздействие, но с запазданием на определенное время. Системы с такими свойствами называются задерживающими устройствами. Частотные характеристики задерживающего устройства имеют вид рис. 58; задержка тем больше, чем круче идет фазо-частотная характеристика ¹⁾. Проще говорить о задерживающих устройствах на временном языке; переходная характеристика идеального задерживающего (на время τ) устройства должна выглядеть так, как показано на рис. 60.

Часто встречается надобность в получении отклика, пропорционального производной или интегралу от воздействия. Схемы, отвечающие этому требованию, называются соответственно дифференцирующими или интегрирующими.

Наконец, в наиболее общем случае можно потребовать, чтобы при определенном воздействии отклик имел некоторую заданную форму

¹⁾ Докажем это. Пусть воздействие

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) e^{i\omega t} d\omega.$$

Отклик должен отличаться от воздействия только запаздыванием и, может быть, постоянным множителем

$$\begin{aligned} y(t) &= ax(t - \tau) = \frac{1}{2\pi} \int S_y(\omega) e^{i\omega t} d\omega = \frac{a}{2\pi} \int S_x(\omega) e^{i\omega(t-\tau)} d\omega = \\ &= \frac{a}{2\pi} \int S_x(\omega) e^{i(\omega t - \varphi)} d\omega = \frac{1}{2\pi} \int S_x(\omega) K(\omega) e^{i\omega t} d\omega, \end{aligned}$$

т. е. характеристика системы должна быть

$$K(\omega) = a e^{-i\varphi},$$

где фазо-частотная характеристика

$$\varphi = \omega\tau$$

— линейная функция частоты.

или чтобы спектр отклика заданным образом отличался от спектра воздействия. Это означает, что мы произвольным образом задаем временную или частотную характеристику системы и задача состоит в том, чтобы сконструировать систему с заданными свойствами. Не любая задача такого рода разрешима: не должны нарушаться условия физической осуществимости, которые мы пока не формулируем. Осуществимая же система может быть построена с любым приближением к поставленным требованиям; чем лучше приближение, тем, как правило, сложнее система.

§ 31. Одиночный контур; частотные характеристики

Одиночный колебательный контур представляет собой цепь, составленную из индуктивности L , емкости C и сопротивления R .

Для вынужденных колебаний в такой цепи характерны явления, возникающие при частоте вынуждающей колебания эдс, точно равной или близкой к значению

$$\omega_0 = \frac{1}{\sqrt{LC}}.$$

Эти явления состоят в том, что сопротивление цепи, вообще комплексное, становится чисто активным или, иначе говоря, что фазовый сдвиг между током и напряжением обращается в нуль; в том, что амплитуда тока в цепи достигает максимального значения; в том, что напряжения на элементах цепи достигают максимума. Все эти явления называются *резонансными* или просто *резонансом*; частота ω_0 — резонансной частотой.

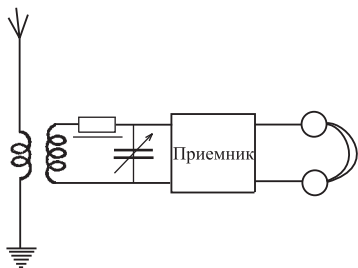


Рис. 61

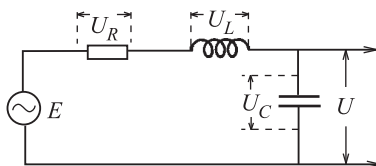


Рис. 62

Одно из основных применений колебательного контура состоит в том, что он включается в состав приемника в качестве избирательной системы. Контур может быть включен, например, как показано на рис. 61, в качестве промежуточного звена между антенной и всей остальной схемой приемника. Настройка контура на несущую частоту принимаемой станции производится конденсатором переменной емко-

сти. Но через контур должен пройти весь спектр сигнала. Поэтому нас интересует частотная характеристика контура.

Выделим контур из схемы рис. 61 и представим его в виде схемы рис. 62. Уравнение напряжений для цепи рис. 62 таково:

$$u_L + u_R + u_C = L \frac{di}{dt} + Ri + \frac{1}{C} \int idt = e, \quad (79)$$

где e — внешняя электродвижущая сила.

Вводя

$$u = u_C = \frac{1}{C} \int idt,$$

получим

$$LC \frac{d^2 u}{dt^2} + RC \frac{du}{dt} + u = e. \quad (80)$$

Обозначая сокращенно:

$$\omega_0^2 = \frac{1}{LC}, \quad 2\alpha = \frac{R}{L},$$

перепишем ур-ние (80) в виде

$$\frac{d^2 u}{dt^2} + 2\alpha \frac{du}{dt} + \omega_0^2 u = \omega_0^2 e. \quad (81)$$

Пусть электродвижущая сила e синусоидальна, т. е.

$$e = E e^{i\omega t}.$$

Тогда можно переписать ур-ние (81) в комплексной форме

$$(\omega_0^2 - \omega^2 + i2\alpha\omega) \dot{U} = \omega_0^2 \dot{E}, \quad (82)$$

где точками обозначены комплексные амплитуды. Из (82) находим коэффициент передачи

$$K = \frac{\dot{U}}{\dot{E}} = \frac{1}{1 - \frac{\omega^2}{\omega_0^2} + i d \frac{\omega}{\omega_0}},$$

где через d обозначена величина

$$d = \frac{2\alpha}{\omega_0} = \frac{R}{\sqrt{L/C}},$$

называемая *затуханием* контура. Величина

$$\rho = \sqrt{\frac{L}{C}}$$

носит название *характеристического сопротивления*. Величина, обратная затуханию, т. е.

$$Q = \frac{1}{d} = \frac{\rho}{R},$$

называется *добротностью* контура. Модуль коэффициента передачи

$$A = |K| = \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2 + \frac{\omega^2}{\omega_0^2}d^2}} \quad (83)$$

выражает отношение амплитуды выходного напряжения U к амплитуде эдс E . Формула (83) представляет амплитудно-частотную характеристику одиночного контура. График зависимости (83) называется *резонансной кривой*. При резонансе, т. е. при $\omega = \omega_0$,

$$A = \frac{1}{d} = Q.$$

Таким образом, добротность Q непосредственно показывает, во сколько раз напряжение при резонансе превосходит эдс. В радиотехнических контурах Q достигает $100 \div 200$ и даже больше. Этими соотношениями объясняется роль контура в радиотехнических схемах: контур не только выделяет нужную частоту, но и повышает напряжение.

Модуль коэффициента передачи достигает максимума при частоте

$$\omega_1 = \omega_0 \sqrt{1 - \frac{1}{2}d^2}.$$

Для контуров с малым затуханием, какие обычно и применяются, различие между ω_1 и ω_0 незначительно и может не приниматься во внимание.

Введя для сокращения обозначение

$$\varepsilon = 1 - \frac{\omega^2}{\omega_0^2},$$

можем переписать (83) в виде

$$A = \frac{1}{\sqrt{\varepsilon^2 + \frac{\omega^2}{\omega_0^2}d^2}}.$$

Обычно нас интересуют соотношения в области частот, близких к резонансной. Полагая $\omega/\omega_0 \approx 1$, получим весьма простое и очень употребительное приближенное выражение для частотной характеристики

$$A \approx \frac{1}{\sqrt{\varepsilon^2 + d^2}}. \quad (84)$$

На частотах, близких к резонансной, для величины ε можно записать следующие соотношения:

$$\varepsilon = 1 - \frac{\omega^2}{\omega_0^2} = \frac{\omega_0^2 - \omega^2}{\omega_0^2} = \frac{(\omega_0 - \omega)(\omega_0 + \omega)}{\omega_0^2} \approx \frac{2(\omega_0 - \omega)}{\omega_0} = 2 \frac{\Delta\omega}{\omega_0}.$$

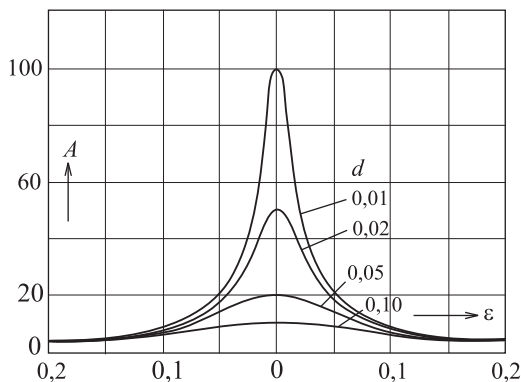


Рис. 63

Если назвать отклонение частоты от резонансной, т. е. $\Delta\omega = \omega_0 - \omega$, — *расстройкой*, а $\Delta\omega/\omega_0$ — относительной расстройкой, то величина ε для $\omega \approx \omega_0$ приблизительно пропорциональна относительной расстройке (на этом основании величину ε обычно и называют просто относительной расстройкой).

Семейство резонансных кривых по формуле (84) для разных значений затухания d построено на рис. 63. Чем меньше d , т. е. чем больше Q , тем острее резонансная кривая, тем выше ее максимум, тем круче ее скаты. Следует помнить, что для больших расстроек формула (84) не годится; в этом случае нужно пользоваться более общей формулой (83).

Найдем сопротивление контура. Переписывая (79) в комплексной форме, получим

$$Z = \frac{\dot{E}}{\dot{I}} = R + i \left(\omega L - \frac{1}{\omega C} \right).$$

Определим фазовый сдвиг между током и эдс. Записав

$$Z = |Z| e^{i\varphi},$$

найдем

$$\operatorname{tg} \varphi = \frac{\omega L - \frac{1}{\omega C}}{R} = Q \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right) \approx -\varepsilon Q.$$

При резонансе

$$\operatorname{tg} \varphi = \varphi = 0$$

и Z чисто активно и имеет наименьшее значение

$$Z_{\min} = R.$$

Эти соотношения характерны для резонанса вообще, как будет видно в дальнейшем. При резонансе реактивные сопротивления компенсируют друг друга $\left(\omega L - \frac{1}{\omega C} = 0 \right)$ и в цепи действует только

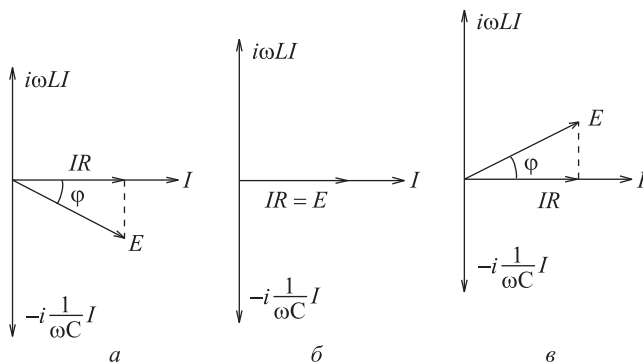


Рис. 64

активное сопротивление. В пояснение приведем векторную диаграмму напряжений для схемы рис. 62 при частоте, ниже резонансной (рис. 64, а), при резонансной частоте (рис. 64, б) и при частоте, выше резонансной (рис. 64, в). Так как падения напряжения на элементах схемы пропорциональны одному и тому же току, то точно такими же векторными диаграммами можно представить и соотношения между сопротивлениями. Например, диаграмме *напряжений* рис. 64, а будет соответствовать диаграмма *сопротивлений* рис. 65. Через X_L и X_C обозначены реактивные сопротивления — индуктивное и емкостное, т. е. $X_L = \omega L$, $X_C = -\frac{1}{\omega C}$. При резонансе векторы iX_L и iX_C равны по длине, но направлены в противоположные стороны, так что сумма их равна нулю. Вектор Z равен геометрической сумме векторов R , iX_L и $-iX_C$.

Рассмотрим теперь частотную характеристику контура с точки зрения требований, предъявляемых избирательной системе. Идеальная частотная характеристика должна была бы иметь прямоугольную форму, т. е. плоскую вершину и отвесные скаты. Резонансная кривая контура, конечно, мало похожа на идеальную характеристику. Она имеет пологие скаты, и нужно как-то условиться о том, что называть полосой пропускания контура. Принято определять полосу пропускания так: проводят горизонтальную прямую на высоте

$$A_0 = \frac{A_{\max}}{\sqrt{2}} = 0,707 A_{\max}$$

и считают полосу пропускания $2\Delta_0\omega$ равной расстоянию (в масштабе оси частот) между точками пересечения этой прямой с резонансной кривой (рис. 66). Полосу пропускания легко найти аналитически, пользуясь уравнением резонансной кривой (84). Составим уравнение

$$A_0 = \frac{1}{\sqrt{\varepsilon_0^2 + d^2}} = \frac{1}{\sqrt{2}} A_{\max} = \frac{1}{\sqrt{2} d}.$$

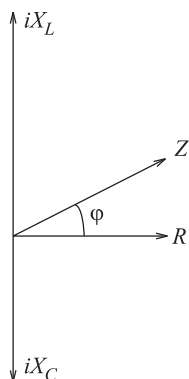


Рис. 65

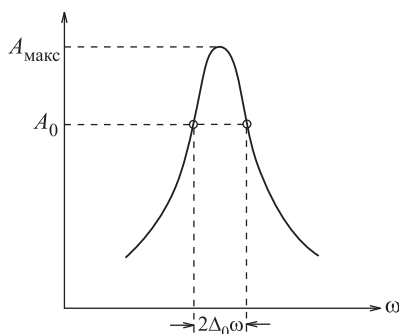


Рис. 66

Возводя в квадрат, получим

$$\varepsilon_0 = d.$$

Здесь

$$\varepsilon_0 \approx \frac{2\Delta_0\omega}{\omega_0}$$

выражает относительную ширину полосы пропускания. Абсолютная ширина полосы пропускания равна

$$2\Delta_0\omega = \omega_0 d \quad (= 2\alpha).$$

Как видим, ширина полосы пропускания пропорциональна затуханию. Для получения заданной полосы (вмещающей весь спектр сигнала) нужно ввести сравнительно большое затухание. С другой стороны, при увеличении затухания резонансная кривая расширяется, скаты ее становятся более пологими, так что избирательность ухудшается. Примирить эти противоречия в одиночном контуре нельзя. Путь улучшения свойств избирательной системы указан в дальнейшем.

Из-за криволинейной формы резонансной кривой контур искажает сигнал. Полное представление о происходящем искажении можно составить на основании формулы (55)

$$S_U(\omega) = K(\omega) S_E(\omega).$$

Беря модули входящих сюда величин, получим аналогичную формулу для амплитудных спектров

$$\Phi_U(\omega) = A(\omega) \Phi_E(\omega).$$

Чтобы найти спектр выходного напряжения Φ_U , нужно помножить спектр эдс Φ_E на частотную характеристику $A(\omega)$. Пусть эдс представляет АМ-сигнал и пусть, для наглядности, ограниченный спектр низкочастотного сигнала сплошной и равномерный. Таким же будет и

спектр в каждой из боковых полос (рис. 67, а). После умножения на ординаты резонансной кривой (рис. 67, б) получается спектр выходного напряжения (рис. 67, в). Рисунок показывает, что в спектре Φ_U изменено соотношение спектральных плотностей по сравнению с исходным спектром Φ_E , а именно: крайние частоты модуляционного спектра (соответствующие высшим частотам в спектре низкочастотного сигнала) оказываются подавленными, а средние частоты преувеличенными. Такое искажение спектра отчетливо воспринимается на слух при приеме музыки и речи: звук становится глухим из-за недостатка высоких частот. При некоторой расстройке, т. е. при настройке контура неточно на несущую частоту, как показано на рис. 68, получается несимметричный спектр и возникают значительные искажения; однако на слух результат может оказаться несколько лучше, чем при точной настройке на несущую; по крайней мере речь получается более разборчивой.

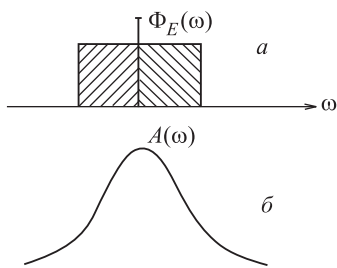


Рис. 67

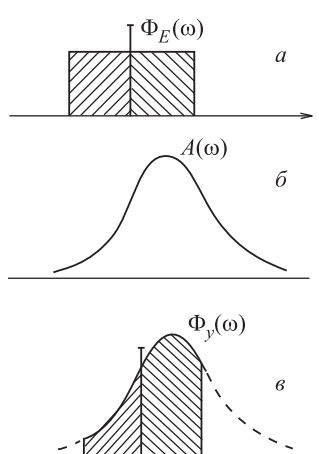


Рис. 68

Мы занимались до сих пор только последовательным контуром. Перейдем к параллельному. Прежде всего рассмотрим схему рис. 69. Уравнение *токов* имеет вид

$$i_C + i_R + i_L = C \frac{du}{dt} + \frac{1}{R}u + \frac{1}{L} \int u dt = i. \quad (85)$$

Заметим, что это уравнение в точности совпадает по строению с уравнением (79) *напряжений* в последовательном контуре. Различие состоит в том, что токи и напряжения поменялись ролями. Из полного сходства уравнений (79) и (85) мы с полным основанием можем заключить о сходстве решений этих уравнений, а также любых характеристик. Можно воспользоваться всеми ранее полученными формулами, в которых нужно лишь заменить величины, основываясь

на сопоставлении уравнений (79) и (85). Легко видеть следующие соответствия:

последовательный контур $uRLCZ$,
 параллельный контур $i\frac{1}{R}CLY \left(= \frac{1}{Z} \right)$.

Таким образом, например, при резонансе в последовательном контуре сопротивление имеет минимум, равный $Z_{\min} = R$, а в параллельном контуре проводимость имеет минимум, равный $Y_{\min} = 1/R$ (сопротивление при этом имеет максимум). Добротность параллельного контура по схеме рис. 69 будет тем больше (а затухание тем меньше), чем меньше проводимость $G = 1/R$, т. е. чем *больше* сопротивление R .

Пара схем, в которых токи и напряжения меняются местами и, следовательно, сопротивление одной схемы ведет себя, как проводимость другой, называются взаимно-обратными. Последовательный и параллельный контуры (рис. 62 и 69) образуют такую пару взаимно-обратных схем.

В вышеописанном сопоставлении важно отметить различие в характере резонанса и в условиях питания. В схеме рис. 62 контур питается *заданным напряжением*, равным эдс источника E (внутреннее сопротивление источника предполагается равным нулю). Резонанс в этой схеме — это резонанс напряжений. Он проявляется в том, что при резонансе напряжение u_C (или равное ему по абсолютной величине u_L) сильно возрастает (в Q раз). В схеме же рис. 69 контур питается *заданным током*. Получаемый резонанс есть резонанс токов; он выражается в том, что токи i_L и i_C возрастают: при резонансе они равны друг другу по абсолютной величине и противоположны по фазе, так что $i = i_R$.

Если бы мы питали параллельный контур рис. 69 заданным напряжением, то напряжения на всех элементах были бы всегда равны между собой и равны приложенному напряжению. Резонанс в таком режиме проявился бы лишь в том, что общий ток имел бы при резонансе минимум.

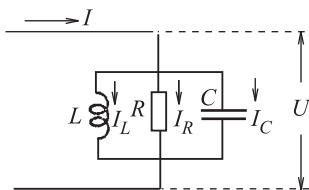


Рис. 69

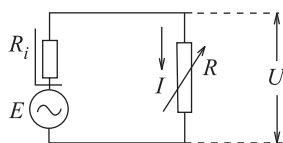


Рис. 70

Режимы заданного напряжения и заданного тока — это не что иное, как режимы *холостого хода* и *короткого замыкания* источника. Рассмотрим схему рис. 70, на которой изображен генератор с эдс e

и внутренним сопротивлением R_i , нагруженный переменным сопротивлением R . Напряжение на нагрузке

$$u = e \frac{R}{R_i + R}.$$

Ток в нагрузке

$$i = \frac{e}{R_i + R}.$$

Эти две формулы являются общими. Но можно рассматривать два предельных режима, а именно:

$$R_i \gg R \text{ и } R_i \ll R.$$

Первый режим называется режимом короткого замыкания (кз), второй — режимом холостого хода (хх). В режиме кз ток равен

$$i \approx \frac{e}{R_i} = i_k.$$

Это — *ток короткого замыкания*. Он является заданным в том смысле, что полностью определяется параметрами самого генератора и не зависит от изменения нагрузки (если всегда соблюдено условие $R_i \gg R$). Напряжение же на сопротивлении нагрузки пропорционально его величине

$$u = e \frac{R}{R_i} = i_k R.$$

В режиме холостого хода имеем

$$u = e.$$

Это — *напряжение холостого хода*. Оно является заданным в том смысле, что не зависит от сопротивления нагрузки (если выполнено условие $R_i \ll R$). Ток же в цепи полностью определяется сопротивлением нагрузки

$$i = \frac{e}{R}.$$

Возвращаясь после этих вспомогательных замечаний к нашим контурам, можно сделать следующее заключение: для использования резонансных явлений в контурах следует включать последовательный контур в качестве нагрузки для источника с малым внутренним сопротивлением (стремясь осуществить режим хх, т. е. режим заданного напряжения), а параллельный контур следует питать от источника с большим внутренним сопротивлением (чтобы приблизиться к режиму кз, т. е. к режиму заданного тока). Напряжение, снимаемое при таких условиях с параллельного контура, пропорционально сопротивлению контура. Следовательно, резонансная кривая параллельного контура выражается непосредственно частотной зависимостью его сопротивления.

Режим, близкий к режиму заданного тока, получается при включении параллельного контура в анодную цепь лампы с большим внут-

ренным сопротивлением, например пентода. Но что получится, если контур включить в анодную цепь триода с внутренним сопротивлением, сравнимым с сопротивлением контура при резонансе? Нам придется рассмотреть общий случай питания и выяснить роль внутреннего сопротивления.

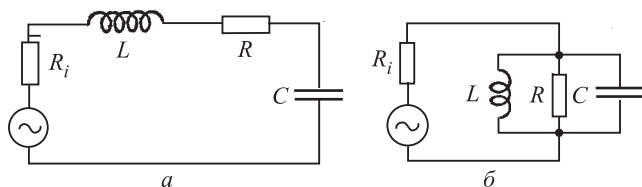


Рис. 71

На рис. 71, *a* и *б* последовательный и параллельный контуры питаются генератором с внутренним сопротивлением R_i .

Для последовательного контура затухание равно:

$$d_1 = \frac{R_i + R}{\rho} = d \left(1 + \frac{R_i}{R} \right), \quad (86)$$

где $d = R/\rho$ — затухание контура как такового. Выражение в скобках дает поправку на затухание, обусловленную влиянием внутреннего сопротивления. Как видим, $d_1 \approx d$ при $R_i \ll R$.

Для параллельного контура внутреннее сопротивление генератора оказывается подключенным к контуру параллельно. Поэтому проводимость $G_i = 1/R_i$ добавляется к проводимости $G = 1/R$, и мы получаем для затухания параллельного контура

$$d_{\text{пар}} = d_{\text{пар}} \left(1 + \frac{G_i}{G} \right) = d_{\text{пар}} \left(1 + \frac{R}{R_i} \right). \quad (87)$$

Таким образом, затухание возрастает с увеличением отношения R/R_i .

Общий вывод всех этих рассуждений состоит в том, что отклонение от предельных режимов (режимов кз и хх) влечет за собой увеличение действующего затухания за счет влияния внутреннего сопротивления источника. Учет этого влияния дают формулы (86) и (87).

Теперь нужно пересмотреть схему параллельного контура. Дело в том, что в последовательном, контуре все три элемента R , L и C включены последовательно, и, составляя схему параллельного контура, мы включили все три элемента параллельно. Это было очень удобно для описания общих свойств параллельного контура, так как дело свелось к совершенно аналогичным уравнениям и оба контура оказались взаимно-обратными системами. Но схема

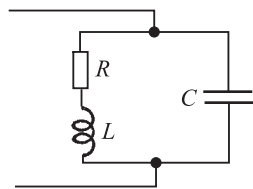


Рис. 72

рис. 69 не соответствует действительности в том смысле, что проводимость изоляции конденсатора обычно очень мала и ею пренебрегают, в то время как активным сопротивлением катушки пренебречь нельзя. Поэтому лучшее приближение к действительности мы получим, заменив ранее рассмотренную схему другой, показанной на рис. 72. Именно в таком виде схема параллельного контура обычно и рассматривается и мы выведем сейчас относящиеся к ней зависимости¹⁾. Найдем прежде всего сопротивление

$$Z = \frac{1}{i\omega C + \frac{1}{R + i\omega L}} = \frac{R + i\omega L}{1 - \omega^2 LC + i\omega RC}.$$

Модуль сопротивления равен

$$|Z| = \sqrt{\frac{R^2 + \omega^2 L^2}{(1 - \omega^2 LC)^2 + \omega^2 R^2 C^2}}$$

или, вводя обычные обозначения,

$$|Z| = \rho \frac{\sqrt{d^2 + \frac{\omega^2}{\omega_0^2}}}{\sqrt{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2 + d^2 \frac{\omega^2}{\omega_0^2}}}.$$

Если нас интересуют только соотношения вблизи резонансной частоты, то отношение ω/ω_0 имеет порядок единицы и можно пренебречь малым слагаемым d^2 (имеющим обычно порядок 10^{-4}) в числителе. Тогда

$$|Z| \approx \rho \frac{\omega}{\omega_0} \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2 + d^2 \frac{\omega^2}{\omega_0^2}}}$$

или, пользуясь обычным приближением (см. (84)):

$$|Z| \approx \rho \frac{1}{\sqrt{\varepsilon^2 + d^2}}. \quad (88)$$

При резонансе, т. е. при $\varepsilon = 0$, сопротивление контура имеет максимум, равный

$$Z_{\text{макс}} \approx \frac{\rho}{d} = \frac{\rho^2}{R}. \quad (89)$$

¹⁾ Разумеется, всегда возможно формально заменить путем соответствующего пересчета последовательное сопротивление эквивалентным параллельным и наоборот. Но это крайне неудобно, так как эквивалентные сопротивления оказываются зависящими от частоты.

Напомним, что

$$\rho = \sqrt{\frac{L}{C}} = \omega_0 L = \frac{1}{\omega_0 C}.$$

Составим теперь полную схему, включая генератор (рис. 73). Выражая (как и раньше для последовательного контура) коэффициент передачи отношением выходного напряжения к эдс, найдем

$$K = \frac{\dot{U}}{\dot{E}} = \frac{Z}{Z + R_i}.$$

В режиме кз, который выгодно осуществить:

$$K \approx \frac{Z}{R_i}.$$

Модуль коэффициента передачи

$$A \approx \frac{|Z|}{R_i}$$

и, подставляя значение $|Z|$ из (88), находим окончательно

$$A \approx \frac{\rho}{R_i} \frac{1}{\sqrt{\varepsilon^2 + d^2}}.$$

Эта формула отличается от формулы (84) для последовательного контура только постоянным множителем ρ/R_i . Таким образом, все, что говорилось раньше о форме резонансных кривых последовательного контура, остается в силе и для параллельных контуров по схемам рис. 72 и 73. Что касается влияния на затухание внутреннего сопротивления генератора (при несоблюдении режима кз), то его легко определить из следующих соображений. При резонансе проводимость контура активна и равна R/ρ^2 (см. (89)). К этой проводимости добавляется проводимость $1/R_i$ цепи генератора. Поступая так же, как при выводе формулы (87), найдем

$$d_{\text{нап}} = d_{\text{нап}} \left(1 + \frac{\rho^2}{RR_i} \right). \quad (90)$$

Во избежание недоразумений напомним, что из трех вышеприведенных формул для действующего затухания (с поправкой на R_i) формула (86) относится к схеме рис. 71, а, формула (87) — к схеме рис. 71, б, а формула (90) — к схеме рис. 73.

Мы рассматривали параллельный контур, в котором индуктивность находится в одной ветви, а емкость — в другой. В радиотехнике применяются иногда *сложные* параллельные контуры, в которых индуктивность или емкость (а в наиболее общем случае и то и другое) содержатся в обеих ветвях контура. Пример схемы такого рода дан на рис. 74. Смысл применения сложных контуров состоит в том, что изменяя соотношение между реактивными сопротивлениями в обеих

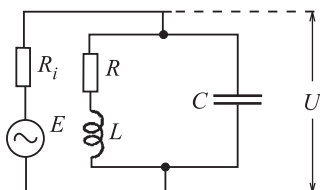


Рис. 73

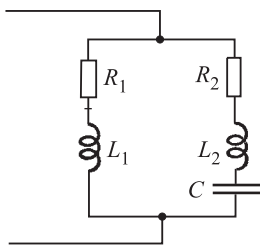


Рис. 74

ветвях, можно изменять сопротивление контура, не меняя его настройки. Опуская сравнительно громоздкие и не имеющие никакого принципиального интереса выкладки, в ходе которых делается ряд упрощений, приведем лишь окончательный результат. Он состоит в том, что резонансная частота приближенно (приближение тем лучше, чем выше добротность) равна

$$\omega_0 = \frac{1}{\sqrt{(L_1 + L_2)C}}.$$

Сопротивление же при резонансе равно

$$Z_{\text{макс}} = n^2 \frac{\rho^2}{R_1 + R_2},$$

где коэффициент n выражает распределение индуктивностей между ветвями, а именно:

$$n = \frac{L_1}{L_1 + L_2}.$$

Таким образом, входное сопротивление контура зависит от n , а резонансная частота не зависит. Нужно еще добавить, что в отличие от простого контура сопротивление сложного контура имеет не только максимум на резонансной частоте ω_0 , но еще и минимум на частоте, при которой происходит последовательный резонанс в ветви, содержащей L и C . Для схемы рис. 74 это произойдет на частоте

$$\omega_1 = \frac{1}{\sqrt{L_2 C}} > \omega_0.$$

§ 32. Одиночный контур; временные характеристики

Найдем переходную характеристику одиночного контура. Переходная характеристика представляет собой (§ 28) отклик системы на воздействие в форме единичной функции. Такого рода условия выполняются в схеме рис. 75, в которой E — единичная эдс. Замыкание ключа производится при $t = 0$. Изменение выходного напряжения во времени и представляет переходную функцию.

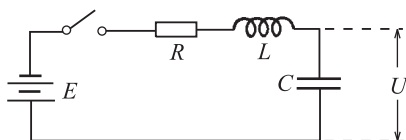


Рис. 75

Уравнение напряжений для схемы рис. 75 имеет вид (см. (80))

$$LC \frac{d^2 u}{dt^2} + RC \frac{du}{dt} + u = e.$$

Деля на LC и переходя к обычным обозначениям, получим

$$\frac{d^2 u}{dt^2} + 2\alpha \frac{du}{dt} + \omega_0^2 u = \omega_0^2 e. \quad (91)$$

Перепишем уравнение (91) в операционной форме

$$(p^2 + 2\alpha p + \omega_0^2) \bar{u} = \omega_0^2 \bar{e},$$

откуда

$$K(p) = \frac{\bar{u}}{\bar{e}} = \frac{\omega_0^2}{p^2 + 2\alpha p + \omega_0^2}. \quad (92)$$

Переходную функцию найдем как оригинал для этого изображения (учитывая, что в нашем случае $e(t) = \sigma(t)$, а $u(t) = h(t)$). Применим формулу (71). Для этого составим сначала характеристическое уравнение

$$H_2(p) = p^2 + 2\alpha p + \omega_0^2 = 0.$$

Корни этого уравнения

$$p_1 = -\alpha + i\omega_1, \quad p_2 = -\alpha - i\omega_1,$$

где

$$\omega_1 = \sqrt{\omega_0^2 - \alpha^2} = \omega_0 \sqrt{4 - \frac{1}{4}d^2}$$

— так называемая *собственная частота*. Заметим, что она всегда меньше резонансной частоты, хотя при малом затухании отличается от резонансной частоты незначительно.

У нас

$$H_1(p) = \omega_0^2, \quad H_2'(p) = 2(p + \alpha),$$

$$H_2'(p_1) = 2i\omega_1, \quad H_2'(p_2) = -2i\omega_1, \quad H_2(0) = \omega_0^2.$$

Подставляя все это в формулу Хевисайда, получаем

$$h(t) = 1 + \frac{\omega_0^2 e^{(-\alpha + i\omega_1)t}}{2(-\alpha + i\omega_1)i\omega_1} - \frac{\omega_0^2 e^{(-\alpha - i\omega_1)t}}{2(-\alpha - i\omega_1)i\omega_1}.$$

Вынося общие множители, приводя к общему знаменателю и пользуясь формулами Эйлера, имеем окончательно

$$h(t) = 1 - e^{-\alpha t} \left(\frac{\alpha}{\omega_1} \sin \omega_1 t + \cos \omega_1 t \right) \quad (t > 0). \quad (93)$$

При малом затухании коэффициент $\alpha/\omega_1 \ll 1$ первым членом в скобках можно пренебречь, и тогда

$$h(t) = 1 - e^{-\alpha t} \cos \omega_1 t \quad (t > 0).$$

График этой функции показан на рис. 76. Пунктиром намечен график единичной функции. Рис. 76 ясно показывает, насколько в рассматриваемом случае отклик отличается от воздействия. Различие обусловлено затухающим колебанием, возникающим в цепи в момент включения. Это есть так называемое *собственное* или *свободное* колебание. Оно происходит с собственной частотой ω_1 , амплитуда его затухает по экспоненциальному закону $e^{-\alpha t}$. Величина α называется *показателем затухания*. Заметим, что при увеличении α собственная частота убывает. При $\alpha = \omega_0$ собственная частота обращается в нуль, а при $\alpha > \omega_0$ (т. е. при $d = 2\alpha/\omega_0 > 2$) собственная частота становится мнимой. Это означает, что переходная функция при таких соотношениях теряет колебательный характер; процесс, происходящий в контуре, делается *апериодическим*. Оба корня характеристического уравнения при этом вещественны. Впрочем, мы рассматриваем обычно контуры с малым затуханием, и апериодический случай нам не встретится.

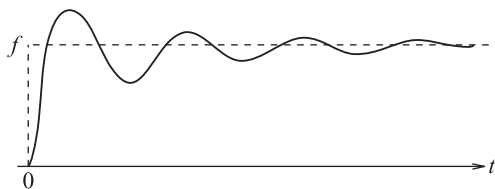


Рис. 76

Импульсную реакцию контура можно найти по формуле (72), но в данном случае проще продифференцировать (93) (так как $g(t) = dh/dt$). Это даст

$$g(t) = \frac{\omega_0^2}{\omega_1} e^{-\alpha t} \sin \omega_1 t$$

или, приближенно (при малом затухании),

$$g(t) = \omega_0 e^{-\alpha t} \sin \omega_0 t.$$

Зная переходную функцию, легко найти отклик на прямоугольный импульс. Аналитическое выражение для прямоугольного импульса длительностью τ можно записать в виде

$$x(t) = \sigma(t) - \sigma(t - \tau).$$

Пользуясь принципом наложения, получим

$$y(t) = h(t) - h(t - \tau).$$

Графики $x(t)$ и $y(t)$ изображены на рис. 77.

Зная переходную функцию, можно найти отклик контура на любое воздействие, пользуясь интегралом Дюамеля (§ 27).

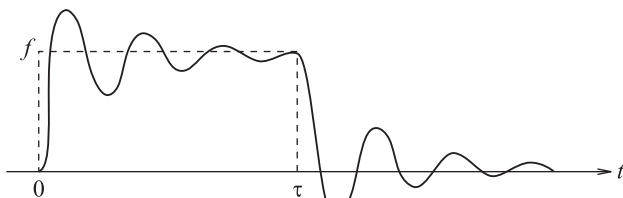


Рис. 77

Наибольший интерес для задач радиотехники представляет отклик контура на внезапно включаемое синусоидальное напряжение, частота которого в общем случае не совпадает с резонансной или собственной частотами контура. Эту задачу можно решить с помощью интеграла Фурье, интеграла Дюамеля или же операционным методом. Но мы воспользуемся классическим методом, подкрепляя его физическими рассуждениями.

Уравнение, которое предстоит решить, запишем в виде

$$\frac{d^2 u}{dt^2} + 2\alpha \frac{du}{dt} + \omega_0^2 u = \omega_0^2 E_m \sin \omega t \quad (t > 0).$$

Полное решение уравнения с правой частью состоит из общего и частного интегралов.

Общий интеграл представляет собой решение однородного уравнения (т. е. уравнения, в правой части которого стоит нуль). По физическому смыслу это решение есть не что иное, как свободное колебание.

Свободное колебание совершается с частотой ω_1 и затухает по закону $e^{-\alpha t}$. Поэтому можно сразу записать для общего интеграла

$$u_1 = A e^{-\alpha t} \sin(\omega_1 t + \psi),$$

где A и ψ — неизвестные пока начальные амплитуда и фаза свободного колебания.

Частный интеграл представляет собой любое решение уравнения с правой частью. В нашем случае частное решение по физическому смыслу представляет собой установившееся вынужденное колебание

$$u_2 = B \sin(\omega t + \varphi).$$

Полное решение уравнения запишется в виде

$$u(t) = u_1(t) + u_2(t) = A e^{-\alpha t} \sin(\omega_1 t + \psi) + B \sin(\omega t + \varphi). \quad (94)$$

Мы предполагаем, что система до включения находилась в покое. Это значит, что в момент включения, т. е. при $t = 0$, напряжение $u(t)$ и его первая производная du/dt равны нулю (так называемые нулевые начальные условия). Дифференцируя (94), получаем

$$\frac{du}{dt} = A e^{-\alpha t} [-\alpha \sin(\omega_1 t + \psi) + \omega_1 \cos(\omega_1 t + \psi)] + B \omega \cos(\omega t + \varphi). \quad (95)$$

Подставляя в (94) и (95) $t = 0$ и приравнивая (на основании начальных условий) нулю, получаем:

$$\begin{aligned} A \sin \psi + B \sin \varphi &= 0, \\ A \left(-\frac{\alpha}{\omega} \sin \psi + \frac{\omega_1}{\omega} \cos \psi \right) + B \cos \varphi &= 0. \end{aligned} \quad (96)$$

Рассмотрим случай малой расстройки ($\omega_0/\omega \approx 1$) и малого затухания ($\alpha/\omega_1 \ll 1$). Первым членом во второй строке (96) при этом можно пренебречь, и мы получаем пару уравнений:

$$\begin{aligned} A \sin \psi + B \sin \varphi &= 0, \\ A \cos \psi + B \cos \varphi &= 0, \end{aligned}$$

которые удовлетворяются при $A = -B$, $\psi = \varphi$. Итак (см. (94)),

$$u(t) = B [\sin(\omega t + \varphi) - e^{-\alpha t} \sin(\omega_1 t + \varphi)].$$

Но B — амплитуда установившегося вынужденного колебания. Эта величина уже известна (см. § 31):

$$B = \frac{E_m}{\sqrt{\varepsilon^2 + d^2}}.$$

Итак, окончательно

$$u(t) \approx \frac{E_m}{\sqrt{\varepsilon^2 + d^2}} [\sin(\omega t + \varphi) - e^{-\alpha t} \sin(\omega_1 t + \varphi)], \quad (97)$$

где φ — начальная фаза, которая в нашем приближении оказалась одинаковой как для вынужденного, так и для свободного колебаний. Из формулы (97) ясно, что напряжение $u(t)$ представляет собой сумму двух колебаний с разными частотами. При таких условиях должны получаться биения с частотой, равной $\Delta\omega = |\omega - \omega_1|$. Но одно из двух колебаний, а именно свободное колебание, с течением времени затухает, поэтому биения ослабевают. По прошествии достаточного времени можно пренебречь вторым членом в (97) и перейти, таким образом, к установившемуся режиму, когда имеется только вынужденное колебание

$$u_2(t) = \frac{E_m}{\sqrt{\varepsilon^2 + d^2}} \sin(\omega t + \varphi).$$

Описанный процесс изображен на рис. 78.

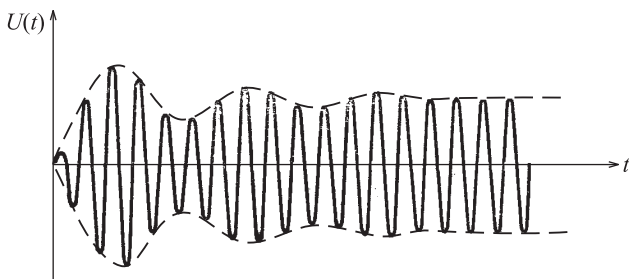


Рис. 78

Найдем закон изменения со временем амплитуды напряжения $u(t)$. Для этого нужно представить сумму двух колебаний в виде одного колебания с изменяющимися амплитудой и фазой. Вычисление облегчается векторной диаграммой рис. 79, на которой показано сложение двух векторов с длинами C_1 и C_2 , расположенных под углами α и β к полярной оси. Задача состоит в определении длины C результирующего вектора. Имеем

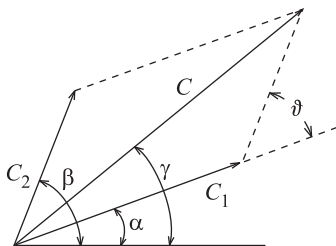


Рис. 79

$$C = \sqrt{C_1^2 + C_2^2 + 2C_1C_2 \cos \vartheta},$$

где $\vartheta = \beta - \alpha$. В нашем случае

$$C_1 \approx \frac{E_m}{\sqrt{\varepsilon^2 + d^2}}, \quad C_2 \approx -\frac{E_m}{\sqrt{\varepsilon^2 + d^2}} e^{-\alpha t}, \quad \vartheta = (\omega - \omega_1)t = \Omega t.$$

Таким образом,

$$C(t) \approx \frac{E_m}{\sqrt{\varepsilon^2 + d^2}} \sqrt{1 + e^{-2\alpha t} - 2e^{-\alpha t} \cos \Omega t}. \quad (98)$$

В пределе при $t \rightarrow \infty$ $C(t)$ стремится к $\frac{E_m}{\sqrt{\varepsilon^2 + d^2}}$. Формула (98) дает закон изменения амплитуды напряжения $u(t)$, т.е. то, что в радиотехнике называется *огibaющей*.

Рассмотрим еще тот важный частный случай, когда частота совпадает с собственной частотой контура. При этом расстройка равна нулю и частота биений $\Omega = |\omega - \omega_1|$ также равна нулю. Формула (98) дает для этого случая

$$C(t) = \frac{E_m}{d} \sqrt{1 + e^{-2\alpha t} - 2e^{-\alpha t}} = \frac{E_m}{d} (1 - e^{-\alpha t}),$$

т.е. при настройке контура на частоту питающего напряжения рост амплитуды происходит монотонно, без колебаний.

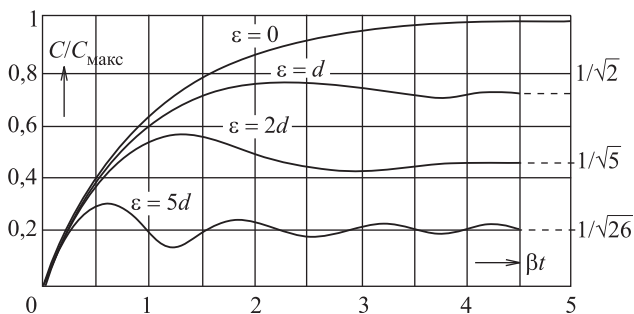


Рис. 80

Все описанные соотношения представлены графиком рис. 80, на котором огибающая $C(t)$ дана в функции времени для различных значений расстройки ε . (Заметим, что для рассматриваемого неустановившегося режима амплитуда колебания есть функция двух переменных — времени и расстройки.)

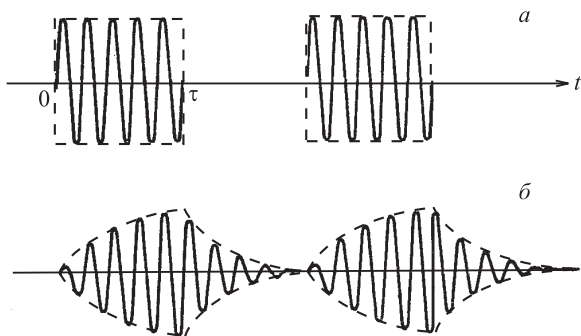


Рис. 81

Полученные соотношения имеют большое техническое значение. Разберем вопрос о приеме импульсных сигналов. Пусть приемник содержит контур, настроенный на несущую частоту. На рис. 81, а изображено воздействие в виде импульсов с прямоугольной огибающей, а на рис. 81, б — соответствующий отклик. Импульсы на выходе контура расплываются, заполняя паузы, и если повышать скорость передачи, то надежный прием сигналов станет невозможным. Для повышения скорости передачи без ущерба для надежности связи нужно изменить параметры контура так, чтобы увеличилась скорость нарастания амплитуды, т. е. наклон огибающей. Принимая установившееся значение амплитуды за единицу, можно записать

$$C(t) = \begin{cases} 1 - e^{-\alpha t} & (0 < t < \tau), \\ e^{-\alpha(t-\tau)} & (t > \tau). \end{cases}$$

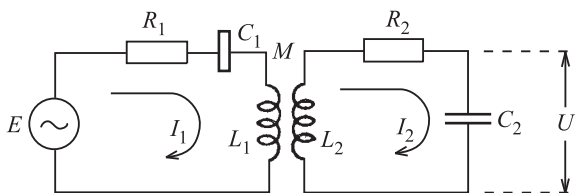


Рис. 83

В радиотехнике часто применяются контуры с *индуктивной* (трансформаторной) *связью*, которые мы и рассмотрим подробнее. Схема двух индуктивно связанных контуров изображена на рис. 83. Степень связи определяется коэффициентом взаимной индукции M . Вводят так называемый *коэффициент связи*, определяемый как

$$k = \frac{M}{\sqrt{L_1 L_2}},$$

причем

$$0 < k < 1.$$

По физическому смыслу коэффициент связи выражает отношение общего магнитного потока (т. е. потока, пронизывающего обе катушки) к полному потоку. Величина $1 - k$ выражает относительную величину потока рассеяния. В энергетических трансформаторах коэффициент связи k очень близок к единице; в радиотехнике, напротив, этот коэффициент обычно мал.

Контуры в общем случае могут быть разные, т. е. иметь различное затухание и настройку. Мы ограничимся исследованием случая двух одинаковых контуров, настроенных на одну и ту же частоту ω_0 , так как и в этом более простом случае все характерные свойства системы связанных контуров обнаружатся вполне отчетливо.

Составим уравнения напряжений для системы рис. 83:

$$\left. \begin{aligned} L_1 \frac{di_1}{dt} + R_1 i_1 + \frac{1}{C_1} \int i_1 dt &= e + M \frac{di_2}{dt} \\ L_2 \frac{di_2}{dt} + R_2 i_2 + \frac{1}{C_2} \int i_2 dt &= M \frac{di_1}{dt} \end{aligned} \right\}. \quad (99)$$

Эти уравнения отличаются от уравнений для одиночных контуров наличием так называемых членов связи $M \frac{di}{dt}$.

Перепишем уравнения (99) в комплексной форме, отбросив индексы у параметров (так как мы полагаем контуры одинаковыми):

$$\left. \begin{aligned} \left(i\omega\Delta + R + \frac{1}{i\omega C} \right) \dot{I}_1 - i\omega M \dot{I}_2 &= \dot{E} \\ \left(i\omega\Delta + R + \frac{1}{i\omega C} \right) \dot{I}_2 - i\omega M \dot{I}_1 &= 0 \end{aligned} \right\}, \quad (100)$$

или, короче,

$$\left. \begin{aligned} Z\dot{I}_1 - i\omega M\dot{I}_2 &= \dot{E} \\ Z\dot{I}_2 - i\omega M\dot{I}_1 &= 0 \end{aligned} \right\}. \quad (101)$$

Исключая из этих двух уравнений $\dot{I}_1$, получаем

$$\dot{I}_2 = \dot{E} \frac{i\omega M}{\omega^2 M^2 + Z^2}.$$

Введем

$$\dot{U} = \frac{1}{i\omega C} \dot{I}_2 = \dot{E} \frac{M}{C} \frac{1}{\omega^2 M^2 + Z^2}$$

и определим коэффициент передачи

$$K = \frac{\dot{U}}{\dot{E}} = \frac{M}{C} \frac{1}{\omega^2 M^2 + Z^2}.$$

Раскрывая выражение Z^2 , получим

$$K = \frac{M}{C} \frac{1}{\omega^2 M^2 + R^2 - \omega^2 L^2 \left(1 - \frac{\omega_0^2}{\omega^2}\right) + i \cdot 2R\omega L \left(1 - \frac{\omega_0^2}{\omega^2}\right)}. \quad (102)$$

Теперь сделаем некоторое упрощение, основанное на том, что нас интересуют соотношения только в узкой полосе частот около частоты ω_0 . Именно, положим

$$\omega L \approx \omega_0 L = \rho.$$

В этом приближении, вынося $\omega^2 L^2$, получаем из (102)

$$K = \frac{k}{k^2 + d^2 - \varepsilon^2 - i \cdot 2d\varepsilon}.$$

Здесь применено обозначение $1 - \frac{\omega_0^2}{\omega^2} \approx -\frac{2\Delta\omega}{\omega_0} = -\varepsilon$ (см. § 31).

Для модуля коэффициента передачи имеем

$$|K| = A = \frac{k}{\sqrt{(d^2 + k^2)^2 + 2(d^2 - k^2)\varepsilon^2 + \varepsilon^4}}. \quad (103)$$

Эта формула выражает амплитудно-частотную характеристику системы связанных контуров. Посмотрим, каков характер этой зависимости. Прежде всего найдем резонансную частоту, т.е. значение ε , при котором модуль коэффициента передачи имеет максимум. Для этого составим производную $\frac{dA}{d\varepsilon}$ и приравняем ее нулю. Получим уравнения, определяющие положение трех экстремальных точек:

$$\varepsilon = 0; \quad d^2 - k^2 + \varepsilon^2 = 0.$$

Если $k > d$, то экстремальная точка, определяемая первым уравнением, есть точка минимума. Второе уравнение определяет положение двух максимумов. Решая это уравнение, находим

$$\varepsilon = \pm \sqrt{k^2 - d^2}$$

или

$$\omega = \frac{\omega_0}{\sqrt{1 \pm \sqrt{k^2 - d^2}}}. \quad (104)$$

Итак, имеются две резонансные частоты, одна из которых ниже, а другая выше ω_0 :

$$\omega_1 = \frac{\omega_0}{\sqrt{1 + \sqrt{k^2 - d^2}}}, \quad \omega_2 = \frac{\omega_0}{\sqrt{1 - \sqrt{k^2 - d^2}}}.$$

Это является важнейшим свойством системы связанных контуров. Резонансные частоты зависят от коэффициента связи k и называются поэтому *частотами связи*. Разнос между частотами ω_1 и ω_2 тем больше, чем больше коэффициент связи k . При уменьшении k частоты связи сближаются, и при $k = d$

$$\omega_1 = \omega_2 = \omega_0.$$

Связь, соответствующая значению $k = d$, называется *критической*. При $k < d$ формула (104) теряет силу. Исследование показывает, что при $k < d$ имеется только один максимум на частоте ω_0 . В этом можно

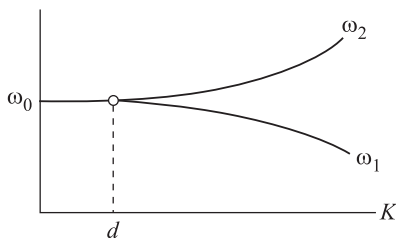


Рис. 84

убедиться, рассматривая формулу (103). Пусть $k > d$. Тогда коэффициент при ε^2 (ε имеет смысл расстройки) отрицателен. Это значит, что при увеличении ε (т. е. при отклонении ω от ω_0 в обе стороны) знаменатель в формуле (103) сначала уменьшается (пока не оказывает заметного влияния член ε^4). Следовательно, A возрастает и при $\varepsilon = 0$, т. е. при $\omega = \omega_0$, мы имеем минимум. При дальнейшем увели-

чении ε положительный член ε^4 начинает оказывать решающее влияние, и A , достигнув максимума, начинает убывать. Если же $k < d$, то все коэффициенты положительны и, следовательно, при любом значении ε модуль коэффициента передачи A меньше максимального значения

$$A_0 = \frac{k}{d^2 + k^2},$$

которое он имеет при $\varepsilon = 0$, т. е. при $\omega = \omega_0$.

Зависимость частот связи от коэффициента связи показана на графике рис. 84. При критической связи, т. е. при $k = d$, получаем из (103)

$$A = \frac{k}{\sqrt{4d^2 + \varepsilon^4}}. \quad (105)$$

Это — уравнение резонансной кривой с одним максимумом при $\varepsilon = 0$, т. е. при $\omega = \omega_0$. Интересно заметить, что и в этом случае, когда характеристика системы связанных контуров утрачивает присущую ей двугорбую форму, резонансная кривая существенно отличается от резонансной кривой одиночного контура. В этом можно убедиться, сравнив формулы (105) и (84). В последней под корнем стоит ε^2 , тогда как в первой — ε^4 . Поэтому кривая, выражаемая формулой (105), имеет более плоскую верхушку и более крутые скаты по сравнению с обычной резонансной кривой. Последняя намечена на рис. 85 пунктирной линией.

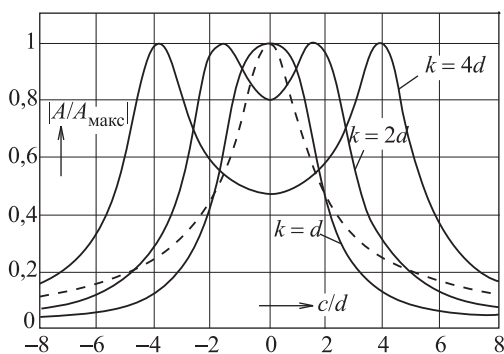


Рис. 85

Все описанные свойства частотной характеристики системы поясняются рис. 85, на котором изображено семейство кривых, построенных по формуле (103) для различных значений коэффициента связи k при фиксированном затухании d . Что касается случая произвольной настройки двух связанных контуров, то, не вдаваясь в исследование получаемых при этом соотношений, заметим лишь, что при расстройке одного из контуров относительно другого разнос частот связи возрастает.

Обсудим теперь свойства системы двух индуктивно-связанных контуров с точки зрения радиосвязи. Кривые рис. 85 показывают, что эта система в качестве избирательного звена обладает значительно более благоприятными свойствами, нежели одиночный контур. Связанные контуры по сравнению с одиночным контуром представляют уже некоторое приближение к идеальной прямоугольной характеристике. Характер этого приближения поясняется рис. 86, на котором даны для сравнения все три характеристики: a — одиночный контур, b — связанные контуры, $в$ — идеальная характеристика.

На рис. 86 все три кривые приведены к одинаковой полосе пропускания. В связи с этим возникает вопрос о том, как определить полосу пропускания для связанных контуров. Естественно воспользоваться

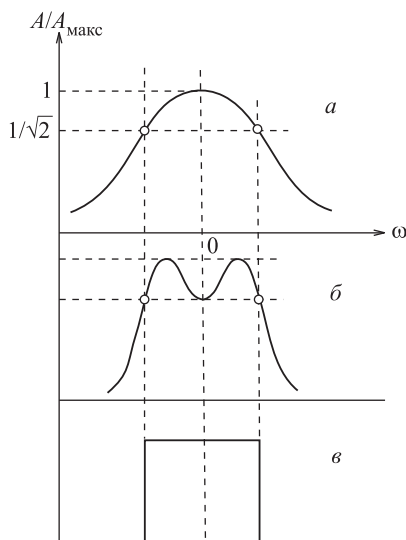


Рис. 86

тем же определением, которым мы пользовались в случае одиночного контура, а именно: проведем горизонтальную прямую на высоте $1/\sqrt{2} = 0,707$ от максимальной ординаты частотной характеристики и определим границы полосы пропускания абсциссами точек пересечения горизонтальной прямой с характеристикой. Короче говоря, полоса пропускания определяется как ширина частотной характеристики на уровне $1/\sqrt{2}$ от максимальной ординаты. Но в случае связанных контуров ширина характеристики зависит не только от затухания, но и от коэффициента связи. При увеличении k разнос частот связи возрастает и характеристика расширяется. При этом, однако, углубляется провал

на частоте ω_0 . Поэтому в качестве дополнительного условия, позволяющего найти однозначное решение задачи, можно потребовать, чтобы и минимум характеристики лежал на той же высоте $1/\sqrt{2}$ от максимальной ординаты (рис. 87). Итак, нужно прежде всего выписать формулы: 1) для ординаты A при произвольной расстройке; 2) для ординаты минимума A_0 ; 3) для максимальной ординаты $A_{\text{макс}}$.

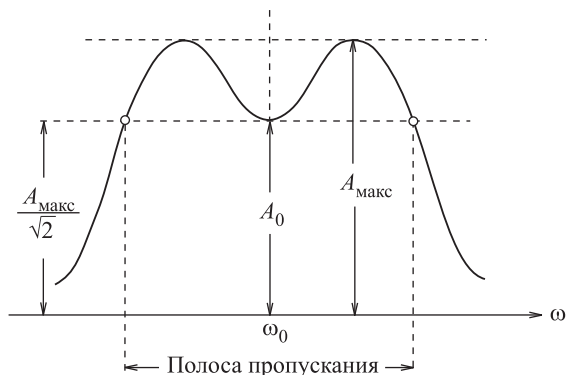


Рис. 87

Первая величина определяется по формуле (103). Вторую получим, положив в (103) $\varepsilon = 0$:

$$A_0 = \frac{k}{k^2 + d^2}.$$

Чтобы найти третью, нужно подставить в (103) значения ε , соответствующие максимумам, т. е.

$$\varepsilon^2 = k^2 - d^2.$$

Это даст

$$A_{\text{макс}} = \frac{1}{2d}. \quad (106)$$

Теперь составим два уравнения:

$$\frac{A}{A_{\text{макс}}} = \frac{2kd}{\sqrt{(d^2 + k^2)^2 + 2(d^2 - k^2)\varepsilon^2 + \varepsilon^4}} = \frac{1}{\sqrt{2}},$$

$$\frac{A_0}{A_{\text{макс}}} = \frac{2kd}{d^2 + k^2} = \frac{1}{\sqrt{2}}.$$

Из второго уравнения находим (выбирая значение $k > d$)

$$k = (1 + \sqrt{2})d.$$

Подставляя это значение в первое уравнение, получаем два решения: $\varepsilon = 0$ (что отвечает одному из наших условий) и

$$|\varepsilon| = 2\sqrt{1 + \sqrt{2}}d \approx 3,1d.$$

Напомним еще раз, что

$$\varepsilon = 1 - \frac{\omega^2}{\omega_0^2} \approx \frac{2\Delta\omega}{\omega_0}.$$

Итак, при том же затухании относительная ширина полосы для связанных контуров получается втрое большей, чем для одиночного контура. Если же взять одинаковую полосу, то затухание связанных контуров можно уменьшить втрое, а при этом скаты характеристики будут соответственно круче, как и показано на рис. 86. К тому же и коэффициент передачи будет при этом в полтора раза больше, так как для одиночного контура

$$A_{\text{макс}} = 1/d,$$

а в формулу (106) для связанных контуров нужно подставить

$$d_1 = \frac{d}{3,1}.$$

В некоторых применениях связанных контуров (например, в передатчиках) задача состоит в том, чтобы получить во втором контуре *наибольшую мощность*, выделяемую на активном сопротивлении. (Это может быть, например, сопротивление излучения антенны.) Мы

должны выяснить, как следует выбрать коэффициент связи, чтобы удовлетворить поставленному условию. Достаточно, очевидно, найти соотношения, при которых получается наибольший ток.

Пусть теперь контуры неодинаковы, так что в уравнениях (101) нужно положить $Z_1 \neq Z_2$. Для тока во втором контуре получим

$$\dot{I}_2 = \dot{E} \frac{i \omega M}{\omega^2 M^2 + Z_1 Z_2}.$$

Пусть оба контура настроены на рабочую частоту, т. е.

$$\omega = \omega_0, \quad X_1 = X_2 = 0, \quad Z_1 Z_2 = R_1 R_2.$$

Тогда

$$\dot{I}_2 = \dot{E} \frac{i \omega_0 M}{\omega_0^2 M^2 + R_1 R_2},$$

или

$$\dot{I}_2 = \frac{E}{\sqrt{\rho_1 \rho_2}} \frac{i k}{d_1 d_2 + k^2}.$$

Легко видеть, что наибольшая амплитуда тока получается при

$$k = \sqrt{d_1 d_2}.$$

Такая связь называется *оптимальной*. Оказывается, что для рассмотренного случая одинаковой настройки обоих контуров оптимальная связь совпадает с критической. При одинаковой настройке обоих контуров (так называемый *полный резонанс*) и при оптимальной связи получается наибольшее возможное значение тока во втором контуре, а следовательно, и наибольшая мощность, выделяемая на активном сопротивлении второго контура.

Займемся теперь временными характеристиками системы связанных контуров. Для этого представим коэффициент передачи в операционной форме. Заменяя в уравнениях (100) $i \omega$ на p , получим вместо (102)

$$K(p) = k \frac{\frac{p^2}{\omega_0^2}}{(1 - k^2) \frac{p^4}{\omega_0^4} + 2d \frac{p^3}{\omega_0^3} + (d + 2) \frac{p^2}{\omega_0^2} + 2d \frac{p}{\omega_0} + 1}. \quad (107)$$

Для того чтобы найти переходную функцию системы, нужно найти оригинал, соответствующий изображению (107). Но здесь возникает затруднение: характеристическое уравнение $H_2(p) = 0$ в нашем случае есть полное уравнение *четвертой* степени, решить которое в общем виде мы не в состоянии. Выход из этого затруднения очень прост: будем решать задачу, положив $d = 0$, в результате чего характеристическое уравнение станет биквадратным. По этому поводу нужно заметить, что пренебрежение активными сопротивлениями является вообще очень полезным и употребительным приемом предварительного ис-

следования сколько-нибудь сложных систем. Пренебрегая активными сопротивлениями при исследовании как частотных, так и временных соотношений, мы достигаем очень существенного упрощения. Достаточно сказать, что из выражений для частотных характеристик при этом устраняется комплексность. Разумеется, система без активных сопротивлений нереальна и решения, которые получаются из упрощенных таким образом уравнений, не соответствуют действительным процессам. Но можно очень легко восстановить истину (хотя бы качественно) на основании простых рассуждений общезначимого характера.

Действительно, если отбросить активные сопротивления, то при резонансе частотная характеристика уйдет в бесконечность, в реальной же системе максимум конечен, и величина его зависит как раз от активного сопротивления. В составе временных характеристик при отбрасывании активных сопротивлений появляются незатухающие колебания. В реальной же системе все свободные колебания должны быть затухающими, причем скорость затухания зависит от активных сопротивлений. Именно такого рода поправки мы должны внести в результат исследования идеализированной системы без потерь (т. е. без активных сопротивлений). Поясим эти соображения на примере одиночного контура.

Возьмем формулу (83) для частотной характеристики и положим в ней $d = 0$. Получим

$$A = \frac{1}{1 - \frac{\omega^2}{\omega_0^2}}.$$

График этой функции показан на рис. 88. При $\omega = \omega_0$ $A_{\text{макс}} \rightarrow \infty$. Но в действительности $A_{\text{макс}} = 1/d$. Значит, кривая должна пройти через точку с координатами $(\omega, 1/d)$, отмеченную на рисунке. Резонансная кривая должна иметь вид, показанный пунктирной линией.

Обратимся к временным характеристикам. В выражении (92) для изображения переходной характеристики положим $\alpha = 0$.

Получим

$$K(p) = \frac{\omega_0^2}{p^2 + \omega_0^2}$$

откуда

$$h(t) = 1 - \cos \omega_0 t$$

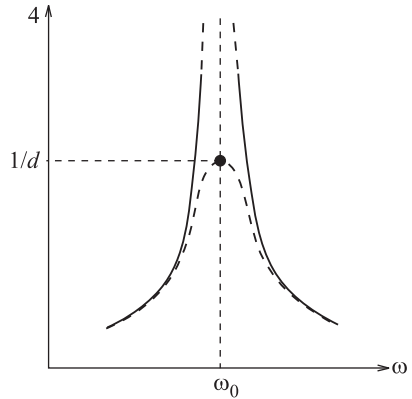


Рис. 88

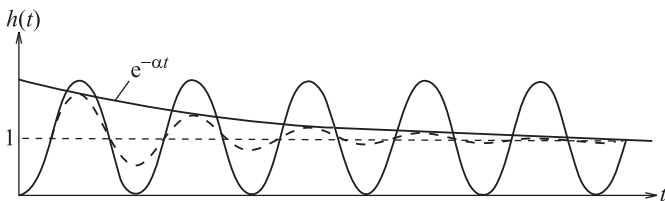


Рис. 89

(рис. 89). Но в реальной системе колебательная составляющая должна затухать по закону $e^{-\alpha t}$. Следовательно, реальная переходная характеристика должна иметь вид, показанный на рис. 89 пунктирной линией.

После этих замечаний вернемся к нашей задаче. Итак, положив в (107) $d = 0$, получим

$$K(p_1) = k \frac{p_1^2}{(1 - k^2)p_1^4 + 2p_1^2 + 1},$$

где $p_1 = p/\omega_0$. Решая биквадратное характеристическое уравнение

$$p_1^4 + \frac{2}{1 - k^2}p_1^2 + \frac{1}{1 - k^2} = 0,$$

найдем четыре корня

$$p_{1i} = \pm \frac{i}{\sqrt{1 \pm k}}.$$

Далее,

$$H_2^l(p_1) = 4p_1 [p_1^2(1 - k^2) + 1].$$

Таким образом,

$$h(t) = \frac{k}{4} \sum_{i=1}^4 \frac{e^{p_{1i}\omega_0 t}}{(1 - k^2)p_{1i}^2 + 1}$$

(переход в изображении от p к $p_1 = p/\omega_0$ означает переход в оригинале к так называемому собственному времени $\tau = \omega_0 t$). Подставляя значения корней, найдем

$$h(t) = \frac{1}{2} (\cos \omega_1 t - \cos \omega_2 t), \quad (108)$$

где

$$\omega_1 = \frac{\omega_0}{\sqrt{1 + k}}, \quad \omega_2 = \frac{\omega_0}{\sqrt{1 - k}}$$

— частоты связи (без поправки на затухание). Итак, переходная функция состоит из двух колебаний с частотами связи. В результате их сложения получаются биения, как показано на рис. 90, а. Частота биений равна

$$\Omega = |\omega_1 - \omega_2| = \omega_0 \left| \frac{1}{\sqrt{1 + k}} - \frac{1}{\sqrt{1 - k}} \right|.$$

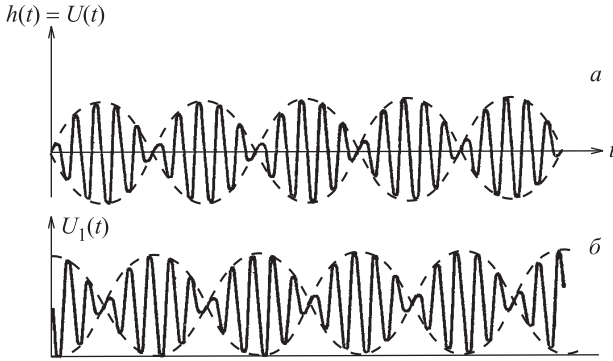


Рис. 90

При малых k получаем приближенную зависимость

$$\Omega \approx k\omega_0.$$

Функция $h(t)$ выражает изменение во времени выходного напряжения (на емкости C_2) при подаче на вход единичной эдс:

$$e(t) = \sigma(t)$$

(см. схему рис. 91). Для полноты физической картины найдем еще напряжение на емкости C_1 в первом контуре. Исключая I_2 из уравнений (см. (101)):

$$\left. \begin{aligned} Z\bar{i}_1 - pM\bar{i}_2 &= \bar{e}, \\ Z\bar{i}_2 - pM\bar{i}_1 &= 0 \end{aligned} \right\},$$

получим

$$\bar{i}_1 = \frac{Z}{Z^2 - p^2 M^2}, \quad \bar{u}_1 = \bar{i}_1 \frac{1}{pC} = \frac{\frac{Z}{pC}}{Z^2 - p^2 M^2}.$$

Подставляя сюда

$$Z = pL + \frac{1}{pC},$$

получим

$$\bar{u}_1(p) = \frac{p_1^2 + 1}{(1 - k^2)p_1^4 + 2p_1^2 + 1}.$$

Оригинал запишется в виде

$$u_1(t) = 1 + \frac{1}{4} \sum_{i=1}^4 \frac{(p_{1i}^2 + 1) e^{p_{1i} \omega_0 t}}{p_{1i}^2 [p_{1i}^2 (1 - k^2) + 1]}.$$

Подставляя значения корней, найдем

$$u_1(t) = 1 - \frac{1}{2} (\cos \omega_1 t + \cos \omega_2 t). \quad (109)$$

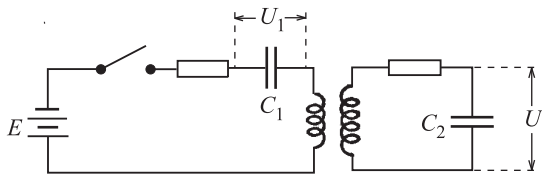


Рис. 91

Для большей наглядности перепишем формулы (108) и (109) в следующем виде:

$$h(t) = u(t) = \sin \frac{\omega_1 - \omega_2}{2} t \sin \frac{\omega_1 + \omega_2}{2} t,$$

$$u_1(t) = 1 - \cos \frac{\omega_1 - \omega_2}{2} t \cos \frac{\omega_1 + \omega_2}{2} t.$$

При малом k приблизительно

$$\frac{\omega_1 - \omega_2}{2} = \frac{k}{2} \omega_0, \quad \frac{\omega_1 + \omega_2}{2} = \omega_0.$$

График $u_1(t)$ изображен на рис. 90, б. Как видим, *огibaющие* сдвинуты по фазе на $\pi/2$. Физический смысл полученной картины состоит в том, что колебательная энергия периодически перекачивается из одного контура в другой; попеременно в одном контуре амплитуда достигает максимума, в то время как в другом контуре она падает до нуля.

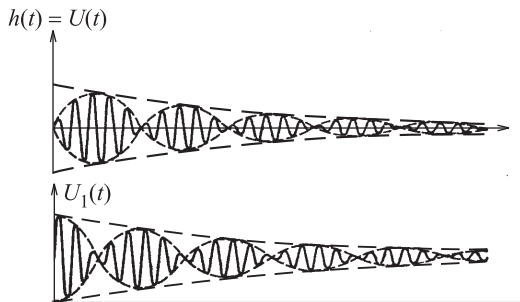


Рис. 92

Если теперь вернуться к реальной системе, то колебания в ней должны затухать и картина рис. 90 изменится, как показано на рис. 92.

§ 34. Системы со многими степенями свободы

В радиотехнике не ограничиваются системами из двух связанных контуров и часто применяют более сложные схемы. Мы приведем сначала некоторые общие соображения по этому поводу. Степень слож-

ности схемы удобно оценивать через *число степеней свободы*. Понятие о степенях свободы — общефизическое понятие; под числом степеней свободы понимается число обобщенных координат, задание которых полностью определяет состояние системы. Применительно к электрическим цепям в качестве обобщенных координат могут рассматриваться токи. Можно определить число степеней свободы для электрической цепи как *число независимых токов, которые могут протекать по цепи*. Можно привести и практическое правило для определения числа степеней свободы: это число равно числу разрывов, которые нужно сделать в цепи, чтобы прекратить любое токопрохождение. Понятно, что речь идет здесь о минимальном (необходимом) числе разрывов.

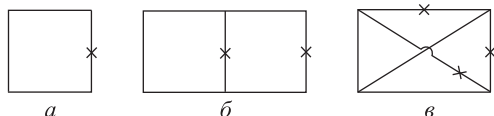


Рис. 93

Из сказанного уже должно быть ясным, что число степеней свободы зависит только от *строения* схемы, т. е. от числа замкнутых контуров, которые она содержит, но не зависит от того, из каких элементов составлена схема. Поэтому для решения вопроса о числе степеней свободы достаточно изобразить только пустой скелет, каркас схемы. На рис. 93 изображены примеры таких скелетов для систем с одной, двумя и тремя степенями свободы. Крестиками отмечены точки разрывов, разрушающих все возможные замкнутые контуры. Показанные на рисунке места разрывов не являются единственно возможными ¹⁾.

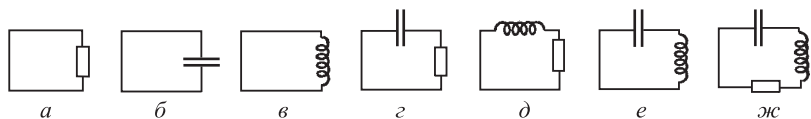


Рис. 94

Обратимся к вопросу о порядке уравнения. Возьмем систему с одной степенью свободы и включим в нее те или иные элементы в различных комбинациях, показанных на рис. 94. Для схемы 94, а порядок уравнения нулевой. Для схем 94, б и 94, в — также; нужно только взять для схемы рис. 94, б за переменную $q = \int i dt$, а для схемы рис. 94, в — $x = di/dt$. Схемы 94, г и 94, д описываются уравнениями

¹⁾ Теории обсуждаемых здесь вопросов относится к топологии — разделу математики, изучающему общие свойства фигур и тел.

первого порядка. Так, для схемы рис. 94, *б* имеем

$$L \frac{di}{dt} + Ri = 0.$$

Схемы рис. 94, *е* и 94, *ж* описываются уравнениями второго порядка. Так, для схемы рис. 94, *ж* уравнение пишется в виде

$$L \frac{di}{dt} + Ri + \frac{1}{C} \int i dt = 0$$

или

$$L \frac{d^2 i}{dt^2} + R \frac{di}{dt} + \frac{1}{C} i = 0.$$

Из этого рассмотрения видно, что порядок уравнения будет нулевым, если система с одной степенью свободы составлена только из однородных элементов. Система описывается уравнением первого порядка, если она содержит сопротивление и реактивный элемент (индуктивность или емкость). И, наконец, система описывается уравнением второго порядка, если она содержит разнородные реактивные элементы (т. е. индуктивность и емкость).

Применим это заключение к системе с несколькими степенями свободы. Составляя уравнения для отдельных контуров (например, по методу контурных токов), получим систему уравнений. Исключая лишние переменные (т. е. составляя уравнение относительно одного какого-либо тока), получим одно уравнение, порядок которого равен сумме порядков уравнений исходной системы. Отсюда сразу следует, что схема любой сложности, составленная из однородных элементов (т. е. только из сопротивлений, или только из емкостей) описывается уравнением нулевого порядка, т. е. не дифференциальным, а алгебраическим. Можно также утверждать, что если n — порядок уравнения, m — число степеней свободы (и то и другое, разумеется, целые числа), то

$$0 \leq n \leq 2m,$$

а также, что

$$n = \sum_{i=1}^m n_i,$$

где n_i — порядок уравнения для i -го контура.

Нужно заметить, что в некоторых случаях результат зависит от выбора контуров. Тогда правильным является наименьшее значение порядка n . Для пояснения этих соотношений на рис. 95 дано несколько примеров.

Система с двумя степенями свободы описывается уравнением второго порядка для случая рис. 95, *а*, третьего порядка — для рис. 95, *б*, четвертого порядка — для рис. 95, *в*.

Быстрое определение порядка уравнения представляет интерес потому, что от порядка уравнения зависит число возможных резонансов или собственных частот системы. Именно, каждой степени свободы

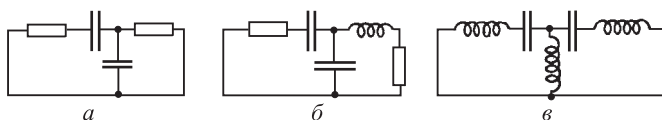


Рис. 95

может соответствовать один резонанс (на частоте, в общем случае не совпадающей с резонансной частотой данного контура, взятого отдельно). В системе с m степенями свободы *могут* наблюдаться m резонансов на различных частотах. Для этого обязательно, во-первых, чтобы каждый из контуров описывался уравнением второго порядка так, чтобы уравнение системы в целом имело наивысший возможный порядок, т. е. $n = 2m$. Второе условие наличия m резонансов состоит в том, что характеристическое уравнение должно иметь — $n/2$ пар сопряженных комплексных корней, мнимые части которых определяют $n/2 = m$ резонансных частот (при малых затуханиях модули мнимых частей — это и есть резонансные частоты).

Таким образом, например, система с двумя степенями свободы, описываемая уравнением четвертого порядка, может иметь два резонанса. Примером такой системы являются связанные контуры, рассмотренные в предыдущем параграфе.

Если образовать систему, например, из трех контуров, т. е. систему с тремя степенями свободы, то на частотной характеристике такой системы можно обнаружить (при выполнении вышеуказанных условий) соответственно три резонанса.

Это открывает возможность дальнейшего усовершенствования избирательных звеньев радиотехнических устройств в смысле дальнейшего приближения их частотных характеристик к идеальной. Так, беря системы из трех или пяти контуров (такие системы применяются в современной радиотехнике), можно получить характеристики вида рис. 96. Эти характеристики займут промежуточное положение между характеристиками рис. 86, б и в.

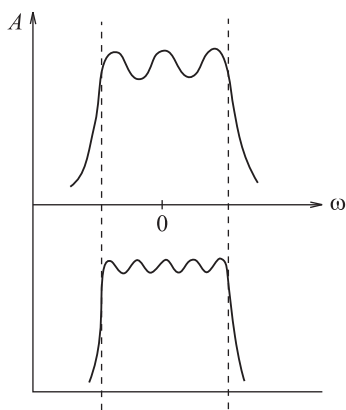


Рис. 96

Теоретическое исследование систем со многими степенями свободы весьма затруднено из-за того, что характеристические уравнения являются алгебраическими уравнениями высоких степеней. Это затруднение возникает уже при исследовании связанных контуров, т. е. системы с двумя степенями свободы (характеристическое уравнение имеет четвертую степень). Поэтому при исследовании и расчете систем со многими степенями свободы приходится либо поль-

зоваться вспомогательными таблицами или графиками, полученными путем численного решения соответствующих уравнений, либо получать желаемые соотношения путем опытного подбора параметров системы. И тот и другой путь приводят к полному решению поставленной инженерной задачи.

Вместе с тем ощущается потребность в таком развитии теории, которое позволило бы в общем виде рассматривать вопросы, относящиеся к системам со многими степенями свободы. Оказывается, что если определенным образом специализировать свойства систем со многими степенями свободы, т. е. наложить на их строение и режим некоторые ограничивающие условия, то построение соответствующей теории, не испытывающей никаких трудностей при увеличении числа степеней свободы, оказывается возможным. Речь идет о теории четырехполусников и вытекающей из нее теории фильтров.

§ 35. Краткие сведения из теории четырехполусников

Современная теория четырехполусников представляет собой обширную и детально разработанную теоретическую дисциплину, широко использующую матричное исчисление. Однако для наших целей достаточны лишь элементарные основания теории четырехполусников, которые и излагаются в этом параграфе.

Объектом изучения являются общие свойства четырехполусников и их соединений. Четырехполусником называется система, имеющая вход и выход (пару входных и пару выходных зажимов, всего четыре «полуса»). Теория до некоторого этапа совершенно игнорирует внутреннее устройство четырехполусника. Наделяя четырехполусник некоторыми вполне общими свойствами, теория изучает лишь соотношения между величинами на его входе и выходе.

Такой обобщенный подход позволяет получить целый ряд заключений универсального характера; эти заключения могут быть затем применены к любым конкретным схемам. Мы пойдем именно этим путем: общие выводы теории четырехполусников применим к изучению электрических фильтров.

Работа четырехполусника, как звена некоторой сложной цепи, полностью определяется соотношениями между четырьмя переменными величинами, а именно: входным током и напряжением и выходным током и напряжением. Если четырехполусник линеен, то между этими четырьмя величинами существуют линейные зависимости, которые и образуют *уравнения четырехполусника*. Классическая форма этих уравнений для установившегося синусоидального режима такова:

$$\left. \begin{aligned} \dot{U}_1 &= a_{11}\dot{U}_2 + a_{12}\dot{I}_2 \\ \dot{I}_1 &= a_{21}\dot{U}_2 + a_{22}\dot{I}_2 \end{aligned} \right\}, \quad (110)$$

где индексы 1 и 2 при переменных величинах $\dot{U}$ и $\dot{I}$ относятся соответственно к входу и выходу (рис. 97). Заметим, что уравнения (110) совершенно универсальны, так как они выражают лишь линейные зависимости между токами и напряжениями.

Из уравнений (110) сразу следует, что с точки зрения соотношений между токами и напряжениями на входе и выходе *всякий* четырехполюсник, независимо от сложности его внутренней схемы, полностью определяется четырьмя коэффициентами a_{11} , a_{12} , a_{21} , a_{22} . Иначе говоря, зная коэффициенты a_{ik} можно рассчитать любой режим четырехполюсника. Коэффициенты a_{ik} могут быть определены из опыта. Так,

$$a_{11} = \left. \frac{\dot{U}_1}{\dot{U}_2} \right|_{I_2=0},$$

т.е. для определения коэффициента a_{11} нужно разомкнуть выходные зажимы (режим хх на выходной стороне четырехполюсника) и измерить отношение входного и выходного напряжений. Коэффициент a_{12} определяется соотношением

$$a_{12} = \left. \frac{\dot{U}_1}{\dot{I}_2} \right|_{U_2=0}.$$

Соответствующий опыт состоит в том, что выходные зажимы четырехполюсника соединяются накоротко; при этом измеряется отношение входного напряжения к току (короткого замыкания) на выходе. Аналогично определяются и остальные коэффициенты. Таким образом, при определении коэффициентов a_{ik} опытным путем нас не интересует внутреннее устройство четырехполюсника, но его необходимо знать при определении коэффициентов a_{ik} расчетным путем.

Обычно рассматриваются так называемые *пассивные* четырехполюсники, т.е. четырехполюсники, не содержащие в себе источников энергии. Для таких четырехполюсников на основании энергетических соображений (которые здесь не приводятся) можно вывести следующее соотношение между коэффициентами a_{ik}

$$\Delta a = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} = 1. \quad (111)$$

Таким образом, равенство единице определителя из коэффициентов a_{ik} выражает условие пассивности четырехполюсника. Из этого следует, что *пассивный* четырехполюсник полностью определяется уже не четырьмя, а только тремя коэффициентами a_{ik} , так как если заданы три коэффициента, то четвертый находится при помощи равенства (111). Дальнейшее упрощение получится, если мы ограничимся рассмотрением *симметричных* четырехполюсников. Симметричным мы



Рис. 97

назовем четырехполюсник, для которого соотношения не изменяются, если поменять местами вход и выход. Такая электрическая симметрия обычно связана с симметрией строения внутренней схемы четырехполюсника. Для симметричного четырехполюсника (по определению) уравнения (110) не должны измениться, если поменять местами индексы 1 и 2 и изменить знаки токов. Это значит, что четырехполюсник рис. 97 теперь питается не слева, а справа. Прodelав такую замену, мы убедимся, что для симметричного четырехполюсника должно быть

$$a_{11} = a_{22}. \quad (112)$$

Это равенство есть условие симметрии четырехполюсника.

Итак, *линейный пассивный симметричный четырехполюсник полностью определяется двумя любыми коэффициентами a_{ik}* . Два остальных коэффициента находятся из условия пассивности (111) и условия симметрии (112). Этот результат сам по себе примечателен: напомним, что мы не делали пока *никаких* предположений о внутреннем устройстве четырехполюсника; его схема может быть как угодно сложна!

Не следует думать, что для описания свойств четырехполюсника нужно пользоваться только коэффициентами a_{ik} ; для этой цели могут служить любые линейно-независимые комбинации из a_{ik} ; или вообще любые независимые параметры, выражающие в той или иной форме соотношения между переменными величинами на входе и на выходе четырехполюсника. Таким образом, можно сказать, что *линейный пассивный симметричный* четырехполюсник полностью определяется двумя независимыми параметрами. (В дальнейшем мы будем иметь в виду только такие четырехполюсники, и все три прилагательных опустим.)

Теперь возникает вопрос о том, как выбрать описывающие свойства четырехполюсника параметры. Выбор велик: четырехполюсник можно описывать множеством способов. Но естественно подчинить выбор требованию удобства исследования. Употребляются преимущественно два способа описания свойств четырехполюсника, а именно:

1. Посредством *сопротивлений холостого хода и короткого замыкания*.
2. Посредством *характеристического сопротивления и постоянной передачи*.

Мы дадим сейчас определения этих параметров и установим связи между ними.

Сопротивлением холостого хода $Z_{1\infty}$ называется входное сопротивление четырехполюсника при режиме хх на выходной стороне (т. е. при разомкнутых выходных зажимах). Сопротивлением короткого замыкания Z_{10} называется входное сопротивление при режиме кз на выходе (т. е. при замкнутых накоротко выходных зажимах).

Экспериментальным путем $Z_{1\infty}$ и Z_{10} определяются из опытов хх и кз. Связь $Z_{1\infty}$ и Z_{10} с коэффициентами a_{ik} легко установить. Для

этого прежде всего найдем общее выражение для *входного сопротивления* четырехполюсника, разделив друг на друга два уравнения (110):

$$Z_1 = \frac{\dot{U}_1}{\dot{I}_1} = \frac{a_{11}\dot{U}_2 + a_{12}\dot{I}_2}{a_{21}\dot{U}_2 + a_{22}\dot{I}_2}.$$

Разделим числитель и знаменатель на $\dot{I}_2$ и заметим, что

$$\frac{\dot{U}_2}{\dot{I}_2} = Z_2$$

есть не что иное, как выходное сопротивление четырехполюсника, т. е. сопротивление нагрузки. Выражение для входного сопротивления запишется в виде

$$Z_1 = \frac{a_{11}Z_2 + a_{12}}{a_{21}Z_2 + a_{22}}. \quad (113)$$

При холостом ходе $Z_2 \rightarrow \infty$. Таким образом, сопротивление хх

$$Z_{1\infty} = \frac{a_{11}}{a_{21}}.$$

При коротком замыкании $Z_2 = 0$; сопротивление кз

$$Z_{10} = \frac{a_{12}}{a_{22}}.$$

Введенные обозначения для сопротивлений хх и кз имеют следующий смысл:

$$Z_{1\infty} = Z_{1|Z_2 \rightarrow \infty}, \quad Z_{10} = Z_{1|Z_2 = 0}.$$

Обратимся к характеристическому сопротивлению¹⁾. Оно определяется как

$$Z_c = \sqrt{Z_{1\infty}Z_{10}},$$

т. е. характеристическое сопротивление равно среднегеометрическому из сопротивлений хх и кз. Через коэффициенты a_{ik} характеристическое сопротивление выражается формулой

$$Z_c = \sqrt{\frac{a_{12}}{a_{21}}}. \quad (114)$$

Характеристическое сопротивление обладает замечательным свойством *повторности*²⁾. Свойство это состоит в том, что если включить характеристическое сопротивление на выход в качестве нагрузки, т. е. положить $Z_2 = Z_c$, то и входное сопротивление окажется равным характеристическому.

¹⁾ Не смешивайте с характеристическим сопротивлением $\rho = \sqrt{L/C}$ колебательного контура.

²⁾ В более общем случае несимметричного четырехполюсника определения характеристического и повторного сопротивления не совпадают.

Действительно, из (113) получаем

$$Z_1 = \frac{a_{11}Z_c + a_{12}}{a_{21}Z_c + a_{22}} = Z_c \frac{a_{11}Z_c + a_{12}}{a_{22}Z_c + a_{21}Z_c^2} = Z_c.$$

Постоянная передачи g определяется как натуральный логарифм отношения напряжений на входе и выходе четырехполюсника при нагрузке его на характеристическое сопротивление, т. е.

$$g = \ln \left. \frac{\dot{U}_1}{\dot{U}_2} \right|_{Z_2=Z_c},$$

или

$$\left. \frac{\dot{U}_1}{\dot{U}_2} \right|_{Z_2=Z_c} = e^g.$$

Выходное напряжение определяется через входное формулой

$$\dot{U}_2 = \dot{U}_1 e^{-g}.$$

Таким образом, величина e^{-g} есть не что иное, как коэффициент передачи четырехполюсника. Нужно только иметь в виду, что определение постоянной g подразумевает вполне определенный выходной режим четырехполюсника, а именно:

$$\frac{\dot{U}_2}{\dot{I}_2} = Z_2 = Z_c.$$

Оказывается, что при этих условиях отношение токов на входе и выходе равно отношению напряжений. Действительно,

$$\dot{U}_2 = \dot{I}_2 Z_2, \quad \dot{U}_1 = \dot{I}_1 Z_1,$$

откуда вообще

$$\frac{\dot{U}_1}{\dot{U}_2} = \frac{\dot{I}_1}{\dot{I}_2} \frac{Z_1}{Z_2}.$$

Но при нагрузке на характеристическое сопротивление

$$Z_1 = Z_2 = Z_c$$

и, следовательно, *при этом условии*

$$\frac{\dot{U}_1}{\dot{U}_2} = \frac{\dot{I}_1}{\dot{I}_2} = e^g.$$

При произвольной нагрузке отношение токов и отношение напряжений, разумеется, не равны.

Постоянная передачи связана простыми соотношениями с ранее введенными параметрами. Из уравнения (110) находим

$$\frac{U_1}{U_2} = a_{11} + \sqrt{a_{12}a_{21}} = e^g.$$

Составим выражение

$$\operatorname{ch} g = \frac{1}{2}(e^g + e^{-g}) = \frac{1}{2} \left(a_{11} + \sqrt{a_{12}a_{21}} + \frac{1}{a_{11} + \sqrt{a_{12}a_{21}}} \right).$$

Приведя к общему знаменателю и воспользовавшись соотношениями (111) и (112), получим

$$\operatorname{ch} g = a_{11}. \quad (115)$$

Далее,

$$\operatorname{sh} g = \sqrt{\operatorname{ch}^2 g - 1} = \sqrt{a_{11}^2 - 1} = \sqrt{a_{12}a_{21}}.$$

Пользуясь выражением (114) для Z_c найдем еще

$$Z_c \operatorname{sh} g = a_{12}, \quad \frac{1}{Z_c} \operatorname{sh} g = a_{21} \quad (116)$$

и, наконец,

$$\operatorname{th} g = Z_c \frac{a_{21}}{a_{11}},$$

или

$$\operatorname{th} g = \sqrt{\frac{Z_{10}}{Z_{1\infty}}}. \quad (117)$$

Рассмотрим физический смысл величины g . При синусоидальных колебаниях отношение напряжений на входе и выходе выражается отношением комплексных амплитуд. Постоянная передачи g при этом будет комплексной величиной, которую можно записать в виде

$$g = b + i a.$$

Мы имеем

$$\frac{\dot{U}_2}{\dot{U}_1} = K_c = e^{-g} = e^{-b} e^{-i a} = A_c e^{-i a}.$$

Множитель

$$A_c = |K_c| = e^{-b}$$

выражает отношение амплитуд напряжений на выходе и на входе. Постоянная b , характеризующая степень уменьшения амплитуды в результате прохождения через четырехполюсник, называется *постоянной затухания*, или просто *затуханием* (не смешивать затухание b

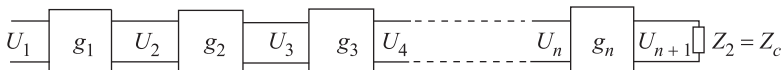


Рис. 98

четырёхполосника с затуханием $d = 1/Q$ колебательного контура!) ¹⁾. Множитель $e^{-i a}$, равный единице по абсолютному значению, выражает сдвиг фазы выходного напряжения по отношению к входному. Величина a носит название *фазовой постоянной*.

В заключение найдем постоянную передачи для цепочки четырехполосников, соединенных друг за другом, как показано на рис. 98. Предполагается, что все четырехполосники обладают одинаковым характеристическим сопротивлением Z_c и что цепочка нагружена на такое же сопротивление. В этом случае получаются очень простые соотношения. Составим выражение для отношения напряжений на выходе и на входе цепочки

$$\begin{aligned} \frac{\dot{U}_{n+1}}{\dot{U}_1} &= e^{-\Gamma} = \frac{\dot{U}_2}{\dot{U}_1} \cdot \frac{\dot{U}_3}{\dot{U}_2} \dots \frac{\dot{U}_n}{\dot{U}_{n-1}} \cdot \frac{\dot{U}_{n+1}}{\dot{U}_n} = \\ &= e^{-g_1} e^{-g_2} \dots e^{-g_n} = \exp\left(-\sum_{k=1}^n g_k\right), \end{aligned}$$

отсюда

$$\Gamma = \sum_{k=1}^n g_k,$$

т. е. постоянная передачи цепочки равна сумме постоянных передачи, составляющих цепочку четырехполосников. Если все четырехполосники одинаковы, то

$$\Gamma = ng, \quad (118)$$

¹⁾ При количественном определении затухания пользуются логарифмическими единицами — *неперами* или *децибелами*. Затухание будет выражено в неперах, если определить его как

$$\ln \frac{U_1}{U_2} = \ln \frac{I_1}{I_2}.$$

Следовательно, затуханию в один непер соответствует коэффициент передачи

$$\frac{U_2}{U_1} = \frac{1}{e} = 0,368,$$

где $e \approx 2,718$ — основание натуральных — неперовых — логарифмов, откуда и название единицы. Затухание будет выражено в децибелах, если определить его как

$$10 \lg \left(\frac{U_1}{U_2} \right)^2 = 20 \lg \frac{U_1}{U_2} = 20 \lg \frac{I_1}{I_2}.$$

Таким образом, затуханию в 10 дБ соответствует коэффициент передачи

$$\frac{U_2}{U_1} = \frac{1}{\sqrt{10}} = 0,316.$$

Так как разница между неперами и децибелами сводится к различию в основаниях логарифмов (неперы определяются на основе натуральных логарифмов, а децибелы — на основе десятичных), то одни единицы могут просто переводиться в другие, а именно:

1 непер = 8,686 децибел; 1 децибел = 0,1151 непера.

т. е. как затухание, так и сдвиг фаз для цепочки из n звеньев в n раз больше, чем для одного звена.

Простые соотношения теории четырехполюсников позволяют строить систему с любым числом степеней свободы при условии, однако, что она может быть разбита на отдельные четырехполюсники с одинаковыми характеристическими сопротивлениями. Это условие не так уж стеснительно. Во всяком случае, на этом пути мы получаем решение многих важных технических задач.

§ 36. Фильтры; общие понятия

Электрическим фильтром называется устройство, пропускающее колебания одних частот и не пропускающее колебания других частот. Такая характеристика требует немедленного уточнения и перевода на инженерный язык. Прежде всего, что значит «пропускает»? Всякий фильтр есть, конечно, некоторый четырехполюсник, характеризующийся коэффициентом передачи. Мы скажем, что фильтр «пропускает» колебания данной частоты, если амплитуда на выходе не уменьшилась по сравнению с амплитудой на входе. Это значит, что для «пропускаемой» частоты

$$|K| = 1,$$

или

$$b = 0$$

(так как $|K| = e^{-b}$ — см. § 35). Для других же частот $|K| < 1$ или $b > 0$, и мы скажем, что эти частоты фильтр «не пропускает» или «задерживает». Ясно, что по существу дело сводится не к полному «задерживанию», а лишь к ослаблению, к уменьшению амплитуды колебания в $|K| = e^{-b}$ раз. Таким образом, фильтр тем лучше выполняет свою функцию «непропускания», чем больше затухание b . Полоса «пропускания» называется в технике *полосой прозрачности*; она определяется как полоса частот, в которой $b = 0$ (или, по крайней мере, достаточно мало). Соответственно полоса «задерживания», называемая *полосой непрозрачности*, определяется как полоса частот, в которой b достаточно велико. Теперь мы можем сформулировать определение: фильтр есть четырехполюсник, у которого затухание b в некоторой полосе — полосе прозрачности (или полосах) — мало отличается от нуля, в остальном же диапазоне частот достаточно велико (значения затухания в полосе прозрачности и в полосе непрозрачности задаются техническими условиями).

Из этого определения сразу же вытекает и способ аналитического определения полосы прозрачности. Для этого запишем выражение для постоянной передачи

$$g = b + ia.$$

В полосе прозрачности должно быть $b = 0$. Значит, g должно быть чисто мнимой величиной (мы рассматриваем здесь и в дальнейшем только идеализированные соотношения для фильтров без потерь).

Возьмем теперь формулу (117):

$$\operatorname{th} g = \sqrt{\frac{Z_{10}}{Z_{1\infty}}}.$$

Известно, что

$$\operatorname{th} i x = i \operatorname{tg} x.$$

Значит, если g — мнимая величина, то и $\operatorname{th} g$ — мнимая величина. Стало быть, подкоренное выражение в правой части должно быть отрицательным. На этом основании аналитическое выражение для полосы прозрачности может быть записано в виде неравенства

$$-\infty < \frac{Z_{10}}{Z_{1\infty}} < 0. \quad (119)$$

Это — очень важный результат, но это и все, что можно сказать, рассматривая вопрос в общем виде. В дальнейшем мы должны уже рассматривать схемы определенной структуры.

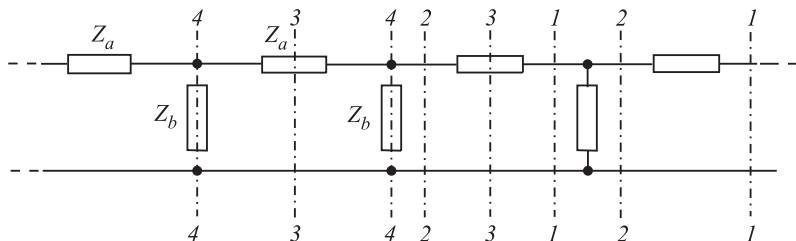


Рис. 99

Основой для построения простейших фильтров служит схема рис. 99, которую будем называть лестничной. Она составлена из одинаковых последовательных сопротивлений Z_a и одинаковых параллельных сопротивлений Z_b . Лестничную схему можно рассматривать как цепочку четырехполюсников. Разделение схемы на отдельные четырехполюсники — *звенья* — можно выполнить различными способами. Так, например, можно выделить звенья, проводя разрезы по пунктирным линиям 1-1 или 2-2. Получаемые при этом звенья имеют строение, показанное на рис. 100, а и б. Это — так называемые Г-звенья (по

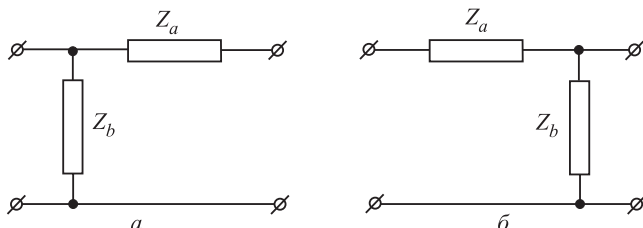


Рис. 100

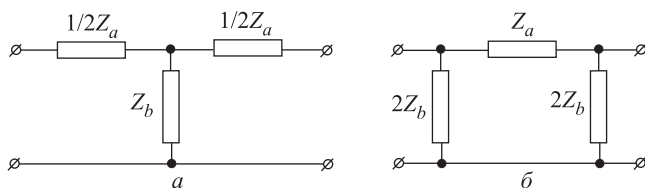


Рис. 101

очертанию буквы Г), а именно правое и левое Г-звенья. Такого вида звенья несимметричны, и для их рассмотрения пришлось бы пользоваться усложненной теорией. Но лестничную схему рис. 99 можно разделить и на симметричные звенья, и притом двумя способами. Для этого нужно провести разрезы, как показано пунктирными линиями 3–3 или 4–4. В первом случае разрезается пополам последовательное сопротивление Z_a , во втором — расщепляется параллельное сопротивление Z_b . Полученные звенья называются Т-звеньями и П-звеньями. На рис. 101, *a* и *б* показаны Т-звено и П-звено. Легко убедиться, что при соединении цепочкой звеньев любого вида восстанавливается строение исходной лестничной схемы. Однако теперь следует обратить внимание на начала и концы цепочек. Именно, если цепочка собрана из Т-звеньев, то она должна начинаться и кончаться последовательным сопротивлением — $\frac{1}{2}Z_a$. Если же цепочка состоит из П-звеньев, то она должна начинаться и кончаться параллельным сопротивлением $2Z_b$. И то и другое показано на рис. 102, *a* и *б*.

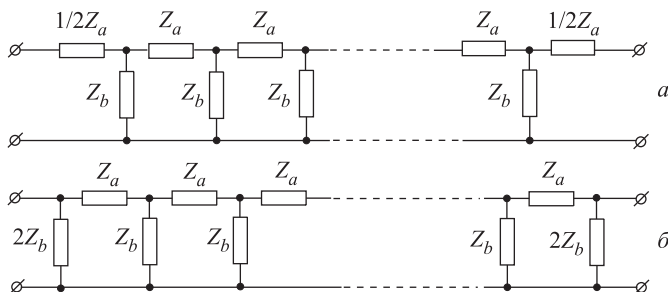


Рис. 102

Займемся теперь вычислением постоянной передачи для Т- и П-звеньев. Нужно прежде всего определить сопротивления холостого хода и короткого замыкания. На основании схем рис. 101, *a* получаем для Т-звена

$$(Z_{10})_T = \frac{1}{2}Z_a + \frac{1}{\frac{1}{Z_b} + \frac{2}{Z_a}} = \frac{1}{2}Z_a + \frac{Z_b}{1 + 2\frac{Z_b}{Z_a}} = \frac{1}{2}Z_a \frac{1 + 4\frac{Z_b}{Z_a}}{1 + 2\frac{Z_b}{Z_a}},$$

$$(Z_{1\infty})_T = \frac{1}{2}Z_a + Z_b = \frac{1}{2}Z_a \left(1 + 2\frac{Z_b}{Z_a}\right).$$

Для постоянной передачи имеем

$$\text{th } g_T = \sqrt{\frac{(Z_{10})_T}{(Z_{1\infty})_T}} = \frac{\sqrt{1 + 4\frac{Z_b}{Z_a}}}{1 + 2\frac{Z_b}{Z_a}}.$$

Для П-звена из рис. 101, б получаем:

$$(Z_{10})_П = \frac{1}{\frac{1}{2Z_b} + \frac{1}{Z_a}} = \frac{2Z_b Z_a}{Z_a + 2Z_b} = 2Z_b \frac{1}{1 + 2\frac{Z_b}{Z_a}},$$

$$(Z_{1\infty})_П = \frac{1}{\frac{1}{2Z_b} + \frac{1}{Z_a + 2Z_b}} = \frac{2Z_b(Z_a + 2Z_b)}{Z_a + 4Z_b} = 2Z_b \frac{1 + 2\frac{Z_b}{Z_a}}{1 + 4\frac{Z_b}{Z_a}},$$

откуда

$$\text{th } g_П = \frac{\sqrt{1 + 4\frac{Z_b}{Z_a}}}{1 + 2\frac{Z_b}{Z_a}},$$

т.е. тот же результат, что и для Т-звена. Это и не удивительно, так как постоянная передачи определяется строением исходной лестничной схемы вне зависимости от того, каким способом она разделена на звенья. К тому же и описанное деление есть *мысленная* операция, а не фактическое разрезание схемы на куски.

Итак, для звеньев обоего вида (Т и П) получили

$$\text{th } g = \frac{\sqrt{1 + 4\frac{Z_b}{Z_a}}}{1 + 2\frac{Z_b}{Z_a}}. \quad (120)$$

Теперь можно сформулировать условие, определяющее полосу прозрачности уже не в общем виде (см. (119)), а для лестничной схемы, составленной из сопротивлений Z_a и Z_b . Как видно из (120), постоянная передачи будет мнимой, если подкоренное выражение отрицательно. Стало быть, выражение для полосы прозрачности запишется в виде неравенства

$$-\infty < 1 + 4\frac{Z_b}{Z_a} < 0,$$

или

$$-\infty < 4\frac{Z_b}{Z_a} < -1,$$

или

$$1 < -4 \frac{Z_b}{Z_a} < \infty. \quad (121)$$

В полосе непрозрачности $b \pm 0$, а постоянная передачи g в общем случае комплексная величина. Чтобы определить затухание в полосе непрозрачности, можно воспользоваться формулой

$$g = b + i a = \operatorname{Ar} \operatorname{th} (x + i y) = \frac{1}{2} \left(\operatorname{Ar} \operatorname{th} \frac{2x}{1+x^2+y^2} + i \operatorname{arc} \operatorname{tg} \frac{2y}{1-x^2-y^2} \right),$$

где x и y — вещественная и мнимая части выражения в правой части (120). Таким образом

$$b = \frac{1}{2} \operatorname{Ar} \operatorname{th} \frac{2x}{1+x^2+y^2},$$

а фазовая постоянная

$$a = \frac{1}{2} \operatorname{arc} \operatorname{tg} \frac{2y}{1-x^2-y^2}.$$

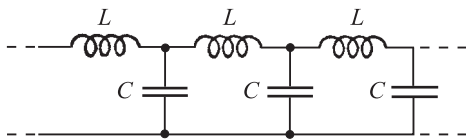


Рис. 103

Перейдем к конкретным схемам. Пусть лестничная схема составлена из индуктивностей L и емкостей C , как показано на рис. 103. Тогда

$$Z_a = i \omega L, \quad Z_b = \frac{1}{i \omega C}, \quad 4 \frac{Z_b}{Z_a} = -\frac{4}{\omega^2 LC} = -\frac{\omega_k^2}{\omega^2},$$

где

$$\omega_k = \frac{2}{\sqrt{LC}}. \quad (122)$$

Подставляя в (121), получаем для полосы прозрачности

$$1 < \frac{\omega_k^2}{\omega^2} < \infty,$$

или, переходя к обратным величинам,

$$0 < \frac{\omega_k^2}{\omega^2} < 1,$$

или же, извлекая корень и умножая на ω_k

$$0 < \omega < \omega_k.$$

Итак, полоса прозрачности для рассматриваемого фильтра простирается от 0 до ω_k . Величина ω_k называется *граничной частотой*.

Фильтр с описанными свойствами называется *фильтром нижних частот*. В полосе прозрачности затухание $b = 0$; вне полосы прозрачности, т.е. при частотах выше граничной частоты затухание конечно.

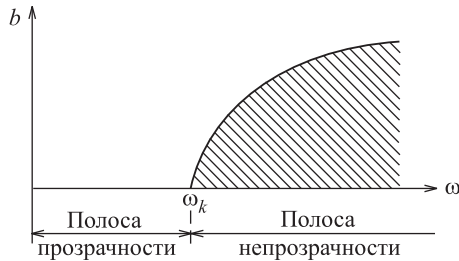


Рис. 104

В полосе непрозрачности $a = 0$ и, следовательно

$$b = \operatorname{Ar} \operatorname{th} \frac{\sqrt{1 - \frac{\omega_k^2}{\omega^2}}}{1 - \frac{\omega_k^2}{2\omega^2}}.$$

Частотная характеристика затухания имеет вид, показанный на рис. 104.

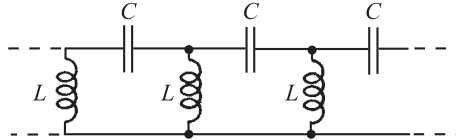


Рис. 105

Рассмотрим теперь фильтр, построенный из последовательно включенных емкостей и параллельно включенных индуктивностей (рис. 105). В этом случае

$$Z_a = \frac{1}{i\omega C}, \quad Z_b = i\omega L, \quad 4\frac{Z_b}{Z_a} = -4\omega^2 LC = -\frac{\omega^2}{\omega_k^2},$$

где

$$\omega_k = \frac{1}{2\sqrt{LC}}$$

(сравни с (122)). Подставляя в (121), получаем для полосы прозрачности

$$1 < \frac{\omega^2}{\omega_k^2} < \infty$$

или

$$\omega_k < \omega < \infty.$$

Фильтр с такими свойствами называется *фильтром верхних частот*. Частотная характеристика затухания показана на рис. 106.

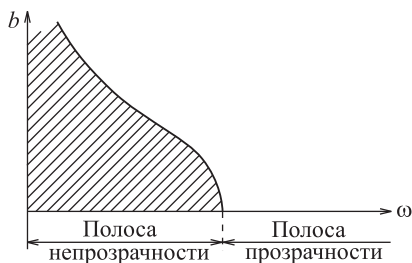


Рис. 106

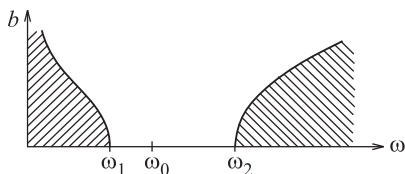


Рис. 107

В радиосвязи большое значение имеет *полосовой фильтр*, отличающийся тем, что полоса прозрачности ограничена сверху и снизу, как показано на частотной характеристике рис. 107. Фильтр с такими свойствами может быть построен по различным схемам. На рис. 108 показана одна из них.

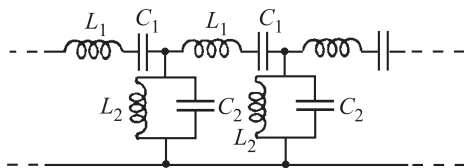


Рис. 108

Найдем граничные частоты ω_1 и ω_2 . Из выражения (121) следует, что граничные частоты определяются уравнениями

$$-4 \frac{Z_b}{Z_a} = 1, \quad -\frac{Z_b}{Z_a} = \infty.$$

В нашем случае:

$$Z_a = i\omega L_1 + \frac{1}{i\omega C_1} = \frac{1}{i\omega C_1}(1 - \omega^2 L_1 C_1),$$

$$Z_b = \frac{1}{i\omega C_2 + \frac{1}{i\omega L_2}} = i\omega L_2 \frac{1}{1 - \omega^2 L_2 C_2}.$$

Положим для упрощения, что последовательные контуры Z_a и параллельные контуры Z_b настроены на одну и ту же частоту

$$\omega_0 = \frac{1}{\sqrt{L_1 C_1}} = \frac{1}{\sqrt{L_2 C_2}}.$$

Запишем

$$Z_a = \frac{1}{i\omega C_1} \left(1 - \frac{\omega^2}{\omega_0^2} \right), \quad Z_b = i\omega L_2 \frac{1}{1 - \frac{\omega^2}{\omega_0^2}}$$

и

$$4 \frac{Z_b}{Z_a} = - \frac{4\omega^2 L_2 C_1}{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2}.$$

Подставляя в первое уравнение границы, получаем

$$4 \frac{\omega^2 L_2 C_1}{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2} = 1,$$

или

$$1 - \frac{\omega^2}{\omega_0^2} = \pm 2\nu \sqrt{L_2 C_1},$$

или

$$\omega^2 \pm 2\omega\omega_0 \sqrt{\frac{L_2}{L_1}} - \omega_0^2 = 0,$$

откуда

$$\omega = \omega_0 \left(\pm \sqrt{\frac{L_2}{L_1}} \pm \sqrt{\frac{L_2}{L_1} + 1} \right).$$

Обозначая

$$\frac{L_2}{L_1} = \frac{C_1}{C_2} = m$$

и сохраняя только положительные значения частоты, находим

$$\omega_1 = \omega_0 \left(\sqrt{1+m} + \sqrt{m} \right), \quad \omega_2 = \omega_0 \left(\sqrt{1+m} - \sqrt{m} \right).$$

Мы получили уже два значения граничной частоты. Что же даст второе, пока не использованное, уравнение границы? Мы имеем

$$-4 \frac{Z_b}{Z_a} = \frac{\omega^2 L_2 C_1}{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2} = \infty,$$

откуда

$$\omega = \omega_0.$$

Это значение не есть граничное — частота ω_0 лежит между ω_1 и ω_2 . Чтобы убедиться в этом, перемножим ω_1 и ω_2

$$\omega_1 \omega_2 = \omega_0^2 \left(\sqrt{1+m} + \sqrt{m} \right) \left(\sqrt{1+m} - \sqrt{m} \right) = \omega_0^2.$$

Таким образом, ω_0 есть среднее геометрическое ω_1 и ω_2 (см. рис. 107).

Поменяв местами последовательные и параллельные контуры, как показано на рис. 109, получим одну из схем *заграждающего фильтра*, имеющего две полосы прозрачности.

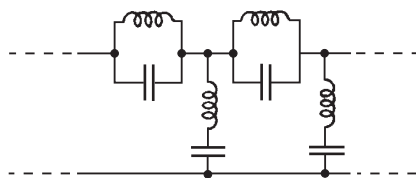


Рис. 109

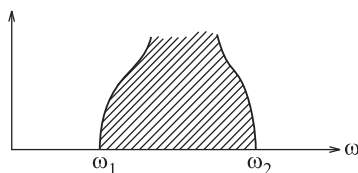


Рис. 110

Мы не будем вычислять граничные частоты, так как выкладки аналогичны предыдущим. Частотная характеристика затухания для заграждающего фильтра изображена на рис. 110.

Остается невыясненным вопрос о роли числа звеньев, так как очевидно, что положение граничных частот определяется только значениями Z_a и Z_b и от числа звеньев не зависит.

Ответ на вопрос дает формула (118): постоянная передачи для цепочки из n одинаковых звеньев, каждое из которых имеет постоянную передачи g , определяется как

$$\Gamma = ng.$$

Следовательно, затухание цепочки в n раз больше затухания одного звена. В полосе прозрачности затухание равно нулю.

Значит, увеличение числа звеньев позволяет увеличить затухание в полосе непрозрачности. При этом возрастает и *крутизна* характеристики затухания около границы, как показано на рис. 111. И то и другое, т.е. как общее увеличение затухания, так и увеличение крутизны характеристики затухания, очень желательны. У идеального фильтра характеристика должна была бы иметь на границе бесконечную крутизну, и затухание в полосе непрозрачности должно было бы быть бесконечно большим. К таким идеальным свойствам можно приближать реальный фильтр путем увеличения количества звеньев.

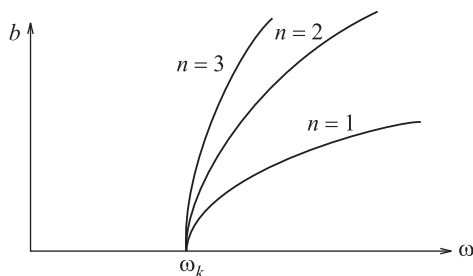


Рис. 111

В заключение заметим, что система связанных контуров с точки зрения фильтров есть не что иное, как звено полосового фильтра. Для того чтобы показать это, удобнее видоизменить схему рис. 83 так,

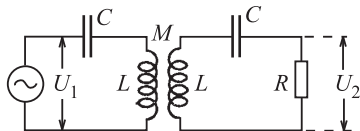


Рис. 112

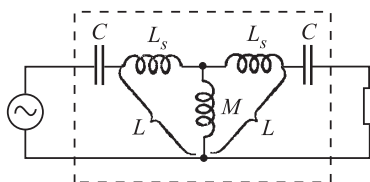


Рис. 113

чтобы выходное напряжение было пропорционально току во втором контуре; пренебрегая затуханием и полагая оба контура одинаковыми, представим схему в виде рис. 112. Эту схему можно также изобразить в виде обычной схемы замещения трансформатора, как показано на рис. 113. На схеме L_s означает индуктивность рассеяния; очевидны следующие соотношения:

$$L = M + L_s, \quad L_s = L - M = L(1 - k) \quad \left(k = \frac{M}{L}\right).$$

Как видно, схема представляет Т-звено полосового фильтра. Найдем полосу прозрачности по общей формуле (121)

$$1 < -4 \frac{Z_b}{Z_a} < \infty.$$

В данном случае

$$Z_a = 2 \left(i\omega L_s + \frac{1}{i\omega C} \right), \quad Z_b = i\omega M,$$

и полоса прозрачности определяется неравенством

$$1 < \frac{2\omega^2 MC}{1 - \omega^2 L_s C} < \infty.$$

Граничные частоты получим, заменив знаки неравенства знаками равенства, что даст два уравнения:

$$1 - \omega^2 L_s C = 0, \quad 2\omega^2 MC = 1 - \omega^2 L_s C.$$

Из первого уравнения находим

$$\omega_1 = \frac{1}{\sqrt{L_s C}} = \frac{1}{\sqrt{LC(1 - k)}} = \frac{\omega_0}{\sqrt{1 - k}},$$

из второго —

$$\omega_2 = \frac{1}{\sqrt{C(L_s + 2M)}} = \frac{1}{\sqrt{LC(1 + k)}} = \frac{\omega_0}{\sqrt{1 + k}}.$$

Таким образом, оказывается, что граничные частоты системы индуктивно связанных контуров, рассматриваемой в качестве полосового фильтра, это не что иное, как частоты связи, найденные в § 33 из совершенно других соображений.

Этот пример показывает, что теория фильтров не противопоставляется теории связанных контуров; она дает более широкую точку зрения на любые применяемые в радиотехнике схемы, обладающие теми или иными избирательными свойствами.

§ 37. Нагрузка фильтра

До сих пор мы с большой легкостью получили структуры и характеристики фильтров с самыми различными свойствами, и, в частности, установили возможность построения полосового фильтра с характеристикой, сколь угодно приближающейся к характеристике идеальной избирательной системы.

Но теперь нужно вспомнить, что исключительная простота теории, которой мы пользовались в предыдущем параграфе, есть результат определенных ограничивающих условий, наложенных на свойства и режим работы рассматриваемых систем. Важнейшим из этих условий является то, что каждое звено и вся цепочка в целом должны быть нагружены на характеристическое сопротивление. Посмотрим, выполнимо ли это требование в действительных условиях.

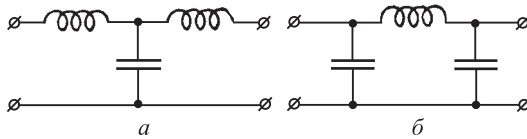


Рис. 114

Будем рассматривать фильтр нижних частот по схеме рис. 103. Найдем характеристическое сопротивление Т-звена (рис. 114, а) и П-звена (рис. 114, б). Для Т-звена (см. § 36)

$$(Z_{10})_T = \frac{1}{2} Z_a \frac{1 + 4 \frac{Z_b}{Z_a}}{1 + 2 \frac{Z_b}{Z_a}}, \quad (Z_{1\infty})_T = \frac{1}{2} Z_a \left(1 + 2 \frac{Z_b}{Z_a} \right),$$

$$(Z_c)_T = \sqrt{Z_{10} Z_{1\infty}} = \frac{1}{2} Z_a \sqrt{1 + 4 \frac{Z_b}{Z_a}}.$$

Подставляя

$$Z_a = i \omega L, \quad Z_b = \frac{1}{i \omega C},$$

получим

$$(Z_c)_T = \frac{1}{2} i \omega L \sqrt{1 - \frac{\omega_k^2}{\omega^2}} = \sqrt{\frac{L}{C}} \sqrt{1 - \frac{\omega^2}{\omega_k^2}}. \quad (123)$$

Действуя аналогичным образом, найдем характеристическое сопротивление для П-звена

$$(Z_c)_\Pi = \frac{2}{i\omega C} \frac{1}{\sqrt{1 - \frac{\omega_k^2}{\omega^2}}} = \sqrt{\frac{L}{C}} \frac{1}{\sqrt{1 - \frac{\omega_k^2}{\omega^2}}}. \quad (124)$$

Из формул (123) и (124) видно, что характеристическое сопротивление обладает следующими свойствами:

- 1) зависит от частоты;
- 2) активно в полосе прозрачности ($\omega < \omega_k$) и реактивно в полосе непрозрачности ($\omega > \omega_k$). Отсюда уже сразу следует, что подобрать для фильтра сопротивление нагрузки с такими свойствами весьма затруднительно. Обычно нагрузкой фильтра (по условиям его использования) является активное сопротивление. Но величина применяемых сопротивлений не зависит (или слабо зависит) от частоты. Поэтому ясно, что при нагрузке фильтра на обыкновенное активное сопротивление основное условие $Z_2 = Z_c$ (так называемое условие согласования) не выполняется, и перед нами встают два вопроса.

1. Как все-таки выбрать $Z_2 = R$, чтобы наименьшим образом уклониться от идеального условия согласования $Z_2 = Z_c$?
2. Каковы последствия неизбежного нарушения условия $Z_2 = Z_c$?

Чтобы разобраться в этом, нужно рассмотреть зависимости (123) и (124).

Как видим, $(Z_c)_\Gamma$ на граничной частоте обращается в нуль, а $(Z_c)_\Pi$ — в бесконечность. При $\omega \rightarrow 0$ оба характеристических сопротивления стремятся к одному и тому же значению

$$\lim_{\omega \rightarrow 0} (Z_c)_\Gamma = \lim_{\omega \rightarrow 0} (Z_c)_\Pi = \frac{1}{2} L \omega_k = \frac{2}{C \omega_k} = \sqrt{\frac{L}{C}}.$$

При $\omega \rightarrow \infty$ из (123) и (124) получаем следующие приближенные выражения:

$$(Z_c)_\Gamma \approx \frac{1}{2} i \omega L, \quad (Z_c)_\Pi \approx \frac{2}{i \omega C}.$$

Эти значения можно получить непосредственно из рассмотрения схем рис. 114.

Графики модулей характеристических сопротивлений показаны на рис. 115, а и б. Заштрихована область, где характеристическое сопротивление реактивно.

Как видим, в большей части полосы прозрачности характеристическое сопротивление для звеньев обоего вида активно и мало отличается от постоянной величины $\sqrt{L/C}$. Это значит, что нагрузив фильтр активным сопротивлением

$$R = \sqrt{\frac{L}{C}}$$

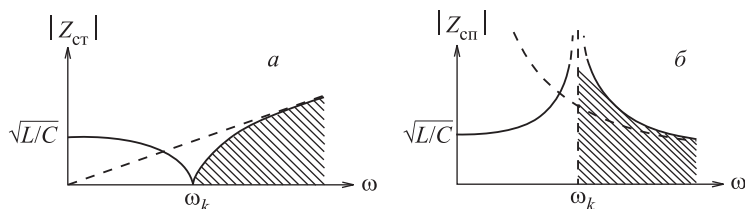


Рис. 115

или близким к этой величине ¹⁾, мы с тем или иным приближением выполним условие согласования $Z_2 = Z_c$ для большей части полосы прозрачности. Это и есть лучшее, что можно сделать при данных обстоятельствах. Согласование ухудшается по мере приближения к граничной частоте. При граничной частоте оно грубо нарушается, так как Z_c обращается в зависимости от типа звена либо в нуль, либо в бесконечность. При нагрузке на сопротивление $Z_2 \pm Z_c$ коэффициент передачи $K = \dot{U}_2/\dot{U}_1$ уже не равен e^{-g} . Значение его можно найти из основных уравнений (110)

$$\left. \begin{aligned} \frac{\dot{U}_1}{\dot{U}_2} &= a_{11} + \frac{1}{Z_2} a_{12} \\ \frac{\dot{I}_1}{\dot{I}_2} &= Z_2 a_{21} + a_{22} \end{aligned} \right\}.$$

Воспользовавшись формулами (115) и (116), получим для отношения напряжений и отношения токов

$$\frac{\dot{U}_1}{\dot{U}_2} = \operatorname{ch} g + \frac{Z_c}{Z_2} \operatorname{sh} g, \quad \frac{\dot{I}_1}{\dot{I}_2} = \operatorname{ch} g + \frac{Z_2}{Z_c} \operatorname{sh} g.$$

Последствия нарушения согласования (или, как говорят, «рассогласования») состоят в том, что характеристика фильтра отклоняется от той идеально простой формы, которую она должна была бы иметь согласно элементарной теории. Так, коэффициент передачи при согласованной нагрузке, который мы обозначали раньше через K_c , равняется по модулю единице на протяжении всей полосы прозрачности, а при $\omega > \omega_k$ так или иначе убывает. Действительный же коэффициент передачи K по мере приближения к границе, т. е. по мере ухудшения согласования, претерпевает колебание, становясь по модулю то меньше единицы, то больше единицы (за счет резонансных явлений).

¹⁾ Можно получить лучшее согласование, если нагрузить фильтр не на $R = \sqrt{L/C}$, т. е. на сопротивление, равное значению характеристического сопротивления при нулевой частоте, а на несколько меньшее сопротивление в случае Т-звена и на несколько большее в случае П-звена. При этом точное согласование будет достигаться не при $\omega \rightarrow 0$, а при некоторой частоте, лежащей ближе к граничной.

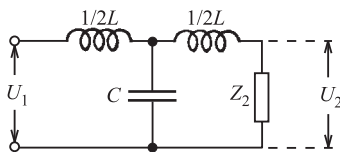


Рис. 116

Пример, приведенный ниже, поможет вникнуть в получаемые соотношения. Возьмем одно Т-звено фильтра нижних частот и нагрузим его произвольным сопротивлением Z_2 (рис. 116). Для коэффициента передачи, не прибегая ни к какой общей теории, нетрудно найти общее выражение

$$K = \frac{1}{1 - 2\frac{\omega^2}{\omega_k^2} + i\frac{\omega L}{Z_2} \left(1 - \frac{\omega^2}{\omega_k^2}\right)}, \quad (125)$$

где

$$\omega_k = \frac{2}{\sqrt{LC}}$$

Пусть нагрузочное сопротивление есть активное сопротивление

$$Z_2 = R = n\sqrt{L/C}.$$

Подставляя это значение в (125) и обозначая для краткости

$$\frac{\omega^2}{\omega_k^2} = x,$$

получим для модуля коэффициента передачи

$$|K| = \frac{1}{\sqrt{(1 - 2x)^2 + \frac{4}{n^2}x(1 - x)^2}}. \quad (126)$$

Частотные характеристики модуля коэффициента передачи по формуле (126) построены на рис. 117 для $n = 1/2$, 1 и 2. Если же взять согласованную нагрузку

$$Z_2 = Z_c \sqrt{\frac{L}{C}} \sqrt{1 - x},$$

то в полосе прозрачности получим, как и должно быть,

$$K_c = 1,$$

а в полосе непрозрачности

$$|K_c| = \frac{1}{2x - 1 + 2\sqrt{x(x - 1)}}.$$

Эта зависимость нанесена на рис. 117 пунктирной линией. Как видим, характеристика при нагрузке на $Z_2 = R$ сильно отличается от

характеристики при согласованной нагрузке $Z_2 = Z_c$ даже при нагрузке на сопротивление, близкое к $R = \sqrt{L/C}$.

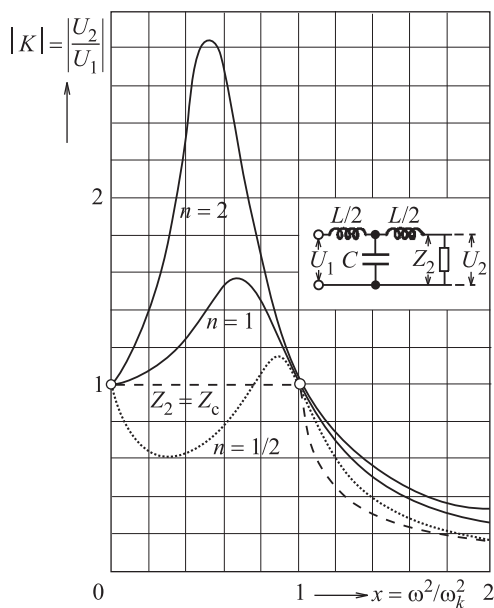


Рис. 117

Влияние неизбежного рассогласования крайне неприятно в двух отношениях. С практической точки зрения характеристики фильтров получаются хуже, чем те, которые могли бы быть получены при согласованной нагрузке. С теоретической же точки зрения невыполнение условия согласования заметно усложняет теорию, утрачивающую компактность и прозрачность, являющиеся главнейшими достоинствами теории в ее элементарной форме. К этому можно добавить, что при расчете реального фильтра следует учитывать и активные сопротивления звеньев, чего мы до сих пор не делали. При этом теоретические соотношения усложняются настолько, что практический расчет фильтров ведется обычно по заранее заготовленным графикам, номограммам и таблицам. Все эти обстоятельства не лишают, однако, элементарную теорию значения; ее ценность состоит в том, что она позволяет очень легко получить общую ориентировку в вопросах, относящихся к устройству и действию фильтров.

Не нужно думать, что из-за рассогласования характеристики фильтров портятся безнадежным образом. Применяя комбинированные фильтры (т.е. фильтры, составленные из неодинаковых звеньев) и несколько усложненные схемы отдельных звеньев, а также прибегая в случае необходимости к электромеханическим фильтрам (об этих

возможностях коротко говорится в последующих параграфах), удается рассчитать и построить фильтры, удовлетворяющие техническим требованиям.

§ 38. Фильтры типа М

Оставаясь в рамках элементарной теории, можно построить более совершенные фильтры ценой некоторого усложнения схем. В частности, на свойствах фильтра благоприятно отражается перераспределение реактивных сопротивлений в последовательных и параллельных плечах лестничной схемы.

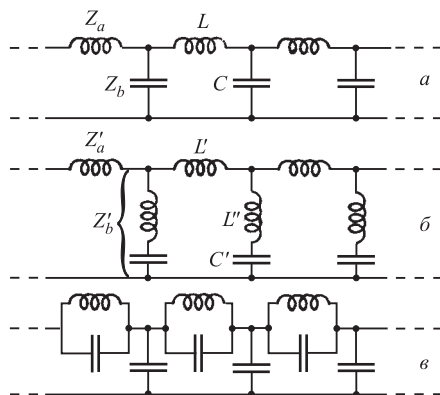


Рис. 118

Возьмем к примеру простейшую схему фильтра нижних частот (рис. 118, а) и перенесем часть индуктивностей в параллельные плечи (рис. 118, б) или, наоборот, часть емкостей в последовательные плечи (рис. 118, в). Рассмотрим схему рис. 118, б. Для расчета ее параметров будем исходить из условия, что характеристическое сопротивление звеньев определенного типа для схем рис. 118, а и б должно быть одинаковым. Пусть это условие относится к звеньям типа Т. Тогда, приравняв характеристические сопротивления

$$(Z_c)_T = (Z'_c)_T \quad (127)$$

(штрихом отмечены величины, относящиеся к конструируемой схеме рис. 118, б), получим

$$Z_a^2 \left(1 + 4 \frac{Z_b}{Z_a} \right) = Z_a'^2 \left(1 + 4 \frac{Z'_b}{Z'_a} \right). \quad (128)$$

Пусть

$$Z'_a = M Z_a, \quad (129)$$

где $0 < M < 1$, т. е. предположим, что в последовательном плече новой схемы оставлена M -я доля последовательного сопротивления Z_a исходной схемы. При этом получим из (128)

$$1 + 4 \frac{Z_b}{Z_a} = M^2 + 4M \frac{Z'_b}{Z_a},$$

откуда

$$Z'_b = \frac{1 - M^2}{4M} Z_a + \frac{1}{M} Z_b. \quad (130)$$

Если

$$Z_a = i\omega L, \quad Z_b = \frac{1}{i\omega C},$$

то по формулам (129) и (130)

$$Z'_a = i\omega ML = i\omega L'; \quad Z'_b = i\omega \frac{1 - M^2}{4M} L + \frac{1}{i\omega MC} = i\omega L'' + \frac{1}{i\omega C'}.$$

Таким образом, все параметры схемы рис. 118, б определены.

Заметим, что, положив $M = 1$, мы возвращаемся к исходной схеме рис. 118, а, так что эту схему можно рассматривать как предельный случай более общей схемы рис. 118, б при $M \rightarrow 1$. Фильтры, полученные из простейших описанным способом, называются *фильтрами типа М*. Простейшие фильтры ($M = 1$) называются фильтрами типа К.

Построим теперь для схемы рис. 118, б П-звено (рис. 119). Характеристическое сопротивление такого звена равно

$$(Z'_c)_\Pi = 2Z'_b \frac{1}{\sqrt{1 + 4 \frac{Z'_b}{Z'_a}}} = 2i\omega ML'' \frac{1 - (1 - M^2) \frac{\omega_k^2}{\omega^2}}{\sqrt{1 - \frac{\omega_k^2}{\omega^2}}}.$$

Как видим, характеристическое сопротивление П-звена зависит от M ($(Z'_c)_\Pi$ не зависит от M , так как по условию (127) оно равно $(Z_c)_\Pi$). На рис. 120 представлены графики зависимости характеристического

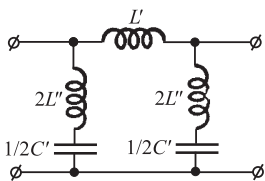


Рис. 119

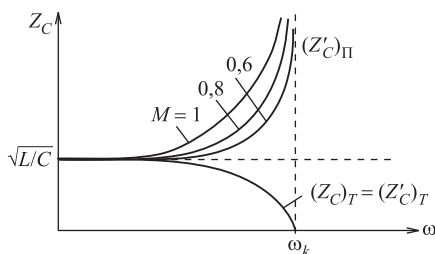


Рис. 120

сопротивления от частоты при различных значениях M . В этих зависимостях существенно то, что при $M < 1$ значения характеристического сопротивления в полосе прозрачности меньше уклоняются от $\sqrt{L/C}$,

чем при $M = 1$. Следовательно, выбирая $M < 1$, можно получить лучшее согласование, чем при $M = 1$. В этом и заключается одно из достоинств фильтров типа M .

При расчете параметров фильтра типа M мы исходили из условия равенства характеристических сопротивлений Т-звеньев и получили семейство зависящих от M характеристик для П-звена. Можно поступить и наоборот: взять в качестве исходного условия

$$(Z'_c)_\Pi = (Z_c)_\Pi.$$

В этом случае получим различные характеристические сопротивления для Т-звена, как показано на рис. 121.

Обратимся к частотным характеристикам затухания. Достаточно беглого взгляда на схемы рис. 118, чтобы отметить существенное раз-

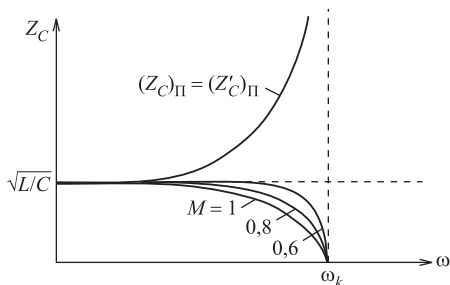


Рис. 121

личие между ними, а именно: в схеме фильтра типа M (рис. 118, б) должно получиться бесконечное значение затухания при частоте, которая является резонансной для плеча $L''C'$, т. е. при

$$\omega' = \frac{2}{\sqrt{L''C'}},$$

так как при резонансе параллельные плечи обладают нулевым сопротивлением и замы-

кают фильтр накоротко. Граничная частота фильтра типа M остается неизменной при изменении M ; она равна

$$\omega_k = \frac{2}{\sqrt{LC}}.$$

Выразим ω' через те же параметры

$$\omega' = \frac{1}{\sqrt{L''C'}} = \frac{2}{\sqrt{(1 - M^2)LC}}.$$

Таким образом,

$$\frac{\omega'}{\omega_k} = \frac{1}{\sqrt{1 - M^2}} > 1,$$

т. е. частота ω' лежит в полосе непрозрачности, и тем ближе к граничной частоте ω_k , чем меньше M . Можно, следовательно, уменьшая M , приблизить резонансную частоту ω' к граничной частоте ω_k и получить, таким образом, более крутой ход характеристики затухания около границы. В этом еще одно достоинство фильтра типа M . Примерный вид характеристики затухания показан на рис. 122.

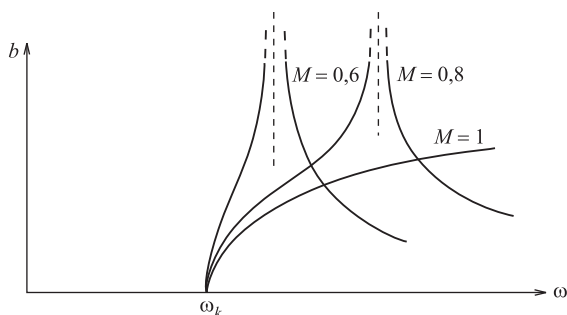


Рис. 122

Вместе с тем характеристика затухания фильтра типа M имеет недостаток: после резонанса на частоте ω' затухание убывает, тогда как при $M = 1$ затухание неуклонно возрастает с повышением частоты, как это показано на рис. 122. Это обстоятельство наводит на мысль построить комбинированный фильтр из различных звеньев, т. е. из звеньев с различным M , включая и $M = 1$. Действительно, такого рода фильтр обладает более благоприятными характеристиками. Так, например, если составить фильтр из трех звеньев с $M = 1$; 0,8 и 0,6, то характеристика затухания такого фильтра (кривая a рис. 123) в широкой области частот, примыкающей к границе, лежит выше характеристики трехзвенного фильтра из одинаковых звеньев с $M = 1$ (кривая $б$).

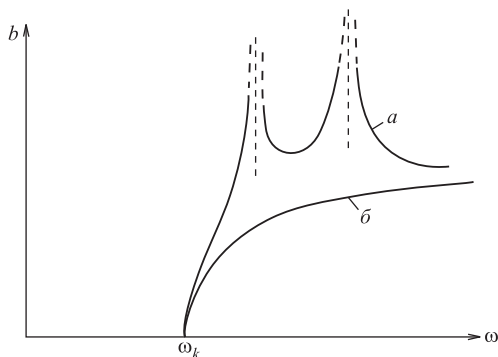


Рис. 123

Мы разбирали в этом параграфе только фильтры нижних частот, но аналогичные соображения относятся и к фильтрам других назначений.

Для того чтобы показать, насколько хороши могут быть характеристики реальных фильтров, на рис. 124 показана схема и характеристика фильтра нижних частот, составленного из нескольких несимметричных M -звеньев. Штрих-пунктирная ломаная линия — заданное значение

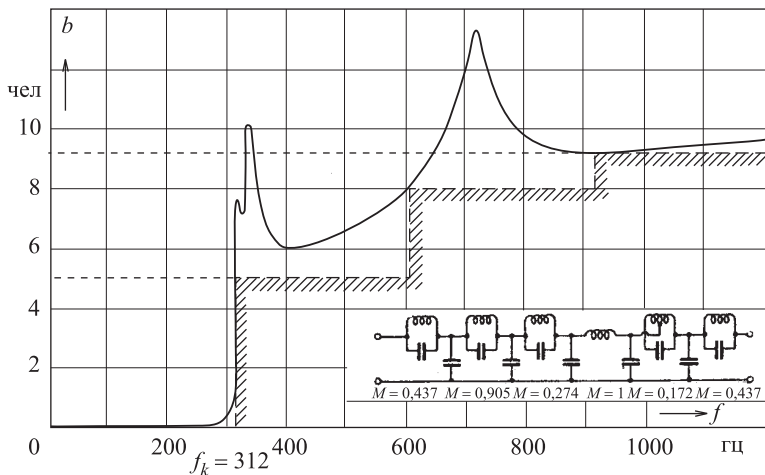


Рис. 124

затухания, послужившее основой расчета. Затухание в полосе прозрачности весьма незначительно, и изменения его неощутимы, кроме как в непосредственной близости от граничной частоты. Экспериментально снятая характеристика в масштабе рисунка неотличима от расчетной.

§ 39. Линия задержки

По основному своему назначению фильтры работают в установившемся режиме, а важнейшие свойства фильтра выражаются непосредственно частотными характеристиками, т. е. зависимостями от частоты коэффициента передачи или затухания. Однако схемы точно такого же вида, как приведенные выше схемы простейших фильтров, используются иногда для совершенно других целей и работают при этом исключительно в нестационарных режимах.

Примером использования схемы типа фильтра в нестационарном режиме является так называемая *линия задержки*. Схема линии задержки ничем не отличается от схемы фильтра нижних частот, но назначение ее состоит в том, чтобы задержать поданное на вход возмущение на некоторое определенное время. При этом требуется обычно, чтобы форма возмущения не подверглась существенному искажению. Таким образом, действие идеальной линии задержки должно было бы описываться соотношением

$$y(t) = x(t - \tau),$$

где x и y , как обычно, воздействие и отклик, τ — время задержки. Само собой разумеется, что выполнить это условие в точности в реальной

системе не удастся; задача состоит в выяснении механизма задержки и всех основных соотношений ¹⁾.

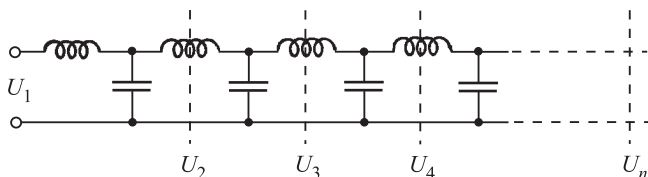


Рис. 125

Нужно найти напряжение как функцию времени на выходе n -го звена схемы рис. 125 при определенном воздействии на вход. В качестве воздействия выберем единичный импульс

$$u_1(t) = \delta(t),$$

так что выходное напряжение $U_n(t)$ будет представлять собой не что иное, как импульсную реакцию n -звенной цепочки. В идеальном случае мы должны были бы получить

$$u_n(t) = \delta(t - \tau_n).$$

В действительности же получается более или менее расплывшийся импульс, так что о запаздывании приходится судить по положению максимума этого импульса.

При решении задачи будем предполагать, что цепочка нагружена на характеристическое сопротивление. Это дает право воспользоваться соотношением

$$\bar{u}_n = \bar{u}_1 e^{-ng}. \quad (131)$$

В этой записи g есть постоянная передачи в операторной форме; развернутое выражение для g мы сейчас выпишем. Мы имеем вообще (§ 35)

$$e^g = \operatorname{ch} g + \operatorname{sh} g = a_{11} + \sqrt{a_{11}^2 - 1}. \quad (132)$$

¹⁾ Имеется в виду задержка на относительно небольшое время τ . Если речь идет о длительной задержке, то широко распространенным методом является запись данного электрического возмущения с последующим воспроизведением. Для этой цели очень хороша магнитная запись. Если же речь идет об очень кратковременных явлениях (или об очень высоких частотах, что одно и то же), то может быть применена запись в форме потенциального рельефа на запоминающих электронных трубках. В этих трубках запись и воспроизведение производятся электронным лучом, и, следовательно, скорость обоих этих процессов может быть очень велика. Таким образом, то, что будет сказано в этом параграфе, отнюдь не исчерпывает техники задерживания, имеющей большое значение; обсуждается лишь одна возможность осуществления задержки, связанная с нестационарными режимами в системах с сосредоточенными параметрами.

Для Т-звена схемы рис. 125

$$a_{11} = \frac{\overline{U}_1}{\overline{U}_2} \bigg|_{\overline{I}_2=0} = 1 + \frac{1}{2} p^2 LC = 1 + 2 \frac{p^2}{\omega_k^2}.$$

Подставляя в (132), получим

$$e^g = 1 + 2 \frac{p^2}{\omega_k^2} + \sqrt{\left(1 + 2 \frac{p^2}{\omega_k^2}\right)^2 - 1}.$$

У нас

$$u_1 = \delta(t), \quad \overline{u}_1 = p$$

и, подставляя все это в (131), получаем изображение решения

$$\overline{u}_n = \frac{p}{\left(1 + 2 \frac{p^2}{\omega_k^2} + \sqrt{\left(1 + 2 \frac{p^2}{\omega_k^2}\right)^2 - 1}\right)^n}.$$

Остается перейти от этого изображения к оригиналу, и задача будет решена. Но возникает чисто техническое затруднение: в справочниках по операционному исчислению не находится подходящей готовой формулы. Искать же оригинал какими-либо общими методами затруднительно. Поэтому постараемся получить изображение решения в другой форме, воспользовавшись одним чисто алгебраическим преобразованием. А именно, применим к выражению

$$e^g = a_{11} + \sqrt{a_{11}^2 - 1}$$

следующее преобразование:

$$a + \sqrt{b} = (\sqrt{c} + \sqrt{d})^2, \quad (133)$$

где

$$c = \frac{1}{2}(a + \sqrt{a^2 - b}),$$

$$d = \frac{1}{2}(a - \sqrt{a^2 - b}).$$

В справедливости равенства (133) можно убедиться подстановкой в него значений c и d . У нас

$$a = a_{11}, \quad b = a_{11}^2 - 1.$$

Подставляя в формулы для c и d , находим

$$c = \frac{1}{2}(a_{11} + 1), \quad d = \frac{1}{2}(a_{11} - 1).$$

Теперь можно записать на основании (133)

$$a_{11} + \sqrt{a_{11}^2 - 1} = \left(\sqrt{\frac{1}{2}(a_{11} + 1)} + \sqrt{\frac{1}{2}(a_{11} - 1)} \right)^2 = \left(\sqrt{1 + \frac{p^2}{\omega_k^2}} + \frac{p}{\omega_k} \right)^2$$

и, наконец,

$$\bar{u}_n = \frac{p}{\left(\sqrt{1 + \frac{p^2}{\omega_k^2}} + \frac{p}{\omega_k} \right)^{2n}}.$$

Введем

$$p_1 = \frac{p}{\omega_k}$$

(нужно помнить, что это соответствует переходу к новому масштабу времени в оригинале, оригинал будет выражен через $t_1 = \omega_k t$); тогда

$$\bar{u}_n = \omega_k \frac{p_1}{\left(\sqrt{1 + p_1^2} + p_1 \right)^{2n}}.$$

Для этого изображения в справочниках можно найти оригинал ¹⁾, а именно

$$\bar{u}_n \rightarrow u_n(t_1) = \omega_k 2n \frac{J_{2n}(t_1)}{t_1},$$

или

$$u_n(t) = 2n \frac{J_{2n}(\omega_k t)}{t} \quad (t > 0), \quad (134)$$

где $J_{2n}(t_1)$ означает функцию Бесселя действительного аргумента t_1 порядка $2n$. Соотношение (134) и представляет собой решение задачи.

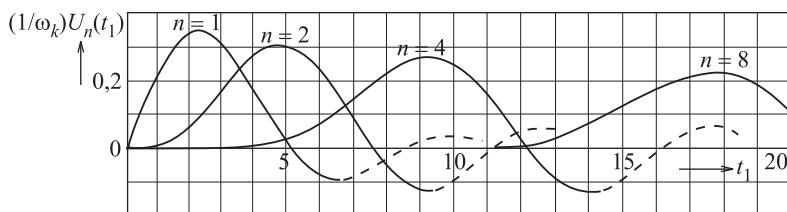


Рис. 126

На рис. 126 построены графики функции $\frac{1}{\omega_k} u_n(t_1)$ для различных значений n . Эти графики представляют изменения напряжения на выходе n -то звена при подаче на вход единичного импульса. Как видим, чем длиннее цепочка, тем позднее развивается напряжение на выходе. Можно считать за момент появления импульса на выходе тот момент, в который выходное напряжение достигает первого максимума. Если

¹⁾ См., например, *Теумин И. И.* Справочник по переходным электрическим процессам. — Связьиздат, 1951; *Диткин В. А., Кузнецов П. И.* Справочник по операционному исчислению. — Гостехиздат, 1951.

обозначить этот момент через τ_n , то на основании рисунка можно установить, что в первом грубом приближении

$$\omega_k \tau_n \approx 2n,$$

или

$$\tau_n \approx n \frac{2}{\omega_k}. \quad (135)$$

Таким образом, время задержки τ_n пропорционально числу звеньев, и дело обстоит так, как если бы ¹⁾ поданный на вход импульс бежал по цепочке с конечной скоростью, равной $\omega_k/2$ *звеньев в секунду*. Величина $2/\omega_k = \sqrt{LC}$ выражает задержку, отнесенную к одному звену. Следовательно, чтобы получить задержку на заданное время τ , нужно при данной граничной частоте ω_k взять число звеньев, равное

$$n = \frac{1}{2} \omega_k \tau.$$

Но мы ничего не говорили да сих пор о выборе ω_k , а между тем эта величина определяет степень расплывания выходного импульса. Для того чтобы получить хотя бы качественное представление о влиянии граничной частоты ω_k , поставим вопрос так: пусть задана задержка τ и пусть *одна и та же* задержка получается в цепочках с *разным* числом звеньев, а следовательно, и с разным значением ω_k . Итак, пусть

$$\tau = 2 \frac{n}{\omega_k} = \text{const.}$$

и число звеньев растет при постоянном τ пропорционально ω_k . Возьмем $n = 1, 2, 4$ и 8 (как на рис. 126) и перестроим график рис. 126 в масштабе $t = t_1/\omega_k$ по оси абсцисс. Получается кривая с тем более сжатым по оси абсцисс масштабом, чем больше ω_k (т.е. чем больше n). Эти кривые построены на рис. 127. Из рисунка видно, что с увеличением n , т.е. с увеличением ω_k , выходной импульс становится все более узким. Задавшись допустимой шириной выходного импульса, определяем требуемые значения ω_k и n .

По поводу соотношений, иллюстрируемых рис. 127, заметим еще, что с увеличением числа звеньев время запаздывания τ стремится к значению $\frac{2n}{\omega_k}$, принятому в качестве первого приближения. Эта тен-

¹⁾ Мы говорим «как если бы» потому, что в действительности в рассматриваемой системе с сосредоточенными постоянными напряжение на выходе появляется в тот же момент, когда напряжение подается на вход. Но вначале напряжение на выходе очень мало и нарастает очень медленно, тем медленнее, чем длиннее цепочка. Поэтому практически дело обстоит так, как если бы импульс на выходе запаздывал относительно импульса на входе. Понятие о конечной скорости распространения электрических возмущений имеет прямой физический смысл только в применении к системам с распределенными параметрами, о которых речь пойдет в следующей главе.

денция обозначена пунктирной линией, проведенной через максимумы выходных импульсов.

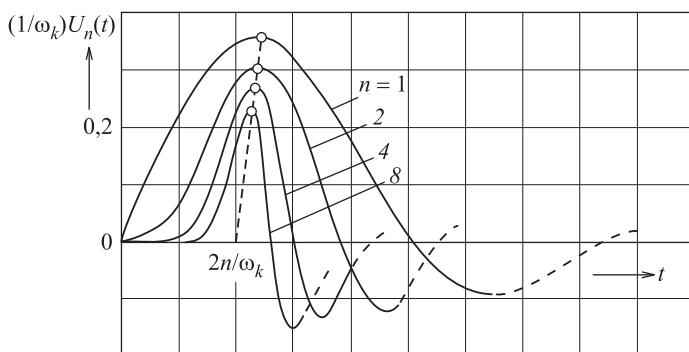


Рис. 127

Рассмотрим теперь частотные характеристики задерживающих схем. Найдем прежде всего характеристику идеальной задерживающей схемы. Как уже говорилось, свойства такой схемы можно выразить соотношением

$$y(t) = x(t - \tau).$$

Это означает, что отклик отличается от произвольного воздействия только запозданием на время τ . Выразим x и y соответствующими интегралами Фурье

$$\int_{-\infty}^{\infty} S_y(\omega) e^{i\omega t} d\omega = \int_{-\infty}^{\infty} S_x(\omega) e^{i\omega(t-\tau)} d\omega.$$

Но

$$S_y = K S_x.$$

Таким образом

$$K = e^{-i\omega\tau},$$

откуда

$$A = |K| = 1; \quad \arg K = -i\omega\tau = -i\varphi,$$

т.е. фазовый сдвиг между откликом и воздействием должен быть прямо пропорционален частоте.

Итак, фазово-частотная характеристика идеальной задерживающей схемы должна представляться прямой, проходящей через начало координат. Наклон этой прямой дает значение времени задержки, которое определяется при этом как

$$\tau = \frac{\varphi}{\omega}.$$

Обратимся к схеме рис. 125. Постоянная передачи для Т-звена этой схемы выражается формулой

$$\text{th } g = \frac{\sqrt{1 - \frac{\omega_k^2}{\omega^2}}}{1 - \frac{1}{2} \frac{\omega_k^2}{\omega^2}}.$$

В полосе прозрачности (т. е. при $\omega \ll \omega_k$) постоянная передачи мнимая, т. е.

$$g = i a$$

и мы имеем

$$\text{tg } a = - \frac{\sqrt{\frac{\omega_k^2}{\omega^2} - 1}}{\frac{1}{2} \frac{\omega_k^2}{\omega^2} - 1}$$

или

$$\text{tg } a = -2 \frac{\omega}{\omega_k} \frac{\sqrt{1 - \frac{\omega^2}{\omega_k^2}}}{1 - 2 \frac{\omega^2}{\omega_k^2}}. \quad (136)$$

При очень низких частотах, т. е. при $\omega \ll \omega_k$, имеем

$$\text{tg } a \approx -2a \frac{\omega}{\omega_k} \approx a$$

и, определяя в этих условиях время задержки как

$$\tau = - \frac{a}{\omega}$$

($-a$ означает фазовый сдвиг, обозначенный выше через φ), находим

$$\tau \approx \frac{2}{\omega_k}.$$

Это — время задержки на одно звено. Но так как постоянная передачи (a в нашем случае — фазовая постоянная) для цепочки из n звеньев в n раз больше, чем для одного звена, то

$$\tau_n = n \frac{2}{\omega_k},$$

что совпадает с (135). Однако следует помнить, что это лишь приближенное выражение. График точной зависимости (136), изображенный на рис. 128, показывает, что действительная фазово-частотная характеристика, рассматриваемая для Т-звена, сильно отклоняется от характеристики идеальной задерживающей схемы. Это расхождение и обуславливает искажение формы импульса, возрастающее по мере удаления от начала цепочки. Легко видеть, что искажение тем меньше, чем выше граничная частота. Форма выходного импульса будет

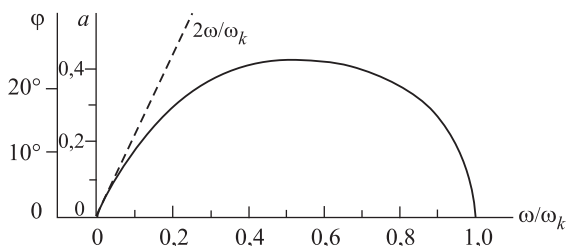


Рис. 128

в пределе приближаться к форме входного импульса (какова бы она ни была) при неограниченном увеличении граничной частоты и одновременно неограниченном увеличении числа звеньев. Подобный предельный переход приводит от фильтра нижних частот к линии без потерь. Этот вопрос будет рассмотрен в § 53.

§ 40. Некоторые схемы с обратной связью

Одно из важнейших в радиотехнике, да и не только в радиотехнике, понятий, вводимое в этом параграфе, — понятие об *обратной связи* ¹⁾. Под обратной связью понимается вообще воздействие выхода системы на ее вход. При наличии обратной связи образуется как бы замкнутое кольцо взаимодействия: прямое направление — вход → выход; обратное направление — выход → вход.

Обратная связь позволяет в широких пределах изменять свойства основной системы. Применение отрицательной (см. ниже) обратной связи в усилителях позволяет уменьшать частотные и нелинейные искажения, уменьшать влияние любых внешних помех, действующих на охваченную обратной связью схему, обеспечивать устойчивость усилителей и повышать стабильность коэффициента усиления. Положительная обратная связь дает возможность увеличить коэффициент усиления. Обратная связь лежит в основе действия электронных стабилизаторов, компрессоров и экспандеров, автоматической регулировки усиления, автоматической подстройки частоты, шумопонижающих устройств. Наконец, обратная связь играет основную роль во всяком генераторе, о чем еще предстоит говорить подробнее в главах, посвященных генерированию колебаний.

В этом параграфе ограничимся рассмотрением основных схем усилителей с обратной связью и выводом выражений для коэффициентов передачи. Вопрос об устойчивости (т.е. о возможности самовозбуж-

¹⁾ Понятие обратной связи — одно из основных в теории автоматического регулирования и управления; в связи с развитием кибернетики оно начинает играть ведущую роль также и в биологии.

дения) рассматривается позднее, так как он имеет прямое отношение к теории генераторов.

В усилителе с обратной связью выход соединен со входом цепью обратной связи. Коэффициент передачи этой цепи принято обозначать через β ; коэффициент передачи самой цепи — через K . Возможно последовательное или параллельное подключение цепи обратной связи ко входу и к выходу усилителя. Основные варианты схемы усилителя с обратной связью показаны на рис. 129.

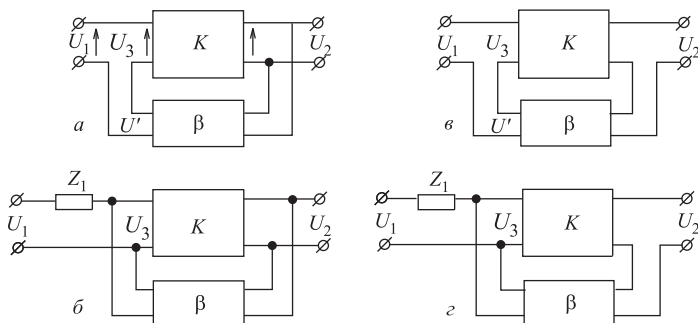


Рис. 129

Условимся

- 1) внутренним сопротивлением источника входного напряжения пренебрегать;
- 2) входное сопротивление усилителя считать бесконечно большим;
- 3) коэффициенты передачи как основной цепи (K), так и цепи обратной связи (β) определять отношениями комплексных амплитуд напряжения на выходе и на входе соответствующих цепей.

В обозначениях рис. 129 имеем:

$$K = \frac{\dot{U}_2}{\dot{U}_3}, \quad \beta = \frac{\dot{U}'}{\dot{U}_2}.$$

В схеме рис. 129, а и б вход цепи обратной связи подключен к выходу усилителя параллельно; в этих схемах осуществляется обратная связь *по напряжению*. В схемах же рис. 129, в и г вход цепи обратной связи включен последовательно по отношению к выходной цепи усилителя. Такого рода связь называется обратной связью *по току*. Ясно, что обратная связь по напряжению перестает действовать при коротком замыкании усилителя, а обратная связь по току — при холостом ходе, т. е. когда отсутствует выходной ток. Заметим, что в схемах рис. 129, а и в напряжение U_3 на входе усилителя равно сумме напряжения источника U_1 и напряжения U' , подаваемого из цепи обратной связи. В схемах же рис. 129, б и г с параллельным подключением выхода цепи обратной связи ко входу усилителя складываются не напряжения, а токи. Для того чтобы найти действующее на входе напряжение, нуж-

но учесть падение напряжения на сопротивлении. Поэтому на схемах рис. 129, б и г введено сопротивление Z_1 . Если бы оно отсутствовало, то выход цепи обратной связи был бы замкнут накоротко (так как по условию источник имеет нулевое сопротивление). Сопротивление Z_1 можно было бы ввести на тех же местах и в схемы рис. 129, а и в, но там оно не играло бы никакой роли, будучи включено последовательно с бесконечно большим входным сопротивлением усилителя.

Найдем коэффициент передачи усилителя с обратной связью по схеме рис. 129, а, т. е. величину

$$K_\beta = \frac{\dot{U}_2}{\dot{U}_1}.$$

Мы имеем следующие соотношения:

$$\frac{\dot{U}_2}{\dot{U}_3} = K, \quad \dot{U}_3 = \dot{U}_1 + \beta \dot{U}_2.$$

Исключая $\dot{U}_3$, находим

$$\dot{U}_2 = K(\dot{U}_1 + \beta \dot{U}_2)$$

или

$$K_\beta = \frac{\dot{U}_2}{\dot{U}_1} = \frac{K}{1 - K\beta}. \quad (137)$$

Итак, обратная связь изменяет коэффициент передачи в $\frac{1}{1 - K\beta}$ раз. Обратную связь принято называть отрицательной, если она уменьшает коэффициент передачи, и положительной, если она его увеличивает. Напомним, что и K и β и их произведение — комплексные величины. Поэтому характер обратной связи зависит от аргумента (т. е. от фазовой постоянной) произведения $K\beta$, которое можно представить в виде

$$K\beta = |K\beta| e^{i\varphi}.$$

В простейшем случае можно считать, что β — действительная величина и аргумент произведения $K\beta$ зависит только от схемы и числа ступеней усилителя. Известно, что в обычной схеме усилителя с заземленным (общим) катодом каждая ступень поворачивает фазу переменного напряжения на π , что соответствует изменению полярности напряжения.

Следовательно, усилитель с нечетным числом ступеней имеет отрицательный коэффициент передачи (так как $e^{i\pi} = -1$), а при четном числе ступеней коэффициент передачи положителен. Итак, при четном числе ступеней усиления получается отрицательная обратная связь, а при нечетном — положительная.

Вернемся к схемам рис. 129, а и б и рассмотрим два часто применяемых варианта, показанных на рис. 130.

В этих схемах цепь обратной связи вырождается в простой делитель напряжения: к выходу усилителя подключено последовательное

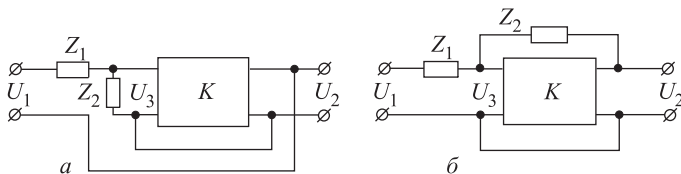


Рис. 130

соединение сопротивлений Z_1 и Z_2 ; напряжение обратной связи снимается с одного из них. Коэффициенты передачи цепи обратной связи для схем рис. 130, *а* и *б* равны соответственно:

$$\beta_1 = \frac{Z_2}{Z_1 + Z_2}, \quad \beta_2 = \frac{Z_1}{Z_1 + Z_2}.$$

Однако эти выражения нельзя непосредственно подставлять в общую формулу (137), так как она получена для схемы рис. 129, *а*, в которой входной ток отсутствует. В схемах рис. 130 это условие не выполняется и придется проделать для них вывод заново.

Для схемы рис. 130, *а* имеем следующие соотношения:

$$\begin{aligned} \dot{I}(Z_1 + Z_2) - \dot{U}_2 &= \dot{U}_1, \\ \dot{U}_2 &= K\dot{U}_3, \quad \dot{U}_3 = \dot{I}Z_2. \end{aligned}$$

Направления напряжений выбраны так, что коэффициент передачи K определен как положительная величина. Исключая из трех приведенных уравнений $\dot{I}$ и $\dot{U}_3$ находим

$$\left[-1 + \frac{1}{K} \left(1 + \frac{Z_1}{Z_2} \right) \right] \dot{U}_2 = \dot{U}_1$$

или

$$K_\beta = \frac{\dot{U}_2}{\dot{U}_1} = \frac{K\beta_1}{1 - K\beta_1}.$$

Для схемы рис. 130, *б* имеем:

$$\begin{aligned} \dot{I}(Z_1 + Z_2) + \dot{U}_2 &= \dot{U}_1, \\ \dot{U}_2 &= K\dot{U}_3, \quad \dot{U}_3 = \dot{U}_1 - \dot{I}Z_1, \end{aligned}$$

откуда

$$\left(1 + \frac{Z_2}{Z_1} \right) \left(\dot{U}_1 - \frac{1}{K} \dot{U}_2 \right) + \dot{U}_2 = \dot{U}_1$$

или

$$K_\beta = \frac{\dot{U}_2}{\dot{U}_1} = \frac{K(1 - \beta_2)}{1 - K\beta_2}.$$

В схеме рис. 130, *а* при отрицательном K коэффициент передачи K_β не может превзойти единицы по абсолютной величине. Примером может служить схема усилительной ступени, известной под назва-

нием катодного повторителя (рис. 131). Для этой схемы (сравните с рис. 130, а):

$$Z_1 = 0, \quad \beta_1 = 1, \quad K = -\mu \frac{R_k}{R_i + R_k} \approx -\mu, \quad K\beta \approx \frac{-\mu}{1 + \mu}.$$

Схема катодного повторителя практически используется как схема с заземленным анодом (рис. 132); при этом коэффициент передачи становится положительным (см. стрелки, показывающие направления напряжений на схемах рис. 131 и 132).

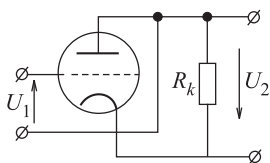


Рис. 131

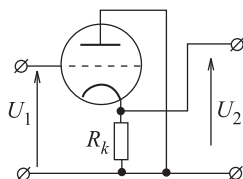


Рис. 132

Схемы с обратной связью могут применяться и в качестве двухполюсников. На рис. 133 показана одна из таких схем. Ее проводимость может быть любой в зависимости от свойств делителя на входе.

Мы имеем:

$$\dot{U}_1 = \dot{U} \frac{Z_1}{Z_1 + Z_2} = \beta \dot{U}, \quad \dot{U}_2 = K \dot{U}_1, \\ \dot{I} = \frac{1}{R_i} (\dot{U} - \dot{U}_2) + \frac{\dot{U}}{Z_1 + Z_2}.$$

Пусть

$$Z_1 + Z_2 \gg R_i.$$

Тогда

$$\dot{I} = \frac{\dot{U}}{R_i} (1 - K\beta)$$

и входная проводимость

$$Y = \frac{\dot{I}}{\dot{U}} = \frac{1}{R_i} (1 - K\beta).$$

Рассмотрим в качестве примера схему рис. 134. Здесь:

$$K = -\mu, \quad Z_1 = i\omega L, \quad Z_2 = \frac{1}{i\omega C}, \quad \frac{\mu}{R_i} = S, \quad \beta = \frac{1}{1 - \frac{\omega_0^2}{\omega^2}}$$

и получаем

$$Y = \frac{1}{R_i} + \frac{S}{1 - \frac{\omega_0^2}{\omega^2}},$$

т. е. входная проводимость за счет обратной связи может быть сделана отрицательной (если $\omega < \omega_0$). Этот результат будет использован в тео-

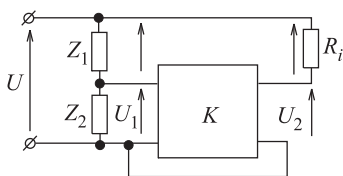


Рис. 133

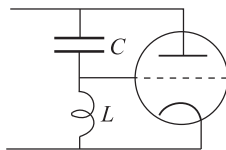


Рис. 134

рии генераторов, Другим применением двухполюсной схемы с обратной связью является так называемая *реактивная лампа* — устройство, обладающее управляемой реактивной проводимостью.

§ 41. Дифференцирующие и интегрирующие схемы

В радиотехнике очень часто возникает надобность в построении четырехполюсников, отклик которых пропорционален производной или интегралу от воздействия. Четырехполюсники с такими свойствами называются соответственно *дифференцирующими* или *интегрирующими* схемами.

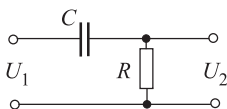


Рис. 135

В наиболее общем случае отклик и воздействие связаны некоторым интегро-дифференциальным выражением, в котором содержатся и производные и интегралы.

Речь идет о построении некоторой вырожденной схемы, в которой требуемые члены играют решающую роль, а влиянием остальных можно с тем или иным приближением пренебречь. Из этих соображений уже следует, что задача дифференцирования или интегрирования решается при помощи электрических схем лишь приближенно, и наша задача состоит в выяснении условий, при которых приближение получается удовлетворительным.

Рассмотрим схему рис. 135. Дифференциальное уравнение, связывающее напряжение на входе и на выходе, имеет вид

$$\frac{du_2}{dt} + \frac{1}{RC}u_2 = \frac{du_1}{dt}. \quad (138)$$

Если

$$\frac{du_2}{dt} \ll \frac{1}{RC}u_2, \quad (139)$$

то приближенно

$$u_2 \approx RC \frac{du_1}{dt}, \quad (140)$$

т.е. схема с тем или иным приближением выполняет дифференцирование.

Нужно сразу заметить, что выполнение условия (139) зависит не только от соотношения параметров схемы, но и от характера функции u_2 , а следовательно, и u_1 . Иначе говоря, одна и та же схема будет удовлетворительно дифференцировать одни функции и не будет дифференцировать другие. Если условие (139) не выполнено, и если имеется противоположное соотношение

$$\frac{du_2}{dt} \gg \frac{1}{RC} u_2, \quad (141)$$

то из (138) получаем

$$u_2 \approx u_1. \quad (142)$$

Заметим, что переход от (139) к (141), а следовательно, и от (140) к (142) происходит непрерывно. Формулы (140) и (142) соответствуют двум крайним случаям. В общем же случае схема выполняет не простое дифференцирование, а более сложное преобразование, описываемое уравнением (138).

Условие (139) характеризует так называемую «медленную» (т. е. медленно изменяющуюся) функцию, условие (140) — «быструю» функцию. Таким образом, можно сказать, что схема рис. 135 дифференцирует медленные функции и не дифференцирует быстрые. Конечно, при прочих равных условиях дифференцирование выполняется тем успешнее, чем *меньше* постоянная времени RC . По этому поводу нужно сразу заметить, что (как это видно из (140)), чем лучше выполняется дифференцирование, тем меньше полезный эффект на выходе схемы.

О быстроте или медленности функций легко судить по их спектральному составу. Поэтому мы переведем описание свойств дифференцирующей схемы рис. 135 на спектральный язык.

Если

$$u = \dot{U} e^{i\omega t},$$

то

$$\frac{du}{dt} = i\omega u,$$

т. е. при дифференцировании происходит умножение на $i\omega$. Отсюда следует, что для идеальной дифференцирующей схемы должно бы быть

$$\dot{U}_2 = i a \omega \dot{U}_1,$$

т. е. коэффициент передачи идеальной дифференцирующей схемы должен иметь вид

$$K_{\partial} = \frac{\dot{U}_2}{\dot{U}_1} = i a \omega,$$

где a — произвольный постоянный множитель.

Составим выражение для коэффициента передачи схемы рис. 135

$$K = \frac{1}{1 + \frac{1}{i\omega RC}}. \quad (143)$$

Если

$$\omega RC \ll 1, \quad (144)$$

то с тем или иным приближением можем записать вместо (143)

$$K \approx i\omega RC = K_{\partial}.$$

Теперь условие хорошего дифференцирования выражается формулой (144) вместо (139). Итак, для синусоидального колебания с частотой ω дифференцирование осуществляется при условии, что частота этого колебания много меньше величины $\frac{1}{RC}$. Чем ниже частота, тем успешнее осуществляется дифференцирование. Теперь легко перейти от синусоидального колебания к любой функции с ограниченным спектром и высказать следующее положение: данная функция будет дифференцироваться схемой рис. 135, если наивысшая частота в спектре функции много меньше $\frac{1}{RC}$.

Эти соотношения поясняются также частотной характеристикой коэффициента передачи. Для схемы рис. 135 имеем

$$|K| = \frac{1}{\sqrt{1 + \frac{1}{\omega^2 R^2 C^2}}}$$

(см. рис. 136). Для идеальной схемы мы имели бы

$$|K_{\partial}| = \omega RC$$

(пунктирная линия на рис. 136). Удовлетворительное дифференцирование происходит в пределах полосы частот от нуля до такого значения, при котором отклонение кривой $|K|$ от прямой $|K_{\partial}|$ еще незначительно.

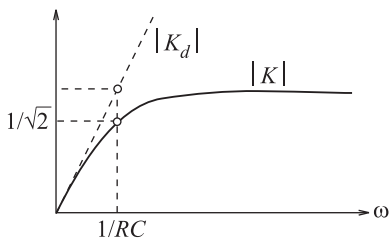


Рис. 136

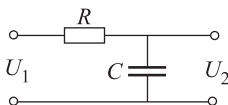


Рис. 137

При $\omega = \frac{1}{RC}$ имеем $|K_{\partial}| = 1$, а $|K| = \frac{1}{\sqrt{2}}$. Это, конечно, очень большое расхождение. Однако уже при $\omega_{\text{макс}} RC \approx 1/2$ можно получить практически удовлетворительное дифференцирование.

Обратимся к схеме рис. 137. Для этой схемы уравнение имеет вид

$$RC \frac{du_2}{dt} + u_2 = u_1.$$

При условии

$$RC \frac{du_2}{dt} \ll u_2$$

имеем

$$u_2 \approx u_1.$$

Если же

$$RC \frac{du_2}{dt} \gg u_2, \quad (145)$$

то приблизительно

$$RC \frac{du_2}{dt} \approx u_1,$$

или

$$u_2 \approx \frac{1}{RC} \int u_1 dt. \quad (146)$$

Таким образом, если выполнено условие (145), то схема рис. 137 интегрирует. Условие (145) определяет быструю функцию. Следовательно, схема рис. 137 хорошо интегрирует быстрые функции и вовсе не интегрирует медленные. При прочих равных условиях интегрирование выполняется тем лучше, чем *больше* постоянная времени RC . Но увеличение постоянной RC уменьшает полезный эффект на выходе интегрирующей схемы (см. (146)).

Рассмотрим свойства интегрирующей схемы с частотной точки зрения. Интегрирование синусоидального колебания равносильно умножению на $\frac{1}{i\omega}$. Следовательно, коэффициент передачи идеального интегрирующего четырехполюсника должен иметь вид

$$K_u = \frac{b}{i\omega}.$$

Для схемы рис. 137 коэффициент передачи

$$K = \frac{1}{1 + i\omega RC}.$$

При условии

$$\omega RC \gg 1$$

можно записать приближенно

$$K \approx \frac{1}{i\omega RC} = K_u.$$

Частотные характеристики коэффициента передачи для схемы рис. 137 изображены на рис. 138. Как видим, при низких частотах схема не интегрирует. Удовлетворительное интегрирование происходит, начиная от частоты $\omega_{\min} \approx \frac{1}{RC}$; качество интегрирования улучшается

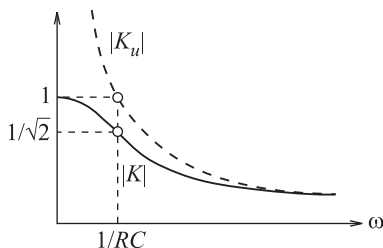


Рис. 138

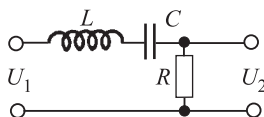


Рис. 139

с повышением частоты. Таким образом, для обеспечения хорошего интегрирования подаваемая на вход интегрирующей схемы функция не должна содержать постоянной составляющей и низких частот.

Как уже говорилось выше, одна и та же схема может вести себя по-разному в зависимости от режима, т. е. от характера воздействующих на нее функций. Для разъяснения этого положения рассмотрим схему рис. 139. Дифференциальное уравнение для нее имеет вид

$$\frac{L}{R} \frac{du_2}{dt} + u_2 + \frac{1}{RC} \int u_2 dt = u_1.$$

Положим, что

$$\left. \frac{L}{R} \frac{du_2}{dt} \right\}_{u_2} \ll \frac{1}{RC} \int u_2 dt.$$

В таком случае

$$\frac{1}{RC} \int u_2 dt \approx u_1$$

или

$$u_2 \approx RC \frac{du_1}{dt},$$

т. е. схема дифференцирует. Пусть теперь

$$\left. \frac{1}{RC} \int u_2 dt \right\}_{u_2} \ll \frac{L}{R} \frac{du_2}{dt}.$$

В этом случае

$$\frac{L}{R} \frac{du_2}{dt} \approx u_1$$

или

$$u_2 \approx \frac{R}{L} \int u_1 dt,$$

т. е. схема интегрирует. На частотном языке эти соотношения описываются при помощи частотной характеристики коэффициента передачи

$$K = \frac{1}{1 + i \left(\omega \frac{L}{R} - \frac{1}{\omega RC} \right)}.$$

При $\omega \rightarrow 0$

$$K \approx i\omega RC = K_{\partial}.$$

При $\omega \rightarrow \infty$

$$K \approx \frac{R}{i\omega L} = K_u.$$

Эти соотношения представлены на графике рис. 140.

Свойства дифференцирующих и интегрирующих схем можно, конечно, представлять не частотными, а временными характеристиками,

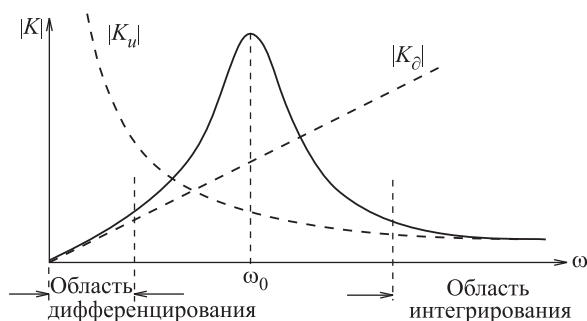


Рис. 140

например переходными характеристиками. Но это оказывается в данном случае менее наглядным, и вот почему: переходная характеристика представляет собой, как известно, отклик системы на воздействие в виде единичной функции. А единичная функция представляет собой сочетание бесконечно быстрого изменения (в момент скачка) и бесконечно медленного изменения (на протяжении всего времени после скачка, где функция остается постоянной). Функция с такими свойствами не может быть ни продифференцирована, ни проинтегрирована с удовлетворительным приближением; поэтому различие между идеальными и реальными характеристиками очень велико, сравнивая их, трудно судить о практических достоинствах схемы. Эти соображения заставляют предпочесть описание свойств дифференцирующих и интегрирующих схем на частотном языке, т.е. при помощи частотных характеристик коэффициента передачи.

Для уяснения действия дифференцирующих и интегрирующих схем полезно рассмотреть пример, приближающийся к действительным условиям применения этих схем. Пусть на дифференцирующую

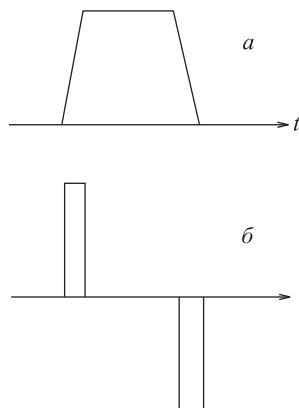


Рис. 141

схему рис. 135 подается трапецеидальный импульс рис. 141, а. График производной такого импульса показан на рис. 141, б. На выходе же дифференцирующей схемы получается различная картина в зависимости от крутизны фронтов импульса. Из рис. 142 видно, что при крутых фронтах происходит значительное искажение формы производной. При малой крутизне форма импульса на выходе приближается к прямоугольной (рис. 142, в).

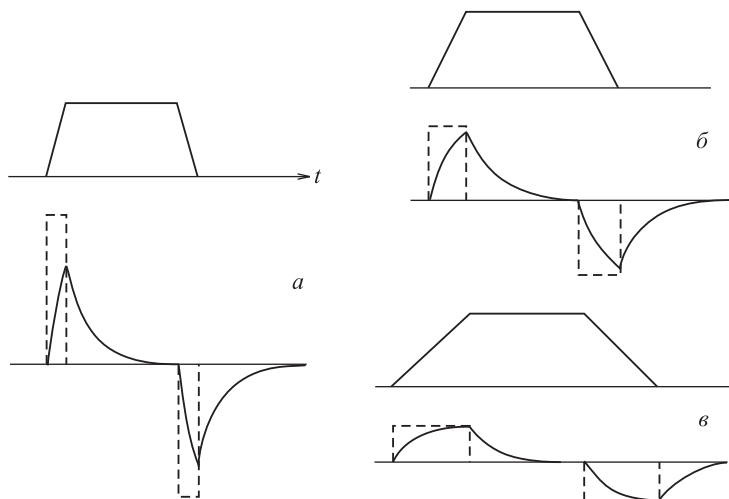


Рис. 142

Рассмотрим одну возможность радикального улучшения качества дифференцирования и интегрирования. Для этого вернемся к простейшим схемам рис. 135 и 137 и объединим их в виде, показанном на рис. 143.

Коэффициент передачи этой схемы равен

$$K = \frac{Z_2}{Z_1 + Z_2}. \quad (147)$$

Если $Z_1 \gg Z_2$, то приблизительно

$$K \approx \frac{Z_2}{Z_1}.$$

Для получения дифференцирования нужно взять $Z_2 = R$, $Z_1 = \frac{1}{i\omega C}$ (или $Z_2 = i\omega L$, $Z_1 = R$); для интегрирования, наоборот, выбираем $Z_2 = \frac{1}{i\omega C}$, $Z_1 = R$ (или $Z_2 = R$, $Z_1 = i\omega L$). Как видим, несовершенство схемы рис. 143 определяется наличием в знаменателе выражения (147) для коэффициента передачи слагаемого Z_2 . Для устранения влияния

этого слагаемого можно прибегнуть к обратной связи. При этом возможны два варианта:

- 1) либо при помощи отрицательной обратной связи увеличивается слагаемое Z_1 настолько, что слагаемым Z_2 можно пренебречь;
- 2) либо при помощи положительной обратной связи вносится отрицательное сопротивление, равное $-Z_2$, так что слагаемое Z_2 компенсируется и выпадает.

Для осуществления первого варианта пригодна схема рис. 130, б, для которой

$$K_\beta = \frac{K(1 - \beta_2)}{1 - K\beta_2}, \quad \beta_2 = \frac{Z_1}{Z_1 + Z_2} \quad (148)$$

или

$$K_\beta = \frac{KZ_2}{Z_1(1 - K) + Z_2}.$$

Сравним это выражение с (147). Как видим, для того, чтобы можно было пренебречь в знаменателе (148) слагаемым Z_2 , достаточно иметь большой отрицательный коэффициент усиления K .

При этом условии

$$K_\beta \approx \frac{1 - \beta_2}{\beta_2} = -\frac{Z_2}{Z_1}.$$

Это означает, что, увеличивая $|K|$, можно получить сколь угодно высокое качество дифференцирования или интегрирования, не меняя значений Z_1 и Z_2 (и, в частности, их отношения, определяющего постоянную времени).

Второй вариант применения обратной связи осуществляется посредством схемы рис. 130, а, для которой

$$K_\beta = \frac{K\beta_1}{1 - K\beta_1}, \quad \beta_1 = \frac{Z_2}{Z_1 + Z_2}$$

или

$$K_\beta = \frac{KZ_2}{Z_1 + Z_2(1 - K)}. \quad (149)$$

Для того чтобы уменьшить второе слагаемое в знаменателе (149) и приблизиться, таким образом, к идеальному дифференцированию или интегрированию, нужно применить усилитель с коэффициентом усиления K , как можно менее отличающимся от $+1$. При этом условии

$$K_\beta \approx \frac{\beta_1}{1 - \beta_1} = \frac{Z_2}{Z_1}.$$

Требуемые соотношения могут быть, например, получены при использовании в качестве усилителя катодного повторителя. Схема дифференцирующего или интегрирующего устройства с катодным повторителем показана на рис. 144.

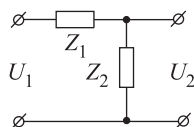


Рис. 143

Коэффициент усиления катодного повторителя тем ближе к единице, а следовательно, качество дифференцирования или интегрирования тем выше, чем меньше отношение R_i/R_k и чем больше коэффициент усиления триода. Мы имеем для катодного повторителя в общем случае

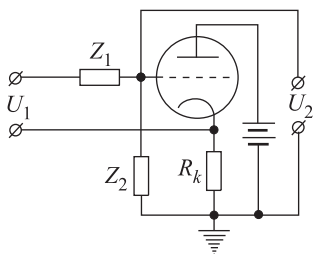


Рис. 144

$$K = \frac{\mu R_k}{R_i + (1 + \mu) R_k}.$$

При $R_i \ll R_k$

$$K \approx \frac{\mu}{1 + \mu},$$

а при $\mu \gg 1$

$$K \approx +1.$$

Оба эти условия выполнить затруднительно. Поэтому на практике применяют более сложные схемы, основанные на тех же принципах, но содержащие большее количество ламп.

§ 42. Электромеханические системы

В радиотехнических устройствах часто применяются электромеханические колебательные системы, как, например, камертонные резонаторы, магнитоstrictionные и пьезоэлектрические (в частности, кварцевые) фильтры и т.п. Действие всех этих устройств основано на преобразовании энергии из одной формы в другую, в частности, на преобразовании электрической энергии в механическую или обратно. Вследствие преобразования энергии возникает определенная связь между механическими и электрическими явлениями, происходящими в преобразователе, так называемая *электромеханическая связь*. Характер этой связи поясним для начала на примере отнюдь не радиотехническом: рассмотрим электродвигатель постоянного тока. Пусть электродвигатель питается от сети напряжением U . Электрическое сопротивление двигателя обозначим через R_0 . При включении двигателя в первый момент (пока он еще не пришел в движение) через него потечет большой пусковой ток

$$I_0 = \frac{U}{R_0}.$$

Но по мере того, как ротор будет набирать скорость, в нем будет наводиться противоэлектродвижущая сила e , пропорциональная скорости. Коэффициент пропорциональности между угловой скоростью Ω и эдс E в выражении

$$E = \mathfrak{M}\Omega$$

характеризует электромеханическую связь и называется *коэффициентом электромеханической связи*. Угловая скорость зависит, с одной

стороны, от развиваемого двигателем вращающего момента, а с другой стороны, от механической нагрузки на валу. Чем больше нагрузка, тем меньше угловая скорость, тем меньше противоэлектродвижущая сила и тем больше ток

$$I = \frac{U - E}{R_0},$$

забираемый двигателем из сети.

Выразим связь между вращающим моментом M и угловой скоростью Ω соотношением

$$M = \tau \Omega$$

и назовем величину

$$\tau = \frac{M}{\Omega}$$

механическим сопротивлением. Далее, вращающий момент пропорционален току и равен

$$M = \mathfrak{M} I.$$

Таким образом,

$$\frac{E}{\Omega} = \frac{M}{I} = \mathfrak{M}. \quad (150)$$

В этом соотношении проявляется один довольно общий физический принцип, называемый принципом взаимности.

Теперь, используя все вышеприведенные соотношения, можно записать следующую цепочку равенств:

$$I = \frac{U - E}{R_0}, \quad IR_0 = U - E, \quad IR_0 + E = U, \\ IR_0 + \mathfrak{M}\Omega = IR_0 + \mathfrak{M} \frac{M}{\tau} = I \left(R_0 + \frac{\mathfrak{M}^2}{\tau} \right) = U$$

и, наконец,

$$R = \frac{U}{I} = R_0 + \frac{\mathfrak{M}^2}{\tau} = R_0 + R'.$$

Это соотношение показывает, что к собственному электрическому сопротивлению R_0 двигателя добавляется еще эквивалентное сопротивление R' , обусловленное реакцией механической части (сопротивление R' называют иногда «сопротивлением движения»). При $\tau \rightarrow \infty$ $R' = 0$, $R = R_0$; это значит, что двигатель заторможен ($\Omega = 0$, $E = 0$) и ток при этом есть пусковой ток. При уменьшении механического сопротивления τ эквивалентное сопротивление возрастает, а ток соответственно убывает. Если двигатель работает холостую, то механическое сопротивление τ_0 обусловлено только собственными потерями двигателя, главным образом трением; эквивалентное сопротивление R' имеет при этих условиях наибольшее значение, а ток — наименьшее (ток холостого хода).

Из приведенных соотношений следует, что электромеханический преобразователь, примером которого является рассмотренный нами

двигатель постоянного тока, можно рассматривать как своего рода трансформатор, но трансформатор особого рода, в котором вторичная сторона имеет механическую природу. Аналогия с трансформатором состоит в том, что сопротивление вторичной стороны может быть пересчитано в первичную цепь; роль коэффициента трансформации при таком пересчете играет коэффициент электромеханической связи $\mathfrak{M}$.

Рассматриваемые нами электромеханические преобразователи обратимы. Это значит, что двигатель, питаемый с электрической стороны и отдающий энергию на механической стороне, может быть обращен и использован в качестве генератора, питаемого с механической стороны и отдающего энергию на электрической стороне. Так, электродвигатель может быть использован в качестве динамомашин, для которой получаются совершенно аналогичные соотношения, основанные на том же общем равенстве (150). Получается, в частности,

$$\mathfrak{r} = \mathfrak{r}_0 + \frac{\mathfrak{M}^2}{R},$$

где $\mathfrak{r}$ — эквивалентное механическое сопротивление, которое приходится преодолевать двигателю, вращающему динамомашину; $\mathfrak{r}_0$ — собственное механическое сопротивление машины; R — сопротивление нагрузки на электрической стороне.

Изложенные общие принципы остаются без изменений, вне зависимости от характера движения. В частности, точно такие же соотношения имеют место и при переменном токе и при колебательном движении; этот случай нас и интересует с точки зрения радиотехнических применений. Придется только изменить смысл величин, входящих в предыдущие выражения, а именно: вместо омических сопротивлений R_1 , R_0 и R' будем говорить о комплексных сопротивлениях Z , Z_0 , и Z' , вместо $\mathfrak{r}$, $\mathfrak{r}_0$ появятся комплексные механические сопротивления $\mathfrak{z}$, $\mathfrak{z}_0$ вместо постоянных токов и напряжений появятся комплексные амплитуды, а вместо момента M и угловой скорости Ω — колебательная сила F и колебательная скорость v , выражаемые также своими комплексными амплитудами. При такой замене все вышеприведенные формулы годятся и для колебательного движения. Так, например, электрическое сопротивление электродинамического громкоговорителя равно

$$Z = Z_0 + \frac{\mathfrak{M}^2}{\mathfrak{z}}, \quad (151)$$

где $\mathfrak{z} = F/v$ — механическое сопротивление подвижной системы громкоговорителя. Сила, действующая на подвижную систему, создается за счет взаимодействия тока, протекающего в подвижной катушке громкоговорителя с постоянным магнитным полем, в котором она помещена. С другой стороны, при колебании диафрагмы в обмотке подвижной катушки создается эдс — весь механизм взаимодействия в общих чертах совершенно такой же, как в двигателе.

При определенной частоте (обычно около 100 Гц) происходит механический резонанс подвижной системы; при этом механическое сопротивление имеет минимум, а электрическое сопротивление (см. (151)) — максимум.

Таким образом, поведение громкоговорителя, рассматриваемого с электрической стороны, ничем не отличается качественно от поведения параллельного электрического контура. Но имеется важное *количественное* отличие, которое и определяет интерес к электромеханическим устройствам. Дело в том, что затухание механической колебательной системы легко может быть сделано на порядок ниже, чем затухание самого лучшего электрического контура; поэтому к электромеханическим системам прибегают в тех случаях, когда нужно получить контуры с очень высокой эквивалентной добротностью.

Второй выгодной особенностью механических колебательных систем является то, что их резонансная частота может сохраняться с высокой степенью постоянства независимо от изменения внешних условий, в частности температуры. Для этого нужно лишь изготовить детали механической колебательной системы из материала с малым температурным коэффициентом расширения.

В качестве примера электромеханической системы, эквивалентной одиночному контуру с высокой добротностью, рассмотрим схему с камертоном рис. 145. Здесь имеется электромагнит 1, питаемый входным напряжением U_1 .

Электромагнит воздействует на одну из ножек стального камертона. В камертоне возникают вынужденные колебания, происходящие симметрично (относительно плоскости симметрии, намеченной штрихпунктирной линией), т. е. колеблются в противоположных направлениях обе ножки. Колебания второй ножки изменяют магнитный поток, пронизывающий обмотку электромагнита 2, и в этой обмотке наводится эдс, поступающая на выход. Интенсивные колебания происходят при резонансе, т. е. когда частота напряжения U_1 совпадает с частотой, на которую настроен камертон. С расстройкой выходное напряжение убывает очень быстро.

Таким образом, система в целом эквивалентна электрическому четырехполюснику, коэффициент передачи которого в зависимости от частоты представляется обычной резонансной кривой, но только очень острой. Внутреннее же устройство этого четырехполюсника характеризуется двукратным преобразованием энергии — сначала из электрической в механическую, а затем обратно, из механической в электри-

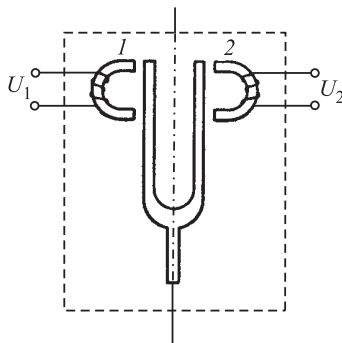


Рис. 145

ческую. Левая часть системы работает как двигатель, правая — как генератор. Если соединить выход и вход схемы рис. 145 через усилитель, то получится камертонный генератор — устройство, применяемое для получения высокостабильных по частоте электрических колебаний звуковой частоты. Связывая между собой несколько четырехполюсников по схеме рис. 145, можно получить различного вида камертонные фильтры и т. д.

Схема рис. 145 пригодна только для низких звуковых частот (не выше порядка 5 кГц, так как построить камертон на более высокие частоты затруднительно). Для более высоких частот (порядка до 100 кГц) применяются магнитострикционные фильтры. Здесь используется явление *магнитострикции*, состоящее, вообще говоря, в том, что изменение механического состояния магнитострикционного тела вызывает изменение его магнитного состояния и обратно.

Следовательно, на основе явления магнитострикции возможно осуществить как преобразователь-двигатель, так и преобразователь-генератор и осуществить, таким образом, электромеханический резонатор, не отличающийся по идее от схемы рис. 145. Различие состоит лишь в конструктивном выполнении, которое мы здесь не обсуждаем.

Наибольшее значение для радиотехники имеют электромеханические системы, построенные на основе явления *пьезоэлектричества*. Явление это, наблюдаемое в некоторых кристаллах, состоит в том, что при механическом сжатии кристалла возникает электрическая поляризация его и, наоборот, при приложении к граням кристалла электрического напряжения происходит механическая деформация кристалла. Известно и применяется большое число пьезоэлектрических кристаллов, например, фосфаты (соли фосфорной кислоты), тартраты (соли виннокаменной кислоты), в том числе сегнетова соль, титанаты и др., но наибольшее значение в радиотехнике сохранил до сих пор один из первых известных пьезокристаллов, а именно кварц (горный хрусталь, SiO_2).

Он обладает замечательными физико-химическими свойствами: тверд, прочен, жаростоек, негигроскопичен, не боится ни кислот, ни щелочей и, что особенно важно, обладает очень малым коэффициентом теплового расширения.

Конструктивное выполнение кварцевого резонатора очень просто: вырезанная из кристалла кварца пластинка (ребра пластинки определенным образом ориентированы относительно осей кристалла) помещается между двумя обкладками плоского конденсатора, обычно с небольшими воздушными зазорами (рис. 146). Кварцевая пластинка обладает определенной резонансной частотой, зависящей только от толщины пластинки h и определяемой по формуле

$$f_0 = 2,8/h,$$

где f_0 — в мегагерцах; h — в миллиметрах. Так, например, для частоты 1 МГц толщина кварцевой пластинки должна составлять 2,8 мм.

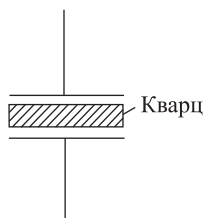


Рис. 146

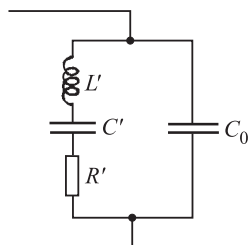


Рис. 147

Теоретический анализ для системы рис. 146 приводит к схеме замещения, показанной на рис. 147. Это — сложный параллельный контур. Емкость C_0 — действительная электрическая емкость кварцевого конденсатора. В другой ветви включены эквивалентные параметры L' , C' и R' , представляющие собой результат пересчета в электрическую цепь механических параметров кварцевой пластинки. Интересно заметить, что L' — очень большая индуктивность (порядка десятка Гн), C' — очень малая емкость (порядка единиц пФ), R' — очень малая величина, так что добротность эквивалентного контура получается порядка 10^3 и даже 10^4 .

Из двухполюсников вида рис. 147 можно составлять различные схемы сложных фильтров, обладающих очень высокими качественными показателями. Кварц широко применяется также для стабилизации частоты ламповых генераторов.

Глава 6

ВОЛНОВЫЕ СИСТЕМЫ

§ 43. Линии; общие понятия

Переходим к исследованию систем с распределенными параметрами. Как уже говорилось в § 28, все реальные системы являются системами с распределенными параметрами. Однако во многих случаях учет этого обстоятельства необязателен, так как можно получить вполне удовлетворительное приближение к действительности, заменяя реальную систему с распределенными параметрами упрощенной моделью в виде системы с сосредоточенными параметрами¹⁾. На возможности такой замены и основан весь материал предыдущей главы. Теперь же будем рассматривать системы, *основные* свойства которых вытекают именно из того, что характеризующие систему параметры распределены. Для этих систем замена моделью с сосредоточенными параметрами приводит к потере основных черт явлений, и, следовательно, такая замена в общем случае недопустима. В радиотехнике системы с распределенными параметрами встречаются в виде линий и кабелей (двухпроводных и коаксиальных), антенн, фидеров, волноводов и объемных резонаторов. Начнем с простейшей формы линии в виде двух параллельных проводов.

Такая линия характеризуется четырьмя распределенными по длине линии параметрами: активным сопротивлением R , индуктивностью L , емкостью C и активной проводимостью между проводниками (обусловленной, например, несовершенством изоляции) G . Каждый отрезок линии, как бы мал он ни был, можно охарактеризовать названными четырьмя параметрами. Если линия однородна, т.е. если параметры распределены по длине линии равномерно, то удобно ввести так называемые *погонные параметры*, отнесенные к единице длины. Обозначим их

¹⁾ Вопрос о возможности применения моделей с сосредоточенными параметрами будет рассмотрен в § 53.

R_0, L_0, C_0 и G_0 . Зная погонные параметры, легко определить параметры отрезка линии заданной длины l . Так, например, сопротивление такого отрезка будет $R = R_0 l$, емкость $C = C_0 l$ и т. д.

Для начала рассмотрим линию без потерь, т. е. положим $R_0 = G_0 = 0$. Это сильно упростит исследование, но не приведет к потере основных закономерностей. К тому же в радиотехнике, имеющей обычно дело с относительно короткими линиями (например, антенны, фидеры и т. п. устройства имеют длины порядка десятков или сотен *метров*, тогда как линии проводной связи имеют длину сотни и тысячи *километров*), такое приближение зачастую достаточно и для технических расчетов.

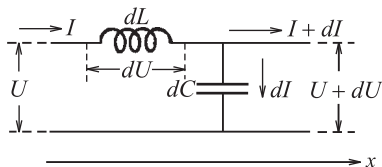


Рис. 148

Возьмем бесконечно короткий отрезок линии без потерь, обозначив его длину через dx . Индуктивность и емкость такого отрезка будут соответственно

$$dL = L_0 dx, \quad dC = C_0 dx,$$

и можно представить схему отрезка, как показано на рис. 148. Даны напряжение u и ток i на входе отрезка. На выходе отрезка напряжение и ток изменятся. Напряжение уменьшится на величину падения напряжения на индуктивности dL , а ток уменьшится на величину тока, ответвляющегося в емкость dC . Можно записать

$$\left. \begin{aligned} di &= \frac{\partial i}{\partial x} dx = -dC \frac{\partial u}{\partial t} \\ du &= \frac{\partial u}{\partial x} dx = -dL \frac{\partial i}{\partial t} \end{aligned} \right\} \quad (152)$$

(заметим, что в первой строке взято напряжение u , а не $u + du$, так как слагаемое du дало бы слагающую тока второго порядка малости. Поэтому безразлично, рисуем мы схему рис. 148 в виде правого или левого Г-звена, или как-либо иначе). Существенно заметить, что в уравнениях (152) фигурируют *частные* производные тока и напряжения. Это обусловлено тем, что ток и напряжение зависят не только от времени, но меняются по длине линии, т. е. зависят также от координаты x , отсчитываемой вдоль линии. Этим определяется вся специфика систем с распределенными постоянными.

Введем погонные индуктивность и емкость и перепишем уравнения (152) в виде

$$\left. \begin{aligned} \frac{\partial i}{\partial x} &= -C_0 \frac{\partial u}{\partial t} \\ \frac{\partial u}{\partial x} &= -L_0 \frac{\partial i}{\partial t} \end{aligned} \right\}. \quad (153)$$

Для того чтобы получить из этой системы уравнения, содержащие только u и i , продифференцируем первое уравнение еще раз по t , а второе еще раз по x . Получим

$$\left. \begin{aligned} \frac{\partial^2 i}{\partial x \partial t} &= -C_0 \frac{\partial^2 u}{\partial t^2} \\ \frac{\partial^2 u}{\partial x^2} &= -L_0 \frac{\partial^2 i}{\partial t \partial x} \end{aligned} \right\}.$$

Исключая из этой пары уравнений $\frac{\partial^2 i}{\partial x \partial t} = \frac{\partial^2 i}{\partial t \partial x}$, найдем уравнение для напряжения

$$\frac{\partial^2 u}{\partial x^2} = L_0 C_0 \frac{\partial^2 u}{\partial t^2},$$

или

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0, \quad (154)$$

где $v = 1/\sqrt{L_0 C_0}$. Это есть так называемое *волновое уравнение*. Аналогично получим уравнение для тока

$$\frac{\partial^2 i}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 i}{\partial t^2} = 0,$$

т. е. изменения напряжения и тока управляются совершенно одинаковыми закономерностями.

Решение волнового уравнения (154) зависит от начальных и граничных условий. Некоторые частные случаи будут рассмотрены в следующих параграфах. Пока установим лишь важнейшее общее свойство решений волнового уравнения. Оно состоит в том, что решением уравнения (154) может быть функция

$$u(x, t) = f\left(t \pm \frac{x}{v}\right), \quad (155)$$

где f — *любая* функция (дважды дифференцируемая). В справедливости этого положения можно убедиться, подставляя решение (155) в уравнение (154). Действительно,

$$\begin{aligned} \frac{\partial u}{\partial x} &= \pm \frac{1}{v} f' \left(t \pm \frac{x}{v} \right), \\ \frac{\partial^2 u}{\partial x^2} &= \frac{1}{v^2} f'' \left(t \pm \frac{x}{v} \right), \\ \frac{\partial^2 u}{\partial t^2} &= f'' \left(t \pm \frac{x}{v} \right). \end{aligned}$$

Решение (155) раскрывает общую картину явлений, происходящих в линии. По физическому смыслу это решение выражает *два возмущения, распространяющихся по линии в противоположных*

направлениях с постоянной конечной скоростью v , зависящей от параметров линии. Эта скорость для воздушной линии практически не отличается от c — скорости света. Но в кабелях с повышенной погонной индуктивностью и с изоляцией, обладающей большой диэлектрической проницаемостью, скорость распространения v может быть значительно ниже скорости света.

Возмущение, распространяющееся с конечной скоростью, называют вообще *волной*. Таким образом решение (155) представляет две движущиеся во встречных направлениях волны. Эти волны называются *бегущими*.

На рис. 149, *а* изображена двухпроводная линия; к некоторому сечению линии подключен источник напряжения, создающий импульс произвольной формы $f(t)$. В некоторый момент $t = t_1$ напряжение вдоль линии будет распределено, как показано на рис. 149, *б*: две волны, расположенные симметрично относительно места возбуждения, разбегаются в противоположные стороны (как указывают стрелки). В более поздний момент, $t = t_2 > t_1$, обе волны еще более удалятся от места возбуждения. За промежуток времени $t_2 - t_1$ каждая из них пробежит путь

$$l = v(t_2 - t_1).$$

Форма волны при этом не изменится.

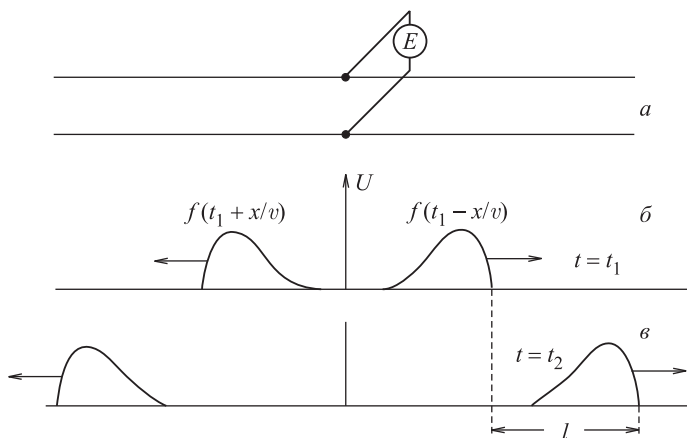


Рис. 149

Рисунок 149 представляет зависимость напряжения от координаты для фиксированного момента времени. Рисунок 149, *б* и 149, *в* — это как бы два моментальных снимка, два последовательных кадра, на которых запечатлены распределения напряжения вдоль линии для двух следующих друг за другом моментов времени. Другими словами, мы рассматриваем функцию $f\left(t - \frac{x}{v}\right)$ двух переменных t и x как

функцию x при фиксированном t . Теперь рассмотрим ту же функцию как функцию t при фиксированном x . На рис. 150 изображена функция

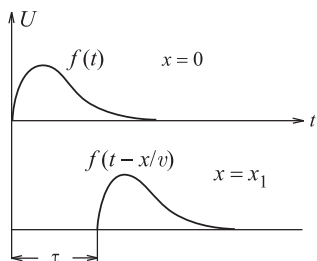


Рис. 150

$f(t)$, которую мог бы записать осциллограф, включенный в месте возбуждения ($x = 0$). Если включить осциллограф в сечении линии, отстоящем от места возбуждения на x_1 , то будем наблюдать тот же процесс, но запаздывающий на время

$$\tau = \frac{x_1}{v}.$$

Таким образом, при помощи линии можно получить идеальную задержку, и затруднение состоит лишь в том, что скорость распространения очень велика, а потому для получения сколько-нибудь значительной задержки потребовались бы очень длинные линии. Например, для такой малой задержки, как $\tau = 1$ мкс, считая $v \approx c$, получим

$$x_1 = v\tau \approx c\tau = 3 \cdot 10^8 \cdot 10^{-6} = 300 \text{ м.}$$

Вопрос этот будет еще рассмотрен в § 53.

Итак, основное свойство систем с распределенными параметрами и, в частности линий, состоит в том, что ток и напряжение являются функциями двух переменных t и x , т. е. зависят не только от времени, но и от координаты. С математической точки зрения это приводит к тому, что ток и напряжение выражаются уравнениями в частных производных по обоим переменным. С физической точки зрения самое существенное состоит в том, что как напряжение, так и ток представляют собой *волны*, распространяющиеся вдоль линии с конечной скоростью v . Поэтому уравнение (154) называется *волновым*, а системы с распределенными параметрами, в которых наблюдаются волны, — *волновыми* системами.

§ 44. Волновые параметры

Будем питать линию источником синусоидального напряжения, развивающим в начале линии напряжение

$$u_1 = \dot{U}_0 e^{i\omega t}.$$

Начало координаты x расположим в начале линии (рис. 151). Напряжение в любом сечении линии будет также синусоидально. На этом основании можно записать

$$\frac{\partial^2 u}{\partial t^2} = -\omega^2 u,$$

и вместо (154) получим волновое уравнение в виде

$$\frac{\partial^2 u}{\partial x^2} + k^2 u = 0,$$

где

$$k = \frac{\omega}{v}$$

— так называемое *волновое число*. Волновое уравнение для рассматриваемого частного случая вырождается в обыкновенное уравнение (т. е. уравнение полных производных). Решение обыкновенного уравнения второго порядка имеет, как известно, вид

$$u = A e^{\alpha_1 x} + B e^{\alpha_2 x},$$

где α_1 и α_2 — корни характеристического уравнения

$$\alpha^2 + k^2 = 0,$$

т. е.

$$\alpha_1 = i k, \quad \alpha_2 = -i k,$$

а постоянные A и B определяются из граничных условий. Итак,

$$u = A e^{i k x} + B e^{-i k x}. \quad (156)$$

По физическому смыслу два члена решения (156) представляют собой две волны, из которых одна движется от начала к концу линии, а вторая — в обратном направлении. Именно, член $B e^{-i k x}$ представляет волну, движущуюся в сторону положительных x , т. е. слева направо на рис. 151.

Предположим, что обратная волна отсутствует (в дальнейшем будет выяснено, при каких условиях это возможно). Тогда решение будет выражено только одним членом

$$u = B e^{-i k x}. \quad (157)$$

Постоянная B определяется из условия в начале линии, а именно

$$u|_{x=0} = u_1 = \dot{U}_0 e^{i \omega t}. \quad (158)$$

Положив в (157) $x = 0$ и сопоставляя с (158), находим

$$B = u_1,$$

так что окончательно

$$u = \dot{U}_0 e^{i(\omega t - kx)}. \quad (159)$$

Это — выражение для синусоидальной волны напряжения, бегущей от начала линии. На рис. 152, *а* и *б* изображены распределения напряжения для двух последовательных моментов $t = t_1$ и $t = t_2$. За время $t_2 - t_1$ волна пробегает путь $l = v(t_2 - t_1)$. Расстояние между двумя



Рис. 151

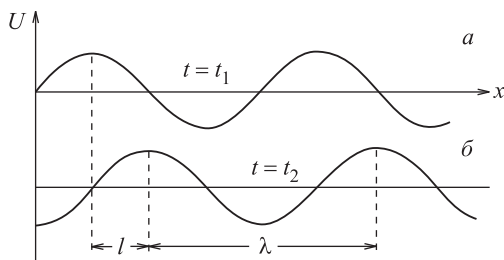


Рис. 152

гребнями или двумя впадинами или вообще между двумя точками с одинаковой фазой колебания называется *длиной волны* и обозначается λ .

На основании соотношения

$$\lambda = vT = \frac{v}{f} = 2\pi \frac{v}{\omega}$$

можно выразить волновое число как

$$k = \frac{\omega}{v} = \frac{2\pi}{\lambda}.$$

Перепишав выражение (159):

$$u = \dot{U}_0 e^{i(\omega t - \varphi)} = u_1 e^{-i\varphi},$$

видим, что в сечении, отстоящем на x от начала линии, напряжение запаздывает по фазе относительно напряжения U_1 в начале линии на угол

$$\varphi = kx = \frac{\omega}{v}x = 2\pi \frac{x}{\lambda}.$$

Если рассматривать отрезок линии как четырехполюсник, то коэффициент передачи такого четырехполюсника есть

$$K = \frac{u(x)}{u(0)} = e^{-i\varphi} = e^{-ikx}, \quad |K| = 1.$$

Величина ikx играет ту же роль, что и постоянная передачи четырехполюсника; применительно к волновым системам будем называть ее *постоянной распространения*, подчеркивая этим наименованием волновой характер происходящих явлений. Если постоянная распространения мнима, то это означает, что амплитуда колебания не убывает по мере того, как волна бежит вдоль линии; изменяется пропорционально расстоянию только фаза колебания.

Найдем теперь соотношение между током и напряжением. Воспользуемся для этого вторым уравнением (153), переписав его в виде

$$i\omega L_0 i = -\frac{du}{dx}. \quad (160)$$

Беря u из (159), найдем

$$\frac{du}{dx} = -i k u. \quad (161)$$

Из (160) и (161) получаем

$$i = \frac{k}{\omega L_0} u.$$

Отношение напряжения к току есть сопротивление. Оно равно в рассматриваемом случае

$$\frac{u}{i} = \frac{\omega L_0}{k} = v L_0 = \sqrt{\frac{L_0}{C_0}} = w.$$

Эта величина называется *волновым сопротивлением*. Следует заметить, что волновое сопротивление, во-первых, активно, а во-вторых, постоянно (т. е. не зависит ни от частоты, ни от координаты).

Величины k , v , w представляют собой основные числовые характеристики происходящих в линии волновых явлений. Мы называем эти величины волновыми параметрами.

§ 45. Отражение; стоячие волны

Когда линия питается источником напряжения, подключенным к началу линии, то возникает волна, бегущая от начала вдоль линии. Но в общем случае (как показывает общая форма решения волнового уравнения) существует и волна, бегущая в обратном направлении. Она является результатом *отражения*.

Отражение есть одно из основных волновых явлений. Мы постоянно встречаемся с ним в повседневной жизни и широко используем его в технике. Известно, например, отражение световых волн от зеркала; знакомо явление эха, представляющее собой отражение звуковых волн. Радиолокация использует отражение радиоволн от тех или иных предметов; действие эхолота основано на отражении от дна звуковых или ультразвуковых волн, распространяющихся в воде, и т. д.

Отражение волн любой физической природы происходит тогда, когда волна встречается с тем или иным препятствием на пути своего распространения, т. е. когда в среде, где распространяется волна, существует какая-либо *неоднородность* (т. е. изменение физических свойств среды).

Те же общие положения относятся к волнам тока и напряжения в линии. В простейшем случае препятствием для распространяющейся по линии волны является *конец* линии. Дойдя до него, волна тока или напряжения отражается. Возникает отраженная волна, бегущая в обратном направлении, т. е. от конца к началу линии.

Таким образом, в общем случае на линии существуют две волны: бегущая от начала к концу — *прямая волна* — и бегущая в обрат-

ном направлении — *отраженная волна*. Прямая и отраженная волны накладываются друг на друга, и происходит *интерференция* волн. Посмотрим, что получается в результате, интерференции синусоидальных волн.

Положим, что амплитуды напряжения для прямой и для отраженной волн одинаковы. Тогда можно записать для прямой волны

$$u_{\text{пр}} = \dot{U}_0 e^{i(\omega t - kx)}, \quad (162)$$

а для отраженной волны

$$u_{\text{отр}} = \dot{U}_0 e^{i(\omega t + kx)}. \quad (163)$$

Сложим оба выражения, чтобы найти результирующее напряжение

$$U = u_{\text{пр}} + u_{\text{отр}} = \dot{U}_0 e^{i\omega t} (e^{ikx} + e^{-ikx})$$

или

$$u = 2\dot{U}_0 \cos kx e^{i\omega t}. \quad (164)$$

Заметим прежде всего, что это выражение представляет колебание, происходящее по всей линии с одинаковой фазой. Это видно из того,

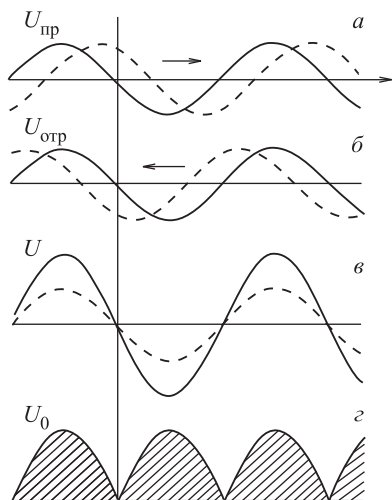


Рис. 153

что в множителе $e^{i\omega t}$, выражающем закон изменения напряжения во времени, отсутствует зависящий от координаты *фазовый сдвиг* (ср. с (162) и (163)). Таким образом, волна (164) не есть бегущая волна; мы называем такую волну *стоячей*. Вторым отличием стоячей волны от бегущей является то, что ее амплитуда зависит от координаты. Амплитуда стоячей волны определяется выражением

$$U_m = |2\dot{U}_0 \cos kx|.$$

Как видим, амплитуда в зависимости от координаты x изменяется от наибольшего значения $2|\dot{U}_0|$ до нуля. Места, где наблюдаются наибольшие значения амплитуды, называются *пучностями*, места

наименьших значений амплитуды — *узлами*. Местоположение узлов и пучностей не зависит от времени: они неподвижны. Если в узле напряжение равно нулю, то это значит, что оно *всегда* равно нулю. В любом же другом сечении линии напряжение изменяется во времени по синусоидальному закону, причем амплитуда колебания зависит от положения сечения: чем ближе к пучности тем амплитуда ближе

к наибольшему значению $2|\dot{U}_0|$ ¹⁾. Описанные соотношения поясняются построениями на рис. 153.

На рис. 153, *а* изображены графики мгновенных значений в функции координаты: *а* — напряжение прямой волны; *б* — напряжение отраженной волны; *в* — напряжение результирующей волны $U = U_{\text{пр}} + U_{\text{отр}}$; это и есть стоячая волна. Пунктирными линиями изображено то же самое, но для более позднего момента времени: прямая волна успела сместиться вправо, отраженная — влево. Напряжение результирующей волны уменьшилось, но узлы и пучности остались на своих местах, что и является характерным признаком *стоячей* волны. Как видно из рисунка, расстояние между соседними пучностями или соседними узлами равно $\lambda/2$.

На рис. 153, *в* показан график *амплитуд* стоячей волны. Штриховка сделана для того, чтобы отличить график амплитуд от графиков мгновенных значений.

В заключение подчеркнем еще раз признаки, различающие бегущую и стоячую волны:

1. В бегущей волне амплитуда колебания постоянна, а в стоячей волне — она периодическая функция координаты.
2. В бегущей волне фаза колебания есть линейная функция координаты, а в стоячей волне фаза постоянна (точнее, фаза постоянна на участке между двумя узлами; вдоль всей линии фаза меняется периодически, принимая попеременно значения 0 и π).

§ 46. Коэффициенты отражения

Выясним количественные характеристики явления отражения. Воспользуемся для этого волновым уравнением синусоидального режима

$$\frac{d^2 u}{dx^2} + k^2 u = 0$$

и его общим решением

$$u = A e^{i k x} + B e^{-i k x}. \quad (165)$$

Рассмотрим конечную линию длиной l , нагруженную на конце произвольным сопротивлением Z_2 (рис. 154).

Постоянные интегрирования A и B найдутся из граничных условий. Граничные условия в данном случае — это условия, определяющие режим в начале и в конце линии, а именно:

$$\begin{aligned} x = 0, \quad u &= u_1, \\ x = l, \quad \frac{u_2}{i_2} &= Z_2, \end{aligned} \quad (166)$$

¹⁾ Теперь можно уточнить сказанное выше по поводу фазы: колебания происходят с одинаковой фазой на протяжении участка между двумя узлами; на соседнем участке фаза отличается на π .

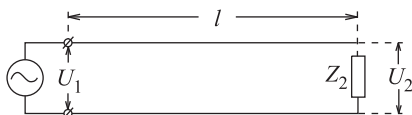


Рис. 154

где u_1 (напряжение в начале линии) и Z_2 (нагрузка на конце линии) — заданные величины. Для того чтобы найти ток, воспользуемся соотношением

$$i = -\frac{1}{i\omega L_0} \frac{du}{dx}.$$

Дифференцируя (165), получаем

$$i = -\frac{1}{w} (A e^{ikx} - B e^{-ikx}). \quad (167)$$

Вернемся к ф-ле (165). Она представляет напряжение на линии как сумму двух напряжений — напряжения прямой волны и напряжения отраженной волны. Можно записать (165) в виде

$$u = u_{\text{пр}} + u_{\text{отр}}.$$

Приписывая тому или другому члену (165) смысл прямой или отраженной волн, мы руководствуемся следующими соображениями: волна, распространяющаяся в направлении положительных x , т.е. от начала к концу линии, имеет запаздывающую фазу, т.е. $\varphi = -kx$, а волна, распространяющаяся в направлении отрицательных x , — опережающую фазу, т.е. $\varphi = kx$. Таким образом, член с множителем e^{-kx} представляет прямую волну, а член с множителем e^{kx} — отраженную. Принимая это во внимание, можно соответственным образом переписать и выражение (167) для тока:

$$i = \frac{1}{w} (u_{\text{пр}} - u_{\text{отр}}).$$

Возьмем теперь отношение напряжения к току в конце линии, где вторым граничным условием (166) это отношение задано. Получим

$$\frac{u_2}{i_2} = Z_2 = w \frac{u_{2\text{пр}} + u_{2\text{отр}}}{u_{2\text{пр}} - u_{2\text{отр}}};$$

отсюда сразу находим

$$\frac{u_{2\text{отр}}}{u_{2\text{пр}}} = \frac{Z_2 - w}{Z_2 + w}.$$

Отношение напряжения отраженной волны к напряжению прямой волны в месте отражения (в нашем случае в конце линии, т.е. при $x = l$) будем называть *коэффициентом отражения напряжения* и обозначать p_U . Итак,

$$p_U = \frac{Z_2 - w}{Z_2 + w}. \quad (168)$$

Важная ф-ла (168) показывает, что коэффициент отражения зависит только от соотношения между сопротивлением нагрузки Z_2 и волновым сопротивлением линии w . Полезно отметить три специальных случая:

1. $Z_2 \rightarrow \infty$. Это случай линии, разомкнутой на конце (холостой ход линии). При этом $p_U = 1$.
2. $Z_2 = 0$. Это — случай линии, замкнутой на конце накоротко (короткое замыкание линии). При этом $p_U = -1$.
3. $Z_2 = w$, т. е. линия нагружена на активное сопротивление, равное волновому. При этом $p_U = 0$, т. е. отражение отсутствует.

По поводу возможных значений коэффициента отражения p_U следует пояснить, что в общем случае он выражает отношение *комплексных* амплитуд отраженной и прямой волн и, следовательно, сам является комплексной величиной. Когда p_U вещественен и положителен, это значит, что оба напряжения (в месте отражения) совпадают по фазе. Когда p_U вещественен и отрицателен, это значит, что оба напряжения находятся в противофазе. Мнимость коэффициента отражения означает сдвиг фаз на $\pi/2$. Коэффициент отражения не может быть больше единицы по абсолютной величине. Противное означало бы, что амплитуда отраженной волны больше амплитуды прямой волны, т. е. что отраженная волна несет большую мощность, чем прямая. Этого не может быть, так как на конце линии источников энергии нет. Случай

$$|p_U| = 1$$

будем называть *полным отражением*. При полном отражении амплитуды отраженной и прямой волн равны. Именно этот случай мы рассматривали в предыдущем § 45, обсуждая возникновение стоячей волны.

Как видим, полное отражение получается, в частности, при холостом ходе и при коротком замыкании линии.

Посмотрим теперь, что происходит, когда линия нагружена на конце конечным (т. е. не равным ни нулю, ни бесконечности) сопротивлением R , не равным волновому сопротивлению w . В этом случае имеем

$$|p_U| = \frac{|R - w|}{R + w}.$$

Эта величина всегда меньше единицы; мы говорим при таких условиях о неполном отражении. При неполном отражении амплитуда отраженной волны меньше, чем амплитуда прямой волны. При наложении обеих волн также происходит интерференция, но амплитуда напряжения в узлах уже не равна нулю; наименьшая амплитуда равна разности амплитуд прямой и отраженной волн. Амплитуда же в пучности равна сумме амплитуд прямой и отраженной волн.

Возьмем ф-лу (165) и перепишем ее в виде

$$u = A e^{i k x} + A e^{-i k x} + (B - A) e^{-i k x} = 2A \cos kx + (B - A) e^{-i k x}.$$

Смысл этого преобразования состоит в том, что мы представили результирующее напряжение в виде суммы чистой стоячей волны (первый член) и бегущей волны (второй член).

Таким образом, можно трактовать неполное отражение как наложение стоячей и бегущей волн.

Положение часто характеризуют отношением минимальной амплитуды результирующего напряжения $U_{\text{мин}}$ к его максимальной амплитуде $U_{\text{макс}}$, т. е. отношением амплитуд в узле и в пучности.

Наименьшая амплитуда равна разности амплитуд прямой и отраженной волн, а наибольшая — их сумме. Таким образом,

$$\frac{U_{\text{мин}}}{U_{\text{макс}}} = \frac{|\dot{U}_{\text{пр}}| - |\dot{U}_{\text{отр}}|}{|\dot{U}_{\text{пр}}| + |\dot{U}_{\text{отр}}|} = \nu.$$

Величина ν называется *коэффициентом бегущей волны*. Этот коэффициент легко выразить через коэффициент отражения:

$$\nu = \frac{1 - \frac{|\dot{U}_{\text{отр}}|}{|\dot{U}_{\text{пр}}|}}{1 + \frac{|\dot{U}_{\text{отр}}|}{|\dot{U}_{\text{пр}}|}} = \frac{1 - |p_U|}{1 + |p_U|}.$$

При отсутствии отражения вся волна — бегущая; при этом $\nu = 1$. При полном отражении получается чисто стоячая волна; бегущая волна при этом отсутствует и $\nu = 0$. Коэффициент бегущей волны — параметр, удобный в том отношении, что он легко определяется непосредственно путем измерения напряжений в узле и в пучности; по коэффициенту бегущей волны можно найти модуль коэффициента отражения. При нагрузке на активное сопротивление R коэффициент бегущей волны равен просто отношению R/w , если $R < w$, или отношению w/R , если $R > w$.

Картина явлений при неполном отражении также может быть пояснена графиками, подобными рис.153; однако картина эта менее наглядна, чем в случае полного отражения, так как из-за наличия бегущей составляющей волны минимальные и максимальные значения напряжения в узлах и пучностях достигаются неодновременно. Поэтому найдем амплитуду результирующей волны непосредственно из ф-лы (165), которая совместно с граничными условиями при $Z_2 = R$ дает

$$|U| = C \sqrt{1 - \left(1 - \frac{R^2}{w^2}\right) \cos^2 k(l-x)},$$

где C — не зависящая от x величина.

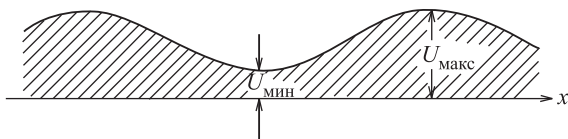


Рис. 155

На рис. 155 дан график этой зависимости для $R = \frac{1}{2}w$. При этом

$$p_U = \frac{R - w}{R + w} = -\frac{1}{3}, \quad \nu = \frac{1 - |p_U|}{1 + |p_U|} = \frac{R}{w} = \frac{1}{2}.$$

Мы говорили до сих пор только о коэффициенте отражения напряжения. Переписывая (167) в виде

$$i = i_{\text{пр}} + i_{\text{отр}}$$

и определяя коэффициент отражения тока как

$$p_I = \frac{i_{2\text{отр}}}{i_{2\text{пр}}},$$

найдем совершенно аналогичным путем, что

$$p_I = \frac{w - Z_2}{w + Z_2} = -p_U.$$

Так, например, при холостом ходе $p_U = 1$, $p_I = -1$.

Такого рода соотношения легко получить путем физических рассуждений. В самом деле, при холостом ходе конец линии разомкнут, ток в конце должен быть равен нулю. Приходит прямая волна, несущая не равный нулю ток; должна образоваться такая отраженная волна, чтобы для результирующей волны выполнялось граничное условие $i = 0$. Это получится только в том случае, если амплитуды отраженной и прямой волн будут равны, а фазы противоположны. Оба эти условия и выражаются равенством $p_I = -1$.

Разберем теперь случай нагрузки на реактивное сопротивление, т. е. положим

$$Z_2 = iX.$$

Коэффициент отражения напряжения будет

$$p_U = \frac{iX - w}{iX + w} = -\frac{w - iX}{w + iX}.$$

Как видим, в числителе и знаменателе стоят сопряженные комплексные величины. Их модули равны, а следовательно, модуль p_U равен единице, независимо от значения X . Итак, при нагрузке на реактивное сопротивление получаем полное отражение. Значение X влияет только на положение узлов и пучностей; расстояние между

соседними узлами или пучностями остается равным $\lambda/2$. В частном случае $X = w$ коэффициент отражения равен

$$p_U = -\frac{1-i}{1+i} = i,$$

т. е. коэффициент отражения оказывается в этом случае чисто мнимым. Это значит, что фаза отраженной волны повернута по фазе относительно прямой волны ровно на $\pi/2$.

В заключение этого параграфа рассмотрим еще вопрос об отражении от неоднородностей. До сих пор мы имели в виду только отражение

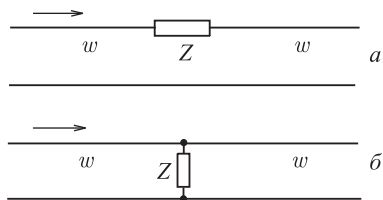


Рис. 156

от конца линии. Но отражение происходит от любой неоднородности, т. е. от места, где изменяются условия распространения волны. Пусть, например, линия с волновым сопротивлением w_1 сопряжена с другой линией, волновое сопротивление которой $w_2 \neq w_1$. Тогда можно считать, что первая линия нагружена входным сопротивлением второй. Если во второй линии нет отражений,

то ее входное сопротивление равно w_2 (об этом подробнее говорится в следующем параграфе) и коэффициент отражения будет

$$p_U = \frac{w_2 - w_1}{w_2 + w_1}.$$

Отражение будет происходить и в том случае, когда однородность линии нарушена включением сосредоточенных сопротивлений, как показано на рис. 156, а и б.

Если на участке линии справа от неоднородности отражений нет, то коэффициент отражения от места включения сосредоточенного сопротивления будет для схемы рис. 156, а

$$p_U = \frac{Z}{Z + 2w},$$

а для схемы рис. 156, б

$$p_U = -\frac{w}{2Z + w}.$$

Эти результаты получаются из общей ф-лы (168), в которой надо положить в первом случае $Z_2 = Z + w$, а во втором

$$Z_2 = \frac{Zw}{Z + w},$$

(так как сопротивление Z и волновое сопротивление правого участка линии включены параллельно).

При рассмотрении отражений от неоднородности естественно возникает вопрос, что делается *справа* от неоднородности (так как коэффициент отражения служит для описания явлений только слева от неоднородности. Мы все время полагаем, что прямая волна идет слева направо). Очевидно, что существует волна, *проходящая* через неоднородность и распространяющаяся далее слева направо. Назовем *коэффициентом прохождения* отношение напряжения проходящей волны U' к напряжению прямой волны $U_{\text{пр}}$, т. е.

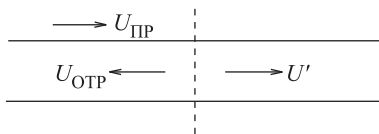


Рис. 157

$$p'_U = \frac{u'}{u_{\text{пр}}}.$$

Прямая, отраженная и проходящая волны ¹⁾ обозначены стрелками на рис. 157; местоположение неоднородности отмечено пунктирной линией. Напряжения по обе стороны некоторого сечения линии должны быть равны (принцип непрерывности). Поэтому

$$u_{\text{пр}} + u_{\text{отр}} = u'.$$

Подставляя это соотношение в определение коэффициента прохождения, получим

$$p'_U = \frac{u'}{u_{\text{пр}}} = \frac{u_{\text{пр}} + u_{\text{отр}}}{u_{\text{пр}}},$$

т. е.

$$p'_U = 1 + p_U.$$

Если входное сопротивление правого участка обозначено через Z_2 , а волновое сопротивление левого участка есть w , то

$$p'_U = 1 + \frac{Z_2 - w}{Z_2 + w} = \frac{2Z_2}{Z_2 + w}.$$

Если, например, неоднородность представляет собой сопряжение двух линий с волновыми сопротивлениями w_1 и w_2 , то (при отсутствии отражений справа)

$$p'_U = \frac{2w_2}{w_2 + w_1}, \quad p_U = \frac{w_2 - w_1}{w_2 + w_1}.$$

Не нужно думать, что при полном отражении проходящая волна отсутствует. Пусть, например $w_2/w_1 \rightarrow \infty$. При этом $p_U \rightarrow 1$ (полное отражение), а $p'_U \rightarrow 2$. Это значит, что проходящая волна имеет амплитуду, вдвое большую, чем прямая волна. Так и должно быть, так как

¹⁾ В оптике аналогичные понятия обозначаются терминами «падающая», «отраженная» и «преломленная» волны.

при полном отражении (при $p_U = 1$) напряжение результирующей волны удваивается (на напряжение прямой волны накладывается равное по амплитуде и совпадающее по фазе напряжение отраженной волны). Если же правый участок линии имеет меньшее волновое сопротивление, то при $w_2/w_1 \rightarrow 0$ коэффициент отражения $p_U \rightarrow -1$, а коэффициент прохождения $p'_U \rightarrow 0$. Для тока получаются обратные соотношения.

§ 47. Входное сопротивление линии

Найдем общее выражение для входного сопротивления линии с волновым сопротивлением w , длиной l , нагруженной на конце произвольным сопротивлением Z_2 .

Будем исходить из общих выражений для напряжения и тока:

$$u = A e^{i k x} + B e^{-i k x}, \quad (169)$$

$$i = -\frac{1}{w} (A e^{i k x} - B e^{-i k x}). \quad (170)$$

Разделив (169) на (170), получим для отношения напряжения к току в любом сечении линии

$$\frac{u}{i} = -w \frac{A e^{i k x} + B e^{-i k x}}{A e^{i k x} - B e^{-i k x}}. \quad (171)$$

Положив в этом выражении $x = 0$, найдем входное сопротивление

$$Z_1 = \frac{u_1}{i_1} = -w \frac{A + B}{A - B} = w \frac{1 + \frac{A}{B}}{1 - \frac{A}{B}}. \quad (172)$$

Отношение A/B определится при помощи граничного условия, задающего нагрузку Z_2 на конце линии. Полагая в (171) $x = l$, получим

$$Z_2 = \frac{u_2}{i_2} = -w \frac{A e^{i k l} + B e^{-i k l}}{A e^{i k l} - B e^{-i k l}}, \quad (173)$$

откуда

$$\frac{A}{B} = \frac{Z_2 - w}{Z_2 + w} e^{-i 2 k l}.$$

Подставляя это значение в (172), получаем после небольшого преобразования следующее выражение для входного сопротивления линии:

$$Z_1 = Z_2 \frac{1 + i \frac{w}{Z_2} \operatorname{tg} k l}{1 + i \frac{Z_2}{w} \operatorname{tg} k l}. \quad (174)$$

В общем случае входное сопротивление линии комплексно.

Рассмотрим три важных частных случая.

1. Холостой ход линии: $Z_2 \rightarrow \infty$. В этом случае

$$Z_1 = Z_{1\infty} = -i w \operatorname{ctg} kl. \quad (175)$$

2. Короткое замыкание линии: $Z_2 = 0$. При этом

$$Z_1 = Z_{10} = i w \operatorname{tg} kl. \quad (176)$$

3. Нагрузка на волновое сопротивление $Z_2 = w$. В этом случае

$$Z_1 = w.$$

Заметим, что при холостом ходе и коротком замыкании линии, т.е. в условиях полного отражения, входное сопротивление получается мнимым. Это означает, что линия не потребляет мощности от источника, что и естественно, так как потери в линии мы пока не учитываем, а в нагрузке мощность не потребляется (так как при хх или кз либо напряжение, либо ток равны нулю). При нагрузке же на волновое сопротивление отражений нет, и по линии течет поток энергии, поступающей в нагрузку. Входное сопротивление при этом действительно. Отсюда можно заключить, что если линия предназначена для передачи энергии, то лучше всего заставить ее работать в согласованном режиме, когда сопротивление нагрузки равно волновому сопротивлению. При этом источник энергии будет поставлен в наиболее выгодные условия, так как ему не придется развивать больших кажущихся мощностей.

Второе обстоятельство, заслуживающее внимания, состоит в том, что при полном отражении (хх и кз) входное сопротивление, будучи реактивным, принимает все возможные значения (от $-\infty$ до $+\infty$). Сопротивление оказывается периодической функцией аргумента kl . Заметим, что сопротивление любых систем с сосредоточенными параметрами выражается алгебраическими функциями. Для систем же с распределенными параметрами характерно появление трансцендентных функций.

Обращение сопротивления в нуль или в бесконечность свидетельствует о наличии в системе резонансных явлений. Так оно и есть: в линии конечной длины возможно бесчисленное множество резонансов, частоты которых находятся между собой в простых кратных отношениях.

На рис. 158 изображены графики модулей входных сопротивлений для разомкнутой (рис. 158, а) и короткозамкнутой (рис. 158, б) линий. Нули сопротивления для разомкнутой линии получаются при

$$kl = (2n + 1) \frac{\pi}{2}$$

или

$$l = (2n + 1) \frac{\lambda}{4} \quad (n = 0, 1, 2, 3, \dots).$$

Для короткозамкнутой линии нули сопротивления получаются при

$$kl = n\pi$$

или

$$l = 2n \frac{\lambda}{4} \quad \left(= n \frac{\lambda}{2} \right).$$

Аналогично находятся значения kl , при которых сопротивление обращается в бесконечность; эти значения лежат посередине между нулями, как видно из рис. 158.

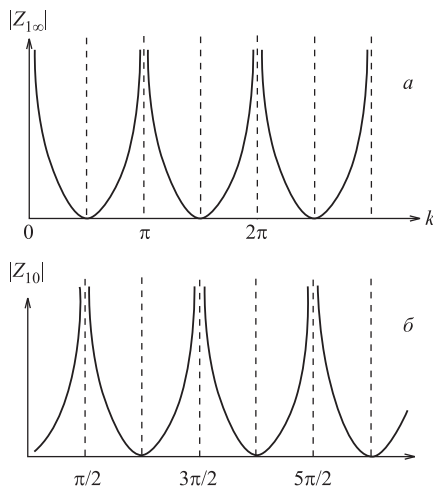


Рис. 158

Зависимость входного сопротивления линии от частоты получает в радиотехнике разнообразные применения. Рассмотрим в качестве примера линию длиной $l = \lambda/4$, замкнутую на конце накоротко. По формуле (176) (в которую нужно под-

ставить $kl = \pi/2$) найдем, что входное сопротивление такой линии бесконечно велико. Это значит, что если входные зажимы линии подключить к любой схеме, то это не повлияет на работу схемы; иными словами, $1/4$ -волновая кз линия ведет себя как *изолятор*. Это открывает интересные технические возможности. Так, например, двухпроводный фидер можно смонтировать на цельнометаллических $1/4$ -волновых стай-

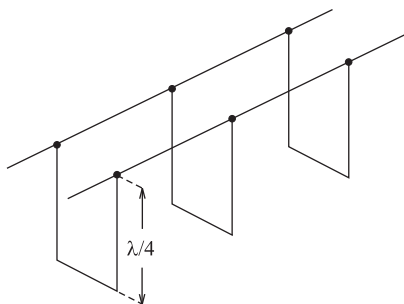


Рис. 159

ках, надежно изолирующих оба провода друг от друга (рис. 159). Конечно, в реальном устройстве (с учетом потерь) сопротивление небесконечно велико, но оно может быть достаточно велико и притом не на

одной частоте, а в пределах некоторой полосы частот. Последнее важно, так как по фидеру передается обычно модулированное колебание, имеющее спектр определенной ширины. 1/4-волновый изолятор обладает еще интересным свойством — избирательностью. Дело в том, что для частот, на которых длина волны отличается от $\frac{4}{2n+1}l$, сопротивление

изолятора невелико. Так, например, случайно попавшие на фидер рис. 159 напряжения низкой частоты окажутся замкнутыми накоротко через стойки. Заметим, что так как основания стоек непосредственно заземлены, то обеспечивается надежная защита от грозových разрядов.

Рассмотрим теперь предельные переходы, которые приведут к важным выводам общезначимого характера. Возьмем формулу (175) для входного сопротивления разомкнутой линии и разложим правую часть в ряд по степеням kl :

$$Z_{1\infty} = -i\omega \left(\frac{1}{kl} - \frac{kl}{3} - \frac{(kl)^3}{45} - \dots \right).$$

Пусть $kl \ll 1$. Тогда, приближенно,

$$Z_{1\infty} \approx -i\omega \frac{1}{kl} = -i\sqrt{\frac{L_0}{C_0}} \frac{1}{\omega\sqrt{L_0C_0}l} = \frac{1}{i\omega C_0l} = \frac{1}{i\omega C}.$$

Таким образом, при $kl \ll 1$ сопротивление разомкнутой линии есть просто емкостное сопротивление; линия ведет себя как *сосредоточенная* емкость $C = C_0l$. Поступим аналогично с формулой (176) для сопротивления короткозамкнутой линии

$$Z_{10} = i\omega \left(kl + \frac{(kl)^3}{3} + \frac{2(kl)^5}{15} + \dots \right),$$

что дает при $kl \ll 1$ приближенное выражение:

$$Z_{10} \approx i\omega kl = i\sqrt{\frac{L_0}{C_0}} \omega\sqrt{L_0C_0}l = i\omega L_0l = i\omega L,$$

т.е. короткозамкнутая линия при $kl \ll 1$ ведет себя как *сосредоточенная* индуктивность $L = L_0l$. Таким образом, два провода короткой разомкнутой линии можно рассматривать как две обкладки конденсатора, обладающего сосредоточенной емкостью C , а короткую замкнутую линию — как виток, обладающий сосредоточенной индуктивностью L . Условием, при котором такая трактовка законна, является неравенство $kl \ll 1$. Но это неравенство можно также переписать в виде

$$l \ll \lambda.$$

Итак, линия может быть заменена сосредоточенным эквивалентом, если длина ее мала по сравнению с длиной волны.

Этот критерий имеет общий характер. Теперь становится ясным, при каких условиях то или иное радиотехническое устройство может

быть представлено принципиальной схемой с сосредоточенными параметрами. Мы вправе представлять параметры (с известным приближением) как сосредоточенные, если размеры устройства и его элементов малы по сравнению с длиной волны. Становится понятным, почему обычные радиотехнические схемы ведут себя иногда так странно на укв. В самом деле, например, самый обыкновенный плоский воздушный конденсатор на частотах, для которых поперечник обкладки сравним с длиной волны, ведет себя уже не как конденсатор, а как разомкнутая линия; его сопротивление может быть и емкостным и индуктивным, очень большим и весьма малым; короче говоря, сопротивление может принимать любые значения в зависимости от $kl = 2\pi \frac{l}{\lambda}$.

§ 48. Методы согласования

Из предыдущего должно быть ясно, что в случае, когда линия используется для передачи энергии высокочастотных колебаний, желательно нагрузить ее на сопротивление, равное волновому, или, как говорят, на согласованную (с волновым сопротивлением) нагрузку. Иначе говоря, желательно получить коэффициент бегущей волны по возможности близкий к единице.

Однако не всегда возможно построить, например, фидер с заданным волновым сопротивлением или, наоборот, подогнать сопротивление нагрузки под волновое сопротивление фидера. В этих случаях применяют специальные согласующие устройства.

В технике сильных токов для согласования сопротивлений пользуются трансформаторами. Применение их основано на том, что, как известно, сопротивление вторичной цепи при пересчете на первичную цепь умножается на квадрат коэффициента трансформации.

Этот метод может применяться и на звуковых частотах и даже на не слишком высоких радиочастотах. На высоких радиочастотах не удается выполнить трансформатор с удовлетворительными показателями (так как влияет емкость обмоток и емкость между обмотками, потери в сердечнике, рассеяние и т. п.). Поэтому для радиотехники высоких частот характерны специальные методы согласования, некоторые из которых будут описаны ниже.

Простой и изящный метод согласования основан на использовании свойств 1/4-волновой линии. Теория этого метода состоит в следующем. Возьмем общую формулу для входного сопротивления линии (174) и положим $l = \frac{l}{4}\lambda$, т. е. $kl = \frac{\pi}{2}$. Получим входное сопротивление 1/4-волновой линии, нагруженной произвольным сопротивлением Z_2 ,

$$Z_1 = \frac{w^2}{Z_2}.$$

Из этого соотношения видно, что $1/4$ -волновая линия представляет собой нечто вроде трансформатора, в котором роль коэффициента трансформации играет волновое сопротивление w . Теперь ясно, что для согласования линии с волновым сопротивлением w при нагрузке $R \neq w$ можно воспользоваться $1/4$ -волновой вставкой, подобрав волновое сопротивление w_1 так, чтобы

$$Z_1 = \frac{w_1^2}{R} = w,$$

т. е.

$$w_1 = \sqrt{Rw}.$$

При этом условии на всей линии вплоть до начала вставки (рис. 160) отражений не будет; по линии будет распространяться чисто бегущая волна. В пределах вставки будет существовать стоячая волна, но это уже не столь важно. К тому же вставка лучше согласована с нагрузкой, чем основная линия (так как w_1 меньше отличается от R , чем w).

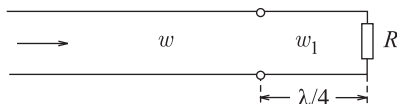


Рис. 160

Описанный метод полностью разгружает основную линию от отраженных волн, но неудобен с чисто технической точки зрения. Дело в том, что нужно построить вставку с заданным волновым сопротивлением, равным среднегеометрическому, из сопротивления нагрузки и волнового сопротивления основной линии. Это может оказаться затруднительным, а потому прибегают к более совершенному методу устранения отражений.

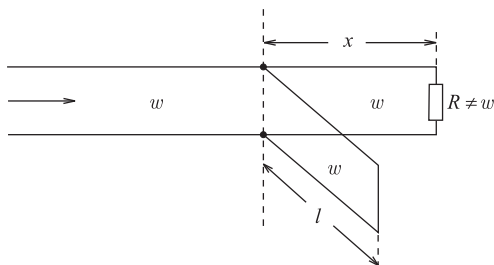


Рис. 161

Метод этот состоит в том, что к некоторому сечению линии (недалеко от конца) подключают короткозамкнутую петлю, представляющую собой отрезок линии с тем же волновым сопротивлением (рис. 161). При помощи такой петли можно скомпенсировать реактив-

ное сопротивление в этом сечении (так как входное сопротивление петли может быть каким угодно), а величину оставшегося активного сопротивления подогнать под волновое сопротивление w . Для выполнения обоих условий можно изменять место включения петли (координату x , рис. 161) и ее длину l . Сопротивление нагрузки R и волновое сопротивление w считаются заданными (разумеется, $R \pm w$).

Проводимость в сечении, где подключена петля,

$$Y = G + iB$$

и наше условие запишется в виде двух равенств:

$$B = 0, \quad G = \frac{1}{w}.$$

Проводимость Y равна сумме проводимости конца линии $Y_{\text{л}}$ (справа от сечения x) и проводимости петли $Y_{\text{п}}$. Для петли имеем (176)

$$Y_{\text{п}} = \frac{1}{i w \operatorname{tg} kl}.$$

Для линии (174)

$$Y_{\text{л}} = \frac{1}{R} \frac{1 + i \frac{R}{w} \operatorname{tg} kx}{1 + i \frac{w}{R} \operatorname{tg} kx}.$$

Результирующая проводимость

$$Y = Y_{\text{п}} + Y_{\text{л}} = \frac{1}{i w \operatorname{tg} kl} + \frac{1}{R} \frac{1 + i \frac{R}{w} \operatorname{tg} kx}{1 + i \frac{w}{R} \operatorname{tg} kx}.$$

Это выражение надо разделить на вещественную и мнимую части и приравнять мнимую часть нулю, а действительную $-1/w$. Получаются два уравнения относительно искомых x и l :

$$G = \frac{1}{R} \frac{1 + \operatorname{tg}^2 kx}{1 + \left(\frac{w}{R}\right)^2 \operatorname{tg}^2 kx} = \frac{1}{w},$$

$$B = -\frac{1}{w \operatorname{tg} kl} + \frac{1}{R} \frac{\left(\frac{R}{w} - \frac{w}{R}\right) \operatorname{tg} kx}{1 + \left(\frac{w}{R}\right)^2 \operatorname{tg}^2 kx} = 0.$$

Из первого уравнения находим

$$\operatorname{tg} kx = \pm \sqrt{R/w},$$

что и определяет одну из искомых величин — координату x сечения, в котором следует включить петлю. Подставляя значение $\operatorname{tg} kx$ во

второе уравнение, получаем

$$\operatorname{tg} kl = \pm \frac{\sqrt{R/w}}{\frac{R}{w} - 1},$$

и вторая искомая величина — длина петли l — определена. Решения для x и l многозначны; на практике берут обычно наименьшее значение (чтобы получить наименьшие размеры устройства).

§ 49. Резонансы в линии

Вернемся к общему решению волнового уравнения для напряжения

$$u = A e^{i k x} + B e^{-i k x}.$$

Используя граничные условия, заданные соотношениями:

$$x = 0, \quad u = u_1; \quad x = l, \quad \frac{u}{i} = Z_2$$

можно, проделав необходимые выкладки, получить следующее выражение для напряжения в любом сечении линии, если задано напряжение на входе u_1 и нагрузка Z_2 на конце линии:

$$u = u_1 \frac{Z_2 \cos k(l-x) + i w \sin k(l-x)}{Z_2 \cos kl + i w \sin kl}. \quad (177)$$

Рассмотрим один из случаев полного отражения; пусть линия разомкнута на конце, т. е. $Z_2 \rightarrow \infty$. В этом случае

$$u = u_1 \frac{\cos k(l-x)}{\cos kl} = \dot{U} \frac{\cos k(l-x)}{\cos kl} e^{i \omega t}. \quad (178)$$

Это — выражение стоячей волны; амплитуда изменяется вдоль линии по закону косинуса. Но важно заметить, что наибольшее значение амплитуды равно $\dot{U}/\cos kl$. Оно зависит от kl , а при определенных значениях kl обращается в бесконечность. При этом во всех точках, кроме узлов, амплитуда напряжения на линии бесконечно велика¹⁾ (при заданном конечном напряжении на входе). Такое положение назовем *резонансом напряжений в линии*.

Резонанс напряжений, как показывает формула (178), получается при

$$\cos kl = 0, \quad kl = (2n+1)\frac{\pi}{2}, \quad l = (2n+1)\frac{\lambda}{4}.$$

Входное сопротивление разомкнутой линии равно

$$Z_1 = -i w \operatorname{ctg} kl;$$

¹⁾ Напомним, что мы рассматриваем пока линию без потерь. В реальной линии с потерями напряжение на линии при резонансе велико, но конечно.

оно обращается при резонансе напряжений в нуль. Ток в разомкнутой линии найдем, дифференцируя (178):

$$i = -\frac{1}{i w} u_1 \frac{\sin k(l-x)}{\cos kl}.$$

Входной ток

$$i_1 = i|_{x=0} = \frac{u_1}{Z_{1\infty}} = -\frac{1}{i w} u_1 \operatorname{tg} kl;$$

при резонансе ток равен бесконечности везде, кроме узлов тока. Так как ток изменяется по закону синуса, то узлы напряжения и узлы тока смещены относительно друг друга на $\lambda/4$.

Соотношения при резонансе можно уяснить при помощи графика. На рис. 162, *а* изображены графики зависимости амплитуды напряжения от координаты. Амплитуда напряжения на входа задана и равна U_0 . Амплитуда напряжения вдоль линии меняется по закону $\cos kl(l-x)$. Кривая 1 представляет распределение амплитуды для такой частоты, когда первый узел напряжения лежит *вне* линии. Она полностью определена: это — косинусоида, максимум которой находится на конце линии ($x = 1$), так как там обязательно образуется пучность напряжения (линия разомкнута на конце). Длина волны косинусоиды определяется частотой $\lambda = (v/f)$, так что для данной частоты узлы располагаются вполне определенным образом и, наконец, косинусоида должна проходить через точку с координатами $x = 0$, $U_m = U_0$, так как напряжение в начале задано, и амплитуда его поддерживается неизменной. Имея все это в виду, посмотрим, как видоизменится график, если повышать частоту. При более высокой частоте волна будет короче; узел приблизится к началу линии (кривая 2); напряжение на линии возрастет. Результат последовательного укорочения волны иллюстрируется кривыми 3 и 4. Наконец, при некоторой вполне определенной частоте узел напряжения совпадет с началом линии. При этом кривая, представляющая распределение амплитуд, должна будет пройти через две точки, лежащие *на одной вертикали* (т. е. через узел и через точку, выражающую заданное входное напряжение); ясно, что кривая сразу же уйдет в бесконечность. Это и есть первый резонанс напряжения; он возникает, как мы убедились, в тех случаях, когда на длине линии l укладывается одна четверть длины волны, так что узел напряжения находится в начале линии (в конце разомкнутой линии — всегда пучность напряжения).

Будем далее повышать частоту, т. е. укорачивать волну. Теперь первый узел напряжения находится уже в пределах линии (кривая 1) и при повышении частоты передвигается вправо (к концу линии), как показано на рис. 162, *б*. Когда узел достигает середины линии, амплитуда напряжения на конце имеет наименьшее значение, равное U_0 (кривая 2). При дальнейшем укорочении волны амплитуда напряжения растет (кривая 3) и снова достигает бесконечности, когда в начале

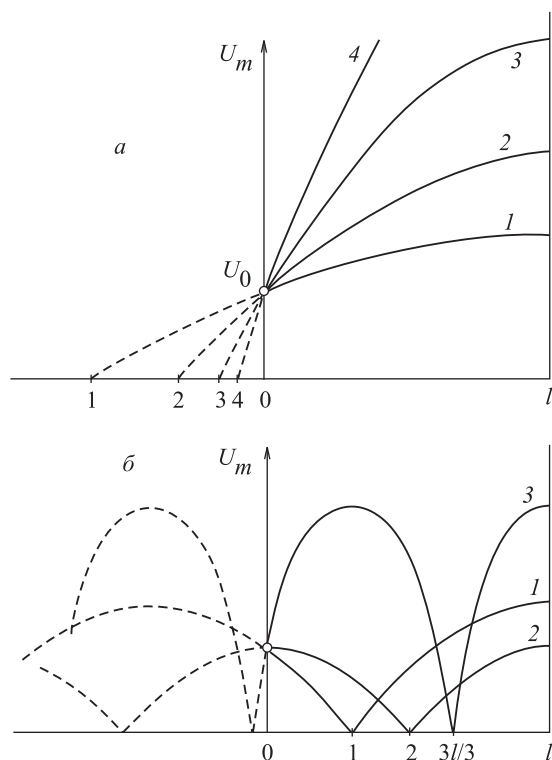


Рис. 162

линии оказывается второй узел напряжения. Первый узел находится при этом на расстоянии $(1/3)l$ от конца линии. Это есть второй резонанс напряжения; он возникает, когда на линии укладывается три четверти длины волны. При дальнейшем повышении частоты на линии будут размещаться два, три и т. д. узлов; каждый раз, когда очередной узел будет проходить через начало линии, будет очередной резонанс напряжения. Число резонансов теоретически бесконечно; практически из-за влияния затухания удается наблюдать ограниченное, но все же довольно большое число резонансов, расположенных в правильной гармонической последовательности, т. е. на кратных частотах. Это позволяет использовать линию в качестве волномера.

Можно видоизменить построение графиков рис. 162 так, чтобы избежать неудобств, связанных с уходом кривых в бесконечность при резонансе. Для этого достаточно рассматривать *отношение* амплитуды напряжения на линии к амплитуде напряжения на входе и полагать наибольшее значение этого отношения постоянным. С физической точки зрения это соответствует такой постановке опыта: для различных значений частоты (или длины волны) *подбирается* такая амплитуда

входного напряжения, которая обеспечивала бы одну и ту же наибольшую амплитуду напряжения на линии. Оказывается, что при резонансе *конечное* напряжение на линии получается при *равном нулю* напряжении на входе. Отношение обоих напряжений равно бесконечности, в чем и состоит явление резонанса напряжений (все это относится, разумеется, к линии без потерь; для реальной линии с потерями при резонансе потребуется хотя и очень малое, но все же конечное напряжение на входе). Итак, построим графики относительной амплитуды, откладывая по оси ординат величину

$$\frac{U_m}{U_0} |\cos kl| = |\cos k(l-x)|.$$

Точно так же поступим и с током: построим функцию

$$\frac{wI_m}{-U_0} |\cos kl| = |\sin k(l-x)|.$$

Как видим, распределения амплитуд тока и напряжения при таком выборе масштаба выражаются просто синусоидой и косинусоидой, что, конечно, очень удобно. Графики рис. 162 в новом масштабе примут вид рис. 163. Можно теперь построить графики относительных амплитуд тока и напряжения для последовательности резонансов напряжения в разомкнутой линии. Графики эти изображены на рис. 164. Заметим, что в разомкнутой линии на конце всегда образуется узел тока, а на входе при резонансе напряжения получается пучность тока. Кривые амплитуд напряжения приведены на рис. 164 сплошными линиями, а кривые амплитуд тока — пунктирными.

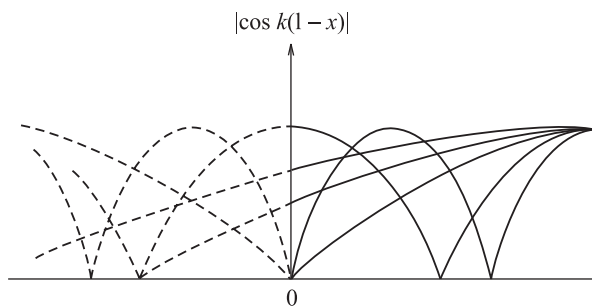


Рис. 163

Рассмотрим теперь соотношения, получаемые, когда на входе линии образуется узел тока. Это получается при

$$\sin kl = 0, \quad kl = n\pi, \quad l = n \frac{\lambda}{2} = 2n \frac{\lambda}{4},$$

т.е. когда на линии укладывается целое число полуволн или четное число четвертей волны. При этом на линии будет бесконечная ампли-

туда тока, и такого рода соотношения можно определить как *резонанс тока в линии*. Ясно, что при резонансе тока для напряжения на входе получается пучность. Входное сопротивление при резонансе тока равно бесконечности. На рис. 165 даны графики относительных амплитуд тока и напряжения для последовательности резонансов тока.

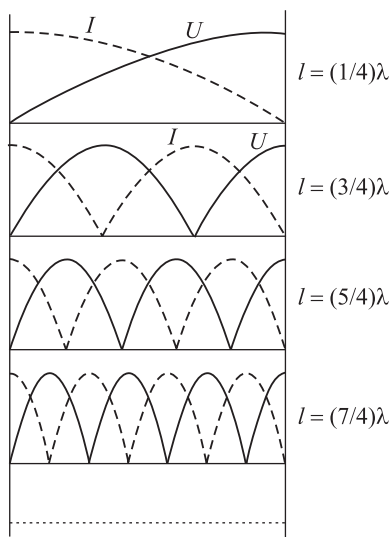


Рис. 164

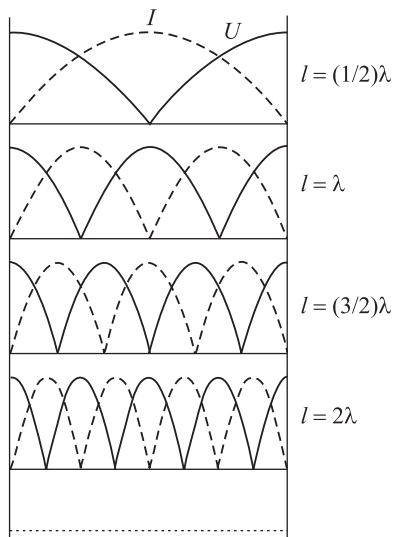


Рис. 165

Итак, резонансы тока и напряжения чередуются по мере повышения частоты. Входное сопротивление обращается попеременно то в нуль, то в бесконечность: $Z_1 \rightarrow \infty$ при резонансе тока и $Z_1 = 0$ при резонансе напряжения.

До сих пор мы разбирали соотношения для разомкнутой линии. Для короткозамкнутой линии из общих формул для тока и напряжения, положив в них $Z_2 = 0$, получим:

$$u = u_1 \frac{\sin k(l-x)}{\sin kl} = \dot{U}_0 \frac{\sin k(l-x)}{\sin kl} e^{i\omega t},$$

$$i = \frac{1}{i w} u_1 \frac{\cos k(l-x)}{\sin kl}.$$

На замкнутом накоротко конце линии всегда получаются узел напряжения и пучность тока. Зависимости для короткозамкнутой линии в некотором смысле *обратны* зависимостям для разомкнутой линии, а именно: кривые рис. 164 и 165 имеют точно такой вид, но ток и напряжение меняются местами; для короткозамкнутой линии пунктирные линии представляют относительную амплитуду напряжения, а сплошные — относительную амплитуду тока. Ничего существенно

нового при этом не обнаруживается; кривые относительных амплитуд легко нарисовать, если начинать их с конца линии в соответствии с заданным на конце условием (хх или кз).

Представляет интерес случай полного отражения при нагрузке линии на реактивное сопротивление. При этом также получается последовательность резонансов. Итак, положим $Z_2 = iX$. Подставим это значение в общую формулу (177), получим

$$u = u_1 \frac{X \cos k(l-x) + w \sin k(l-x)}{X \cos kl + w \sin kl}.$$

Пусть

$$X = C \sin k\Delta l, \quad w = C \cos k\Delta l,$$

тогда

$$u = u_1 \frac{\cos k(l-x) \sin k\Delta l + \sin k(l-x) \cos k\Delta l}{\cos kl \sin k\Delta l + \sin kl \cos k\Delta l},$$

или

$$u = u_1 \frac{\sin k(l + \Delta l - x)}{\sin k(l + \Delta l)}.$$

Таким образом, соотношения в линии, нагруженной на реактивное сопротивление, получаются такие же, как в короткозамкнутой линии, длина которой увеличена на

$$\Delta l = \frac{1}{k} \arctg \frac{X}{w}.$$

Знак добавки Δl зависит от знака X ; индуктивная нагрузка ($X > 0$) соответствует удлинению линии, емкостная нагрузка ($X < 0$) — укорочению линии. Графики для рассматриваемого случая строятся точно так же, как и для короткозамкнутой линии, но узел напряжения располагается на расстоянии Δl от конца линии, как показано для индуктивной нагрузки ($X > 0$, $\Delta l > 0$) на рис. 166. На рисунке показаны относительные амплитуды тока и напряжения при первом резонансе напряжений, для которого

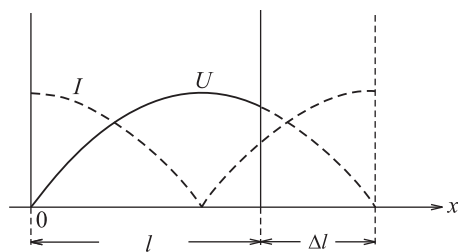


Рис. 166

индуктивной нагрузки ($X > 0$, $\Delta l > 0$) на рис. 166. На рисунке показаны относительные амплитуды тока и напряжения при первом резонансе напряжений, для которого

$$\cos k(l + \Delta l) = 0, \quad l + \Delta l = \frac{\lambda}{2}.$$

Такого рода соотношения используются в радиотехнике. Так, например, в антенну включают иногда сосредоточенную индуктивность, называемую *удлинительной катушкой*. Это позволяет получить в вертикальном проводе антенны резонанс при фактической длине провода антенны меньше наименьшей резонансной длины (равной четверти длины волны).

Из всего сказанного ясно, что отрезок линии с полным отражением может заменить колебательный контур. Но для того, чтобы определить добротность линии как резонансного контура, придется рассмотреть линию с потерями.

§ 50. Линия с потерями

Для вывода уравнения линии с потерями нужно рассмотреть бесконечно малый отрезок линии с включенными в него активными элементами (рис. 167). Составим уравнение сразу для установившегося синусоидального режима, записав соотношения

$$\left. \begin{aligned} du &= -i dZ \\ di &= -u dY \end{aligned} \right\}, \quad (179)$$

где

$$\left. \begin{aligned} dZ &= dR + i dX = (R_0 - i \omega L_0) dx = Z_0 dx \\ dY &= dG + i dB = (G_0 + i \omega C_0) dx = Y_0 dx \end{aligned} \right\}. \quad (180)$$

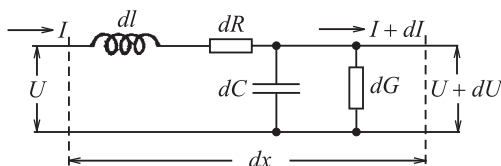


Рис. 167

Из (179) и (180) получаем

$$\frac{d^2 u}{dx^2} - Z_0 Y_0 u = 0 \quad (181)$$

и аналогичное уравнение для тока

$$\frac{d^2 i}{dx^2} - Z_0 Y_0 i = 0.$$

Решение уравнения (181) имеет вид

$$u = A e^{\alpha_1 x} + B e^{\alpha_2 x},$$

где α_1 и α_2 — корни характеристического уравнения

$$\alpha^2 - Z_0 Y_0 = 0.$$

Таким образом,

$$\alpha_1 = \gamma, \quad \alpha_2 = -\gamma,$$

где

$$\gamma = \sqrt{Z_0 Y_0}$$

— постоянная распространения.

Полученное выражение представляет собой обобщение соотношений § 44.

При отсутствии потерь

$$Z_0 = i \omega L_0, \quad Y = i \omega C_0,$$

и

$$\gamma = \sqrt{Z_0 Y_0} = \sqrt{-\omega^2 L_0 C_0} = i k.$$

Теперь же в выражение для γ нужно подставить комплексные значения Z_0 и Y_0 . Это дает

$$\begin{aligned} \gamma = \sqrt{Z_0 Y_0} &= \sqrt{(R_0 + i \omega L_0)(G_0 + i \omega C_0)} = \\ &= \sqrt{R_0 G_0 - \omega^2 L_0 C_0 + i \omega (L_0 G_0 + R_0 C_0)} \quad (182) \end{aligned}$$

и, следовательно, γ есть комплексная величина

$$\gamma = \alpha + i \beta,$$

где β и α сложным образом зависят от всех четырех погонных параметров линии. Для случая малых потерь можно воспользоваться приближением. Перепишем (182) в виде

$$\gamma = i \omega \sqrt{L_0 C_0} = \sqrt{1 - \frac{R_0}{\omega L_0} \frac{G_0}{\omega C_0} - i \left(\frac{G_0}{\omega C_0} + \frac{R_0}{\omega L_0} \right)}.$$

Будем считать

$$\frac{R_0}{\omega L_0}, \quad \frac{G_0}{\omega C_0}$$

малыми величинами и разложим радикал по формуле Ньютона, сохранив лишь члены не выше первого порядка малости. Это даст

$$\gamma = i k \left[1 - i \frac{1}{2} \left(\frac{G_0}{\omega C_0} + \frac{R_0}{\omega L_0} \right) \right]$$

или

$$\gamma = \frac{1}{2} \left(G_0 w + \frac{R_0}{w} \right) + i k = \alpha + i \beta.$$

Таким образом, мнимая часть постоянной распространения остается (в первом приближении) равной $k = \omega/v$, т. е. мы имеем дело с волнами, распространяющимися с той же скоростью

$$v = \frac{1}{\sqrt{L_0 C_0}}.$$

Вещественная часть постоянной распространения

$$\alpha = \frac{1}{2} \left(G_0 w + \frac{R_0}{w} \right)$$

выражает затухание на единицу длины линии.

Для волны в линии без отражения получаем (сохраняя один член в общем решении) бегущую *затухающую* волну напряжения, выражаемую формулой

$$u = u_1 e^{-\gamma x} = \dot{U}_0 e^{-\alpha x} e^{i(\omega t - kx)}$$

(начало координат — в начале линии).

Рассмотрим теперь конечную линию с полным отражением в качестве резонатора. Напряжение в такой линии

$$u = A e^{\gamma x} + B e^{-\gamma x},$$

а ток

$$i = -\frac{1}{Z_0} \frac{du}{dx} = -\frac{\gamma}{Z_0} (A e^{\gamma x} - B e^{-\gamma x}).$$

Входное сопротивление

$$Z_1 = \frac{u_1}{i_1} = -\frac{Z_0}{\gamma} \frac{A+B}{A-B} = -\sqrt{\frac{Z_0}{Y_0}} \frac{\frac{A}{B} + 1}{\frac{A}{B} - 1}.$$

Чтобы найти отношение A/B , воспользуемся граничным условием для конца линии. Пусть, в частности, линия замкнута на конце накопителя. Тогда

$$u_2 = A e^{\gamma l} + B e^{-\gamma l} = 0,$$

откуда

$$\frac{A}{B} = -e^{-2\gamma l}.$$

Подставив это значение в выражение для входного сопротивления, получим

$$Z_1 = \sqrt{\frac{Z_0}{Y_0}} \operatorname{th} \gamma l.$$

Для величины $\sqrt{Z_0/Y_0}$ примем приближенно

$$\sqrt{\frac{Z_0}{Y_0}} = \sqrt{\frac{R_0 + i\omega L_0}{G_0 + i\Delta C_0}} \approx \sqrt{\frac{L_0}{C_0}} = w.$$

Постоянная распространения γ , а следовательно, и входное сопротивление Z_1 зависят от частоты. Перепишем выражение для постоянной распространения, введя резонансную частоту ω_0 первого резонанса (при котором $\lambda = 4l$):

$$\gamma l = (\alpha + i k) l = \alpha l + i 2\pi \frac{l}{\lambda} = \alpha l + i \frac{\pi}{2} \frac{\lambda_0}{\lambda} = \alpha l + i \frac{\pi}{2} \frac{\omega}{\omega_0}.$$

Теперь выражение для входной проводимости

$$Y_1 = \frac{1}{Z_1} \approx \frac{1}{w} \operatorname{cth} \gamma l = \frac{1}{w} \operatorname{cth} \left(\alpha l + i \frac{\pi}{2} \frac{\omega}{\omega_0} \right)$$

разложим в ряд Тэйлора около значения $\omega = \omega_0$, т. е. запишем

$$Y_1(\omega) = Y_1(\omega_0) = Y_1'(\omega_0) \frac{\omega - \omega_0}{1!} + Y_1''(\omega_0) \frac{(\omega - \omega_0)^2}{2!} + \dots$$

Интересуясь только ближайшей окрестностью первой резонансной частоты, т. е. полагая расстройку

$$\Delta\omega = \omega - \omega_0$$

малой, можем в первом приближении ограничиться первыми двумя членами ряда. Тогда имеем:

$$Y_1(\omega_0) = \frac{1}{w} \operatorname{cth} \left(\alpha l + i \frac{\pi}{2} \right) = \frac{1}{w} \operatorname{th} \alpha l,$$

$$Y_1'(\omega_0) = i \frac{\pi}{2\omega_0} \frac{1}{w \operatorname{ch}^2 \beta l}.$$

Здесь использованы формулы:

$$\operatorname{cth}(x + y) = \frac{1 + \operatorname{th} x \operatorname{th} y}{\operatorname{th} x + \operatorname{th} y},$$

$$\operatorname{th} i x = i \operatorname{tg} x.$$

Таким образом,

$$Y_1(\omega) \approx \frac{1}{w} \operatorname{th} \alpha l + i \frac{\pi}{2} \frac{1}{w \operatorname{ch}^2 \beta l} \frac{\omega - \omega_0}{\omega_0}.$$

При небольшом затухании можно положить приближенно

$$\operatorname{th} \alpha l \approx \alpha l, \quad \operatorname{ch} \alpha l \approx 1$$

и

$$Y_1(\omega) \approx \frac{\alpha l}{w} + i \frac{\pi}{2\omega} \frac{\Delta\omega}{\omega_0}$$

или, вводя обычное обозначение относительной расстройки,

$$\varepsilon = \frac{2\Delta\omega}{\omega_0}, \quad Y_1(\omega) \approx \frac{\alpha l}{w} + i \frac{\pi}{4} \frac{\varepsilon}{w}.$$

Определим теперь модуль входного сопротивления линии:

$$|Z_1| = \frac{1}{|Y_1|} \approx \frac{1}{\sqrt{\left(\frac{\alpha l}{w}\right)^2 + \left(\frac{\pi}{4} \frac{\varepsilon}{w}\right)^2}}. \quad (183)$$

Это выражение нужно сопоставить с выражением для входного сопротивления параллельного контура, которое было выведено ранее:

$$|Z| = \rho \frac{1}{\sqrt{\varepsilon^2 + d^2}} = \frac{1}{\sqrt{\left(\frac{d}{\rho}\right)^2 + \left(\frac{\varepsilon}{\rho}\right)^2}}. \quad (184)$$

Сравнивая формулы (183) и (184), видим, что короткозамкнутая линия с потерями при небольшом затухании αl ведет себя в окрестности частоты первого резонанса аналогично параллельному контуру со следующими эквивалентными параметрами:

$$\rho_{\text{экв}} = \frac{4}{\pi} w, \quad d_{\text{экв}} = \frac{4}{\pi} \alpha l.$$

Практически важно, что величина αl может быть сделана очень малой, так что добротность эквивалентного контура достигает порядка 10^3 .

Напомним, что линия и контур ведут себя сходным образом лишь в окрестности резонансной частоты. При значительной расстройке частотные характеристики сопротивления уже сильно расходятся. На частоте $2\omega_0$ сопротивление линии имеет минимум, а на частоте $3\omega_0$ получается очередной резонанс напряжения, как показано на рис. 168.

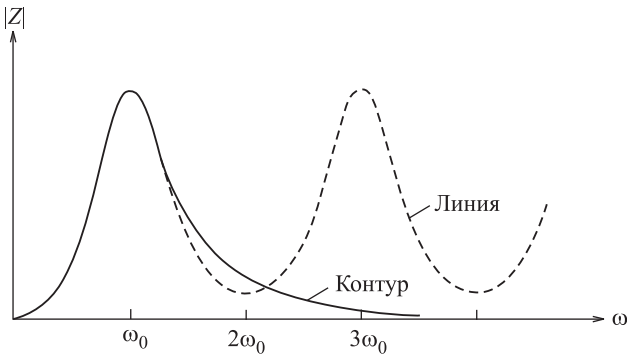


Рис. 168

Нужно добавить, что вовсе не обязательно использовать первый резонанс линии. Линия может работать в качестве резонатора при резонансах любой кратности. Так как при этом на линии получается несколько узлов, то говорят иногда, что линия при таких условиях работает «на обертоне». Этот термин заимствован из музыкальной акустики и относится, собственно говоря, к колебаниям струн, которые во многом аналогичны колебаниям в линии.

Вернемся теперь к вопросу о распределении напряжения вдоль линии при резонансе. В § 49 рассматривался абстрактный случай линии без потерь. Для такой линии при резонансе напряжение становится бесконечно большим повсюду, кроме узлов, где оно равно нулю. Рассмотрим теперь резонансные явления в реальной линии с потерями. В этом случае получаются, как и всегда при учете потерь, несколько более сложные соотношения.

Возьмем снова короткозамкнутую линию с граничными условиями

$$u(0) = u_1, \quad u(l) = 0.$$

Определяя при помощи этих условий постоянные интегрирования в общем решении

$$u = A e^{\gamma x} + B e^{-\gamma x},$$

найдем выражение для напряжения в любом сечении линии

$$u = u_1 \frac{\operatorname{sh} \gamma(l-x)}{\operatorname{sh} \gamma l}. \quad (185)$$

Будем полагать, как и выше,

$$\gamma = \alpha + i k.$$

Воспользовавшись соотношением

$$\operatorname{sh}(a + i b) \equiv \operatorname{sh} A \cos b + i \operatorname{ch} a \sin b,$$

запишем (185) в развернутом виде:

$$u = u_1 \frac{\operatorname{sh} \alpha(l-x) \cos k(l-x) + i \operatorname{ch} \alpha(l-x) \sin k(l-x)}{\operatorname{sh} \alpha l \cos kl + i \operatorname{ch} \alpha l \sin kl}. \quad (186)$$

Легко видеть, что формулы (185) и (186) представляют собой обобщение ранее приведенной (§ 49) формулы для короткозамкнутой линии без потерь, а именно

$$u = u_1 \frac{\sin k(l-x)}{\sin kl}.$$

Мы вернемся к этой формуле, если положим в (186) $\alpha = 0$.

Формула (186) может быть упрощена в предположении, что затухание мало, т. е. что $\alpha l \ll 1$. Тогда

$$u \approx u_1 \frac{\alpha(l-x) \cos k(l-x) + i \sin k(l-x)}{\alpha l \cos kl + i \sin kl}.$$

Пренебрегая поправкой на затухание, будем считать, что резонанс (как и при отсутствии потерь) возникает на частотах, для которых

$$kl = n\pi.$$

При этом

$$u \approx u_1 \frac{\alpha(l-x) \cos k(l-x) + i \sin k(l-x)}{(-1)^n \alpha l}.$$

В пучности напряжения $\sin k(l-x) \approx 1$, $\cos k(l-x) \approx 0$; таким образом, амплитуда напряжения в пучности при резонансе равна приблизительно

$$U_{m \text{ макс}} \approx U_{1m} \frac{1}{\alpha l}.$$

В узле напряжения, наоборот, $\sin k(l-x) \approx 0$, $\cos k(l-x) \approx 1$; поэтому амплитуда напряжения в узле при резонансе равна приблизительно

$$U_{m \text{ мин}} \approx U_{1m} \frac{l-x}{l}.$$

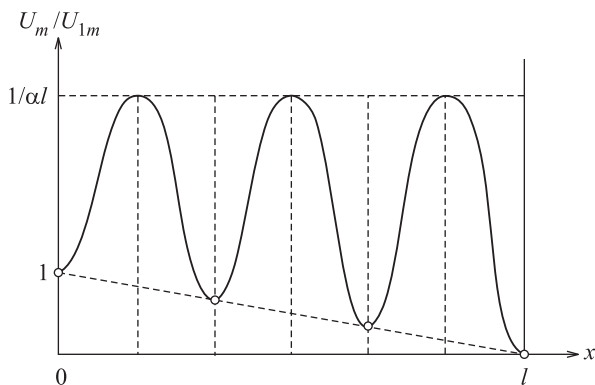


Рис. 169

Таким образом, напряжение в узлах убывает пропорционально удалению узла от начала линии. В первом узле, т. е. в начале линии,

$$U_{m \text{ мин}} = U_{1m}.$$

В последнем узле, т. е. в конце линии, по условию

$$U_{m \text{ мин}} = 0.$$

При наличии одного узла посередине ($l = \lambda$) амплитуда напряжения в нем равна

$$U_{m \text{ мин}} = \frac{1}{2} U_{1m}.$$

Для двух узлов на линии (т. е. при $l = \frac{3}{2}\lambda$) примерное распределение напряжения на линии показано на рис. 169.

Итак, в реальной линии амплитуда напряжения везде конечна; напряжение в пучности тем больше, чем меньше затухание.

§ 51. Неустановившиеся процессы в линии

Мы разбирали до сих пор вопросы, относящиеся к установившемуся синусоидальному режиму линии. Займемся теперь исследованием нестационарных режимов. Ограничимся при этом случаем линии без потерь.

Основой исследования послужит волновое уравнение

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0.$$

Применяя к этому уравнению преобразование Лапласа

$$\bar{u}(p) = p \int_0^{\infty} u(t) e^{-pt} dt,$$

получаем для *изображения* напряжения обыкновенное дифференциальное уравнение второго порядка

$$\frac{d^2 \bar{u}}{dx^2} - \frac{p^2}{v^2} \bar{u} = 0 \quad (187)$$

(начальные условия предполагаются нулевые, т.е. $u(0) = u'(0) = 0$).

Таким образом, от функции

$$u = u(x, t)$$

(оригинал) мы перешли к функции

$$\bar{u} = \bar{u}(x, p)$$

(изображение). Найдя изображение, нужно будет снова вернуться к оригиналу, т.е. определить напряжение как функцию координаты и времени. Такова обычная процедура при применении операционного метода. Рассматриваемая задача отличается от более простых задач теории цепей с сосредоточенными параметрами тем, что мы получаем для изображения *дифференциальное* уравнение, тогда как в простейших задачах изображение находится из алгебраического уравнения¹⁾. Решение уравнения (187) имеет вид

$$\bar{u} = A e^{\frac{p}{v}x} + B e^{-\frac{p}{v}x}. \quad (188)$$

Строение этого выражения показывает, что мы имеем дело с двумя возмущениями: запаздывающим и опережающим. Формула (188) есть, в сущности, не что иное, как операционная запись решения вида $f\left(t \pm \frac{x}{v}\right)$, о котором говорилось в § 43 (см. формулу (155)).

Для изображения тока из операционного соотношения

$$\bar{i} = -\frac{1}{pL_0} \frac{d\bar{u}}{dx},$$

(см формулу (153)) получаем

$$i = -\frac{1}{w} \left(A e^{\frac{p}{v}x} - B e^{-\frac{p}{v}x} \right). \quad (189)$$

Решим для начала несколько простых задач на включение постоянного напряжения. Пусть в момент $t = 0$ на вход линии подключается источник единичного постоянного напряжения с нулевым внутренним

¹⁾ Надо, впрочем, заметить, что и уравнение линии можно свести к алгебраическому уравнению, применяя преобразование Лапласа дважды: один раз интегрируя по t , второй раз — по x . Получим тогда второе изображение

$$\bar{\bar{u}}(q, p) \rightarrow \bar{u}(x, p) \rightarrow u(x, t).$$

Относительно $\bar{\bar{u}}$ уравнение будет алгебраическим, относительно $\bar{u}$ — обыкновенным дифференциальным, относительно u — уравнением в частных производных. Для наших целей применение двойного преобразования Лапласа излишне.

сопротивлением. Таким образом, на входе задается напряжение, изменяющееся во времени по закону единичной функции

$$u_1 = \sigma(t).$$

Изображение такого напряжения есть единица:

$$\bar{u}_1 = 1.$$

Пусть отражение отсутствует и решение (188) содержит только один член, выражающий возмущение, распространяющееся от начала линии, т. е.

$$\bar{u} = B e^{-\frac{p}{v}x}.$$

Граничное условие состоит в том, что при $x = 0$, $\bar{u} = 1$. Используя это условие, находим, что $B = 1$ и

$$\bar{u} = e^{-\frac{p}{v}x},$$

откуда

$$u = \sigma\left(t - \frac{x}{v}\right).$$

На рис. 170, а изображена линия и питающий ее источник. На рис. 170, б показано распределение напряжения по линии в некоторый момент t_1 ; на рис. 170, в — то же, для более позднего момента t_2 . Возмущение характеризуется отвесным фронтом, передвигающимся со скоростью v . Перед фронтом напряжение равно нулю, позади фронта оно везде равно единице. Такого вида возмущение называют иногда *единичной волной*; конечно, ничего «волнообразного» в нем нет, но термин *волна* применяется вообще в более широком смысле: он характеризует некоторое возмущение, движущееся с конечной скоростью (см. также § 43).

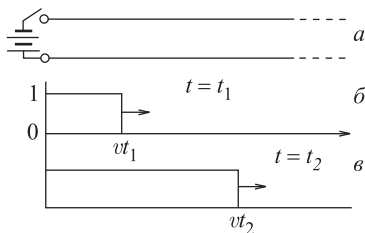


Рис. 170

Для тока получается решение в форме точно такой же волны:

$$i = \frac{1}{w} \sigma\left(t - \frac{x}{v}\right).$$

Разделив напряжение на ток, можно найти сопротивление; оно оказывается равным

$$u/i = w.$$

Это значение не зависит от x .

Теперь займемся явлениями отражения. Основное определение коэффициента отражения остается в силе, но если записать

$$\bar{p}_U = \frac{\bar{u}_{\text{отр}}}{\bar{u}_{\text{пр}}},$$

то $\bar{p}_U$ будет теперь означать некоторое операционное выражение и определяться общей формулой

$$\bar{p}_U = \frac{Z_2(p) - w}{Z_2(p) + w}, \quad (190)$$

где $Z_2(p)$ — операционное выражение сопротивления нагрузки на конце линии. Формулой (190) нам еще предстоит воспользоваться, но пока заметим, что в двух важных предельных случаях — холостого хода и короткого замыкания линии — коэффициент отражения напряжения будет равен соответственно $+1$ и -1 . Для синусоидального режима знак $(-)$ означал обращение *фазы* при отражении; теперь знак $(-)$ означает изменение *знака* (полярности) волны.

Пусть имеется разомкнутая на конце линия, для которой

$$\bar{p}_U = p_U = 1.$$

Единичная волна, дойдя до конца линии, отразится с тем же знаком и двинется в обратный путь. На рис. 171, *а* показана единичная волна, возникшая при включении напряжения на входе, на рис. 171, *б* — отраженная волна, на линии остается напряжение от предыдущего

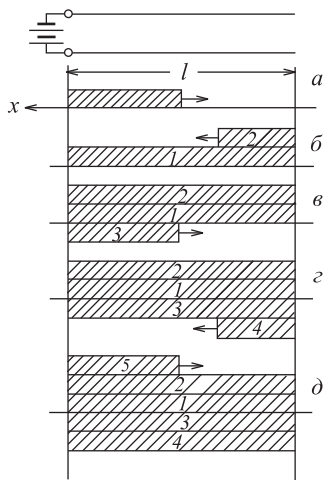


Рис. 171

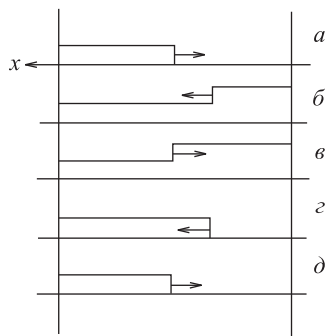


Рис. 172

прохождения волны. Когда отраженная волна дойдет до *начала* линии, то здесь снова произойдет отражение. Но нужно учесть, что мы полагаем сопротивление источника равным нулю. Поэтому от начала линии волна отражается как от короткозамкнутого конца, т.е. с коэффициентом отражения, равным -1 . Это означает перемену знака, так что после отражения от начала возникает отрицательная волна, движущаяся слева направо (рис. 171, *в*). После отражения от конца с коэффициентом отражения $+1$ волна, не меняя своего отрицатель-

ного знака, идет снова справа налево (рис. 171, $г$). После очередного отражения от начала с переменной знака появляется положительная волна (рис. 171, $д$) и т. д.

Картина распределения напряжения на линии станет более ясной, если просуммировать напряжения, оставшиеся на линии после всех предшествующих проходов положительных и отрицательных волн.

Результат такого суммирования показан на рис. 172. Здесь важно заметить, что проходящая справа налево отрицательная волна 4 (рис. 171, $г$) полностью снимает с линии напряжение (рис. 172, $г$) и восстанавливает таким образом исходное положение. Рисунок 172, $д$ повторяет картину рис. 172, $а$, и мы видим, что при описанных условиях в линии происходит *периодический* процесс, совершающийся в четыре такта, показанных на рис. 172, $а$, $б$, $в$ и $г$. Аналогичным путем можно показать, что в короткозамкнутой линии также происходит периодический процесс, но уже в два такта.

Вышеописанная картина явления получения путем последовательного физического рассуждения. Покажем теперь, как этот результат может быть получен аналитически. Перепишем еще раз общие выражения (188) и (189) для тока и напряжения:

$$\begin{aligned}\bar{u} &= A \exp\left(\frac{p}{v}x\right) + B \exp\left(-\frac{p}{v}x\right), \\ \bar{i} &= -\frac{1}{w} \left(A \exp\left(\frac{p}{v}x\right) - B \exp\left(-\frac{p}{v}x\right) \right).\end{aligned}$$

Пользуясь граничными условиями:

$$x = 0, \quad u = 1, \quad x = l, \quad \bar{i} = 0,$$

находим

$$\bar{u} = \frac{\exp\left(\frac{p}{v}(l-x)\right) + \exp\left(-\frac{p}{v}(l-x)\right)}{\exp\left(\frac{p}{v}l\right) + \exp\left(-\frac{p}{v}l\right)}. \quad (191)$$

Обычным способом нахождения оригинала является применение формулы Хевисайда (71). Но если воспользоваться ею в данном случае, то получился бы оригинал в форме ряда Фурье, что очень неудобно. Поэтому применим специализированный прием, состоящий в разложении изображения по показательным функциям. Это приведет к решению в наиболее естественной и наглядной форме.

Перепишем (191) в виде

$$\bar{u} = \left(\exp\left(-p\frac{x}{v}\right) + \exp\left(-p\frac{2l-x}{v}\right) \right) \left(1 + \exp\left(-2\frac{p}{v}l\right) \right)^{-1}$$

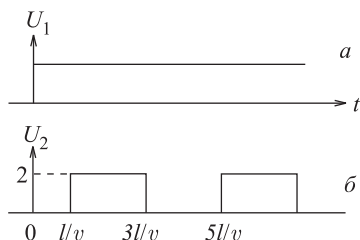
и разложим второй множитель по формуле Ньютона

$$\begin{aligned}\bar{u} &= \left(\exp\left(-p\frac{x}{v}\right) + \exp\left(-p\frac{2l-x}{v}\right) \right) \times \\ &\times \left(1 - \exp\left(-2\frac{p}{v}l\right) + \exp\left(-4\frac{p}{v}l\right) - \exp\left(-6\frac{p}{v}l\right) + \dots \right) = \\ &= \exp\left(-p\frac{x}{v}\right) + \exp\left(-p\frac{2l-x}{v}\right) - \exp\left(-p\frac{2l+x}{v}\right) - \\ &\quad - \exp\left(-p\frac{4l-x}{v}\right) + \exp\left(-p\frac{4l+x}{v}\right) + \dots\end{aligned}$$

Из этого выражения легко получается оригинал

$$\begin{aligned}u &= \sigma\left(t - \frac{x}{v}\right) + \sigma\left(t - \frac{2l-x}{v}\right) - \sigma\left(t - \frac{2l+x}{v}\right) - \\ &\quad - \sigma\left(t - \frac{4l-x}{v}\right) + \sigma\left(t - \frac{4l+x}{v}\right) + \dots\end{aligned}$$

Это и есть решение задачи, выражающее u как сумму запаздывающих относительно друг друга единичных волн. При $x = 0$ получается, как и должно быть, $u_1 = \sigma(t)$. Положив $x = l$, найдем напряжение на разомкнутом конце



$$\begin{aligned}u_2 &= 2\sigma\left(t - \frac{l}{v}\right) - 2\sigma\left(t - \frac{3l}{v}\right) + \\ &\quad + 2\sigma\left(t - \frac{5l}{v}\right) - \dots\end{aligned}$$

Рис. 173

График этой зависимости дан на рис. 173, б. На рис. 173, а изображен график входного напряжения u_1 . Эти графики полезно сопоставить с картиной рис. 172.

Описанные явления могут быть использованы для формирования прямоугольных импульсов заданной длительности. Длительность импульсов напряжения равна $\tau = \frac{2l}{v}$ и может выдерживаться с высокой степенью точности и постоянства. Эта возможность используется в современной импульсной технике.

Для усвоения своеобразия нестационарных явлений в линии рассмотрим изменения входного тока при включении на вход линии единичного напряжения. В случае разомкнутой линии имеем для входного

тока (см. вывод формулы (191))

$$\bar{i}_1 = \frac{1}{w} \frac{\exp\left(\frac{pl}{v}\right) - \exp\left(-\frac{pl}{v}\right)}{\exp\left(\frac{pl}{v}\right) + \exp\left(-\frac{pl}{v}\right)}$$

или

$$\begin{aligned} w\bar{i}_1 &= \left(1 - \exp\left(-2\frac{pl}{v}\right)\right) \left(1 + \exp\left(-2\frac{pl}{v}\right)\right)^{-1} = \left(1 - \exp\left(-2\frac{pl}{v}\right)\right) \times \\ &\times \left(1 - \exp\left(-2\frac{pl}{v}\right) + \exp\left(-4\frac{pl}{v}\right) - \exp\left(-6\frac{pl}{v}\right) + \dots\right) = \\ &= 1 - 2\exp\left(-2\frac{pl}{v}\right) + 2\exp\left(-4\frac{pl}{v}\right) - 2\exp\left(-6\frac{pl}{v}\right) + \dots, \end{aligned}$$

откуда

$$wi_1 = \sigma(t) - 2\sigma\left(t - \frac{2l}{v}\right) + 2\sigma\left(t - \frac{4l}{v}\right) - 2\sigma\left(t - \frac{6l}{v}\right) + \dots \quad (192)$$

(рис. 174). Таким образом, при включении *постоянного* напряжения получается *переменный ток*, среднее значение которого равно нулю.

Для короткозамкнутой линии граничные условия:

$$x = 0, \quad \bar{u} = 1; \quad x = l, \quad \bar{u} = 0.$$

С помощью этих условий находим из (188) и (189)

$$w\bar{i}_1 = \frac{\exp\left(\frac{pl}{v}\right) + \exp\left(-\frac{pl}{v}\right)}{\exp\left(\frac{pl}{v}\right) - \exp\left(-\frac{pl}{v}\right)},$$

или

$$w\bar{i}_1 = 1 + 2\exp\left(-2\frac{pl}{v}\right) + 2\exp\left(-4\frac{pl}{v}\right) + 2\exp\left(-6\frac{pl}{v}\right) + \dots,$$

откуда

$$wi_1 = \sigma(t) + 2\sigma\left(t - \frac{2l}{v}\right) + 2\sigma\left(t - \frac{4l}{v}\right) + \dots, \quad (193)$$

(рис. 174), т. е. входной ток нарастает неограниченно, изменяясь ступенями.

По поводу полученных результатов возникает следующий чисто физический вопрос. Как видно из формул (192) и (193) и рис. 174 и 175 в первой стадии явления (при $0 < t < 2l$) мы получаем как для разомкнутой, так и для короткозамкнутой линии одинаковый постоянный ток $i_1 = 1/w$ и, следовательно, в обоих случаях линия потребляет мощность, равную

$$I_1^2 w = 1/w$$

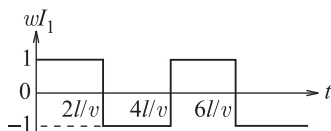


Рис. 174

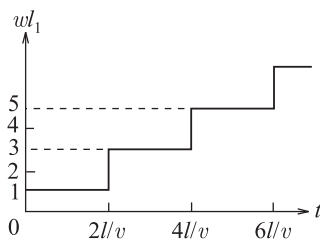


Рис. 175

(напомним, что напряжение равно единице). Спрашивается, куда девается эта мощность, принимая во внимание, что потерь в линии нет? Дело в том, что волна тока и напряжения, распространяясь по линии, постепенно *заряжает* ее энергией, запасаемой в электрическом поле распределенной емкости и магнитном поле распределенной индуктивности. Поэтому на протяжении первой стадии явления (непосредственно после включения) *всякая* линия ведет себя как активное сопротивление, равное w , независимо от длины линии и условий нагрузки на конце. Пока волна не дошла до конца и, отразившись от него, не вернулась к началу, условия нагрузки на конце (да и само наличие конца!) никак не могут повлиять на явления в начале линии. Лишь по прошествии времени $\tau = (2l)/v$, т.е. времени, за которое волна успевает пробежать по линии туда и обратно, конец дает о себе знать, и положение так или иначе определяется. Это соображение помогает многое понять.

В заключение разберем вопрос об отражении единичной волны от сосредоточенного реактивного сопротивления, включенного в конце линии. Предварительно заметим, что если при синусоидальном режиме отражение от реактивного сопротивления приводит к зависящему от частоты изменению *фазы* отраженной волны, то в нестационарном режиме при отражении от реактивного сопротивления происходит изменение *формы* волны ¹⁾.

Пусть линия нагружена на емкость C . Для операторного коэффициента отражения имеем по формуле (190)

$$\bar{p}_U = \frac{\frac{1}{pC} - w}{\frac{1}{pC} + w}.$$

Отраженная волна

$$\bar{u}_{\text{отр}} = \bar{p}_U \bar{u}_{\text{пр}}.$$

¹⁾ Конечно, можно было бы и нестационарные процессы свести к наложению стационарных и изложить весь материал этого параграфа на основе метода интеграла Фурье. Но, действуя таким образом, мы совершенно утратили бы физическую наглядность рассуждения.

Положим

$$u_{\text{пр}} = \sigma \left(t - \frac{x}{v} \right), \quad \bar{u}_{\text{пр}} = \exp \left(-\frac{p}{v} x \right)$$

(единичная волна рис. 176, а). Тогда

$$\bar{u}_{\text{отр}} = \frac{\frac{1}{pC} - w}{\frac{1}{pC} + w} \exp \left(-\frac{p}{v} x \right),$$

или

$$\bar{u}_{\text{отр}} = \left(\frac{\frac{1}{wC}}{p + \frac{1}{wC}} - \frac{p}{p + \frac{1}{wC}} \right) \exp \left(-\frac{p}{v} x \right),$$

откуда

$$u_{\text{отр}} = \left(1 - 2 \exp \left(-\frac{t - \frac{x}{v}}{wC} \right) \right) \sigma \left(t - \frac{x}{v} \right)$$

(отраженная волна распространяется справа налево).

Если добавить единичное напряжение от прямой волны, то картина распределения напряжения на линии после первого отражения будет иметь вид рис. 176, б.

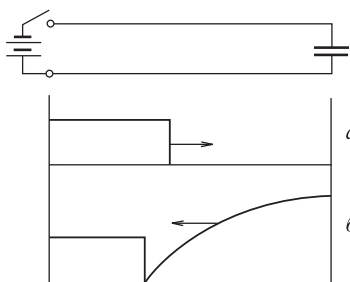


Рис. 176

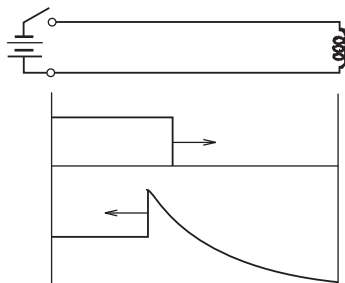


Рис. 177

Аналогичным образом решается задача для случая нагрузки на индуктивность. Получаемый результат представлен графиками рис. 177.

Эти соотношения можно толковать так: в первый момент емкость ведет себя как нулевое сопротивление; коэффициент отражения равен -1 и напряжение нулевое. При $t \rightarrow \infty$ емкость ведет себя как бесконечно большое сопротивление (т. е. для постоянного тока сопротивление емкости бесконечно велико); при этом $p_U \rightarrow +1$ и напряжение на конце стремится к удвоенному напряжению прямой волны. Для индуктивности, наоборот, сопротивление в первый момент бесконечно велико, а с течением времени стремится к нулю. Эти представления вытекают из предельных соотношений, обсужденных в конце § 28.

§ 52. Распространение импульсов по линии

В предыдущем параграфе описаны закономерности, управляющие распространением и отражением единичных волн. На этой основе могут быть исследованы явления при распространении импульсов и вообще возмущений любой формы, так как любое возмущение можно представить в виде конечной или бесконечной суммы запаздывающих единичных волн (см. § 27). Так, например, прямоугольный импульс произвольной высоты h и длительности τ может быть представлен как результат наложения двух единичных волн — положительной и запаздывающей относительно нее отрицательной — следующим образом:

$$f(x, t) = h \left[\sigma \left(t - \frac{x}{v} \right) - \sigma \left(t - \tau - \frac{x}{v} \right) \right].$$

Рассмотрим распространение прямоугольного импульса напряжения по разомкнутой линии длиной l (рис. 178). Импульс создается путем замыкания ключа в момент $t = 0$ с последующим размыканием его в момент $t = \tau$ ¹⁾.

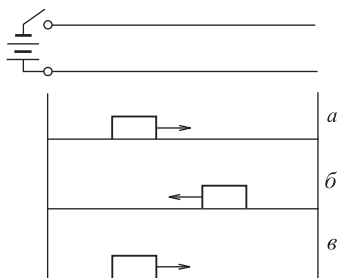


Рис. 178

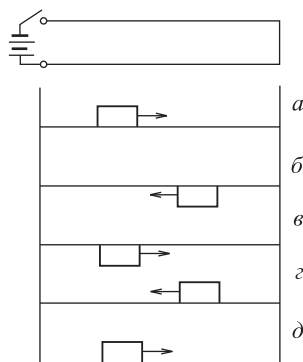


Рис. 179

После первого отражения от разомкнутого конца ($p_U = 1$) импульс бежит в обратном направлении (рис. 178, б), не изменив ни величины, ни знака, ни формы. Ключ в начале линии уже разомкнут, поэтому импульс отражается от начала линии таким же образом (рис. 178, в) и т. д. Как видим, импульс совершает периодическое движение в два такта.

Если линия замкнута накоротко (рис. 179), то импульс при отражении от конца меняет знак (рис. 179, б), сохраняет отрицательный знак при отражении от начала (рис. 179, в) и меняет знак на положительный

¹⁾ Нужно иметь в виду, что схема с батареями и ключом — не более как модель. В действительности питающий источник представляет собой то или иное электронное устройство с конечным внутренним сопротивлением, которое нужно учитывать при расчете отражений от начала.

при втором отражении от короткозамкнутого конца (рис. 179, *з*). Таким образом, и в этом случае происходит периодическое движение импульса, совершающееся в четы-ре такта.

Если линия замкнута на сопротивление $R \pm w$, то $|p_U| < 1$ и при отражении от конца линии импульс уменьшается; его высота составляет $|p_U| h$ (рис. 180). После второго отражения от конца высоты импульса будет уже $|p_U|^2 h$ (рис. 180, *з*) и т. д., так что высота импульса убывает по геометрической прогрессии и импульс постепенно затухает.

Влияние потерь в линии проявляется в том, что высота импульса непрерывно убывает в процессе его распространения по закону $e^{-\alpha x}$.

Если линия нагружена на сопротивление, имеющее реактивную составляющую, то при отражении импульс меняет форму. Искаженную форму импульса после n -кратного отражения можно найти из операционного выражения

$$\bar{u}_n = \bar{p}_U^n u_{\text{пр}},$$

где $\bar{u}_{\text{пр}}$ — изображение первоначального импульса,

$$\bar{p}_U = \frac{Z_2(p) - w}{Z_2(p) + w}.$$

Разберем теперь явления, происходящие в более сложной линии, содержащей неоднородности. Пусть в линию включено сосредоточенное сопротивление Z , как показано на рис. 181, *а*. На конце линии пусть включено сопротивление Z_2 , неравное волновому. Волна u , следующая от начала линии, отражается от неоднородности с коэффициентом отражения

$$p = \frac{Z}{Z + 2w}.$$

Образуется проходящая волна $p'u$; коэффициент прохождения равен

$$p' = 1 + p = \frac{2(Z + w)}{Z + 2w}.$$

Волна, достигающая конца линии, отражается с коэффициентом отражения

$$p_2 = \frac{Z_2 - w}{Z_2 + w}.$$

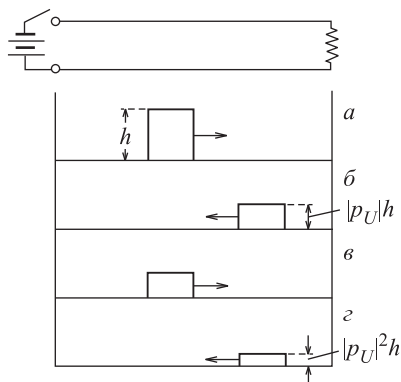


Рис. 180

Волны, движущиеся справа налево и достигающие начала линии, отражаются от него с коэффициентом отражения

$$p_1 = \frac{Z_i - w}{Z_i + w}.$$

где Z_i — внутреннее сопротивление источника, питающего линию. Коэффициент прохождения через неоднородность для волн, движущихся в обоих направлениях, один и тот же.

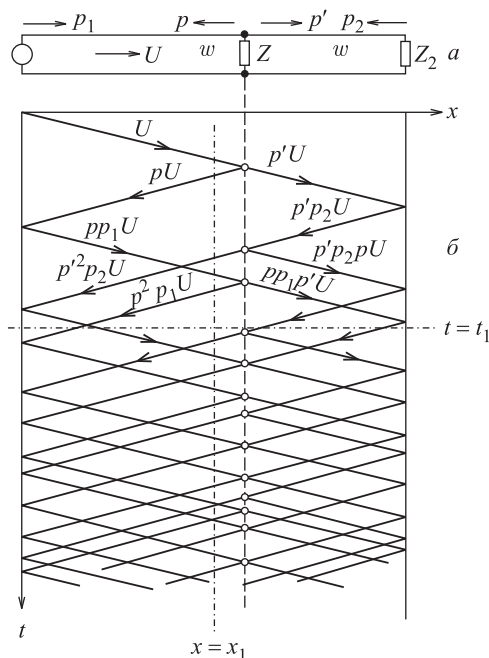


Рис. 181

Как видим, на линии образуется сложная волновая картина. Разобраться в этой картине помогает построение, которое назовем *волновой сеткой* (рис. 181, б). По горизонтальной оси откладывается координата, отсчитываемая вдоль линии. По вертикальной оси вниз откладывается время. Если при $t = 0$ из начала линии был послан импульс u , то его движение изобразится наклонной прямой (в левом верхнем углу рисунка), причем наклон прямой выражает скорость распространения¹⁾. Когда импульс достигает неоднородности, то возникают два импуль-

¹⁾ Такого рода график очень похож на график движения поездов, составляемый железнодорожным диспетчером. Разница состоит в том, что волны не знают остановок!

са: отраженный импульс pu , движущийся в обратном направлении, и проходящий импульс $p'u$, движущийся в прежнем направлении. Этот проходящий импульс, достигнув конца линии, отразится и образует импульс $p'r_2u$, который, двигаясь справа налево, достигнет неоднородности и породит снова *два* импульса; p'^2r_2u и $p'r_2pu$ и т. д.

Волновая сетка позволяет получить ответ на два вопроса, которые могут нас интересовать, а именно:

1. Каково распределение напряжения на линии в некоторый момент $t = t_1$?
2. Что происходит с течением времени в некотором сечении линии с координатой $x = x_1$?

Для ответа на первый вопрос нужно провести горизонтальную прямую с ординатой $t = t_1$. Точки пересечения этой горизонтали с наклонными прямыми сетки укажут мгновенное положение фронтов всех имеющихся на линии импульсов. Для ответа на второй вопрос нужно провести вертикальную прямую с абсциссой $x = x_1$. Точки пересечения этой вертикали с наклонными прямыми сетки укажут моменты вступления в данное сечение следующих друг за другом импульсов.

Построение рис. 181 обнаруживает одно очень важное обстоятельство, а именно: при наличии неоднородности в линии происходит *размножение* волн. Каждый раз, когда импульс встречается с неоднородностью, возникают *два* импульса — проходящий и отраженный. Поэтому с течением времени число отдельных импульсов неограниченно растет, и картина быстро усложняется настолько, что проследить тем же методом за ее дальнейшим развитием становится невозможным. Одновременно происходит уменьшение отдельных импульсов за счет многократного отражения. Можно сказать, что первоначальная электрическая *волна* превращается в мелкую и беспорядочную электрическую *рябь*. При таких условиях проследить за каждым отдельным импульсом уже нельзя, но это не значит, что исследование режима линии вообще невозможно. При описанных условиях нужно лишь изменить метод исследования. Мы имеем дело с массовым явлением, с совокупностью весьма большого числа объектов (отдельных импульсов или волн). Такого рода явления изучаются *статистическими методами*.

Мы не в состоянии различить поведение отдельных составляющих, но можем говорить о средней мощности, о статистическом спектре мощности и о других статистических величинах, и это нас вполне устраивает (по крайней мере, в такой же степени, как нас устраивает описание состояния некоторого объема газа через его давление, температуру и т. д.).

Описанные явления имеют для техники большое значение. Дело в том, что однородная линия — это абстракция. Всякая реальная линия имеет случайные неоднородности, обусловленные технологией как изготовления, так и монтажа. Возьмем к примеру коаксиальный кабель. Даже если не говорить о грубых неоднородностях, появляющихся в виде вмятин при небрежной перевозке и укладке, всякие

отклонения, например, от номинальных диаметров, от круглой формы, от соосности и т.п. представляют собой неоднородности. А такие отклонения в производстве неизбежны (хотя, разумеется, совершенствуя технологию и методы контроля, можно уменьшать допуски на конструктивные размеры). В свете сказанного выше легко представить себе последствия наличия неоднородностей. На каждой неоднородности будут образовываться отражения, которые, отразившись, в свою очередь, от предшествующих неоднородностей, образуют затухающий «хвост», следующий за основным возмущением. Этот специфический вид искажений носит название *попутного потока*¹⁾. Борьба с ним — нелегкое дело, а между тем попутный поток может сильно вредить на сколько-нибудь длинных линиях.

§ 53. Сопоставление линии и фильтра

Нижеследующие рассуждения имеют целью перекинуть мост между системами с сосредоточенными параметрами и системами с распределенными параметрами. Это имеет как теоретическое, так и практическое значения.

Различие между обоим рода систем уже неоднократно обсуждалось (см. § 30, 43, 47). Вместе с тем из всего предшествующего материала отчетливо видны и определенные черты сходства. Сейчас задача состоит в том, чтобы показать, что систему с распределенными параметрами можно рассматривать как предельный случай системы с сосредоточенными параметрами и, наоборот, что последнюю можно рассматривать как вырожденную форму первой. Речь идет о переходах такого же типа, как переход от дискретной зависимости к непрерывной, от функции к последовательности, от суммы к интегралу и т.п.

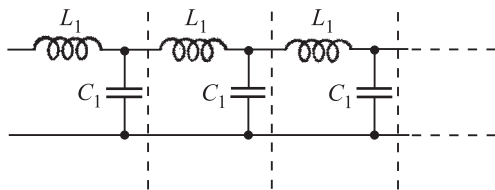


Рис. 182

Возьмем отрезок простейшего фильтра нижних частот (рис. 182) и разобьем его произвольным образом на звенья, хотя бы на Г-звенья, как намечено пунктирными линиями на рис. 182. Будем теперь уве-

¹⁾ Характер явления можно пояснить аналогией с акустическим явлением реверберации (послезвучания), обусловленным многократным и постепенно ослабевающим отражением звука от стен, пола и потолка помещения.

личивать число звеньев n , однако, так, чтобы общая индуктивность и общая емкость отрезка, т. е.

$$L = nL_1, \quad C = nC_1,$$

остались неизменными. При $n \rightarrow \infty$ индуктивность и емкость в каждом звене будут стремиться к нулю:

$$L_1 = \frac{L}{n} \xrightarrow{n \rightarrow \infty} 0, \quad C_1 = \frac{C}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Таким образом, в пределе фильтр будет представляться как совокупность бесконечного числа звеньев, каждое из которых содержит бесконечно малую индуктивность dL в последовательном плече и бесконечно малую емкость dC в параллельном плече. Но каждое такое звено есть не что иное, как элемент линии (см. § 43 и рис. 148); остается лишь положить

$$dL = L_0 dx, \quad dC = C_0 dx.$$

Итак, предельный переход от фильтра к линии состоит в неограниченном увеличении числа звеньев фильтра при соответствующем уменьшении индуктивности и емкости в каждом звене ¹⁾.

Посмотрим, что делается при таком переходе с параметрами фильтра. Граничная частота равна (122)

$$\omega_k = \frac{2}{\sqrt{L_1 C_1}}.$$

При предельном переходе L_1 и C_1 стремятся к нулю, стало быть $\omega_k \rightarrow \infty$, т. е. линия прозрачна для всех частот (речь идет, разумеется, об идеализированной линии § 43). Далее, характеристическое сопротивление для симметричных звеньев фильтра нижних частот выражается формулами:

$$(Z_c)_T = \sqrt{\frac{L_1}{C_1}} \sqrt{1 - \frac{\omega^2}{\omega_k^2}}, \quad (Z_c)_П = \sqrt{\frac{L_1}{C_1}} \frac{1}{\sqrt{1 - \frac{\omega^2}{\omega_k^2}}}.$$

Так как при предельном переходе $\omega_k \rightarrow \infty$, то

$$\lim_{\omega_k \rightarrow \infty} (Z_c)_T = \lim_{\omega_k \rightarrow \omega} (Z_c)_П = \sqrt{\frac{L_1}{C_1}} = \frac{L_0}{C_0} = w$$

и, следовательно, характеристическое сопротивление Z_c переходит в пределе в не зависящее от частоты волновое сопротивление w . Таким образом, согласование нагрузки, встречающее непреодолимые затруд-

¹⁾ Различие между системами с сосредоточенными и распределенными параметрами можно также выразить через различие в числе степеней свободы, а именно: система с сосредоточенными параметрами имеет конечное число степеней свободы, а система с распределенными параметрами — бесконечное.

нения в случае фильтра, легко осуществляется в линии: достаточно нагрузить ее активным сопротивлением $R = w$.

Проделаем предельный переход для постоянной передачи. Для фильтра (120)

$$\operatorname{th} g = \frac{\sqrt{1 + 4 \frac{Z_b}{Z_a}}}{1 + 2 \frac{Z_b}{Z_a}}.$$

Подставляя

$$Z_a = i \omega L_1, \quad Z_b = \frac{1}{i \omega C_1},$$

получим

$$\operatorname{th} g = \frac{\sqrt{1 - \frac{\omega_k^2}{\omega^2}}}{1 - \frac{\omega_k^2}{2\omega^2}} \approx i \cdot 2 \frac{\omega}{\omega_k} \rightarrow 0.$$

Так как $\operatorname{th} g$ неограниченно убывает с возрастанием ω_k , то можно положить

$$\operatorname{th} g \approx g \approx i \cdot 2 \frac{\omega}{\omega_k}$$

(приближенные равенства становятся точными в пределе). Но g — это постоянная передачи одного звена. Для цепочки из n звеньев при $n \rightarrow \infty$ постоянная передачи равна

$$\Gamma = ng = i 2n \frac{\omega}{\omega_k} = i \omega \sqrt{n L_1 n C_1} = i \omega \sqrt{LC} = i \omega \sqrt{L_0 C_0} l = i \frac{\omega}{v} l = i kl.$$

Вообще же говоря, постоянная передачи фильтра переходит в пределе в умноженную на l постоянную распространения линии.

Аналогичным образом можно продумать предельные переходы и для любых других параметров.

Теперь становится ясным, что действительную линию, т.е. систему с распределенными параметрами, можно в известных условиях и с известным приближением заменить моделью в виде системы с сосредоточенными параметрами, именно, в виде фильтра нижних частот (с потерями или без потерь, в зависимости от постановки задачи). Такое устройство называется *искусственной линией*. Каждое звено искусственной линии заменяет отрезок действительной линии длиной Δl ; отсюда следует, что индуктивность L_1 и емкость C_1 каждого звена искусственной линии определяются соотношениями:

$$L_1 = L_0 \Delta l, \quad C_1 = C_0 \Delta l.$$

То же относится и к активным параметрам (R и G), если они введены в схему.

Основной вопрос при построении искусственной линии — это вопрос о том, как выбрать длину отрезка Δl , представляемую од-

ним звеном искусственной линии. Ответ на этот вопрос таков: так как искусственная линия строится для имитации волновых процессов в действительной линии, то Δl должно быть достаточно малым по сравнению с длиной волны λ . То же требование может быть выражено и иначе: рабочие частоты должны быть достаточно малы по сравнению с граничной частотой искусственной линии. При выполнении вышеуказанного условия на искусственной линии можно наблюдать явления, во всем сходные с явлениями на действительной линии: распространение, отражение, прохождение волн, бегущие и стоячие волны и т. д. Нужно только помнить, что искусственная линия есть система с сосредоточенными параметрами, и волн в собственном смысле слова в ней нет и быть не может. Поэтому речь идет не о тождественности, а лишь о сходстве явлений в действительной и искусственной линиях. Однако при надлежащем подборе параметров и режимов приближение может быть достаточно хорошим для того, чтобы из экспериментов, произведенных на искусственной линии, сделать правильные (и не только качественные, но и количественные) заключения о работе действительной линии.

С изложенной точки зрения линия задержки, рассмотренная в § 39, есть искусственная линия; при анализе было показано, что результат улучшается при увеличении числа звеньев и соответственном повышении граничной частоты. Это и есть постепенное приближение к свойствам действительной линии, представляющей собой идеальную систему задержки, так как по ней электрические возмущения *на самом деле* распространяются с конечной скоростью, не меняя формы ¹⁾.

Ясно, чем определяется интерес к искусственным линиям. Ведь нельзя поместить не только внутри того или иного аппарата, но и в стенах лаборатории линию длиной, скажем, 100 метров. Но искусственный эквивалент такой линии может быть сделан весьма компактным.

В заключение добавим, что, основываясь на сходстве явлений в фильтре и в линии, часто говорят о фильтрах, пользуясь волновыми понятиями и терминами. Так, например, влияние несогласованности нагрузки трактуют как результат «отражения» от конца фильтра. Конечно, можно пользоваться такими представлениями, не забывая, однако, об их условном смысле.

§ 54. Волноводы; общие понятия

Волновод представляет собой трубу из проводящего материала, служащую для передачи энергии при посредстве электромагнитной волны, распространяющейся внутри волновода.

Для того чтобы разобраться в происходящих в волноводе явлениях, нужно прежде всего вспомнить, что энергия передается по ли-

¹⁾ Точнее говоря, можно создать такие условия, что при распространении по линии изменение формы возмущения будет достаточно малым.

нии электромагнитным полем, существующим в окружающем линию пространстве. Если векторы напряженности электрического поля $\mathbf{E}$ и напряженности магнитного поля $\mathbf{H}$ (будем в дальнейшем называть их для краткости «электрический вектор» и «магнитный вектор») лежат в плоскости, нормальной к направлению линии, то имеется поток энергии, направленный вдоль линии и выражаемый вектором Пойнтинга

$$\mathbf{P} = \frac{c}{4\pi} [\mathbf{E}\mathbf{H}]^1)$$

Для примера на рис. 183, изображена двухпроводная линия, по которой производится передача энергии переменным током. На рис. 183, *а* изображено продольное сечение поля. Вертикальные прямые изображают электрические силовые линии. Магнитные силовые линии образуют замкнутые контуры, охватывающие провода. Эти контуры лежат в плоскости, перпендикулярной к плоскости чертежа. Поэтому на рис. 183, *а* показаны только точки и крестики, означающие, как обычно, острия и хвосты стрелок. На рис. 183, *б* представлено сечение поля по АБ. Электрические силовые линии изображены сплошными линиями, магнитные — пунктирными. Векторы $\mathbf{E}$ и $\mathbf{H}$ в любой точке поля взаимно-перпендикулярны и лежат в плоскости поперечного сечения линии; вектор $\mathbf{P}$, следовательно, везде направлен вдоль линии (мы рассматриваем линию без потерь и без излучения).

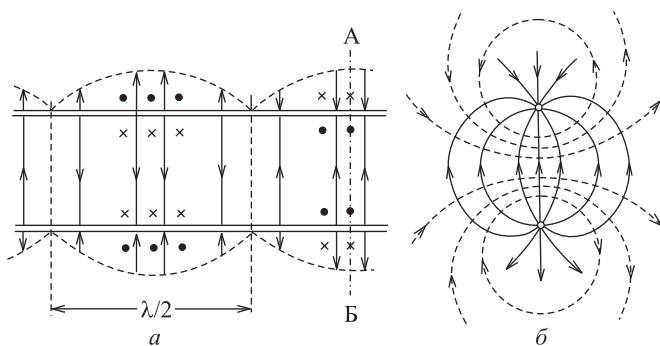


Рис. 183

На рис. 184 изображено поле коаксиальной линии в продольном (рис. 184, *а*) и поперечном (рис. 184, *б*) разрезах. Заметим, что в силу симметрии по сравнению с двухпроводной линией поле имеет более простое строение. Существенно, что поле полностью заключено внутри линии (силовые линии не выходят наружу), в чем состоит одно из основных преимуществ коаксиальной линии: она не создает внешнего

¹⁾ Здесь и далее величины электромагнитного поля выражены в абсолютных единицах.

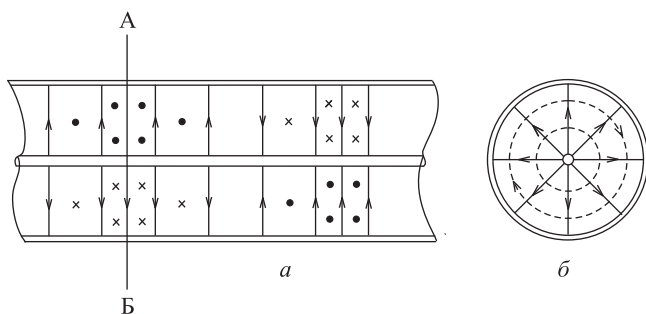


Рис. 184

(мешающего другим передачам) поля и сама защищена от воздействия посторонних полей; внешний трубчатый проводник играет роль экрана.

Переход от коаксиальной линии к волноводу, казалось бы, очень прост: достаточно устранить центральный проводник. Останется только внешний трубчатый проводник, который и можно рассматривать как волновод. Но здесь возникает чисто физическое затруднение: если устранить второй проводник, то непонятно, как должны расположиться силовые линии электрического поля. Ведь в обычной линии, независимо от конструктивного выполнения, силовые линии начинаются на одном проводе и заканчиваются на другом. Что же получится при отсутствии второго провода?¹⁾

С другой стороны, известно, что электромагнитная волна может распространяться в пространстве вообще без всяких проводов. Наличие стенок волновода есть просто ограничение некоторой части пространства, препятствующее рассеянию электромагнитной волны. Нам отнюдь не покажется странным, например, что свет может распространяться не только в открытом пространстве, но и внутри трубы.

Для того чтобы разобраться в вопросе, придется вникнуть в механизм распространения, т. е. рассмотреть явление во времени.

Представим себе коаксиальную линию, центральный провод которой в некотором месте заканчивается, так что коаксиальная линия переходит в волновод (рис. 185). Линия питается слева источником переменного напряжения; по линии слева направо бежит волна напряжения. На рис. 185, *а* изображено начало процесса. Линии электрического поля приближаются к концу центрального проводника. На рис. 185, *б* они уже дошли до конца. На рис. 185, *в* показано, что линии начинают вытягиваться и деформироваться. Слева подходят все новые группы линий (рис. 185, *г*); деформация линий, продолжающих удерживаться за конец центрального проводника, достигает крайних пределов. Наконец, происходит срыв (рис. 185, *д*): линии замыкаются *на себя* и образуют

¹⁾ Очень интересно заметить, что это затруднение казалось в свое время непреодолимым даже такому ученому, как Хевисайд.

петли, начинающиеся и кончающиеся на стенке трубы. Процесс срыва линий продолжается. На рис. 185, *e* можно видеть уже две поперечные электрического поля в волноводе (в продольном разрезе). Длина волны в волноводе Λ в общем случае отличается от длины волны в коаксиальной линии λ . Подробнее это важное обстоятельство обсуждается в следующем параграфе.

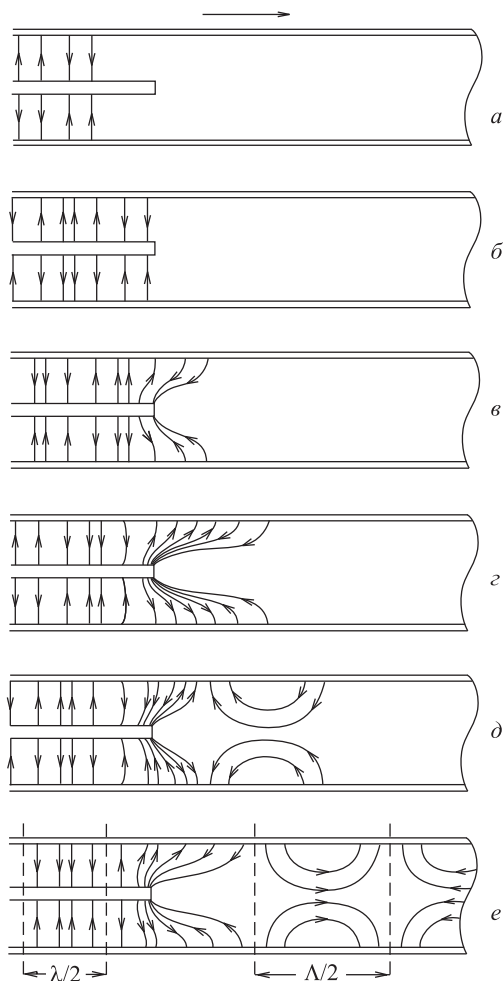


Рис. 185

Таким образом, распространение электромагнитной волны в волноводе возможно. Но более подробный анализ показывает, что распространение волны будет иметь место лишь при определенных услови-

ях и что явления распространения в волноводе имеют существенные особенности. Все это выясняется в результате исследования решений уравнений электромагнитной волны с соответствующими граничными условиями, к чему мы и переходим в следующем параграфе.

§ 55. Элементарная теория прямоугольного волновода

Основой теории волноводов, как и любого теоретического вопроса, относящегося к электромагнитному полю, являются уравнения Максвелла.

Для среды, диэлектрическая проницаемость ϵ которой равна единице, магнитная проницаемость μ — также единице, а проводимость σ — нулю (таковы значения названных параметров для пустоты и воздуха), уравнения электромагнитного поля в векторной форме имеют вид:

$$\operatorname{div} \mathbf{E} = 0, \quad (194)$$

$$\operatorname{div} \mathbf{H} = 0, \quad (195)$$

$$\frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \operatorname{rot} \mathbf{H}, \quad (196)$$

$$\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} = -\operatorname{rot} \mathbf{E}, \quad (197)$$

где $\mathbf{E}$ и $\mathbf{H}$ — электрический и магнитный векторы; c — скорость света в пустоте.

Будем рассматривать только установившийся синусоидальный режим. Для этого случая дифференцирование по времени заменяется умножением на $i\omega$, и вместо (196) и (197) можно записать:

$$i \frac{\omega}{c} \mathbf{E} = \operatorname{rot} \mathbf{H}, \quad (196')$$

$$i \frac{\omega}{c} \mathbf{H} = -\operatorname{rot} \mathbf{E}. \quad (197')$$

Из системы уравнений (196') и (197') можно получить уравнения, содержащие только $\mathbf{E}$ или только $\mathbf{H}$. Чтобы получить эти уравнения, найдем $\mathbf{H}$ из (197')

$$\mathbf{H} = -\frac{1}{ik} \operatorname{rot} \mathbf{E},$$

где $k = \omega/c$, и подставим в (196'); это даст

$$\mathbf{E} = \frac{1}{k^2} \operatorname{rot} \operatorname{rot} \mathbf{E} = \frac{1}{k^2} [\operatorname{grad} \operatorname{div} \mathbf{E} - \nabla^2 \mathbf{E}].$$

Но первый член в скобках равен нулю на основании (194). Таким образом, получается уравнение

$$\nabla^2 \mathbf{E} + k^2 \mathbf{E} = 0. \quad (198)$$

Точно такое же уравнение получается и для магнитного вектора **Н**. Уравнение (198) есть *векторная форма волнового уравнения*.

Для перехода от векторной формы к скалярной нужно прежде всего выбрать систему координат. Возьмем простейшую систему прямоугольных координат x, y, z . В этом случае можно разложить вектор **Е** на три составляющие E_x, E_y и E_z (иначе говоря, представить вектор тремя проекциями на координатные оси). Каждая составляющая поля зависит от координат и от времени. Но так как мы рассматриваем синусоидальный колебательный режим, то закон изменения каждой составляющей во времени выражается одним и тем же множителем $e^{i\omega t}$. Опустив этот множитель, мы придадим величинам E_x, E_y и E_z смысл *комплексных амплитуд* составляющих электрического вектора. Именно в этом смысле и понимаются в дальнейшем составляющие поля.

Итак, введем составляющие электрического вектора по направлениям трех координатных осей. Уравнение (198) распадается при этом на три одинаковых по форме уравнения относительно каждой из составляющих. Символ ∇^2 («набла квадрат») будет означать просто сумму вторых частных производных по всем трем координатам. Вот эти уравнения:

$$\left. \begin{aligned} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) E_x + k^2 E_x &= 0 \\ \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) E_y + k^2 E_y &= 0 \\ \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) E_z + k^2 E_z &= 0 \end{aligned} \right\}. \quad (199)$$

Решения этих уравнений являются, конечно, функциями трех переменных x, y и z . Но в данном случае можно искать решение в виде произведения трех функций, каждая из которых есть функция только одного аргумента, т. е.

$$E = X(x) Y(y) Z(z). \quad (200)$$

Дальнейший путь таков: решение (200) подставляется в (199); получаются отдельные уравнения для функций X, Y и Z ; эти уравнения интегрируются; в общие интегралы вводятся граничные условия.

Но можно значительно сократить математические выкладки и одновременно продвинуться в понимании проблемы, если уже на данном этапе опираться на физические представления.

Прежде всего обсудим граничные условия. Будем полагать, что стенки волновода изготовлены из идеально проводящего материала. В таком случае касательная (т. е. направленная параллельно стенке) составляющая вектора электрического поля около самой стенки долж-

на равняться нулю. Иными словами, электрический вектор должен быть нормален к стенке волновода на ее поверхности.

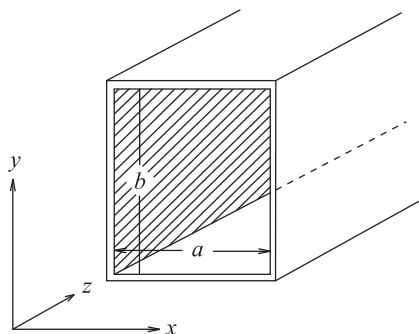


Рис. 186

Возьмем волновод прямоугольного сечения и расположим оси координат, как показано на рис. 186; ось z направим вдоль волновода. Размеры сечения волновода обозначим через a и b . В соответствии со сказанным выше граничные условия запишутся теперь следующим образом:

$$\left. \begin{aligned} E_x &= 0 \quad \text{при } y = 0, \quad y = b \\ E_y &= 0 \quad \text{при } x = 0, \quad x = a \\ E_z &= 0 \quad \text{при } x = 0, \quad x = a, \quad y = 0, \quad y = b \end{aligned} \right\}. \quad (201)$$

Мы предполагаем, что волновод неограничен в направлении оси z («бесконечный» волновод). В таком случае в направлении оси z , вдоль волновода, может (мы выясним, при каких условиях) распространяться бегущая волна. Эта волна и выражается функцией $Z(z)$ в формуле (200); можно записать

$$Z(z) = C e^{-\gamma z},$$

где γ — неизвестная пока постоянная распространения. Функции X и Y представляют зависимость амплитуды от координат точки в сечении волновода. В силу полной симметрии уравнений (199) относительно x, y и z функции X и Y также представляют волны, но распространяющиеся поперек волновода, т. е. перпендикулярно к стенкам, а так как стенки бесконечно проводящие, то в направлении осей x и y могут существовать только *стоячие* волны, амплитуды которых выражаются тригонометрическими функциями координат. Для касательных составляющих на стенках получаются узлы, для нормальных — пучности. Исходя из этих общих представлений, можем сразу записать для уравнений (199) решения вида (200), удовлетворяющие граничным

условиям (201) следующим образом:

$$\left. \begin{aligned} E_x &= A_x \cos m\pi \frac{x}{a} \sin n\pi \frac{y}{b} e^{-\gamma z} \\ E_y &= A_y \sin m\pi \frac{x}{a} \cos n\pi \frac{y}{b} e^{-\gamma z} \\ E_z &= A_z \sin m\pi \frac{x}{a} \sin n\pi \frac{y}{b} e^{-\gamma z} \end{aligned} \right\}. \quad (202)$$

Здесь m и n — любые положительные целые числа (некоторые ограничения выяснятся позднее), A_x , A_y и A_z — произвольные постоянные. Эти постоянные определяются, если задать дополнительные граничные условия, например поле во входном сечении волновода. Иначе говоря, постоянные A_x , A_y и A_z зависят от условий возбуждения волновода, и пока не будем их определять.

Найдем вторые производные E_x по координатам:

$$\frac{\partial^2 E_x}{\partial x^2} = -\left(\frac{m\pi}{a}\right)^2 E_x, \quad \frac{\partial^2 E_x}{\partial y^2} = -\left(\frac{n\pi}{b}\right)^2 E_x, \quad \frac{\partial^2 E_x}{\partial z^2} = \gamma^2 E_x.$$

Подставляя эти значения в первое уравнение (199), найдем

$$\gamma = \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 - k^2}.$$

Постоянная распространения будет мнимой, т.е. волна будет распространяться вдоль волновода без затухания, при условии

$$k > \pi \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2},$$

или

$$\omega > \pi c \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2} = \omega_k.$$

При этом

$$\gamma = ik \sqrt{1 - \frac{\omega_k^2}{\omega^2}} = i\beta.$$

При $\omega < \omega_k$ постоянная распространения вещественна, т.е. имеется затухание. Итак, волновод обладает свойствами фильтра верхних частот: в нем распространяются без затухания только колебания с частотой выше некоторой граничной частоты ω_k . Граничная частота ω_k , как показывает формула для ω_k тем выше, чем меньше размеры волновода.

Рассмотрим соотношения в полосе прозрачности, т.е. при $\omega > \omega_k$. Постоянная распространения при этом мнима, т.е. $\gamma = i\beta$. Это означает, что вектор $\mathbf{E}$, не изменяя своей величины при распространении волны вдоль волновода, изменяется по фазе. Фазный сдвиг равен

$$\varphi = \beta z = 2\pi \frac{z}{\Lambda},$$

где Λ — длина волны в волноводе. Из предыдущего выражения следует, что

$$\Lambda = \frac{2\pi}{\beta}.$$

Подставляя значение β , получаем

$$\Lambda = \frac{2\pi}{\sqrt{k^2 - \left[\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2\right]}} = \frac{2\pi}{\sqrt{k^2 - \frac{\omega_k^2}{c^2}}},$$

или, так как $\Lambda = 2\pi c/\omega$,

$$\Lambda = \frac{\lambda}{\sqrt{1 - \frac{\omega_k^2}{\omega^2}}}.$$

С другой стороны, очевидно, что длина волны пропорциональна скорости распространения, т. е.

$$\frac{\Lambda}{\lambda} = \frac{v}{c},$$

откуда

$$v = \frac{c}{\sqrt{1 - \frac{\omega_k^2}{\omega^2}}}, \quad (203)$$

где v — скорость распространения волны в волноводе.

Важное соотношение (203) показывает, что скорость распространения волны в волноводе зависит от частоты; что она всегда *больше* скорости света, т. е. скорости распространения электромагнитной волны в открытом пространстве; что скорость распространения v обращается в бесконечность при граничной частоте ω_k ; что скорость v становится мнимой при $\omega < \omega_k$. Последнее означает, что при частотах, ниже граничной, волновое состояние в волноводе вообще не может быть осуществлено; имеется лишь синфазное и быстро затухающее по длине волновода колебание вектора $\mathbf{E}$. Что касается скоростей, превышающих скорость света, то нужно пояснить, что речь идет о *фазовой* скорости. Фазовая скорость есть скорость перемещения в пространстве точек, в которых наблюдается определенная фаза колебания (например, максимум или нуль). Фазовая скорость есть понятие кинематическое; никаких физических ограничений на нее не накладывается. *Групповая* же скорость или скорость распространения энергии ни при каких обстоятельствах не может превзойти скорости света. Зависимость фазовой скорости от частоты колебаний носит в физике название *дисперсии*.

Наряду с граничной частотой ω_k часто пользуются также граничной или критической длиной волны

$$\lambda_k = \frac{2\pi c}{\omega_k} = \frac{2}{\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2}}. \quad (204)$$

Как видим, граничная длина волны имеет порядок периметра сечения волновода. Отсюда следует, что волноводы (приемлемых размеров) могут применяться лишь при очень высоких частотах, для которых длина волны имеет порядок дециметра и меньше (что соответствует частотам порядка 10^3 МГц и выше). Таким образом, волноводы применяются только для передачи энергии электромагнитных колебаний на сверхвысоких частотах; они составляют непрерывную принадлежность современной аппаратуры СВЧ.

§ 56. Картина поля в прямоугольном волноводе; типы волн

Для построения картины электромагнитного поля в прямоугольном волноводе нужно найти составляющие магнитного вектора. Для этого послужит уравнение (197')

$$\mathbf{H} = -\frac{1}{ik} \operatorname{rot} \mathbf{E}.$$

В проекциях на координатные оси x, y и z это уравнение может быть записано в виде следующих трех уравнений:

$$\left. \begin{aligned} H_x &= -\frac{1}{ik} \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \right) \\ H_y &= -\frac{1}{ik} \left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \right) \\ H_z &= -\frac{1}{ik} \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \right) \end{aligned} \right\}, \quad (205)$$

где E_x, E_y и E_z — составляющие электрического вектора, определяемые формулами (202). Выполнив дифференцирование, получаем

$$\left. \begin{aligned} H_x &= B_x \sin m\pi \frac{x}{a} \cos n\pi \frac{y}{b} e^{-\gamma z} \\ H_y &= B_y \cos m\pi \frac{x}{a} \sin n\pi \frac{y}{b} e^{-\gamma z} \\ H_z &= B_z \cos m\pi \frac{x}{a} \cos n\pi \frac{y}{b} e^{-\gamma z} \end{aligned} \right\}, \quad (206)$$

где B_x, B_y и B_z — постоянные, не содержащие новых неизвестных величин. Например,

$$B_x = -\frac{1}{ik} \left(\frac{n\pi}{b} A_z + \gamma A_y \right)$$

и т. д. Теперь можно уже рассматривать строение поля для волн различных типов.

Рассмотрим прежде всего волны, отличающиеся тем, что магнитный вектор лежит в плоскости поперечного сечения волновода, т. е. волны, у которых составляющая магнитного вектора, направленная вдоль

волновода, равна нулю, $H_z = 0$. Волны такого типа называются *поперечно-магнитными* и обозначаются символом TM (другое название этих волн — электрические; обозначение E). Для различных значений множителей m и n обозначение волны записывается в виде TM_{mn} . Индексы m и n определяют сложность поля. Простейшие волны имеют наименьшие значения m и n . Заметим, что для волн TM ни m , ни n не могут равняться нулю. Это следует из того, что при $m = 0$ составляющая $H_x = 0$, а при $n = 0$ $H_y = 0$, как это следует из (206). Так как для волн TM , по определению, $H_z = 0$, выходит, что при m или n , равном нулю, остается только одна неравная нулю составляющая магнитного вектора — либо H_x , либо H_y . Но это невозможно, так как магнитные силовые линии должны образовывать замкнутые контуры в плоскости xy , а для этого необходимо наличие обеих составляющих — H_x и H_y . Итак, волны TM_{0n} и TM_{m0} не могут существовать. Следовательно, простейшим видом волны TM является волна TM_{11} ($m = n = 1$). Рассмотрим поле этой волны.

Выпишем формулы для составляющих электрического и магнитного векторов в некотором определенном сечении с координатой $z = z_1$. Общий множитель $e^{-\gamma z_1}$, представляющий собой некоторую комплексную постоянную, пока опустим:

$$\left. \begin{aligned} E_x &= A_x \cos \pi \frac{x}{a} \sin \pi \frac{y}{b} \\ E_y &= A_y \sin \pi \frac{x}{a} \cos \pi \frac{y}{b} \\ E_z &= A_z \sin \pi \frac{x}{a} \sin \pi \frac{y}{b} \\ H_x &= B_x \sin \pi \frac{x}{a} \cos \pi \frac{y}{b} \\ H_y &= B_y \cos \pi \frac{x}{a} \sin \pi \frac{y}{b} \\ H_z &= 0 \end{aligned} \right\}. \quad (207)$$

Для того чтобы получить общее представление о строении поля, посмотрим, каковы значения составляющих на средних линиях сечения, т. е. при $x = a/2$, $y = b/2$. Для электрического вектора:

$$\left. \begin{aligned} x &= \frac{a}{2}, & y &= \frac{b}{2} \\ E_x &= 0, & E_x &= A_x \cos \pi \frac{x}{a} \\ E_y &= A_y \cos \pi \frac{y}{b}, & E_y &= 0 \\ E_z &= A_z \sin \pi \frac{y}{b}, & E_z &= A_z \sin \pi \frac{x}{a} \end{aligned} \right\}.$$

Графики этих зависимостей изображены на рис. 187. На рис. 188 изображены направления электрического вектора на средних линиях

сечения волновода $x = a/2$ и $y = b/2$. На вертикальной линии $x = a/2$ электрический вектор состоит из двух составляющих: E_y и E_z . Посредине, т.е. при $y = b/2$, $E_y = 0$ и электрический вектор определяется только одной составляющей E_z . При возрастании y составляющая E_z убывает, а составляющая E_y растет: вектор $\mathbf{E}$ изменяет свой наклон, пока около стенки (при $y = b$) он не станет перпендикулярно к стенке. Составляющая E_z в этом месте обращается в нуль. Аналогично строится вектор $\mathbf{E}$ вдоль горизонтали $y = b/2$. Здесь электрический вектор образован двумя составляющими $-E_x$ и E_z .

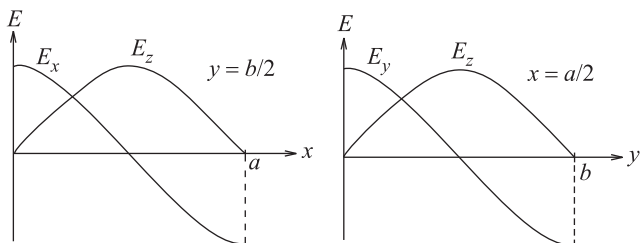


Рис. 187

Формулы (207) позволяют построить вектор $\mathbf{E}$ для любой точки поперечного сечения волновода.

Картина магнитного поля проще. На средних линиях сечения:

$$x = \frac{a}{2}, \quad y = \frac{b}{2},$$

$$H_x = B_x \cos \pi \frac{y}{b}, \quad H_y = B_y \cos \pi \frac{x}{a}.$$

Остальные составляющие равны нулю. Магнитный вектор лежит в плоскости поперечного сечения волновода; его величина и направление на средних линиях показаны на рис. 189.

Теперь необходимо выяснить зависимость всех составляющих электрического и магнитного векторов от координаты z . Общий характер этой зависимости определяется множителем $e^{-\gamma z}$. Мы имеем дело с периодическим изменением вдоль волновода всех составляющих поля. Но фаза этого изменения может оказаться различной для различных составляющих, так как постоянные A_x, A_y, A_z, B_x, B_y в общем случае — комплексные величины. Все перечисленные постоянные могут быть выражены через какую-либо одну; выведем соответствующие выражения.

Прежде всего заметим, что из условия $H_z = 0$ на основании формул (205) сразу получаем

$$\frac{\partial E_y}{\partial x} = \frac{\partial E_x}{\partial y}.$$

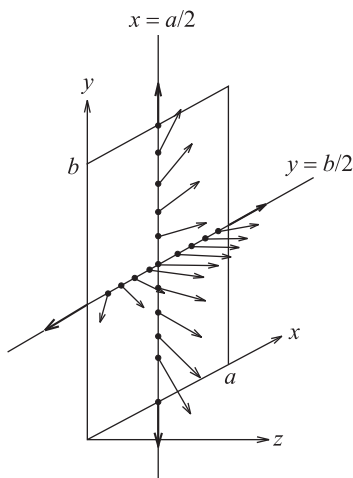


Рис. 188

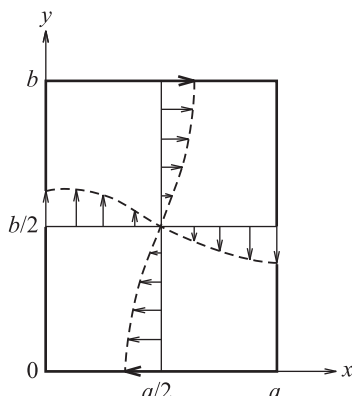


Рис. 189

Составляя и приравнявая производные, находим

$$\frac{m}{a} A_y = \frac{n}{b} A_x.$$

Далее, из уравнения

$$\mathbf{E} = \frac{1}{i k} \operatorname{rot} \mathbf{H},$$

принимая во внимание, что $H_z = 0$, получим в проекциях:

$$\left. \begin{aligned} E_x &= -\frac{1}{i k} = \frac{\partial H_y}{\partial z} \\ E_y &= \frac{1}{i k} = \frac{\partial H_x}{\partial z} \\ E_z &= \frac{1}{i k} = \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right) \end{aligned} \right\}. \quad (208)$$

Дифференцируя (206), найдем (принимая во внимание, что в полосе прозрачности постоянная распространения — мнимая величина, т. е. $\gamma = i\beta$) из первых двух строк (208):

$$B_y = \frac{k}{\beta} A_x, \quad B_x = -\frac{k}{\beta} A_y$$

и, наконец, из третьей строки (208)

$$A_z = -i \frac{\pi}{k} \left(\frac{m}{a} B_y - \frac{n}{b} B_x \right) = -i \frac{\pi}{\beta} \left(\frac{m}{a} A_x + \frac{n}{b} A_y \right).$$

Таким образом, все постоянные могут быть выражены через одну, например A_x . Пусть A_x — вещественна (выбор начальной фазы произволен). В таком случае A_y, B_x, B_y также вещественны, а постоянная

A_z — мнимая. Это значит, что составляющие E_x, E_y, H_x, H_y в любом сечении волновода имеют одинаковую фазу, а составляющая E_z сдвинута по отношению к ним по фазе на $\pi/2$. Другими словами, если в некотором сечении волновода E_x и E_y имеют максимальное значение, то $E_z = 0$, и, наоборот, там, где E_z имеет максимум, $E_x = E_y = 0$ ¹⁾.

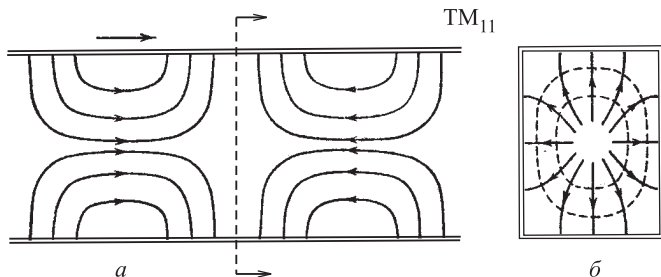


Рис. 190

Этих соображений уже достаточно, чтобы построить всю картину поля для волны TM_{11} , как показано на рис. 190. Поле изображено в продольном (рис. 190, а) и поперечном (рис. 190, б) сечениях волновода. Сплошные линии изображают электрические силовые линии, а пунктирные — магнитные. Картина, представленная на рис. 190, а, движется вдоль волновода (в направлении оси z) со скоростью v .

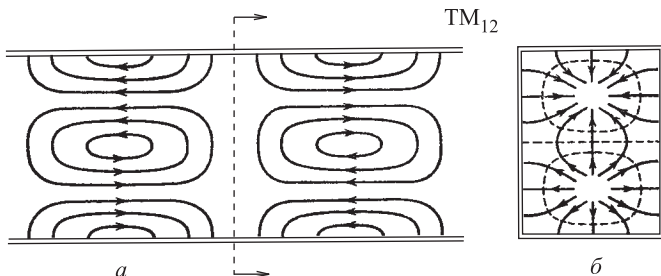


Рис. 191

Следующая по сложности поля будет волна TM_{12} или TM_{21} (перестановка индексов означает всего лишь перемену местами осей x и y или поворот сечения волновода на 90°). Для волны TM_{12} поле имеет структуру, как бы составленную из двух полей TM_{11} по направлению оси y , как показано на рис. 191.

Волна TM_{22} имеет поле, показанное на рис. 192, только в поперечном разрезе (продольный разрез такой же, как на рис. 191). Это

¹⁾ Рисунки 187 и 188 относятся к сечению, в котором ни одна из составляющих электрического вектора не равна нулю. Очевидно, что в этом сечении ни одна из составляющих не достигает наибольшего значения.

поле можно себе представить составленным из двух полей TM_{12} или из четырех полей TM_{11} . Кстати сказать, намеченные пунктиром плоскости эквипотенциальны. Поэтому можно расположить в этих плоскостях идеально проводящие перегородки, не изменив условий распространения волны и строения ее поля. Аналогично получаются поля возрастающей сложности для волн более высоких порядков.

Обратимся теперь к другому типу волн — так называемым *поперечно-электрическим* волнам TE (другое название: магнитные волны; обозначение — H). Эти волны отличаются тем, что электрический вектор лежит в плоскости поперечного сечения волновода; иначе говоря, $E_z = 0$. Для волн TE возможно существование волн при m или n , равном нулю. Дело в том, что при этом остается только одна составляющая электрического вектора — E_x или E_y , но для электрического поля это возможно: электрическое поле будет в этом случае представляться прямыми параллельными силовыми линиями, начинающимися на одной стенке волновода и оканчивающимися на противоположной.

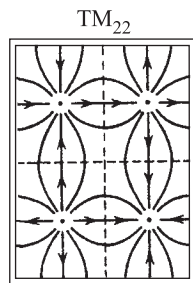


Рис. 192

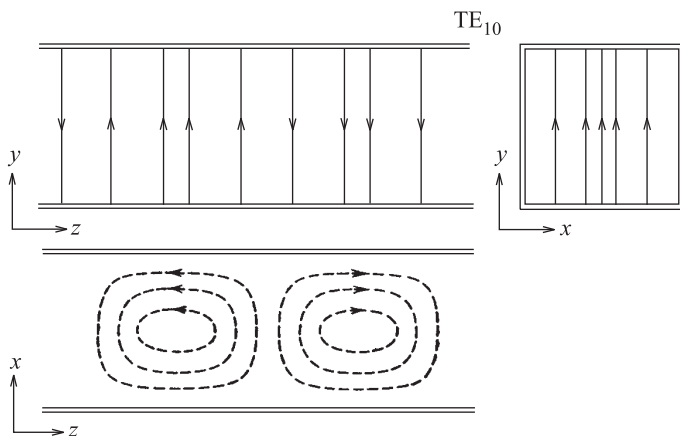


Рис. 193

Составляющие поля волны TE_{10} в некотором сечении (если опустить множитель $e^{-\gamma z}$, аналогично тому, как это сделано в формулах (207)) выражаются следующим образом:

$$\left. \begin{aligned} E_x &= 0, & H_x &= B_x \sin \pi \frac{x}{a} \\ E_y &= A_y \sin \pi \frac{x}{a}, & H_y &= 0 \\ E_z &= 0, & H_z &= B_z \cos \pi \frac{x}{a} \end{aligned} \right\}. \quad (209)$$

Как видим, магнитный вектор лежит в плоскости xz . Поэтому картина поля для волны TE_{10} дана на рис. 193 в трех проекциях. На рис. 194 показано поле волны TE_{20} , которое, как и выше, можно составить из двух полей TE_{10} .

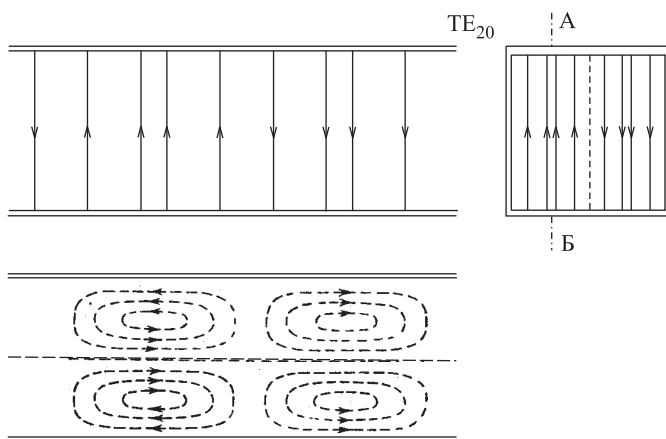


Рис. 194

Несколько сложнее картина поля для волны TE_{11} . Составляющие поля выражаются в этом случае формулами:

$$\left. \begin{aligned} E_x &= A_x \cos \pi \frac{x}{a} \sin \pi \frac{y}{b}, & H_x &= B_x \sin \pi \frac{x}{a} \cos \pi \frac{y}{b} \\ E_y &= A_y \sin \pi \frac{x}{a} \cos \pi \frac{y}{b}, & H_y &= B_y \cos \pi \frac{x}{a} \sin \pi \frac{y}{b} \\ E_z &= 0, & H_z &= B_z \cos \pi \frac{x}{a} \cos \pi \frac{y}{b} \end{aligned} \right\}.$$

Построение по этим формулам дает картину, показанную на рис. 195. Заметим, что, основываясь на симметрии картины поля, можно, разделив мысленно сечение волновода диагональными плоскостями, как показано на рис. 195 справа, получить картину поля для одной из простейших волн в волноводе треугольного сечения. Прямое решение этой задачи было бы довольно затруднительно.

Наряду с поперечными волнами возможны также *продольно-электрические* волны LE и *продольно-магнитные* волны LM . Как показывает название, волны LE характеризуются тем, что электрический вектор лежит в плоскости продольного сечения волновода (т.е. E_x или E_y равно нулю). Для волн LM вектор магнитного поля лежит в продольной плоскости (т.е. H_x или H_y равно нулю). В частности, волна TE_{10} может рассматриваться как волна LM , так как для нее

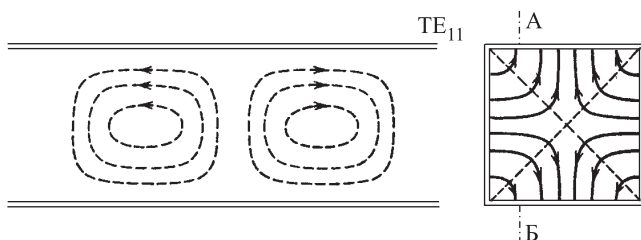


Рис. 195

$H_y = 0$ и магнитный вектор лежит в плоскости xz . Более сложные продольные волны рассматривать не будем.

В заключение вернемся к вопросу о граничной длине волны. Раньше (§ 55) было получено следующее соотношение (204):

$$\lambda_k = \frac{2}{\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2}}.$$

Это выражение показывает, что граничная длина волны зависит не только от размеров волновода, но и от индексов m и n ; чем сложнее волна, т. е. чем больше m и n , тем меньше λ_k (и тем выше граничная частота ω_k). Эту зависимость можно представить в более наглядной форме, если перейти к квадратному волноводу, т. е. положить

$$b = a.$$

В этом случае

$$\lambda_k = a \frac{2}{\sqrt{m^2 + n^2}}.$$

В табл. 11 дано отношение граничной длины волны λ к стороне квадрата a при различных значениях m и n :

Таблица 11

mn	01	11	02	12	22	03	13	23	04
λ_k/a	2,0	1,4	1,0	0,9	0,72	0,67	0,63	0,55	0,50

Таким образом, при заданной граничной частоте нужно брать тем большие размеры волновода, чем сложнее волна. В этом заключается одна из причин того, что на практике предпочитают работать с простейшими волнами. С другой стороны, интересно, что если работать на частоте немного выше граничной, то волны более сложной структуры, даже если они и возникнут, будут быстро затухать, так как для них рабочая частота окажется ниже граничной.

представляет i -й по порядку корень уравнения

$$J_n(x) = 0.$$

В табл. 12 даны корни бесселевых, функций, т. е. величины ν_{ni}

Таблица 12

Порядок функций n	Порядок корня i				
	1	2	3	4	5
0	2,41	5,52	8,65	11,79	14,93
1	3,83	7,02	10,17	13,32	16,47
2	5,14	8,42	11,60	14,80	17,96
3	6,38	9,76	13,02	16,22	19,41

Граничная длина волны для волн $TM_{(ni)}$ в круглом волноводе определяется по формуле

$$\lambda_k = 2\pi \frac{a}{\nu_{ni}}.$$

Простейшую волну получим при $n = 0$, $i = 1$. Из таблицы находим $\nu_{01} = 2,41$. Далее, из теории бесселевых функций известно, что

$$J'_0(x) = -J_1(x).$$

Таким образом, для волны рассматриваемого типа получаем из (210) следующие значения составляющих:

$$\left. \begin{aligned} E_r &= -A_r J_1\left(\frac{2,41}{a}r\right), & H_\varphi &= -B_\varphi J_1\left(\frac{2,41}{a}r\right) \\ E_z &= A_z J_0\left(\frac{2,41}{a}r\right), & E_\varphi &= H_r = H_z = 0 \end{aligned} \right\}. \quad (211)$$

Так как составляющая E_φ отсутствует, то в плоскости поперечного сечения имеется чисто радиальное электрическое поле. Отсутствие радиальной составляющей H_r указывает на то, что магнитные силовые линии образуют окружности, лежащие в плоскости поперечного сечения. Волна такого типа есть волна $TM_{(01)}$ ¹⁾ ее часто называют волной E_0 . Картина поля показана на рис. 197; как видно, волна $TM_{(01)}$ — это та самая волна, возникновение которой рассматривалось в § 54.

Следующая по сложности — волна $TM_{(02)}$. Поле этой волны отличается от предыдущего тем, что составляющие как электрического, так и магнитного поля обращаются в нуль при определенном значении $r < a$ (т. е. на некоторой цилиндрической поверхности внутри волновода),

¹⁾ Индекс ni для круглого волновода — это не те же самые, что индекс mn для прямоугольного. Во избежание путаницы будем ставить в скобки индексы при обозначениях волн в круглых волноводах.

и меняют знак. При волне $TM_{(03)}$ составляющие поля дважды меняют знак, однако общая структура поля остается неизменной для всех волн вида $TM_{(0i)}$. Формулы (211) также остаются без изменения, не считая того, что нужно каждый раз подставлять соответствующее значение корня ν_{0i} : 2,41 для $TM_{(01)}$, 5,52 для $TM_{(02)}$, 8,65 для $TM_{(03)}$ и т. д.

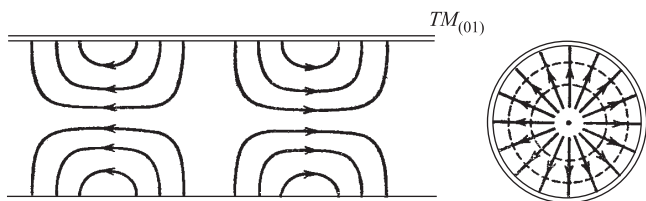


Рис. 197

Более сложное строение поля получается при $n \neq 0$. В этом случае имеются все пять составляющих поля (см. (210)).

На рис. 198 изображена картина поля для волны $TM_{(11)}(E_1)$. Заметим, что в поперечных разрезах поля волн $TM_{(01)}$ и $TM_{(11)}$ (рис. 197 и 198) совершенно такие же, как поля коаксиальной линии и двухпроводной экранированной линии.

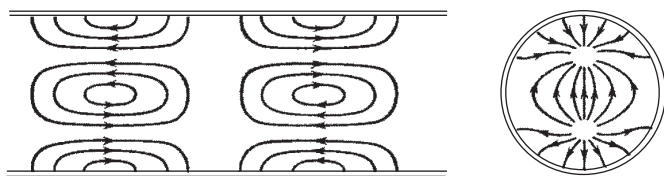


Рис. 198

В круглом волноводе возможны также поперечно-электрические волны TE . Составляющие поля волн TE выражаются следующими формулами:

$$\left. \begin{aligned} E_r &= A_r \sin n\varphi J_n\left(\frac{\mu}{a}r\right), & H_r &= B_r \cos n\varphi J'_n\left(\frac{\mu}{a}r\right) \\ E_\varphi &= A_\varphi \cos n\varphi J'_n\left(\frac{\mu}{a}r\right), & H_\varphi &= B_\varphi \sin n\varphi J_n\left(\frac{\mu}{a}r\right) \\ E_z &= 0, & H_z &= B_z \cos n\varphi J_n\left(\frac{\mu}{a}r\right) \end{aligned} \right\}. \quad (212)$$

В этих формулах $\mu = \mu_{ni}$ — корни уравнения

$$J'_n(x) = \frac{dJ_n(x)}{dx} = 0.$$

Значения μ_{ni} даны в табл. 13.

Таблица 13

Порядок функций n	Порядок корня i			
	1	2	3	4
0	3,83	7,02	10,17	13,32
1	1,84	5,33	8,54	11,71
2	3,05	6,71	9,97	—
3	4,20	8,02	11,34	—

Граничные длины волн $TE_{(ni)}$ определяются по формуле

$$\lambda_k = 2\pi \frac{a}{\mu_{ni}}.$$

Самая длинная граничная волна соответствует наименьшему корню. Сопоставляя табл. 12 и 13, видим (это — несколько неожиданный результат), что самая большая граничная длина волны получается для волны $TE_{(11)}$, хотя поле волны $TE_{(01)}$ имеет более простое строение.

Для волны $TE_{(01)}(H_0)$ составляющие поля таковы:

$$\left. \begin{aligned} E_\varphi &= A_\varphi J_1\left(\frac{3,83}{a}r\right), & H_r &= B_r J_1\left(\frac{3,83}{a}r\right) \\ E_r &= E_z = H_\varphi = 0, & H_z &= B_z J_0\left(\frac{3,83}{a}r\right) \end{aligned} \right\},$$

т. е. электрические силовые линии представляют окружности, лежащие в плоскости поперечного сечения волновода; магнитные силовые линии в поперечном сечении расположены радиально, а в продольном образуют замкнутые контуры. Картина поля волны $TE_{(01)}$ показана на рис. 199.

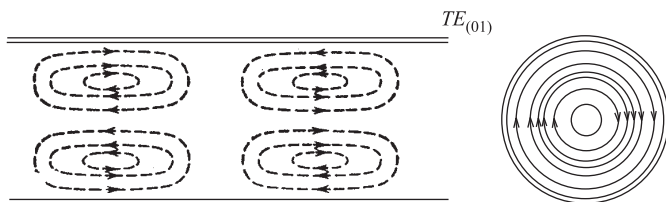


Рис. 199

Для волны $TE_{(11)}(H_1)$ поле строится из пяти не равных нулю составляющих, определяемых формулами (212). Картина поля показана на рис. 200. Более сложные волны в практике редко используются, и их рассматривать не будем. В случае надобности картина поля для любой волны $TM_{(ni)}$ или $TE_{(ni)}$ может быть построена по формулам (210) или (212).

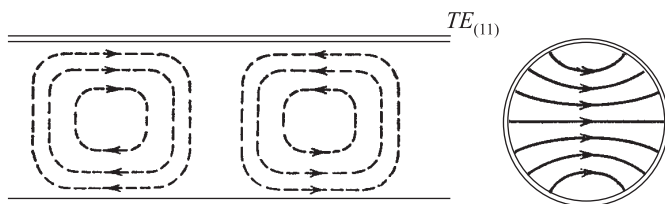


Рис. 200

В заключение упомянем о волноводах с *эллиптическим* сечением. Такие волноводы в практике не применяются, так как технология изготовления труб с эллиптическим сечением очень сложна, а никаких преимуществ эллиптические волноводы не имеют. Но теория эллиптических волноводов имеет большое прикладное значение, так как эллиптический волновод можно рассматривать как деформированный (сжатый с боков) круглый волновод. Поэтому теория эллиптического волновода позволяет рассчитать эффект отклонения от надлежащей формы сечения *круглого* волновода.

§ 58. Возбуждение и фильтрация волн в волноводе

Для возбуждения волны требуемого типа нужно создать в начальном сечении волновода электромагнитное поле соответствующего строения.

В § 54 было рассмотрено возникновение волны в круглом волноводе. Из описания явления ясно, каково должно быть устройство возбуждителя: он представляет собой стержень (штырь), расположенный по оси волновода в его начале. Напряжение от источника подается между возбуждающим стержнем и стенкой волновода (рис. 201). При этом образуется волна $TM_{(01)}$. В начальном сечении радиальное электрическое поле может быть выражено через приложенное напряжение; сопоставляя это выражение с формулой для E_r из (210), можем определить сначала постоянную A_r , а затем и остальные неизвестные постоянные A_z и B_φ . В прямоугольном волноводе при аналогичных условиях возбуждается волна $TM_{(11)}$.

Если ввести в круглый волновод два стержня, направленных вдоль волновода, то возбуждятся волна $TM_{(11)}$ (рис. 202). При помощи большего числа попарно соединенных между собой стержней можно возбудить более сложные волны TM .

Волны TM можно возбуждать не через посредство электрической составляющей, а через посредство магнитной составляющей. Иначе говоря, действие возбуждителя может определяться не приложенным к возбуждителю напряжением, а протекающим по возбуждителю током. Выбор того или иного способа возбуждения зависит от свойств ис-

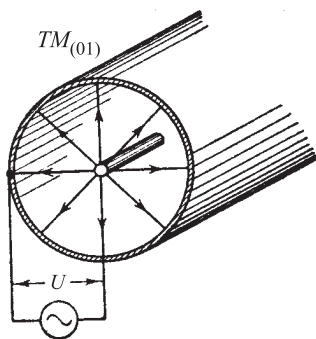


Рис. 201

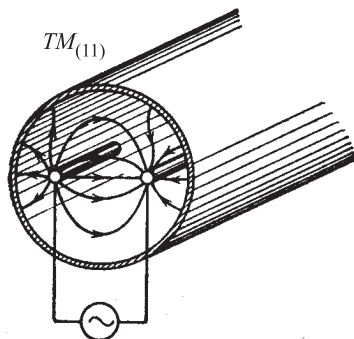


Рис. 202

точника (от величины внутреннего сопротивления) и от особенностей общей схемы.

На рис. 203 показан употребительный способ возбуждения волновода петлей (витком), лежащей в плоскости продольного сечения. При прохождении по петле тока возникает магнитное поле, одна силовая линия которого показана на рис. 203. Такое строение магнитного поля соответствует волне $TM_{(01)}$, которая и возбуждается. Волна $TM_{(11)}$ может возбуждаться, если расположить петлю посередине волновода, как показано на рис. 204¹⁾.

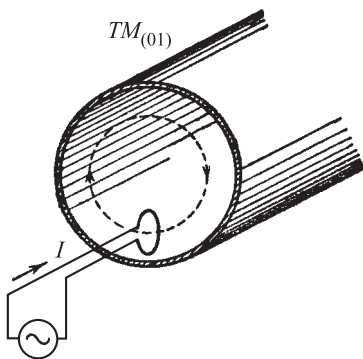


Рис. 203

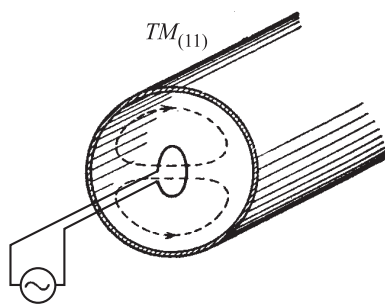


Рис. 204

Волны TE могут возбуждаться петлями, лежащими в плоскости поперечного сечения волновода. На рис. 205 изображен магнитный возбудитель волны $TE_{(01)}$ в виде одного витка, центр которого лежит на оси волновода.

¹⁾ Практически стенка волновода обычно заземлена; петли делаются одним проводом, конец которого запаивается на стенке волновода.

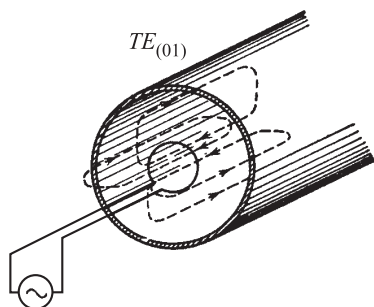


Рис. 205

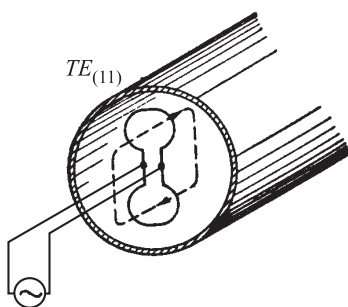


Рис. 206

На рис. 206 представлен возбуждатель в виде двух симметрично расположенных и параллельно включенных петель. При помощи такого возбуждателя можно создать волну $TE_{(11)}$. Эту же волну можно возбудить, просто присоединяя источник к двум диаметрально противоположным точкам стенки волновода (рис. 207). Механизм возбуждения

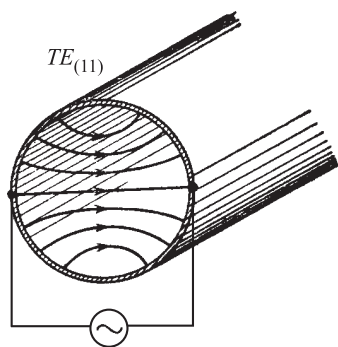


Рис. 207

при этом можно по желанию толковать как с точки зрения магнитного, так и с точки зрения электрического поля. В первом случае стенка волновода рассматривается как два полувитка; к этому представлению можно придти, если расширять петли на рис. 206, пока они не совпадут со стенкой. Со второй точки зрения внешнее поле источника смещает заряды, так что в сечении волновода образуется электрическое поле, показанное на рис. 207. Обе точки зрения одинаково хорошо объясняют образование поля, характерного для волны $TE_{(11)}$. Те же устройства, которые применяются для возбуждения в волно-

водах волн различных типов, могут служить и для вывода энергии из волновода с целью передачи ее в другие звенья высокочастотного тракта.

Вследствие несовершенства возбуждителей наряду с требуемой волной возбуждаются в той или иной мере волны других типов; основная волна оказывается засоренной наложенными на нее волнами других типов. Это во многих случаях нежелательно, и возникает задача *фильтрации*, т. е. выделения волны основного типа и подавления всех прочих волн.

Фильтрация волн в волноводе осуществляется особыми приемами, основанными на том, что все волны определенным образом *поляризованы*. Иначе говоря, для каждого типа волны как электрический,

так и магнитный векторы имеют вполне определенное направление в каждой точке сечения волновода. Волновой фильтр, не пропускающий волну некоторого определенного типа, можно построить в виде системы проводников, направление которых совпадает с направлением электрических силовых линий. В таком случае возникает поляризация, создающая поле обратного направления, и волна данного типа будет погашена, тогда как волны других типов (иначе, поляризованные) пройдут через волновой фильтр беспрепятственно или с малым ослаблением. Действие волнового фильтра можно объяснить и иначе: он представляет собой систему проводников, образующих контуры, сцепляющиеся с магнитными силовыми линиями. В контурах наводятся токи, создающие магнитное поле противоположного направления.

Для примера на рис. 208 изображены схематически некоторые фильтры. Все они состоят из провололочек, расположение которых воспроизводит строение электрического поля для волны соответствующего типа.

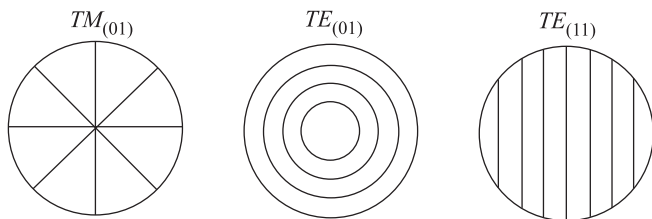


Рис. 208

На тех же основаниях можно построить и преобразователи волн, при помощи которых волна одного типа может быть превращена в волну другого типа. Пример конструкции такого рода преобразователя дан на рис. 209. Устройство предназначено для преобразования волны $TM_{(01)}$ в волну $TE_{(01)}$. Преобразователь состоит из провололочек, загнутых зигзагом с таким расчетом, что касательные участки образуют круговое кольцо в плоскости поперечного сечения волновода. Электрическое поле волны $TM_{(01)}$ наводит в радиальных участках провололочек токи. Эти токи, протекая по касательным участкам, образуют круговой ток, создающий магнитное поле волны $TE_{(01)}$ (см. рис. 205).

Из приспособлений волноводной техники упомянем еще о разного рода *диафрагмах*. Диафрагмы представляют собой перегородки, перекрывающие часть сечения волновода. В зависимости от формы диафрагмы она может вести себя как сопротивление того или иного характера. Для примера на рис. 210 показаны диафрагмы в прямоугольном волноводе с волной TE_{01} . Диафрагма рис. 210, а ведет себя как параллельно включенное индуктивное сопротивление, диафрагма рис. 210, б — как емкостное сопротивление (электрический вектор волны TE_{01} во входном сечении волновода направлен вертикально).

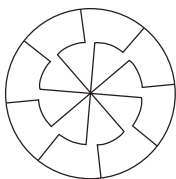


Рис. 209

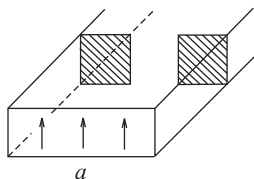
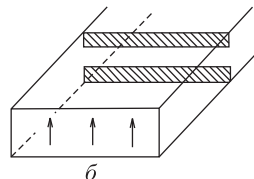


Рис. 210



Перегородка из материала с большим удельным сопротивлением, расположенная вдоль волновода, может применяться для увеличения затухания. Затухание легко регулировать, если выполнить перегородку в виде заслонки, вдвигаемой внутрь волновода через продольную щель в стенке волновода.

В целом система волновода со всеми приспособлениями для управления его работой и наблюдения за ней очень своеобразна и никак не похожа по внешнему виду на устройства для обычных радиочастот.

§ 59. Потери в волноводе

В предыдущей теории предполагалось, что стенки волновода имеют бесконечную проводимость. При этом потери в волноводе отсутствуют, и волна в нем распространяется без затухания.

В реальном волноводе имеются потери, обусловленные конечной проводимостью материала стенок. Очередная задача состоит в том, чтобы вычислить затухание в волноводе.

Если обозначить мощность, проходящую через данное сечение волновода через P_0 , то при наличии затухания мощность будет убывать вдоль волновода по закону

$$P = P_0 e^{-2\alpha z}$$

(множитель 2 в показателе появляется потому, что мощность пропорциональна квадрату электрического или квадрату магнитного вектора). Отсюда

$$\frac{dP}{dz} = -2\alpha P = -W,$$

где через W обозначена мощность, теряемая на единице длины волновода. Таким образом,

$$\alpha = \frac{W}{2P} \quad (213)$$

и для нахождения затухания α нужно вычислить проходящую через некоторое сечение мощность P , а также теряемую в стенках волновода мощность W . Начнем с первой из двух этих величин.

Будем рассматривать волны TM в прямоугольном волноводе.

Поток мощности, переносимый электромагнитной волной, выражается вектором Пойнтинга

$$\mathbf{P} = \frac{c}{4\pi} [\mathbf{E}\mathbf{H}].$$

Поток направлен вдоль волновода и зависит от составляющей вектора P по оси z

$$P_z = \frac{c}{4\pi} (E_x H_y - E_y H_x)$$

(мы полагаем затухание малым, а поле таким же, каково оно было при отсутствии затухания). Подставляя значения составляющих поля, получаем

$$P_z = \left(A_x B_y \cos^2 m\pi \frac{x}{a} \sin^2 n\pi \frac{y}{b} - A_y B_x \sin^2 m\pi \frac{x}{a} \cos^2 n\pi \frac{y}{b} \right).$$

Чтобы найти полную мощность, нужно проинтегрировать это выражение по всей площади поперечного сечения, т. е. по x от 0 до a и по y от 0 до b . Это даст:

$$P = \frac{c}{4\pi} \left(A_x B_y \int_0^a \cos^2 m\pi \frac{x}{a} dx \int_0^b \sin^2 n\pi \frac{y}{b} dy - A_y B_x \int_0^a \sin^2 m\pi \frac{x}{a} dx \int_0^b \cos^2 n\pi \frac{y}{b} dy \right).$$

Каждый из интегралов равен половине верхнего предела и

$$P = \frac{c}{16\pi} ab (A_x B_y - A_y B_x).$$

Значения A и B , выраженные через A_x (см. § 56), равны:

$$A_y = \frac{n}{m} \frac{a}{b} A_x; \quad B_y = \frac{k}{\beta} A_x, \quad B_x = -\frac{n}{m} \frac{a}{b} \frac{k}{\beta} A_x.$$

Подставляя эти значения в формулу для P , находим окончательно ¹⁾

$$P = \frac{\omega_k^2}{16\pi^3 c} \frac{a^2}{m^2} ab \frac{k}{\alpha} A_x^2, \quad (214)$$

где

$$\omega_k^2 = c^2 \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right], \quad \beta^2 = k^2 - \frac{\omega_k^2}{c^2}.$$

По физическому смыслу ω_k есть граничная частота, β — фазовая постоянная.

¹⁾ Определенная таким образом величина есть не средняя, а пиковая мощность. Для нахождения средней мощности нужно было бы делить (214) на два. Но в этом нет надобности, так как и W будет определять также.

Вычислим теперь мощность, теряемую в стенках. Прежде всего заметим, что речь идет о колебаниях очень высокой частоты, при которой

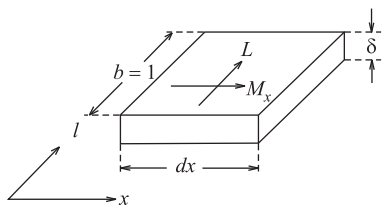


Рис. 211

определяющую роль играет *поверхностный эффект*. Плотность тока имеет наибольшее значение на поверхности и быстро убывает по мере углубления в вещество стенки. Для расчетов пользуются эквивалентной толщиной поверхностного слоя, представляя дело так, как если бы весь ток был сосредоточен в слое толщиной δ , в пределах которого плотность тока одинакова и

равна плотности тока на поверхности. Эквивалентная толщина δ зависит от частоты и проводимости; теория поверхностного эффекта дает для этой величины значение

$$\delta = \frac{1}{\sqrt{2\pi\sigma\omega}}.$$

Вырежем из горизонтальной стенки волновода полоску толщиной δ , длиной $b = 1$ и шириной dx (рис. 211). Обозначая плотность тока через j , будем иметь для тока

$$dI = j dx$$

(в этом случае j — поверхностная плотность).

Проводимость полоски равна

$$dG = \sigma \delta dx.$$

Мощность, теряемая в полоске,

$$dW = \frac{(dI)^2}{dG} = j^2 \frac{dx}{\delta\sigma}.$$

Нужно найти плотность тока. Для ее определения послужит закон полного тока, применяя который к данному случаю, можно записать

$$j = \frac{1}{4\pi} H_{x0}.$$

Итак,

$$dW = \frac{1}{16\pi^2} |H_{x0}|^2 \frac{dx}{\delta\sigma}.$$

Величина H_{x0} в этих формулах выражает составляющую поля H_x около самой стенки, т.е. при $y = 0$ (или $y = b$). Из общих формул получаем

$$H_{x0} = B_x \sin m\pi \frac{x}{a},$$

так что

$$dW = \frac{1}{16\pi^2\delta\sigma} B_x^2 \sin^2 m\pi \frac{x}{a} dx.$$

Для вычисления мощности, теряемой в горизонтальной стенке, нужно проинтегрировать это выражение по x в пределах от 0 до a , т. е.

$$W_a = \frac{B_x^2}{16\pi^2\delta\sigma} \int_0^a \sin^2 m\pi \frac{x}{a} dx = \frac{a}{32\pi^2\delta\sigma} B_x^2,$$

или, подставляя сюда значения B_x и δ ,

$$W_a = \frac{a}{32\pi^2} \sqrt{\frac{2\pi\omega}{\sigma}} \left(\frac{n}{m} \frac{a}{b} \frac{k}{\beta} \right)^2 A_x^2.$$

Совершенно аналогично найдем мощность, теряемую в вертикальной стенке:

$$W_b = \frac{b}{32\pi^2\delta\sigma} B_y^2,$$

или

$$W_b = \frac{b}{32\pi^2} \sqrt{\frac{2\pi\omega}{\sigma}} \left(\frac{k}{\beta} \right)^2 A_x^2.$$

Полная мощность, теряемая во всех четырех стенках прямоугольного волновода, определится как

$$W = 2(W_a + W_b) = \frac{1}{16\pi^2} \sqrt{\frac{2\pi\omega}{\sigma}} \left(\frac{k}{\beta} \right)^2 A_x^2 \left(\frac{n^2}{m^2} \frac{a^3}{b^2} + b \right). \quad (215)$$

Остается согласно (213) разделить (215) на удвоенное значение (214). Получаем

$$\alpha \left(\frac{\pi c}{2} \sqrt{\frac{2\pi}{\sigma}} \right) \left[\frac{\frac{1}{b} \left(\frac{n}{b} \right)^2 + \frac{1}{a} \left(\frac{m}{a} \right)^2}{\omega_k^2} \right] \left(\sqrt{\frac{\omega}{1 - \frac{\omega_k^2}{\omega^2}}} \right).$$

В этом выражении первый множитель зависит от физических констант, второй — от размеров волновода (a, b) и типа волны (m, n), третий — от частоты. Общий характер зависимости затухания от частоты определяется тем, что при $\omega = \omega_k$ затухание бесконечно велико. Оно быстро падает при повышении частоты и достигает на некоторой частоте минимума; затем затухание сравнительно медленно растет неограниченно с увеличением частоты.

На основании тех же физических представлений можно подсчитать затухание и для волн TE . Отличие этого случая от ранее рассмотренного состоит в том, что имеются поперечные токи, т. е. токи, пути которых лежат в плоскости поперечного сечения волновода.

Подобным же образом могут быть выполнены расчеты и для круглого волновода. Не делая выкладок, приведем лишь готовые результаты

для четырех простейших волн ¹⁾:

$$\begin{aligned}
 TM_{(01)} \quad \alpha &= \frac{0,0567}{a^{3/2}} \sqrt{\rho} \frac{y^{3/2}}{\sqrt{y^2 - 1}}, \\
 TM_{(11)} \quad \alpha &= \frac{0,0715}{a^{3/2}} \sqrt{\rho} \frac{y^{3/2}}{\sqrt{y^2 - 1}}, \\
 TE_{(01)} \quad \alpha &= \frac{0,0715}{a^{3/2}} \sqrt{\rho} \frac{y^{-1/2}}{\sqrt{y^2 - 1}}, \\
 TE_{(11)} \quad \alpha &= \frac{0,0496}{a^{3/2}} \sqrt{\rho} \frac{y^{-1/2} + 0,429 y^{3/2}}{\sqrt{y^2 - 1}}.
 \end{aligned}$$

В эти формулы вместо удельной проводимости σ введена более употребительная величина — удельное сопротивление ρ . Если ρ выражено в Ом · см, то α получается в неперах на сантиметр. В вышеприведенных

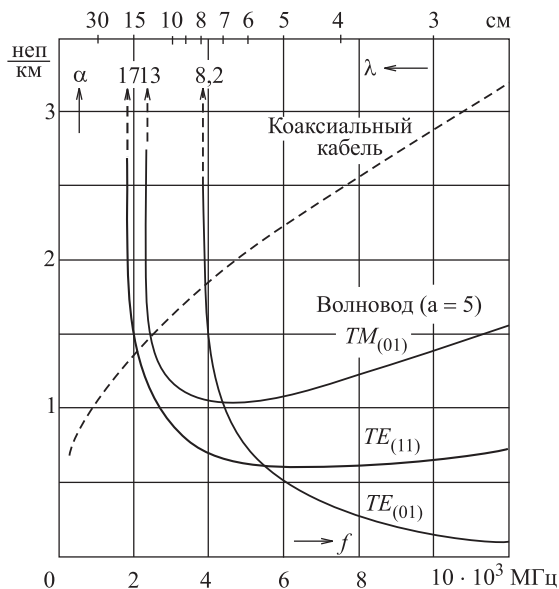


Рис. 212

формулах a — радиус волновода, $y = \omega/\omega_k$. Зависимость затухания от частоты для всех волн такая же, как указано выше; исключение составляет только волна $TE_{(01)}$: для нее затухание монотонно убывает с частотой. Это интересное свойство может быть объяснено следую-

¹⁾ Из книги Введенского Б. А., Аренберга А. Г. Радиоволноводы. — Гостехиздат, 1946. — С. 119.

щим образом. Волна $TE_{(01)}(H_0)$ — единственная волна, для которой отсутствует касательная составляющая магнитного поля H_φ . Поэтому токи в стенках волновода наводятся только продольной составляющей H_z . Эти токи замыкаются по круговым контурам, лежащим в плоскости поперечного сечения волновода (т. е. расположенным так же, как электрические силовые линии). Пусть частота увеличивается, а длина волны соответственно убывает. При этом происходит постепенный переход к условиям распространения, которые имелись бы в открытом пространстве (так как отношение поперечника волновода к длине волны растет с повышением частоты). Но в открытом пространстве поле электромагнитной волны чисто поперечно. Следовательно, с повышением частоты составляющая H_z должна стремиться к нулю, а вместе с нею и токи, наводимые ею в стенках волновода.

На рис. 212 представлена зависимость затухания от частоты для трех типов волн, а именно: $TM_{(01)}$, $TE_{(11)}$ и $TE_{(01)}$ в одном и том же круглом волноводе диаметром 10 см. Граничные длины волн названных типов составляют соответственно 13, 17 и 8,2 см. Затухание выражено в неперах на километр. Для сравнения на том же рисунке пунктиром дана частотная характеристика затухания в коаксиальном кабеле такого же диаметра.

Из приведенных данных следует, что для передачи энергии по круглому волноводу на значительные расстояния интерес представляют в первую очередь волны $TE_{(01)}$ и $TE_{(11)}$ (H_0 и H_1).

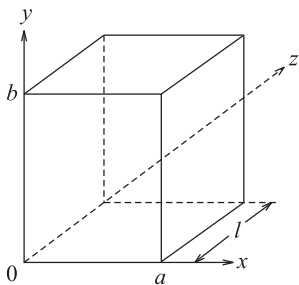
§ 60. Объемные резонаторы

Резонансные явления играют в технике свч столь же важную роль, как и в радиотехнике обычных частот. Но такой простейший резонатор, как колебательный контур, составленный из сосредоточенных индуктивности и емкости, на свч не применим. Во-первых, его просто нельзя осуществить в обычных конструктивных формах — он вырождается в один разомкнутый виток, а во-вторых, открытый контур на свч сильно излучает. По этим причинам в технике свч применяются резонаторы особого рода, а именно объемы той или иной формы, ограниченные со всех сторон проводящими стенками. Такого рода устройства называются *объемными* (или *полыми*) *резонаторами*. Резонанс в них имеет волновую природу.

Рассмотрим волновую картину в прямоугольном параллелепипеде с ребрами, длиной a , b и l , которые направим соответственно по осям x , y и z (рис. 213). Представление о поле внутри такого параллелепипеда можно составить проще всего, основываясь на уже известных соотношениях для прямоугольного волновода. Именно, рассматриваемый параллелепипед можно получить, если взять прямоугольный волновод с сечением ab и, вырезав из него кусок длиной l , закрыть проводящими стенками торцы отрезка. Ясно, что при этом условия для распространения волн в направлении всех трех осей будут одинаковы в том смысле,

что во всех трех направлениях могут существовать только стоячие волны. В этом только и состоит отличие явлений в прямоугольном объемном резонаторе от явлений в прямоугольном волноводе, в котором вдоль оси z может существовать бегущая волна.

Представляя составляющие поля произведением трех функций, каждая из которых зависит только от одной координаты, и принимая во внимание, что каждая из этих функций выражает стоячую волну, приходим к заключению, что каждая составляющая поля должна выражаться произведением трех тригонометрических функций (т. е. синусов или косинусов). Решение должно удовлетворять следующим граничным условиям:



$$\begin{array}{lll} \text{при } x = 0, & x = a & E_y = E_z = 0, \\ \text{при } y = 0, & y = b & E_x = E_z = 0, \\ \text{при } z = 0, & z = l & E_x = E_y = 0. \end{array}$$

Рис. 213

Рассмотрим частные виды волн, для которых одна из составляющих равна нулю.

Воспользуемся теми же обозначениями, которые мы применяли при исследовании волноводов. Обозначим символом TM поле, для которого $H_z = 0$. Для этого случая можно записать следующие выражения для составляющих поля:

$$\left. \begin{array}{l} E_x = A_x \cos m\pi \frac{x}{a} \sin n\pi \frac{y}{b} \sin i\pi \frac{z}{l} \\ E_y = A_y \sin m\pi \frac{x}{a} \cos n\pi \frac{y}{b} \sin i\pi \frac{z}{l} \\ E_z = A_z \sin m\pi \frac{x}{a} \sin n\pi \frac{y}{a} \cos i\pi \frac{z}{l} \\ H_x = B_x \sin m\pi \frac{x}{a} \cos n\pi \frac{y}{b} \cos i\pi \frac{z}{l} \\ H_y = B_y \cos m\pi \frac{x}{a} \sin n\pi \frac{y}{b} \cos i\pi \frac{z}{l} \\ H_z = 0 \end{array} \right\}. \quad (216)$$

Уже известно, что по меньшей мере две составляющие магнитного вектора должны быть не равны нулю. В данном случае речь идет о составляющих H_x и H_y .

Как видно из формул (216), m и n не могут быть равны нулю. Поэтому простейшее строение поля соответствует комбинации индексов mni , выражаемой тройкой чисел 110. Комбинации 000, 001, 010, 100, 101 и 011 не могут быть осуществлены.

Для полей типа TE , для которых $E_z = 0$, имеем следующие значения составляющих:

$$\left. \begin{aligned} E_x &= A_x \cos m\pi \frac{x}{a} \sin n\pi \frac{y}{b} \sin i\pi \frac{z}{l} \\ E_y &= A_y \sin m\pi \frac{x}{a} \cos n\pi \frac{y}{b} \sin i\pi \frac{z}{l} \\ E_z &= 0 \\ H_x &= B_x \sin m\pi \frac{x}{a} \cos n\pi \frac{y}{b} \cos i\pi \frac{z}{l} \\ H_y &= B_y \cos m\pi \frac{x}{a} \sin n\pi \frac{y}{b} \cos i\pi \frac{z}{l} \\ H_z &= B_z \cos m\pi \frac{x}{a} \cos n\pi \frac{y}{b} \sin i\pi \frac{z}{l} \end{aligned} \right\}. \quad (217)$$

Из этих формул видно, что как m , так и n может быть равно нулю, но не одновременно; i не может равняться нулю (так как при $i = 0$ обращается в нуль электрический вектор).

Итак, простейшие поля вида TE соответствуют комбинациям индексов 011 и 101.

Выбор направления осей, разумеется, произволен; все три направления в параллелепипеде равноправны. Поэтому можно построить поле, для которого *любая* составляющая равна нулю. Для этого достаточно лишь поменять местами x , y и z в формулах (216) или (217).

Теперь определим собственные частоты резонатора. Собственными частотами назовем такие значения частоты, при которых удовлетворятся волновое уравнение (198):

$$\nabla^2 \mathbf{E} + k^2 \mathbf{E} = 0.$$

С физической точки зрения это означает, что в объемном резонаторе могут существовать свободные колебания только вполне определенных частот. Эти частоты зависят от строения поля (т. е. от m , n и i) и от размеров резонатора (т. е. от a , b и l). Чтобы найти собственные частоты, составим вторые производные для любой составляющей, хотя бы для E_x (см. (216)):

$$\frac{\partial^2 E_x}{\partial x^2} = -\left(\frac{m\pi}{a}\right)^2 E_x, \quad \frac{\partial^2 E_x}{\partial y^2} = -\left(\frac{n\pi}{b}\right)^2 E_x, \quad \frac{\partial^2 E_x}{\partial z^2} = -\left(\frac{i\pi}{l}\right)^2 E_x.$$

Подставляя эти значения в уравнение (199)

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) E_x + k^2 E_x = 0,$$

получим

$$\omega = \pi c \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{i}{l}\right)^2}.$$

Это и есть уравнение собственных частот, одинаковое для всех видов полей. Возьмем для примера кубический резонатор ($a = b = l$). Так как для простейших полей один из индексов равен нулю, а остальные два — единице, получаем

$$\omega = \frac{\pi c}{a} \sqrt{2}$$

или

$$f = \frac{c}{\sqrt{2} a}.$$

Так, например, при $a = 5$ см, наименьшая собственная частота будет

$$f = \frac{3 \cdot 10^{10}}{\sqrt{2} \cdot 5} \approx 4,3 \cdot 10^9 \text{ Гц} = 4300 \text{ МГц}.$$

Объемный резонатор, как и волновод, может возбуждаться возбудителем либо магнитного типа (виток, петля), либо электрического типа (стержень, штырь). В первом случае режим возбуждения определяется током, протекающим по витку, во втором случае — напряжением, поданным на стержень.

Резонансные явления в объемном резонаторе состоят в том, что при заданном режиме возбуждения (конечный ток или напряжение) электромагнитное поле в резонаторе зависит от частоты и достигает очень больших значений, когда частота возбуждения совпадает с одной

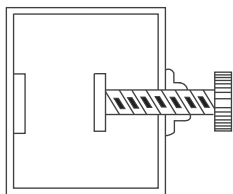


Рис. 214

из собственных частот резонатора. При этом, конечно, поле оказывает обратное воздействие на возбудитель, в результате чего входное сопротивление возбудителя при резонансе также получает экстремальное значение. Таким образом, по отношению к цепи возбудителя объемный резонатор ведет себя в окрестности какой-либо из резонансных частот как обычный контур. Очень существенно, что благодаря малым потерям добротность эквивалентного контура оказывается очень высокой.

Мы рассматривали только прямоугольный объемный резонатор. На тех же общих физических основаниях строятся объемные резонаторы самых различных форм, из которых чаще других применяются цилиндрические и тороидальные. Выбор формы резонатора определяется общей конструктивной компоновкой данного устройства свч.

Укажем коротко на некоторые способы настройки объемных резонаторов. Очень распространена настройка посредством подвижного поршня, перемещение которого изменяет один из размеров резонатора. Не обязательно, чтобы поршень перекрывал все сечение резонатора; аналогичный результат получается и в конструкции, показанной схематически на рис. 214. При ввинчивании винта резонансная частота понижается.

Интересен метод настройки, состоящий в перемещении внутри объемного резонатора некоторого проводящего тела (например, металли-

ческого шарика). Если это тело поместить в область поля, где наиболее густо расположены электрические силовые линии, то резонансная частота понизится (так как возрастет эквивалентная емкость). Если же проводящее тело переместить в область с наибольшей густотой магнитных линий, то резонансная частота повысится (так как уменьшается эквивалентная индуктивность).

Из всего сказанного ясно, что вопросы механического конструирования и технологии играют в устройствах свч существенную роль. В то время как устройство, работающее на обычных радиочастотах, может быть собрано из типовых деталей при помощи одного лишь паяльника, устройство свч представляет собой специализированную механическую конструкцию.

§ 61. Сопоставление волновода и линии

При изучении линий мы говорили о волнах напряжения и тока, распространение которых подчиняется одномерным волновым уравнениям. Мы не пользовались при этом уравнениями электродинамики и не исследовали электромагнитного поля в окружающем линию пространстве.

Когда же мы занялись волноводами, то начали с уравнений Максвелла и, получив из них трехмерное волновое уравнение, свели задачу к исследованию поля электромагнитной волны в пространстве. При этом сразу обнаружились особенности волн в волноводах: дисперсия фазовой скорости и наличие граничных частот, зависящих как от размеров волновода, так и от типа волны.

Возникает естественный вопрос: если подход к линиям и волноводам столь различен и если налицо существенное различие в их свойствах, то означает ли это, что имеется фундаментальное *физическое различие явлений* в линиях и волноводах?

Этот важный вопрос заслуживает внимательного рассмотрения.

Нужно сразу сказать, что как волноводы, так и линии *можно* трактовать с единой точки зрения на основе уравнений электродинамики. Но для линий это не является необходимым, так как специфические условия их работы позволяют применить к ним упрощенный подход, которым мы и пользовались. При этом упрощенном рассмотрении теряются, однако, некоторые детали, которые мы сейчас восстановим, применив к обычным линиям общий электродинамический подход.

Явления в линии исчерпывающим образом описываются строением электромагнитного поля в окружающем линию пространстве. На рис. 182 и 183 даны картины поля для двухпроводной и коаксиальной линий (без потерь). Было отмечено — и это имеет решающее значение — что векторы $\mathbf{E}$ и $\mathbf{H}$ лежат в плоскости поперечного сечения линии. Волна, удовлетворяющая такому условию, называется *поперечной* и может быть обозначена *ТЕМ*.

Отметим различие с волнами в волноводе, которые не могут быть чисто поперечными: в волноводе либо вектор $\mathbf{E}$ может лежать в плоскости поперечного сечения, и тогда вектор $\mathbf{H}$ обязательно имеет продольную составляющую (волны TE), либо наоборот (волны TM).

Определим постоянную распространения для волны TEM .

Для этой волны

$$E_z = H_z = 0.$$

Воспользуемся уравнениями Максвелла. Имеем (194)

$$\operatorname{div} \mathbf{E} = 0$$

или в проекциях

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0.$$

Но так как $E_z = 0$, то

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = 0. \quad (218)$$

Возьмем теперь уравнение (197)

$$\mathbf{H} = -\frac{1}{ik} \operatorname{rot} \mathbf{E},$$

из которого получаем для составляющей H_z (205)

$$H_z = -\frac{1}{ik} \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \right).$$

Но так как $H_z = 0$, то

$$\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} = 0. \quad (219)$$

Дифференцируя повторно (218) по x и (219) по y и складывая, находим

$$\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial^2 E_x}{\partial y^2} = 0. \quad (220)$$

Это — известное из теории поля уравнение Лапласа. Такому же уравнению удовлетворяют и остальные три составляющие поля (т. е. E_y , H_x и H_y).

Таким образом, в плоскости поперечного сечения поле волны TEM является потенциальным.

Волновое уравнение

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) E_x + k^2 E_x = 0$$

распадается на два уравнения: на уравнение Лапласа (220) и на уравнение

$$\frac{\partial^2 E_x}{\partial z^2} + k^2 E_x = 0. \quad (221)$$

Решение уравнения (221) может быть записано в виде

$$E_x = f(x, y) e^{-\gamma z},$$

где $f(x, y)$ — решение уравнения Лапласа (220), а γ — постоянная распространения, получаемая из характеристического уравнения для (221), т. е.

$$\gamma^2 + k^2 = 0,$$

откуда

$$\gamma = i k = i \frac{\omega}{c}.$$

Итак, поперечная волна TEM распространяется с фазовой скоростью c . Эта скорость постоянна (т. е. не зависит от частоты); дисперсия фазовой скорости отсутствует; не существует граничной частоты, т. е. возможна волна любой частоты.

Мы выяснили, таким образом, сущность различия между линией и волноводом: в линии возможна поперечная волна TEM , называемая также *основной* волной; эта волна в отличие от волн в волноводе обладает уже известными из элементарной теории простыми свойствами.

Но из общей электродинамической теории линий следует, что основная волна — не единственный вид волны, которая может существовать на линии. Возможны и более сложные волны — волны высших типов. Эти волны уже не являются поперечными волнами TEM , а обладают всеми свойствами волноводных волн: дисперсией фазовой скорости и наличием граничных частот. Для нахождения этих волн нужно решить общее волновое уравнение

$$\nabla^2 \mathbf{E} + k^2 \mathbf{E} = 0$$

при соответствующих граничных условиях. Эти условия состоят в том, что электрический вектор нормален к поверхности проводов (имеется в виду все время линия без потерь; в этом случае поверхности проводов эквипотенциальны). Вычисления сравнительно сложны и здесь не приводятся. Укажем лишь для примера, что для коаксиальной линии самая длинная граничная волна (т. е. самая низкая граничная частота) получается для волны TE_{11} электрическое поле которой показано на рис. 215.

Граничная длина волны для волны TE_{11} в коаксиальной линии равна приблизительно

$$\lambda_0 \approx \pi(a + b),$$

где a и b — радиусы внутреннего и внешнего проводов. Таким образом, самые низкие граничные частоты лежат в диапазоне сантиметровых волн, и возможность возникновения волноводного типа волн не имеет

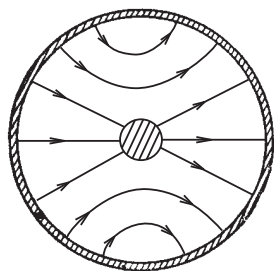


Рис. 215

практического значения в тех условиях, в которых обычно работают линии, в том числе и коаксиальные.

Остается пояснить, что мы пользуемся в элементарной теории линий напряжением и током не потому, что волны в линиях отличны по своей физической природе от волн в волноводах, а потому, что простые условия поперечной основной волны позволяют описать электромагнитное поле интегральными величинами. Именно, мы определяем *напряжение* интегралом

$$U = \int_L \mathbf{E} dl,$$

где интеграл берется по произвольной траектории между двумя проводниками. Это есть так называемый линейный интеграл напряженности электрического поля. Аналогично ток определяется через закон полного тока

$$I = \int_L \mathbf{H} dl,$$

где интеграл берется по замкнутому контуру, охватывающему провод. Введение интегральных величин U и I избавляет нас от необходимости рассматривать строение электромагнитного поля в пространстве, окружающем линию; одномерное волновое уравнение линии может быть при помощи вышеуказанных интегральных операций *выведено* из уравнений электромагнитного волнового поля. Легко понять, что это методика не применима к волноводам, так как линейные интегралы, выражающие U и I , не зависят от пути интегрирования только в случае безвихревого (потенциального) поля. Таково, как было показано, чисто поперечное поле волны *ТЕМ*. В волноводе же поле всегда имеет вихревой характер, и понятие напряжения (как разности потенциалов) теряет свой простой смысл. К тому же в волноводе негде локализовать как напряжение, так и ток.

§ 62. Разложение волн в волноводе на плоские волны

Введем определение *плоской волны*. Плоской называется волна, поле которой зависит только от одной координаты z , отсчитываемой в направлении распространения, т. е. производные всех составляющих поля по x и по y равны нулю.

Отсюда сразу следует (см. (205) и (208)), что

$$E_z = H_z = 0,$$

т. е. плоская волна является поперечной и представляет собой простейшую форму волн *ТЕМ*. Поверхности равных фаз представляют собой плоскости, нормальные к направлению распространения, откуда и название этого типа волны. Волновые уравнения для плоской волны

становятся одномерными и принимают вид

$$\frac{d^2 E_x}{dz^2} + k^2 E_x = 0, \quad (222)$$

аналогичный вид имеют уравнения для E_y , H_x и H_y .

Решение уравнения (222) имеет вид

$$E_x = A_x e^{-i k z} \quad (222a)$$

(рассматривается только волна, бегущая в положительном направлении оси Z).

Введем теперь более общее выражение для плоской волны, распространяющейся в направлении Z' , как угодно расположенном относительно координатных осей X, Y, Z . Введем новую переменную

$$z' = ax + by + cz,$$

отсчитываемую вдоль оси Z' новой координатной системы $X'Y'Z'$. Величины a, b и c представляют собой так называемые направляющие косинусы, т. е. косинусы углов между новой осью Z' и старыми осями X, Y и Z . При этом

$$a^2 + b^2 + c^2 = 1,$$

что следует непосредственно из того, что квадрат длины вектора равен сумме квадратов его проекций.

Для любой из составляющих плоской волны, распространяющейся в направлении оси Z' , можно записать

$$E = A e^{-i k z'},$$

или

$$E = A e^{-i k(ax+by+cz)}. \quad (223)$$

Это и есть выражение для плоской волны, движущейся в произвольном направлении, задаваемом в системе прямоугольных координат XYZ направляющими косинусами a, b и c .

Плоские волны обладают, как видим, весьма простыми свойствами. Возникает идея представить волну в волноводе как результат наложения (суперпозиции) плоских волн. Другими словами, попытаемся разложить волну в волноводе на плоские составляющие. Можно показать, что такое разложение всегда возможно.

Рассмотрим простейший из известных случаев, а именно, разложение на плоские составляющие волны TE_{10} в прямоугольном волноводе. Для этого вида волны составляющие электрического вектора выражаются следующими формулами:

$$E_x = E_z = 0; \quad E_y = A_y \sin \pi \frac{x}{a} e^{-\gamma z}$$

(здесь и ниже a — размер волновода в направлении x), причем для бегущей волны постоянная распространения мнимая, так что можно записать

$$\gamma = i \beta.$$

Пользуясь формулой Эйлера, представим составляющую E_y в виде

$$E_y = A_y \frac{1}{2i} \left[\exp \left(-i \left(-\frac{\pi}{a} x + \beta z \right) \right) - \exp \left(-i \left(\frac{\pi}{a} x + \beta z \right) \right) \right],$$

или

$$E_y = A'_y \left[\exp \left(-i k \left(-\frac{\pi}{ka} x + \frac{\beta}{k} z \right) \right) - \exp \left(-i k \left(\frac{\pi}{ka} x + \frac{\beta}{k} z \right) \right) \right].$$

Сопоставляя это выражение с формулой (223), видим, что волна TE_{10} разложена на две плоские волны, направления распространения которых Z' и Z'' лежат в плоскости xz и образуют с осью Z углы, косинус которых равен

$$\cos \vartheta = \frac{\beta}{k},$$

как показано на рис. 216.

Сумма квадратов направляющих косинусов должна быть равна единице. Из этого условия следует

$$\frac{\pi^2}{a^2} + \beta^2 = k^2,$$

откуда

$$\beta = \sqrt{k^2 - \frac{\pi^2}{a^2}}$$

или

$$\beta = k \sqrt{1 - \frac{\pi^2}{k^2 a^2}} = k \sqrt{1 - \frac{\omega_k^2}{\omega^2}},$$

где

$$\omega_k = \frac{\pi c}{a}.$$

Мы нашли, таким образом, фазовую постоянную β и показали наличие дисперсии фазовой скорости для волны в волноводе.

Столь же просто можно получить и другие соотношения, характеризующие волны в волноводе.

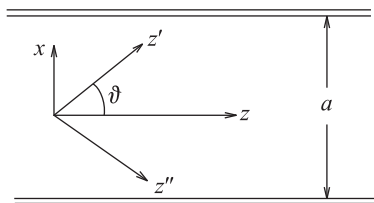


Рис. 216

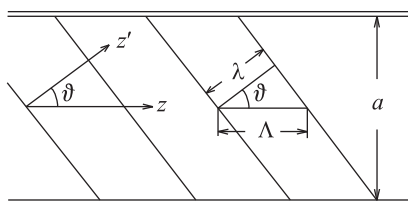


Рис. 217

На рис. 217 показаны параллельные линии, представляющие следы плоскостей равных фаз для одной из плоских составляющих. Расстояние между этими плоскостями в направлении z' распространения

плоской волны равно, очевидно, длине волны λ . Но расстояние между теми же плоскостями в направлении оси волновода z равно длине волны в волноводе Λ . Из прямоугольного треугольника на рис. 217 получаем

$$\lambda = \Lambda \cos \vartheta$$

или

$$\Lambda = \frac{\lambda}{\cos \vartheta} = \frac{k\lambda}{\beta},$$

откуда

$$\frac{\Lambda}{\lambda} = \frac{v}{c} = \frac{k}{\beta} = \frac{1}{\sqrt{1 - \frac{\omega_k^2}{\omega^2}}}.$$

Из того же рис. 217 можно получить и выражение для скорости движения энергии, т. е. для групповой скорости. В плоской волне энергия перемещается со скоростью c по направлению z' . Но вдоль волновода скорость движения энергии выражается проекцией c на направление z . Таким образом, групповая скорость равна

$$u = c \cos \vartheta.$$

Эта скорость всегда меньше скорости света. А так как фазовая скорость v согласно предыдущему равна

$$v = \frac{c}{\cos \vartheta},$$

то имеем общее соотношение

$$uv = c^2.$$

По мере понижения частоты и приближения ее к граничной угол ϑ растет. При граничной частоте $\vartheta = \pi/2$; при этом обе плоские волны распространяются поперек волновода и волна, бегущая вдоль волновода, исчезает; фазовая скорость v обращается в бесконечность, а групповая скорость u — в нуль. Ниже граничной частоты волновое состояние поля в волноводе невозможно, так как не могут быть удовлетворены граничные условия.

В самом деле: при $\omega < \omega_k$ $\lambda > \lambda_k = \frac{2\pi c}{\omega_k} = 2a$, т. е. половина длины волны больше размера волновода, и на стенках волновода не могут образоваться требуемые узлы касательной составляющей электрического поля E_y ¹⁾.

Подводя итоги, можно сказать, что, разложив волну в волноводе на плоские волны и рассматривая геометрию поля одной из плоских

¹⁾ Формально это заключение следует непосредственно из соотношения

$$\sin \vartheta = \frac{\omega_k}{\omega} = \frac{\pi}{ka} = \frac{\lambda}{2a} \leq 1.$$

составляющих, можно очень просто и наглядно вывести основные свойства волн в волноводе.

Выше уже говорилось о том, что разложение на плоские волны всегда возможно. Не приводя общего метода такого разложения, заметим лишь, что более сложные волны вида TE_{mn} в прямоугольном волноводе разлагаются на *четыре* плоские составляющие.

В круглом волноводе приходится брать бесконечное число бесконечно малых плоских составляющих, так что волна представляется не конечной суммой плоских составляющих, а соответствующим интегралом.

В заключение коснемся одного общего вопроса, который может возникнуть в связи с предыдущими рассуждениями, а именно вопроса о *реальности* плоских составляющих сложного волнового поля. Действительно ли они существуют? Заметим, что точно в таком же смысле можно ставить вопрос о реальности боковых полос в модуляционном спектре (кстати говоря, этот вопрос вызвал в свое время длительную и острую дискуссию), и вообще вопрос о реальности синусоидальных составляющих, на которые можно разложить периодическую (ряд Фурье) или даже непериодическую (интеграл Фурье) функцию.

Речь идет вообще о реальности любого разложения, представляющего некоторую функцию конечной или бесконечной суммой составляющих или слагаемых.

С математической точки зрения всякая формула разложения представляет собой тождество, правая и левая части которого совершенно равноправны. Вопрос о реальности составляющих нужно понимать в смысле возможности экспериментального выделения и наблюдения отдельных составляющих. Как известно, синусоидальные составляющие гармонического анализа в этом смысле реальны — они могут быть выделены и наблюдаемы. Что же касается плоских составляющих волны в волноводе, то в принципе и они могли бы быть выделены соответствующими пространственными фильтрами и наблюдаться порознь. Если это не делается, то только потому, что в этом нет надобности.

Часть III

НЕЛИНЕЙНЫЕ И ПАРАМЕТРИЧЕСКИЕ ЯВЛЕНИЯ В РАДИОТЕХНИКЕ

Глава 7

ВВЕДЕНИЕ

§ 63. Характеристика предмета

Эта часть посвящена нелинейным и параметрическим явлениям, используемым в радиотехнике.

Такая терминология относится к физическим особенностям изучаемых явлений, особенностям, находящим отражение в математических уравнениях, которые описывают поведение тех или иных радиотехнических устройств.

Явления, с которыми приходится иметь дело в радиотехнике, описываются математическими уравнениями следующих трех видов:

- 1) линейные уравнения с постоянными коэффициентами,
- 2) линейные уравнения с переменными коэффициентами,
- 3) нелинейные уравнения.

Свойства и способы решения этих трех видов уравнений существенно различны; этим и обусловлено традиционное деление курса радиотехники на «линейную» и «нелинейную» части.

Линейные уравнения с постоянными коэффициентами — самый простой тип уравнений, теория которых детально разработана и для решения которых существуют мощные общие методы (например, методы интегралов Фурье и Дюамеля). Эти общие методы основаны на применении принципа наложения (суперпозиции).

Нужно заметить, что применимость принципа наложения служит определением линейной системы, т. е. системы, описываемой линейным уравнением как с постоянными, так и с переменными коэффициентами.

Общий вид однородного линейного уравнения с постоянными коэффициентами таков:

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = 0,$$

где $a_n, a_{n-1}, \dots, a_1, a_0$ — постоянные величины, откуда и название этого типа уравнений.

Линейные уравнения с переменными коэффициентами характеризуются тем, что коэффициенты уравнения *зависят от аргумента t* , т. е. являются заданными функциями времени. Теория этих уравнений более сложна и менее разработана. Этим главным образом оправдывается отнесение явлений, описываемых уравнениями с переменными коэффициентами, к «нелинейной» радиотехнике. С физической точки зрения уравнения с переменными коэффициентами описывают поведение систем, параметры которых так или иначе изменяются во времени; такие системы и происходящие в них явления будем кратко называть *параметрическими*. Примеры уравнений с переменными коэффициентами:

$$\frac{d^2 y}{dt^2} + ty = 0,$$

$$\frac{d^2 y}{dt^2} + (a + 2q \cos 2t) y = 0.$$

Последнее уравнение носит название уравнения Матье; оно играет большую роль в теоретической радиотехнике.

Наконец, нелинейные уравнения отличаются тем, что их коэффициенты *зависят от функции* или ее производных. Это наиболее трудный для решения тип уравнений; некоторые общие приемы будут рассмотрены ниже применительно к тем или иным радиотехническим проблемам.

Примеры нелинейных уравнений:

$$\frac{d^2 y}{dt^2} + ay^2 = 0,$$

$$\frac{d^2 y}{dt^2} + \frac{dy}{dt} y = 0,$$

$$\frac{d^2 y}{dt^2} + f\left(\frac{dy}{dt}\right) + y = 0.$$

Последнее уравнение имеет довольно общий характер; оно встречается еще в теории лампового генератора.

Математическая сторона дела интересует нас потому, что при помощи математического анализа можно провести количественное исследование различных радиотехнических устройств, и, в конечном счете, построить систему расчета этих устройств. Что же касается физической и технической сторон дела, то нужно сказать, что к нелинейным и параметрическим явлениям относятся важнейшие для радиотехники явления и процессы. Вот краткий перечень этих процессов: генерирование электрических колебаний, умножение и деление частоты, выпрямление и детектирование, модуляция и преобразование частоты.

Итак, многие важные радиотехнические процессы относятся к нелинейным и параметрическим явлениям. Они описываются нелинейными уравнениями или уравнениями с переменными коэффициентами. Эти уравнения и составляют математический аппарат «нелинейной» радиотехники.

§ 64. Обзор методов нелинейной теории

Как указывалось в предыдущем параграфе, принцип наложения, составляющий основу линейной теории, не применим для исследования явлений в нелинейных системах. Поясним коротко это положение.

Начнем со случая линейной системы, описываемой линейным алгебраическим уравнением

$$y = ax.$$

Здесь x и y — функции времени; x означает воздействие на систему, а y — ее отклик. Принцип наложения состоит в том, что если на систему оказывается воздействие

$$x = x_1 + x_2,$$

то отклик системы будет

$$y = ax = a(x_1 + x_2) = y_1 + y_2,$$

т. е. действия обоих слагаемых x_1 и x_2 независимы. Это можно выразить краткой формулой: действие суммы причин равно сумме действий каждой причины, отдельно взятой. Так, например, ток в линейной цепи при действии нескольких эдс определяется суммой токов, порождаемых каждой из эдс в отсутствие других.

Принцип наложения лежит в основе общих методов решения многочисленных задач линейной теории. Мы настолько привыкаем к постоянному применению этого принципа, что, переходя к исследованию нелинейных проблем, нужно специально напомнить о неприменимости принципа наложения к нелинейным явлениям. Упустив это обстоятельство из вида, можно наделать грубых ошибок.

Поясним это на простом примере. Пусть дана нелинейная, в частности, квадратичная зависимость

$$y = ax^2.$$

Пусть далее

$$x = x_1 + x_2.$$

Тогда

$$y_1 = ax_1^2, \quad y_2 = ax_2^2,$$

но

$$y = a(x_1 + x_2)^2 = ax_1^2 + ax_2^2 + 2ax_1x_2 \neq y_1 + y_2,$$

т. е. квадрат суммы не равен сумме квадратов. Поэтому, например, при действии в нелинейной цепи суммы нескольких эдс уже нельзя

свести решение задачи к суммированию токов, порождаемых каждой эдс в отдельности; приходится учитывать сразу все составляющие эдс.

Понятно, что положение еще более усложнится при переходе от нелинейных алгебраических соотношений к нелинейным дифференциальным уравнениям.

Вынужденный отказ от принципа наложения очень осложняет исследования. К сожалению, таких мощных и универсальных методов исследования, какими располагает линейная теория, нелинейная теория пока не знает. Исследование нелинейных проблем проводится при помощи ряда специализированных приемов и методов, имеющих различную силу и область применения. Ниже дана краткая предварительная характеристика этих методов; в дальнейшем нам предстоит познакомиться с ними в их действии применительно к различным радиотехническим проблемам.

1. Алгебраическая задача нахождения $y(t)$ по заданному $x(t)$ и по заданной нелинейной характеристике

$$y = f(x)$$

может быть всегда решена с любой степенью точности путем графического построения.

2. Та же задача может быть решена и аналитически, т.е. может быть найдена функция

$$y(t) = f[x(t)],$$

если дано аналитическое выражение характеристики системы

$$y = f(x).$$

При этих условиях возможно (но далеко не всегда) решение и дифференциального уравнения.

3. Нужно заметить, что очень часто характеристика системы известна из опыта, т.е. задана графиком или табличными данными. Нахождение аналитического выражения, достаточно точно представляющего действительную зависимость, является, таким образом, очень важной операцией в нелинейной теории. Эту операцию называют *аппроксимацией* нелинейной характеристики.

Аппроксимация должна по возможности удовлетворять следующим требованиям: а) аналитическое выражение должно представлять аппроксимируемую зависимость с достаточной точностью, б) оно должно допускать выполнение над ним требуемых математических операций и в) оно должно быть как можно более кратким.

Существующие методы аппроксимации всем этим пожеланиям одинаково хорошо, к сожалению, не отвечают. Применяются следующие главнейшие методы: а) аппроксимация степенным многочленом, б) аппроксимация ломаной линией, в) аппроксимация трансцендентными функциями.

Первый метод состоит в том, что функция

$$y = f(x)$$

представляется в виде

$$y = a_0 + a_1x + a_2x^2 + \dots + a_nx^n = \sum_{k=0}^n a_kx^k.$$

Число членов степенного многочлена определяется требуемой точностью. Во всяком случае приближение должно быть достаточным, чтобы правильно отобразить хотя бы основные качественные черты явления. Это замечание относится, конечно, ко всякой аппроксимации.

Аппроксимация ломаной линией, называемая также кусочно-линейной аппроксимацией, состоит в том, что график действительной зависимости заменяется ломаной, тем лучше представляющей характеристику, чем больше берется точек излома. Аппроксимация ломаной особенно удобна для представления характеристик выпрямителей, работы лампы «с отсечкой» и т. п.

Этот вид аппроксимации интересен тем, что позволяет воспользоваться специальным методом решения нелинейного дифференциального уравнения — так называемым «методом сшивания». Сущность его состоит в том, что поведение нелинейной системы представляется различными линейными уравнениями для различных интервалов значений переменной величины, в пределах которых нелинейная характеристика представлена отрезком прямой.

Решения линейных уравнений нужно затем «сшить» на границах интервалов, т. е. приравнять конечные значения для одного интервала начальным значениям для другого.

Метод аппроксимации трансцендентными функциями, например экспоненциальными или гиперболическими, позволяет получить очень компактные и в то же время дающие очень хорошее приближение аппроксимирующие выражения, но, к сожалению, дальнейшая аналитическая обработка этих выражений сильно затруднена.

4. Большую роль в нелинейной радиотехнике играют так называемые «квазилинейные» методы. Сущность их сводится вкратце к тому, что уравнение нелинейной системы заменяется линейным по форме уравнением; однако в качестве коэффициентов этого уравнения вводятся некоторые *средние* параметры, определенные с учетом имеющейся в действительности нелинейности. Поэтому средние значения параметров оказываются зависящими от величин, характеризующих режим системы. Квазилинейные методы особенно удобны для исследования *установившихся* режимов нелинейных систем, когда колебания близки к синусоидальным.

5. При исследовании *режимов установления* применяется также относящийся к квазилинейным *метод медленно меняющихся амплитуд*. Общая идея этого метода состоит в том, что, исследуя вопросы установления колебаний в том или ином радиотехническом устройстве,

мы интересуемся законом изменения амплитуды и частоты, предполагая, что эти величины изменяются медленно. Это обстоятельство позволяет заметно упростить задачу.

При известных предположениях (подробнее об этом говорится ниже) мы получаем обычно уравнение, порядок которого ниже порядка исходного уравнения. Метод медленно меняющихся амплитуд применяется главным образом в случаях, когда заранее известно, что форма колебаний близка к синусоидальной.

6. Наиболее общим методом является, пожалуй, графический метод построения интегральных кривых данного нелинейного дифференциального уравнения. Он известен под названием *метода построения поля направлений*. Сущность и техника применения этого метода будут изложены ниже. Пока заметим лишь, что метод позволяет решать нелинейные задачи наиболее общего характера; он годится как для установившихся, так и для неустойчивых режимов, как для синусоидальных, так и для резко несинусоидальных релаксационных колебаний. Удобный вариант метода, сводящий нахождение наклона интегральной кривой к чисто геометрическому построению, известен под названием *построения Льенара*.

7. Наконец, общим методом, позволяющим (иногда с исключительной простотой) сделать важные заключения общезначимого характера, является *энергетический метод*. Сущность его состоит в рассмотрении энергетического баланса нелинейной системы, т.е. в прямом применении закона сохранения энергии. Этот метод не всегда позволяет разъяснить детали, но дает надежные общие соотношения, относящиеся к существу явления.

Глава 8

НЕЛИНЕЙНЫЕ ЦЕПИ И ОСНОВНЫЕ НЕЛИНЕЙНЫЕ ПРОЦЕССЫ

В этой главе рассматриваются те вопросы нелинейной теории, математическое описание которых может быть (хотя бы в первом приближении) сведено к недифференциальным (т.е. алгебраическим и трансцендентным) нелинейным уравнениям. Вопросы же, для рассмотрения которых требуется решение нелинейных дифференциальных уравнений и являющиеся, очевидно, более сложными, разбираются в последующих главах.

§ 65. Нелинейная цепь

Простейшая нелинейная цепь содержит, кроме источника эдс, *нелинейное сопротивление*, т.е. элемент, сопротивление которого не является постоянной величиной, а зависит от тока. Предполагается, что реактивных элементов, т.е. емкостей и индуктивностей, в цепи нет. Поэтому уравнение такой нелинейной цепи будет уравнением алгебраическим (а не дифференциальным).

Свойства нелинейного сопротивления выражаются непосредственно зависимостью тока i , протекающего через сопротивление, от напряжения u на его концах. Эта зависимость носит название *вольтамперной характеристики* и может быть записана как

$$i = f(u). \quad (224)$$

Задавшись напряжением в функции времени, можем найти ток также как функцию времени. Это делается или графически, или аналитически. Графическое построение показано на рис. 218. В левой части рисунка изображены, сверху — характеристика нелинейного сопротивления $i = f(u)$, снизу — $u = f(t)$ (выбрано синусоидальное напряжение). График $i = f(t)$ получается построением, аналогичным построению третьей проекции по двум данным (рис. 218, справа). Заметим, что форма кривой тока отличается от формы кривой напряжения. Это типично для нелинейной цепи и имеет очень большое значение.

Формы кривых напряжения и тока были бы подобны только в случае линейной цепи, когда вольтамперная характеристика представляется прямой линией.

На рис. 218 на синусоидальное напряжение наложено еще постоянное напряжение U_0 . Оно определяет постоянную составляющую тока I_0 в отсутствие колебаний. Точка на характеристике с координатами (U_0, I_0) называется *рабочей точкой*. Изменяя U_0 , можно перемещать рабочую точку по характеристике. Выбор рабочей точки определяется желаемым результатом, как будет видно из дальнейшего. Отметим, что постоянная составляющая тока при колебаниях I' отличается от тока покоя I_0 . Это также типично для нелинейной цепи.

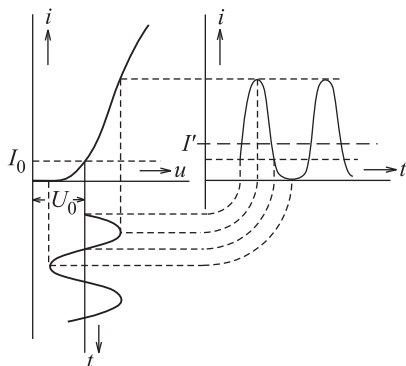


Рис. 218

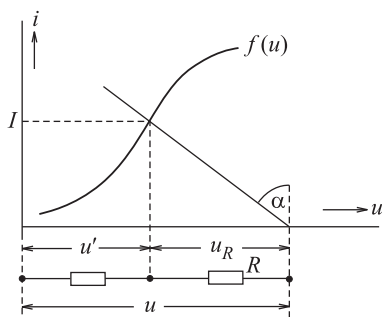


Рис. 219

Легко решается графически вопрос о распределении напряжений в цепи, в которой последовательно включены обычное (линейное, т. е. постоянное) сопротивление R и нелинейное сопротивление, определенное своей вольтамперной характеристикой. Соответствующее построение показано на рис. 219.

Вольтамперная характеристика нелинейного сопротивления построена, как обычно, в координатах (i, u) . Из точки на оси абсцисс, соответствующей приложенному к цепи напряжению, проведена прямая под углом α к вертикальной оси. Эта прямая изображает вольтамперную характеристику линейного сопротивления R ; ее наклон $\operatorname{tg} \alpha$ непосредственно выражает величину сопротивления и численно равен R .

Координаты точек пересечения прямой с вольтамперной характеристикой определяют, во-первых, ток в цепи, а во-вторых, распределение напряжений на элементах цепи, как показано на рис. 219 (u_R означает падение напряжения на линейном сопротивлении, а u' — на нелинейном).

Аналогично решается задача о распределении напряжений в случае соединения двух нелинейных сопротивлений.

Если число нелинейных сопротивлений, включенных в цепь, больше двух, то построение целесообразно изменить. На рис. 220 показаны три вольтамперные характеристики. Если нелинейные сопротивления включены параллельно, то вольтамперная характеристика результирующего сопротивления получается поординатным сложением характеристик. Если же сопротивления включены последовательно, то результирующая характеристика получается *поабсциссным* сложением всех характеристик.

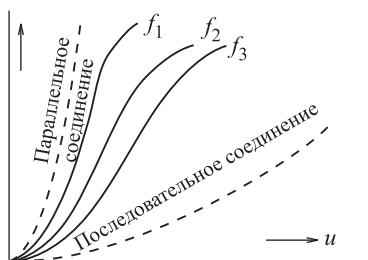


Рис. 220

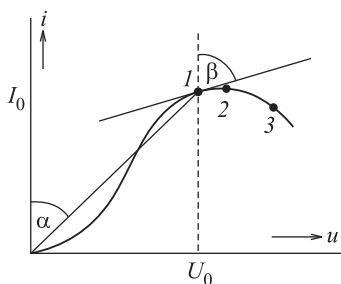


Рис. 221

Теперь займемся вопросом об *определении* нелинейного сопротивления. Простое линейное сопротивление определяется как отношение падения напряжения на нем к протекающему через него току, т. е.

$$R = U/I.$$

Это — постоянная величина, и определение не требует никаких дополнений. Но нелинейное сопротивление — величина переменная, и ее можно определить по-разному. На рис. 221 изображена вольтамперная характеристика нелинейного сопротивления. Если подать на сопротивление постоянное напряжение U_0 , то получим постоянный ток I_0 ; отношение этих величин можно назвать *сопротивлением постоянному току*

$$R_{=} = \frac{U_0}{I_0} (= \operatorname{tg} \alpha).$$

Графически $R_{=}$ определяется наклоном секущей; ясно, что эта величина зависит от приложенного напряжения U_0 .

Положим теперь, что на постоянное напряжение наложено небольшое переменное напряжение Δu , вызывающее приращение тока Δi . Вводят определение *сопротивления переменному току*, как предела

$$R_{\sim} = \lim_{\Delta U \rightarrow 0} \frac{\Delta u}{\Delta i} = \frac{du}{di} = \frac{1}{f'(u)} (= \operatorname{tg} \beta).$$

Определенную таким образом величину называют также *дифференциальным сопротивлением*. Графически сопротивление переменному току выражается наклоном касательной к вольтамперной характеристике

ке в рабочей точке, как показано на рис. 221. Ясно, что сопротивления постоянному и переменному току не равны (они равны только для линейного сопротивления). Более того: сопротивление постоянному току всегда конечно и положительно, тогда как сопротивление переменному току может быть равно нулю, бесконечности (точка 2 на рис. 221, в которой касательная к кривой $f(u)$ горизонтальна, т. е. конечному приращению напряжения соответствует нулевое приращение тока) и даже становится отрицательным (точка 3 на рис. 221, в которой положительному приращению напряжения соответствует отрицательное приращение тока). Понятие отрицательного сопротивления играет в радиотехнике очень большую роль. Как увидим дальше, отрицательное сопротивление может рассматриваться как источник энергии, тогда как положительное сопротивление является ее потребителем.

Кроме названных сопротивлений постоянному и переменному току, вводится еще одно определение, а именно определение *среднего сопротивления*. Оно учитывает нелинейность особым образом. Положим, что через нелинейное сопротивление протекает синусоидальный ток. Вследствие нелинейности напряжение на сопротивлении будет несинусоидально. Однако напряжение будет периодическим, и основная частота его будет равна частоте тока. Если взять теперь амплитуду первой гармоники напряжения и отнести ее к амплитуде тока, то это отношение и определит то, что мы называем средним сопротивлением. Таким образом, определение среднего сопротивления связано с нахождением первой гармоники периодического отклика нелинейной системы на синусоидальное воздействие. К сожалению, среднее сопротивление не может быть определено графически по вольтамперной характеристике. Его находят чаще всего аналитически, располагая аппроксимирующим характеристику выражением.

Все три приведенные выше определения могут применяться не только к сопротивлению, но и к другим параметрам, постоянным в линейной системе и переменным в нелинейной. Так, например, можно говорить о проводимости для постоянного тока, проводимости для переменного тока и средней проводимости. В радиотехнике широко применяется понятие *средней крутизны* электронной лампы. Эта величина определяется как отношение амплитуды первой гармоники анодного тока к амплитуде синусоидального напряжения на управляющей сетке. Чтобы составить себе представление о зависимости средней крутизны от амплитуды управляющего напряжения, рассмотрим рис. 222. На этом рисунке представлена характеристика триода, аппроксимированная ломаной линией. Рабочая точка выбрана посередине наклонного участка. При увеличении амплитуды управляющего напряжения крутизна остается постоянной, пока сеточное напряжение не выходит за пределы прямолинейного участка характеристики. При этом анодный ток (переменная составляющая) синусоидален. При дальнейшем увеличении амплитуды напряжения на сетке синусоида анодного тока отсекается сверху и снизу. В пределе получаем в анодной цепи импуль-

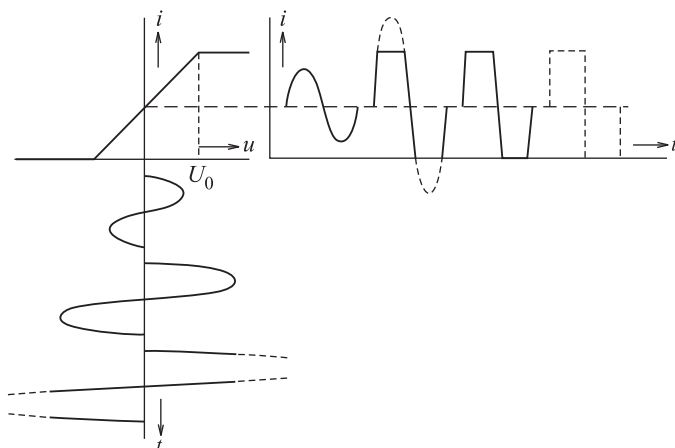


Рис. 222

сы прямоугольной формы: амплитуда первой гармоники анодного тока остается неизменной при неограниченном росте амплитуды напряжения на сетке $U_{ст}$. Значит, средняя крутизна S_{cp} стремится к нулю, как показано на рис. 223. В действительности характеристика лампы представляется кривой линией, а потому средняя крутизна изменяется плавно. В дальнейшем нам предстоит вычислить среднюю крутизну при степенной аппроксимации характеристик триода.

Речь шла до сих пор о нелинейном активном сопротивлении, а также о проводимости и крутизне, имеющей размерность проводимости. Но в радиотехнике имеют широкое применение также нелинейные *реактивные* элементы, т. е. нелинейная емкость и нелинейная индуктивность.

Обычная — линейная — емкость определяется соотношением

$$q = Cu,$$

где q — заряд, u — напряжение. Таким образом,

$$C = \frac{q}{u}.$$

Это — постоянная величина, не зависящая от напряжения. В нелинейном же случае заряд выражается некоторой нелинейной функцией напряжения

$$q = f_C(u)$$

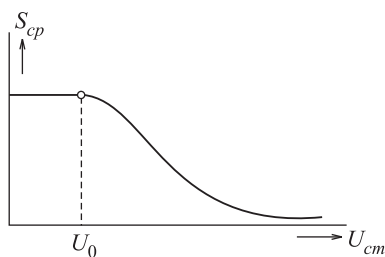


Рис. 223

и можно ввести определение дифференциальной емкости

$$C_{\sim} = \frac{dq}{du} = f'_C(u).$$

Определенная таким образом нелинейная емкость уже не постоянная величина, а функция приложенного напряжения.

Аналогично для линейной индуктивности

$$\psi = Li,$$

где ψ — потокосцепление; i — ток. Для нелинейной же индуктивности

$$\psi = f_L(i)$$

и дифференциальная нелинейная индуктивность есть

$$L_{\sim} = \frac{d\psi}{di} = f'_L(i).$$

Очень важным отличием нелинейных реактивных элементов от нелинейного сопротивления является то, что сами по себе дифференциальные емкость и индуктивность ни при каких обстоятельствах не могут быть отрицательными. Иначе говоря, например, в случае нелинейной емкости при *увеличении* напряжения не может происходить *уменьшение* заряда. Таким образом характеристические функции $f_C(u)$ и $f_L(i)$ являются монотонными функциями своих аргументов. Обычно эти функции нечетны.

§ 66. Преобразование спектра в нелинейной цепи

В § 65 уже отмечалось, что синусоидальное напряжение вызывает в нелинейной цепи несинусоидальный ток. Вообще, форма тока отличается от формы напряжения. И, наоборот, если задан синусоидальный ток, то получается несинусоидальное напряжение.

Если рассмотреть вопрос со спектральной точки зрения, то оказывается, что спектр отклика нелинейной системы всегда отличается по составу от спектра воздействия; обычно спектр отклика богаче. Это общее положение, широко используемое в радиотехнике, поясним на нескольких простых примерах.

Пусть на нелинейное сопротивление с характеристикой

$$i = f(u)$$

действует синусоидальное напряжение

$$u = U_m \sin \omega t.$$

Найдем ток, для чего развернем характеристику в степенной ряд

$$i = \sum_{k=0}^{\infty} a_k u^k = \sum_{k=0}^{\infty} a_k U_m^k \sin^k \omega t.$$

При возведении синуса в k -ю степень получаем ряд колебаний кратных частот, причем наивысшая кратность равна k . Так, квадрат синуса дает постоянную составляющую (нулевую частоту) и вторую гармонику; куб дает первую и третью гармоники и т. д. Ниже приведены степени синуса и косинуса:

$$\begin{aligned}\sin^2 x &= \frac{1}{2}(1 - \cos 2x), \\ \sin^3 x &= \frac{1}{4}(3 \sin x - \sin 3x), \\ \sin^4 x &= \frac{1}{8}(3 - 4 \cos 2x + \cos 4x), \\ \sin^5 x &= \frac{1}{16}(10 \sin x - 5 \sin 3x + \sin 5x), \\ &\dots\dots\dots \\ \cos^2 x &= \frac{1}{2}(1 + \cos 2x), \\ \cos^3 x &= \frac{1}{4}(3 \cos x + \cos 3x), \\ \cos^4 x &= \frac{1}{8}(3 + 4 \cos 2x + \cos 4x), \\ \cos^5 x &= \frac{1}{16}(10 \cos x + 5 \cos 3x + \cos 5x), \\ &\dots\dots\dots\end{aligned}$$

Из сказанного следует, что при воздействии синусоидального напряжения на нелинейное сопротивление мы получаем в общем случае ток с бесконечным спектром, содержащим все гармоники. В частном случае характеристика может представляться не бесконечным рядом, а конечным степенным многочленом. В этом случае и ток будет иметь ограниченный спектр, причем номер наивысшей гармоники будет определяться степенью многочлена. Так, например, при квадратичной характеристике вида

$$i = a_0 + a_1 u + a_2 u^2$$

в составе тока не будет гармоник выше второй.

Рассмотрим теперь случай, когда на нелинейное сопротивление действуют два напряжения с разными частотами, т. е. пусть

$$u = u_1 + u_2 = U_{m1} \sin \omega_1 t + U_{m2} \sin \omega_2 t.$$

Подставляя это выражение в формулу характеристики, получим для тока

$$i = f(u_1 + u_2) = \sum_{k=0}^{\infty} a_k (U_{m1} \sin \omega_1 t + U_{m2} \sin \omega_2 t)^k.$$

Возводя двучлен в скобках в k -ю степень и группируя затем члены суммы, можно убедиться, что в составе тока образуются частоты, вы-

ражаемые общим соотношением

$$n\omega_1 \pm m\omega_2,$$

где n и m — любые целые числа, не исключая нуля. При $m = 0$ получаем последовательность гармоник $n\omega_1$; при $n = 0$ — гармоник $m\omega_2$. Частоты, отвечающие не равным нулю значениям m и n , называются *комбинационными частотами*. Они могут быть как выше, так и ниже исходных частот ω_1 и ω_2 .

При вычислении комбинационных спектров полезна таблица произведений степеней синуса и косинуса, часть которой приведена ниже:

$$\begin{aligned}\sin x \cos x &= \frac{1}{2} \sin 2x, \\ \sin x \cos^2 x &= \frac{1}{4} (\sin x + \sin 3x), \\ \sin x \cos^3 x &= \frac{1}{8} (2 \sin 2x + \sin 4x), \\ &\dots\dots\dots \\ \cos x \sin^2 x &= \frac{1}{4} (\cos x - \cos 3x), \\ \cos x \sin^3 x &= \frac{1}{8} (2 \sin 2x - \sin 4x), \\ \sin^2 x \cos^2 x &= \frac{1}{8} (1 - \cos 4x), \\ &\dots\dots\dots\end{aligned}$$

§ 67. Умножение частоты

Примером использования обогащения спектра нелинейной системой может служить *умножение частоты*. Задача умножения частоты состоит в получении частоты, кратной исходной. Умножение частоты применяется в тех случаях, когда непосредственное генерирование требуемой частоты почему-либо неудобно.

Идея умножения крайне проста; она заключается в том, что напряжение исходной частоты действует в нелинейной цепи, в которой получается богатый гармониками ток. Гармоника желаемой кратности выделяется фильтром.

Возможная схема умножения показана на рис. 224. Синусоидальное напряжение с частотой ω приложено к сетке триода; контур в выходной цепи настроен на n -ю гармонику, которая и выделяется на контуре; остальные гармоники отфильтровываются, так как контур представляет для них малое сопротивление. На сетку триода подается регулируемое отрицательное смещение, достаточное, чтобы запереть триод. В резуль-

тате этого происходит так называемая «работа с отсечкой».

Аппроксимируем характеристику триода ломаной линией, как показано на рис. 225; для наших целей это достаточное приближение. Выберем рабочую точку левее излома характеристики. При таком положении рабочей точки и синусоидальном напряжении на сетке в анодной цепи будут проходить лишь короткие импульсы тока.

Остальная часть синусоиды (намеченная пунктиром) будет отсечена — отсюда и название этого режима. Режим вполне характеризуется *углом отсечки*, т. е. тем значением аргумента ωt (отсчитанным от момента, когда анодный ток имеет максимальное значение), при котором ток

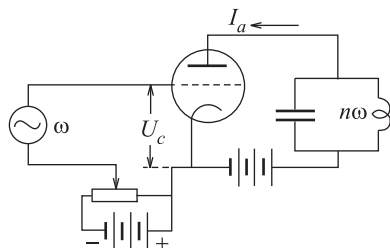


Рис. 224

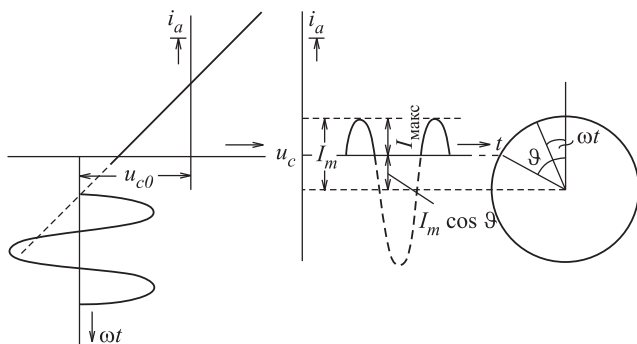


Рис. 225

обращается в нуль. Зная угол отсечки, можно разложить выражение для анодного тока в ряд Фурье. Анодный ток как функцию времени можно представить в виде

$$i_a = I_m(\cos \omega t - \cos \vartheta) \quad (i_a > 0)$$

(обозначения указаны на рис. 225). Для амплитуды n -й гармоники имеем по известной формуле

$$\begin{aligned} i_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} i_a(\omega t) \cos n\omega t d\omega t = \frac{2}{\pi} I_m \int_0^{\vartheta} \cos nx (\cos x - \cos \vartheta) dx = \\ &= \frac{1}{\pi} I_m \left[\frac{\sin(n+1)\vartheta}{n+1} + \frac{\sin(n-1)\vartheta}{n-1} - \frac{2 \cos \vartheta \sin n\vartheta}{n} \right] = I_m F_n(\vartheta). \end{aligned}$$

Практически может быть задано не I_m (определяемое амплитудой напряжения на сетке), а $I_{\text{макс}}$ — пиковое значение анодного тока.

Между этими величинами существует очевидная связь (см. рис. 225):

$$I_{\text{макс}} = I_m(1 - \cos \vartheta).$$

Выражая теперь амплитуду n -й гармоники анодного тока через пиковое значение анодного тока имеем

$$i_n = I_{\text{макс}} \frac{F_n(\vartheta)}{1 - \cos \vartheta} = I_{\text{макс}} f_n(\vartheta).$$

Функции $f_n(\vartheta)$ угла отсечки для различных n представлены графиками рис. 226. Эти графики

показывают, как следует выбирать угол отсечки для получения наибольшей амплитуды желаемой гармоники. Так, например, при утроении частоты следует брать $\vartheta = 40^\circ$, при учетверении — 30° и т. д.

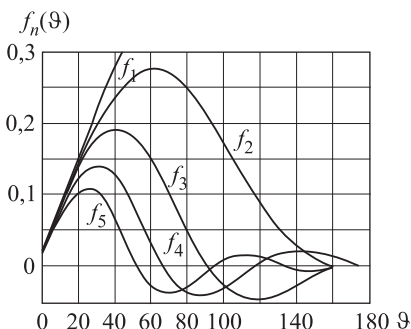


Рис. 226

§ 68. Линеаризация и нелинейные искажения

Нелинейные зависимости играют в радиотехнике двоякую роль. В целом ряде случаев нелинейность определяет самое существо процессов и используется для получения желаемого результата. Так обстоит дело, например, при умножении частоты, при детектировании и во многих других важных для радиотехники случаях. Но иногда нелинейность не только не требуется для осуществления того или иного процесса, но является вредной и нежелательной. Так обстоит, например, дело при усилении.

Типичное для нелинейной системы искажение формы и, следовательно, обогащение спектра колебаний необходимо для умножения частоты. Но при усилении задача состоит как раз в обратном — в сохранении по возможности исходной формы колебания. Всякие изменения формы колебания, происходящие вследствие нелинейности тех или иных элементов схемы, называются *нелинейными искажениями*. Необходимо стремиться к их уменьшению, т. е. к устранению влияния нелинейности.

Строго линейных систем не существует; всякая реальная система нелинейна. Таким образом, линейными мы называем системы, нелинейностью которых можно при данных обстоятельствах пренебречь. Возможность пренебречь нелинейностью зависит не только от характеристики системы, но и от режима работы. Так, например, при усилении стремятся расположить рабочую точку посередине прямолинейного участка характеристики лампы, а амплитуду управляющего

напряжения ограничивают так, чтобы она не выходила за пределы прямолинейного участка. В действительности характеристика всегда кривая; говоря о прямолинейном участке, мы подразумеваем участок с достаточно малой кривизной.

Замена нелинейной зависимости приближенной линейной, а также меры, направленные к улучшению приближения, называют вообще *линеаризацией*.

Понятие о линеаризации поясним сначала на частном примере. Пусть дана квадратичная характеристика

$$i = au^2.$$

Эта нелинейная зависимость никак не может быть представлена приближенной линейной. Но положим, что напряжение u состоит из постоянной слагающей U_0 и переменной слагающей Δu , которая нас и интересует. Тогда

$$i = au^2 = a(U_0 + \Delta u)^2 = a(U_0^2 + 2U_0 \Delta u + \Delta u^2).$$

Как видим, ток имеет три составляющие: постоянную

$$I_0 = aU_0^2,$$

линейную

$$i_1 = 2aU_0 \Delta u$$

и квадратичную

$$i_2 = a \Delta u^2.$$

Постоянная составляющая нас обычно не интересует. Переменную составляющую тока запишем в виде

$$\Delta i = i_1 + i_2 = 2aU_0 \Delta u + a \Delta u^2 = 2aU_0 \Delta u \left(1 + \frac{1}{2} \frac{\Delta u}{U_0}\right).$$

Второе слагаемое можно отбросить при условии

$$\frac{\Delta u}{U_0} \ll 1,$$

т. е. линеаризация возможна с тем лучшим приближением, чем меньше Δu и чем больше U_0 . Этот результат поясняется графиком рис. 227, на котором представлена зависимость $i = au^2$. Из рисунка видно, что увеличение U_0 переводит рабочую точку на участок параболы с меньшей кривизной. Одновременно возрастает крутизна.

Уменьшение Δu позволяет линеаризировать любую нелинейную зависимость. Чтобы показать это, положим, что задана нелинейная характеристика общего вида

$$i = f(u).$$

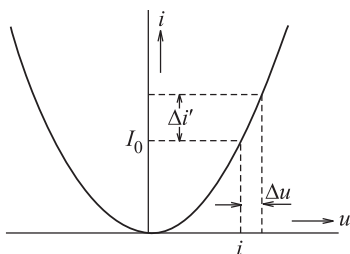


Рис. 227

Разложим выражение для тока в ряд Тейлора около точки U_0 :

$$i = f(U_0) + f'(U_0)\Delta u + f''(U_0)\frac{\Delta u^2}{2} + f'''(U_0)\frac{\Delta u^3}{6} + \dots$$

Если Δu — достаточно малая величина, то членами высших порядков можно пренебречь и записать линеаризованное выражение для тока в виде

$$i \approx f(U_0) + f'(U_0) \Delta u.$$

Постоянная составляющая определяется ординатой характеристики в рабочей точке, а крутизна — производной, т. е. наклоном касательной к характеристике в той же точке.

Может оказаться, что для данной характеристики $f(u)$ касательная к характеристике в рабочей точке горизонтальна, т. е. $f'(U_0) = 0$. В таком случае нужно сместить рабочую точку, изменив U_0 . Так, например, в только что рассмотренном случае квадратичной зависимости мы имели:

$$f(u) = au^2, \quad f'(U_0) = 2aU_0,$$

и если бы мы взяли $U_0 = 0$, то получили бы $f'(U_0) = 0$.

Нелинейные искажения в почти линейной системе нужно каким-то образом оценивать и измерять. Для этого существует несколько методов; все они основаны на том, что нелинейность вызывает обогащение спектра, которое не происходит в строго линейной системе.

Первый метод состоит в том, что на вход системы задается синусоидальное воздействие, а на выходе системы определяется действующее значение всех гармоник, кроме первой. Отношение этой величины к действующему значению первой гармоники (основной частоты) называется *коэффициентом гармоник*:

$$k = \frac{\sqrt{I_2^2 + I_3^2 + \dots}}{I_1},$$

где через I_k обозначены амплитуды гармоник. Считают, что нелинейные искажения практически не ощущаются, если коэффициент гармоник не превосходит 1%. Техническими условиями на аппаратуру задается обычно коэффициент гармоник не свыше 5%.

Другой метод состоит в подаче на вход системы одновременно двух синусоидальных напряжений с различными частотами. При этом, как было показано в § 66, образуется богатый спектр комбинационных частот. Коэффициент, характеризующий нелинейные искажения, определяется так же, как и коэффициент гармоник.

В недавнее время предложен новый способ определения нелинейных искажений. Он основан на измерении суммарной мощности в некоторой полосе частот при подаче на вход воздействия со сплошным спектром, из которого предварительно вырезана фильтром та же самая полоса частот. Если бы система была линейна, то частоты, вырезанные на входе, отсутствовали бы и на выходе. В результате же нелинейности в пределах вырезанной полосы появляются комбинационные частоты и гармоники — так называемый *продукт нелинейности*. Этот продукт и измеряется. Новый метод ближе соответствует действительным условиям работы аппаратуры и дает более правильную оценку нелинейных искажений.

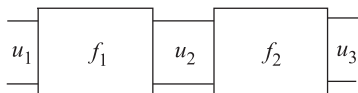


Рис. 228

Следует еще указать, что нелинейные искажения можно корректировать аналогично тому, как корректируются частотные искажения. Положим, что имеется нелинейный четырехполюсник с характеристикой

$$u_2 = f_1(u_1).$$

Добавим к этому четырехполюснику второй, служащий в качестве корректора (рис. 228). Его характеристика

$$u_3 = f_2(u_2)$$

должна быть подобрана так, чтобы нелинейные искажения были устранены, т. е. чтобы

$$u_3 = Au_1.$$

Это условие дает

$$\frac{u_3}{u_1} = \frac{f_2[f_1(u_1)]}{u_1} = A$$

или

$$f_2[f_1(u_1)] = Au_1,$$

т. е. функция f_2 должна быть обратной ¹⁾ функции f_1 . Например, если

$$f_1(u_1) = au_1^2,$$

то должно быть

$$f_2(u_2) = b\sqrt{u_2}.$$

¹⁾ Напомним определение обратной функции. Если

$$y = f(x),$$

то решая это уравнение относительно x , получим

$$x = \varphi(y).$$

Функция φ обратна функции f . При этом

$$\frac{\varphi[f(x)]}{x} = \frac{f[\varphi(y)]}{y} = 1.$$

При таких соотношениях получим

$$u_3 = b\sqrt{u_2} = b\sqrt{a} u_1 = Au_1.$$

§ 69. Выпрямление

Объектом выпрямления является переменный ток; цель выпрямления состоит в получении постоянного тока (например, для питания радиотехнических устройств).

Простейшая схема выпрямления показана на рис. 229. В схему включен нелинейный элемент, относительно которого предполагается, что он обладает конечной проводимостью в одном направлении и нулевой (или очень малой) проводимостью в другом. Направление,

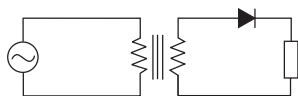


Рис. 229

в котором выпрямитель пропускает ток, называется *прямым*, противоположное — *обратным*. Соответственно обозначаются и проводимости или сопротивления. Отношение прямой проводимости к обратной называется *коэффициентом выпрямления*.

Реальная и идеализированные характеристики выпрямителя показаны на рис. 230, а и б. Если в схеме рис. 229 при характеристике рис. 230, б действует синусоидальное напряжение, то в цепи проте-

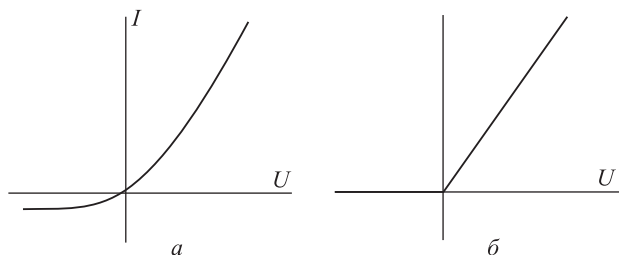


Рис. 230

кают импульсы тока, отвечающие положительным полупериодам напряжения (рис. 231). Это есть так называемое *однополупериодное* или *однотактное* выпрямление. В цепи протекает пульсирующий, т. е. не меняющий направления, но изменяющийся по величине ток. Он содержит требуемую постоянную составляющую, но, кроме того, также и переменную составляющую со сложным спектром. Постоянная составляющая равна

$$I_0 = \frac{1}{\pi} I_m.$$

Спектр переменной составляющей определим, разложив в ряд Фурье выражение для тока, имеющее вид

$$i = I_m \cos \omega t \quad \left(-\frac{T}{4} < t < \frac{T}{4} \right)$$

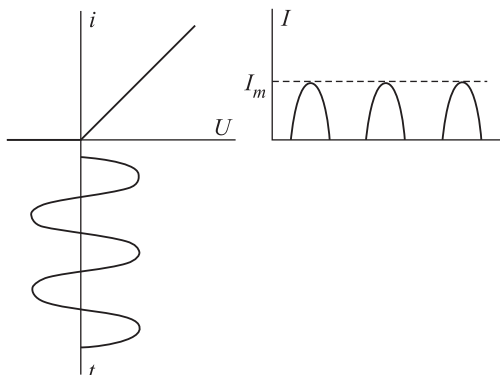


Рис. 231

или, вводя $x = \omega t$,

$$i = I_m \cos x \quad \left(-\frac{\pi}{2} < x < \frac{\pi}{2}\right).$$

На протяжении остальной доли периода ток равен нулю (рис. 232). Таким образом,

$$\begin{aligned} I_n &= \frac{1}{n} I_m \int_{-\pi/2}^{\pi/2} \cos x \cos nx \, dx = \frac{1}{\pi} I_m \left[\frac{\sin(n+1)x}{n+1} + \frac{\sin(n-1)x}{n-1} \right] \Bigg|_0^{\pi/2} = \\ &= \begin{cases} \frac{1}{2} I_m & \text{при } n = 1, \\ \frac{2}{\pi} I_m \frac{1}{n^2 - 1} & \text{при } n \text{ четном,} \\ 0 & \text{при } n \text{ нечетном.} \end{cases} \end{aligned}$$

Итак, в составе выпрямленного тока имеется первая гармоника с амплитудой

$$I_1 = \frac{1}{2} I_m$$

и, кроме того, все четные гармоники, амплитуды которых равны:

$$I_2 = \frac{2}{3\pi} I_m, \quad I_4 = \frac{2}{15\pi} I_m, \dots$$

Гораздо чаще применяется *двухполупериодное (двухтактное)* выпрямление по схеме рис. 233. Действие выпрямителей в обеих ветвях схемы таково, что результирующую характеристику можно представить, как показано на рис. 234. Постоянная составляющая при двухполупериодном выпрямлении, очевидно, вдвое больше, чем при однополупериодном; в этом состоит одно из преимуществ двухполупериодного выпрямления. Что касается спектра тока, то первая гармоника в нем отсутствует. Это и понятно, так как основная частота выпрямленного

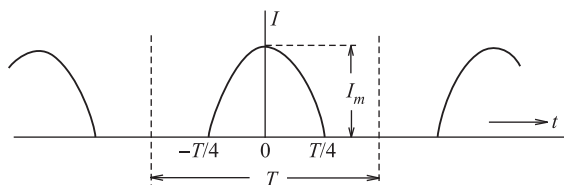


Рис. 232

тока (при условии симметрии обеих ветвей схемы) вдвое больше частоты питающего напряжения и, следовательно, наименьшая частота в составе тока отвечает второй гармонике напряжения. Спектральный состав выпрямленного тока таков:

$$I_0 = \frac{2}{\pi} I_m, \quad I_2 = \frac{4}{3\pi} I_m, \quad I_4 = \frac{4}{15\pi} I_m, \dots$$

Для очистки выпрямленного тока от переменной составляющей применяются обычные фильтры нижних частот. Типичная схема выпрямителя с двойным диодом изображена на рис. 235. Здесь выясняется еще

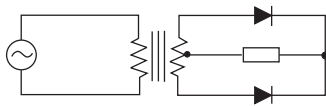


Рис. 233

одно преимущество двухполупериодного выпрямления. Дело в том, что основная частота переменной составляющей при двухполупериодном выпрямлении вдвое выше, чем при однополупериодном. Это позволяет соответственно повысить граничную частоту фильтра, а следовательно,

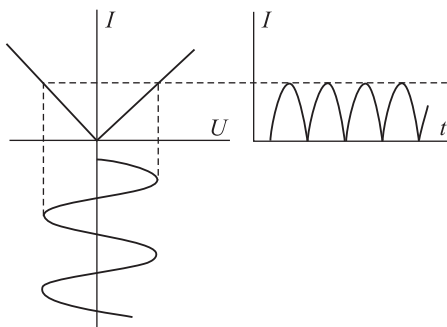


Рис. 234

так как при этом уменьшается величина пульсаций и возрастает их частота. Пульсация тем меньше, чем больше число фаз. Нужно пояснить, что мы не рассматриваем в этом параграфе работу выпрямителя в целом, включая фильтр. Схема выпрямителя с фильтром описывается

нелинейным дифференциальным уравнением, решение которого представляет более сложную задачу. Этот вопрос будет рассмотрен в § 94.

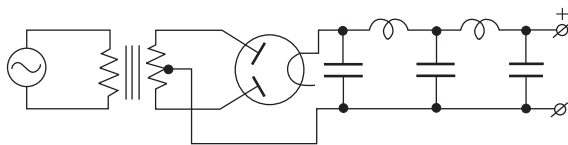


Рис. 235

Следует еще добавить, что выпрямление часто применяется для измерения переменных напряжений и токов посредством электроизмерительных приборов постоянного тока (магнитоэлектрической системы). Все современные многошкальные универсальные приборы постоянного и переменного тока являются детекторными приборами, т. е. снабжены выпрямительными элементами, что и позволяет пользоваться одним и тем же прибором для измерений как на постоянном, так и на переменном токе. Ламповый вольтметр — основной радиоизмерительный прибор — также представляет собой магнитоэлектрический прибор в комбинации с ламповой выпрямительной схемой. Однако требования к выпрямителю для измерительных целей несколько иные, чем требования к выпрямителю для целей питания. Дело в том, что измерительные приборы переменного тока должны измерять действующие значения измеряемых напряжений или токов, которые могут быть и несинусоидальными. А это условие выполняется лишь в том случае, если характеристика выпрямителя квадратична. Подробно этот вопрос изложен в курсе радиоизмерений.

§ 70. Детектирование

Объектом детектирования является модулированное колебание; цель детектирования состоит в извлечении из модулированного колебания сигнала, т. е. колебания низкой модулирующей частоты. В этом параграфе рассмотрим только детектирование колебаний, модулированных по амплитуде.

Процесс детектирования амплитудно-модулированных (АМ) колебаний имеет много общего с процессом выпрямления; различие состоит главным образом в объектах и желаемых результатах. Простейшая схема детектора содержит нелинейный элемент с такой же характеристикой, как и характеристика выпрямителя на рис. 230. Разберем сначала качественную картину явлений при детектировании. Пусть амплитудно-модулированное напряжение действует в цепи с нелинейным элементом — детектором, — обладающим ломаной характеристикой (рис. 236). Ток в цепи детектора представляет собой импульсы высокой частоты с огибающей, воспроизводящей форму огибающей модулиро-

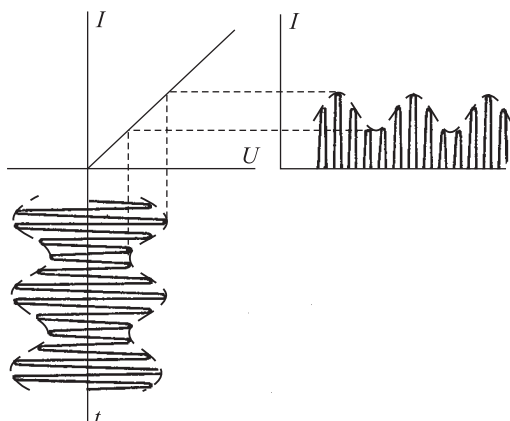


Рис. 236

ванного колебания. В составе тока получаются несущая частота и ее гармоники. Кроме того, имеется постоянная составляющая, которая нас обычно не интересует. И, наконец, в составе тока появляется низкочастотная составляющая, которую мы стремимся выделить.

Нужно подчеркнуть, что в составе модулированного колебания низкочастотной *составляющей* нет; как известно, например, при синусоидальной модуляции модулированное колебание состоит из трех синусоидальных колебаний: колебания несущей частоты и двух спутников, имеющих также высокие частоты.

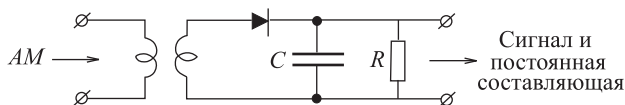


Рис. 237

Для выделения низкочастотной составляющей, появляющейся в результате детектирования, нужно отфильтровать высокие частоты и постоянную составляющую. Последняя устраняется обычно сама собой, так как не проходит через усилительный тракт, содержащий трансформаторы и переходные емкости. Высокие частоты могут быть отфильтрованы простейшим звеном, состоящим из емкости и сопротивления. Схема детектора принимает вид, показанный на рис. 237. По виду она не отличается от схемы выпрямителя, но выполняет другую задачу: выделяет низкую частоту сигнала, тогда как схема выпрямителя предназначена для устранения переменной составляющей и получения постоянного тока. Это различие определяет и различие в требованиях к фильтрам. Фильтрующее звено в схеме детектора на рис. 237 должно устранить высокую частоту, но сохранить низкую частоту. Для этого необходимо, чтобы сопротивление R было много

больше емкостного сопротивления при высокой (несущей) частоте ω_0 и много меньше емкостного сопротивления при низкой частоте сигнала Ω . Это условие можно записать в виде неравенств

$$\frac{1}{\omega_0 C} \ll R \ll \frac{1}{\Omega C}.$$

К тому же заключению можно придти и иным путем. Можно рассматривать звено RC как интегрирующую схему, назначение которой состоит в усреднении тока. Однако схема должна усреднить ток по высокой частоте, но не по низкой, так как полное усреднение дало бы на выходе лишь постоянный ток. Эффект интегрирования или усреднения определяется постоянной времени звена, точнее соотношением между постоянной времени

$$\tau = RC$$

и периодом тока. Усреднение происходит тогда, когда постоянная времени значительно превосходит период. Если же, напротив, постоянная времени значительно меньше периода, то усреднения не происходит и напряжение на выходе звена следует за напряжением на входе. Исходя из этих соображений, можно записать условие правильной работы звена RC в следующей форме:

$$T_0 \ll \tau \ll T_1,$$

где T_0 и T_1 — соответственно периоды несущей частоты и частоты сигнала. Это неравенство совпадает с предыдущим, если принять во внимание, что

$$T_0 = \frac{2\pi}{\omega_0}, \quad T_1 = \frac{2\pi}{\Omega}.$$

Часто применяется двухтактное детектирование. При этом несущая частота подвергается двухполупериодному выпрямлению, частота удваивается и ее легче отделить.

Рассмотрим теперь процесс детектирования аналитически. Положим, что детектирование двухтактное и что характеристика детектора может быть записана в виде $i = k|u|$ (рис. 238). Возьмем случай синусоидальной модуляции частотой Ω несущей частоты ω_0 . Имеем следующее выражение для модулированного напряжения:

$$u = U_m(1 + m \sin \Omega t) \sin \omega_0 t.$$

Для того чтобы найти ток в цепи детектора, мы должны взять модуль (абсолютное значение) этого выражения. Для этого нужно найти модуль каждого из сомножителей, так как модуль произведения равен произведению модулей сомножителей. Заметим, что коэффициент модуляции m — величина положительная, не превосходящая единицы. Поэтому множитель $(1 + m \sin \Omega t)$ не

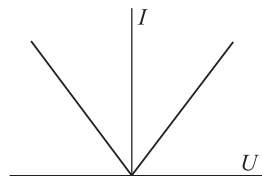


Рис. 238

может быть меньше нуля, а стало быть, остается найти только модуль последнего множителя. Таким образом,

$$i = kU_m(1 + m \sin \Omega t) |\sin \omega_0 t|.$$

Но $|\sin \omega_0 t|$ — это не что иное, как выпрямленная синусоида при двухполупериодном выпрямлении. Разложение этой функции в ряд Фурье известно

$$|\sin \omega_0 t| = \frac{2}{\pi} \left(1 - 2 \sum_{k=1}^{\infty} \frac{\cos 2k \omega_0 t}{4k^2 - 1} \right).$$

Итак,

$$\begin{aligned} i &= \frac{2}{\pi} kU_m(1 + m \sin \Omega t) \left(1 - 2 \sum_{k=1}^{\infty} \frac{\cos 2k \omega_0 t}{4k^2 - 1} \right) = \\ &= \frac{2}{\pi} kU_m \left[1 + m \sin \Omega t - 2 \sum_{k=1}^{\infty} \frac{\cos 2k \omega_0 t}{4k^2 - 1} - \right. \\ &\quad \left. - 2m \sum_{k=1}^{\infty} \frac{\sin \Omega t \cos 2k \omega_0 t}{4k^2 - 1} \right]. \end{aligned}$$

Как видим, цель детектирования достигнута: мы получили в составе тока слагаемое $(1 + m \sin \Omega)$, т. е. модулирующий сигнал. Кроме того, в составе тока имеется ряд высоких частот: четные гармоники несущей частоты с частотами $2k\omega_0$ (первая сумма) и спутники с частотами $2k\omega_0 \pm \Omega$ (вторая сумма). Спектры модулированного колебания до и после детектирования изображены на рис. 239.

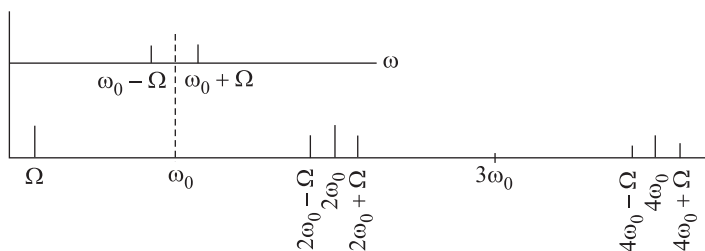


Рис. 239

Мы рассматривали до сих пор детектирование при характеристике детектора в виде ломаной линии (рис. 236 или 238). Детектор с такого рода характеристикой называется *линейным*. Само собой разумеется, что термин «линейный» имеет в виду не характеристику детектора в целом, а лишь форму ветвей характеристики по одну сторону (рис. 236) или по обе стороны (рис. 238) точки излома.

Из рис. 236 ясно, что только линейный детектор не искажает формы огибающей. Для того чтобы показать, как влияет форма характеристики детектора на результат детектирования, рассмотрим в качестве примера детектирование амплитудно-модулированного колебания *квадратичным* детектором.

Пусть характеристика детектора дана соотношением

$$i = ku^2.$$

Подставляя в эту формулу выражение синусоидально модулированного по амплитуде колебания, находим

$$i = kU_m^2(1 + m \sin \Omega t)^2 \sin^2 \omega_0 t = \frac{k}{2}U_m^2 \left[1 + \frac{m^2}{2} + 2m \sin \Omega t - \frac{m^2}{2} \cos 2\Omega t - \left(1 + \frac{m^2}{2} \right) \cos 2\omega_0 t - m \sin (2\omega_0 - \Omega)t + m \sin (2\omega_0 + \Omega)t + m^2 \cos 2(\omega_0 - \Omega)t + m^2 \cos 2(\omega_0 + \Omega)t \right].$$

Как видим, спектр получается ограниченным; он содержит только пять высоких частот, а именно:

$$2\omega_0, \quad 2\omega_0 \pm \Omega, \quad 2(\omega_0 \pm \Omega).$$

Но сигнал оказывается искаженным: кроме частоты Ω , в спектре детектированного колебания появляется также частота 2Ω , т. е. вторая гармоника сигнала, искажающая его форму. Спектры модулированного колебания до и после квадратичного детектирования показаны на рис. 240.

Таким образом оказывается, как и следовало ожидать, что детектор с квадратичной характеристикой не пригоден для детектирования

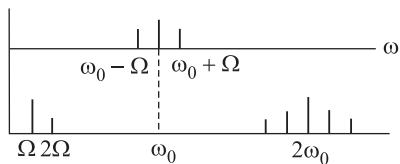


Рис. 240

амплитудно-модулированных колебаний. Тем не менее квадратичное детектирование играет в радиотехнике большую роль. Оно применяется, например, для детектирования *биений*, когда желательно получить в результате детектирования чисто синусоидальное колебание разностной частоты. Именно так обстоит дело в гетеродинных генераторах звуковой частоты. Принцип действия этих генераторов состоит в том, что колебания двух генераторов высокой частоты складываются, образуя биения. Биения детектируются, и получается колебание низкой частоты, равной разности двух высоких частот. Если сделать один из генераторов высокой частоты перестраивающимся, то можно получить плавно изменяющуюся в широких пределах низкую частоту. Так, например, чтобы получить низкую частоту, изменяющуюся в интервале 0–10 кГц, можно взять два генератора высокой частоты: один, скажем,

с частотой 100 кГц, а другой — перестраивающийся от 100 до 110 кГц, т. е. всего на 10 %.

Очевидным преимуществом гетеродинных генераторов звуковой частоты является простота управления: весь диапазон перекрывается поворотом одной ручки, управляющей небольшим конденсатором переменной емкости в контуре одного из генераторов высокой частоты.

Разберем вопрос о детектировании биений. Для начала положим, что амплитуды обоих колебаний высокой частоты равны, т. е.

$$u = U_m(\sin \omega_1 t + \sin \omega_2 t).$$

Преобразуя это выражение, получим

$$u = 2U_m \sin \frac{\omega_1 + \omega_2}{2} t \cos \frac{\omega_1 - \omega_2}{2} t = 2U_m \sin \omega_0 t \sqrt{\frac{1}{2}(1 + \cos \Omega t)},$$

где

$$\omega_0 = \frac{\omega_1 + \omega_2}{2}, \quad \Omega = \omega_1 - \omega_2,$$

т. е. через ω_0 обозначена средняя высокая частота, а через Ω — низкая частота биений. Форма подводимого к детектору напряжения показана на рис. 241.

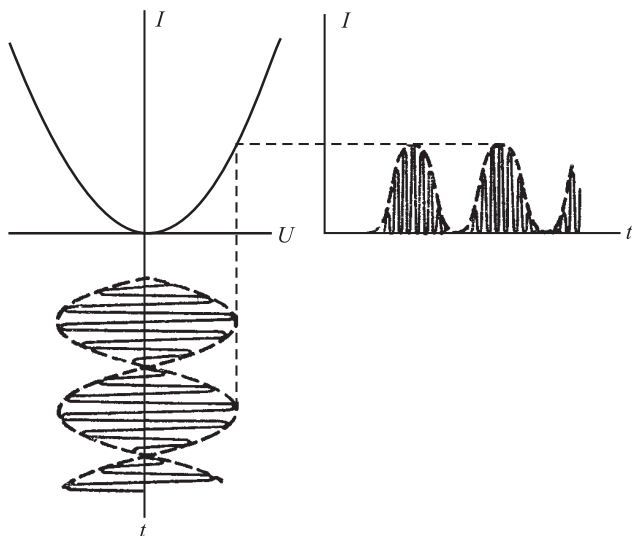


Рис. 241

Следует заметить, что биения можно рассматривать, как колебания высокой частоты ω_0 , модулированные частотой биений. Но, что существенно, модуляция несинусоидальная, модулирующий множитель имеет вид

$$\sqrt{\frac{1}{2}(1 + \cos \Omega t)}.$$

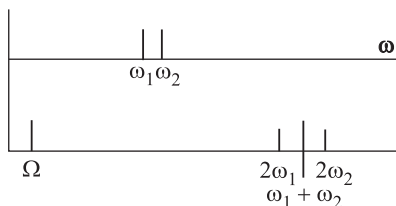


Рис. 242

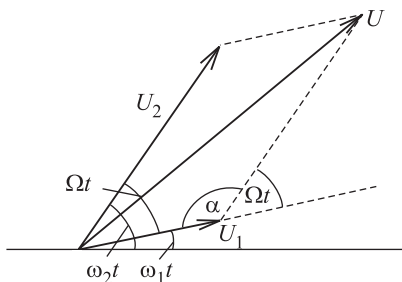


Рис. 243

Если теперь применить квадратичное детектирование, то получится

$$\begin{aligned} i &= ku^2 = 2kU_m^2(1 + \cos \Omega t) \sin^2 \omega_0 t = kU_m^2(1 + \cos \Omega t)(1 - \cos 2\omega_0 t) = \\ &= kU_m^2 \left[1 + \cos \Omega t - \frac{1}{2} \cos 2\omega_1 t - \frac{1}{2} \cos 2\omega_2 t - \cos (\omega_1 + \omega_2)t \right], \end{aligned}$$

т. е. чисто синусоидальное колебание разностной частоты Ω , удвоенные исходные высокие частоты и их суммы. Спектр биений до и после квадратичного детектирования показан на рис. 242. Заключение о целесообразности квадратичного детектирования биений остается в силе и в том более общем случае, когда амплитуды обоих колебаний высокой частоты оказываются неравными. Пусть

$$u = U_1 \sin \omega_1 t + U_2 \sin \omega_2 t.$$

На рис. 243 представлена векторная диаграмма, поясняющая сложение двух напряжений с неравными частотами. Амплитуда результирующего напряжения

$$U_m = \sqrt{U_1^2 + U_2^2 - 2U_1U_2 \cos \alpha}.$$

Но, как видно из рисунка,

$$\alpha = \pi - \Omega t$$

и, следовательно,

$$-\cos \alpha = \cos \Omega t.$$

Таким образом,

$$U_m = U_1 \sqrt{1 + \frac{U_2^2}{U_1^2} + 2 \frac{U_2}{U_1} \cos \Omega t}.$$

Эта формула представляет собой выражение для изменяющейся амплитуды, т. е. для огибающей биений при неравных амплитудах составляющих колебаний. При квадратичном детектировании, в результате возведения в квадрат, радикал пропадет, и мы получим и в этом случае чисто синусоидальное колебание разностной частоты Ω .

Следует заметить, что если амплитуды обоих напряжений сильно различаются, т. е.

$$U_2 \ll U_1$$

или

$$U_2 \gg U_1,$$

то синусоидальное напряжение низкой частоты может быть получено и при линейном детектировании. Это следует из того, что, например, при $U_2 \ll U_1$

$$U_m = U_1 \sqrt{1 + \frac{U_2^2}{U_1^2} + 2 \frac{U_2}{U_1} \cos \Omega t} \approx \approx U_1 \left(1 + 2 \frac{U_2}{U_1} \cos \Omega t \right)^{1/2} \approx U_1 \left(1 + \frac{U_2}{U_1} \cos \Omega t \right).$$

Нужно еще добавить, что квадратичное детектирование играет большую роль в схемах модуляции и преобразования частоты, о чем речь будет ниже.

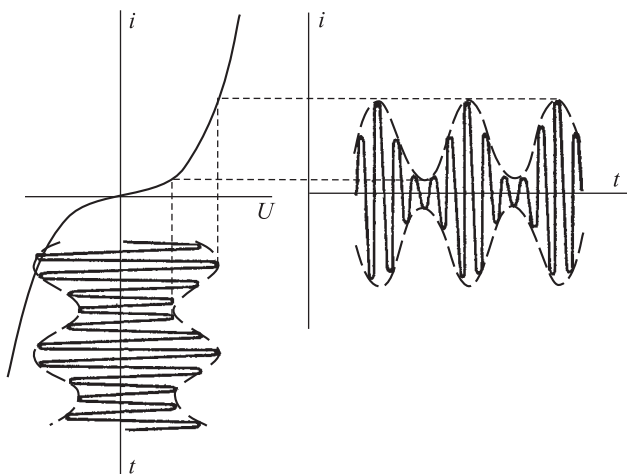


Рис. 244

Рассмотрим теперь свойства характеристик детекторов. Характеристика детектора должна быть нелинейной — это очевидно. Но этого мало. Если бы мы выбрали нелинейную характеристику вида

$$i = ku^3,$$

то никакого детектирования не получили бы, так как колебание, изменив свою форму, осталось симметричным относительно оси времени (рис. 244). Для детектирования годятся только такие нелинейные зависимости, которые несимметричны относительно оси U и которые вследствие этого дают постоянную составляющую тока при приложении

переменного напряжения. Зависимости такого рода должны содержать четную часть.

Напомним определения.

Четной функцией называется функция, удовлетворяющая условию

$$f(x) = f(-x).$$

Нечетная функция удовлетворяет условию

$$f(x) = -f(-x).$$

Примеры четных функций: x^2 , x^4 , $\cos x$ и т. п.

Примеры нечетных функций: x , x^3 , $\sin x$ и т. п.

Степенная функция x^n четна или нечетна в зависимости от того, четен или нечетен показатель степени n .

Детектор детектирует, а выпрямитель выпрямляет лишь поскольку его характеристика содержит четную часть. Любая функция может быть разложена на четную и нечетную части.

Пусть

$$f(x) = f_{\text{ч}}(x) + f_{\text{н}}(x),$$

где $f_{\text{ч}}$ — четная часть; $f_{\text{н}}$ — нечетная. Изменим знак аргумента. Тогда по определению получим

$$f(-x) = f_{\text{ч}}(x) - f_{\text{н}}(x).$$

Складывая и вычитая, найдем:

$$f_{\text{ч}} = \frac{1}{2}[f(x) + f(-x)],$$

$$f_{\text{н}} = \frac{1}{2}[f(x) - f(-x)].$$

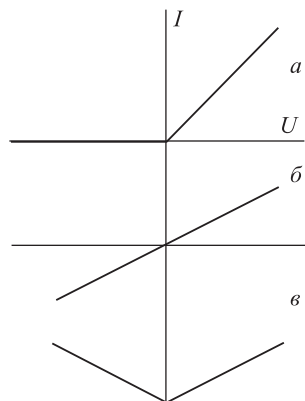


Рис. 245

В качестве примера приведем разложение на четную и нечетную части ломаной характеристики рис. 245, а. Аналитически эту характеристику можно записать в виде

$$i = \begin{cases} ku & U > 0, \\ 0 & U < 0. \end{cases}$$

Четная часть характеристики будет

$$i_{\text{ч}} = \frac{k}{2}|u|,$$

а нечетная

$$i_{\text{н}} = \frac{k}{2}u$$

(рис. 245, б и в). Нетрудно видеть, что, сложив (б) и (в), мы получим (а). Таким образом, аналитически ломаную характеристику на рис. 245, а можно представить формулой

$$i = \frac{k}{2}(|u| + u),$$

которая годится для всех значений u как положительных, так и отрицательных.

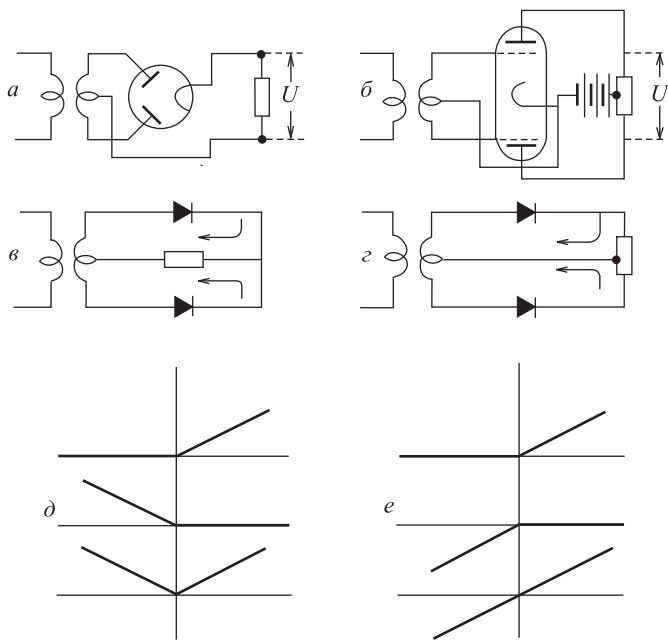


Рис. 246

Нечетная часть в характеристике детектора или выпрямителя может содержаться — она не производит действия. Естественно, однако, пользоваться выпрямителями и детекторами с чисто четными характеристиками. Такими характеристиками обладают все симметричные двухтактные схемы — нечетные члены характеристики устраняются в этих схемах в результате того, что на обе ветви схемы выпрямленное или детектируемое напряжение подается с противоположными знаками (в противофазе), а токи обеих ветвей складываются. В этом смысле схема выпрямителя или детектора обладает свойствами, противоположными свойствам схемы двухтактного усилителя. В последнем напряжение также подается в обе ветви в противофазе, но снимаемое напряжение пропорционально не сумме, а разности токов. Таким образом, двухтактная схема выпрямителя использует только четную часть

характеристики нелинейного элемента, а двухтактная схема усилителя использует только нечетную часть характеристики. Вот почему в двухтактной схеме усилителя выпадают все четные гармоники. Само собой разумеется, что из нечетных гармоник в случае усиления стремятся сохранить только первую — это соответствует неискаженному усилению, когда результирующая характеристика — прямая. В пояснение этих соображений на рис. 246, *а* и *б* показаны схемы двухтактного выпрямителя на двойном диоде и двухтактного усилителя на двойном триоде. На рис. 246, *в* и *г* даны те же схемы в обобщенном виде. Наконец, рис. 246, *д* и *е* показывают построение результирующих характеристик в предположении, что характеристика отдельного элемента — ломаная.

§ 71. Детектирование ЧМ и ИМ

В этом параграфе кратко излагаются общие принципы детектирования, применяемые при других (кроме АМ) видах модуляции, в частности при частотной модуляции (ЧМ) и некоторых видах импульсной модуляции (ИМ).

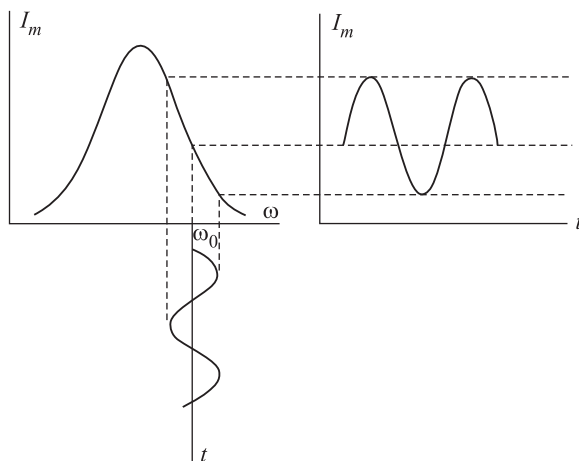


Рис. 247

Обычный прием детектирования частотно-модулированных колебаний состоит в том, что ЧМ колебание сначала преобразовывается в АМ колебание, которое затем детектируется, как описано в § 70. В качестве простейшего преобразователя ЧМ в АМ может служить расстроенный относительно несущей частоты контур. При отсутствии модуляции, т.е. при $\omega = \omega_0$ рабочая точка располагается на скате резонансной кривой контура, как показано на рис. 247. В нижней части этого рисунка изображен закон изменения частоты во времени. При изменении частоты изменяется амплитуда тока I_m высокой частоты в контуре,

как показано в правой части рисунка. Но изменение амплитуды — это и есть амплитудная модуляция. Таким образом, сняв с контура напряжение, остается продетектировать и отфильтровать его обычным образом. Принципиальная схема детектирования ЧМ принимает вид, показанный на рис. 248.

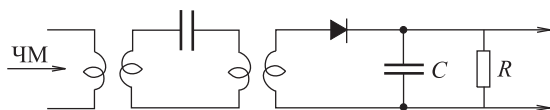


Рис. 248

В современной технике применяются более совершенные схемы частотных детекторов. Дело в том, что во избежание нелинейных искажений характеристика частотного детектора должна быть прямой. Другими словами, крутизна частотного детектора, выражаемая производной $\frac{dI_m}{d\omega}$, должна быть постоянной в пределах рабочего участка. При широкополосной ЧМ удовлетворить этому требованию при помощи простого контура затруднительно. Поэтому и применяются симметричные схемы, обладающие более благоприятными характеристиками.

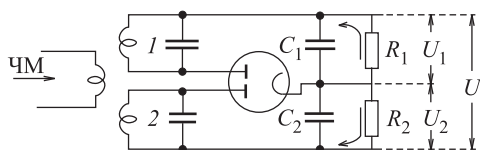


Рис. 249

Пример такой схемы приведен на рис. 249. Два контура 1 и 2 настраиваются обычно на крайние значения частоты, т. е. на $\omega_0 + \Delta\omega$ и $\omega_0 - \Delta\omega$. Каждый из контуров преобразует ЧМ в АМ, как описано выше. АМ колебания детектируются соответствующими половинами двойного диода. Комбинации RC служат фильтрами. Низкочастотные напряжения U_1 и U_2 возникают на сопротивлениях R_1 и R_2 . Эти напряжения пропорциональны ординатам резонансных кривых соответственно контуров 1 и 2. Но так как напряжения на R_1 и R_2 противоположны по знаку (см. стрелки на схеме, указывающие направления токов), то с выхода схемы снимается *разность* напряжений U_1 и U_2 . Характеристика частотного детектора, т. е. зависимость выходного напряжения u от частоты ЧМ колебания на входе, получается путем вычитания двух резонансных кривых, как показано на рис. 250. Частотные детекторы описанного типа относятся к числу так называемых *дискриминаторов* (различителей).

Перейдем к импульсной модуляции. Основные виды модуляции импульсов следующие: 1) по высоте (по «амплитуде») — АИМ, 2) по частоте следования — ЧИМ, 3) по длительности — ДИМ и 4) по фазе — ФИМ.

Детектирование АИМ сигнала ничем по существу не отличается от детектирования обычного амплитудно-модулированного колебания. Различие состоит лишь в составе высокочастотной части спектра, которая все равно должна быть отфильтрована.

То же относится и к ДИМ, хотя это на первый взгляд неочевидно. Но нужно учесть, что составляющая низкой частоты после обычного детектирования ИМ сигнала пропорциональна *площади* импульса. Следовательно, безразлично, что именно будет изменяться в огибающей импульса при модуляции — высота или длительность. Таким образом, для детектирования как АИМ, так и ДИМ пригодна обычная схема детектора, например схема рис. 237.

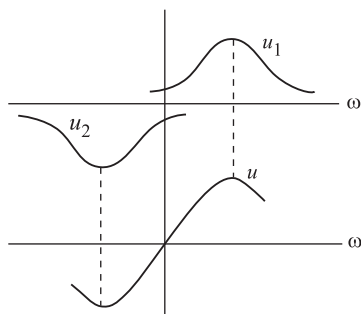


Рис. 250

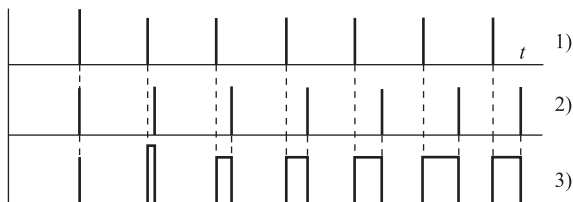


Рис. 251

Иначе обстоит дело с фазо-импульсной модуляцией. При ФИМ, как известно, площадь импульса остается неизменной. Модуляция состоит лишь в смещении импульса относительно тактовой точки. Поэтому для детектирования ФИМ нужно прибегнуть к предварительному преобразованию ФИМ в АИМ или ДИМ или вообще в такую форму, в которой изменяется площадь импульса. Для преобразования ФИМ в ДИМ может быть, например, применена схема, содержащая электронное реле, включаемое тактовыми импульсами и выключаемое фазо-модулированными импульсами (или наоборот), как показано на рис. 251, на котором 1 — включающие тактовые импульсы, 2 — выключающие импульсы ФИМ, 3 — импульсы на выходе электронного реле.

Вопросы, связанные с детектированием при импульсно-кодовой модуляции, как более специальные, здесь не рассматриваются.

§ 72. Инерционная нелинейность

В предыдущем параграфе свойства нелинейной цепи были выражены зависимостью тока от напряжения

$$i = f(u). \quad (224)$$

Если напряжение заданным образом изменяется во времени, т.е. задана функция

$$u = f_1(t),$$

то можно найти и ток как функцию времени

$$i = f[f_1(t)] = f_2(t).$$

Эти три формулы соответствуют трем графикам на рис. 218.

Таким образом, соотношение (224) выражает связь между *мгновенными* значениями тока и напряжения. В условиях нелинейной цепи обычное понятие сопротивления неприменимо. В § 65 приведены употребительные определения сопротивления, пригодные для нелинейной цепи. Если же мы попытались бы определить сопротивление обычным образом как отношение *мгновенного* значения напряжения к *мгновенному* значению тока, то получили бы переменную величину, зависящую от *мгновенного* значения напряжения (а следовательно, и от времени).

Известен и другой вид нелинейности с существенно отличными свойствами — так называемая *инерционная нелинейность*. Она характеризуется тем, что обычным образом определяемое сопротивление зависит не от мгновенного значения напряжения, а от среднеквадратичного (т.е. действующего или эффективного значения). Можно также говорить о зависимости от амплитудного значения. В случае синусоидального значения между действующим значением и амплитудой имеется простая общеизвестная связь

$$U = \frac{1}{\sqrt{2}} U_m.$$

В случае же любого несинусоидального периодического напряжения пропорциональность между действующим и максимальным напряжениями сохраняется, но коэффициент пропорциональности зависит от формы кривой напряжения.

Главное состоит в том, что для периодического напряжения как действующее, так и максимальное значения остаются *постоянными*. А так как при инерционной нелинейности сопротивление зависит именно от этих значений напряжения (а не от мгновенного значения), то пока они остаются неизменными, остается постоянным и сопротивление. Для простоты и определенности будем полагать, что мы имеем дело с синусоидальным напряжением и что инерционная нелинейность

характеризуется зависимостью сопротивления от амплитуды приложенного напряжения.

Рассмотрим пример, который поможет уяснить существо дела. Известно, что сопротивление металлов зависит от температуры: для большинства металлов оно в той или иной степени возрастает с температурой. Но температура металлического проводника зависит от выделяемой в нем мощности, т. е. от тока и напряжения. Значит, сопротивление непостоянно, а зависит от приложенного напряжения; следовательно мы имеем дело с нелинейной цепью. Спрашивается, какова эта нелинейность?

Сопротивление непосредственно зависит от температуры. Пока температура остается постоянной, сопротивление также остается неизменным. Если бы температура изменялась, следуя за изменениями напряжения, то мы имели бы обычную нелинейность (т. е. сопротивление зависело бы от мгновенного значения напряжения). Но в обычных условиях температура не успевает следовать за изменениями мгновенного значения напряжения; в этом и проявляется *инерционность* металлического проводника.

Температура проводника определяется тепловым балансом: температура остается неизменной, если мощность электрического тока, выделяемая в проводнике в виде тепла, равна количеству тепла, отдаваемого проводником внешней среде за счет теплопроводности. Но мгновенная мощность тока колеблется; на протяжении периода она дважды принимает максимальное значение и дважды обращается в нуль. Поэтому и температура должна колебаться на протяжении периода. Эти колебания, однако, тем меньше, чем больше *постоянная времени* проводника!

Речь идет о *тепловой* постоянной времени, которая прямо пропорциональна теплоемкости проводника и обратно пропорциональна теплопроводности, определяющей отвод тепла во внешнюю среду. Постоянная времени и является количественной мерой инерционности. Если постоянная времени велика по сравнению с периодом напряжения, то температура проводника, а следовательно, и его сопротивление остаются практически постоянными при условии, что не изменяется средняя мощность¹⁾. Но средняя мощность может быть выражена через квадрат амплитуды напряжения. Итак, сопротивление металлического проводника в описанных условиях зависит не от мгновенного значения напряжения, а от амплитуды напряжения. Таким образом, мы имеем в этом примере дело с типичной инерционной нелинейностью.

Столь подробные рассуждения приведены здесь для того, чтобы пояснить, что при описании нелинейных свойств тех или иных цепей

¹⁾ Так, например, температура нити обычной лампы накаливания, питаемой переменным током, остается практически постоянной. Небольшие колебания температуры (а следовательно, и яркости) тем меньше, чем мощнее лампа, т. е. чем больше теплоемкость нити.

следует быть внимательным. Говоря о зависимости тех или иных параметров от тока или от напряжения, недостаточно сказать: «параметр X зависит от напряжения»; нужно обязательно указать, от какого именно значения напряжения — мгновенного, или амплитудного, или действующего — зависит данный нелинейный параметр. Как мы видели, это уточнение очень существенно.

В дальнейшем мы будем, говоря о нелинейности, подразумевать зависимость от мгновенных значений; случай инерционной нелинейности будет всякий раз оговариваться особо.

Наиболее существенным свойством инерционной нелинейности является то, что она не приводит к искажению формы колебания. Так как при неизменной амплитуде инерционное нелинейное сопротивление остается постоянным, то при синусоидальном приложенном напряжении в цепи получается синусоидальный ток.

§ 73. Ограничение и стабилизация

К числу простейших практических применений нелинейных сопротивлений относятся ограничение и стабилизация токов и напряжений.

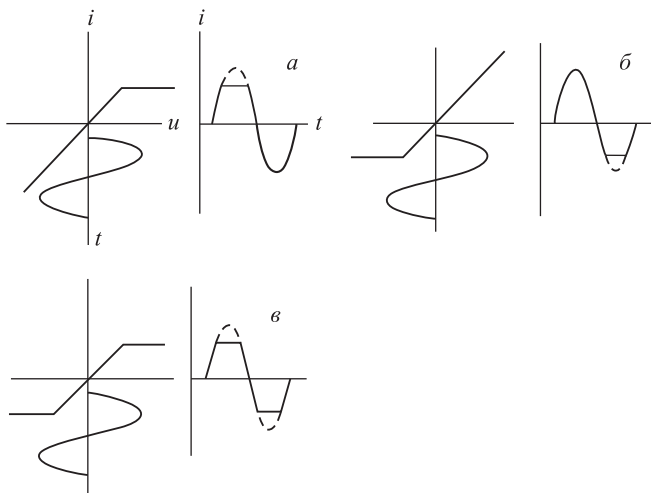


Рис. 252

Ограничение тока есть нелинейная операция, в результате которой мгновенные значения тока не выходят за заданные пределы, называемые *порогами* или *уровнями ограничения*. Возможно ограничение «сверху», когда ток не должен превосходить заданного уровня, и ограничение «снизу», когда фиксируется наименьшее значение тока. На рис. 252, *а* и *б* показаны операции ограничения сверху и снизу и требуемые для этого нелинейные характеристики цепи. Операции ограничения сверху и снизу могут быть совмещены, как показано

на рис. 252, в; получается двустороннее ограничение. Говоря попросту, ограничитель срезает все мгновенные значения, выходящие за порог ограничения.

Практические схемы ограничителей выполняются часто с применением диодов — вакуумных или полупроводниковых. Идеализированная характеристика диода показана на рис. 253.

Эту характеристику можно рассматривать как характеристику ограничителя тока снизу с пороговым значением $I = 0$ (заметим, что такое ограничение есть просто выпрямление, рассмотренное в § 70). Для того чтобы произвольным образом сместить характеристику диода, последовательно с ним включают дополнительное постоянное, так называемое *смещающее* напряжение E . Действие этого напряжения равносильно

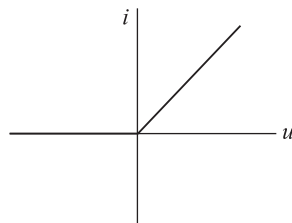


Рис. 253

смещению характеристики рис. 253 (или смещению рабочей точки) вправо или влево в зависимости от полярности смещающего напряжения. Для ограничения напряжения применяются последовательные и параллельные схемы. Действие последовательной схемы состоит в том, что при напряжении, превосходящем пороговое, диод оказывается запертым, ток не проходит и напряжение на выходе отсутствует. В параллельной схеме используется шунтирующее действие диода, включенного параллельно относительно большому сопротивлению нагрузки; когда диод заперт, напряжение на сопротивлении имеется, когда же он отпирается, напряжение падает практически до нуля. Основные схемы ограничителей на диодах показаны на рис. 254. Острые стрелки в условном изображении диода указывает направление тока (для вакуумного диода это есть направление от анода к катоду). На всех схемах имеется сопротивление R , назначение которого состоит в защите диода от чрезмерных токов (сопротивление открытого диода считается очень малым).

Для целей ограничения могут также применяться триоды в надлежащим образом подобранном режиме. Различные ограничители широко применяются в импульсной технике. При практическом расчете схем ограничителей приходится учитывать влияние паразитных параметров, что заметно усложняет дело.

Обратимся теперь к задаче стабилизации токов и напряжений. Она состоит в поддержании постоянства тока или напряжения питания радиотехнических устройств в изменяющихся условиях. Под изменением условий понимается обычно изменение напряжения источника ¹⁾.

¹⁾ Изменение условий может также состоять в изменении сопротивления нагрузки. Но этот случай не дает ничего нового; изменение сопротивления можно свести к эквивалентному изменению напряжения.

С принципиальной точки зрения можно разделить существующие стабилизаторы на две группы. В одних непосредственно используются нелинейные свойства тех или иных элементов схемы. Другие же представляют собой системы с обратной связью, т. е. по существу являются автоматическими регуляторами.

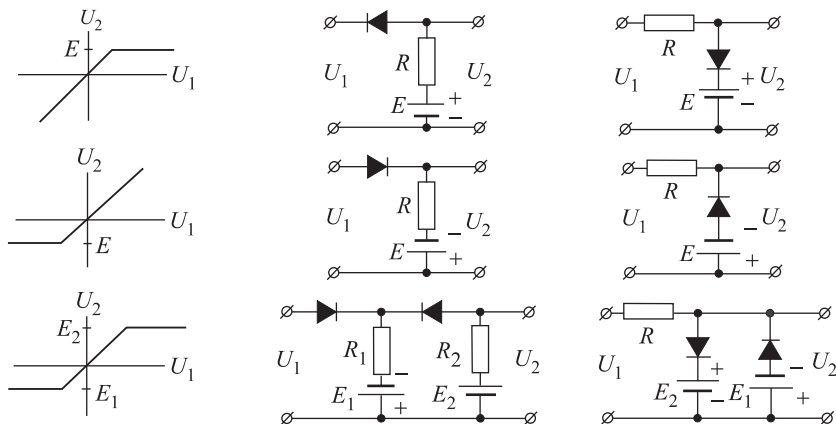


Рис. 254

Если речь идет о стабилизации *постоянного* напряжения (или тока), то задача не отличается по существу от задачи ограничения; например, при изменении входного напряжения (напряжения источника) выходное напряжение (напряжение питания) должно оставаться постоянным. Разница — не принципиальная, а практическая — состоит лишь в том, что обычно не требуют строгого постоянства стабилизованного напряжения; довольствуются тем, что относительные изменения выходного напряжения в заданное число раз меньше, чем относительные изменения входного напряжения.

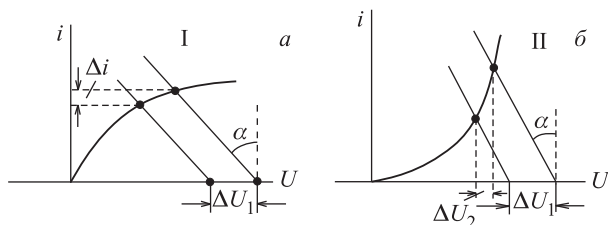


Рис. 255

Для целей стабилизации могут применяться любые нелинейные элементы с вольтамперными характеристиками двух типов, показанных на рис. 255, *а* и *б*; назовем их для краткости выпуклой (I) и вогнутой (II). Выпуклой характеристикой обладает, например, вакуумный

диод, вогнутой характеристикой — газоразрядная лампа, тиритовое сопротивление (содержащее карбид кремния). Рабочим является участок характеристики за изгибом, т. е. там, где наклон по возможности мал для выпуклой характеристики и по возможности велик для вогнутой. Простейшие схемы для стабилизации напряжения показаны на рис. 256, *а* и *б*. В схеме рис. 256, *а* с последовательным включением применяется нелинейный элемент с выпуклой характеристикой; в схеме рис. 256, *б* нелинейный элемент с вогнутой характеристикой включен параллельно нагрузке. Действие схемы рис. 256, *а* состоит в том, что при повышении входного напряжения и увеличении тока падение напряжения на нелинейном элементе резко возрастает, так что большая часть приращения входного напряжения падает на нелинейном элементе.

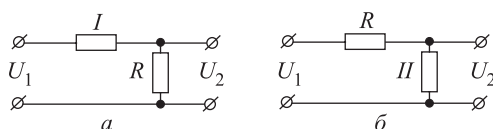


Рис. 256

Впрочем, действие обеих схем лучше всего поясняется графическим построением, аналогичным рис. 219. На рис. 255 проведены прямые под углом $\alpha = \text{arctg } R$. Входное напряжение получает приращение ΔU_1 . В результате этого (рис. 255, *а*) ток получает малое приращение ΔI , а выходное напряжение — малое приращение $\Delta U_2 = R \Delta I$ (рис. 256, *а*). В схеме рис. 256, *б* напряжение снимается с нелинейного элемента; построение, показывающее соотношение между ΔU_1 и ΔU_2 , дано на рис. 255, *б*.

Что касается стабилизации тока, то здесь дело обстоит еще проще: для стабилизации тока достаточно включить элемент с выпуклой характеристикой последовательно, а элемент с вогнутой характеристикой параллельно нагрузке.

Обратимся к стабилизации *переменного* напряжения. Стабилизация при помощи нелинейных элементов приводит в этом случае к искажению формы напряжения; при синусоидальном напряжении на входе мы получаем на выходе напряжение, резко отличающееся от синусоидального, что по ряду соображений очень нежелательно. Если требуется сохранить форму кривой входного напряжения, то нужно прибегнуть к *инерционной* нелинейности. При этом схемы рис. 256 сохраняются, а характеристики нелинейных элементов могут быть выпуклыми и вогнутыми, но в отличие от рис. 255 они будут выражать уже зависимость не мгновенных значений тока и напряжения, а *амплитуд* тока и напряжения. Примером инерционного нелинейного элемента с выпуклой характеристикой типа I является лампа накаливания или ее специальная форма — барретор, представляющий собой железную нить, помещенную в наполненный водородом баллон. Вогнутую харак-

теристику имеют полупроводниковые термосопротивления — термисторы. И у барреторов и у термисторов постоянная времени определяется тепловыми параметрами и может изменяться в широких пределах.

Не нужно, однако, думать, что инерционность может зависеть только от тепловых явлений. В более сложных схемах постоянные времени могут определяться электрическими параметрами тех или иных цепей. В пояснение на рис. 257 приведены (в сильно упрощенном виде) две схемы, относящиеся к числу систем с обратной связью. Схема рис. 257, а представляет собой безынерционный стабилизатор *постоянного* напряжения. Сопротивление триода изменяется вследствие изменения напряжения смещения, подаваемого по цепи обратной связи. Схема рис. 257, б представляет собой инерционный стабилизатор *переменного* напряжения. Эта схема отличается тем, что напряжение смещения получается путем выпрямления переменного напряжения, так что напряжение смещения на сетке, а следовательно, и сопротивление триода зависят от *амплитуды* переменного напряжения. Постоянная времени в этой схеме определяется произведением RC .

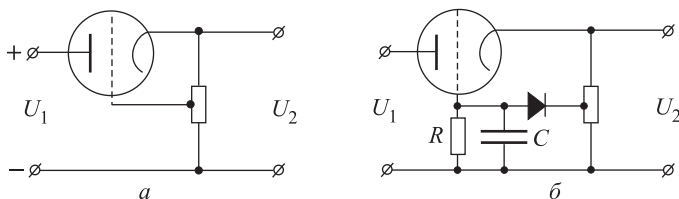


Рис. 257

В заключение заметим, что система АРУ (автоматическая регулировка усиления) также может рассматриваться как стабилизатор, так как ее назначение состоит в том, чтобы при больших изменениях уровня входного сигнала получить малые изменения уровня сигнала на выходе. Система АРУ в целом есть нелинейная система; однако форма сигнала не должна искажаться. Отсюда следует, что система АРУ должна обладать инерционной нелинейностью. Однако важно отметить, что в системе АРУ (и в других родственных системах, как-то: в компрессорах и экспандерах, шумопонижающих устройствах и т. п.) вопрос о выборе постоянных времени, характеризующих инерционность системы, приобретает решающее значение. В самом деле, постоянная времени должна быть велика по сравнению с периодом наименьшей частоты сигнала (в противном случае произойдет искажение формы сигнала), и одновременно постоянная времени должна быть мала по сравнению с периодом наиболее быстрого изменения уровня сигнала (в противном случае система не будет успевать следить за изменениями уровня, т. е. выполнять свое прямое назначение). Удовлетворить обоим требованиям иногда довольно затруднительно.

§ 74. Модуляция

Цель процесса модуляции — перенос спектра передаваемого (низкочастотного) сигнала в область высоких частот для того, чтобы он мог быть передан радиотехническими средствами, т.е. посредством излучения электромагнитных волн. Такое преобразование необходимо потому, что эффективное излучение может быть получено только при высоких радиочастотах. Кроме того, модуляция необходима, например, при частотном уплотнении. В этом случае берется столько несущих частот, сколько требуется образовать независимых каналов. Каждая несущая модулируется своим (т.е. данного канала) низкочастотным сигналом. Разнос между несущими определяется шириной модуляционного спектра.

Модуляция состоит в том, что в результате определенного воздействия сигнала на колебание высокой частоты он оказывается запечатленным в изменениях тех или иных параметров колебания высокой частоты. Высокая частота как бы несет в себе сигнал, сохраняя все его свойства, так что он может быть снова выделен на приемном конце. Отсюда и выразительное название *несущая* частота. Процесс выделения сигнала называется детектированием; он рассмотрен в предыдущих параграфах. Модуляция в своей основе есть *линейный* процесс, осуществляемый в линейной системе с переменными параметрами.

Для начала поясним механизм амплитудной модуляции на примере крайне простой схемы, применявшейся в свое время в маломощных телефонных передатчиках. Схема представлена на рис. 258. Генератор высокой частоты связан индуктивно с антенной. В антенну последовательно включен обычный угольный микрофон. Как известно, под воздействием звуковых колебаний мембрана микрофона колеблется и, надавливая на заключенный в микрофоне угольный порошок, изменяет его проводимость.

Предположим, что микрофон строго линеен, т.е. что его проводимость есть линейная функция звукового давления. Тогда для проводимости цепи антенны имеем

$$Y = Y_0 + kP.$$

Пусть звуковое давление синусоидально

$$P = P_m \sin \Omega t.$$

Тогда для проводимости можем записать

$$Y = Y_0 + kP_m \sin \Omega t = Y_0 + Y_1 \sin \Omega t = Y_0(1 + m \sin \Omega t),$$

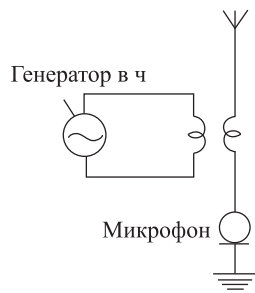


Рис. 258

где $m = Y_1/Y_0$ — коэффициент модуляции проводимости. Пусть, далее, электродвижущая сила, наводимая генератором в антенне, равна

$$e = E_m \sin \omega_0 t,$$

где ω_0 — высокая (несущая) частота. Ток в антенне равен произведению из эдс на проводимость

$$i = eY = E_m Y_0 (1 + m \sin \Omega t) \sin \omega_0 t.$$

Мы получили обычное выражение амплитудно-модулированного колебания для случая синусоидальной модуляции частотой Ω .

Рассмотренная схема есть *линейная схема с переменным параметром*. Переменным параметром является проводимость.

Возвращаясь к выражению для амплитудно-модулированного колебания, перепишем его в более общем виде

$$i = I_m [1 + mf(t)] \sin \omega_0 t, \quad (225)$$

где $f(t)$ — произвольный передаваемый сигнал (предполагается лишь, что $f(t) < 1$). Как видим, процесс модуляции состоит в *перемножении* двух функций времени: $1 + mf(t)$ и $I_m \sin \omega_0 t$. Следовательно, модулятор в принципе должен представлять собой перемножающее устройство, т.е. устройство с двумя входами и одним выходом, действие которого состоит в том, что при подаче на два входа двух функций $x(t)$ и $X_0(t)$ на выходе получается

$$y(t) = x(t) X_0(t).$$

Операция умножения на заданную функцию $X_0(t)$ линейна, как это видно из следующего соотношения:

$$y(t) = [x_1(t) + x_2(t)] X_0(t) = x_1(t) X_0(t) + x_2(t) X_0(t) = y_1(t) + y_2(t).$$

Устройства, назначение которых состоит в перемножении двух функций времени, применяемые для модуляции и преобразования частоты, называются обычно смесителями. Термин этот нехорош, так как не выражает существа дела. Будем кратко называть устройство для перемножения двух функций *перемножителем*.

Итак, для получения обычной АМ нужно взять перемножитель и подать на его входы $x(t) = 1 + mf(t)$ и $X_0(t) = I_m \sin \omega_0 t$. В результате перемножения получится модулированное колебание (225).

Заметим, что выражение (225) можно переписать в виде

$$i = I_m \sin \omega_0 t + m I_m f(t) \sin \omega_0 t.$$

Первый член есть колебание несущей частоты, разложение второго члена дает боковые полосы модуляционного спектра. Наряду с обычной АМ применяется также передача без несущей частоты, получаемая при помощи так называемой *балансной* модуляции. Для получения АМ колебания без несущей частоты при помощи перемножителя достаточ-

но устранить постоянную составляющую модулирующей функции, т. е. взять просто $x(t) = f(t)$. При этом на выходе перемножителя получим

$$y(t) = x(t) X_0(t) = I_m f(t) \sin \omega_0 t.$$

В спектре этого колебания несущая частота отсутствует; спектр состоит только из двух боковых полос.

Обратимся к вопросу о техническом осуществлении перемножителя. Перемножитель по определению должен представлять собой устройство с двумя входами и одним выходом. Напряжение на выходе должно быть пропорционально произведению напряжений на обоих входах (речь идет о мгновенных значениях напряжения).

Хорошего перемножителя в радиотехнике до сих пор нет. При помощи электрических схем, и притом очень простых, можно осуществлять не только сложение и вычитание напряжений, но и такие операции, как дифференцирование и интегрирование. Без всяких затруднений выполняются логические операции («и», «или», «не»). А вот такая, казалось бы, простая операция, как перемножение, вызывает большие затруднения. Они связаны главным образом с тем, что требуется получить достаточно точное перемножение для напряжений с широким спектром (т. е. для быстро изменяющихся напряжений) и при произвольных знаках перемножаемых напряжений.

Было предложено громадное количество различных перемножающих устройств. Нет возможности не только описать их здесь, но даже и просто перечислить. Ограничимся лишь несколькими примерами.

Пусть имеется четырехполюсник, у которого выходное напряжение u_2 пропорционально входному напряжению u_1 , т. е.

$$u_2 = k u_1.$$

Если бы сделать коэффициент k *управляемым*, т. е. изменяющимся пропорционально некоторому напряжению u_3 ,

$$k = a u_3,$$

то мы имели бы

$$u_2 = a u_1 u_3$$

и задача построения перемножителя была бы решена.

Эта идея осуществляется путем применения многосеточных ламп (гексодов или гептодов). В этих лампах, как известно, крутизна по третьей (так называемой «сигнальной») сетке зависит от напряжения на первой («гетеродинной» сетке). А так как ток в аноде пропорционален произведению крутизны на напряжение на управляющей сетке, то, подавая два перемножаемых напряжения на первую и третью сетки, можно получить на анодной нагрузке напряжение, пропорциональное произведению обоих входных напряжений. При этом предполагается, что крутизна линейно зависит от напряжения на первой сетке. К сожалению, ни одно из этих условий не выполняется в действительности,

а потому перемножение при помощи гептода имеет весьма приближенный характер. Тем не менее (за неимением лучшего) такие перемножители довольно широко применяются в качестве маломощных модуляторов и преобразователей частоты.

Упомянем еще об одном довольно сложном перемножающем устройстве, позволяющем, однако, получить весьма точное перемножение. Используется сложная импульсная модуляция: прямоугольные импульсы модулируются одним входным напряжением по амплитуде, а вторым —

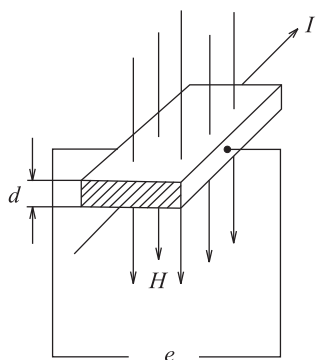


Рис. 259

по длительности. Таким образом, *площадь* импульса, равная произведению основания прямоугольника на его высоту, оказывается пропорциональной произведению обоих модулирующих напряжений. Остается получить напряжение, пропорциональное площади импульсов, для чего применяется интегрирующая схема с надлежащим образом выбранной постоянной времени.

К таким замысловатым решениям приходится прибегать потому, что неизвестно ни одного физического явления, в котором одна электрическая величина была бы точно пропорциональна произведению двух других, также электрических величин в силу общего закона, управляющего этим явлением. В явлениях, подчиняющихся закону индукции, закону Био-Савара, закону Кулона, мы имеем требуемую точную зависимость между тремя величинами, но одна из них — механическая величина (сила, скорость).

В результате длительных поисков внимание исследователей было привлечено за последнее время к явлению, называемому *эффектом Холла*. Вкратце, явление состоит в следующем. Если пропускать вдоль пластинки из полупроводника ток i и одновременно воздействовать на пластинку поперечным полем H , то на боковых гранях пластинки наблюдается разность потенциалов e , называемая *эдс Холла* (рис. 259). При этом

$$e = R \frac{1}{d} i H,$$

где R — постоянная Холла, зависящая от физических свойств материала; d — толщина пластинки.

Как видим, одна из величин, образующих произведение, имеет магнитную природу. Однако возможно построить электромагнитную систему, в которой поле H будет в широких пределах пропорционально напряжению, подведенному к обмотке. Вторая перемножаемая величина есть, к сожалению, ток, а не напряжение; это вызывает определенные затруднения. Однако в общем преимущества перемножающих устройств, основанных на эффекте Холла, весьма значительны.

Разработка таких устройств усиленно ведется в настоящее время, и можно ожидать их широкого распространения.

§ 75. Преобразование частоты

Рассмотренное выше преобразование спектра при АМ состоит в том, что низкочастотный спектр сигнала смещается по шкале частот в область радиочастот. Такое смещение может рассматриваться как частный случай более общей линейной операции, называемой в радиотехнике *преобразованием частоты*.

Под преобразованием частоты подразумевается вообще смещение спектра по шкале частот в ту или другую сторону, т. е. в область как более высоких, так и более низких частот. Предполагается, что при таком смещении или переносе спектра он не подвергается искажению.

Для преобразования частоты применяется общий прием, состоящий в том, что данное колебание (подлежащее преобразованию) умножается на синусоидальное колебание, доставляемое отдельным генератором — *гетеродином*. Этот прием называют часто *гетеродинированием*.

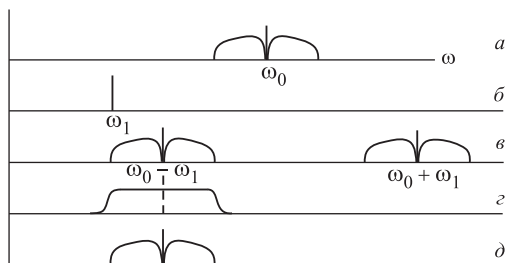


Рис. 260

Сущность преобразования состоит в том, что при перемножении тригонометрических функций разных аргументов получаются два члена, один из которых зависит от суммы, другой — от разности аргументов, например,

$$\cos x \cos y \equiv \frac{1}{2} [\cos (x + y) + \cos (x - y)].$$

Рассмотрим для примера преобразование, состоящее в смещении модуляционного спектра из области радиочастот в область промежуточных частот (порядка сотен килогерц). На рис. 260, а изображен исходный спектр, имеющий несущую частоту ω_0 и две сплошные боковые полосы. Рисунок 260, б изображает спектр колебания гетеродина — одну спектральную линию с частотой ω_1 . На рис. 260, в представлен спектр на выходе перемножителя; он состоит из двух модуляционных спектров, имеющих несущие частоты $\omega_0 - \omega_1$ и $\omega_0 + \omega_1$. На рис. 260, г

изображена частотная характеристика полосового фильтра промежуточной частоты (ФПЧ) и, наконец, на рис. 260, δ — спектр на выходе фильтра; на рис. 261 изображена скелетная схема устройства, выполняющего все описанные действия, в результате которых спектр исходного сигнала оказывается перемещенным с частоты ω_0 на частоту $\omega_0 - \omega_1$. Буквы на рис. 261 соответствуют рис. 260.

Заметим теперь, что если плавно изменять частоту гетеродина ω_1 , то преобразованный спектр будет плавно же перемещаться по шкале частот, так как средняя частота спектра

$$\omega_{np} = \omega_0 - \omega_1.$$

На возможности такого перемещения основан принцип действия *супергетеродинного* приемника. Принцип этот состоит в том, что на

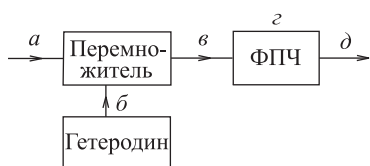


Рис. 261

в этой схеме это то, что фильтры настроены раз навсегда на одну определенную полосу. В супергетеродинном приемнике не фильтры (или контуры) настраиваются на частоту принимаемого сигнала, а, наоборот, *принимаемый сигнал настраивается на частоту УПЧ* путем, вышеописанного перемещения спектра сигнала по шкале частот. Это

на входе стоит перемножитель с гетеродином по схеме рис. 261; настройка осуществляется изменением частоты гетеродина. Усиление производится на промежуточной частоте, причем усилитель промежуточной частоты (УПЧ) имеет большое число ступеней, снабженных фильтрами промежуточной частоты. Самое главное

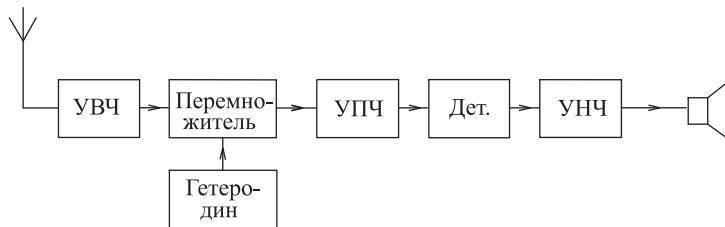


Рис. 262

дает неоценимые преимущества. Во-первых, настройка приемника производится одной ручкой (управляющей частотой гетеродина), независимо от числа ступеней усиления. Во-вторых, фильтры ступеней *УПЧ*, не подвергающиеся перестройке, могут быть сделаны по сложной схеме, обеспечивающей минимальные искажения при максимальной избирательности. Нужно, впрочем, заметить, что схема супергетеродинного приемника обычно усложняется введением *УВЧ* — усиления высокой частоты (до перемножителя), однако и в таком виде эта схема является

наилучшей схемой приемника. Она показана в скелетной форме на рис. 262. Кроме уже упомянутых элементов, на ней указаны детектор (Дет.) и усилитель низкой частоты (УНЧ).

Во многих радиотехнических устройствах как на передающем, так и на приемном концах системы радиосвязи прибегают к многократному преобразованию частоты.

В заключение заметим еще, что возможен специальный случай преобразования частоты, когда гетеродин настраивается в точности на частоту ω_0 . При этом модуляционный спектр смещается в область низких частот ($\omega_{np} = 0$). Это есть случай так называемого синхронного детектирования, о котором речь будет позднее (§ 101).

§ 76. Нелинейный способ получения АМ

Как показано в § 74, модуляция есть по существу линейная операция, выполняемая в линейной системе с переменным параметром. Операция модуляции есть не что иное, как операция *перемножения*.

Но большинство практических схем АМ работает по другому принципу. Дело в том, что построение правильно действующего линейного перемножителя затруднительно. Поэтому прибегают к косвенному способу перемножения, сущность которого состоит в том, что *сумму* двух перемножаемых напряжений заставляют действовать в нелинейной цепи. При этом в составе тока получается составляющая, пропорциональная произведению обоих напряжений. Прежде всего покажем качественно, что получение таким путем АМ колебания возможно. Возьмем два напряжения: напряжение несущей частоты $u_1 = U_m \sin \omega_0 t$ и модулирующее напряжение $u_2 = 1 + m \sin \Omega t$. Включим источники обоих напряжений последовательно в цепь с нелинейным элементом. Пусть характеристика этого элемента представлена ломаной линией. Форма тока в цепи может быть получена обычным построением, показанным на рис. 263. Нетрудно видеть, что в составе тока получается АМ колебание. Кроме него, получаются еще постоянная составляющая, составляющая низкой частоты и ряд высокочастотных составляющих. Все эти лишние составляющие должны быть отфильтрованы.

Разберем теперь подробнее механизм нелинейного перемножения. Пусть характеристика нелинейного элемента представлена степенным рядом

$$i = a_0 + a_1 u + a_2 u^2 + a_3 u^3 + a_4 u^4 + \dots$$

В нашем случае

$$u = u_1 + u_2,$$

так что

$$\begin{aligned} i &= a_0 + a_1(u_1 + u_2) + a_2(u_1 + u_2)^2 + a_3(u_1 + u_2)^3 + a_4(u_1 + u_2)^4 + \dots = \\ &= a_0 + a_1 u_1 + a_1 u_2 + a_2 u_1^2 + a_2 u_2^2 + \boxed{2a_2 u_1 u_2} + a_3 u_1^3 + a_3 u_2^3 + \\ &\quad + 3a_3 u_1^2 u_2 + 3a_3 u_1 u_2^2 + a_4 u_1^4 + \dots \end{aligned}$$

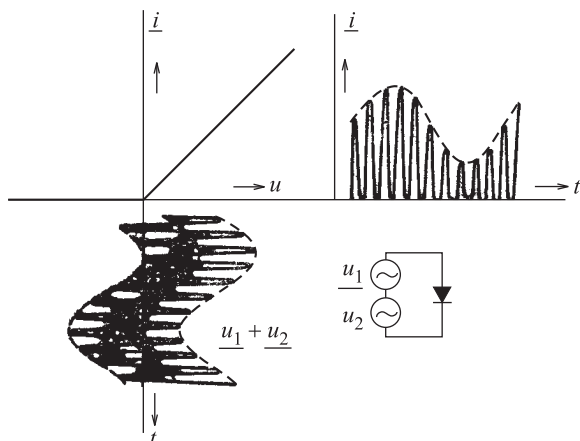


Рис. 263

Как видно, нужный член, содержащий произведение $u_1 u_2$, получается в составе тока *только* за счет квадратичного члена вольт-амперной характеристики нелинейного элемента. Все остальные члены дают лишние составляющие, так что мы имеем нечистое перемножение, и получаемый ток требует последующей очистки, хотя бы путем фильтрации.

Однако возможно путем последовательного усложнения схемы добиться исключения ряда лишних членов и получить, таким образом, более совершенное перемножение.

Обратимся к рис. 264. Исходная схема с двумя напряжениями u_1 и u_2 , действующими в нелинейной цепи, показана на рис. 264, а. Совместим две такие схемы, как показано на рис. 264, б, так, чтобы они имели общую ветвь, в которую включено напряжение u_1 . Напряжение, снимаемое со схемы, равно

$$u = R(i_1 - i_2).$$

Заметим, что напряжение u_2 подано в схему дважды и при этом так, что в верхней половине действует суммарное напряжение $u_1 + u_2$ (мгновенные направления напряжений обозначены стрелками), а в нижней половине — разностное $u_1 - u_2$. Для токов получаем:

$$\begin{aligned} i_1 &= a_0 + a_1(u_1 + u_2) + a_2(u_1 + u_2)^2 + a_3(u_1 + u_2)^3 + a_4(u_1 + u_2)^4 + \dots, \\ i_2 &= a_0 + a_1(u_1 - u_2) + a_2(u_1 - u_2)^2 + a_3(u_1 - u_2)^3 + a_4(u_1 - u_2)^4 + \dots \end{aligned}$$

Разность токов

$$\begin{aligned} i_1 - i_2 &= 2a_1u_2 + 4a_2u_1u_2 + 6a_3u_1^2u_2 + 2a_3u_2^3 + \\ &\quad + 8a_4u_1^3u_2 + 8a_4u_1u_2^3 + \dots \end{aligned} \quad (226)$$

(Предполагается, что схема симметрична, т. е. что коэффициенты a_k в разложениях токов i_1 и i_2 одинаковы.) Как видим, в результате вычитания целый ряд членов выпал. Сохранились только члены, содержащие нечетные степени u_2 . Если бы характеристика нелинейного элемента была строго квадратичной, т. е.

$$i = ku^2,$$

то мы получили бы чистое перемножение по формуле

$$k(u_1 + u_2)^2 - k(u_1 - u_2)^2 = 4ku_1u_2.$$

Схемы типа рис. 264, б называются *балансными*. Они применяются, в частности, для получения модулированного колебания, не содержащего несущей частоты. Действительно, если u_1 есть напряжение несущей частоты, то из (226), как видим, оно выпадает. Такого рода модуляция называется *балансной модуляцией*.

Идея усовершенствования схемы рис. 264, а, приводящая к балансной схеме рис. 264, б, сводится к тому, что, изменяя знак одного из напряжений и производя вычитание токов, мы избавляемся от целого ряда излишних членов. По этому пути можно идти и далее. Возьмем теперь две одинаковые балансные схемы, изменим в одной из них знаки *обоих* напряжений u_1 и u_2 на обратные и образуем напряжение, равное сумме напряжений, получаемых на каждой из схем в отдельности. Такой результат достигается в схеме рис. 265, а, представляющей собой совмещение двух балансных схем. Снимаемое с этой схемы напряжение можно представить в виде

$$u = u' + u'',$$

где u' и u'' — напряжения, развиваемые каждой из совмещенных балансных схем в отдельности. Для этих двух напряжений мы можем записать (см. (226)):

$$u' = R(i'_1 - i'_2) = 2R(a_1u_2 + 2a_2u_1u_2 + 3a_3u_1^2u_2 + \\ + a_3u_2^3 + 4a_4u_1^3u_2 + 4a_4u_1u_2^3 + \dots),$$

$$u'' = R(i''_1 - i''_2) = 2R(-a_1u_2 + 2a_2u_1u_2 - 3a_3u_1^2u_2 - \\ - a_3u_2^3 + 4a_4u_1^3u_2 + 4a_4u_1u_2^3 + \dots),$$

Выражение для u'' отличается от выражения для u' тем, что знаки *обоих* напряжений u_1 и u_2 заменены на обратные. В результате сложения получаем

$$u = u' + u'' = 4R(a_2u_1u_2 + 2a_4u_1^3u_2 + 2a_4u_1u_2^3 + \dots). \quad (227)$$

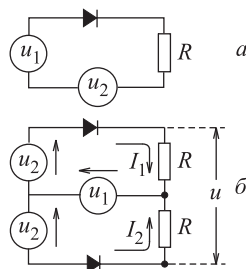


Рис. 264

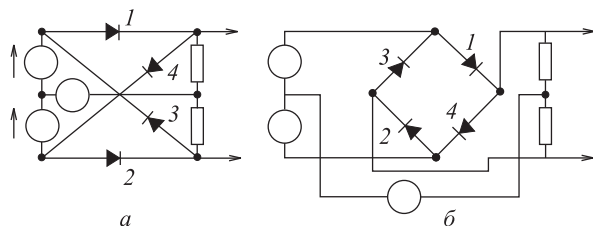


Рис. 265

Это выражение показывает, что достигнуто дальнейшее устранение лишних членов. В частности, если характеристика нелинейного элемента выражается многочленом степени не выше третьей, т. е.

$$i = a_0 + a_1 u + a_2 u^2 + a_3 u^3,$$

то схема рис. 265, а (при условии полной ее симметрии) дает идеальное перемножение. Можно перерисовать эту схему в виде рис. 265, б, как ее чаще всего изображают. Получается как бы замкнутое кольцо из четырех нелинейных элементов. Поэтому схему рис. 265 называют *кольцевым модулятором*. Она имеет широкое применение в качестве модулятора и преобразователя частоты.

Следует заметить, что дальнейшее улучшение схемы тем же способом, последовательное применение которого привело нас сперва к балансной, а затем к кольцевой схеме, невозможно. В самом деле, формула (227) показывает, что в составе напряжения, снимаемого с кольцевой схемы, сохраняются только члены, содержащие произведения нечетных степеней u_1 и u_2 . Следовательно, при изменении знака одного или обоих напряжений *все* члены выражения (227) сохраняют одинаковый знак и ни один из них не может быть устранен в отдельности сложением или вычитанием. Поэтому построение более еложной схемы путем совмещения двух кольцевых схем ничего не дает.

§ 77. Практические схемы модуляции

Амплитудная модуляция. Для обычной амплитудной модуляции применяются главным образом две схемы: *сеточной модуляции* и *анодной модуляции*. Схема сеточной модуляции показана на рис. 266. Напряжение несущей частоты u_1 и модулирующее напряжение u_2 подаются на сетку триода. Блокировочная емкость C_1 обеспечивает прохождение высокой частоты в обход обмотки низкочастотного трансформатора. Модуляция получается за счет нелинейности характеристики триода $i_a = f(u_c)$. В составе анодного тока получаются лишние составляющие. Их отфильтровывает контур LC , настроенный на несущую частоту. Ширина полосы пропускания контура берется немного больше ширины модуляционного спектра, т. е. двойной ширины спектра сигнала. При таком условии контур отфильтровывает как постоянную

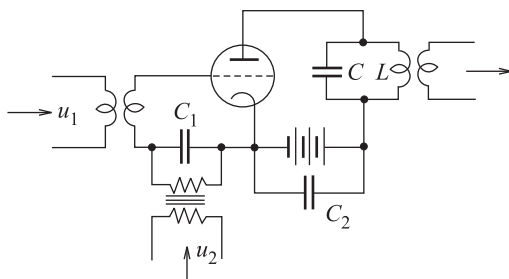


Рис. 266

составляющую и низкую частоту, так и высокочастотные составляющие, засоряющие модуляционный спектр.

Схема анодной модуляции показана на рис. 267. Ее отличие от схемы сеточной модуляции состоит в том, что напряжение несущей частоты u_1 и модулирующее напряжение u_2 действуют в разных цепях лампы Λ_1 ; напряжение u_1 приложено непосредственно к сетке, а напряжение u_2 подается в анодную цепь. Так как действие напряжения в анодной цепи, грубо говоря, в μ раз слабее действия такого же напряжения в цепи сетки, то модулирующее напряжение предварительно усиливается вспомогательной лампой Λ_2 . С принципиальной точки зрения действие схем рис. 266 и 267 одинаково, но схема анодной модуляции рис. 267 имеет чисто технические преимущества, рассматриваемые в курсе передатчиков.

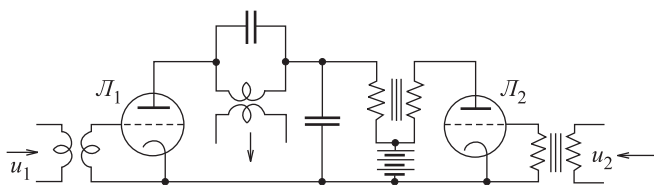


Рис. 267

Частотная модуляция. Для того чтобы получить частотную модуляцию, необходимо изменять частоту задающего генератора, что проще всего сделать, изменяя настройку контура задающего генератора. В простейшей схеме радиотелефонной модуляции это достигается включением конденсатора микрофона непосредственно в контур задающего генератора. В современной технике используется главным образом схема, действие которой основано на применении так называемой *реактивной лампы*. Реактивная лампа представляет собой устройство, входное сопротивление которого практически чисто реактивно, и изменяет свою величину в соответствии с изменениями подводимого модулирующего напряжения.

Реактивная лампа является частным случаем двухполюсника с обратной связью (см. § 40). Составим схему рис. 268. Ко входу подключен делитель $Z_1 Z_2$. Напряжение обратной связи снимается с Z_1 и подается на сетку. Мы имеем:

$$\dot{U}_1 = \frac{Z_1}{Z_1 + Z_2} \dot{U} = \beta \dot{U}, \quad \dot{U}_2 = -\mu \dot{U}_1, \quad \dot{I} = \frac{1}{R_i} (\dot{U} + \dot{U}_2),$$

где μ — коэффициент усиления триода; R_i — его внутреннее сопротивление. При условии

$$Z_1 + Z_2 \gg R_i$$

получаем для входной проводимости схемы

$$Y = \frac{\dot{I}}{\dot{U}} = \frac{1}{R_i} (1 + \mu\beta) = \frac{1}{R_i} + S\beta = \frac{1}{R_i} + S \frac{Z_1}{Z_1 + Z_2}.$$

Если выбрать $Z_2 \gg Z_1$, то приближенно

$$Y = \frac{1}{R_i} + S \frac{Z_1}{Z_2}.$$

На схеме рис. 268 делитель составлен из емкости и сопротивления, так что

$$Z_1 = R, \quad Z_2 = \frac{1}{i\omega C}, \quad \frac{Z_1}{Z_2} = i\omega RC$$

и в этом случае

$$Y = \frac{1}{R_i} + i\omega RCS = \frac{1}{R_i} + i\omega C_{\text{ЭКВ}},$$

где

$$C_{\text{ЭКВ}} = RCS,$$

означает эквивалентную емкость. Таким образом, двухполюсник рис. 268 представляется схемой замещения рис. 269.

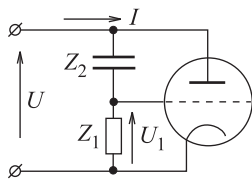


Рис. 268

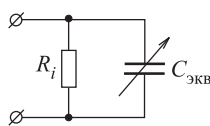


Рис. 269

Итак, схема рис. 268 имеет емкостную составляющую проводимости. При этом существенно, что эквивалентная емкость $C_{\text{ЭКВ}}$ является *управляемым* параметром. В самом деле: $C_{\text{ЭКВ}}$ пропорциональна крутизне S , а крутизна зависит от напряжения на сетке. Предположим, что выбран квадратичный участок характеристики триода, т. е.

$$i_a = a_0 + a_1 u_1 + a_2 u_1^2.$$

Тогда

$$S = \frac{di_a}{du_1} = a_1 + 2a_2u_1,$$

т. е. крутизна линейным образом зависит от напряжения на сетке.

Для получения частотной модуляции достаточно подключить схему рис. 268 к задающему контуру генератора и подать модулирующее напряжение на сетку реактивной лампы.

Укажем на возможность улучшения свойств схемы рис. 268. Дело в том, что проводимость этой схемы имеет, кроме емкостной, еще и активную составляющую $1/R_i$, что крайне нежелательно, так как эта составляющая вносит в контур дополнительные потери. От активной составляющей можно легко избавиться, так как обратная связь позволяет в принципе получить *любой* характер входной проводимости при надлежащем выборе сопротивлений Z_1 и Z_2 . Иначе говоря, можно

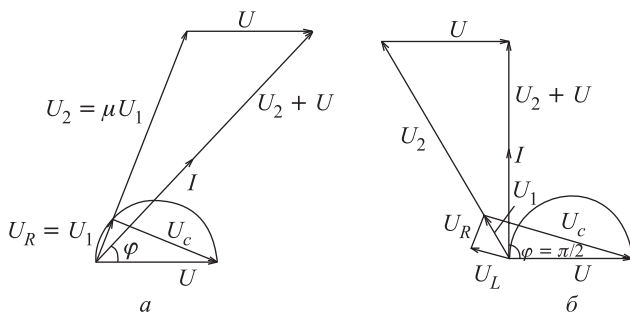


Рис. 270

получить *любой* фазный сдвиг между током и напряжением (от 0 до 2π). У нас получился сдвиг меньше $\pi/2$, требуется фазный сдвиг ровно $\pi/2$ (чтобы входное сопротивление было чисто реактивным). Для этого достаточно добавить в Z_1 индуктивность, т. е. выбрать

$$Z_1 = R + i\omega L.$$

Тогда

$$Y = \frac{1}{R_i} + S \frac{Z_1}{Z_2} = \frac{1}{R_i} + i\omega RCS - \omega^2 LCS,$$

и, выбрав L так, чтобы удовлетворить равенству

$$\omega^2 LC = \frac{1}{R_i S} = \frac{1}{\mu},$$

получаем полную компенсацию активной проводимости (при данной частоте ω). Механизм компенсации поясняется векторными диаграммами рис. 270, а и б; на первом $Z_1 = R$, на втором $Z_1 = R + i\omega L$, Z_2 в обоих случаях равно $\frac{1}{i\omega C}$.

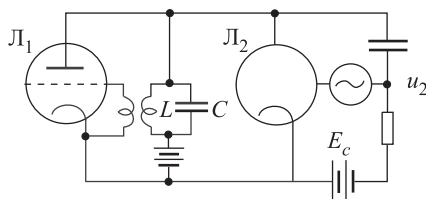


Рис. 271

Принципиальная схема модулятора ЧМ с реактивной лампой показана на рис. 271. На этой схеме L_1 — генераторная лампа, LC — задающий контур, L_2 — подключенная параллельно контуру реактивная лампа, u_2 — модулирующее напряжение, подаваемое на сетку реактивной лампы, E_c — напряжение смещения для установки рабочей точки.

Импульсная модуляция. Мы не будем останавливаться на АИМ, так как, во-первых, для этого вида модуляции пригодны те же способы, что и для обычной АМ, а во-вторых, АИМ мало употребительна в настоящее время вследствие своей малой помехоустойчивости. Наибольшее применение имеют ДИМ и ФИМ, механизм которых и рассмотрим. Простейший способ получения последовательности импульсов, модулированных по фазе или по длительности, основан на наложении

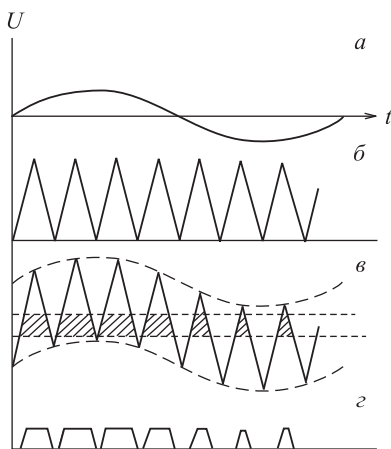


Рис. 272

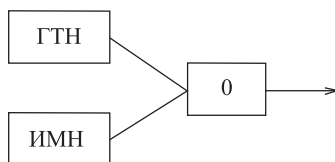


Рис. 273

модулирующего напряжения на периодическое напряжение треугольной формы с основной частотой, равной частоте следования немодулированной последовательности импульсов. Полученная сумма напряжений подвергается дальнейшей обработке: ограничению, дифференцированию и т. п. Начнем с модуляции по длительности. На рис. 272, а изо-

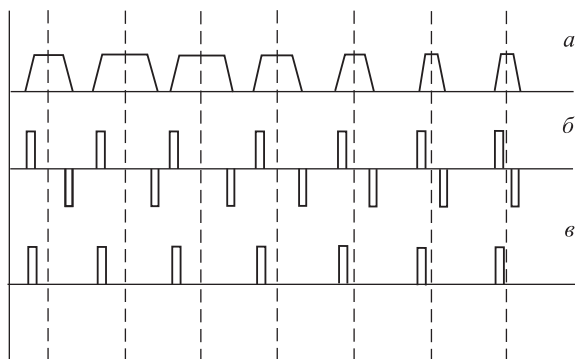


Рис. 274

бражено модулирующее напряжение, на рис. 272, б — треугольное напряжение, на рис. 272, в — результат сложения обоих напряжений. Если теперь ограничить результирующее напряжение сверху и снизу, как показано пунктирными линиями на рис. 272, в, то получатся трапецеидальные импульсы, модулированные по длительности. Крутизной скатов легко управлять, изменяя соотношение между размахом треугольного напряжения и полосой ограничения. Таким образом, скелетная схема ДИМ состоит из генератора треугольного напряжения (ГТН на рис. 273), источника модулирующего напряжения (ИМН) и ограничителя (О). Схемы названных элементов никаких принципиальных особенностей не имеют, и мы ограничиваемся скелетной схемой. Для получения ФИМ можно использовать ДИМ как промежуточный продукт. Возьмем последовательность импульсов, модулированных по длительности, в том виде, в каком эта последовательность получена вышеописанным способом, и продифференцируем ее (рис. 274, б). Получим прямоугольные импульсы противоположной полярности; положительные импульсы соответствуют переднему скату дифференцируемых трапецеидальных импульсов, отрицательные — заднему скату. Если теперь пропустить полученную последовательность через одноконтный выпрямитель, который отсекает отрицательные импульсы, то получится (рис. 274, в) последовательность прямоугольных¹⁾ импульсов, модулированных по фазе, т. е. смещенных относительно тактовых точек, отме-

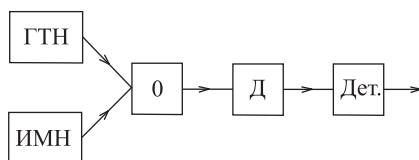


Рис. 275

ля (О). Схемы названных элементов никаких принципиальных особенностей не имеют, и мы ограничиваемся скелетной схемой. Для получения ФИМ можно использовать ДИМ как промежуточный продукт. Возьмем последовательность импульсов, модулированных по длительности, в том виде, в каком эта последовательность получена вышеописанным способом, и продифференцируем ее (рис. 274, б). Получим прямоугольные импульсы противоположной полярности; положительные импульсы соответствуют переднему скату дифференцируемых трапецеидальных импульсов, отрицательные — заднему скату. Если теперь пропустить полученную последовательность через одноконтный выпрямитель, который отсекает отрицательные импульсы, то получится (рис. 274, в) последовательность прямоугольных¹⁾ импульсов, модулированных по фазе, т. е. смещенных относительно тактовых точек, отме-

¹⁾ Практически импульсы не могут быть прямоугольными. Их форма также приближается к трапецеидальной, т. е. скаты импульсов имеют конечную крутизну, определяемую шириной полосы пропускания всей схемы модулятора.

ченных на рис. 274 вертикальными линиями. Таким образом, скелетная схема ФИМ отличается от схемы ДИМ тем, что к ней добавляется (рис. 275) дифференцирующее звено (Д) и детектор. Варианты и подробности схем импульсной модуляции рассматриваются в соответствующих специальных курсах.

§ 78. Специальные электронные приборы

В современной радиотехнике можно отметить два различных пути к решению одной и той же технической задачи. Один путь — *схемный*, когда для получения желаемого результата составляется схема из обычных элементов (сопротивлений, конденсаторов, диодов, триодов и т. п. деталей). Второй путь — *аппаратурный*, когда для той же цели конструируется специальный электронный прибор. Предпочесть то или иное решение можно только с учетом всех технико-экономических и эксплуатационных соображений. Ограничимся описанием двух типов специальных электронно-лучевых трубок, применяемых для импульсной модуляции.

Одна из этих трубок — *циклофон* — предназначена для получения импульсно-фазовой модуляции. Прибор состоит из электронной пушки 1 (рис. 276), дающей узкий электронный луч. Отклоняющая система 2 состоит из двух пар взаимно-перпендикулярных пластин. Обе пары питаются синусоидальными напряжениями, сдвинутыми относительно друг друга по фазе на 90° . На рис. 276 показан генератор, снабженный фазовращателем для обеспечения требуемого сдвига фазы. При описанных условиях внутри отклоняющей системы образуется вращающееся электрическое поле, под действием которого луч также вращается, совершая один оборот за период отклоняющего напряжения. Вращающийся луч попадает на маску 3. Путь луча на маске отмечен пунктирной окружностью. В маске сделаны косые щели. Когда луч пробегает по щели, он проходит через маску и ток луча замыкается через коллектор 4. При этом в цепи коллектора получается импульс. Частота следования импульсов равна fz , где f — частота отклоняющего генератора, z — число щелей в маске.

Пусть теперь к трубке подведено модулирующее напряжение, действующее таким образом, что при его изменении пропорционально изменяется диаметр окружности, описываемой лучом на маске. Щели прорезаны косо (под углом к радиусам); поэтому при изменении диаметра кругового пути луча импульсы будут появляться позже или раньше, т. е. окажутся модулированными по фазе. Это пояснено графически на том же рис. 276. Круговые пути луча спрямлены; 1 — путь луча при отсутствии модулирующего напряжения, импульсы находятся при этом в тактовых точках. При смещении траектории луча в положение 2 получается запаздывание импульсов, в положение 3 — опережение.

Управление диаметром траектории луча можно осуществить по-разному: модулировать генератор отклоняющего напряжения или подавать

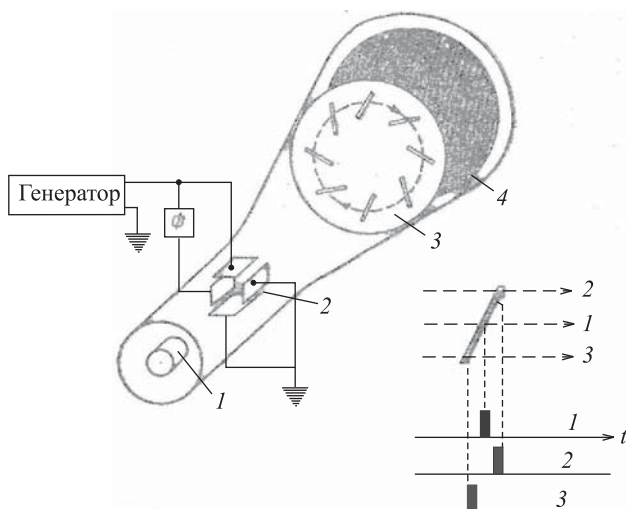


Рис. 276

модулирующее напряжение на ускоряющий электрод. Если же трубка используется для многоканальной системы с временным разделением, то самое удобное — подать модулирующие напряжения каждого канала на специальные отклоняющие пластины, расположенные непосредственно у каждой из щелей.

В описанном приборе можно получить и модуляцию импульсов по длительности. Для этого достаточно только заменить в маске косые щели отверстиями треугольной формы.

Как видим, в циклофоне, одном сравнительно простом приборе, объединены импульсный генератор и модулятор, или даже несколько модуляторов отдельных каналов.

Другой, заслуживающий внимания прибор — это *кодирующая трубка* для получения импульсно-кодовой модуляции. Прибор состоит из электронной пушки 1 (рис. 277) и двух пар отклоняющих пластин 2 и 3, расположенных как в обычной осциллографической трубке. На горизонтально-отклоняющие пластины 2 подается пилообразное напряжение развертки. Напряжение сигнала подается на вертикально-отклоняющие пластины 3. Луч попадает на маску 4, в которой пробиты отверстия. Расположение отверстий вдоль горизонтальных строк соответствует двоичной записи чисел. На рисунке изображены отверстия, соответствующие числам (в трехзначной двоичной записи) 001, 010, 011, 100, 101, 110, 111. Напряжение развертки заставляет луч пробежать по строке. Проходя через отверстие, луч замыкается на коллектор 5, и во внешней цепи коллектора появляется импульс. Отклонение по вертикали, производимое управляющим напряжением сигнала, переводит луч со строки на строку, так что каждому мгновенному значению

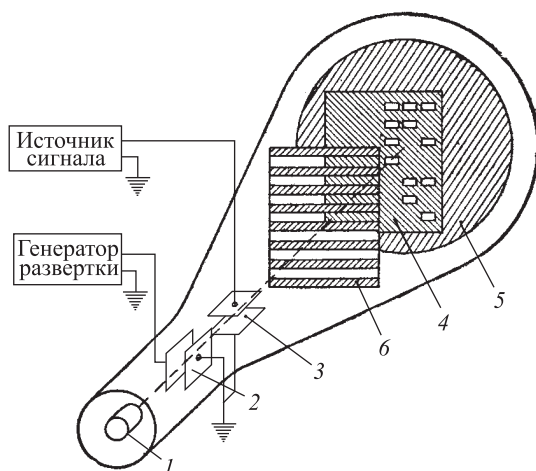


Рис. 277

напряжения сигнала соответствует своя строка маски, а следовательно, и определенное кодовое обозначение. В этом собственно и состоит действие кодирующей трубки.

Мы не касались пока одной важной детали, а именно квантования. Если подводимое к вертикально-отклоняющим пластинам напряжение сигнала заранее квантовано, и притом так, что каждому разрешенному уровню соответствует своя строка на маске, то описание не требует дополнений. Возможно, однако, возложить операцию квантования на кодирующую трубку.

С этой целью ее устройство дополняется решеткой 6, горизонтальные стержни которой перекрывают промежутки между строками. Если непрерывно изменяющееся напряжение сигнала наводит луч на промежуток между строками, то на пути оказывается стержень решетки. Используя эффект вторичной эмиссии от квантующей решетки, можно получить на вспомогательном электроде (не показанном на рисунке) потенциал, используемый для дополнительного перемещения луча по вертикали. Луч смещается до тех пор, пока не сойдет со стержня решетки и не попадет на строку ¹⁾.

Описанное устройство представляет собой автоматический регулятор, выполняющий операцию квантования. Таким образом, в кодирующей трубке происходят все требуемые для получения ИКМ операции: сигнал квантуется и преобразуется в двоичные кодовые комбинации.

¹⁾ Нужно заметить, что существуют специальные коды (например, код Грея), позволяющие расположить строки маски вплотную друг к другу (без промежутков). В этом случае квантование осуществляется само собой без каких-либо дополнительных приспособлений.

Извне к прибору подводится только (не считая сигнала) пилообразное развертывающее напряжение тактовой частоты.

Приведенные описания могут уже дать некоторое представление об остроумных и компактных конструкциях специальных электронных приборов. Дальнейшие разработки таких приборов ведутся в настоящее время. Из новейших приборов можно упомянуть о так называемых характронах и тайпатронах, дающих на экране локатора вместе с изображением цели ее условный шифр в виде букв и цифр.

ГЕНЕРИРОВАНИЕ КОЛЕБАНИЙ

§ 79. Автоколебания

Всякое радиотехническое устройство представляет собой цепь, составленную из линейных и нелинейных звеньев. Вдоль этой цепи проходят электрические колебания, передаваемые от звена к звену в усиленном или так или иначе преобразованном виде в зависимости от назначения данного звена. Существенно, что в большинстве звеньев колебания возникают лишь под воздействием колебаний в предшествующем звене. Так, например, на выходе данной ступени усилителя колебания возникают лишь при подаче на вход данной ступени изменяющегося напряжения с выхода предшествующей ступени. При таких обстоятельствах мы называем колебания *вынужденными*.

Но если следовать от конца к началу сложной радиотехнической цепи, то всегда дойдем до звена, в котором колебания не являются вынужденными, в котором они впервые зарождаются. Такого рода самостоятельно возникающие колебания носят название *автоколебаний*. Звено, в котором возникают автоколебания, называется по физической терминологии *автоколебательной системой*. С точки зрения техники это есть *генератор* колебаний. Автоколебательная система подчинена, разумеется, закону сохранения энергии; всякий электрический генератор можно рассматривать как преобразователь энергии источника постоянного тока в энергию электрических колебаний.

Рассмотрим типичный механизм автоколебаний; основные черты этого механизма мы найдем во многих конкретных схемах генераторов. Начнем со свободных колебаний в простейшем контуре. Если бы контур был без потерь, то в нем происходили бы незатухающие колебания. Но во всяком реальном контуре имеются потери; кроме того, мы рассматриваем генератор как источник колебательной энергии, и, следовательно, некоторое количество энергии отбирается во внешние цепи, в нагрузку генератора. Поэтому простой контур не может слу-

жить генератором; запас энергии, который был ему первоначально сообщен, быстро исчерпается, а колебания затухнут. Следовательно, для получения незатухающих колебаний необходимо все время пополнять запас энергии в контуре по мере ее расходования.

Итак, в состав генератора, кроме колебательного контура, должен входить *источник энергии*. Этим источником является батарея или выпрямитель — вообще источник постоянного напряжения. Для того чтобы поддерживать колебания в контуре, необходимо, чтобы энергия подавалась в контур периодически отдельными порциями в такт с уже происходящими колебаниями; поток энергии должен быть пульсирующим. Орган, управляющий поступлением энергии в контур, назовем *клапаном*. Для того чтобы работа клапана была согласована с колебаниями в контуре, нужно, чтобы она, в свою очередь, управлялась самим контуром. Механизм управления клапаном со стороны контура называется *обратной связью*. Таким образом, получается замкнутая цепь взаимодействия, которая и необходима для получения автоколебаний. Скелетная схема автоколебательной системы показана на рис. 278.

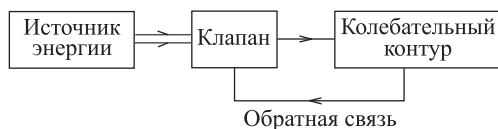


Рис. 278

Покажем практическое осуществление этой схемы на примере простейшего лампового генератора, схема которого представлена на рис. 279. Потери энергии в контуре LC , обусловленные наличием сопротивления R , восполняются за счет импульсов анодного тока. Источником энергии служит анодная батарея. Роль клапана играет сетка триода, напряжение на сетке управляет анодным током. В свою очередь, напряжение на сетке определяется колебаниями в контуре. Цепь взаимодействия замыкается через катушку L_1 в анодной цепи триода, индуктивно связанную с катушкой контура L ; коэффициент взаимной индукции обозначен через M . Таким образом, все перечисленные выше элементы автоколебательной системы в схеме рис. 279 налицо. Нужно заметить, что хотя эта схема *может* генерировать колебания, но не обязательно *будет* их генерировать, если не выполнить некоторые необходимые условия. В дальнейшем нам предстоит сделать подробный анализ работы генератора. Нужно будет получить ответ на ряд основных вопросов, как, напри-

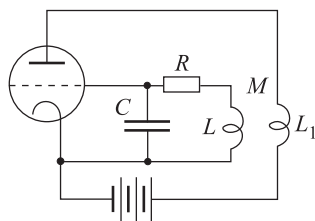


Рис. 279

мер, 1) при каких условиях возникнут автоколебания? 2) какова будет амплитуда установившихся автоколебаний?

На второй вопрос нельзя ответить, не разобравшись в той роли, которую играет нелинейность автоколебательной системы.

Предварительные соображения по этому поводу даны в следующем параграфе.

§ 80. Энергетика автоколебаний; динамическая устойчивость

Мощность, *теряемая* в сопротивлении R (рис. 279), равна

$$P_- = \frac{d^2}{2} \frac{U_m^2}{R},$$

т. е. пропорциональна квадрату амплитуды напряжения на контуре. Обозначим мощность, *поступающую* в контур за счет источника энергии, через P_+ . Если бы вся схема генератора, включая работу клапана и действие обратной связи, была линейна, то и мощность P_+ росла бы с квадратом амплитуды. Изобразим графически зависимости теряемой и получаемой контуром мощностей от амплитуды, выбрав для

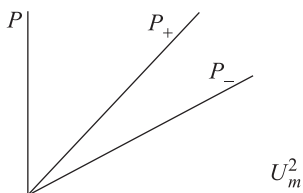


Рис. 280

амплитуд квадратичный масштаб. В таком масштабе оба графика будут представлены для линейной системы прямыми линиями, проходящими через начало координат (рис. 280). Если, как показано на рисунке, прямая P_+ лежит выше прямой P_- , то это означает, что контур получает больше энергии, чем теряет, т. е. что запас энергии в контуре возрастает, а это, в свою очередь, означает рост амплитуды колебания.

Если же, наоборот, прямая P_+ лежала бы ниже P_- , то это означало бы, что расход энергии превосходит приход, и при таком условии в контуре были бы возможны только затухающие колебания.

В условиях рис. 280 амплитуда росла бы неограниченно. Практически этого не может быть: всякая реальная схема нелинейна. В рассматриваемом случае нелинейность может быть обусловлена, например, насыщением триода или влиянием сеточного тока. Благодаря влиянию нелинейности получаемая мощность P_+ растет медленнее, чем с квадратом амплитуды, график искривляется и принимает вид рис. 281. При некотором значении амплитуды графики P_+ и P_- пересекаются. Точка пересечения отвечает *балансу энергии*, т. е. точному равенству теряемой и получаемой энергии. Очевидно, что только при условии баланса энергии запас энергии в контуре остается неизменным, а следовательно, амплитуда колебаний — постоянной. Таким образом, абсцисса точки пересечения на рис. 281 непосредственно определяет установившуюся амплитуду колебания U_0 .

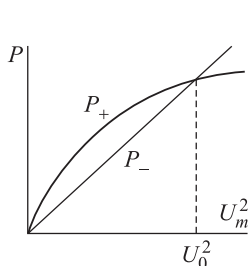


Рис. 281

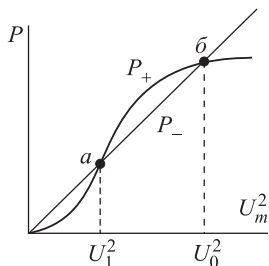


Рис. 282

Итак, наличие нелинейности есть обязательное свойство автоколебательной системы, способной генерировать колебания с неизменной (установившейся) амплитудой.

Баланс энергии может достигаться и не при одном значении амплитуды. Рассмотрим рис. 282. Здесь имеются две точки пересечения; которая из них определяет установившуюся амплитуду колебания? Введем понятие об *устойчивости* колебательного движения, которую будем называть *динамической* устойчивостью (в отличие от статической устойчивости, которая характеризует не движение, а состояние равновесия). Будем называть колебательное движение устойчивым, если имеется тенденция к сохранению установившейся амплитуды колебания. Пусть некоторое внешнее воздействие изменило амплитуду колебания, а затем это воздействие было прекращено, и система была предоставлена самой себе. Если теперь в системе восстановится первоначальное значение амплитуды, то можно сказать, что система динамически устойчива. Если же, напротив, амплитуда будет и далее изменяться в том же направлении уже без внешнего воздействия, то система динамически неустойчива. Легко видеть, что понятия динамической и статической устойчивости во многом сходны. Из приведенного определения сразу вытекает и способ испытания на устойчивость: достаточно вывести систему из данного состояния и посмотреть, как она будет себя вести в дальнейшем.

Применим этот способ к исследованию системы, соотношения в которой характеризуются графиками рис. 282. Вопрос сводится теперь к тому, устойчивы или неустойчивы режимы, определяемые точками пересечения *a* и *б*. Начнем с точки *a* и предположим, что по какой-либо причине амплитуда стала меньше U_1 . В этой области прямая P_- лежит выше кривой P_+ , т.е. система теряет энергии больше, чем получает. Поэтому амплитуда будет убывать и далее, пока колебания совсем не затухнут. Если же взять амплитуду больше U_1 , то $P_+ > P_-$ и амплитуда будет возрастать. Отсюда следует, что точка *a*, хотя и отвечает условию баланса энергии, но не является устойчивой: при малейшем изменении амплитуды появится тенденция к дальнейшему ее изменению в ту же сторону. Если амплитуда станет больше U_1 , то она будет затем

нарастать и далее, пока не достигнет значения U_0 . Дальнейший рост амплитуды на этом прекратится, так как при $U_m > U_0$ мы имеем снова $P_- > P_+$. Таким образом, точка b является устойчивой точкой, и ее абсцисса определяет установившуюся амплитуду U_0 .

Итак, различие между системами, свойства которых отображены графиками рис. 281 и 282, состоит в том, что первая самостоятельно раскачивается до установившейся амплитуды U_0 , вторая же нуждается в первоначальном толчке, который создал бы амплитуду, превосходящую U_1 , после чего дальнейшее нарастание амплитуды до U_0 произойдет уже автоматически. Первая система называется системой с *мягким самовозбуждением* или просто мягкой системой; вторая — системой с *жестким самовозбуждением* или жесткой системой. Следует еще заметить, что различие между мягкой и жесткой системой сводится к тому, устойчива или неустойчива система в состоянии покоя (т. е. когда амплитуда равна нулю). Рассуждая, как выше, относительно начала координат (которое, ведь, также является точкой баланса), легко установить, что жесткая система имеет устойчивое состояние покоя, а мягкая — неустойчивое.

§ 81. Условие самовозбуждения генератора

Если не задаваться вопросом о величине установившейся амплитуды, а интересоваться лишь тем, возбуждаться в данной системе автоколебания или нет, то такая постановка вопроса приведет к линейной задаче.

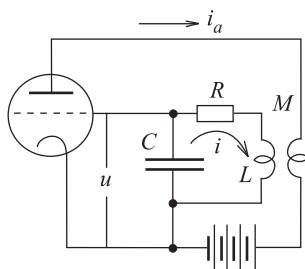


Рис. 283

Действительно, мы интересуемся при этом лишь начальной стадией явления, когда колебания в генераторе только что возникли, т. е. когда амплитуды еще очень малы, тогда как нелинейность проявляется лишь при значительных амплитудах. Поэтому математическая формулировка задачи о самовозбуждении приведет к *линейному* дифференциальному уравнению. Составим такое уравнение для генератора по схеме рис. 283. Если обозначить i — ток в контуре, i_a — переменная составляющая анодного тока, $u = u_c$ — напряжение на контуре, равное напряжению на сетке, то будем иметь уравнение напряжений в контуре

$$L \frac{di}{dt} + Ri + \frac{1}{C} \int i dt = M \frac{di_a}{dt}. \quad (228)$$

Здесь в правой части стоит эдс, наводимая в контуре катушкой связи. С другой стороны,

$$i_a = S_0 u, \quad u = \frac{1}{C} \int i dt,$$

где S_0 — крутизна триода в рабочей точке. Учитывая эти соотношения, получаем из (228)

$$LC \frac{d^2 u}{dt^2} + RC \frac{du}{dt} + u = MS_0 \frac{du}{dt}$$

или

$$\frac{d^2 u}{dt^2} + (2\alpha - \omega_0^2 MS_0) \frac{du}{dt} + \omega_0^2 u = 0, \quad (229)$$

где, как обычно,

$$\alpha = \frac{R}{2L}, \quad \omega_0 = \frac{1}{\sqrt{LC}}$$

соответственно показатель затухания и резонансная частота. Как видно, при наличии обратной связи показатель затухания изменяется: он убывает, так как величина $\omega_0^2 MS_0$ положительна. Показатель α характеризует скорость затухания колебаний, т. е. скорость расходования запаса энергии в контуре. Увеличивая обратную связь, можно достичь равенства

$$2\alpha - \omega_0^2 MS_0 = 0.$$

При этом результирующий (т. е. с учетом действия обратной связи) показатель затухания обратится в нуль. Это означает, что колебания уже не будут затухать. Отрицательный же показатель затухания означает, что колебания будут *нарастать* по экспоненциальному закону. Итак, условие

$$2\alpha - \omega_0^2 MS_0 < 0$$

есть условие возникновения колебаний, т. е. *условие самовозбуждения* (система предполагается мягкой). Вводя затухание d контура, можно переписать это условие в виде

$$M > \frac{d}{\omega_0 S_0}.$$

Таким образом, чтобы генератор возбудился при данных затухании и настройке контура и при данном триоде, нужно увеличивать обратную связь (т. е. коэффициент взаимоиנדукции M между контурной катушкой и катушкой обратной связи).

Условие самовозбуждения можно истолковать еще несколько иначе. Перепишем уравнение (229) в виде

$$L \frac{d^2 u}{dt^2} + \left(R - \frac{MS_0}{C} \right) \frac{du}{dt} + \frac{1}{C} u = 0.$$

Теперь можно представить дело так, что обратная связь вносит в контур *отрицательное сопротивление*

$$R' = -\frac{MS_0}{C}$$

и условие самовозбуждения с этой точки зрения состоит в том, что результирующее сопротивление становится отрицательным, т. е.

$$R + R' = R - \frac{MS_0}{C} < 0. \quad (230)$$

Такая трактовка вопроса очень распространена в радиотехнике.

§ 82. Комплексная обратная связь

Обратная связь в генераторе вносит в контур отрицательное активное сопротивление. Такого рода обратную связь называют *положительной* в отличие от *отрицательной обратной связи*, широко применяемой в усилителях и других радиотехнических устройствах

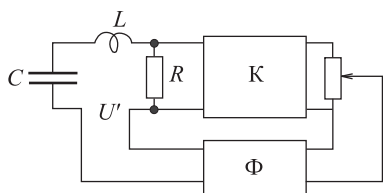


Рис. 284

для исправления характеристик и уменьшения искажений. Но обратная связь может рассматриваться и в более общем виде. Возьмем к примеру схему рис. 284. Здесь имеется контур LCR . Усиленное напряжение снимается с выхода усилителя, и через делитель напряжения и *фазовращатель* Φ вводится обратно в контур; обусловленное обратной

связью дополнительное напряжение может иметь произвольные фазу и амплитуду. Это и есть наиболее общий случай обратной связи. На рис. 285, *а* показана векторная диаграмма напряжений в контуре. Если разделить все напряжения на общий ток, то можно перейти от

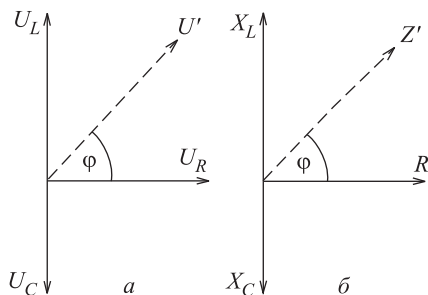


Рис. 285

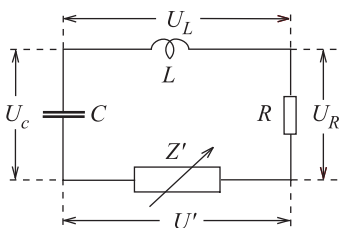


Рис. 286

диаграммы напряжений к совершенно подобной ей диаграмме сопротивлений, как это делается в теории переменных токов. Диаграмма сопротивлений изображена на рис. 285, *б*. Она соответствует эквивалентной схеме рис. 286, на которой обозначены также все напряжения. Теперь видно, что действие обратной связи в общем случае может быть истолковано как внесение в схему *произвольного* комплексного

сопротивления. Характер этого сопротивления зависит от фазы обратной связи. При $\varphi = 0$ имеем $Z' = R'$, результирующее активное сопротивление возрастает; это есть отрицательная обратная связь. При $\varphi = \pi$ получаем $Z' = -R'$, т.е. обратная связь вносит отрицательное активное сопротивление; это есть положительная обратная связь. Если при этом R' по абсолютному значению больше R , то результирующее сопротивление отрицательно, и в системе возникают автоколебания. Таким образом, при положительной обратной связи условие

$$|R'| > R$$

есть условие самовозбуждения.

Рассмотрим теперь случай произвольного угла φ . Пусть значение φ заключено между $\pi/2$ и π (рис. 287). Тогда вносимое обратной связью сопротивление Z' комплексно и может быть выражено как

$$Z' = R' + iX',$$

что с геометрической точки зрения означает построение двух проекций вектора Z' , как показано на рисунке. Если $|R'| > R$, то система самовозбудится. Наличие X' , как видно из диаграммы, равносильно увеличению индуктивного сопротивления X_L (или, что то же самое, уменьшению емкостного сопротивления X_C). Иначе говоря, при описываемых условиях обратная связь как бы увеличивает индуктивность или емкость. Следствием этого является изменение *настройки* контура и при возбуждении в системе колебаний они будут иметь меньшую частоту, чем резонансная частота

$$\omega_0 = 1/\sqrt{LC}$$

самого контура. Если $\pi < \varphi < \frac{3}{2}\pi$, то обратная связь будет повышать частоту колебаний. Отсюда следует, что во избе-

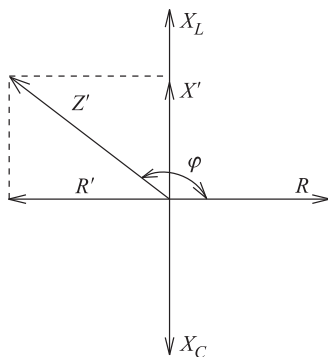


Рис. 287

жание влияния обратной связи на частоту нужно позаботиться, чтобы в генераторе фаза обратной связи была точно равна π ($\varphi = 180^\circ$; «чистая» положительная обратная связь)¹⁾. С другой же стороны, все описанные соотношения можно использовать для построения схемы частотной модуляции. Если фаза лежит в пределах $-\pi/2 < \varphi < \pi/2$, то влияние обратной связи на настройку сохраняется, но автоколебания

¹⁾ В генераторе по схеме рис. 283 это условие выполнено. Если схема не генерирует, то первое, что нужно сделать, это поменять концы у одной из катушек (не меняя ее положения), так как возможно, что катушка включена неправильно, и вместо положительной обратной связи получается отрицательная.

уже невозможны, так как составляющая R' дает отрицательную обратную связь, т. е. увеличивает результирующее сопротивление контура.

§ 83. Инженерные критерии устойчивости

Вопрос о возможности возникновения автоколебаний в некоторой схеме интересен не только с точки зрения построения генераторов. Генератор есть источник колебаний; автоколебания являются основной его функцией. Другие же звенья радиотехнических устройств, как, например, усилители, преобразователи частоты, детекторы и т. п., работают в режиме вынужденных колебаний, и возникновение в этих звеньях автоколебаний в большинстве случаев не только не желательно, но совершенно недопустимо. Поэтому вопрос о возможности возникновения автоколебаний в данной схеме имеет в радиотехнике довольно широкое значение. Вопрос этот ставится как вопрос об устойчивости в том смысле, что схема устойчива, если в нормальном для нее режиме автоколебания в ней не возникают, и наоборот. Таким образом, всякий генератор ¹⁾ принципиально должен быть, неустойчивым, а все остальные схемы, назначением которых не является генерирование колебаний, должны быть устойчивыми. Следует пояснить, что речь идет об устойчивости начального состояния системы, а не об устойчивости уже существующих автоколебаний (см. § 80).

Общий вопрос об устойчивости, т. е. о том, могут или нет в данной системе возбудиться автоколебания, решается на основании некоторых общих признаков, выражаемых при помощи так называемых *критериев устойчивости*. Существует несколько критериев устойчивости, по-разному отображающих единые свойства системы. В основе всех известных критериев лежит исследование корней характеристического уравнения системы.

Дело в том, что решение обыкновенного дифференциального уравнения с постоянными коэффициентами

$$\sum_{k=0}^n a_k \frac{d^k y}{dt^k} = 0$$

имеет вид

$$y(t) = \sum_{k=0}^n A_k e^{\gamma_k t},$$

где A_k — постоянные интегрирования, определяемые из начальных условий, а γ_k — корни характеристического уравнения. Система будет устойчивой, если вещественные части всех корней отрицательны. Физически это означает, что все колебания, возникшие в системе по тем или иным причинам, будут затухающими. Возможно, не решая

¹⁾ В этом параграфе здесь и дальше имеется в виду генератор с мягким возбуждением.

характеристического уравнения, т. е. не находя корней, установить, выполняется ли вышеуказанное требование. Это можно сделать при помощи критерия Рауса–Гурвица, основанного на свойствах определителей, составленных из коэффициентов a_k .

Критерий Рауса–Гурвица имеет чисто аналитический характер. Существуют и другие критерии, которым придана форма графических построений, что делает их более пригодными для инженерной практики. Приведем здесь без доказательства описание двух наиболее известных и распространенных критериев устойчивости, а именно *критерия Найквиста* и *критерия Михайлова*.

Начнем с некоторых предварительных соображений. Пусть имеется некоторый четырехполюсник, например, усилитель с коэффициентом передачи

$$K = \frac{\dot{U}_2}{\dot{U}_1}.$$

Пусть, далее, выход и вход усилителя соединены между собой цепью обратной связи с коэффициентом передачи

$$\beta = \frac{\dot{U}_3}{\dot{U}_2}.$$

При таких условиях вполне возможно, что схема будет генерировать, т. е. что в ней возбудятся автоколебания. Но если в замкнутой схеме существуют *установившиеся* синусоидальные колебания, то обязательно должно выполняться условие

$$K_1 = K\beta = \frac{\dot{U}_2}{\dot{U}_1} \frac{\dot{U}_3}{\dot{U}_2} = \frac{\dot{U}_3}{\dot{U}_1} = 1,$$

так как $U_1 = U_3$ по той простой причине, что это одно и то же напряжение (см. схему рис. 288, а).

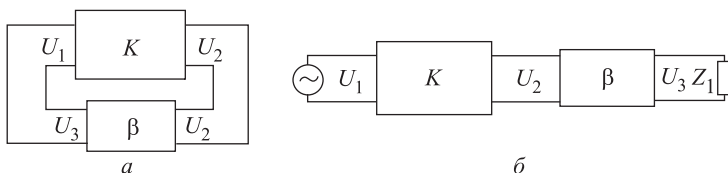


Рис. 288

Коэффициент передачи есть комплексная величина, которая может быть записана в виде

$$K_1 = A e^{i\varphi}.$$

Поэтому равенство

$$K_1 = 1$$

распадается на два:

$$A = 1, \quad \varphi = 0.$$

Эти два равенства известны под названием *условий баланса амплитуд* и *баланса фаз*. Важно заметить, что эти равенства являются *точными*, и что в установившемся режиме *оба* равенства обязательно выполняются. Условие баланса амплитуд представляет собой уравнение, из которого может быть найдена установившаяся амплитуда колебаний; из условия баланса фаз находится частота колебаний.

Разорвем теперь цепь обратной связи, как показано на рис. 288, б и попытаемся воспроизвести в такой *разомкнутой* схеме тот же режим, который имелся в замкнутой схеме рис. 288, а. Для этого нужно подать на вход переменное напряжение от независимого генератора. На выходном же конце нужно нагрузить схему сопротивлением, равным входному сопротивлению четырехполюсника K . Предположим теперь, что частота генератора в схеме рис. 288, б плавно изменяется от 0 до ω . Для каждого значения частоты можно определить комплексное отношение

$$K_1(\omega) = \frac{\dot{U}_3}{\dot{U}_1}.$$

Будем строить K_1 как вектор на комплексной плоскости и следить за траекторией конца этого вектора при изменении частоты. Эту траекторию называют *амплитудно-фазовой* характеристикой. Построение показано на рис. 289. Во многих случаях коэффициент передачи при $\omega = 0$ и при $\omega = \infty$ равен нулю. В этих случаях амплитудно-фазовая характеристика образует замкнутый контур: точки начала и конца обхода контура совпадают с началом координат. Каждая точка амплитудно-фазовой (аф) характеристики отвечает определенному значению частоты (рис. 290).

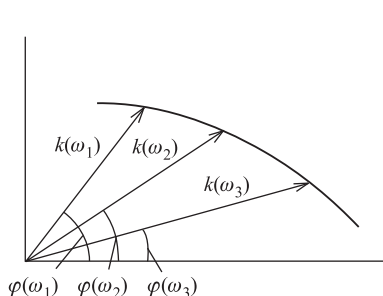


Рис. 289

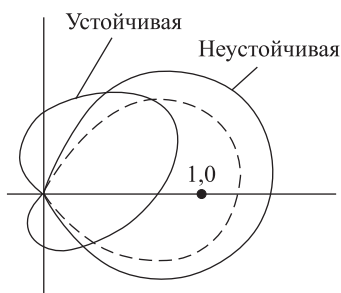


Рис. 290

Если установившиеся автоколебания возможны, то аф характеристика проходит через точку с полярными координатами $A = 1$, $\varphi = 0$ — так называемая «точка (1,0)», отвечающая условию баланса фаз и амплитуд. При этом будут генерироваться колебания с частотой, отвечающей положению вектора K_1 вдоль вещественной оси ($\varphi = 0$).

Если при $\varphi = 0$ $A < 1$, то это означает, что колебания в системе могут быть только затухающими. Если же $A > 1$, то это значит, что колебания не только возникнут, но будут нарастать. Отсюда и следует критерий Найквиста: замкнутая система будет неустойчивой, т.е. в ней возникнут автоколебания, если аф характеристика разомкнутой системы охватывает точку $(1, 0)$. На рис. 290 показаны примеры аф характеристик для устойчивой и неустойчивой систем.

Амплитудно-фазовая характеристика может быть снята экспериментально по схеме рис. 288, б либо вычислена, если схема и параметры четырехполюсника известны.

Амплитудно-фазовые характеристики не только могут дать ответ на вопрос об устойчивости системы, но позволяют также определить частоту и амплитуду установившихся колебаний. Мы не затрагивали до сих пор вопроса о том, при каких амплитудах входного напряжения снимается аф характеристика. Если бы исследуемый четырехполюсник был линейным, то коэффициент передачи K не зависел бы от амплитуды. При таких условиях $A > 1$ при $\varphi = 0$ означало бы неограниченное нарастание амплитуды автоколебаний, так как *установившийся* режим с неизменной амплитудой может быть только при $A = 1$. В действительности всякая система нелинейна, а потому по мере увеличения амплитуды коэффициент передачи убывает. Контур аф характеристики при этом стягивается (пунктир на рис. 290), и если первоначально точка $(1, 0)$ была охвачена этим контуром ($A > 1$), то рано или поздно контур пройдет через точку $(1, 0)$. Амплитуда, при которой это случится, и есть установившееся значение амплитуды. Одновременно определится и значение частоты, генерируемой системой в установившемся режиме. Теперь можно уточнить вышеприведенную формулировку критерия Найквиста в том смысле, что если речь идет об устойчивости системы в состоянии покоя (т.е. когда колебаний еще нет), то аф характеристика должна сниматься при самых малых амплитудах входного напряжения.

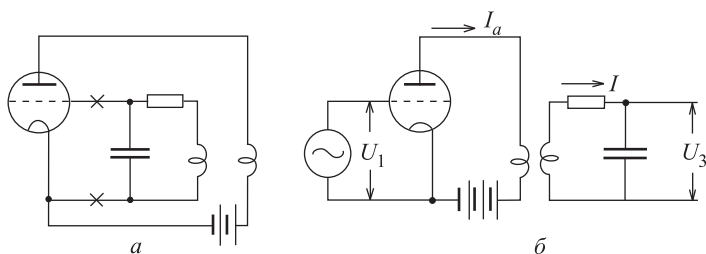


Рис. 291

В качестве примера выведем при помощи критерия Найквиста условие самовозбуждения для генератора с контуром в цепи сетки, уже рассмотренного выше (см. § 81). На рис. 291, а дана замкнутая схема

генератора, на рис. 291, б — та же схема в разомкнутом виде (цепь обратной связи разорвана в точках, отмеченных на рис. 291, а крестиками). Найдем коэффициент передачи для схемы рис. 291, б. Имеем:

$$\begin{aligned} \dot{I}_a &= S\dot{U}_1, \quad \dot{E} = i\omega M\dot{I}_a, \\ \dot{I} &= \frac{\dot{E}}{R + i\left(\omega L - \frac{1}{\omega C}\right)}, \quad \dot{U}_3 = \frac{\dot{I}}{i\omega C}, \end{aligned}$$

откуда

$$K_1 = \frac{\dot{U}_3}{\dot{U}_1} = \frac{MS}{C} \frac{1}{R + i\left(\omega L - \frac{1}{\omega C}\right)}.$$

Найдем модуль и аргумент этого выражения:

$$\begin{aligned} A = |K_1| &= \frac{MS}{C} \frac{1}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}, \\ \varphi &= \arctg \frac{\frac{1}{\omega C} - \omega L}{R}. \end{aligned} \quad (231)$$

Уравнение аф характеристики в полярных координатах имеет вид

$$A = \frac{MS}{RC} \frac{1}{\sqrt{1 + \operatorname{tg}^2 \varphi}} = \frac{MS}{RC} \cos \varphi.$$

Таким образом, график аф характеристики в рассматриваемом случае имеет вид окружности, проходящей через начало координат и расположенной симметрично относительно вещественной оси (рис. 292).

Диаметр окружности равен

$$A_0 = \frac{MS}{RC}.$$

Легко видеть, что условие охвата точки (1, 0) выполняется при

$$A_0 = \frac{MS}{RC} > 1,$$

а это и есть условие самовозбуждения, полученное другим путем в § 81. Для определения генерируемой частоты нужно приравнять φ нулю. Это дает (см. (231))

$$\omega = \omega_0 = 1/\sqrt{LC}.$$

В более общем случае будем иметь нелинейное выражение, в котором A будет функцией как частоты, так и амплитуды. Составляя два уравнения, выражающие условие баланса фаз и амплитуд, т. е.

$$A = 1, \quad \varphi = 0,$$

найдем из них две искомые величины — амплитуду и частоту установившихся колебаний.

Отличие критерия Михайлова от критерия Найквиста состоит прежде всего в том, что он приспособлен для исследования устойчивости замкнутых систем. Составляется уравнение данной системы с учетом обратной связи в операторной форме и приводится к виду

$$D(p)x = 0,$$

где x — колебательная величина (ток, напряжение), относительно которой составлено уравнение. Затем подстановкой $p = i\omega$ образуется комплексная величина $D(i\omega)$, представляемая вектором на комплексной плоскости. Частоте придают все значения от 0 до ∞ . При этом конец вектора D описывает траекторию, называемую *годографом Михайлова*. Вид этого годографа и определяет устойчивость системы на основании следующего правила: для того чтобы система была устойчивой, необходимо и достаточно, чтобы при изменении частоты от 0 до ∞ вектор D , беря начало на положительной вещественной полуоси, вращался только против часовой стрелки, нигде не обращаясь в нуль. При этом для системы, описываемой уравнением n -го порядка, вектор поворачивается при изменении частоты в указанном интервале на угол $n\frac{\pi}{2}$.

Это правило и есть формулировка критерия Михайлова.

На рис. 293 изображены годографы для устойчивых систем различных порядков.

Покажем применение критерия Михайлова на примере, в качестве которого возьмем все тот же генератор по схеме рис. 291, а. Для этого генератора в § 81 было выведено следующее уравнение:

$$\frac{d^2 u}{dt^2} + \frac{1}{L} \left(R - \frac{MS_0}{C} \right) \frac{du}{dt} + \omega_0^2 u = 0$$

или в операторной форме

$$(p^2 + 2\delta p + \omega_0^2)\bar{u} = D(p)\bar{u} = 0,$$

где для сокращения обозначено

$$2\delta = \frac{1}{L} \left(R - \frac{MS_0}{C} \right).$$

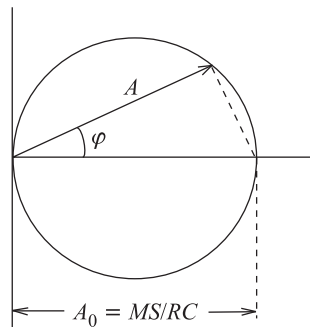


Рис. 292

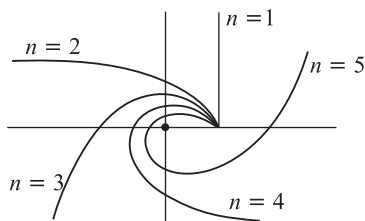


Рис. 293

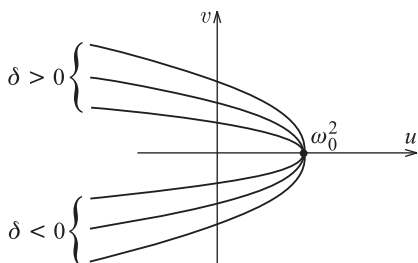


Рис. 294

По физическому смыслу δ есть результирующий (т.е. с учетом действия обратной связи) показатель затухания. Заменяя p на $i\omega$, получим

$$D(i\omega) = \omega_0^2 - \omega^2 + i2\omega\delta = u + i v.$$

Уравнение годографа удобнее в данном случае выразить в прямоугольных координатах u и v . Мы имеем:

$$u = \omega_0^2 - \omega^2, \quad v = 2\delta\omega.$$

Исключая ω , получим

$$v = 2\delta\sqrt{\omega_0^2 - u}.$$

Это есть уравнение параболы. Мы рассматриваем только одну ветвь параболы, отвечающую положительным значениям ω . Учитывая это, получим для различных значений δ семейство кривых, представленное на рис. 294. Кривые, лежащие выше вещественной оси, удовлетворяют условию устойчивости; они соответствуют положительным значениям δ . Система теряет устойчивость при $\delta = 0$, т.е. при

$$R = \frac{MS_0}{C}.$$

Кривые, лежащие ниже оси, соответствуют уже отрицательным значениям δ ; применяя к ним вышеуказанное правило, видим, что эти кривые выражают неустойчивые состояния системы, так как вектор поворачивается (при следовании его конца по кривой) по часовой стрелке.

§ 84. Установившийся режим генератора

В этом параграфе будем исследовать установившийся режим лампового генератора, пользуясь элементарной квазилинейной теорией.

Сущность квазилинейной теории генератора состоит в том, что нелинейность, ограничивающая амплитуду автоколебаний, учитывается в уравнении генератора особым образом, а именно: в качестве одного из параметров в уравнение вводится *средняя крутизна триода*,

величина не постоянная, а функция *амплитуды* колебаний. Однако для неизменной амплитуды, т.е. для незатухающих и ненарастающих колебаний средняя крутизна также остается неизменной. Отсюда следует, что в *установившемся* режиме со средней крутизной можно обращаться как с постоянной величиной, а стало быть, уравнение генератора принимает форму линейного уравнения с постоянными коэффициентами. Отсюда и названия «квазилинейная (как бы линейная) теория», «квазилинейное уравнение». Из сказанного следует, что квазилинейная теория приспособлена в первую очередь для исследования установившихся режимов. Впрочем, с ее помощью можно разрешить и более сложные вопросы.

Будем вычислять методами квазилинейной теории установившуюся амплитуду колебаний в ламповом генераторе. Объектом исследования послужит опять генератор с контуром в цепи сетки. Составим уравнение напряжений для контура

$$L \frac{di}{dt} + Ri + \frac{1}{C} \int i dt = M \frac{di_a}{dt}. \quad (232)$$

Предположим, что напряжение и ток в контуре синусоидальны. Это обычное для квазилинейной теории предположение. Оно означает, что мы пренебрегаем влиянием гармоник анодного тока. Практически это предположение оправдывается тем, что контур генератора, обладая большой добротностью, а следовательно, высокой избирательностью, отфильтровывает гармоники. Таким образом, можно в правой части заменить анодный ток его первой гармоникой. После этого возможен переход к комплексной форме уравнения (232) (путем замены дифференцирования по времени умножением на $i\omega$)

$$\left(i\omega L + R + \frac{1}{i\omega C} \right) \dot{I} = i\omega M \dot{I}_{a1}$$

или, вводя $\dot{U} = \frac{1}{i\omega C} \dot{I}$,

$$\left(\omega_0^2 - \omega^2 + i\omega \frac{R}{L} \right) \dot{U} = i\omega \frac{M \dot{I}_{a1}}{LC}.$$

Здесь $\dot{I}_{a1}$ — амплитуда первой гармоники анодного тока.

Воспользуемся определением средней крутизны

$$S_c = \left| \frac{\dot{I}_{a1}}{\dot{U}} \right| = \frac{\dot{I}_{a1}}{\dot{U}} \quad 1)$$

и получим

$$\omega_0^2 - \omega^2 + i\omega \frac{1}{L} \left(R - \frac{MS_c}{C} \right) = 0.$$

¹⁾ Так как i_{a1} и u совпадают по фазе.

Приравнявая нулю порознь действительную и мнимую части, получим два уравнения. Из первого находим частоту установившихся колебаний

$$\omega = \omega_0.$$

Второе же уравнение имеет вид

$$R - \frac{MS_c}{C} = 0. \quad (233)$$

Это соотношение очень похоже по внешнему виду на формулу (230), но имеет совершенно другой смысл. *Неравенство* (230)

$$R - \frac{MS_0}{C} < 0,$$

выведенное в § 81, выражает условие самовозбуждения. Входящая в него величина S_0 есть *начальная крутизна*, или крутизна в рабочей точке, определяемая как

$$S_0 = \left(\frac{di_a}{du} \right)_{u=E_0},$$

где E_0 — абсцисса рабочей точки на характеристике лампы. Выражение же (233) представляет собой *равенство*, обязательно выполняемое при установившихся колебаниях с конечной амплитудой. Входящая в (233) величина S_c есть *средняя крутизна*, зависящая не только от выбора рабочей точки, но и от амплитуды. Но если S_c есть функция амплитуды, то (233) может рассматриваться как уравнение, решая которое можно найти установившуюся амплитуду. Именно это мы сейчас и сделаем.

Выберем подходящую аппроксимацию характеристики лампы. Достаточно представить характеристику многочленом пятой степени

$$i_a = a_0 + a_1 u + a_3 u^3 + a_5 u^5. \quad (234)$$

Для получения дифференциальной крутизны

$$S = \frac{di_a}{du}$$

достаточно было бы продифференцировать (234) по u . Но нам нужно найти среднюю крутизну. Положим для этого

$$u = U_m \cos \omega t.$$

Тогда

$$i_a = a_0 + a_1 U_m \cos \omega t + a_3 U_m^3 \cos^3 \omega t + a_5 U_m^5 \cos^5 \omega t.$$

Пользуясь таблицей степеней косинуса (с. 316), получим

$$\begin{aligned} i_a = a_0 + a_1 U_m \cos \omega t + \frac{1}{4} a_3 U_m^3 (3 \cos \omega t + \cos 3\omega t) + \\ + \frac{1}{16} a_5 U_m^5 (10 \cos \omega t + 5 \cos 3\omega t + \cos 5\omega t). \end{aligned}$$

Отсюда амплитуда первой гармоники (т. е. коэффициент при $\cos \omega t$)

$$I_{a1} = a_1 U_m + \frac{3}{4} a_3 U_m^3 + \frac{5}{8} a_5 U_m^5.$$

Средняя крутизна по определению

$$S_c = \frac{I_{a1}}{U_m} = a_1 + \frac{3}{4} a_3 U_m^2 + \frac{5}{8} a_5 U_m^4. \quad (235)$$

Итак, мы получили выражение для средней крутизны как функции амплитуды. Из (233) получаем

$$S_c - \frac{RC}{M} = 0 \quad (236)$$

или

$$\frac{5}{8} a_5 U_m^4 + \frac{3}{4} a_3 U_m^2 + a_1 - \frac{RC}{M} = 0. \quad (237)$$

Это есть биквадратное уравнение, легко разрешаемое относительно интересующей нас амплитуды U_m .

Не останавливаясь на этом, представим уравнение (236) в графической форме. Будем откладывать оба члена левой части (236), т. е. величины S_c и $\frac{RC}{M}$ по оси ординат, а амплитуду напряжения

U_m — по оси абсцисс. $\frac{RC}{M}$ не зависит от U_m , это — постоянная величина, график которой есть горизонтальная прямая. Эту прямую будем называть *прямой обратной связи*, так как ее расположение зависит от глубины обратной связи, выражаемой через коэффициент взаимоиндукции M . Средняя крутизна зависит от амплитуды, и график этой зависимости будет выражаться некоторой кривой, форма которой зависит от знаков и величин коэффициентов a_k . Положим для начала, что оба коэффициента a_3 и a_5 отрицательны. Тогда средняя крутизна, имеющая при $U_m = 0$ значение

$$S_c = a_1 = S_0,$$

с увеличением U_m монотонно убывает. График для этого случая показан на рис. 295. Если прямая обратной связи лежит выше кривой S_c (пунктир на рис. 295), то равенство (236) не выполняется ни при каких значениях амплитуды, и автоколебания вообще невозможны. Если же увеличить обратную связь, то прямая обратной связи опустится и будет пересекаться с кривой S_c в некоторой точке, как показано на рис. 295. Точка пересечения отвечает равенству (236); абсцисса этой точки непосредственно дает установившуюся амплитуду U_0 . Если обратная связь будет увеличена, то прямая обратной связи опустится еще ниже, точка пересечения сдвинется вправо. Это означает, что амплитуда установившихся колебаний увеличится. Если же мы будем уменьшать обратную связь, то точка пересечения будет двигаться влево, пока не дойдет до оси ординат. При этом $S_c = S_0$, и это положение соответствует границе

области самовозбуждения; амплитуда при этом обращается в нуль; при дальнейшем уменьшении обратной связи условие самовозбуждения не будет уже выполнено и автоколебания не возникнут.

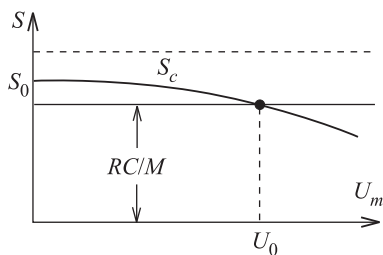


Рис. 295

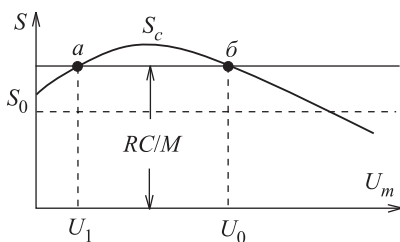


Рис. 296

Вся описанная картина относится к *мягкой* системе. В самом деле, так как

$$S_c < S_0,$$

то, если выполнено условие самовозбуждения

$$\frac{RC}{M} - S_0 < 0,$$

обязательно возникают автоколебания с вполне определенной установившейся амплитудой.

Рассмотрим теперь другой случай: пусть $a_3 > 0$, $a_5 < 0$. При малых амплитудах квадратичный член (см. формулу (235)) преобладает, а потому вначале S_c *растет* с увеличением амплитуды. В дальнейшем же член, содержащий четвертую степень амплитуды, а потому растущий быстрее квадратичного, оказывает решающее влияние, и средняя крутизна, достигнув максимума, начинает убывать. Кривая средней крутизны для рассматриваемого случая показана на рис. 296. Пусть прямая обратной связи занимает положение, показанное на рисунке сплошной прямой. Условие самовозбуждения

$$S_0 > \frac{RC}{M}$$

при этом не выполнено, и автоколебания не возникнут. Однако имеются две точки пересечения кривой S_c с прямой обратной связи. Для обсуждения состояния системы в этих точках введем термины *избыток* и *недостаток крутизны*, относящиеся к неравенствам

$$S_c \gtrless \frac{RC}{M}.$$

Эти термины имеют прямую связь с энергетическими соотношениями: избыток крутизны $\left(S_c > \frac{RC}{M}\right)$ означает, что система получает

энергии больше, чем теряет, а недостаток крутизны $\left(S_c < \frac{RC}{M}\right)$ — наоборот, что расход энергии больше прихода.

Рассмотрим состояние системы в точке a . Если амплитуда по какой-либо причине больше U_1 , то будем иметь избыток крутизны, и амплитуда будет возрастать и далее. Если же амплитуда станет меньше U_1 , то вследствие недостатка крутизны она будет убывать и далее, пока колебания не затухнут. Таким образом, точка a — неустойчивая точка. В точке же b увеличение амплитуды вызывает недостаток крутизны, а уменьшение амплитуды — избыток крутизны. Таким образом, точка b — устойчивая точка; абсцисса этой точки определяет установившуюся амплитуду U_0 . Графики рис. 296 изображают свойства жесткой системы: для того чтобы получить установившиеся колебания, нужно дать системе такой начальный толчок, чтобы возникли колебания с амплитудой, большей U_1 . Дальнейшее увеличение амплитуды вплоть до установившегося значения U_0 произойдет автоматически. Графики рис. 295 и 296 изображают по существу те же энергетические соотношения, что и графики рис. 281 и 282. Если увеличить обратную связь настолько, что будет выполнено условие самовозбуждения (пунктирная прямая на рис. 296), то система делается мягкой, независимо от формы кривой средней крутизны.

Можно получить как мягкий, так и жесткий режим при одном и том же триоде путем соответствующего выбора рабочей точки. Если рабочая точка выбрана в точке перегиба характеристики

$$i_a = f(u_c),$$

то средняя крутизна будет монотонно убывать с увеличением амплитуды (рис. 295). Если же взять рабочую точку на загибе характеристики, то средняя крутизна будет сначала возрастать, а затем убывать. Аналитически условие получения того или иного режима выражается знаком коэффициента при члене третьей степени в разложении характеристики в степенной ряд. Если обозначить напряжение смещения через E_0 (абсцисса рабочей точки), то по формуле Тейлора

$$i = f(E_0 + u) = f(E_0) + f'(E_0)\frac{u}{1!} + f''(E_0)\frac{u^2}{2!} + f'''(E_0)\frac{u^3}{3!} + \dots$$

Таким образом,

$$a_3 = \frac{1}{3!}f'''(E_0),$$

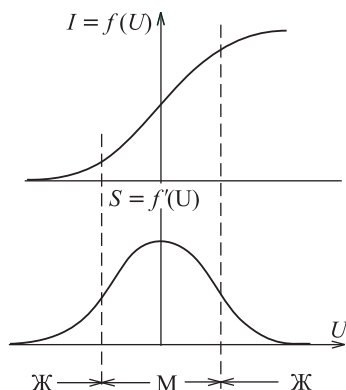


Рис. 297

т. е. коэффициент a_3 (см. (235)), знак которого определяет мягкость или жесткость системы, пропорционален третьей производной функции $f(U)$ в рабочей точке $U = E_0$.

Равенство

$$f'''(E_0) = 0$$

определяет границу между мягким и жестким режимом. Но при $f''' = 0$ имеем максимум или минимум f'' , что соответствует точке перегиба f' . С другой же стороны, $f'(u_c)$ есть не что иное, как дифференциальная крутизна. Итак, границы между мягким и жестким режимом проходят через точки перегиба характеристики крутизны

$$S = f'(u_c) = \frac{di_a}{du_c},$$

как показано на рис. 297 вертикальными прямыми.

§ 85. Установление колебаний в генераторе

После рассмотрения установившегося режима генератора переходим к процессу установления: проследим развитие колебательного явления от момента включения схемы и до того времени, когда амплитуда и частота практически уже не будут меняться, т. е. когда можно считать колебание уже установившимся.

Задача сводится к тому, чтобы проследить изменение во времени величин, характеризующих колебание, т. е. в первую очередь амплитуды, а затем и частоты. Форму колебаний будем полагать близкой к синусоидальной.

Задача будет решаться методом, известным под названием *метода медленно меняющихся амплитуд*. По существу своему этот метод относится к квазилинейным; возможность его применения к данной задаче основана на том, что амплитуда, хотя и не постоянная, как при установившемся режиме, предполагается *медленно* изменяющейся. Критерий «медленности» будет дан ниже.

Запишем уравнение генератора в виде

$$L \frac{di}{dt} + Ri + \frac{1}{C} \int i dt = M \frac{di_a}{dt}. \quad (238)$$

Этим уравнением мы уже пользовались в §§ 81 и 84. Но в § 81 мы полагали $i_a = S_0 u$ и свели, таким образом, уравнение (238) к линейному уравнению, чтобы найти условие самовозбуждения. В § 84 мы составили уравнение для первой гармоники в комплексной форме, введя в уравнение среднюю крутизну; задача состояла в исследовании установившегося режима. Теперь же придадим уравнению (238) общую нелинейную форму.

Зависимость анодного тока от напряжения на сетке представляется нелинейной функцией

$$i_a = f(u).$$

Запишем правую часть (238) в виде

$$M \frac{di_a}{dt} = M \frac{di_a}{du} \frac{du}{dt}.$$

Но

$$\frac{di_a}{du} = \frac{df(u)}{du} = S(u)$$

— дифференциальная крутизна, выражаемая графически наклоном касательной к характеристике триода в данной точке (т. е. в точке с абсциссой u). Вводя, как и прежде,

$$u = \frac{1}{C} \int i dt,$$

получим вместо (238) исходное уравнение генератора в виде

$$\frac{d^2 u}{dt^2} + \frac{1}{L} \left[R - \frac{M}{C} S(u) \right] \frac{du}{dt} + \omega_0^2 u = 0. \quad (239)$$

Положим, что рассматривается мягкая система. В этом случае нелинейная характеристика может аппроксимироваться степенным многочленом третьей степени

$$i_a = a_0 + a_1 u + a_3 u^3.$$

Дифференцируя, находим выражение для крутизны

$$S = \frac{di_a}{du} = a_1 + 3a_3 u^2.$$

Подставляя это выражение в (238), получаем нелинейное уравнение мягкой системы в виде

$$\frac{d^2 u}{dt^2} + (2\delta_0 + \gamma u^2) \frac{du}{dt} + \omega_0^2 u = 0. \quad (240)$$

Здесь

$$\delta_0 = \alpha - \frac{1}{2} \omega_0^2 M S_0$$

— начальный показатель затухания (с учетом действия обратной связи),

$$\gamma = 3\omega_0^2 M a_3.$$

Способ решения уравнения (240) состоит в том, что мы предполагаем решение в виде

$$u = A(t) \sin \omega t. \quad (241)$$

Величина A есть переменная «амплитуда» колебания; функция $A(t)$ выражает то, что в радиотехнике называется *огibaющей*. Неизвестную пока частоту ω будем считать постоянной¹⁾.

¹⁾ Это лишь первое приближение. В действительности в процессе установления меняются как амплитуда, так и частота колебаний. Но изменения частоты

Продифференцируем предполагаемое решение (240) дважды:

$$u' = A' \sin \omega t + A\omega \cos \omega t,$$

$$u'' = A'' \sin \omega t + 2A'\omega \cos \omega t - A\omega^2 \sin \omega t$$

и подставим значения u и его производных в (240)

$$A'' \sin \omega t + 2A'\omega \cos \omega t - A\omega^2 \sin \omega t + \\ + (2\delta_0 + \gamma A^2 \sin^2 \omega t)(A' \sin \omega t + A\omega \cos \omega t) + \omega_0^2 A \sin \omega t = 0.$$

Произведя тригонометрические преобразования и приравняв по-порознь нулю коэффициенты при $\sin \omega t$ и $\cos \omega t$, получим два уравнения:

$$\left. \begin{aligned} A'' + \left(2\delta_0 + \frac{3}{4}\gamma A^2\right) A' + (\omega_0^2 - \omega^2)A &= 0 \\ 2A' + \left(2\delta_0 + \frac{1}{4}\gamma A^2\right) A &= 0 \end{aligned} \right\}. \quad (242)$$

При этом члены, содержащие тройной аргумент ($3\omega t$), отброшены в силу исходного предположения о синусоидальности колебания (см. формулу (241)).

Уравнения (242) могут быть еще упрощены. Дело в том, что члены этих уравнений имеют различный порядок малости. Мы предполагаем, что функция $A(t)$, выражающая изменение амплитуды, есть *медленная* функция. Это значит, что изменения амплитуды совершаются значительно медленнее, чем само колебательное изменение напряжения, или, проще говоря, что относительное изменение амплитуды за период мало. Функция $f_1(t)$ считается медленной по сравнению с какой-либо другой функцией $f_2(t)$, если ее производная $f_1'(t)$ много меньше производной $f_2'(t)$ (для сравнения нужно предварительно нормировать производные, например, разделив их на среднеквадратичные значения самих функций).

Производная функции, выражающей колебание, т. е. $u = A \sin \omega t$, не превосходит по абсолютной величине ωA . Производная же огибающей $A(t)$ обозначена A' , и условие медленности $A(t)$ может быть записано в виде

$$|A'| \ll \omega A.$$

Другими словами A' есть величина первого порядка малости по сравнению с ωA . Но если так, то вторая производная функции $A(t)$ будет уже величиной второго порядка малости по сравнению со второй производной u , имеющей порядок $\omega^2 A$.

На основании этих соображений можно вычеркнуть из первого уравнения (242) член A'' , как имеющий второй порядок малости по сравнению с последним членом.

обычно невелики; учтя эти изменения, получим во втором приближении лишь незначительную поправку на закон изменения амплитуды.

В результате получим вместо (242):

$$\left. \begin{aligned} \left(2\delta_0 + \frac{3}{4}\gamma A^2\right) A' + (\omega_0^2 - \omega^2)A &= 0 \\ 2A' + \left(2\delta_0 + \frac{1}{4}\gamma A^2\right) A &= 0 \end{aligned} \right\}. \quad (243)$$

Уравнения, получаемые путем вычеркивания членов высших порядков малости, называются *укороченными* уравнениями. Укорачивание уравнений является общим приемом метода медленно меняющихся амплитуд. Он позволяет понизить порядок уравнений и упростить их строение. Впрочем, в рассматриваемом случае эти преимущества останутся неиспользованными, так как мы воспользуемся ниже только вторым из двух уравнений (243), не подвергшимся укорочению.

Уравнения (243) являются нелинейными, так как коэффициенты зависят от амплитуды. Найдем из этих уравнений прежде всего установившиеся амплитуду и частоту.

В установившемся режиме $A' = 0$, $A = A_0$. Первое уравнение дает

$$\omega = \omega_0,$$

а из второго находим

$$A_0 = \sqrt{-\frac{8\delta_0}{\gamma}} \quad (244)$$

или, выражая δ_0 и γ через исходные параметры,

$$A_0 = \frac{2}{\sqrt{3}} \sqrt{\frac{a_1}{a_3} - \frac{2\alpha}{\omega_0^2 M a_3}} = \frac{2}{\sqrt{3a_3}} \sqrt{S_0 - \frac{RC}{M}}. \quad (245)$$

Это то же значение, которое мы нашли бы из уравнения (237), положив $a_5 = 0$. Как видим, амплитуда колебаний определяется нелинейностью, выражаемой в нашем случае коэффициентом a_3 ; амплитуда зависит также от собственного затухания контура, начальной крутизны и глубины обратной связи.

Из формулы (245) видно, что амплитуда тем меньше, чем больше нелинейность (т. е. чем больше коэффициент a_3); амплитуда тем больше, чем с большим запасом выполнено условие самовозбуждения, т. е. условие

$$S_0 > \frac{RC}{M}.$$

Теперь найдем закон установления амплитуды, т. е. функцию $A(t)$, воспользовавшись вторым уравнением (243). Для решения этого уравнения умножим его почленно на A , после чего можно записать

$$\frac{dA^2}{dt} + \left(2\delta_0 + \frac{1}{4}\gamma A^2\right) A^2 = 0.$$

Разделим переменные

$$\frac{dA^2}{\left(2\delta_0 + \frac{1}{4}\gamma A^2\right) A^2} = -dt.$$

Разложим дробь в левой части на простейшие

$$\frac{dA^2}{A^2} - \frac{d\left(2\delta_0 + \frac{1}{4}\gamma A^2\right)}{2\delta_0 + \frac{1}{4}\gamma A^2} = -2\delta_0 dt.$$

Интегрируя, получим

$$\ln A^2 - \ln \left(2\delta_0 + \frac{1}{4}\gamma A^2\right) = -2\delta_0 t + C,$$

или

$$\ln \left(\frac{1}{4}\gamma + 2\delta_0 \frac{1}{A^2}\right) = 2\delta_0 t - C.$$

Отсюда

$$\frac{\gamma}{8\delta_0} + \frac{1}{A^2} = \frac{1}{2\delta_0} e^{-C} e^{2\delta_0 t} = B e^{2\delta_0 t}$$

и, наконец (см. (244)),

$$A(t) = \frac{A_0}{\sqrt{1 + D e^{2\delta_0 t}}}.$$

В этом выражении $D = BA_0^2$ — постоянная интегрирования, определяемая начальными условиями. В зависимости от значения D амплитуда в первый момент может быть меньше или больше установившегося значения A_0 или даже в точности равняться ему (при $D = 0$). Величина δ_0 есть начальный результирующий показатель затухания. Она во всяком случае отрицательна — в этом состоит условие самовозбуждения. Учитывая это, мы видим, что с течением времени экспоненциальный член под

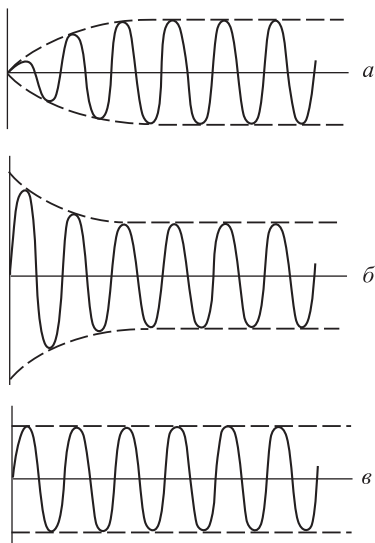


Рис. 298

радикалом убывает, и мы получаем в пределе (при $t \rightarrow \infty$) установившееся значение амплитуды, выражаемое формулой (244).

На рис. 298 изображен процесс установления колебаний для различных начальных условий. Физический смысл тех и иных начальных условий легче раскрыть, опираясь на графическое представление процесса установления, как это будет сделано в § 88.

§ 86. Генератор с инерционной нелинейностью

В генераторе, установившийся режим которого рассмотрен в § 84, рост амплитуды автоколебаний ограничивается нелинейностью характеристики триода $i_a = f(u_c)$. Эта характеристика выражает зависимость между мгновенными значениями анодного тока и напряжения на сетке. Поэтому даже при чисто синусоидальном напряжении на сетке переменная составляющая анодного тока *принципиально* не синусоидальна. Напряжение на контуре (а следовательно, и на сетке) практически синусоидально только за счет фильтрующего действия самого контура. Переход к квазилинейной форме уравнения состоит в том, что мы составляем уравнение для первых гармоник тока и напряжения и вводим среднюю крутизну, определяемую как отношение амплитуд первых гармоник тока и напряжения.

Нелинейность в автоколебательной системе обязательна. Возможно, однако, построить генератор на другой основе, введя в состав автоколебательной системы нелинейный параметр, значение которого зависит не от мгновенных, а от действующих (или амплитудных) значений тока или напряжения. Такого рода нелинейность называется *инерционной* (см. § 72).

Примером устройства, обладающего инерционной нелинейностью, может служить *термистор* (термосопротивление). Сопротивление термистора зависит от его температуры, а температура определяется балансом мощностей — мощности, выделяемой в термисторе в форме джоулева тепла, и мощности, отдаваемой термистором в окружающую среду через посредство конвекции и излучения. Скорость процесса теплообмена между термистором и окружающей средой характеризуется постоянной времени, которая тем больше, чем больше теплоемкость термистора и чем меньше теплопроводность в окружающую среду. Постоянная времени имеет порядок единиц и даже десятков секунд, т. е. на много порядков превосходит период высокочастотного колебания. Поэтому, если нагрев термистора производится током высокой частоты, то температура термистора на протяжении периода не может следовать за изменениями мгновенной мощности и остается постоянной с высокой степенью точности. Таким образом, можно принять, что сопротивление термистора есть функция *амплитуды* (а не мгновенного значения) тока или напряжения, причем важно, что это предположение может быть выполнено с любой степенью точности (так как технически легко увеличить постоянную времени до весьма значительных величин). А это, в свою очередь, означает, что получаемые автоколебания синусоидальны со столь же высокой степенью точности. Это является существенной особенностью систем с инерционной нелинейностью.

Рассмотрим теперь генератор, отличающийся от ранее рассмотренного тем, что триод предполагается идеально линейным, т. е. характеристика его предполагается линейной с постоянной крутизной S_0 . Необходимая для ограничения амплитуды нелинейность вводится в си-

стему путем включения в контур в качестве активного сопротивления термистора с зависящим от температуры, а следовательно, от амплитуды колебания сопротивлением R_T . Термисторы бывают металлические и полупроводниковые; у первых сопротивление с температурой растет, у вторых — падает; первые нужно включать в контур последовательно, вторые — параллельно. Будем в обоих случаях считать, что сопротивление R_T выражает эквивалентное последовательное сопротивление контура. В таком случае можно воспользоваться комплексной формой квазилинейного уравнения из § 84, переписав его в виде

$$\omega_0^2 - \omega^2 + i \omega \frac{1}{L} \left(R_T - \frac{MS_0}{C} \right) = 0.$$

Для генератора с термистором, описываемого этим соотношением, частота колебаний ω равна ω_0 при любой амплитуде (а следовательно, не меняется в процессе установления колебаний). Условие самовозбуждения запишется в виде

$$R_T(0) < \frac{MS_0}{C},$$

где $R_T(0)$ — сопротивление термистора при отсутствии колебаний, определяемое только температурой окружающей среды. Амплитуда в установившемся режиме находится из уравнения

$$R_T(U_m) = \frac{MS_0}{C}.$$

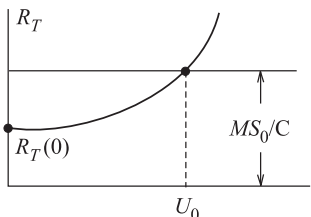


Рис. 299

Характеристика термистора $R_T(U_m)$ должна, разумеется, быть известна. Уравнение можно решить графически, как показано на рис. 299. Возможность получения с помощью инерционной нелинейности синусоидальных установившихся автоколебаний имеет большое значение. Во-первых, во многих случаях важно иметь хорошую форму колебаний, свободную от гармоник. Во-вторых, наличие гармоник вызывает смещение частоты. Этот эффект тем заметнее, чем меньше добротность контура: применение инерционной нелинейности позволяет получить синусоидальные колебания и в случае контура с невысокой добротностью. Последнее особенно важно для так называемых RC -генераторов (см. § 91), цепи которых, составленные только из емкостей и сопротивлений, эквивалентны апериодическим контурам.

§ 87. Фазовая плоскость

Обычное представление колебательных явлений в какой-либо системе состоит в том, что все изменяющиеся во времени величины (например, токи, напряжения) выражаются соответствующими функциями времени. Графики этих функций — это *осциллограммы*.

Но можно поступить и иначе, а именно: взять в качестве прямоугольных координат мгновенные значения переменных величин, совокупность которых полностью определяет состояние системы. Получим таким образом некоторое пространство, каждая точка которого представляет состояние системы в данный момент времени. Такое пространство называется *фазовым*.

Число измерений фазового пространства определяется порядком дифференциального уравнения системы. Известно, что для решения уравнения n -го порядка нужно задать в качестве начальных условий значения для начального момента функции и ее производных до $(n - 1)$ -го порядка, т.е. всего n независимых величин. Нам придется иметь дело преимущественно с уравнениями второго порядка. В этом случае число измерений фазового пространства равно двум, т.е. фазовое пространство в этом случае есть *фазовая плоскость*. Колебательные явления в системе, описываемой уравнением второго порядка, могут быть представлены диаграммами на фазовой плоскости.

Рассмотрим, к примеру, свободные колебания в линейном контуре без потерь. Дифференциальное уравнение для этого случая имеет вид

$$L \frac{di}{dt} + \frac{1}{C} \int i dt = 0$$

или

$$\frac{d^2 q}{dt^2} + \omega_0^2 q = 0. \quad (246)$$

Решение этого уравнения таково

$$q = M \sin \omega_0 t + N \cos \omega_0 t, \quad (247)$$

где M и N — произвольные постоянные, определяемые начальными условиями. Для тока имеем

$$i = \frac{dq}{dt} = M\omega_0 \cos \omega_0 t - N\omega_0 \sin \omega_0 t. \quad (248)$$

Возьмем ток и заряд в качестве координат фазовой плоскости и выразим одну величину непосредственно через другую, т.е. исключим *параметр* t . Для этого разделим (248) на ω_0 , возведем в квадрат и сложим с квадратом (247). Получим уравнение

$$\frac{1}{\omega_0^2} i^2 + q^2 = M^2 + N^2,$$

которое есть уравнение семейства эллипсов (рис. 300). Заметим, что, изменив масштаб, например отложив по оси абсцисс $\omega_0 q$ вместо q , получим семейство окружностей. Каждая точка на одном из эллипсов изображает состояние системы в данный момент (так как координаты точки суть мгновенные значения тока и заряда). Поэтому точка на диаграмме называется *изображающей* точкой. С течением времени эта точка перемещается, описывая линию, называемую *фазовой траекто-*

рией. Если на фазовой плоскости изображен периодический процесс, то фазовая траектория будет представлять собой замкнутый контур, обегаемый изображающей точкой за время, равное одному периоду. Семейство фазовых траекторий, отображающее возможные в данной системе колебательные явления, называют «фазовым портретом» данной системы. Так, рис. 300 является фазовым портретом линейного контура без потерь.

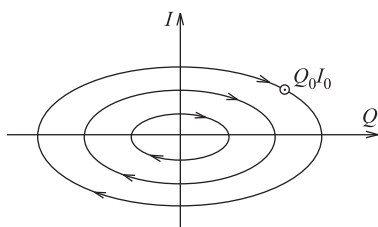


Рис. 300

На рис. 300 выбраны координаты q и $i = dq/dt$. При таком выборе координат изображающая точка движется по часовой стрелке, как отмечено на рис. 300.

Мы получили фазовый портрет на рис. 300 аналитическим путем. Покажем, как производится построение фазового портрета по осциллограммам.

На рис. 301, а изображены осциллограммы заряда и тока. Если теперь в моменты, отмеченные и занумерованные на рисунке, взять отсчеты мгновенных значений q и i (т.е. ординаты соответствующих кривых) и перенести их в качестве координат изображающей точки на рис. 301, б, то можно построить фазовую траекторию. Построение подтверждает, что изображающая точка движется по часовой стрелке.

Поясним теперь роль начальных условий. Начальными условиями для уравнения второго порядка вида (246) являются, как известно, значения функции (q) и ее первой производной ($i = dq/dt$) в момент, принимаемый за начало отсчета времен (в частности, в момент $t = 0$). Следовательно, задание начальных условий определяет исходные координаты изображающей точки. На рис. 300 свободное колебание представляется эллипсом, проходящим через точку, отображающую своими координатами Q_0 и I_0 начальные условия.

Рассмотрим контур с потерями. Уравнение такого контура имеет вид

$$\frac{d^2 q}{dt^2} + 2\alpha \frac{dq}{dt} + \omega_0^2 q = 0. \quad (249)$$

Решение этого уравнения при начальных условиях:

$$q(0) = 0, \quad i(0) = I_0$$

таково:

$$q = \frac{I_0}{\omega_1} e^{-\alpha t} \sin \omega_1 t, \quad (250)$$

где $\omega_1 = \sqrt{\omega_0^2 - \alpha^2}$ — собственная частота. Найдем ток:

$$i = \frac{I_0}{\omega_1} e^{-\alpha t} (-\alpha \sin \omega_1 t + \omega_1 \cos \omega_1 t).$$

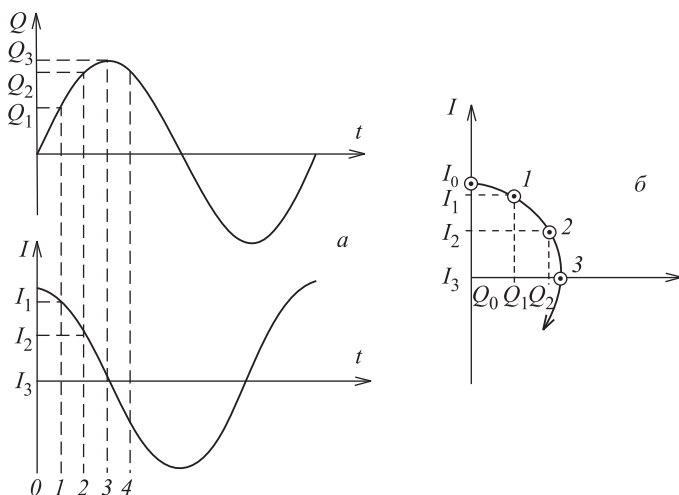


Рис. 301

При малом показателе затухания α можно пренебречь первым членом в скобках и принять

$$i \approx I_0 e^{-\alpha t} \cos \omega_1 t. \quad (251)$$

Для исключения времени t из (250) и (251) представим зависимость между током и зарядом в полярных координатах. Запишем:

$$\begin{aligned} i &= I_0 e^{-\alpha t} \cos \omega_1 t = \rho \cos \vartheta, \\ \omega_1 q &= I_0 e^{-\alpha t} \sin \omega_1 t = \rho \sin \vartheta, \end{aligned}$$

т. е. положим

$$\rho = I_0 e^{-\alpha t}, \quad \vartheta = \omega_1 t,$$

откуда

$$\rho = I_0 \exp \left(-\frac{\alpha}{\omega_1} \vartheta \right). \quad (252)$$

Это и есть уравнение фазовой траектории в полярных координатах, представляющее свертывающуюся логарифмическую спираль (рис. 302, а). При других начальных условиях получим другие спирали. Все они будут вложены друг в друга и все будут сходиться в точке $i = q = 0$. Каждая спираль изображает затухающее колебание. Семейство же спиралей с различными начальными условиями образует фазовый портрет контура с потерями.

В заключение данного параграфа рассмотрим еще случай контура, в котором действием обратной связи затухание сделано отрицательным. Подобная система описывается уравнением, отличающимся от (249) только знаком при члене с первой производной заряда по времени. Решение уравнения представится формулами (250) и (251),

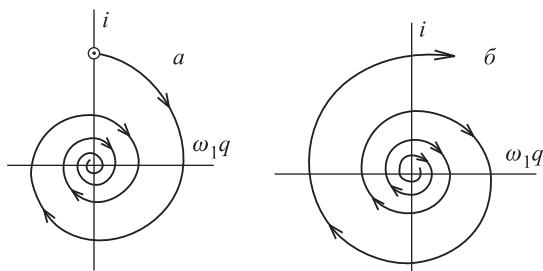


Рис. 302

а уравнение фазовой траектории — формулой (252), причем во всех трех формулах надо лишь заменить знак $(-)$ на $(+)$. Это означает, что мы имеем колебания с экспоненциально нарастающей амплитудой, а фазовая траектория представляет собой раскручивающую спираль (рис. 302, б).

Если бы система действительно была линейной, то амплитуда нарастала бы неограниченно; этому соответствует неограниченный рост радиуса спирали.

§ 88. Фазовые портреты генераторов

В предыдущем параграфе рассмотрены в качестве примеров линейные системы. Между тем нас интересуют явления в генераторах, являющихся заведомо нелинейными системами, так как в линейной системе автоколебания с установившейся амплитудой, как уже известно, невозможны.

Очередная задача состоит в построении фазовых портретов генераторов. Начнем со случая мягко возбуждающегося генератора, как более простого. Фазовый портрет должен дать полную картину возможных режимов генератора: установившегося режима и процессов установления колебаний при любых начальных условиях.

Не будем выводить уравнений фазовых траекторий; ограничимся качественным рассуждением. Для этого нужно прежде всего пояснить, как отображается на фазовых диаграммах то или иное значение затухания.

Для простого контура уравнение спирали имеет вид

$$\rho = I_0 \exp \left(-\frac{\alpha}{\omega_1} \vartheta \right).$$

Возьмем два значения радиуса ρ , соответствующих изменению ϑ на 2π , т. е. радиусы двух соседних витков спирали (рис. 303):

$$\rho_1 = I_0 \exp \left(-\frac{\alpha}{\omega_1} \vartheta \right), \quad \rho_2 = I_0 \exp \left(-\frac{\alpha}{\omega_1} (\vartheta + 2\pi) \right)$$

и составим отношение этих радиусов

$$\frac{\rho_1}{\rho_2} = \exp\left(\frac{\alpha}{\omega_1} 2\pi\right) = e^{\alpha T} \approx e^{\pi d},$$

где d — затухание контура. Таким образом, радиус спирали изменяется от витка к витку в геометрической прогрессии¹⁾ знаменатель которой зависит от затухания. То же относится и к раскручивающейся спирали.

Рассмотрим теперь явления в мягко возбуждающемся генераторе. В начальной стадии самовозбуждения, когда колебания еще малы, явления могут с удовлетворительной точностью описываться линейным уравнением (229). Если условие самовозбуждения выполнено, то показатель затухания отрицателен и процесс самовозбуждения представляется раскручивающейся спиралью (рис. 302). Однако по мере роста амплитуды колебаний начинает сказываться нелинейность. Ее влияние проявляется в том, что затухание становится меньше по абсолютной величине; радиус спирали замедляет свой рост. Наконец, достигается энергетический баланс. При этом затухание становится равным нулю, дальнейший рост амплитуды прекращается. Изображающая точка описывает замкнутый контур в виде эллипса (или окружности при соответствующем выборе масштабов). Это и есть установившийся режим генератора.

Если задать генератору начальные условия, представляемые некоторой точкой *вне* контура, соответствующего установившемуся режиму, то затухание будет положительным и изображающая точка будет перемещаться по свертывающейся спирали, радиус которой будет убывать все медленнее, пока траектория не вернется на тот же контур установившегося режима. Все это изображено на рис. 304, представляющем фазовый портрет генератора с мягким возбуждением. Замкнутый контур, отображающий установившееся периодическое колебание (жирная линия на рис. 304), называется в теории колебаний *предельным циклом*. Для фазового портрета на рис. 304 характерно, что любая фазовая траектория, где бы она не начиналась, навертывается в конце концов на предельный цикл. Это означает, что всякое нарушение установившегося режима выравнивается автоматически после устранения вызвавшей это нарушение причины. При таких условиях предельный

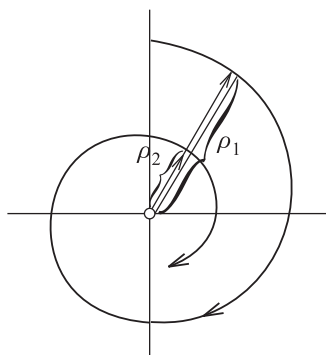


Рис. 303

¹⁾ Свойство логарифмической спирали. Для так называемой архимедовой спирали радиус изменяется по арифметической прогрессии.

цикл называется *устойчивым*. Легко видеть теперь, что три случая установления колебаний, изображенные на рис. 298, соответствуют на рис. 304 положениям исходной точки фазовой траектории: *a* — в начале координат, *в* — на предельном цикле и *б* — вне его.

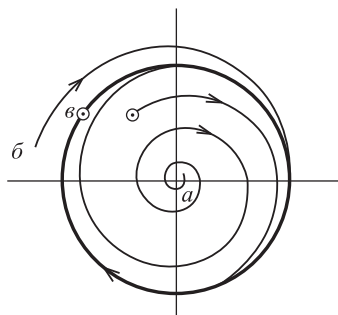


Рис. 304

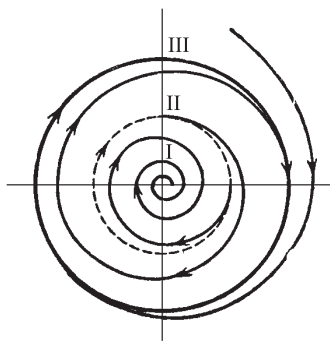


Рис. 305

Рассмотрим генератор с жестким возбуждением. Этот случай отличается от только что рассмотренного тем, что при малых амплитудах затухание положительно и фазовые траектории сворачиваются к началу координат. Существует значение амплитуды, при котором затухание первый раз обращается в нуль. Этому значению амплитуды соответствует свой предельный цикл. Однако цикл этот неустойчив. Фазовые траектории не наворачиваются на него, а наоборот, сходя с предельного цикла, либо свертываются во внутрь к началу координат, либо развертываются в области вне неустойчивого предельного цикла. В этой области затухание отрицательно, а амплитуда растет, пока затухание снова не обратится в нуль. Здесь образуется второй предельный цикл, на этот раз уже устойчивый. Отличие устойчивого предельного цикла от неустойчивого с точки зрения фазовых диаграмм состоит в том, что фазовые траектории стремятся к устойчивому циклу и сбегают с неустойчивого. Все сказанное поясняется фазовым портретом генератора с жестким возбуждением, изображенного на рис. 305. Фазовая плоскость может быть разбита на три области: I — внутри неустойчивого цикла, II — между устойчивым и неустойчивым циклами и III — вне устойчивого цикла. Автоколебания не возбуждятся, если начальные условия изображаются точкой в области I, и возбуждятся, если исходная точка лежит в области II или III. Разница между двумя последними случаями будет состоять в том, что в первом амплитуда будет устанавливаться, возрастая до установившегося значения, а во втором, — убывая до того же значения. Различие между мягким и жестким возбуждением генератора, представляемое рис. 304 и 305, ранее было пояснено рис. 295 и 296 и рис. 281 и 282. Полезно сравнить между собой все три пары рисунков, так как они по-разному отображают одни и те же соотношения.

§ 89. Графические методы

Фазовые диаграммы позволяют просто и наглядно представить в графической форме различные колебательные явления, а также свойства тех или иных систем (пассивных цепей, автоколебательных систем).

Но роль фазовых диаграмм этим не ограничивается. Построения на фазовой плоскости служат не только в качестве иллюстрации — они составляют также основу графических методов решения нелинейных уравнений. При этом существенно, что графическое решение не требует упрощающих предположений, например предположения малой нелинейности, синусоидального решения и т. п.

Когда говорится о решении нелинейной задачи путем графического построения на фазовой плоскости, то имеется в виду, что решение получается в виде фазового портрета. Фазовый портрет дает все необходимые сведения о поведении системы, описываемой данным нелинейным уравнением: на нем видны все возможные режимы при любых начальных условиях; в частности, видны предельные циклы, если они существуют; по нему можно судить об устойчивости предельных циклов; форма предельных циклов выражает форму колебаний и т. д.

Одним из наиболее универсальных графических методов является метод *построения поля направлений*. Он позволяет построить фазовый портрет системы, описываемой нелинейным дифференциальным уравнением второго порядка следующего вида:

$$q'' + \varphi(q') + \psi(q) = 0. \quad (253)$$

Введем

$$i = \frac{dq}{dt}.$$

Уравнение (253) перепишется в виде

$$\frac{di}{dt} + \varphi(i) + \psi(q) = 0. \quad (254)$$

Разделим (254) на i

$$\frac{di}{dq} = -\frac{\varphi(i) + \psi(q)}{i}. \quad (255)$$

В этом соотношении $\frac{di}{dq}$ означает наклон касательной к фазовой траектории в данной точке, т. е. в точке с координатами q и i . Беря любую точку, можно, зная ее координаты, вычислить правую часть (255) и определить таким образом *направление* фазовой траектории в данной точке. Вычислив значения $\frac{di}{dq}$ для многих точек, можно построить на фазовой плоскости так называемое *поле направлений*, как показано на рис. 306. Ориентируясь на стрелочки, показывающие

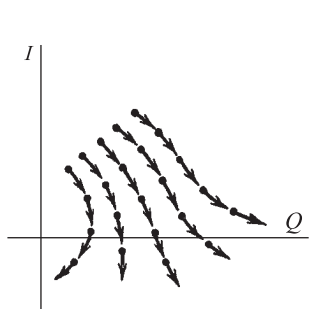


Рис. 306

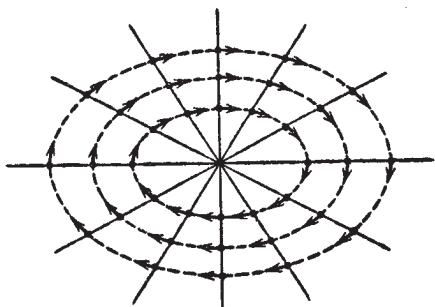


Рис. 307

направление фазовой траектории в данной точке, можно построить семейство траекторий, т.е. фазовый портрет, с любой желаемой степенью точности, которая, очевидно, повышается с увеличением числа точек. Удобно вести построение, опираясь на линии, вдоль которых направление траекторий (т.е. наклон касательной к траектории) остается неизменным. Эти линии называются *изоклинами* (линиями равных наклонов). Уравнение изоклин имеет вид

$$\frac{\varphi(i) + \psi(q)}{i} = \text{const.}$$

Поясним это на простейшем примере линейной системы, описываемой уравнением

$$\frac{di}{dt} + \omega_0^2 q = 0.$$

В этом случае:

$$\varphi(i) = 0, \quad \psi(q) = \omega_0^2 q,$$

и уравнение изоклин имеет вид

$$\frac{q}{i} = \text{const},$$

что представляет собой уравнение пучка прямых, проходящих через начало координат. Построение соответствующего фазового портрета показано на рис. 307 — это портрет контура без потерь.

В ряде случаев можно ограничиться уравнением менее общего вида, чем (253),

$$q'' + \varphi(q') + q = 0. \quad (256)$$

Именно к такой форме можно привести уравнение большинства генераторов, с которыми мы имеем дело в радиотехнике. Графическое решение уравнения (256) может быть выполнено без вспомогательных вычислений путем чисто графического построения, известного под на-

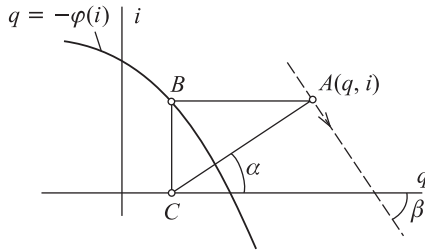


Рис. 308

званием *построения Льенара*. Обоснование этого построения состоит в следующем. Вместо (255) имеем

$$\frac{di}{dq} = -\frac{\varphi(i) + q}{i}.$$

Построим на фазовой плоскости кривую

$$q = -\varphi(i) \quad (257)$$

(рис. 308). Из произвольной рис. 308 точки A с координатами (q, i) проведем параллельно оси q прямую до пересечения с кривой. Из точки пересечения B опускаем перпендикуляр на ось q . Основание перпендикуляра обозначим через C . Имеем:

$$\overline{BA} = q + \varphi(i), \quad \overline{BC} = i, \quad \frac{\overline{BC}}{\overline{BA}} = \frac{i}{q + \varphi(i)}.$$

Но последнее отношение есть не что иное, как $\operatorname{tg} \alpha$ — тангенс угла, образуемого прямой CA с осью q . Для построения же прямой, тангенс угла наклона которой к оси q равен

$$\operatorname{tg} \beta = \frac{di}{dq} = -\frac{\varphi(i) + q}{i} = -\operatorname{ctg} \alpha,$$

нужно построить перпендикуляр к CA в точке A . Этот перпендикуляр и дает направление фазовой траектории в точке A . Итак, построение Льенара сводится к следующему: 1) после приведения уравнения к виду (256) строится характеристическая кривая (257); 2) строится прямоугольный треугольник ABC ; 3) строится перпендикуляр к гипотенузе AC , который и дает искомое направление траектории в точке A .

Покажем построение Льенара на примере лампового генератора, нелинейное уравнение которого имеет вид (см. § 85)

$$u'' + \frac{1}{L} \left(R - \frac{MS}{C} \right) u' + \omega_0^2 u = 0$$

или

$$u'' + f(u)u' + \omega_0^2 u = 0 \quad (258)$$

(так как крутизна S зависит от напряжения u). Для приведения (258) к виду (256) введем прежде всего безразмерное («собственное») время

$$\tau = \omega_0 t.$$

Тогда:

$$\frac{du}{dt} = \frac{du}{d\tau} \frac{d\tau}{dt} = \omega_0 \frac{du}{d\tau}; \quad \frac{d^2 u}{dt^2} = \omega_0^2 \frac{d^2 u}{d\tau^2},$$

и мы получим

$$u'' + \frac{1}{\omega_0} f(u) u' + u = 0. \quad (259)$$

Здесь штрихи означают дифференцирование по τ . Теперь введем новую переменную

$$s = \int u d\tau \quad (260)$$

и определим некоторую функцию $F(x') = F(u)$ соотношением

$$F(x') = \frac{1}{\omega_0} \int f(u) du.$$

Проинтегрируем (259) по τ

$$\int u'' d\tau + \frac{1}{\omega_0} \int f(u) u' d\tau + \int u d\tau = u' + \frac{1}{\omega_0} \int f(u) du + \int u d\tau.$$

Заменив u на x' на основании (260), получим

$$x'' + F(x') + x = 0,$$

т. е. уравнение, в точности совпадающее с (256). Координатами фазовой плоскости будут теперь служить величины x и $x' = u$, т. е. напряжение и интеграл от напряжения. Построим характеристическую кривую

$$x = -F(u).$$

Мы имеем

$$\begin{aligned} F(u) &= \frac{1}{\omega_0} \int f(u) du = \frac{1}{\omega_0 L} \int \left(R - \frac{MS}{C} \right) du = \\ &= \frac{1}{\omega_0 L} \left[Ru - \frac{M}{C} \int S(u) du \right]. \end{aligned} \quad (261)$$

Рассмотрим случай мягкого возбуждения, для которого

$$S(u) = a_1 + 3a_3 u^2 = S_0 + 3a_3 u^2 \quad (262)$$

(см. § 85). Подставляя (262) в (261), находим

$$F(u) = \frac{1}{\omega_0 L} \left[\left(R - \frac{MS_0}{C} \right) u - \frac{M}{C} a_3 u^3 \right].$$

Кривая

$$x = -F(u)$$

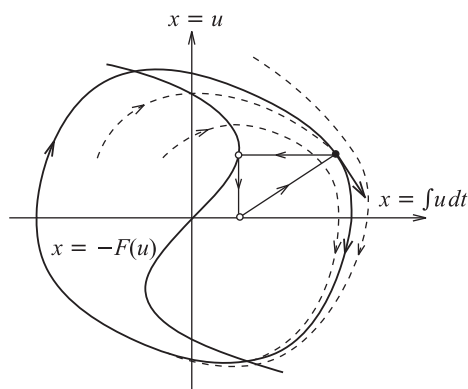


Рис. 309

есть, таким образом, кубическая парабола, график которой нанесен на рис. 309. Дальнейшее построение производится графическим способом, как показано на том же рисунке. Мы получаем уже известный фазовый портрет мягкого возбуждения генератора. Однако следует заметить, что мы узнали нечто большее, чем то, что дала квазилинейная теория: мы получили *форму* колебания, которая в нелинейной системе должна, конечно, отличаться от синусоидальной. На портрете рис. 309 это находит свое отражение в том, что предельный цикл отличается от эллипса.

§ 90. Некоторые схемы генераторов

Говоря о ламповом генераторе, мы до сих пор пользовались только одной классической схемой с контуром в цепи сетки и с индуктивной обратной связью. Это делалось для того, чтобы продемонстрировать различные подходы и методы на одном и том же объекте исследования.

Рассмотрим теперь некоторые другие схемы генераторов и воспользуемся случаем для разъяснения одного принципиального вопроса, относящегося к обратной связи. Схема генератора с контуром в цепи сетки повторена на рис. 310; на рис. 311 изображена схема генератора с контуром в цепи анода. Не отличается принципиально от обеих этих схем с автотрансформаторной связью схема, показанная на рис. 312; она часто называется трехточечной.

Схема с автотрансформаторной связью рис. 312 является частным случаем общей трехточечной схемы рис. 313. Эта схема содержит сложный контур из трех реактивных сопротивлений: X_1 , X_2 и X_3 . Частота генерируемых колебаний определяется из условия резонанса

$$X_1 + X_2 + X_3 = 0. \quad (263)$$

Найдем условие самовозбуждения для схемы рис. 313. Коэффициент усиления

$$K = -\mu$$

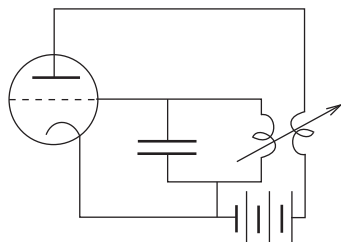


Рис. 310

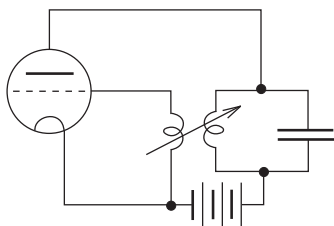


Рис. 311

(если пренебречь внутренним сопротивлением триода). Коэффициент обратной связи

$$\beta = \frac{X_2}{X_1 + X_2} = \frac{1}{1 + \frac{X_1}{X_2}}.$$

Так как K и β — вещественны, то для получения надлежащей фазы обратной связи нужно, чтобы коэффициент обратной связи был отрицателен, а для этого достаточно выполнить два условия:

$$\frac{X_1}{X_2} < 0, \quad (264)$$

$$\left| \frac{X_1}{X_2} \right| > 1. \quad (265)$$

Условия (263), (264) и (265) выполняются в следующих двух случаях:

$$\begin{aligned} X_1 < 0, \quad X_2 > 0, \quad X_3 > 0, \\ X_1 > 0, \quad X_2 < 0, \quad X_3 < 0 \end{aligned}$$

и схема рис. 313 принимает вид рис. 314, а и б соответственно, т. е. для самовозбуждения необходимо, чтобы X_2 и X_3 имели одинаковый знак, а X_1 — противоположный ему.

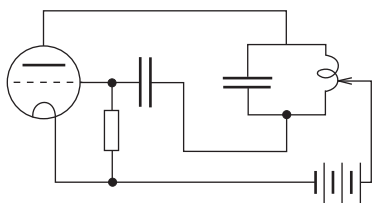


Рис. 312

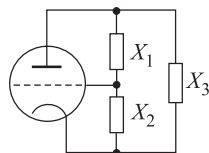


Рис. 313

Можно получить из схемы рис. 313 несколько более общую схему, включив в качестве X_1, X_2 и X_3 не индуктивности или емкости, а контуры, как показано на рис. 315. Реактивное сопротивление каждого из контуров имеет индуктивный характер (т.е. положительно) для

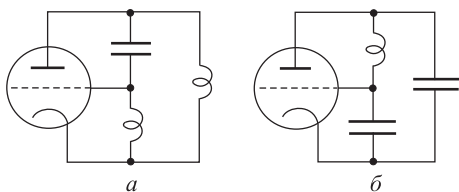


Рис. 314

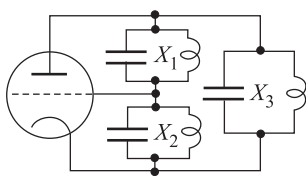


Рис. 315

частот, ниже резонансной и емкостный характер (т.е. отрицательно) для частот выше резонансной. Обозначая резонансные частоты через ω_1, ω_2 и ω_3 , можем записать два варианта условий, при которых схема рис. 315 будет генерировать колебания с частотой ω :

$$\begin{aligned} \omega_1 < \omega, \quad \omega_2 > \omega, \quad \omega_3 > \omega, \\ \omega_1 > \omega, \quad \omega_2 < \omega, \quad \omega_3 < \omega. \end{aligned}$$

Частота ω_1 может быть и больше и меньше ω_3 . Существенно лишь, чтобы частота ω_1 была либо наивысшей, либо наинизшей из трех. Значение генерируемой частоты ω находится из соотношения (263).

Мы интересовались до сих пор только фазовыми соотношениями. Вернемся к общему условию самовозбуждения

$$K\beta > 1.$$

Найдем, полагая, что условие (264) выполнено,

$$\frac{-\mu}{1 - \left| \frac{X_1}{X_2} \right|} > 1$$

или

$$\left| \frac{X_1}{X_2} \right| < \mu + 1.$$

С учетом (265) получаем окончательно

$$1 < \left| \frac{X_1}{X_2} \right| < \mu + 1.$$

Это последнее условие обеспечивает выполнение требуемых для самовозбуждения соотношений как по фазе, так и по амплитуде.

Схемы рис. 310–315 обладают тем общим признаком, что в них имеется явная цепь обратной связи, причем глубина обратной связи может регулироваться изменением того или иного параметра схемы — коэффициента связи в схемах рис. 310 и 311, коэффициента трансформации в схеме рис. 312, емкости сетка–анод в схеме рис. 314, а. Но в некоторых генераторах обратная связь действует по скрытым, неявным каналам, которые не могут быть изображены на схеме в виде отдельной цепи обратной связи. Но обратная связь существует и в та-

ких генераторах потому, что, как было показано в § 79, она является необходимым звеном всякой автоколебательной системы. Для полного понимания действия генераторов с неявной обратной связью раскроем ее механизм.

В качестве примеров генераторов с неявной обратной связью рассмотрим вкратце *динатронный* и *транзитронный* генераторы.

Действие динатронного генератора основано на так называемом *динатронном эффекте*, состоящем в том, что при попадании электронов на анод возникает вторичная эмиссия; электрон, прилетающий на анод из катода, выбивает из анода новые электроны. Если на один первичный электрон приходится несколько вторичных, то составляющая тока, обусловленная вторичными электронами, преобладает над током первичной эмиссии. В результате этого анодный ток не только убывает, но может даже изменить знак. Очевидно, что каждый первичный электрон может выбить тем больше вторичных, чем больше скорость первичного электрона. Скорость же электрона зависит от ускоряющего напряжения. В качестве динатрона может быть использован любой подходящий триод.

Следует заметить, что электрод, называемый в триоде анодом, при использовании триода в качестве динатрона не является уже анодом; назначение этого электрода состоит в испускании вторичных электронов; чтобы подчеркнуть это обстоятельство, анод триода, работающего в динатронном режиме, называют иногда *динодом*. На сетку триода в динатронном режиме задается высокий положительный потенциал, так что в динатроне она является анодом. Схема динатронного генератора на триоде изображена на рис. 316.

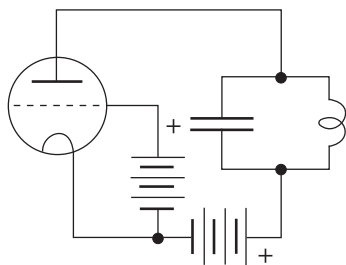


Рис. 316

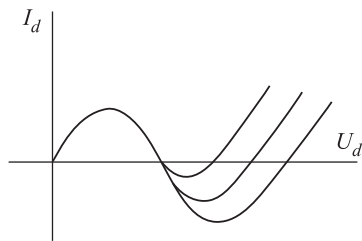


Рис. 317

Характеристики динатрона, т.е. зависимость тока I_d в цепи динода от напряжения U_d на нем, показаны на рис. 317. Три кривые рис. 317 относятся к разным напряжениям на аноде; чем выше анодное напряжение, тем правее располагается характеристика. Как видим, вольтамперная характеристика динатрона имеет падающий участок, крутизна которого зависит от тока эмиссии. В триоде ток эмиссии можно менять только путем изменения накала. Поэтому в качестве

динатрона предпочитают применять тетрод, в котором током эмиссии можно управлять посредством потенциала на первой сетке.

Механизм обратной связи в динатронном генераторе крайне прост: он состоит в том, что напряжение на контуре, будучи подано непосредственно на динод, управляет током динода. Таким образом, обратная связь действует по той же цепи, по которой осуществляется питание контура. Надлежащие фазовые соотношения обеспечиваются работой на падающем участке характеристики.

Рассмотрим транзитронный генератор. Тетрод с ограниченной полной эмиссией (т. е., например, с вольфрамовым катодом) включен по схеме рис. 318. На вторую сетку подано отрицательное напряжение, и ток в ней отсутствует. Ток первой сетки I_1 в сумме с анодным током I_a должен равняться току эмиссии. Таким образом, любое изменение тока I_1 может происходить только за счет перераспределения тока эмиссии между первой сеткой и анодом. При определенных соотношениях увеличение положительного напряжения на первой, а следовательно, и на второй сетках (они соединены между собой через батарею) может вызвать увеличение анодного тока, а следовательно, уменьшение тока I_1 . Таким образом, характеристика зависимости тока первой сетки от напряжения на ней может иметь падающий участок. Обратная связь в транзитронном генераторе состоит в том, что напряжение на контуре управляет перераспределением токов. Обратная связь действует одновременно по двум цепям, а именно, по цепям обеих сеток.

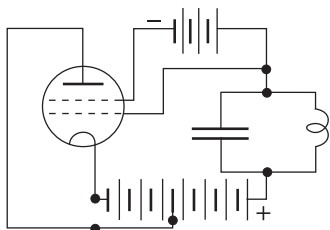


Рис. 318

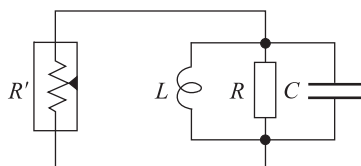


Рис. 319

Практические схемы транзитронных генераторов собираются на пентодах или гексодах, в которых первая сетка служит для ограничения полного тока эмиссии.

Нужно признать, что обычная трактовка генератора с помощью понятия обратной связи оказывается неудобной в применении к динатронному и транзитронному генераторам и вообще к генераторам, не имеющим явных и самостоятельных внешних цепей обратной связи. Для этого класса генераторов более удобна другая трактовка, основанная на применении понятия отрицательного сопротивления.

Если генератор содержит задающий колебательный контур, то вся остальная (кроме контура) схема, содержащая усилитель, те или иные

обратные связи (явные и скрытые), источники питания и прочее, может быть представлена как некоторое отрицательное сопротивление и обобщенная схема генератора может быть изображена, как показано на рис. 319.

Условие самовозбуждения записывается при этом просто как

$$R' > R.$$

Легко видеть, что схемы динактронного (рис. 316) и транзитронного (рис. 318) генераторов непосредственно приводятся к схеме рис. 319. Как динактрон, так и транзитрон являются отрицательными сопротивлениями, так как используются падающие участки их характеристик, на которых положительным приращениям напряжения соответствуют отрицательные приращения тока (а такое соотношение и является *определением* отрицательного сопротивления).

Но более того: и обычные схемы генераторов можно рассматривать с той же точки зрения. Так, например, в схемах рис. 314 можно выделить часть, представляющую собой двухполюсник с усилителем и обратной связью, обладающей свойствами отрицательного сопротивления, как было показано в § 81.

Понятие обратной связи и отрицательного сопротивления отнюдь не противопоставляются друг другу; в сущности это две формы одних и тех же соотношений.

§ 91. RC -генераторы синусоидальных колебаний

В рассмотренных до сих пор генераторах генерируемая частота определяется контуром из LC , представляющим собой колебательную систему в составе генератора. Возникает вопрос, при каких условиях возможно получить тот же результат, применяя схемы, составленные только из сопротивлений и емкостей.

Как уже известно, уравнение генератора имеет вид уравнения второго порядка, в котором действием обратной связи коэффициент при первой производной обращается в нуль. Следовательно, если составить схему из R и C , описываемую таким же уравнением, то это будет означать, что построенная схема обладает такими же свойствами, как и схема обычного генератора. (Такое заключение основывается на предположении, что составляемые дифференциальные уравнения правильно отображают реальные свойства схемы.)

Следовательно, для построения генератора RC нужно прежде всего составить такую схему из сопротивлений и емкостей, которая описывалась бы дифференциальным уравнением второго порядка, аналогичным уравнению обычного контура. Но, как известно, простая цепь, содержащая только R и C , описывается уравнением первого порядка. Стало

быть, мы должны перейти от простой цепи к сложной, обладающей по меньшей мере двумя степенями свободы ¹⁾.

Рассмотрим схему рис. 320. Уравнение, связывающее напряжение на входе и на выходе, имеет вид

$$u_2' + 2\alpha u_2 + \omega_0^2 \int u_2 dt = \frac{1}{R_1 C_2} u_1, \quad (266)$$

где:

$$\omega_0^2 = \frac{1}{R_1 R_2 C_1 C_2},$$

$$2\alpha = \frac{1}{R_1 C_1} + \frac{1}{R_2 C_2} + \frac{1}{R_1 C_2}.$$

Уравнение (266) совершенно аналогично уравнению цепи из последовательно включенных R , C и L . Если теперь ввести в схему усилитель и дать обратную связь с выхода на вход, то получится схема, способная генерировать синусоидальные (точнее, близкие к синусоидальным) колебания (рис. 321).

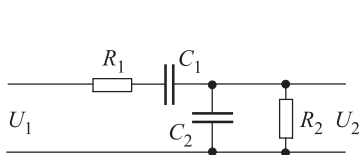


Рис. 320

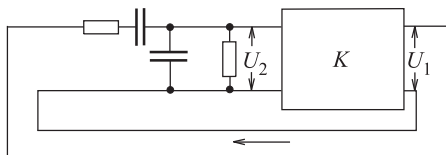


Рис. 321

Составим уравнение схемы рис. 321. Пусть коэффициент усиления усилителя есть

$$K = \frac{u_1}{u_2} \quad (267)$$

(заметим, что мы сохранили обозначения рис. 320, так что для усилителя u_1 является выходным напряжением, u_2 — входным). Выражая на основании (267) u_1 через u_2 в правой части (266) и перенося все в левую часть, получим

$$u_2'' + \left(\frac{1}{R_1 C_1} + \frac{1}{R_2 C_2} + \frac{1-K}{R_1 C_2} \right) u_2' + \frac{1}{R_1 R_2 C_1 C_2} u_2 = 0$$

— обычное уравнение генератора. Условие самовозбуждения получим, если коэффициент при первой производной

$$\frac{1}{R_1 C_1} + \frac{1}{R_2 C_2} + \frac{1-K}{R_1 C_2} < 0.$$

¹⁾ Число степеней свободы в электрической цепи определяется числом независимых токов, которые могут протекать по цепи. Число степеней свободы может также определяться минимальным числом разрывов цепи, прекращающих всякое токопротекание.

Очевидно, что для выполнения условия самовозбуждения нужно иметь достаточно большой коэффициент усиления K .

По поводу схемы рис. 321 следует еще заметить, что соотношение (267) выполняется только при четном числе ступеней усиления, так

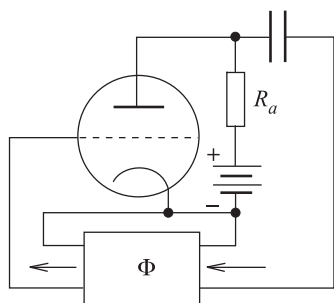


Рис. 322

как каждая ступень производит обращение фазы на 180° . Поэтому при нечетном числе ступеней получим отрицательную обратную связь вместо требуемой положительной. Стало быть, необходимое число ламп в усилителе схемы рис. 321 не менее двух, и эту схему часто называют схемой двухлампового RC -генератора.

К построению схемы RC -генератора можно подойти и иначе. Всякая схема, приводимая к скелетной схеме автоколебательной системы рис. 278, будет генерировать установившиеся по частоте и амплитуде колебания, если устойчиво выполняются условия баланса амплитуд и фаз. Первое условие определяет амплитуду, второе — частоту установившихся колебаний. Условие баланса амплитуд обеспечивается подбором надлежащего усиления. Для выполнения же условия баланса фаз необходимо, чтобы при данной частоте (т.е. при той частоте, которую генерирует или должен генерировать генератор) получалась положительная обратная связь. Так, схема рис. 322, в которой обратная связь подается на сетку триода через фазовращающий четырехполюсник Φ , может генерировать частоту, для которой фаза напряжения при прохождении через четырехполюсник сдвигается *ровно* на 180° (так как сама лампа осуществляет поворот фазы еще на 180°). Эти соображения указывают путь построения однолампового RC -генератора. В качестве фазовращателя может быть использована лестничного типа схема, состоящая из нескольких звеньев RC , как показано на рис. 323. Фаза поворачивается на некоторый угол при переходе через каждое звено. Однако в каждом звене совершается поворот фазы на угол меньше 90° . Таким образом, для получения сдвига на 180° необходима цепочка не менее чем из трех звеньев. В результате схема однолампового RC -генератора принимает вид рис. 324. Полагая

$$i_a = Su$$

и составляя уравнение схемы, можно вывести условие самовозбуждения

$$R_a S > 29 + 23 \frac{R_a}{R} + 4 \frac{R_a^2}{R^2}$$

и выражение для частоты

$$\omega_0 = \frac{1}{RC} \sqrt{\frac{R}{6R + 4R_a}}.$$

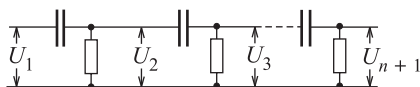


Рис. 323

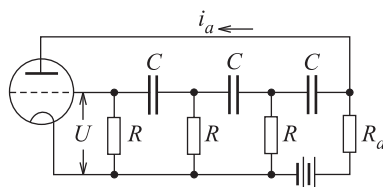


Рис. 324

Следует пояснить, что для схемы с тремя звеньями получается уравнение третьего порядка. Из трех корней характеристического уравнения один отрицателен и действителен, два же остальных комплексны; они и определяют частоту и затухание.

§ 92. Релаксационные автоколебания

Мы рассматривали до сих пор генераторы колебаний, близких по форме к синусоидальным. Такие автоколебания можно назвать *почти-гармоническими*. Но в радиотехнике применяются также автоколебания, резко отличные по форме от синусоидальных; такие автоколебания, отличающиеся иногда внезапными скачкообразными изменениями тока или напряжения, или скорости изменения этих величин, — что изображается графически осциллограммой в виде ломаной линии — называются *релаксационными* автоколебаниями. Существенное различие формы почти-гармонических и релаксационных автоколебаний обусловлено различием в устройстве и действии генераторов этих видов колебаний. В состав генератора почти гармонических колебаний входит обычно контур, составленный из L и C (или эквивалентная контуру схема из R и C), т.е. колебательная система, содержащая два (по крайней мере) реактивных элемента, способных накапливать энергию. Колебательный процесс в контуре состоит с энергетической точки зрения в том, что запасенная в контуре энергия то сосредоточивается полностью в магнитном поле индуктивности, то переходит в электрическое поле емкости. Такая перекачка энергии происходит периодически; если потери отсутствуют, то колебание совершается по строго синусоидальному закону; при наличии же потерь амплитуда убывает со временем по экспоненциальному закону в связи с тем, что запас энергии все время убывает. Если же потери энергии восполняются, как во всяком генераторе, за счет энергии источника, то колебания делаются незатухающими, а форма их лишь постольку отличается от синусоидальной, поскольку это зависит от ограничивающей амплитуду нелинейности.

Совершенно иной механизм генератора релаксационных колебаний. Он характеризуется тем, что имеется единственный (в простейшем случае) накопитель энергии. Происходящий в релаксационном генераторе периодический процесс состоит в том, что накопитель энергии наби-

рает возрастающий запас энергии, пока не будет достигнут некоторый определенный высший уровень. Затем некоторый клапан разряжает накопитель на цепь разряда, пока запас энергии не достигнет определенного низшего уровня. Здесь клапан снова срабатывает и включает накопитель на зарядку. Сказанное поясняется блок-схемой рис. 325, на которой И — источник энергии, Н — накопитель, К — клапан, ОС — обратная связь, РЦ — разрядная цепь. Из приведенного общего описания уже следует, что для релаксационного генератора характерно наличие клапана, могущего занимать два фиксированных положения, одно из которых соответствует включению накопителя на заряд, а второе — на разряд. Клапан управляется обратной связью с накопителя; действие обратной связи состоит в том, что переключение клапана производится при достижении двух фиксированных уровней энергии — высшего и низшего. Отсюда следует, что размах релаксационных колебаний не зависит от энергетического баланса, а определяется полностью перепадом между указанными двумя уровнями, при которых срабатывает клапан.

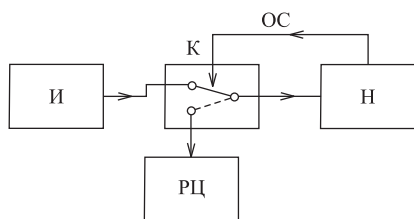


Рис. 325

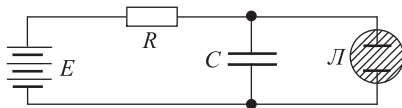


Рис. 326

Все эти общие свойства релаксационного генератора легко уяснить себе, рассмотрев простейший пример — релаксационный генератор с газоразрядной лампой. Схема генератора изображена на рис. 326. Источник энергии — батарея E — заряжает через сопротивление R конденсатор C , играющий в этой схеме роль накопителя. В качестве клапана применена газоразрядная лампа L ; она же является цепью, на которую разряжается накопитель. Действие схемы таково: после выключения батареи конденсатор начинает заряжаться. Когда напряжение достигает значения U_2 — «напряжения зажигания» лампы L — в последней произойдет ионизация заполняющего ее газа и сопротивление лампы, которое до этого было весьма большим, резко упадет до малой величины r . Конденсатор начнет быстро разряжаться на это сопротивление. Когда же напряжение упадет до значения U_1 , при котором ионизация прекращается, сопротивление лампы снова резко возрастет и возобновится процесс зарядки конденсатора. Ход явления представлен осциллограммой рис. 327. Размах колебаний зависит только от перепада напряжения, т. е. от разности $U_2 - U_1$. Частота зависит от постоянных времени цепей зарядки и разрядки, а также от питаю-

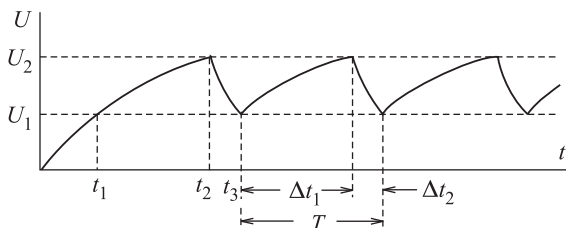


Рис. 327

щего напряжения. Определим период колебания. Он складывается из двух отрезков Δt_1 и Δt_2 (см. рис. 327), первый из которых определяет время зарядки, второй — разрядки. Для процесса зарядки имеем

$$u = E \left(1 - e^{-t/\tau_1} \right),$$

где $\tau_1 = RC$ — постоянная времени зарядной цепи. Можем составить два равенства:

$$U_1 = E \left(1 - e^{-t_1/\tau_1} \right), \quad U_2 = E \left(1 - e^{-t_2/\tau_1} \right)$$

(см. рис. 327), из которых находим:

$$t_1 = \tau_1 \ln \frac{E}{E - U_1}, \quad t_2 = \tau_1 \ln \frac{E}{E - U_2},$$

а отсюда и время зарядки

$$\Delta t_1 = t_2 - t_1 = RC \ln \frac{E - U_1}{E - U_2}.$$

Это выражение показывает, что чем больше E при данных U_1 и U_2 , тем меньше отношение под знаком логарифма отличается от единицы и тем меньше время зарядки. Процесс разрядки выражается формулой

$$u = U_2 \exp \left(-\frac{1}{\tau_2} (t - t_2) \right),$$

где $\tau_2 = rC$ — постоянная времени разрядной цепи. При $t = t_3$ достигается значение напряжения $U = U_1$. Мы получаем

$$U_1 = U_2 \exp \left(-\frac{1}{\tau_2} (t_3 - t_2) \right) = U_2 \exp \left(-\frac{\Delta t_2}{\tau_2} \right),$$

откуда

$$\Delta t_2 = rC \ln \frac{U_2}{U_1}.$$

Теперь получим период колебания, складывая Δt_1 и Δt_2 :

$$T = \Delta t_1 + \Delta t_2 = C \left(R \ln \frac{E - U_1}{E - U_2} + r \ln \frac{U_2}{U_1} \right).$$

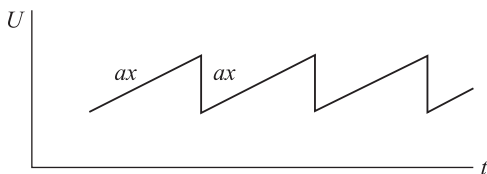


Рис. 328

Как видим, период колебания зависит от постоянных времени цепей зарядки и разрядки, т. е. от величин RC и rC , от соотношения напряжений зажигания и потухания U_2 и U_1 и от питающего напряжения E . Если взять E и R достаточно большими, а r и C достаточно малыми, то форма колебания приближается к показанной на рис. 328. Напряжение такой формы называется «пилообразным» и имеет в радиотехнике широкое применение в качестве *развертывающего* напряжения в электронной осциллокопии, в телевидении, радиолокации и др. Развертывающему напряжению предъявляют обычно требования высокой степени линейности на участке прямого хода (п. х., рис. 328) и минимальной длительности обратного хода (о. х.). Рассмотренная простейшая схема рис. 326 плохо удовлетворяет обоим этим требованиям. Линейность прямого хода можно значительно улучшить, включив вместо сопротивления R ограничитель тока, например, диод, работающий в режиме насыщения. Однако для генерирования пилообразного развертывающего напряжения предпочитают пользоваться специальными, более совершенными релаксационными генераторами.

§ 93. Релаксационные генераторы

Одной из наиболее ранних схем релаксационных генераторов является симметричный *мультивибратор*¹⁾, не утративший значения и до наших дней. Схема симметричного мультивибратора показана на рис. 329. Обе половины схемы можно рассматривать как два отдельных генератора, связанных между собой так, что один из генераторов выполняет функции клапана, управляющего работой другого.

Переходя к описанию работы схемы, начнем с такого ее состояния, когда токи в анодных цепях обеих ламп в точности равны. Это есть состояние равновесия, однако неустойчивого. В самом деле, пусть ток первой лампы немного возрастет. При этом напряжение на лампе станет меньше первоначального и конденсатор C_1 начнет разряжаться. Разрядный ток, протекая по сопротивлению R_{c2} , создаст отрицательное напряжение на сетке второй лампы. Вследствие этого напряжение на

¹⁾ Т. е. генератор колебаний со многими частотами (имеется в виду богатство спектра). Мультивибратор применялся в свое время для градуировки волномеров; опорными точками служили частоты многочисленных гармоник.

аноде возрастет, а анодный ток уменьшится. Благодаря повышению напряжения на аноде начнется подзарядка конденсатора C_2 . Зарядный ток, протекая по сопротивлению R_{c1} , вызовет положительное напряжение на сетке первой лампы, а следовательно, дальнейшее увеличение тока первой лампы.

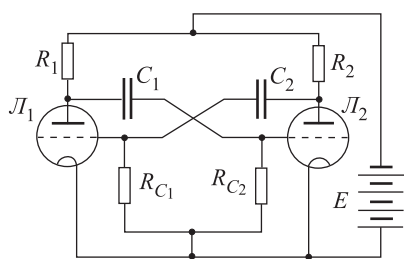


Рис. 329

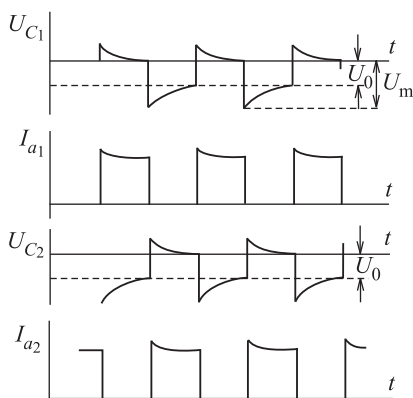


Рис. 330

Это рассуждение показывает, что рассмотренное исходное состояние, характеризующееся равенством токов и напряжений в обеих половинах схемы, действительно неустойчиво.

Рост тока первой лампы и нарастание отрицательного напряжения на сетке второй лампы происходит очень быстро — скачком. В результате первая лампа окажется насыщенной, а вторая — запертой большим отрицательным напряжением. Теперь конденсатор C_1 продолжает разряжаться через $Л_1$ и R_{c2} ; разрядный ток, а с ним и отрицательное напряжение на сетке $Л_2$ убывают, пока лампа $Л_2$ не отопрется. В этот момент напряжение на ней падает, конденсатор C_2 начинает разряжаться и происходит скачкообразное изменение токов и напряжений, описанное выше, но с переменной ролей обеих половин схемы. Изменения во времени анодных токов и напряжений на сетках в обеих половинах схемы изображены на рис. 330. Вследствие симметрии схемы длительность импульсов анодного тока в обеих половинах одинакова. Таким образом, форма анодного тока в каждой лампе приближается к прямоугольной. Впрочем, изменив параметры, можно получить также форму, приближающуюся к треугольной. Форма колебания зависит от постоянных времени разрядки и зарядки. Из схемы рис. 329 видно, что постоянная времени разрядки равна

$$\tau_p = C \left(R_c + \frac{R_i R}{R_i + R} \right),$$

а постоянная времени зарядки

$$\tau_3 = C \left(R + \frac{R_c r_c}{R_c + r_c} \right),$$

где C — емкость; R — сопротивление в цепи анода; R_i — сопротивление лампы (катод–анод); R_c — сопротивление в цепи сетки; r_c — сопротивление лампы (катод–сетка). Обычно $\tau_p \gg \tau_3$; при этом форма колебаний близка к прямоугольной. Зная постоянную времени, можно найти период колебания. За половину периода напряжение на сетке убывает от наибольшего значения U_m до напряжения запираения U_0 . Таким образом,

$$U_0 = U_m \exp \left(\frac{1}{2} \frac{T}{\tau_p} \right),$$

откуда

$$T = 2\tau_p \ln \frac{U_m}{U_0}.$$

Напряжение запираения U_0 — это параметр лампы. Что же касается U_m , то в начале разрядки, т. е. непосредственно после скачка тока, оно равно

$$U_m = E \frac{1}{1 + \frac{R_i}{R} + \frac{R_i}{R_c}}.$$

Итак, период зависит от параметров лампы, от параметров схемы и от питающего напряжения.

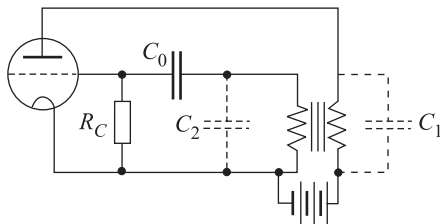


Рис. 331

В качестве второго примера релаксационного генератора рассмотрим вкратце так называемый *блочкинг-генератор*, применяемый главным образом для получения весьма коротких импульсов (длительностью до долей микросекунды). Схема блочкинг-генератора изображена на рис.331. Она состоит из триода и трансформатора, связывающего цепи анода и сетки. Емкости, образующие с индуктивностями обмоток колебательные контуры, в схеме не предусмотрены; однако необходимо учитывать распределенные емкости, эквиваленты которых намечены на схеме пунктиром.

Схему рис. 331 можно рассматривать как вырожденную схему обычного генератора с контурами в цепях анода и сетки; для схемы рис. 331

характерно, что процессы в цепях имеют аperiодический характер и перемежаются резкими скачками (эти скачки, называемые в иностранной литературе «блокинг», и обуславливают наименование рассматриваемого типа генератора).

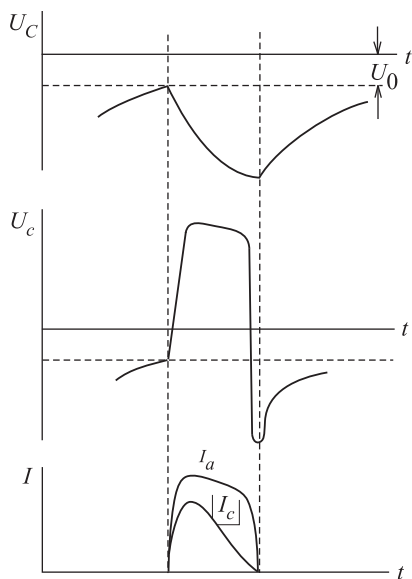


Рис. 332

Действие схемы рассмотрим с момента, когда конденсатор заряжен, на сетке — отрицательный потенциал относительно катода и лампа заперта. Конденсатор C_0 разряжается через сопротивление R_c ; постоянная времени процесса разрядки равна $R_c C_0$. Когда напряжение достигнет значения U_0 , лампа отперется; возникнет анодный ток, а благодаря сильной положительной обратной связи на сетке появится положительное напряжение, обуславливающее дальнейший быстрый рост анодного тока. Этот процесс, протекающий с очень большой скоростью, приводит к тому, что напряжение на аноде падает почти до нуля (это и есть «блокинг»). Когда результирующее напряжение на сетке (равное сумме отрицательного напряжения на C_0 и положительного напряжения, вводимого обратной связью) упадет до нуля, появится сеточный ток и конденсатор начнет быстро заряжаться через промежуток сетка–катод с постоянной времени $r_c C_0$. В этой стадии процесса напряжение на сетке положительно и превосходит напряжение на аноде, крутизна лампы при этом мала, и изменения напряжения на сетке не оказывают большого влияния на анодный ток. Но в дальнейшем, после того, как напряжение на сетке, достигнув наибольшего значения, начинает убывать, крутизна увеличивается. Нако-

нец, достигаются соотношения, благоприятные для очередного скачка: уменьшение напряжения на сетке вызывает уменьшение анодного тока, что вызывает дальнейшее уменьшение сеточного напряжения, и лампа с очень большой скоростью запирается.

После этого весь описанный цикл повторяется. На рис. 332 показано изменение во времени напряжения на конденсаторе u_C на сетке u_c и сеточного и анодного токов i_c и i_a .

Условие самовозбуждения блокинг-генератора может быть записано в виде (при условии, что C_1 и C_2 очень малы по сравнению с C_0)

$$kS_a - \frac{1}{nR_i} - \frac{n}{R_c} - nS_c + \frac{1}{r_c} > 0,$$

где S_a и S_c — соответственно крутизна по аноду и по сетке; R_i и r_c — внутренние сопротивления катод–анод и катод–сетка; n — коэффициент трансформации; k — коэффициент связи (близкий к единице в рассматриваемом случае). Минимальная длительность импульсов зависит от индуктивности рассеяния трансформатора и паразитной емкости C_1 . Период следования импульсов определяется в основном постоянной времени разрядки конденсатора C_0 ; приближенно $T \approx R_c C_0$.

Глава 10

НЕЛИНЕЙНЫЕ СИСТЕМЫ ПОД ВНЕШНИМ ВОЗДЕЙСТВИЕМ

В гл. 1 рассматривались цепи с нелинейными сопротивлениями и главнейшие радиотехнические процессы, основанные на явлениях в таких цепях. В настоящей главе рассматриваются более сложные явления, происходящие в нелинейных цепях с реактивными сопротивлениями, а также вынужденный режим некоторых автоколебательных и потенциально-автоколебательных систем. Последний термин означает систему, могущую при известных обстоятельствах генерировать колебания, но находящуюся в исходном положении (при отсутствии внешнего воздействия) в устойчивом состоянии («недовозбужденная» система).

§ 94. Работа выпрямителя

Выпрямительная схема всегда содержит того или иного устройства фильтр, служащий для ослабления переменной составляющей выпрямленного напряжения. Таким образом, в реальной выпрямительной схеме должны содержаться реактивные сопротивления. Простейшая схема выпрямителя представлена на рис. 333. Найдем выходное напряжение u .

Положим для простоты, что характеристика выпрямителя задана в виде следующей зависимости проводимости от напряжения:

$$Y = \begin{cases} 0 & \text{при } u_{\text{в}} < 0 \\ \infty & \text{при } u_{\text{в}} > 0 \end{cases}.$$

(В действительности прямая и обратная проводимости конечны, но различаются в хорошем выпрямителе во много раз.)

Если сопротивление выпрямителя постоянно в некотором интервале напряжений, то это значит, что в *пределах этого интервала* явления описываются линейным уравнением. В соседнем же интервале может действовать тоже линейная, но уже другая зависимость. В нашем слу-

чае работа схемы описывается следующими линейными уравнениями:

$$\left. \begin{aligned} u &= E & [u_B > 0] \\ u + Ri &= 0 & [u_B < 0] \end{aligned} \right\}.$$

Напряжение на выпрямителе равно, очевидно,

$$u_B = E - u.$$

Беря

$$E = E_m \sin \omega t$$

и выражая в (268) ток через напряжение, получим:

$$\left. \begin{aligned} u &= E_m \sin \omega t & [E > u] \\ \frac{du}{dt} + \frac{1}{RC}u &= 0 & [E < u] \end{aligned} \right\}. \quad (269)$$

Периодическое решение нелинейной задачи получим, учитывая, что напряжение u не может измениться скачком; следовательно, нужно значение напряжения внутри одного интервала (т.е. при $E > u$) приравнять соответствующему значению напряжения внутри другого интервала (т.е. при $E < u$) на общей их границе (т.е. при $E = u$). Это и есть метод сшивания.

Качественная картина явления, совершенно очевидная заранее, представлена на рис. 334. Напряжение изменяется по синусоидальному закону до $t = t_1$, в этот момент скорость убывания напряжения на конденсаторе становится меньше скорости убывания синусоидального напряжения E . Выпрямитель запирается, и конденсатор разряжается на сопротивление R . Процесс

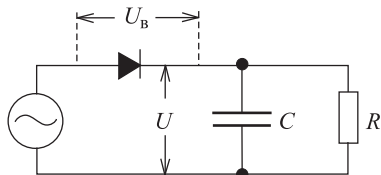


Рис. 333

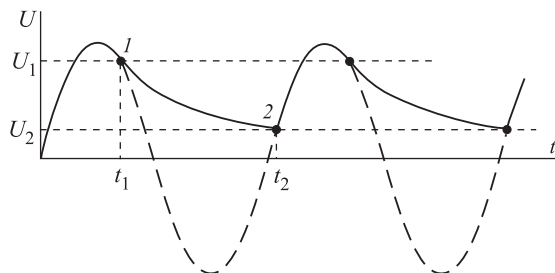


Рис. 334

разряда продолжается, пока в момент $t = t_2$ величины u и E не сравняются. В этот момент выпрямитель отпирается, и дальнейшее изменение напряжения происходит по синусоидальному закону вплоть до $t = t_1 + T$.

Процесс разряда происходит по экспоненциальному закону

$$u = U_1 e^{-\alpha(t-t_1)}, \quad (270)$$

где

$$\alpha = \frac{1}{RC}.$$

Неизвестными являются координаты точек 1 и 2, т. е. величины U_1 , U_2 , t_1 и t_2 . Для их определения нужно располагать четырьмя уравнениями.

Первое уравнение связывает U_1 и t_1

$$U_1 = E_m \sin \omega t_1. \quad (271)$$

Второе уравнение получим, приравняв в этой же точке производные экспоненты (270) и синусоиды (269). Это дает

$$-\alpha U_1 = \omega E_m \cos \omega t_1. \quad (272)$$

Далее имеем

$$U_2 = E_m \sin \omega t_2, \quad (273)$$

с другой стороны должно быть

$$U_2 = U_1 e^{-\alpha(t_2-t_1)}. \quad (274)$$

Система уравнений (271)–(274) определяет искомые величины. Из (271) и (272) находим

$$\operatorname{tg} \omega t_1 = -\frac{\omega}{\alpha}.$$

Выражая $\sin$ через tg и подставляя в (271), получаем

$$U_1 = \frac{E_m}{\sqrt{1 + \frac{\alpha^2}{\omega^2}}}.$$

Для t_2 , приравняв правые части (273) и (274), получим аналитически неразрешимое трансцендентное уравнение. Но графически это уравнение разрешается очень просто: из точки 1 с уже известными координатами U_1, t_1 нужно провести экспоненту (270) до пересечения с синусоидой в точке 2. Координаты этой точки и дают искомые значения U_2 и t_2 . В зависимости от значения постоянной времени $1/\alpha = RC$ будем иметь ход напряжения в пределах, указанных на рис. 335.

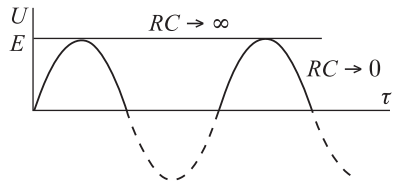


Рис. 335

Режим $RC \rightarrow \infty$ употребляется, в частности, при построении *пик-вольтметров* — приборов, измеряющих максимальное значение напряжения.

§ 95. Резонанс в нелинейной цепи

Явление резонанса в цепи, содержащей нелинейные элементы, имеет существенные особенности. Рассмотрим резонанс в последовательной цепи, содержащей нелинейную емкость.

Составим уравнение напряжений для последовательной цепи из R , L и C , считая, что R и L — постоянные величины, а C зависит от заряда

$$Lq'' + Rq' + \frac{1}{C(q)}q = E_m \sin \omega t$$

или

$$q'' + 2\alpha q' + F(q) = B \sin \omega t. \quad (275)$$

Будем считать, что $F(q)$ — нечетная функция, и положим

$$F(q) = \omega_0^2(q + \lambda q^3).$$

Таким образом, степень нелинейности во всех последующих соотношениях определяется параметром λ ; при $\lambda \rightarrow 0$ мы возвращаемся к обычным линейным зависимостям.

Будем искать установившееся решение (275) в виде

$$q = a \sin \omega t + b \cos \omega t, \quad (276)$$

где a и b — постоянные величины. Амплитуда заряда

$$A = Q_m = \sqrt{a^2 + b^2}.$$

Подставляя решение (276) в (275) и приравнявая порознь коэффициенты при $\sin \omega t$ и $\cos \omega t$, получим два уравнения:

$$\left. \begin{aligned} (\omega_0^2 - \omega^2)a - 2\alpha\omega b + \frac{3}{4}\lambda\omega_0^2 a A^2 &= B \\ (\omega_0^2 - \omega^2)b + 2\alpha\omega a + \frac{3}{4}\lambda\omega_0^2 b A^2 &= 0 \end{aligned} \right\}.$$

Умножая поочередно на a и b , складывая и вычитая, получим уравнение, содержащее только $A^2 = a^2 + b^2$,

$$\left(\frac{3}{4}\lambda\right)^2 \omega_0^4 A^6 + \frac{3}{2}\lambda\omega_0^2(\omega_0^2 - \omega^2)A^4 + [(\omega_0^2 - \omega^2) + \omega_0^2\omega^2 d^2]A^2 - B^2 = \varphi(A^2, \omega^2) = 0. \quad (277)$$

Если решить это уравнение относительно A , то можно получить амплитуду заряда как функцию амплитуды внешней эдс B и частоты ω , т. е. выражение для резонансной кривой.

Но здесь возникает затруднение, связанное с тем, что уравнение (277) — кубическое относительно A^2 . Поэтому его удобнее решать не относительно A^2 , а относительно ω^2 , так как для ω^2 уравнение (277)

является квадратным (а для ω — биквадратным). Перепишав (277) в виде

$$\omega^4 - 2\omega_0^2 \left(1 + \frac{3}{4}\lambda A^2 - \frac{1}{2}d^2 \right) \omega^2 + \omega_0^4 \left(1 + \frac{3}{4}\lambda A^2 \right)^2 - \frac{B^2}{A^2} = 0,$$

найдем решение этого уравнения:

$$\omega^2 = \omega_0^2 \left[1 + \frac{3}{4}\lambda A^2 - \frac{1}{2}d^2 \pm \sqrt{\frac{B^2}{A^2} - d^2 \left(1 + \frac{3}{4}\lambda A^2 \right) + \frac{1}{4}d^4} \right].$$

То обстоятельство, что мы нашли не амплитуду как функцию частоты, а, наоборот, частоту как функцию амплитуды, не меняет сути дела: такая замена с точки зрения графика сводится лишь к перемене местами осей абсцисс и ординат.

Высокая степень уравнения для A указывает на то, что решение может быть многозначным. Поэтому полезно исследовать зависимость между A^2 и ω^2 путем рассмотрения крайних положений касательных к кривой, выражающей эту зависимость.

Производную можно выразить через частные производные неявной функции (277)¹⁾.

$$\frac{dA^2}{d\omega^2} = -\frac{\frac{d\varphi}{d\omega^2}}{\frac{d\varphi}{dA^2}}. \quad (278)$$

Касательная горизонтальна, когда производная обращается в нуль, т. е. когда

$$\frac{d\varphi}{d\omega^2} = 0.$$

Это равенство определяет положение максимумов резонансной кривой. Касательная принимает вертикальное положение, когда производная (278) обращается в бесконечность, т. е. при условии

$$\frac{d\varphi}{dA^2} = 0.$$

Найдем обе частные производные; приравняв их нулю, получим уравнения:

$$\omega^2 - \omega_0^2 \left(1 + \frac{3}{4}\lambda A^2 - \frac{1}{2}d^2 \right) = 0, \quad (279)$$

$$\frac{27}{16}\lambda^2\omega_0^4 A^4 + 3\lambda\omega_0^2(\omega_0^2 - \omega^2)A^2 + (\omega_0^2 - \omega^2)^2 + \omega_0^2\omega^2 d^2 = 0. \quad (280)$$

¹⁾ Рассматривая зависимость между A^2 и ω^2 (вместо зависимости между A и ω), мы не меняем характера особых точек кривой; изменяется лишь масштаб графика, т. е. линейный масштаб заменяется квадратичным.

Уравнение (279) есть уравнение кривой, на которой располагаются максимумы семейства резонансных кривых. Уравнение (280) есть уравнение кривой, представляющей место точек, где резонансная кривая проходит вертикально.

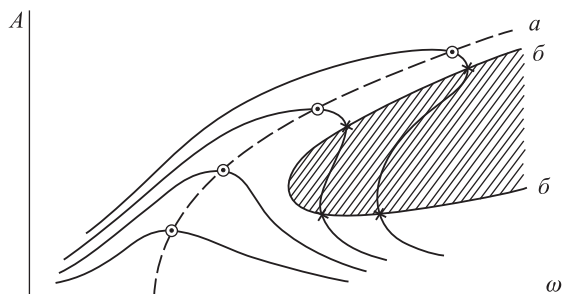


Рис. 336

Семейство резонансных кривых изображено на рис. 336. Как видим, кривые наклонены вправо, причем наклон тем больше, чем больше максимальная амплитуда, так как при этом сильнее проявляется нелинейность. (Заметим, что при $\lambda < 0$ получился бы наклон влево.) На графике нанесены кривые, уравнения которых выражены формулами (279) (кривая a) и (280) (кривая $б$). Максимумы резонансных кривых отмечены кружками; точки, в которых касательные к резонансным кривым вертикальны, отмечены крестиками.

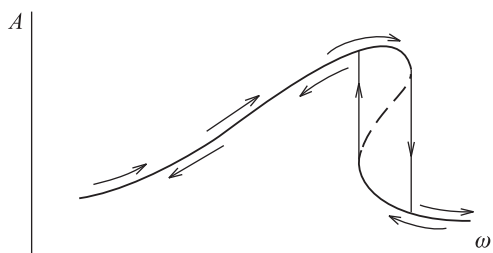


Рис. 337

При экспериментальном снятии нелинейных резонансных кривых типа рис. 336 происходят следующие явления. При повышении частоты амплитуда колебаний нарастает, достигает максимума, затем убывает, пока не будет достигнуто значение, соответствующее точке с вертикальной касательной. Здесь происходит срыв: амплитуда колебаний скачком убывает до значения, соответствующего нижней ветви кривой (рис. 337), а в дальнейшем убывает монотонно. Если же снимать кривую путем уменьшения частоты, то сначала амплитуда нарастает монотонно, а затем срывается на верхнюю ветвь кривой, как показано

на том же рис. 337. Таким образом, скачок амплитуды происходит при разных значениях частоты в зависимости от того, повышается частота или понижается. Участок резонансной кривой между точками срыва (пунктирная линия на рис. 337) соответствует неустойчивым режимам, не могущим наблюдаться при снятии резонансной кривой. Следовательно, область, ограничиваемая кривой б, есть область неустойчивости; на рис. 336 эта область заштрихована.

Аналогичные явления наблюдаются и при нелинейной индуктивности, например при наличии в цепи катушки индуктивности с ферромагнитным сердечником. Резонансные явления в такой цепи называют обычно *феррорезонансом*.

В заключение заметим, что резонанс в нелинейной цепи возможен и на дробной частоте, т. е. на частоте, равной некоторой доле частоты внешнего напряжения. Другими словами, резонанс в нелинейной цепи может применяться для деления частоты. Так, например, в технике имеют применение феррорезонансные делители частоты в отношении 1 : 3.

§ 96. Триггеры

В радиотехнике все более широкое применение получают устройства, имеющие одно или несколько устойчивых состояний, из которых эти устройства могут быть выведены внешними напряжениями, обычно импульсными.

В процессе развития различные разновидности таких устройств получали наименование электронных реле, кipp-реле, спусковых схем, курковых схем, ждущих схем.

За последнее время установился термин *триггер*. Триггером называется устройство, имеющее два устойчивых состояния; действием внешнего импульсного напряжения триггер может быть переведен из одного устойчивого состояния в другое. При этом для всех применений триггеров существенно, что переход из одного устойчивого состояния в другое совершается очень быстро; часто говорят, что схема «опрокидывается», подчеркивая этим выражением скачкообразный характер происходящих в триггере явлений.

Схемы триггеров очень многочисленны и разнообразны. Начнем с упрощенного описания схемы, основанной на применении нелинейного усилителя с положительной обратной связью (рис. 338).

Усилитель Y — нелинейный усилитель постоянного тока. Его свойства выражаются характеристикой рис. 339, которая представляет зависимость выходного напряжения u_2 от входного напряжения u_1 . Через E обозначено внешнее управляющее напряжение.

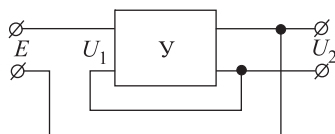


Рис. 338

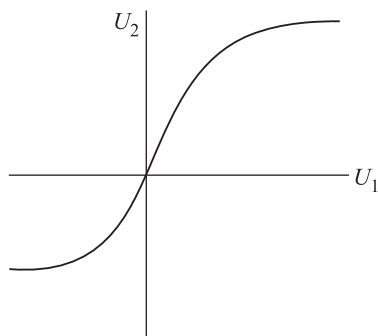


Рис. 339

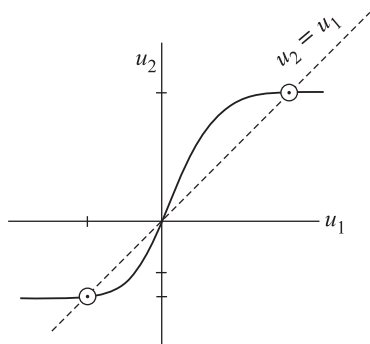


Рис. 340

Рассмотрим сначала поведение схемы при отсутствии внешнего напряжения E (входные зажимы замкнуты накоротко). Пусть на входе усилителя появилось некоторое постоянное напряжение. Будучи усилено, оно снова попадает на вход с тем же знаком (положительная обратная связь). Таким образом, напряжение возрастает, пока дальнейший его рост не будет ограничен нелинейностью. Установившийся режим будет устойчивым. Для нахождения устойчивого состояния нужно решить совместно два уравнения:

выражение характеристики усилителя

$$u_2 = f(u_1)$$

и очевидное равенство

$$u_2 = u_1.$$

Графическое построение решения показано на рис. 340. Два устойчивых состояния триггера обозначены кружками.

Из предыдущего рассуждения следует, что начало координат, хотя и является точкой, удовлетворяющей вышеприведенной системе уравнений (так как $u_2 = u_1 = 0$), но представляет неустойчивое состояние. Достаточно сколь угодно малого флуктуационного напряжения, чтобы триггер оказался в одном из двух своих устойчивых состояний (в зависимости от знака флуктуации).

Пусть теперь на вход триггера подано управляющее напряжение e . При этом, очевидно,

$$u_1 = e + u_2.$$

Желательно построить характеристику триггера в целом, т.е. зависимость выходного напряжения u_2 от управляющего напряжения e . Можно записать

$$u_1 = f(u_1) = f(e + u_2) = F(e).$$

Нельзя получить в общем виде явное выражение для u_2 . Но из приведенного выражения видно, что абсцисса искомой функции F меньше абсциссы известной функции f на величину u_2 . Отсюда следует способ

графического построения функции F . Он состоит в том, что из абсцисс кривой $f(e)$ вычитаются абсциссы прямой $u_2 = e$ (или, что то же, добавляются абсциссы прямой $u_2 = -e$). Это построение показано на рис. 341.

Устойчивые состояния отмечены теперь точками, лежащими на пересечениях характеристики с осью ординат (т. е. $e = 0$).

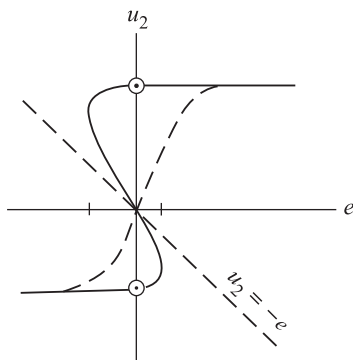


Рис. 341

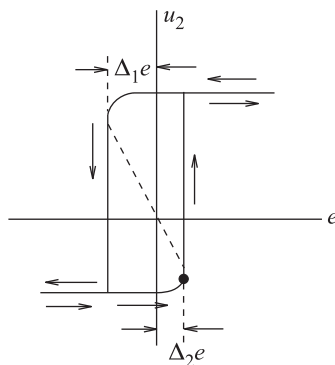


Рис. 342

Проследим за поведением триггера при изменениях управляющего напряжения. Если, начиная с верхнего устойчивого состояния, уменьшать e , то u_2 будет изменяться, следуя по верхней ветви характеристики. При некотором отрицательном значении $e = \Delta_1 e$ касательная к характеристике делается вертикальной (т. е. $\frac{du_2}{de} = \infty$ или $\frac{de}{du_2} = 0$). В этой точке происходит опрокидывание схемы: напряжение u_2 скачком перебрасывается на нижнюю ветвь характеристики. Аналогичный скачок происходит при $e = \Delta_2 e$ при увеличении управляющего напряжения, начиная от нижнего устойчивого состояния. Направления возможных изменений отмечены стрелками на рис. 342.

Итак, чтобы перебросить триггер из одного устойчивого состояния в другое, достаточно подать на вход напряжение соответствующей полярности, превосходящее по абсолютной величине Δe . Так как переброс совершается скачкообразно, то управляющее напряжение может представлять собой короткий импульс.

Отметим, что характеристика рис. 342 имеет типичный *гистерезисный* характер.

Приведем примеры схем триггеров с усилителями на триодах. На рис. 343, *a* показана одна из простейших (но не лучших) схем. Она представляет собой двухступенный усилитель, выход которого соединен со входом.

Требуемая форма характеристики (рис. 339) обеспечивается тем, что при больших отрицательных напряжениях запирается одна лампа, а при больших положительных напряжениях — другая. Схема совер-

шенно симметрична; безразлично, какую ступень считать первой, а какую — второй. Чтобы подчеркнуть это обстоятельство, часто представляют эту схему, как показано на рис. 343, б.

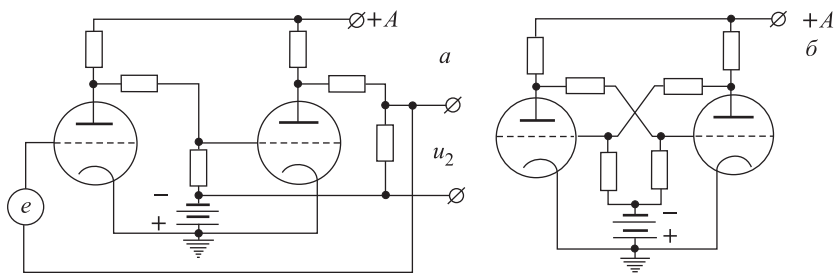


Рис. 343

Способ управления можно изменить. Так, управляющее напряжение можно подвести к обеим сеткам одновременно (включив источник управляющего напряжения в общую цепь обеих сеток). Тогда триггер будет опрокидываться всякий раз при поступлении управляющих импульсов одной и той же полярности и не будет реагировать на импульсы противоположной полярности.

Описание действия триггера очень полезно сравнить с описанием действия мультивибратора (см. § 93), от которого схема рис. 343, б отличается только тем, что анод каждой лампы связан с сеткой другой не через конденсатор, а через сопротивление.

Начнем с равновесного состояния, т. е. такого, при котором анодные токи и напряжения на анодах обеих ламп равны. Это состояние не-

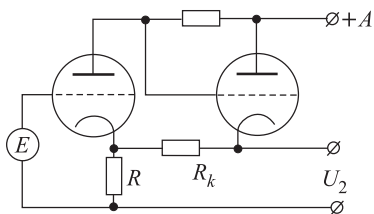


Рис. 344

устойчиво. Действительно, пусть ток в правой лампе возрастет. Напряжение на аноде уменьшится, в результате чего возрастет отрицательное напряжение на аноде левой лампы и уменьшится отрицательное напряжение на сетке, а следовательно, произойдет дальнейшее увеличение анодного тока правой лампы. Схема опрокинется и левая лампа окажется запертой. Это состояние устойчиво.

Если в этом состоянии подать на сетку левой лампы достаточно большой положительный управляющий импульс, то схема опрокинется и перейдет в другое устойчивое состояние, когда заперта правая лампа.

Отличие от мультивибратора, работающего в генераторном режиме, состоит, как видим, в том, что мультивибратор благодаря связям через емкость вообще не имеет устойчивых состояний.

Вариант схемы триггера, отличающийся применением катодной связи, показан на рис. 344. Эта схема также представляет собой двухступенный усилитель. Напряжение положительной обратной связи на сетку первой лампы снимается с сопротивления R .

Одной из главнейших технических характеристик триггера является скорость опрокидывания. Она определяется главным образом постоянными времени цепей усилителя. Постоянные времени, в свою очередь, зависят в основном от паразитных емкостей. Современная техника предъявляет триггерным устройствам очень высокие требования в отношении быстродействия. В рамках нашего курса нет возможности заняться обстоятельно этим важным вопросом; он изучается в курсе импульсной техники. Отметим лишь, что схемы вроде рис. 343 или 344 позволяют получить время опрокидывания порядка десятых и сотых долей микросекунды.

Как уже говорилось, устройство триггеров разнообразно. В качестве примера устройства, основанного на принципе, совершенно отличном от описанного выше, упомянем о магнитном триггере. Здесь используются элементы с *прямоугольной петлей гистерезиса* (ППГ).

Устройство представляет собой замкнутый сердечник с двумя обмотками. Характеристика намагничивания, т. е. зависимость магнитной индукции B от намагничивающего поля H (которое определяется ампер-витками намагничивающей обмотки, играющей роль управляющей), имеет прямоугольную форму, подобную характеристике рис. 342. Триггер имеет два устойчивых состояния, соответствующих намагничиванию сердечника до насыщения в двух противоположных направлениях. Импульс надлежащей полярности, поданный в управляющую обмотку, перебрасывает триггер из одного состояния в другое. При этом происходит изменение потока, наводящее импульсное напряжение во второй обмотке. Преимуществом магнитных триггеров является их долговечность и компактность. Сердечники выполняются в виде прессованных ферритовых колечек диаметром всего 3–4 мм.

В заключение отметим некоторые применения триггеров.

Формирование импульсов. Опрокидывание триггера используется для формирования импульсов с весьма крутыми фронтами. Процесс формирования поясняется рис. 345. Точками отмечены значения e , при которых происходит опрокидывание триггера.

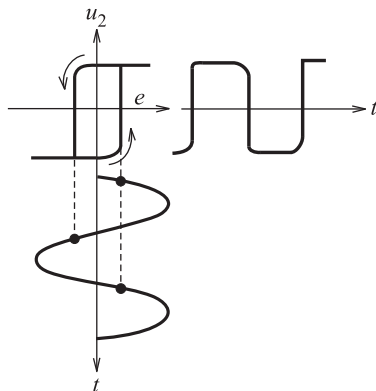


Рис. 345

Запоминание. Наличие устойчивых состояний, в которых триггер может пребывать неопределенно долго, используется для построения запоминающих устройств, в частности в вычислительных машинах. Триггеры, естественно, пригодны для запоминания информации в форме двоичных чисел; одному устойчивому состоянию приписывается символ 0, другому — 1.

Счет импульсов. Соединив несколько триггеров цепочкой, т. е. так, чтобы выходной импульс предшествующего триггера служил управляющим импульсом для последующего, получаем весьма простую схему счетчика импульсов, ведущего счет в двоичном исчислении. Эту возможность рассмотрим несколько подробнее.

Пусть триггеры, из которых собран счетчик, обладают следующими свойствами:

1) триггер опрокидывается из любого данного состояния в другое импульсами определенной полярности (которую назовем положительной) и не реагирует на импульсы противоположной полярности;

2) при переходе из состояния 1 в состояние 0 на выходе триггера вырабатывается импульс положительной полярности, а при обратном переходе (т. е. от 0 к 1) — импульс отрицательной полярности.

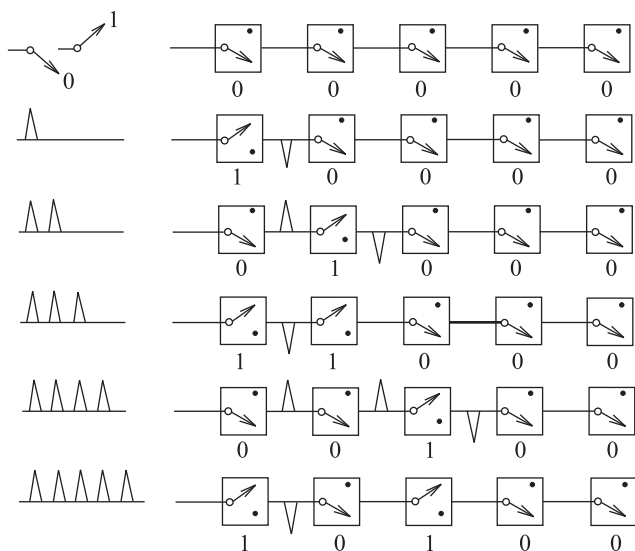


Рис. 346

Цепочка триггеров изображена схематически на рис. 346. Каждый триггер представлен квадратиком, стрелка указывает состояние, в котором он находится. Импульсы на входах и выходах триггеров представлены треугольниками, направленными вершиной вверх для положительной полярности и вниз для отрицательной.

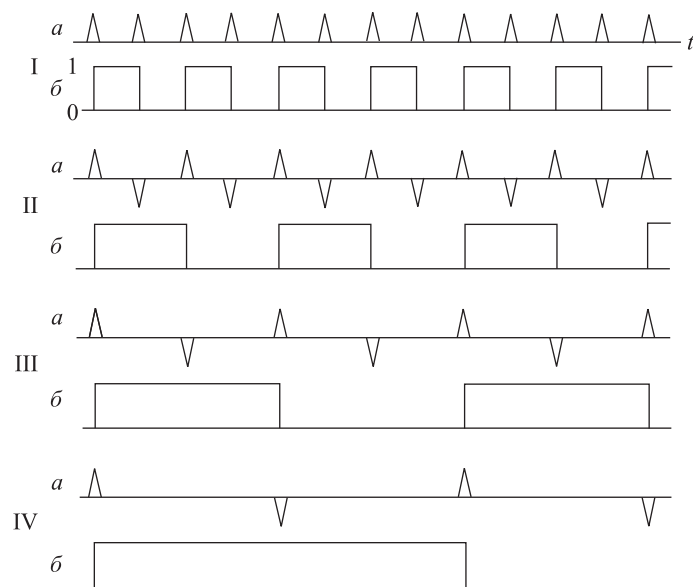


Рис. 347

Первая строка рис. 346 изображает исходное состояние счетчика; все триггеры находятся в состоянии 0. Вторая строка: поступивший на вход счетчика положительный импульс перебрасывает первый триггер в состояние 1. Отрицательный импульс, выданный на выход первого триггера, т. е. на вход второго, не производит на последний действия. Третья строка: следующий (второй) импульс на входе перебрасывает первый триггер в положение 0. Появившийся при этом на входе второго триггера положительный импульс перебрасывает его в положение 1. Дальнейшее видно из рисунка.

Заметим теперь, что при описанном способе действия число поступивших на входе импульсов оказывается записанным в двоичной форме на триггерной цепочке, если читать это число слева направо в порядке возрастания разрядов (т. е., проще говоря, на первом триггере записываются единицы, на втором — двойки, на третьем — четверки и т. д.).

Возможно построение и десятичного счетчика, но эту возможность рассматривать не будем.

Деление частоты. Если на вход вышеописанного двоичного счетчика поступает периодическая последовательность импульсов, то частота следования в каждом последующем триггере уменьшается вдвое, а в цепочке из n триггеров — в 2^n раз.

Это совершенно ясно из рис. 347, на котором изображено следующее: Ia — входные импульсы, Ib — изменение состояний первого триггера, IIa — импульсы на выходе первого триггера и на входе второго, IIб — изменение состояний второго триггера и т. д.

§ 97. Автоколебательная система под внешним воздействием

Выведем некоторые общие соотношения, которые в дальнейшем будут использованы при рассмотрении ряда технических вопросов. Речь пойдет об автоколебательной системе, на которую действует внешняя электродвижущая сила. Всякая автоколебательная система нелинейна;

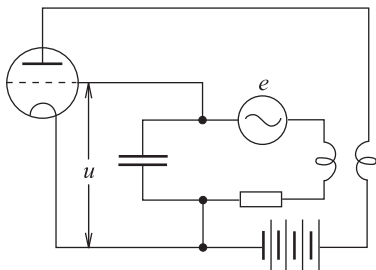


Рис. 348

поэтому при воздействии на нее внешней эдс возникает ряд своеобразных нелинейных явлений, используемых в радиотехнике.

Составим уравнение для схемы рис. 348, представляющей собой генератор с контуром в цепи сетки. Отличие этой схемы от ранее рассматривавшейся состоит в том, что в контуре действует еще внешняя эдс $e = E_m \cos \omega t$. Обозначив напряжение на сетке через u , получим

$$u'' + 2\delta u' + \omega_0^2 u = \omega_0^2 E_m \cos \omega t, \quad (281)$$

где, как и раньше, обозначено:

$$\omega_0^2 = \frac{1}{LC}, \quad 2\delta = 2\alpha - \omega_0^2 MS.$$

Пусть характеристика лампы аппроксимирована многочленом 3-й степени, тогда:

$$S = a_1 + 3a_3 u^2, \quad 2\delta = 2\delta_0 + \gamma u^2,$$

где:

$$2\delta_0 = 2\alpha - \omega_0^2 MS_0, \quad \gamma = 3\omega_0^2 Ma_3.$$

Используя эти соотношения и обозначения, перепишем (281) в виде

$$u'' + (2\delta_0 + \gamma U^2) u' + \omega_0^2 u = \omega_0^2 E_m \cos \omega t. \quad (282)$$

Решение уравнения (282) будем искать в виде

$$u = A \sin(\omega t + \varphi), \quad (283)$$

где A и φ — неизвестные амплитуда и фаза, являющиеся в общем случае функциями времени, что следует иметь в виду при дифференцировании. Так, первая производная (283) будет

$$u' = A' \sin(\omega t + \varphi) + A(\omega + \varphi') \cos(\omega t + \varphi).$$

Подставим решение (283) в (282); при подстановке $u^2 u'$ отбросим члены, содержащие третью гармонику. Отбросим также член, содержащий A'' (так как A — медленная функция по сравнению с $\sin \omega t$).

После этого разложим правую часть (282) по формуле

$$\cos \omega t \equiv \cos \varphi \cos (\omega t + \varphi) + \sin \varphi \sin (\omega t + \varphi)$$

и приравняем порознь коэффициенты при $\sin (\omega t + \varphi)$ и $\cos (\omega t + \varphi)$ в правой и левой частях. Это даст нам два уравнения:

$$\left. \begin{aligned} \left(2\delta_0 + \frac{3}{4}\gamma A^2 \right) A' + (\omega_0^2 - \omega_1^2)A &= \omega_0^2 E_m \sin \varphi \\ 2\omega_1 A' + \left(2\delta_0 + \frac{1}{4}\gamma A^2 \right) \omega_1 A &= -\omega_0^2 E_m \cos \varphi \end{aligned} \right\}. \quad (284)$$

Здесь обозначено

$$\omega_1 = \frac{d}{dt}(\omega t + \varphi) = \omega + \varphi'.$$

Будем в дальнейшем рассматривать главным образом установившийся режим, когда

$$A' = \varphi' = 0.$$

При этом условии получаем из (284) уравнения для установившихся амплитуды и фазы:

$$\left. \begin{aligned} (\omega_0^2 - \omega^2)A &= \omega_0^2 E_m \sin \varphi \\ \left(2\delta_0 + \frac{1}{4}\gamma A^2 \right) \omega A &= -\omega_0^2 E_m \cos \varphi \end{aligned} \right\}. \quad (285)$$

При отсутствии внешней эдс, т. е. при $E_m = 0$, первое уравнение даст частоту автоколебаний

$$\omega = \omega_0,$$

а второе уравнение — уже известное (см. § 85) значение установившейся амплитуды автоколебаний генератора с мягким возбуждением

$$A_0 = \sqrt{-\frac{8\delta_0}{\gamma}} = \sqrt{\frac{4}{3} \frac{a_1 - R \frac{C}{M}}{a_3}}. \quad (286)$$

Возводя уравнения (285) в квадрат и складывая, получим

$$\left[(\omega_0^2 - \omega^2)^2 + \omega^2 \left(2\delta_0 + \frac{1}{4}\gamma A^2 \right) \right] A^2 = \omega_0^4 E_m^2. \quad (287)$$

Это соотношение позволяет найти амплитуду установившихся колебаний, которая зависит: 1) от амплитуды внешней эдс E_m ; 2) от расстройки $\omega_0^2 - \omega^2$; 3) от начального затухания $\delta_0 = \alpha - \frac{1}{2}\omega_0^2 MS_0$ и от 4) нелинейности характеристики, выражаемой коэффициентом $\gamma = 3\omega_0^2 Ma_3$. Если найти из (287) амплитуду A как функцию частоты, то получим выражение для своего рода резонансных кривых.

Здесь, однако, мы столкнемся с тем же затруднением, что и в предыдущем параграфе при исследовании нелинейного резонанса: относительно A мы получаем уравнение 6-й степени. Это затруднение преодолевается тем же приемом — обращением интересующей нас зависимости.

При некоторых соотношениях решения уравнения (287) неустойчивы. Границу области неустойчивости можно определить как место точек, в которых касательные к резонансным кривым вертикальны, аналогично тому, как это делалось в предыдущем параграфе.

Положим, что расстройка мала, т. е. $\omega \approx \omega_0$, и обозначим

$$\Delta\omega = \omega_0 - \omega.$$

Воспользовавшись, кроме того, соотношением (286), можно преобразовать (287) к виду

$$\left[\frac{\Delta\omega^2}{\delta_0^2} + \left(1 - \frac{A^2}{A_0^2} \right)^2 \right] \frac{A^2}{A_0^2} = \frac{\omega_0^2 E_m^2}{4\delta_0^2 A_0^2}$$

или, обозначая для краткости:

$$\frac{\Delta\omega}{\delta_0} = x, \quad \frac{A^2}{A_0^2} = y, \quad \frac{\omega_0^2 E_m^2}{4\delta_0^2 A_0^2} = E$$

к виду

$$y = [x^2 + (1 - y)^2] - E = F(x, y) = 0. \quad (288)$$

Таким образом, уравнение резонансных кривых (287) приведено к виду (288); в этой форме уравнение связывает переменные x и y . Первая переменная пропорциональна расстройке, вторая есть квадрат относительной амплитуды (т. е. амплитуды A при наличии внешней эдс, отнесенной к амплитуде A_0 автоколебаний при отсутствии внешней эдс). График y как функции x изображает в известном масштабе резонансные кривые. Уравнение (288) является кубическим относительно y и квадратным относительно x . Поэтому для вычислений удобнее разрешить его относительно x . Геометрический смысл подобного обращения был разъяснен выше.

Исследуем производную

$$\frac{dy}{dx} = - \frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}.$$

Уравнение координат максимумов получим, приравняв производную нулю, что даст условие $dF/dx = 0$. Имеем

$$\frac{\partial F}{\partial x} = 2yx = 0.$$

Отсюда следует, что максимумы располагаются на вертикальной прямой $x = 0$, что соответствует отсутствию расстройки, т. е. $\omega = \omega_0$;

ось x , уравнение которой есть $y = 0$, является асимптотой семейства резонансных кривых.

Производная dy/dx обращается в бесконечность (т. е. касательные к резонансным кривым вертикальны) в точках, для которых

$$\frac{\partial F}{\partial y} = 0.$$

Это равенство является границей области устойчивых решений. Дифференцируя (288) по y , находим

$$\frac{\partial F}{\partial y} = x^2 + (1 - y)^2 - 2y(1 - y) = x^2 + (1 - y)(1 - 3y) = 0. \quad (289)$$

Это — уравнение эллипса с центром в точке $x = 0$, $y = 2/3$ и с вертикальной полуосью, равной $1/3$. Таким образом, верхняя горизонтальная касательная к эллипсу проходит на высоте $y = 1$. Область внутри

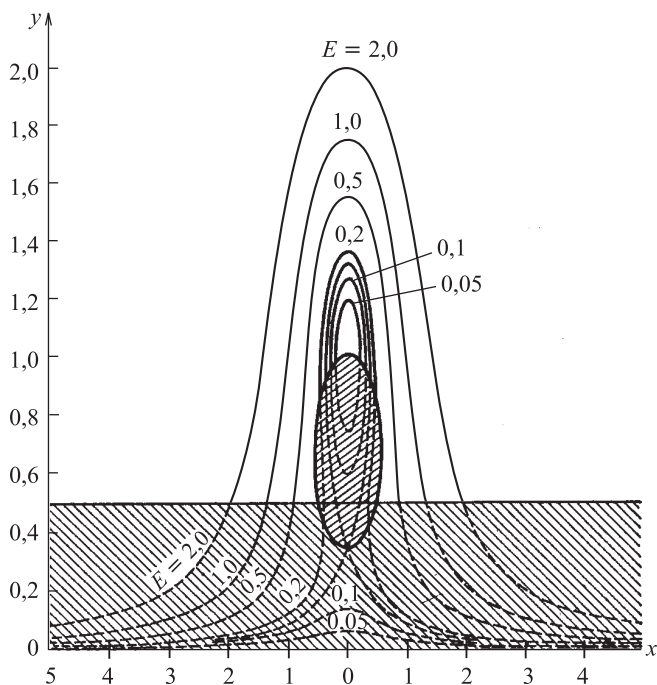


Рис. 349

эллипса соответствует неустойчивым режимам. Подробное исследование устойчивости показывает, что, кроме уже найденного, существует еще одно условие устойчивости, а именно:

$$y > \frac{1}{2}, \quad \text{т. е.} \quad A > \frac{1}{\sqrt{2}} A_0.$$

Таким образом, при малых амплитудах «вынужденных» колебаний режим оказывается неустойчивым. Горизонтальная прямая

$$y = \frac{1}{2} \quad (290)$$

является нижней границей области устойчивости.

Графики функции (288) совместно с граничными линиями (289) и (290) представлены на рис. 349. Области неустойчивости заштрихованы.

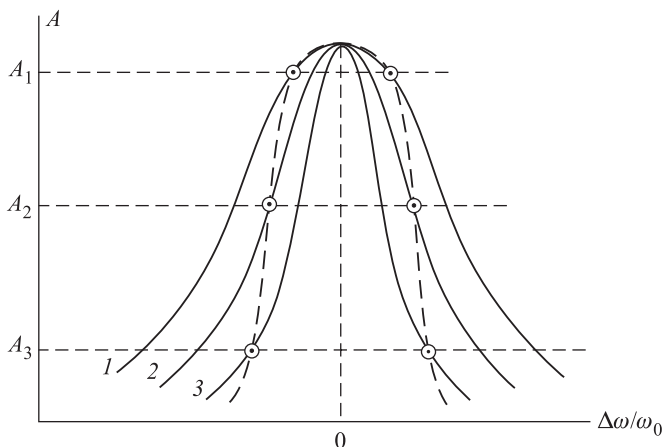


Рис. 350

Иногда для пояснения характера зависимости (287) прибегают к нестрогому, но зато очень наглядному приему. Формулу (287) переписывают в виде

$$A = \frac{E_m}{\sqrt{\left(\frac{\omega_0^2 - \omega^2}{\omega_0^2}\right)^2 + \frac{\omega^2}{\omega_0^2} \left(\frac{2\delta_0}{\omega_0} + \frac{1}{4} \frac{\gamma}{\omega_0} A^2\right)^2}}$$

или, полагая расстройку малой,

$$A \approx \frac{E_m}{\sqrt{\varepsilon^2 + \left(d_0 + \frac{1}{4} \frac{\gamma}{\omega_0} A^2\right)^2}}, \quad (291)$$

где $\varepsilon = \frac{2\Delta\omega}{\omega_0}$ — обычное обозначение относительной расстройки; $d_0 = \frac{2\delta_0}{\omega_0}$ — начальное затухание.

Последняя формула рассматривается как выражение для резонансной кривой, отличающееся от обычного тем, что затухание зависит от амплитуды (второй член под корнем).

При такой трактовке можно подойти к построению зависимости (291), исходя из семейства обычных резонансных кривых при различном затухании. Такое семейство изображено на рис. 350. При больших амплитудах (значение A_1 на рис. 350) затухание велико; такому значению затухания соответствует резонансная кривая 1. При меньшей амплитуде A_2 затухание соответственно меньше, и мы переходим на более узкую резонансную кривую 2. Еще меньшей амплитуде A_3 и еще меньшему затуханию соответствует резонансная кривая 3. Соединяя отмеченные на рисунках точки пунктирной линией, получаем очерта-ние резонансной кривой.

Заметим в заключение, что условие устойчивости на основании первого уравнения (285) может быть выражено через угол φ , определяющий сдвиг фазы между внешней эдс и свободным автоколебанием, в виде

$$-\frac{\pi}{2} < \varphi < \frac{\pi}{2}.$$

§ 98. Захватывание

К числу нелинейных явлений, имеющих большое значение в радиотехнике, относится явление *захватывания*.

Оно состоит в том, что если на генератор воздействует периодическая эдс внешнего источника, то при определенных соотношениях генератор генерирует не ту частоту, на которую настроен его задающий контур, а частоту внешней эдс. Получается своеобразный вынужденный режим; в этом смысле и говорят, что внешний источник *захватывает* генератор, навязывая ему свою частоту.

Рассмотрим сначала качественную картину явления. Пусть в схему обычного лампового генератора введен источник внешней синусоидальной эдс e с частотой ω (рис. 348). При отсутствии внешней эдс генерируется частота ω_0 , на которую настроен контур генератора. Пусть теперь источник эдс e включен и пусть частота постепенно изменяется, приближаясь к ω_0 . Пока расстройка, т. е. разность $\omega - \omega_0$ велика, нелинейность не дает себя знать и в схеме независимо существуют два колебания с частотами ω и ω_0 . Интерференция этих колебаний порождает биения, проявляющиеся в периодическом изменении амплитуды результирующего колебания. При сближении частот ω и ω_0 начинают наблюдаться специфические для данной нелинейной системы явления. Частота генератора изменяется, сближаясь с частотой внешнего напряжения. При определенном значении $\omega = \omega_1$ частота генератора делается равной частоте внешней эдс, и это равенство сохраняется вплоть до значения $\omega = \omega_2$. При дальнейшем изменении ω генератор генерирует уже частоту, отличную от ω , и постепенно

возвращается к исходному значению ω_0 . Все это показано на рис. 351, на котором по оси ординат отложена частота ω' , которую фактически

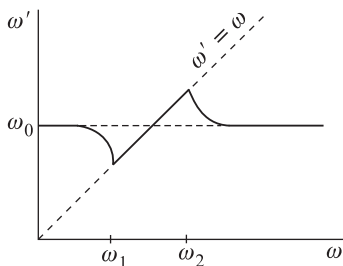


Рис. 351

генерирует генератор. Интервал частот от ω_1 до ω_2 называется *полосой захватывания*. В полосе захватывания биений нет. Режим в полосе захватывания можно рассматривать как вынужденное колебание, так как оно происходит с частотой источника внешней эдс.

Найдем ширину полосы захватывания и выясним, от чего она зависит.

Для решения этой задачи обратимся к картине явлений, представленной резонансными кривыми рис. 349. Устойчивый режим, представленный на рис. 349 сплошными линиями, — это и есть режим захватывания, когда происходят вынужденные колебания с частотой внешней эдс. Переход в неустойчивую область означает возникновение биений. Таким образом, ширина полосы захватывания (интервал $\omega_2 - \omega_1$ на рис. 351) есть не что иное, как ширина резонансной кривой на границе области неустойчивости. Исходя из этих соображений, легко определились ширину полосы захватывания.

Уравнение резонансных кривых рис. 349 есть

$$y[x^2 + (1 - y)^2] - E = 0. \quad (292)$$

При малых значениях внешней эдс, т. е. при малых E , кривые опираются на вершину эллипса на высоте $y = 1$. Подставляя это значение в (292), находим

$$x^2 = E$$

или, раскрывая сокращенные обозначения (см. § 97),

$$\frac{2\Delta\omega}{\omega_0} = \frac{E_m}{A_0}.$$

Величина слева — это и есть относительная ширина полосы захватывания. Она равна отношению амплитуды E_m внешней эдс к амплитуде A_0 напряжения автоколебания при отсутствии внешней эдс. Считая величину A_0 постоянным параметром рассматриваемой системы, заключаем, что ширина полосы захватывания тем больше, чем сильнее внешнее воздействие.

При больших E_m резонансные кривые опираются уже не на эллипс, а на горизонтальную прямую $y = 1/2$. Подставляя это значение в (292), найдем

$$x^2 = 2E - \frac{1}{4}$$

или

$$\frac{2\Delta\omega}{\omega_0} = \sqrt{2\frac{E_m^2}{A_0^2} - \frac{\delta_0^2}{\omega_0^2}}.$$

Второй член под корнем выражает квадрат начального (т. е. с учетом действия обратной связи) затухания системы. Это малая величина, так что можно положить в рассматриваемом случае

$$\frac{2\Delta\omega}{\omega_0} = \sqrt{2} \frac{E_m}{A_0}.$$

Для промежуточных значений E_m , когда резонансные кривые опираются на скаты эллипса, можно принять

$$\frac{2\Delta\omega}{\omega_0} = k \frac{E_m}{A_0},$$

где, как мы установили,

$$1 < k < \sqrt{2}.$$

Точное значение k может быть найдено, в случае надобности, путем совместного решения уравнений (288) и (289).

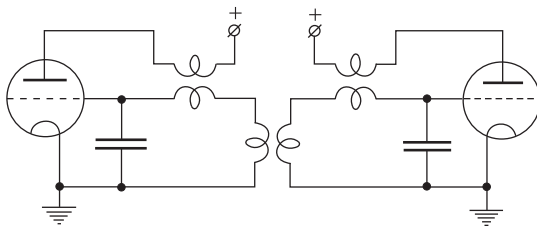


Рис. 352

Явление захватывания используется в радиотехнике для синхронизации генератора внешней эдс или для взаимной синхронизации двух генераторов. Если два генератора связаны друг с другом, как показано на схеме рис. 352, то каждый из них наводит в контуре другого эдс, являющуюся для этого другого генератора внешней. Работа схемы описывается системой из двух уравнений (285), причем правые части играют роль членов связи. Если оба генератора одинаковы, но настроены на разные частоты, то они могут взаимно синхронизироваться на некоторой промежуточной частоте, имеющей в общем случае два возможных значения. Если же один из генераторов является более мощным, то он может рассматриваться как источник независимой эдс и, следовательно, захватывает менее мощный генератор, заставляя его работать на своей частоте.

§ 99. Регенеративный приемник

Схема простейшего регенеративного приемника, изображенная на рис. 353, сводится к схеме рис. 348. Однако в отличие от ранее рассмотренного случая регенеративный приемник характеризуется тем, что он работает в недовозбужденном режиме, хотя и вблизи порога возбуждения, так как при этом получается наибольшее усиление. Уравнение регенеративного приемника совпадает с уравнением (287)

$$(\omega_0^2 - \omega^2)^2 + \omega^2 \left(2\delta_0 + \frac{1}{4}\gamma A^2 \right) A^2 = \omega_0^4 E_m^2. \quad (293)$$

Рассмотрим вопрос о чувствительности приемника при точной настройке на принимаемую частоту, т. е. при $\omega_0 = \omega$. При этом условии из (293) получим

$$\left(2\delta_0 + \frac{1}{4}\gamma A^2 \right) A = \omega_0 E_m.$$

Увеличивая обратную связь, можно довести приемник почти до самовозбуждения. При этом начальный показатель затухания $\delta_0 = \alpha - \frac{1}{2}\omega_1^2 MS$ близок к нулю и

$$\frac{1}{4}\gamma A^3 = \omega_0 E_m,$$

откуда

$$A = \left(4 \frac{\omega_0}{\gamma} E_m \right)^{1/3} = k E_m^{1/3}. \quad (294)$$

Найдем чувствительность (или, что то же, усиление) приемника, определив ее как

$$\nu = \frac{dA}{dE_m}.$$

Дифференцируя (294), получаем

$$\nu = k E_m^{-2/3}.$$

Как видим, при E_m , стремящемся к нулю, чувствительность стремится к бесконечности. Таким образом, регенеративный приемник обладает переменной чувствительностью, тем большей, чем слабее принимаемый сигнал. Это является, конечно, выгодным свойством рассматриваемого приемника.

Практически предельно достижимая чувствительность ограничивается собственным шумом приемника, который вблизи порога самовозбуждения становится очень сильным из-за большого усиления.

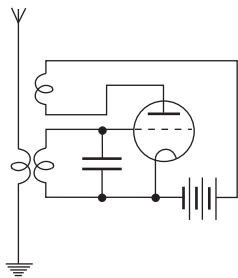


Рис. 353

Регенеративный приемник сыграл в свое время большую роль благодаря исключительной простоте и высокой чувствительности. Однако он обладает очень существенными недостатками: неустойчивостью, плохой избирательностью. К тому же при работе на самом пороге возбуждения часто возникает генерация, создающая помеху для близко расположенных приемников. Поэтому регенеративный приемник был вытеснен более совершенными приемниками, а регенеративная схема сохранила значение лишь в качестве регенеративного фильтра.

§ 100. Импульсная синхронизация

В § 98 рассматривалась синхронизация генератора синусоидальным внешним напряжением.

В этом параграфе будет рассмотрена синхронизация периодически короткими импульсами. Этот случай допускает очень наглядное качественное объяснение процесса синхронизации, и мы воспользуемся этой возможностью; аналитического исследования импульсной синхронизации делать не будем.

Мы исходим по-прежнему из схемы рис. 348, но предполагаем теперь, что внешняя эдс представляет собой периодическую последовательность коротких импульсов произвольной формы. Для объяснения механизма синхронизации представим явление в виде диаграмм на фазовой плоскости. Прежде всего дадим графическое представление действия импульса напряжения.

Возьмем в качестве переменных, откладываемых по координатным осям, те же величины, что и ранее (§ 89), а именно:

$$x = \int u dt, \quad x' = u.$$

Пусть импульс характеризуется тем, что напряжение начинает изменяться от некоторого начального значения (в частности, равно нулю), а затем по прошествии небольшого времени снова возвращается к исходному значению. Таким образом, по окончании импульса напряжение $u = x'$ остается неизменным, а интеграл от напряжения получает приращение, определяемое площадью импульса. Эти соотношения поясняются на рис. 354 для случая прямоугольного импульса. Если теперь представить те же соотношения на фазовой плоскости в координатах x, x' , то изображающая точка, находившаяся первоначально в положении 1, перейдет под действием импульса в новое положение 2, описав при этом намеченную пунктиром траекторию (рис. 355). Время перехода равно длительности импульса. Будем считать, что импульс настолько короток, что изображающая точка перебрасывается из 1 в 2 практически мгновенно. Что касается формы траектории, то она зависит, конечно, от формы импульса. Но нас, собственно, не интересует вид траектории; для дальнейшего существенно лишь то, что изображающая точка смещается в конечном счете параллельно оси

x (так как напряжение до и после импульса одинаково) на отрезок, пропорциональный площади импульса. Это смещение изображающей точки и выражает действие импульса. Обозначим его через h .

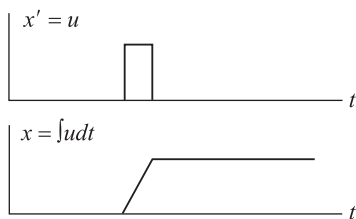


Рис. 354

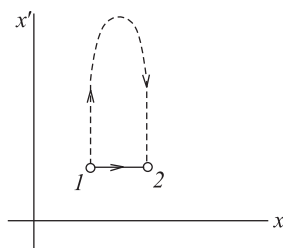


Рис. 355

Теперь рассмотрим действие синхронизирующих импульсов на работающий генератор. При отсутствии импульсов режим генератора выражается устойчивым предельным циклом. При наличии же импульсов должен образоваться замкнутый цикл, включающий в себя скачок, обусловленный действием импульса. Такой цикл может образоваться лишь из отрезка спирали, как показано на рис. 356. По оси абсцисс отложено $\omega_0 x$ вместо x для того, чтобы уравнивать масштабы по обеим осям. При таком согласованном масштабе предельный цикл имеет форму окружности. Будем считать, что изображающая точка движется по спирали с постоянной угловой скоростью и что поворот на 2π соответствует периоду автоколебаний

$$T_0 = \frac{2\pi}{\omega_0}.$$

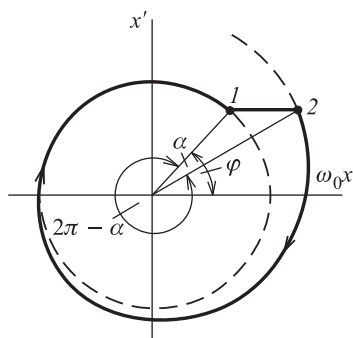


Рис. 356

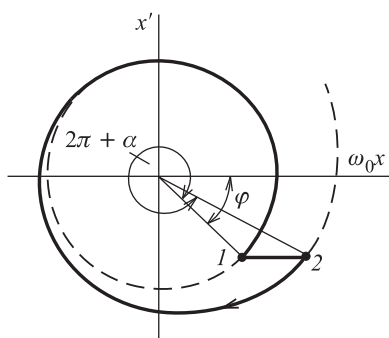


Рис. 357

Тогда на рис. 356 сразу видно, что изображающая точка обходит замкнутый цикл, отмеченный сплошной линией, за время $T < T_0$. Это следует из того, что угловое перемещение точки равно $2\pi - \alpha < 2\pi$, а следовательно, на такое перемещение затрачивается меньшее время.

Участок скачка (1–2) проходится, по нашему предположению, мгновенно. Рисунок 356 соответствует, стало быть, случаю, когда импульсы синхронизируют генератор на частоте $\omega > \omega_0$. Фазовая диаграмма для случая синхронизации на частоте $\omega < \omega_0$ изображена на рис. 357. Здесь изображающая точка проходит за цикл по спирали угловой путь $2\pi + \alpha > 2\pi$ и затрачивает на это время $T > T_0$. Подобного рода диаграммы позволяют ответить на вопрос о ширине полосы захватывания. Будем увеличивать угол φ , определяющий фазу синхронизирующего импульса относительно свободного колебания генератора. Наибольшее значение угла α , т.е. наибольшая расстройка $\Delta\omega = |\omega - \omega_0|$, будет, очевидно, при $|\varphi| = \pi/2$. Соотношения, получаемые при этом условии, поясняются рис. 358. Мы имеем

$$\operatorname{tg} \alpha = \frac{\omega_0 h}{A_0},$$

где A_0 — амплитуда автоколебаний, соответствующая предельному циклу. В действительности, α — малый угол (на рисунках быстрота свертывания спирали грубо преувеличена). Поэтому $\operatorname{tg} \alpha \approx \alpha$. Но α есть мера изменения периода (против периода автоколебаний, равного 2π в угловой мере).

Таким образом,

$$\frac{\alpha}{2\pi} = \frac{\Delta\omega}{\omega_0}$$

и, следовательно,

$$\frac{\Delta\omega}{\omega_0} \approx \frac{\omega_0}{2\pi} \frac{h}{A_0}.$$

Это есть наибольшая относительная расстройка в одну сторону. Полная же относительная ширина полосы захватывания равна

$$\frac{2\Delta\omega}{\omega_0} \approx \frac{\omega_0}{\pi} \frac{h}{A_0} \approx \frac{\omega h}{\pi A_0}.$$

Можно еще преобразовать этот результат, учитывая, что $\frac{1}{\pi}\omega h$ есть не что иное, как амплитуда первой гармоники в разложении периодической последовательности коротких импульсов в ряд Фурье. Эта составляющая, имеющая частоту ω , и производит синхронизирующее действие. Обозначая

$$\frac{1}{\pi}\omega h = E_1,$$

получим для полосы захватывания

$$\frac{2\Delta\omega}{\omega_0} \approx \frac{E_1}{A_0},$$

т.е. тот же результат, что и в § 98.

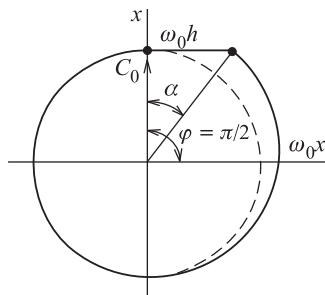


Рис. 358

Заметим в заключение, что связь между расстройкой и углом φ , очевидная из рис. 356 и 357, — это та же самая связь, которая определяется уравнениями (285).

§ 101. Синхронизация релаксационного генератора

Разберем очень простой механизм синхронизации релаксационного генератора внешним синусоидальным напряжением. В качестве генератора выберем схему с тиратроном (рис. 359). Она отличается от схемы релаксационного генератора с двухэлектродной газоразрядной лампой (рис. 326) тем, что напряжение зажигания U_2 зависит от напряжения U на сетке тиратрона T : оно тем выше, чем больше отрицательное напряжение на сетке. Напряжение потухания U_1 практически неизменно. Заметим, что в схему введен ограничитель тока OT , так что зарядный ток предполагается постоянным, а следовательно, напряжение на конденсаторе в процессе зарядки нарастает линейно.

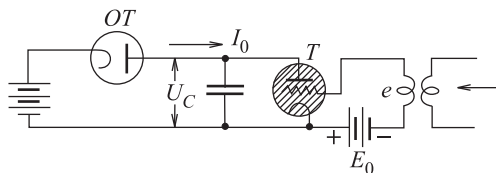


Рис. 359

При отсутствии внешнего напряжения процесс в генераторе протекает, как было описано в § 92: конденсатор заряжается до напряжения зажигания U_2 ; при этом значении напряжения возникает большая ионная проводимость, и конденсатор быстро (практически мгновенно) разряжается до напряжения потухания U_1 ; затем возобновляется зарядка конденсатора.

Включим теперь на сетку внешнее синхронизирующее синусоидальное напряжение

$$e = E_m \sin \omega t.$$

Напряжение зажигания U_2 будет теперь переменной величиной. Положим, что напряжение зажигания изменяется пропорционально напряжению на сетке. Будем иметь

$$u_2(t) = k(E_0 + E_m \sin \omega t) = U_0 + U_m \sin \omega t,$$

где U_0 — напряжение зажигания при отсутствии переменного внешнего напряжения (т. е. при постоянном напряжении e). Зажигание происходит при

$$u_c = u_2,$$

т. е. когда напряжение на конденсаторе достигнет напряжения зажигания. В этот момент происходит скачок напряжения. Напряжение на

конденсаторе растет во время зарядки по линейному закону

$$u_c = U_1 + \frac{I_0}{C}t.$$

График колебания построен на рис. 360. Как видно, это периодический процесс с периодом колебания, равным периоду внешнего напряжения T и неравным периоду свободных колебаний генератора T_0 . Иначе говоря, внешнее напряжение захватывает частоту генератора; происходит синхронизация генератора внешним напряжением.

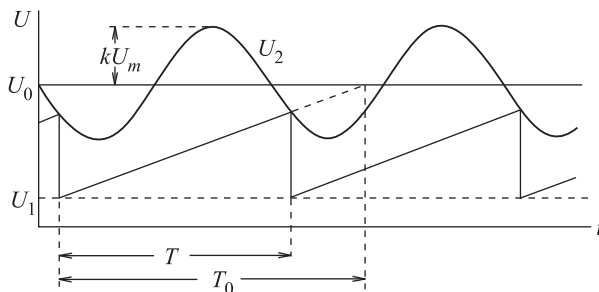


Рис. 360

Выведем основные соотношения для установившегося синхронизированного колебания. Если вести отсчет времени от скачка напряжения, то напряжение зажигания нужно записать в виде

$$u_2 = U_0 - U_m \sin(\omega t + \varphi).$$

Следующий скачок произойдет при

$$u_c = U_0 - U_m \sin(\omega T + \varphi) = U_0 - U_m \sin \varphi.$$

Но, с другой стороны, в момент $t = T$

$$u_c = U_1 + \frac{I_0}{C}T = U_1 + \alpha T.$$

Таким образом,

$$U_1 + \alpha T = U_0 - U_m \sin \varphi.$$

Воспользовавшись соотношением

$$U_1 + \alpha T_0 = U_0,$$

получим

$$\alpha(T_0 - T) = U_m \sin \varphi$$

или

$$\frac{T_0 - T}{T_0} = \frac{U_m}{U_0 - U_1} \sin \varphi. \quad (295)$$

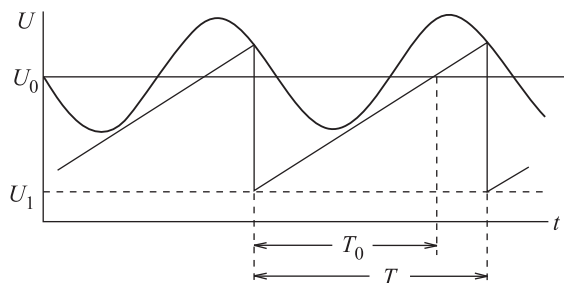


Рис. 361

При $\sin \varphi < 0$ период получает положительное приращение, т. е. увеличивается. Соответствующий график изображен на рис. 361. Если считать, что относительное изменение периода невелико, то его можно приравнять относительному изменению частоты, т. е. принять (см. (295))

$$\frac{\omega_0 - \omega}{\omega_0} \approx \frac{U_m}{U_0 - U_1} \sin \varphi.$$

Крайние значения $\sin \varphi$ равны ± 1 ; отсюда находим полную ширину полосы захватывания

$$\frac{2\Delta\omega}{\omega_0} \approx \frac{U_m}{\frac{1}{2}(U_0 - U_1)}.$$

Исследование устойчивости синхронизированного режима показывает, что условие устойчивости выражается через фазу синхронизации точно так же, как в § 97, а именно:

$$-\frac{\pi}{2} < \varphi < \frac{\pi}{2}.$$

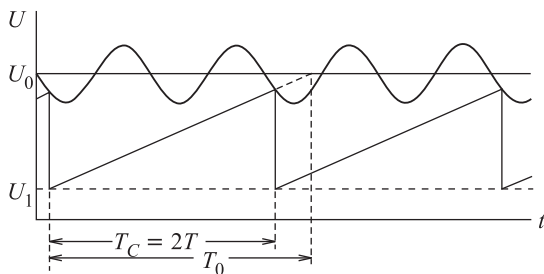


Рис. 362

Выведенные соотношения требуют еще уточнения. Во-первых, нужно заметить, что равенство

$$u_c = U_0 - U_m \sin \varphi,$$

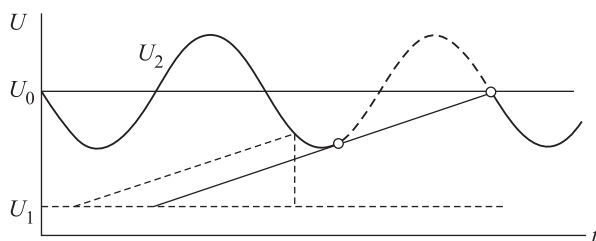


Рис. 363

из которого мы определяем момент скачка напряжения, может удовлетворяться не только при $t = T$, но и при $t = nT$, где n — любое целое число. Таким образом, в формуле (295) следует писать в общем случае

$$T_c = nT,$$

где T_c — период синхронизированных колебаний. На рис. 362 в пояснение сказанного приведен график синхронизированного колебания при $n = 2$. Это называется иногда *синхронизацией на субгармонике*. Во-вторых, при заданных U_0 , U_1 , α , U_m и T некоторые режимы неосуществимы, так как прямая U_c касается синусоиды U_2 и не может попасть на известный участок синусоиды. Это показано на рис. 363, на котором недостижимый участок синусоиды отмечен пунктиром.

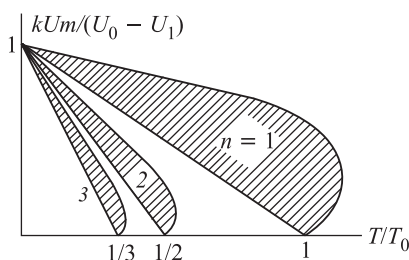


Рис. 364

Принимая все это во внимание, можно построить график, представляющий области возможных режимов синхронизации при различных кратностях n (рис. 364); эти области на рисунке заштрихованы.

ПАРАМЕТРИЧЕСКИЕ ЯВЛЕНИЯ

§ 102. Синхронное детектирование

Исследование синхронного детектирования начнем с простейшего примера синхронного выпрямления. Пусть имеется цепь, состоящая из источника переменного напряжения E , сопротивления нагрузки R и синхронного ключа K (рис. 365). Ключ имеет независимое управление, однако синхронное с частотой источника. Действие ключа состоит в том, что он попеременно то замыкает цепь на время, равное половине периода переменного напряжения, то размыкает ее на такое же время. Таким образом, проводимость цепи равна либо нулю, либо $1/R$, в зависимости от того, разомкнут или замкнут ключ. При таких обстоятельствах можно получить в цепи выпрямленный ток (рис. 366).

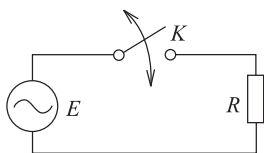


Рис. 365

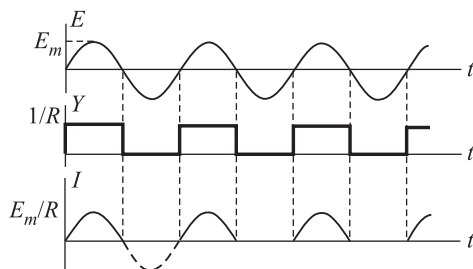


Рис. 366

Необходимо сразу же подчеркнуть, что такое выпрямление принципиально отличается от обычного. При обычном выпрямлении проводимость цепи есть также переменная величина, но она *зависит*

от напряжения на выпрямителе. Таким образом, схема обычного выпрямителя есть *нелинейная* цепь. В рассмотренном же здесь случае синхронного выпрямителя переменная *проводимость зависит от времени*, следовательно, цепь имеет переменные параметры и может быть описана *линейным* уравнением с переменными коэффициентами.

Особенность синхронного выпрямления заключается в том, что при независимой (хотя и синхронной) работе ключа фаза коммутации произвольна. Легко представить себе, что если фаза коммутации изменится на 180° , то изменится полярность выпрямленного напряжения на сопротивлении R . Если же сдвинуть фазу на 90° или 270° , то выпрямления вообще не произойдет (рис. 367). Как видим, при сдвиге на 90° , т.е. во времени на $T/4$, ток в цепи не имеет постоянной составляющей. Такого рода зависимость выпрямляющего действия от фазы характерна для синхронного выпрямления и для синхронного детектирования. Следствия, вытекающие из этого свойства, будем обсуждать дальше.

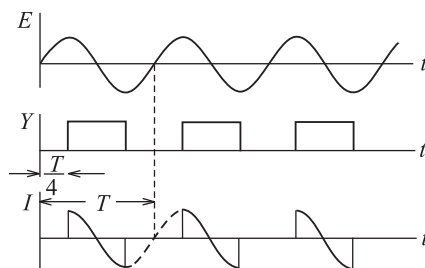


Рис. 367

Рассматривая синхронное выпрямление, мы полагали, что синусоидальное напряжение источника умножается на периодически изменяющуюся проводимость, график которой имеет вид прямоугольных импульсов (см. график функции Y на рис. 366 и 367). Это предположение было сделано только для наглядности, так как легко представить себе такой закон изменения проводимости, как результат периодического замыкания и размыкания ключа. Но для получения эффекта выпрямления такая зависимость проводимости от времени вовсе не обязательна. Проводимость может изменяться по любому периодическому закону, в частности по синусоидальному.

Все сказанное выше по поводу влияния фазы сохраняет силу для любого периодического закона.

Обратимся теперь к синхронному детектированию АМ колебания. Задача детектирования состоит в том, чтобы восстановить низкочастотный модулирующий сигнал. Пусть дано модулированное напряжение вида

$$e = E_0[1 + mf(t)] \sin \omega_0 t. \quad (296)$$

Подадим это напряжение на цепь, в которую включено переменное сопротивление (рис. 368), изменяющееся периодически с той же частотой ω_0 , так что проводимость цепи может быть записана в виде

$$Y = Y_0(1 + n \sin \omega_0 t), \quad (297)$$

т. е. предположим, что проводимость меняется синхронно и синфазно с напряжением несущей частоты. Для тока в цепи получим, перемножив (296) и (297),

$$i = eY = E_0 Y_0 \left\{ [1 + mf(t)] \sin \omega_0 t - \frac{n}{2} [1 + mf(t)] \cos 2\omega_0 t + \frac{n}{2} [1 + mf(t)] \right\}.$$

Первое слагаемое представляет собой исходное модулированное колебание с высокочастотным спектром, расположенным симметрично около частоты ω_0 . Второе слагаемое имеет аналогичный (но в $\pi/2$ раз меньший по плотности) спектр, расположенный около частоты $2\omega_0$. И, наконец, третье слагаемое является модулирующим низкочастотным сигналом. Высокочастотные слагаемые легко отделяются соответствующими фильтрами, которые должны предусматриваться в схеме любого детектора. Не менее легко освободиться и от постоянной составляющей. После такой очистки получим составляющую тока



Рис. 368

$$i = \frac{1}{2} nm E_0 Y_0 f(t), \quad (298)$$

что представляет желаемый результат детектирования.

Посмотрим теперь, как повлияет на результат нарушение синфазности. Пусть в отличие от предыдущего

$$Y = Y_0[1 + n \sin (\omega_0 t + \varphi)]. \quad (299)$$

Перемножая (296) и (299), найдем

$$i = eY = E_0 Y_0 [1 + mf(t)][1 + n \sin (\omega_0 t + \varphi)] \sin \omega_0 t.$$

Выполнив действия и отбросив высокочастотные и постоянную составляющие, найдем, что интересующая нас составляющая тока, содержащая сигнал $f(t)$, равна

$$i = \frac{1}{2} nm E_0 Y_0 f(t) \cos \varphi. \quad (300)$$

Как видим, выгоднее всего вести детектирование при $\varphi = 0$ (или π), при этом (300) переходит в (298). Если же $\varphi = \pm\pi/2$, то детектирования не получается.

Наконец, необходимо выяснить, что получится при нарушении условия синхронности. Здесь обнаружится, что синхронный детектор

обладает избирательностью. Для исследования этого свойства рассмотрим конкретную схему с простейшим фильтром в виде RC -звена (рис. 369). Предполагается, что сопротивление r , с которого снимается напряжение u_1 достаточно мало, так что присоединение параллельно ему звена RC не изменяет режима основной схемы синхронного выпрямителя. Мы имеем

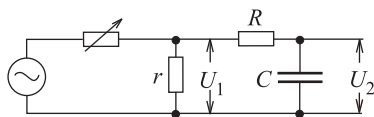


Рис. 369

$$u_1 = ir = eYr.$$

Положим, что

$$\begin{aligned} e &= E_0 \sin \omega_0 t, \\ Y &= Y_0(1 + n \sin \omega t), \end{aligned} \quad (301)$$

где ω — произвольная частота. Для комплексной амплитуды напряжения на выходе при некоторой частоте Ω имеем

$$\dot{U}_2 = \dot{U}_1 \frac{1}{1 + i\Omega RC}. \quad (302)$$

Найдем u_1

$$\begin{aligned} u_1 &= eYr = E_0 Y_0 r (1 + n \sin \omega t) \sin \omega_0 t = \\ &= E_0 Y_0 r \left[\sin \omega_0 t + \frac{n}{2} \cos(\omega_0 - \omega)t - \frac{n}{2} \cos(\omega_0 + \omega)t \right]. \end{aligned}$$

Таким образом, на входе звена RC имеем три напряжения с частотами ω , $\omega_0 - \omega$ и $\omega_0 + \omega$ и комплексными амплитудами соответственно $E_0 Y_0 r$, $i \frac{n}{2} E_0 Y_0 r$ и $-i \frac{n}{2} E_0 Y_0 r$.

Подставляя эти величины в формулу (302), получим следующие значения комплексных амплитуд выходного напряжения u_2 для трех перечисленных выше составляющих:

$$\frac{E_0 Y_0 r}{1 + i\omega_0 RC}, \quad \frac{i \frac{n}{2} E_0 Y_0 r}{1 + i(\omega_0 - \omega)RC}, \quad \frac{-i \frac{n}{2} E_0 Y_0 r}{1 + i(\omega_0 + \omega)RC}.$$

При небольшой расстройке, т. е. при $\omega \approx \omega_0$ и при большой постоянной времени, т. е. при

$$\omega_0 RC \gg 1,$$

можно пренебречь первой и третьей величинами по сравнению со второй, так что приблизительно

$$|\dot{U}_2| \approx \frac{E_0 Y_0 r n}{2RC} \frac{1}{\sqrt{(\omega_0 - \omega)^2 + \frac{1}{R^2 C^2}}}.$$

Но это выражение совпадает с выражением для обычной резонансной кривой. Особенность рассматриваемого случая состоит в том, что

роль затухания играет величина $1/RC$. Следовательно, чем больше постоянная времени RC , тем острее «резонансная кривая» синхронного детектора. Однако максимум чувствительности получается, когда разностная частота $\omega_0 - \omega$ обращается в нуль, т.е. при синхронном режиме.

Использовать асинхронный режим для приема как угодно модулированных сигналов нельзя. Действительно, пусть вместо синусоидального напряжения (301) имеется амплитудно-модулированное напряжение

$$e = E_0[1 + mf(t)] \sin \omega_0 t.$$

Составляя произведение

$$\begin{aligned} u_1 &= eYr = E_0Y_0r[1 + mf(t)](1 + n \sin \omega t) \sin \omega_0 t = \\ &= E_0Y_0r[1 + mf(t)] \left[\sin \omega_0 t + \frac{n}{2} \cos(\omega_0 - \omega)t - \frac{n}{2} \cos(\omega_0 + \omega)t \right], \end{aligned}$$

видим, что спектр напряжения u_1 воспроизводит спектр e на частотах ω_1 , $\omega_0 - \omega$ и $\omega_0 + \omega$ в качестве несущих, как показано на рис. 370, а, получается не детектирование, а преобразование частоты. При малых расстройках, когда $\omega_0 - \omega$ меньше ширины спектра модулирующей функции, возникают искажения. И лишь при $\omega = \omega_0$, т.е. при синхронном режиме, получается правильное детектирование (рис. 370, б).

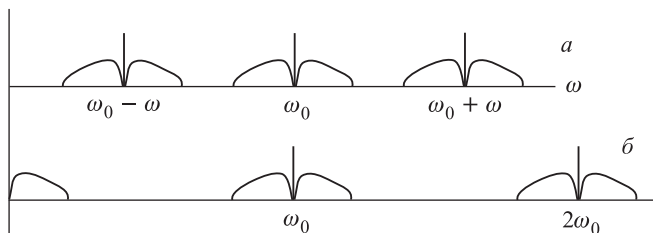


Рис. 370

В заключение заметим, что если входное напряжение e — периодическое, то, изменяя частоту ω , можно поочередно выделить и измерить гармонические составляющие e . При этом синхронный детектор работает, в сущности, как гармонический анализатор. Разрешающая способность такого анализатора, т.е. способность различить и измерить порознь близко расположенные на шкале частот гармонические составляющие, зависит от постоянной времени RC и может быть сделана очень большой.

§ 103. Фазовое детектирование и разделение

Отмечавшаяся в предыдущем параграфе чувствительность синхронного детектора к фазе детектируемого напряжения может иметь технические применения.

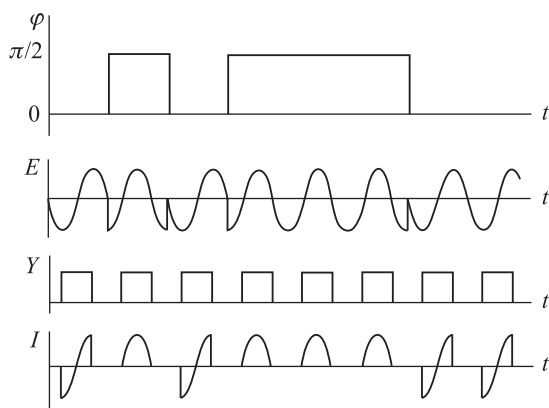


Рис. 371

Для начала рассмотрим прием радиотелеграфных сигналов при фазовой модуляции. Речь идет о сигнале, в котором при нажатии и отжатии телеграфного ключа фаза несущей частоты меняется скачком, как показано на рис. 371. График первой строки показывает изменение фазы при передаче буквы «А» кодом Морзе. Во второй строке изображено модулированное напряжение. В третьей строке дан график проводимости — предполагается, что применено простейшее устройство с синхронным ключом рис. 365. В четвертой строке изображен получаемый ток. После усреднения, т.е. фильтрации, мы воспроизведем с тем или иным приближением переданный телеграфный сигнал. Нужно сразу пояснить, что на рисунке на каждую телеграфную точку приходится только один период несущей частоты, что не соответствует действительности (такое соотношение выбрано только для того, чтобы не потерять наглядности рисунка). В действительности несущая частота может быть, скажем, в 10^4 раз больше частоты манипуляции, т.е. на одну точку приходится порядка 10^4 периодов несущей частоты. Если модулировать изменением фазы не на $\pi/2$, а на π , то получится картина, показанная на рис. 372 и не требующая пояснений. Итак, цепь с периодически изменяющейся проводимостью может применяться в качестве фазового детектора.

Рассмотрим теперь фазовое разделение. Задача разделения сигналов возникает при многоканальной связи, когда по одной линии передается одновременно несколько разных сигналов. На приемном конце многоканальной системы связи необходимо разделить поступающую смесь сигналов и выделить сигнал, принадлежащий каждому данному каналу. При частотной системе многоканальной связи для передачи сигналов, принадлежащих различным каналам, применяют отдельные несущие (или поднесущие) для каждого канала. При этом можно расположить спектры сигнала каждого канала на шкале частот так, что они не будут перекрываться. Разделение на приемном конце произ-

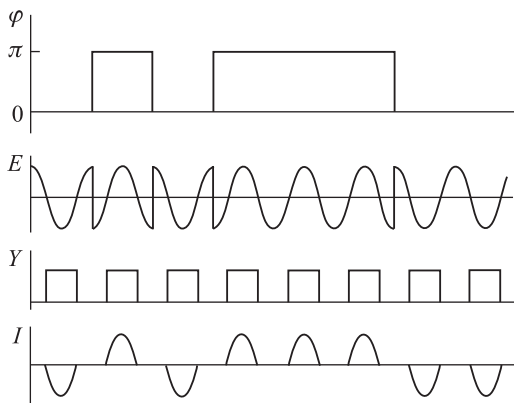


Рис. 372

водится обычными полосовыми фильтрами (так называемая «расфилтровка»). При фазовом разделении в качестве несущих применяются колебания одинаковой частоты, но сдвинутые по фазе. Рассмотрим вопрос о возможности фазового разделения аналитически.

Пусть сигнал одного канала есть $f_1(t)$, а другого — $f_2(t)$.

Промодулируем одним сигналом напряжение несущей частоты вида

$$e_1 = E_{01} \sin \omega_0 t,$$

а для другого сигнала возьмем несущее напряжение вида

$$e_2 = E_{02} \sin (\omega_0 t + \psi).$$

Тогда на приемном конце получим сложное модулированное напряжение

$$\begin{aligned} e &= E_{01}[1 + m_1 f_1(t)] \sin \omega_0 t + E_{02}[1 + m_2 f_2(t)] \sin (\omega_0 t + \psi) = \\ &= a(t) \sin \omega_0 t + b(t) \sin (\omega_0 t + \psi). \end{aligned} \quad (303)$$

Подадим это напряжение в цепь, проводимость которой меняется по закону

$$Y = Y_0(1 + n \sin \omega_0 t). \quad (304)$$

Для нахождения тока перемножим (303) и (304)

$$\begin{aligned} i &= eY = Y_0(1 + n \sin \omega_0 t)[a(t) \sin \omega_0 t + b(t) \sin (\omega_0 t + \psi)] = \\ &= Y_0 \left[a(t) \sin \omega_0 t + b(t) \sin (\omega_0 t + \psi) + \frac{n}{2} a(t) - \frac{n}{2} a(t) \cos 2\omega_0 t + \right. \\ &\quad \left. + \frac{n}{2} b(t) \cos \psi - \frac{n}{2} b(t) \cos \psi \cos 2\omega_0 t + \frac{n}{2} b(t) \sin \psi \sin 2\omega_0 t \right]. \end{aligned}$$

Как видим, кроме составляющих высоких частот, группирующихся около ω и $2\omega_0$, в составе тока имеются составляющие, пропорциональные $a(t)$ и $b(t)$. Однако нам нужно, чтобы в результате обработки модулированного напряжения приемным устройством сигналы $a(t)$ и $b(t)$ были выделены порознь. Предыдущая формула показывает, что мы достигнем желаемого результата при $\psi = \pi/2$. Тогда после отфильтровки высокочастотных составляющих получим

$$i = \frac{n}{2} Y_0 a(t).$$

Итак, для того чтобы при помощи синхронного детектирования разделить два канала, необходимо взять несущие колебания в виде:

$$e_1 = E_{01} \sin \omega_0 t, \quad e_2 = E_{02} \cos \omega_0 t.$$

На приемном конце нужно поставить два синхронных детектора, проводимости которых изменяются также по закону синуса и косинуса. При этих условиях получим на выходах фильтров разделенные сигналы $a(t)$ и $b(t)$, как показано на скелетной схеме рис. 373. Из всего сказанного следует, что описанным путем можно осуществить всего лишь двухканальную систему передачи. Если же изменение фазы используется в дополнение к любому другому способу разделения (например, частотному или временному), то возможное число каналов может быть удвоено.

Следует заметить в заключение, что осуществление синхронного детектирования и разделения наталкивается на ряд специфических технических трудностей, которые здесь обсуждать не будем.

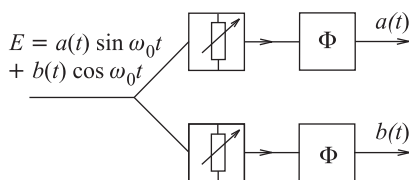


Рис. 373

§ 104. Параметрическое возбуждение

Рассмотрим механизм колебаний, могущих возникнуть в контуре, если какой-либо его параметр (например, емкость или индуктивность) периодически изменяется. Нужно подчеркнуть, что мы имеем здесь дело с явлением совершенно особого рода. Обычно колебания в контуре возникают при действии той или иной электродвижущей силы. В рассматриваемом же случае никаких электродвижущих сил нет — схема опыта показана на рис. 374.

Поэтому возникает вопрос, откуда берется энергия, необходимая для возбуждения и поддержания колебаний. Общие энергетические рассуждения и в данном случае позволяют раскрыть механизм явления.

Дело в том, что для изменения емкости (будем в дальнейшем рассматривать именно этот частный случай) необходимо, вообще говоря,

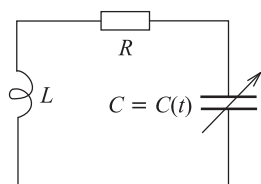


Рис. 374

совершить определенную работу против сил притяжения между обкладками. Представим себе, что колебания в контуре уже существуют и что на конденсаторе имеется переменное напряжение. Будем раздвигать обкладки в момент, когда напряжение максимально, и сближать их, когда напряжение равно нулю. Тогда за каждый цикл мы будем совершать определенную работу, т.е. вкладывать в систему некоторую энергию в механической форме.

В системе происходит преобразование механической энергии в электрическую, и таким образом осуществляется приток энергии в систему извне.

Для того чтобы пояснить механизм преобразования механической энергии в электрическую, напомним известные соотношения электростатики. Пусть изолированный конденсатор емкостью C заряжен до напряжения U между обкладками. Энергия, запасенная в электрическом поле конденсатора, равна

$$\mathfrak{E} = \frac{U^2 C}{2}.$$

При этом механическая сила, с которой обкладки притягиваются друг к другу, равна

$$F = \mathfrak{E}/l,$$

где l — расстояние между обкладками.

Если раздвинуть обкладки на Δl , то будет совершена механическая работа

$$\Delta A = F \Delta l = \mathfrak{E} \frac{\Delta l}{l}.$$

Соответствующее приращение должна получить и потенциальная энергия электрического поля. Перепишем выражение для энергии в виде

$$\mathfrak{E} = \frac{Q_0^2}{2C},$$

где $Q_0 = UC$ — начальный заряд, остающийся при описываемом опыте неизменным (так как конденсатор изолирован). Емкость конденсатора равна

$$C = \frac{\varepsilon S}{l},$$

где ε — диэлектрическая проницаемость; S — площадь обкладок. Используя это соотношение, получим

$$\mathfrak{E} = \frac{Q_0^2 l}{2\varepsilon S}.$$

Если изменится l , то энергия получит приращение

$$\Delta\mathfrak{E} = \frac{Q_0^2}{2\varepsilon S}\Delta l = \frac{Q_0^2}{2C}\frac{\Delta l}{l} = \mathfrak{E}\frac{\Delta l}{l},$$

равное, как видим, совершенной работе ΔA . Увеличение энергии конденсатора выразится в повышении напряжения на нем на величину

$$\Delta U = U\frac{\Delta l}{l}.$$

Таким образом, относительное изменение напряжения равно относительному изменению расстояния между обкладками ¹⁾.

Применим теперь энергетические соотношения к выводу условия возбуждения параметрических колебаний в контуре рис. 374.

Полная энергия $\mathfrak{E}$ в колебательном контуре равна сумме энергий, запасенных в магнитном поле индуктивности ($\mathfrak{E}_L$) и в электрическом поле емкости ($\mathfrak{E}_C$). При отсутствии потерь и внешнего воздействия запас энергии остается неизменным; энергия лишь периодически перекачивается из индуктивности в емкость и обратно, в чем, в сущности, и состоит процесс свободных колебаний. При наличии потерь, т. е. с учетом активного сопротивления R , запас энергии убывает. Потеря энергии $\mathfrak{E}_-$, отнесенная ко времени и усредненная за период, дает теряемую мощность, равную

$$P_- = \left(\overline{\frac{d\mathfrak{E}_-}{dt}} \right) = \frac{1}{2}I_m^2 R,$$

где черта означает усреднение за период, а

$$I_m = \omega_0 Q_m = \omega_0 C U_m$$

— амплитуда тока, которую можно также выразить через амплитуду заряда Q_m и через амплитуду напряжения на конденсаторе U_m . Таким образом,

$$P_- = \frac{1}{2}\omega_0^2 C^2 R U_m^2 = \omega_0^2 R C \mathfrak{E} = \omega_0 d \mathfrak{E}, \quad (305)$$

где d — затухание контура. Колебания могут быть поддержаны, если теряемая мощность восполняется извне. Это и достигается периодическим изменением емкости. Мы имеем

$$\mathfrak{E} = \mathfrak{E}_L + \mathfrak{E}_C.$$

¹⁾ Не следует забывать, что все выше приведенные зависимости получены из линеаризованных соотношений и представляют собой первое приближение, справедливое при $\frac{\Delta l}{l} \ll 1$.

Такой же порядок малости имеют $\frac{\Delta U}{U}$ и $\frac{\Delta \mathfrak{E}}{\mathfrak{E}}$.

Найдем приращение энергии, обусловленное только изменением емкости,

$$d_C \mathfrak{E} = \frac{\partial \mathfrak{E}_C}{\partial C} dC,$$

потому что $\mathfrak{E}_L$ не зависит от C .

Так как

$$\mathfrak{E}_C = \frac{u^2 C}{2},$$

то

$$d_C \mathfrak{E} = \frac{1}{2} u^2 dC.$$

Разделив обе части на dt , получим

$$\frac{d_C \mathfrak{E}}{dt} = \frac{1}{2} u^2 \frac{dC}{dt}. \quad (306)$$

Мы предполагаем, что имеется синусоидальное колебание, так что

$$u = U_m \cos \omega_0 t.$$

Подставляя это значение в (306), найдем

$$\frac{d_C \mathfrak{E}}{dt} = \frac{1}{4} U_m^2 (1 + \cos 2\omega_0 t) \frac{dC}{dt}. \quad (307)$$

Теперь заметим, что среднее значение $\frac{d_C \mathfrak{E}}{dt}$ есть не что иное, как мощность P_+ , поступающая в контур. Для того чтобы она не была равна нулю, необходимо в формуле

$$C(t) = C_0 + C_1 f(t)$$

положить

$$f(t) = \sin 2\omega_0 t,$$

а это означает, что емкость должна изменяться с частотой, вдвое большей частоты колебаний в контуре.

Тогда

$$\frac{dC}{dt} = 2\omega_0 C_1 \cos 2\omega_0 t. \quad (308)$$

Действительно, подставив (308) в (307), получим:

$$\begin{aligned} \frac{d_C \mathfrak{E}}{dt} &= \frac{1}{2} U_m^2 \omega_0 C_1 (\cos 2\omega_0 t + \cos^2 2\omega_0 t), \\ P_+ &= \left(\overline{\frac{d_C \mathfrak{E}}{dt}} \right) = \frac{1}{4} U_m^2 \omega_0 C_1 = \frac{1}{4} U_m^2 \omega_0 C_0 \frac{C_1}{C_0} = \frac{1}{2} n \omega_0 \mathfrak{E}. \end{aligned} \quad (309)$$

Для того чтобы колебания не затухли, необходимо, очевидно, выполнение неравенства

$$P_+ > P_-.$$

Отсюда с помощью формул (305) и (309) находим

$$n > 2d. \quad (310)$$

Итак, условие самовозбуждения параметрических колебаний состоит в том, что относительное изменение емкости $n = C_1/C_0$ должно быть больше удвоенного затухания контура.

Заметим, что мы рассматривали задачу о параметрическом возбуждении колебаний как линейную; мы нашли приближенное условие самовозбуждения при $\omega/\omega_0 = 2$, но ничего не можем пока сказать об установившейся амплитуде колебаний. Установившийся режим может быть определен только при учете нелинейности. Так как в схеме рис. 374 нелинейных элементов нет, то амплитуда колебаний теоретически может нарастать неограниченно. Практически в такого рода схеме напряжение быстро достигает таких больших значений, что пробиваются конденсаторы. Поэтому в схему приходится вводить нелинейные сопротивления специально для ограничения амплитуды колебаний.

В заключение укажем, что элементарный вывод, основанный на энергетических соотношениях, позволил ценою некоторых упрощений легко получить правильный результат, выражаемый формулой (310). Однако полная картина возбуждения параметрических колебаний сложнее. Оказывается, что можно возбудить колебания не только при отношении частоты изменения емкости ω к частоте колебаний ω_0 , равном двум, но и при других значениях этого отношения, а именно:

$$\frac{\omega}{\omega_0} = 2, 1, \frac{2}{3}, \frac{1}{2}, \dots, \frac{2}{k} \quad (k = 1, 2, 3, \dots).$$

При $k = 1$, т. е. при $\omega/\omega_0 = 2$, колебания возбуждаются легче всего, т. е. при наименьшем относительном изменении емкости. К этому случаю и относится формула (310). Общая теория параметрического возбуждения основывается на исследовании решений уравнения Матьё (см. приложение 4).

§ 105. Деление частоты ¹⁾

В радиотехнике часто бывает необходимо увеличить или уменьшить частоту колебания точно в некоторое — обычно целое — число раз. В первом случае говорят об умножении частоты, во втором — о делении. Число, характеризующее изменение частоты колебания, называется соответственно коэффициентом умножения или деления частоты: ниже оно обозначается через n .

Об умножении частоты говорилось ранее (§ 67); вернемся к нему сейчас, чтобы сопоставить его с процессом деления. В § 66 было показано, что любой нелинейный элемент, находящийся под синусои-

¹⁾ Этот параграф написан И. Х. Ризкиным.

дальним воздействием, создает гармоники частоты воздействия. Если характеристика нелинейного элемента, используемого для умножения, выражается функцией $y = f(x)$, где x — воздействие, а y — отклик, то при

$$x = A \cos \omega_{\text{вх}} t$$

будем иметь

$$y = \sum_k A_k \cos (k\omega_{\text{вх}} t + \varphi_k).$$

Иначе говоря, отклик будет суммой (конечной или бесконечной — в зависимости от характеристики $f(x)$) колебаний с частотами, кратными $\omega_{\text{вх}}$ — частоте входного воздействия.

Можно пояснить образование гармоник в подобном умножителе следующим образом. При прохождении синусоидального колебания через нелинейный элемент оно искажается, причем искажения, внесенные на протяжении одного периода, всегда одинаковы. Поэтому период отклика, т. е. величины y , будет равным (см. например рис. 231) или меньшим (см. например, рис. 234) периода воздействия. Ясно, что при

таком искажении возрасти период не может, следовательно, невозможно и деление частоты.

Для того чтобы поделить частоту, необходимо, очевидно, либо исказить некоторые периоды входного колебания, оставляя другие периоды вообще неискаженными (рис. 375), либо исказить колебание в течение разных периодов различным образом (рис. 375).

Очевидно, для того, чтобы по-разному исказить колебание в течение различных периодов, нужно

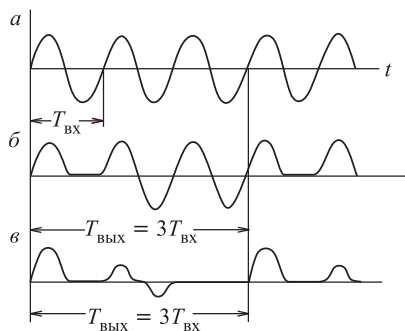


Рис. 375

прежде всего суметь эти периоды *различать*. Поэтому делитель частоты должен содержать некоторое устройство для различения периодов, которое и управляет процессом искажения входного колебания.

В реальных схемах делителей, основанных на описанном принципе, операция различения осуществляется благодаря устройству, которое запоминает, каковы были предыдущие периоды воздействия, и в зависимости от этого управляет искажением колебаний в течение последующих периодов. Иногда функции искажения и запоминания осуществляются одним и тем же узлом схемы.

Однако деление частоты можно осуществить и иным способом, не прибегая к явному искажению колебания делимой частоты. Способ этот состоит в том, что в состав делителя вводится некоторая колебательная система, создающая колебание с частотой, почти равной

поделенной. Таким образом, искаженное колебание уже как бы «заготовлено» заранее, но оно еще не совсем такое, как требуется, так как его частота лишь приблизительно, а не точно равна поделенной.

Сущность действия таких делителей заключается в том, что они как бы сравнивают это колебание с колебанием частоты, точно равной поделенной, и регулируют частоту первого до тех пор, пока она не примет требуемого значения.

Соответственно этим двум способам деления частоты можно различать две большие группы схем делителей.

Предыдущие рассуждения показывают, что реализация процесса деления частоты в радиотехнических устройствах связана, вообще говоря, с гораздо большими трудностями, чем умножение частоты. Поэтому не удивительно, что схемы делителей значительно сложнее схем умножителей.

Рассмотрим теперь некоторые схемы делителей.

Делители частоты с запоминающими устройствами. Напомним, что запоминающим называется устройство, которое может в течение некоторого отрезка времени — иногда неопределенно большого — хранить в той или иной форме сообщенную ему информацию. В ряде случаев эта информация представляет собой просто число, например, 0 или 1. Самый процесс хранения информации может иметь двоякий характер: динамический или статический.

В первом случае запасенная информация (например, импульс, символизирующий число 1) циркулирует в некоторой цепи в течение всего времени хранения. В другом случае процесс запоминания информации не связан с ее циркуляцией, передачей из одного узла в другой, преобразованием формы и т. д., а сводится к изменению статических состояний некоторых элементов. Примером статического запоминающего устройства является триггер. Динамическое запоминание осуществляется, например, линией задержки, в которой в течение некоторого времени «циркулирует» поданный на вход импульс.

Возвратимся теперь к делителям и на примере деления в два раза ($n = 2$) частоты следования импульсов покажем, как просто решается задача различения и искажения при применении динамического запоминающего устройства — линии задержки. Блок-схема делителя представлена на рис. 376, а. Здесь ЛЗ — линия задержки с временем за-

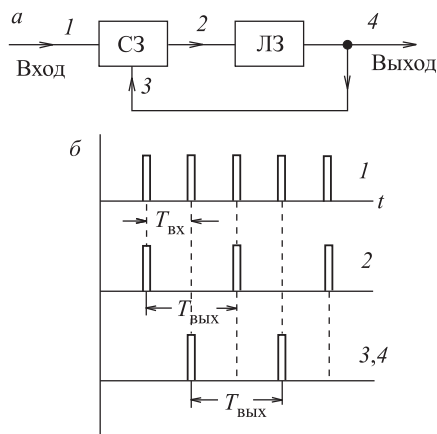


Рис. 376

держки $\tau = T_{\text{вх}}$, где $T_{\text{вх}}$ — период входного процесса, представляющего собой последовательность коротких импульсов; СЗ — так называемая «схема запрета», имеющая два входа (1 и 3), из которых второй называется запрещающим, и один выход (2). Импульсы подводятся к обоим входам и снимаются с выхода. Основное свойство СЗ состоит в том, что импульс на выходе 2 появляется лишь тогда, когда на входе 1 импульс имеется, а на входе 3 — отсутствует. Импульс на выходе 2 не появляется, либо если нет импульса на входе 1, либо если он есть, но одновременно имеется также импульс на входе 3. Таким образом, воздействие в точке 3 запрещает прохождение импульса со входа 1 на выход 2.

Ясно, что схема рис. 376, а действительно различает периоды воздействия и искажает входную последовательность импульсов так, что на выход проходит лишь каждый второй импульс: период на выходе $T_{\text{вых}} = 2T_{\text{вх}}$. На рис. 376, б изображены временные диаграммы для точек 1, 2 и 3 (или 4). Когда в точке 1 появился первый импульс входной последовательности, линия задержки «свободна», поэтому на запрещающем входе импульс отсутствует. Следовательно, входной импульс пройдет на выход 2 схемы запрета и появится на выходе ЛЗ через время $\tau = T_{\text{вх}}$. В этот же момент на входе 1 появится второй импульс входной последовательности, который не сможет пройти на выход 2 из-за запрещающего действия импульса на входе 3. Запасенный в ЛЗ импульс № 1 позволяет отличить первый импульс от второго и управляет соответствующим искажением входной последовательности. Очевидно, третий (или любой нечетный) импульс будет в положении первого, а четвертый (или любой четный) — в положении второго, и мы получаем:

$$T_{\text{вых}} = 2T_{\text{вх}}, \quad \omega_{\text{вых}} = \frac{1}{2}\omega_{\text{вх}}, \quad n = 2.$$

Равенство $\tau = T_{\text{вх}}$ не обязательно должно выполняться точно — иначе делитель работал бы лишь при одной частоте $\omega_{\text{вх}} = \frac{2\pi}{\tau}$. Дело в том, что СЗ работает и тогда, когда импульсы на входах 1 и 3 несколько сдвинуты друг относительно друга.

Примером делителя частоты со статическими запоминающими устройствами является описанная в § 96 триггерная цепочка.

Заметим, что, хотя выше рассматривалось лишь деление частоты следования импульсов, подобный метод может быть использован и для деления частоты синусоидального колебания. В частности, синусоидальное колебание можно сначала преобразовать в импульсное, а затем подвести его ко входу делителя.

Делители частоты со сравнением. Рассмотрим теперь устройства, использующие принцип сравнения. Для определенности будем иметь в виду синусоидальные колебания и схему делителя с автоподстройкой частоты — смысл этого термина выяснится ниже.

Блок-схема такого делителя приведена на рис. 377. Генератор Γ создает колебание частоты $\omega_0 \approx \frac{1}{n}\omega_{\text{вх}}$ и в отсутствие входного напряжения — с частотой $\omega_{\text{вх}}$. Сравнивающее устройство СУ сопоставляет колебания частот $\omega_{\text{вх}}$ и ω_0 и, в зависимости от результата этого сопоставления, создает электрический сигнал, воздействующий на элемент УЭ, управляющий частотой ω_0 генератора Γ . При точном равенстве частот ω_0 и $\frac{1}{n}\omega_{\text{вх}}$ сигнал на выходе СУ таков, что УЭ не меняет частоты ω_0 . Отклонение частоты ω_0 от значения $\frac{1}{n}\omega_{\text{вх}}$ сразу приведет в действие управляющий элемент, который скорректирует это отклонение.

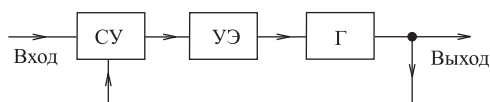


Рис. 377

Управляющий элемент УЭ может быть устроен по-разному; это может быть, например, реактивная лампа. Сравнивающие устройства также разнообразны; один из способов их построения состоит в том, что частота ω_0 первоначально умножается в n раз, а затем в специальном преобразователе частоты образуется комбинационная (разностная) частота $\omega_{\text{вх}} - n\omega_0$. Возникновение колебаний комбинационной частоты является важнейшим моментом всего процесса. Остальные спектральные составляющие, порождаемые в СУ, отсеиваются фильтром нижних частот. Пока $\omega_{\text{вх}} - n\omega_0 \neq 0$ или, что то же самое, пока $\omega_0 \neq \frac{1}{n}\omega_{\text{вх}}$ на выходе СУ имеется сигнал с переменной составляющей, которая и приводит в действие УЭ. При $\omega_0 = \frac{1}{n}\omega_{\text{вх}}$ комбинационная частота обращается в нуль, на выходе фильтра имеется лишь постоянная составляющая, так что процесс регулирования частоты генератора прекращается.

Итак, в делителе со сравнением искаженный сигнал, как уже говорилось, заранее присутствует благодаря наличию автоколебательной системы — в данном случае генератора Γ . Этот заготовленный сигнал не вполне совпадает с требуемым, так как $\omega_0 \neq \frac{1}{n}\omega_{\text{вх}}$. Сравнивая имеющийся сигнал с требуемым в СУ, схема управляет искажением колебания генератора, приближая ω_0 к $\frac{1}{n}\omega_{\text{вх}}$. Результатом сравнения является колебание с частотой $\omega_{\text{вх}} - n\omega_0$.

Рассмотренные делители частоты с автоподстройкой позволяют поделить частоту в десятки и сотни раз и широко используются на практике. Некоторые их варианты могут применяться и для деления частоты импульсных колебаний.

Примером деления частоты методом сравнения является рассмотренная в конце § 101 синхронизация релаксационного генератора на субгармонике.

Другой возможный подход к делению частоты методом сравнения обсуждается в следующем параграфе.

§ 106. Регенеративное деление частоты

Рассмотрим скелетную схему рис. 378. На вход подается напряжение u_1 . Оно поступает на перемножитель Π , к которому подводится также напряжение u_4 . Действие перемножителя определяется тем, что выходное напряжение u_2 пропорционально произведению двух входных напряжений u_1 и u_4 . Напряжение u_2 фильтруется фильтром Φ , усиливается усилителем $У$ и поступает на выход. Выходное напряжение u_3 через фазовращатель $\PhiВ$ подается обратно на перемножитель. Схемы типа рис. 378 называются *регенеративными делителями*.

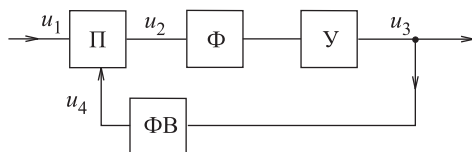


Рис. 378

Покажем, что в такой схеме возможно деление частоты в отношении $1 : 2$. Пусть на входе подано напряжение

$$u_1 = \sin 2\omega t.$$

Предположим, что деление частоты происходит, так что напряжение u_4 имеет частоту, вдвое меньшую частоты напряжения u_1 , т. е.

$$u_4 = \sin(\omega t + \varphi),$$

где φ — неизвестная пока фаза. Тогда

$$u_2 = u_1 u_4 = \sin 2\omega t \sin(\omega t + \varphi) = \frac{1}{2} \cos(\omega t - \varphi) - \frac{1}{2} \cos(3\omega t - \varphi).$$

Фильтр погасит составляющую тройной частоты, так что на выходе фильтра получим напряжение $\frac{1}{2} \cos(\omega t - \varphi)$.

Пусть усилитель дает усиление по напряжению вдвое и не вносит дополнительного фазового сдвига.

Тогда

$$u_3 = \cos(\omega t - \varphi).$$

Если фазовращатель поворачивает фазу напряжения на ψ , то имеем

$$u_4 = \cos(\omega t - \varphi + \psi).$$

Приравнявая два выражения для одного и того же напряжения u_4 , находим, что должно быть

$$2\varphi = \psi + 90^\circ.$$

Это условие выполняется, в частности, при $\varphi = 0$, $\psi = -90^\circ$.

Приведенное рассуждение показывает лишь, что деление частоты при помощи регенеративного делителя возможно. Но это рассуждение не позволяет найти условия, при которых деление действительно происходит. Оно не позволяет также обнаружить существенные черты явления, одна из которых состоит в том, что деление начинается только при вполне определенном значении входного напряжения u_1 . Для того чтобы разобраться в этих важных подробностях, рассмотрим какую-либо конкретную схему, например показанную на рис. 379. В этой

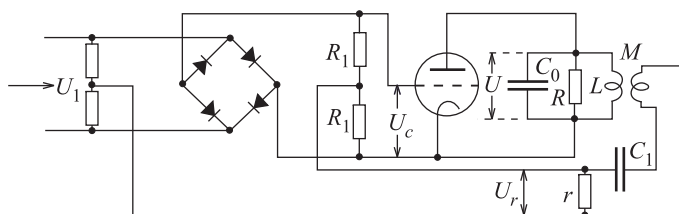


Рис. 379

схеме в качестве перемножителя применен кольцевой модулятор (см. § 76); будем полагать, что он производит чистое перемножение поданных на него напряжений. Эта схема описывается, вообще говоря, довольно сложным нелинейным уравнением с переменными коэффициентами. Ограничимся линейным приближением. При таком упрощении мы теряем возможность исследования установившегося режима деления, в частности, определения амплитуды напряжения поделенной частоты и ширины полосы деления. Но можем найти условие возникновения процесса деления. Составим уравнение токов для контура в анодной цепи триода

$$C_0 \frac{du}{dt} + \frac{1}{R} u + \frac{1}{L} \int u dt = S u_c, \quad (311)$$

где S — крутизна. Эдс, наводимая в катушке обратной связи, равна

$$e = M \frac{di_L}{dt} = \frac{M}{L} u. \quad (312)$$

Эта эдс действует в цепи $C_1 r$; напряжение на r равно

$$u_r \approx r C_1 \frac{de}{dt} \quad (313)$$

(при условии, что напряжение на емкости много больше U_1).

Комбинация rC_1 выполняет роль фазовращателя. Напряжение подается на перемножитель; это и есть то напряжение, которое в предыдущем было обозначено через u_4 . Далее, напряжение, снимаемое с кольцевого модулятора, равно (см. § 76)

$$u_2 = u_c = 4R_1 a_2 u_1 u_4. \quad (314)$$

Из (312), (313) и (314) получаем

$$u_c = 4a_2 r R_1 C_1 \frac{M}{L} u_1 \frac{du}{dt} = k u_1 \frac{du}{dt}.$$

Подставляя это значение в уравнение (311) и объединяя члены, содержащие $\frac{du}{dt}$, получим уравнение с переменным коэффициентом

$$(C_0 - kS u_1) \frac{du}{dt} + \frac{1}{R} u + \frac{1}{L} \int u dt = 0.$$

Это — уравнение токов в контуре, в котором емкость изменяется в соответствии с входным напряжением. Если входное напряжение изменяется по закону

$$u_1 = U_m \sin 2\omega t,$$

то для действующей емкости имеем

$$C(t) = C_0 - kS U_m \sin 2\omega t = C_0 \left(1 - \frac{kS}{C_0} U_m \sin 2\omega t \right).$$

Теперь становится ясным, что механизм регенеративного деления частоты сводится к параметрическому возбуждению колебаний. Условие возбуждения

$$n > 2d$$

в нашем случае выглядит так:

$$\frac{kS}{C_0} U_m > 2d$$

или

$$U_m > \frac{1}{2} \frac{d}{a_2 \omega_0^2 r R_1 C_1 M S}. \quad (315)$$

Напомним смысл входящих в эту формулу величин: ω_0 — резонансная частота контура, равная $(LC_0)^{-1/2}$; d — затухание контура; a_2 — коэффициент при u^2 в выражении для характеристики нелинейного элемента кольцевого модулятора; R_1 — сопротивление нагрузки модулятора (при выводе формулы (314) это сопротивление предполагалось малым по сравнению с внутренним сопротивлением самого модулятора), r и C_1 — сопротивление и емкость в цепи обратной связи (напомним, что должно быть $rC_1 \ll 1/\omega_0$), M — коэффициент взаимной индукции.

Формула (315) позволяет найти минимальную амплитуду входного напряжения, при которой схема делителя приходит в действие.

В предыдущем анализе мы стремились подобрать наиболее благоприятную фазу обратной связи, для чего в цепь обратной связи был введен соответствующий фазовращатель. Это не обязательно; делитель будет работать и при других фазовых соотношениях, хотя для возбуждения потребуются соответственно увеличенное входное напряжение. В практических схемах специальный фазовращатель обычно не ставится.

Возможны многочисленные варианты схем регенеративных делителей. Укажем здесь на существенное видоизменение схемы рис. 378, приводящее к делителю, способному выполнить деление частоты в любое число раз. Для этого в цепь обратной связи вводится умножитель частоты в $n + 1$ раз (рис. 380). При работе делителя на выходе получается частота ω , если на вход подается частота $n\omega$. На выходе умножителя в цепи обратной связи получается частота $(n + 1)\omega$. Эта частота попадает на вход перемножителя. При перемножении колебаний с частотами ω и $(n + 1)\omega$ получаются: суммарная частота $(2n + 1)\omega$ и нужная разностная частота $(n + 1)\omega - n\omega = \omega$, которая выделяется фильтром и поступает на выход.

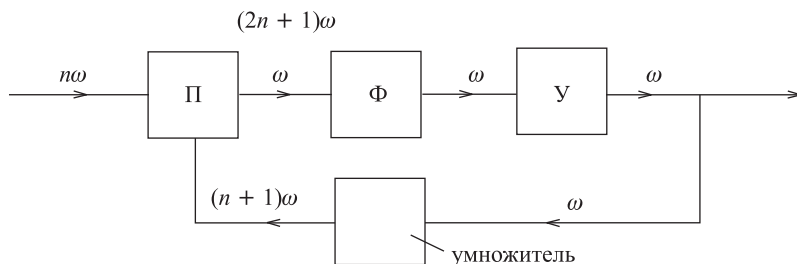


Рис. 380

Возможно, конечно, получить деление высокой кратности и другим путем, а именно, включая делители друг за другом ступенями. При таком использовании делителей, делящих, например, частоту в отношении $1 : 2$, можно получить деление в отношении $1 : 2^m$, где m — число ступеней.

§ 107. Усиление как параметрический процесс

На протяжении десятков лет главным видом усилителя был ламповый усилитель, основанный на применении вакуумного триода. Но за последнее время появился ряд новых видов усилителей, использующих самые различные физические явления. Явления эти настолько разнообразны, что изучение их особенностей заслоняет основные принципы усиления. Возникает потребность в более общем описании свойств усилителей, позволяющем выделить основные черты процесса усиления. Условимся говорить только об усилителях мощности. Конечно, так

называемые усилители напряжения или усилители тока имеют свою специфику, но для наших целей рассмотрение связанных с ней вопросов излишне; к тому же названные усилители в принципе могут рассматриваться как вырожденные частные случаи усилителя мощности.

Итак, определим сначала усилитель. Усилителем будем называть четырехполюсник (устройство, имеющее вход и выход), отклик которого (на выходе) пропорционален воздействию (на входе)¹⁾, причем мощность отклика больше мощности воздействия.

Мощность на выходе может быть больше мощности на входе, разумеется, только за счет добавления мощности от стороннего источника (источника питания). Этот источник является неотъемлемой составной частью усилителя. Процесс усиления сводится по существу к управлению потоком энергии, поступающим от источника. Входное воздействие выполняет роль управляющего сигнала.

Механизм управления состоит в том, что те или иные параметры цепи, соединяющей источник с выходом усилителя, изменяются во времени в соответствии с изменениями входного воздействия. Таким образом, всякий усилитель в принципе является системой с переменными параметрами или, по принятой нами терминологии, параметрической системой. Самый же процесс усиления является параметрическим процессом.

Рассмотрим вкратце энергетику усилителя.

Введем обозначения: P_1 — входная мощность; эта мощность расходуется на управление; P_2 — выходная мощность; P_0 — мощность, отдаваемая источником питания; P_n — мощность потерь (превращаемая в тепло в элементах схемы усилителя); P_1 и P_0 — составляют приход мощности, а P_2 и P_n — ее расход. По закону сохранения энергии

$$P_0 + P_1 = P_2 + P_n. \quad (316)$$

Коэффициент полезного действия (кпд) усилителя определяется как отношение выходной мощности ко всей мощности, поступающей в усилитель, т. е.

$$\eta = \frac{P_2}{P_0 + P_1}.$$

Обычно

$$P_1 \ll P_0$$

и можно принять для КПД приближенное выражение

$$\eta \approx \frac{P_2}{P_0}.$$

¹⁾ Условие пропорциональности не выполняется в точности в реальном усилителе; однако вопрос об искажениях мы здесь не рассматриваем.

Из (316) видно, что кпд всегда меньше единицы. Действительно,

$$\frac{P_2 + P_n}{P_0 + P_1} = 1; \quad \eta = 1 - \frac{P_n}{P_0 + P_1} \approx 1 - \frac{P_n}{P_0}.$$

Важнейшей характеристикой усилителя является коэффициент усиления. Определим его просто как отношение выходной мощности к входной

$$\zeta = \frac{P_2}{P_1}.$$

Этот коэффициент всегда больше единицы (по определению усилителя). Кпд и коэффициент усиления друг от друга не зависят. При данной выходной мощности коэффициент усиления тем больше, чем меньше мощность, затрачиваемая на управление переменными параметрами усилителя, а кпд тем больше, чем меньше мощность потерь.

Рассмотрим теперь разновидности усилителей по способу управления и питания. Простейший усилитель можно схематически представить в виде рис. 381. На этом рисунке R — переменное сопротивление, Y — управляющий механизм, изменяющий сопротивление R в соответствии с изменениями входного напряжения u_1 (предполагается, что механизм Y потребляет мощность P_1), E — источник *постоянного* напряжения, R_n — сопротивление нагрузки, с которого снимается выходное напряжение u_2 .

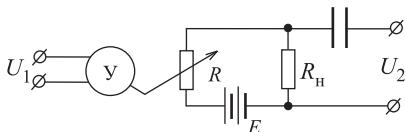


Рис. 381

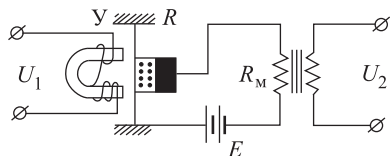


Рис. 382

Одним из первых усилителей по схеме рис. 381 является *микрофонный усилитель*, применявшийся задолго до изобретения электронных ламп. Этот усилитель представляет собой комбинацию телефона с угольным микрофоном. Входное напряжение питает обмотку электромагнита, который притягивает стальную диафрагму (рис. 382). Диафрагма, колеблясь, сжимает угольные зерна и изменяет их сопротивление, вследствие чего изменяется ток в цепи микрофона и батареи.

К этому же виду относится и обычный усилитель на вакуумном триоде. Эта схема (рис. 383) не требует пояснений.

К числу новых типов усилителей с управляемым сопротивлением относится *криотрон*. Действие криотрона основано на влиянии магнитного поля на сверхпроводящие свойства металлов при очень низких температурах. Устройство криотрона очень просто. На металлический стержень наложена управляющая обмотка, создающая магнитное поле. Стержень включен в цепь источника постоянного напряжения. Устрой-

ство поддерживается при температуре, близкой к температуре перехода в сверхпроводящее состояние. Эта температура зависит от магнитного поля, а потому при его увеличении металл стержня переходит из сверхпроводящего состояния в обычное, его проводимость очень сильно изменяется. Соответственно меняется и падение напряжения на стержне. В качестве материала стержня удобно применить тантал, так как температура перехода для него близка к 4,2 К — температуре жидкого гелия.

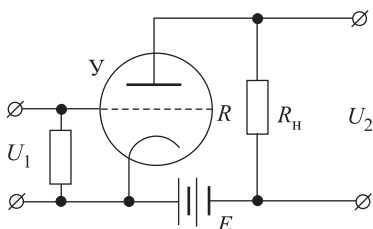


Рис. 383

Характеристики криотрона настолько круты, что он применяется главным образом не как обычный пропорциональный усилитель, а как устройство релейного действия.

Усиление по схеме рис. 381 можно назвать «прямым», в отличие от другого способа, который назовем «модуляционным» или «усилением на несущей частоте». Схема такого рода усилителя показана на рис. 384.

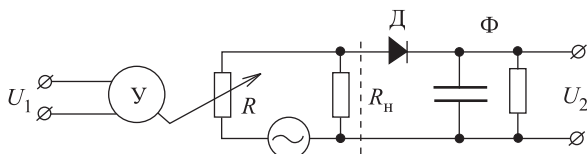


Рис. 384

Эта схема отличается от схемы рис. 381 прежде всего тем, что источник *постоянной* эдс заменен источником *переменной* эдс (например, синусоидальной).

В результате изменения сопротивления R в цепи возникает *модулированный* по амплитуде ток, а с нагрузочного сопротивления снимается модулированное напряжение. Чтобы получить выходное напряжение, подобное входному, нужно это модулированное напряжение протектировать, для чего в схеме предусмотрен детектор D и фильтрующее RC -звено Φ с надлежащим образом выбранной постоянной времени. Таким образом, схема усилителя отчетливо разделяется на две части: слева от вертикального пунктира (рис. 384) — модулятор, справа — детектор¹⁾. Система модулятор–детектор образует усилитель; точнее говоря, эта система *может* обладать коэффициентом усиления, большим единицы.

¹⁾ Часто требуется получить усиление на высокой частоте; в этом случае детектирующая часть схемы рис. 384 может отсутствовать.

Приведем в качестве примера схему давно известного и широко применяемого магнитного усилителя, один из вариантов схемы которого приведен на рис. 385.

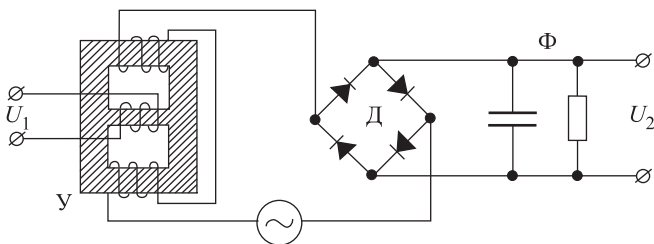


Рис. 385

В этой схеме управляемым параметром является индуктивность. Механизм управления состоит в том, что управляющая обмотка, наложенная на средний стержень стального сердечника, создает магнитный поток, изменяющий проницаемость сердечника, а следовательно, и индуктивность рабочих обмоток. Эти последние наложены на крайние стержни сердечника и питаются источником синусоидального напряжения. В качестве детектора Д на рис. 385 изображена мостовая схема, что не имеет, конечно, принципиального значения.

Прежде чем приводить другие примеры, остановимся на вопросе о сущности механизма управления. Этот механизм может быть двух типов. Первый характеризуется тем, что управляемый параметр X изменяется пропорционально управляющему напряжению u_1 , действующему в отдельной цепи, так что

$$X = ku_1.$$

Примером может служить обычный триод, в котором сопротивление катод-анод изменяется пропорционально напряжению катод-сетка.

Второй тип механизма управления основан на использовании элементов с нелинейными характеристиками. В этом случае напряжение питания и управляющее напряжение действуют в *одной и той же* цепи. Роль управляющего напряжения сводится к смещению рабочей точки на нелинейной характеристике.

Напомним, что в нелинейном случае проводимость, емкость и индуктивность определяются дифференциальными выражениями:

$$G = \frac{di}{du}, \quad C = \frac{dq}{du}, \quad L = \frac{d\psi}{di}.$$

Здесь u — напряжение; i — ток; q — заряд; ψ — потокосцепление, т. е. $\psi = n\Phi$, где Φ — магнитный поток, n — число витков.

Рассмотрим нелинейную проводимость. Пусть вольтамперная характеристика задана степенным рядом

$$i = f(u) = a_0 + a_1 u + a_2 u^2 + \dots$$

Тогда проводимость

$$G = \frac{di}{du} = f'(u) = a_1 + 2a_2 u + \dots,$$

т.е. зависит от напряжения. Если вольтамперная характеристика квадратична, то проводимость зависит от напряжения линейно. Если постоянное напряжение питания u_0 и управляющее напряжение u_1 действуют в одной и той же цепи, то:

$$u = U_0 + u_1, \quad G = a_1 + 2a_2 U_0 + 2a_2 u_1 = G_0 + \alpha u_1,$$

т.е. проводимость линейно зависит от управляющего напряжения.

Пусть питающее напряжение u_0 не постоянно, а переменное, например синусоидально («модуляционный» способ усиления). При квадратичной характеристике

$$i = f(u) = f(u_0 + u_1) = a_0 + a_1(u_0 + u_1) + a_2(u_0^2 + u_1^2 + 2u_0 u_1),$$

и мы видим, что в составе тока появляется член, пропорциональный произведению $u_0 u_1$, который и представляет модулированное колебание. Излишние члены, зависящие от u_0 и u_1 и их квадратов, могут быть устранены применением балансных и кольцевых схем модуляторов.

Вернемся к схеме магнитного усилителя (рис. 385). Здесь используется нелинейная индуктивность, т.е. нелинейная зависимость потокоцепления от тока.

Мы имеем

$$L = \frac{d\psi}{di} = n^2 S \frac{dB}{dH},$$

где S — площадь сечения сердечника; $\Phi = BS$; B — индукция; $H = ni$ — магнитодвижущая сила.

Итак, в конечном счете нелинейность определяется непостоянством магнитной проницаемости

$$\mu = \frac{dB}{dH}.$$

На рис. 386 показана нелинейная характеристика $B = f(H)$, т.е. кривая намагничивания. Величина магнитной проницаемости выражается наклоном касательной к кривой намагничивания в рабочей точке. Магнитодвижущие силы, создаваемые рабочими обмотками, питаемыми источником e , и управляющей обмоткой, питаемой входным напряжением u_1 , действуют в одной и той же магнитной цепи. Магнитодвижущая сила H_1 зависящая от u_1 , смещает рабочую точку и изменяет индуктивность. Отметим, что наличие гистерезиса не меняет существа дела, но сильно осложняет явления и их описание.

Рассмотрим еще несколько примеров. Наряду с управляемой индуктивностью для построения модуляционного усилителя может применяться управляемая емкость. Таковы усилители с использованием так называемых *варикондов* — конденсаторов с нелинейной зависимостью между напряжением и зарядом. Нелинейность в этих конденсаторах определяется свойствами диэлектрика. Нужно заметить, что мы получим нелинейный конденсатор и в том случае, когда диэлектрическая проницаемость диэлектрика постоянна, а от напряжения зависят какие-либо другие параметры, определяющие емкость. Так, например, воздушный конденсатор с гибкими обкладками нелинеен, так как с изменением напряжения меняется сила притяжения между обкладками, а следовательно, и зазор между ними. В этом примере приложенное напряжение изменяет геометрию конденсатора. Практического значения такой нелинейный конденсатор не имеет, так как инерция механических частей не позволяет применять его при быстро изменяющихся напряжениях ¹⁾.

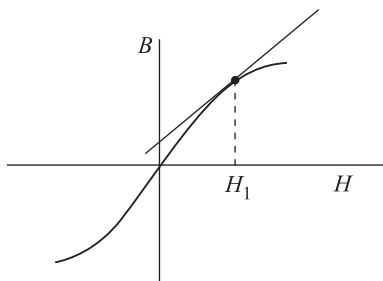


Рис. 386

За последнее время появились усилители, в которых в качестве управляемого параметра используется емкость запорного слоя в кристаллическом диоде. Запорный слой образуется на $p-n$ переходе в результате рекомбинации носителей зарядов противоположного знака, имеющих по обе стороны перехода, и проникающих в пограничную зону вследствие диффузии. Роль обкладок конденсатора играют проводящие p - и n -области. Расстояние между обкладками определяется толщиной запорного слоя, которая зависит от приложенного напряжения. Если плюс подан на p -область, а минус — на n -область, то толщина слоя убывает; при обратной полярности толщина слоя возрастает.

Существенно, что емкость запорного слоя не зависит от частоты, по крайней мере, вплоть до очень высоких частот, порядка 10^{12} Гц. Емкость запорного слоя можно выразить формулой

$$C = a \left(1 - \frac{u}{\varphi} \right)^{-b},$$

где u — приложенное напряжение; φ — контактная разность потенциалов (около 0,5 В); a и b — постоянные, зависящие от типа диода

¹⁾ Говоря более определенно: механические свойства системы не сказываются лишь в области частот ниже частот механического резонанса; последние лежат обычно в диапазоне звуковых частот.

($b < 1$). Относительное изменение емкости равно

$$\frac{\Delta C}{C} \approx b \frac{\Delta u}{\varphi - u} \approx -b \frac{\Delta u}{u},$$

т. е. при $\varphi \ll u$ и $\frac{\Delta u}{u} \ll 1$ может считаться пропорциональным изменению напряжения.

В связи со специфическими трудностями усиления на свч в этом диапазоне получили применение усилительные устройства, основанные на взаимодействии электронного пучка с электромагнитным полем — квазистационарным или волновым.

К числу простейших устройств этой группы относится *клистрон*. Принцип его действия вкратце сводится к следующему. Электронный пучок проходит через управляющее электрическое поле, действующее вдоль направления движения электронов, и, следовательно, ускоряющее или замедляющее электроны, в зависимости от знака поля. По выходе из зоны действия управляющего поля электронный пучок оказывается *модулированным по скорости*. Затем пучок вступает в *пространство дрейфа* (или пространство группирования), где внешние поля на электрон не воздействуют; здесь более быстрые электроны догоняют более медленные и образуются «электронные сгустки». Таким образом, модуляция электронного пучка по скорости сама собой преобразуется в пространстве дрейфа в *модуляцию по плотности*. Модулированный по плотности электронный пучок попадает затем в резонатор и возбуждает его, отдавая часть своей энергии.

Как видим, и в этой схеме ясно выступают все основные черты параметрического механизма. Источник питания — в данном случае источник постоянного ускоряющего напряжения — отдает свою мощность электронному пучку. Управляющее напряжение модулирует пучок, так что мощность пучка получает переменную составляющую. Эта переменная мощность больше мощности, затрачиваемой на управление, так как она получена за счет мощности источника питания.

Следует отметить, что усиление ведется в узкой полосе, так как и создание управляющего поля, и извлечение усиленной мощности из модулированного пучка осуществляется резонаторами высокой добротности. Первый резонатор называется обычно группирователем, второй — улавливателем.

Устройство двухрезонаторного клистрона показано схематически на рис. 387. На этом рисунке 1 и 2 — первый и второй резонаторы, 3 — катод, 4 — коллектор (т. е. электрод, на которой попадают отработанные электроны), 5 — трубка дрейфа, 6 и 7 — входная и выходная петли, 8 — ускоряющий электрод.

Процесс группирования, т. е. образование электронных сгустков, поясняется рис. 388. На этом рисунке в нижней части изображено поле в группирователе. В верхней части рисунка графиком в координатах расстояние — время изображено движение электронов. Предполагается

для простоты, что электроны выходят из катода через равные промежутки времени. Получив одинаковые скорости после прохождения ускоряющего поля, электроны поступают в группирователь, где их скорости увеличиваются или уменьшаются, в зависимости от знака поля в группирователе (см. синусоиду в нижней части рисунка). График движения электронов, выходящих из группирователя, представляется прямыми с разными наклонами; более крутые линии соответствуют более быстрым электронам. Благодаря различному наклону прямые образуют сходящиеся пучки, что и является графическим пояснением формирования электронных сгустков.

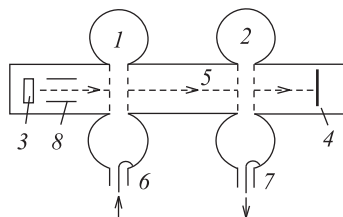


Рис. 387

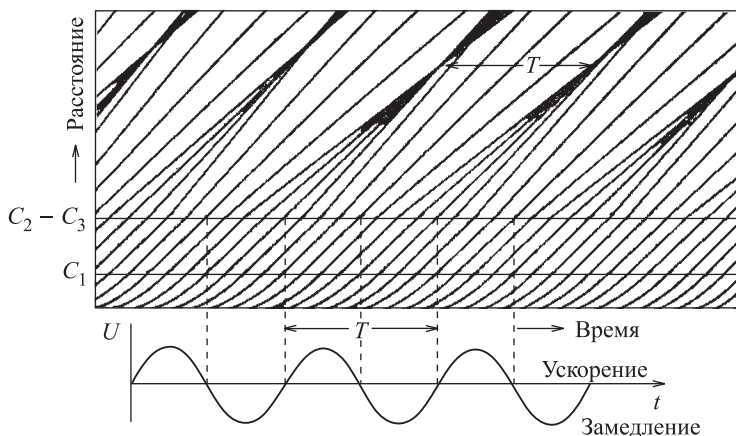


Рис. 388

Для клистрона характерно, что взаимодействие электрона с управляющим полем кратковременно. Электрон проходит через группирователь за время, малое по сравнению с периодом колебания. Другими словами, управляющее поле не успевает заметно измениться за время его воздействия на электрон, и мы имеем здесь дело с квазистационарным полем. Для получения большего усиления желательно было бы увеличить время взаимодействия, но для этого нужно, чтобы на протяжении пути электрона на него воздействовало поле неизменной фазы. Этого можно достичь, заставив электрон двигаться в поле бегущей электромагнитной волны при условии, что скорости движения электрона и волны сравнимы. Но разогнать электрон до скорости электромагнитной волны (имеющей в обычных условиях порядок скорости света в пустоте) очень трудно. Для разгона электронов до субсвето-

вых скоростей, что необходимо в ядерных исследованиях, приходится, как известно, строить гигантские ускорители вроде циклотронов, синхрофазотронов и др. Поэтому в усилительных устройствах с целью уравнивания скоростей электрона и волны прибегают не к ускорению электрона, а, наоборот, к замедлению волны. Основанный на этом принципе прибор носит название *лампы бегущей волны (ЛБВ)*.

Простейший способ замедления состоит в увеличении пробегаемого волной пути. С этой целью волну заставляют следовать вдоль направляющего проводника, свернутого в винтовую спираль. При этом скорость распространения вдоль оси спирали, разумеется, меньше, чем в направлении образующей спирали проводника; степень замедления определяется геометрией спирали, а именно отношением длины витка к шагу спирали.

Схема этого варианта конструкции ЛБВ дана на рис. 389, на котором: 1 — спиральная замедляющая катушка; 2 и 3 — входная и выходная спирали; 4 — фокусирующая катушка; 5 — катод; 6 — ускоряющий электрод; 7 — коллектор.

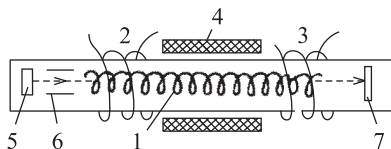


Рис. 389

В ЛБВ также происходит модуляция электронного пучка по плотности; электронные сгустки формируются в местах тормозящей фазы поля волны. Сгустки движутся со скоростью, несколько превышающей скорость волны (речь идет, конечно, о результирующей скорости в направлении движения электронов); в результате этого электроны испытывают тормозящее действие на всем своем пути. Электроны отдают свою кинетическую энергию волне; волна, следовательно, усиливается по мере распространения. Деталей устройств для ввода и вывода волны в прибор здесь не рассматриваем. Интересно заметить, что если клистронный усилитель может иметь коэффициент усиления порядка десяти (10 дБ), то в ЛБВ коэффициент усиления может достигать миллиона (60 дБ). Важным свойством ЛБВ является то, что она может работать в очень широком диапазоне частот.

К числу наиболее новых принципов усиления относится использование закономерностей квантовой физики. Существенной особенностью квантовых усилителей является то, что для усиления используется не кинетическая энергия, а внутренняя энергия частиц (атомов, ионов и молекул).

Как известно, внутренняя энергия сложной частицы не может иметь любых значений. Она может принимать только вполне определенные дискретные значения, называемые энергетическими уровнями. Между

частицей и внешним волновым полем возможен энергетический обмен: переход с нижнего на верхний уровень сопровождается поглощением кванта энергии из внешнего поля; наоборот, переход с верхнего на нижний уровень сопровождается излучением кванта энергии. Количественные соотношения при таких переходах выражаются соотношением

$$W_2 - W_1 = h\nu,$$

где W_1 и W_2 — энергетические уровни; ν — частота, а h — постоянная Планка ($h = 6,61 \cdot 10^{-27}$ эрг · с). Как видим, величина перепада между уровнями непосредственно определяет частоту поглощаемой или излучаемой волны. В диапазоне свч перепад должен быть относительно мал (по сравнению с теми, которые обуславливают, например, излучение квантов видимого света, соответствующих частоте порядка 10^{14} Гц).

Вероятности переходов частицы с нижнего энергетического уровня на верхний и обратно — одинаковы. Преобладание поглощения над излучением или наоборот определяется только соотношением количеств частиц на различных уровнях, или, как говорят, «населенностью» уровней.

При отсутствии взаимодействия с полем вещество находится в равновесном состоянии, характеризующемся тем, что число частиц на низших уровнях больше, чем на высших. Если в среду, содержащую частицы вещества, направить электромагнитную волну с частотой, соответствующей величине перепада (пусть для простоты имеются всего лишь два энергетических уровня — нижний и верхний), то энергия волны будет поглощаться и за счет этой энергии некоторое количество частиц перейдет на верхний уровень. Вещество будет теперь находиться в неравновесном (возбужденном) состоянии. Действие волны при этом эквивалентно действию цепи питания в усилителе.

Если теперь в среду с возбужденным веществом направить слабую волну той же частоты, то поле этой волны вызовет переход частиц с верхнего уровня на нижний и излучение, совпадающее по частоте и фазе с воздействующей волной, и, следовательно, усиливающее эту волну.

Возможно развязать «цепи» питания и усиления. В одной из разработанных систем это осуществляется путем применения вещества с *тремя* энергетическими уровнями $W_1 < W_2 < W_3$. Возбуждение вещества производится волной, частота которой соответствует переходу $W_1 \rightarrow W_3$. Усилению же подвергается волна с другой (меньшей) частотой, соответствующей переходу $W_3 \rightarrow W_2$.

Описанный принцип усиления очень сильно отличается от всех ранее известных. Тем не менее, можно без натяжки обсуждать действие такого усилителя в обычных терминах. Именно: источником питания является источник излучения на частоте $\nu_{13} = \frac{1}{h}(W_3 - W_1)$. За счет мощности этого источника осуществляется перевод молекул с уровня

W_1 на уровень W_3 . Источник входного сигнала — это источник слабого излучения на частоте $\nu_{32} = \frac{1}{h}(W_3 - W_2)$. Управление осуществляется входным сигналом; управляемым параметром является число молекул, переводимых с уровня W_3 на уровень W_2 . Это число в известных пределах пропорционально мощности усиливаемой волны, как оно и должно быть для обеспечения требуемой в любом усилителе пропорциональности между мощностью на входе и мощностью на выходе.

Описанный принцип осуществлен в форме так называемого *парамагнитного усилителя*. Здесь используются энергетические уровни, определяемые магнитными моментами ядер или ионов. Дело в том, что эти энергетические уровни имеют перепады, как раз соответствующие интересующим нас частотам ($1000 \div 10\,000$ МГц). Величина перепада, т. е. настройка на желаемую частоту, легко регулируется внешним магнитным полем. Отметим, что для повышения эффективности усилителя желательно, чтобы равновесное распределение частиц по уровням было как можно более неравномерным (т. е. чтобы отношение числа частиц на низших и на высших уровнях было велико). Распределение по уровням тем неравномернее, чем ниже температура. Осуществленные образцы парамагнитных усилителей работают при температуре около 4 К (-269°C), достигаемой охлаждением их жидким гелием.

Особо важным свойством квантовых усилителей является низкий уровень шума. Внутренние шумы в любой физической системе определяются дискретным строением вещества и тем обстоятельством, что движение частиц подчиняется законам случая. Поэтому, например, электрический ток, рассматриваемый как поток электронов, всегда претерпевает случайные изменения — флуктуации — около своего среднего значения, т. е. и число и скорость электронов, проходящих за одинаковые промежутки времени через сечение потока, случайным образом изменяются.

В квантовом усилителе процессы также имеют дискретную природу, но квант энергии, отдаваемый одной частицей при переходе на другой уровень, во много раз меньше энергии, отдаваемой одним электроном в обычном усилителе. Поэтому, хотя в квантовом усилителе и имеются флуктуации, но мощность их (определяющая уровень собственных шумов) может быть в сотни и тысячи раз меньше, чем в обычном усилителе.

Квантовые усилители имеют большое будущее в тех важных отраслях радиотехники, где требуется усиление очень слабых сигналов.

В этом параграфе мы трактовали процесс усиления как параметрический процесс на том основании, что во всяком усилителе имеется переменный параметр, управляющий потоком мощности, поступающей в нагрузку. Но необходимо подчеркнуть, что в современной радиотехнике термин *параметрический усилитель* относится к определенному классу устройств, основанных на одном специальном принципе. Общие пояснения этого принципа даны в приложении 5.

§ 108. Некоторые генераторы свч

Проблему генерирования колебаний в современной радиотехнике приходится ставить и решать заново применительно к диапазону свч. Это обусловлено теми же причинами, которые приводят к описанным в предыдущем параграфе своеобразным решениям проблемы усиления. Ни одна схема усилителя или генератора средних радиочастот не годится для работы на свч. Можно вообще сказать, что понятие *схемы*, как некоторой совокупности деталей, определенным образом соединенных между собой, отмирает при переходе к диапазону свч. В этом диапазоне мы имеем дело не со схемами, а с устройствами, в которых осуществляется взаимодействие между пучками частиц (электронов или молекул) и электромагнитным полем. Этим и определяется специфика техники свч.

В принципе возможно, разумеется получить генератор свч на основе любого усилителя свч, добавив к нему внешнюю цепь обратной связи, соединяющую выход усилителя с его входом. Так, например, взяв схему клистронного усилителя рис. 387 и видоизменив ее, как показано на рис. 390, добавлением коаксиальной линии, связывающей улавливатель с группирователем, получаем схему клистронного генератора.

Точно так же, соединив вход и выход усилителя на ЛБВ (рис. 389), получим генератор на ЛБВ.

Однако в этом параграфе мы не будем рассматривать этот вопрос, а приведем несколько примеров генераторов с *внутренними* обратными связями, осуществляемыми через пучок или через волновое поле.

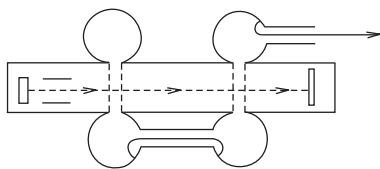


Рис. 390

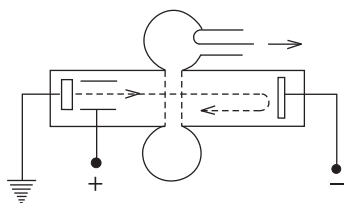


Рис. 391

В качестве первого примера рассмотрим клистронный генератор с применением *отражательного* клистрона. Устройство этого прибора показано схематически на рис. 391. Отличие от ранее описанного так называемого *пролетного* клистрона состоит в том, что пучок электронов, модулированный по скорости, попадает после группирователя в пространство с тормозящим полем, создаваемым специальным отражательным электродом, имеющим относительно катода отрицательный потенциал. Под действием этого поля электроны замедляются, а затем изменяют направление своего движения и снова попадают в группирователь. Наличие постоянного поля не отражается на процессе группирования, т. е. образования сгустков, так что возвращенный назад

пучок оказывается уже модулированным по плотности. Процесс образования сгустков поясняется рис. 392. Теперь группирователь работает уже в качестве улавливателя. Энергия группирующего поля возрастает за счет энергии сгустков, и система самовозбуждается. Надлежащую фазу легко подобрать, изменяя напряжение на отражателе. Как видим, в этом примере обратная связь осуществляется через электронный пучок.

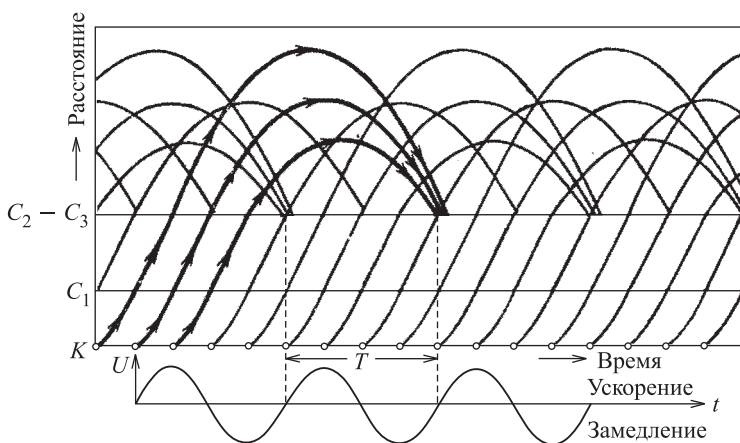


Рис. 392

В качестве второго примера опишем генератор на *лампе обратной волны* (ЛОВ), называемой также *карсинотроном*. В этой лампе, как и в ЛВБ, имеется электромагнитная волна, распространяющаяся по замедляющей системе, и электронный пучок, взаимодействующий с волной.

Различие же состоит в том, что волна и пучок движутся во встречных направлениях. Для того чтобы получить усиление волны, приходится в этом случае прибегать к прерывистому или перемежающемуся взаимодействию поля волны с электронами пучка, т.е. нужно организовать взаимодействие так, чтобы поле действовало на электроны не на всем пути, а только в тех местах, где оно имеет надлежащую фазу (т.е. там, где имеется тормозящее поле). Поясним это на примере устройства, показанного схематически на рис. 393 ¹⁾.

Волновод прямоугольного сечения изогнут, как показано на рисунке. По волноводу справа налево распространяется волна TE_{01} . Направление вектора электрического поля показано стрелками. Пучок

¹⁾ Этот рисунок не нужно рассматривать как изображение некоторой практической конструкции; он служит лишь для пояснения принципа действия ЛОВ.

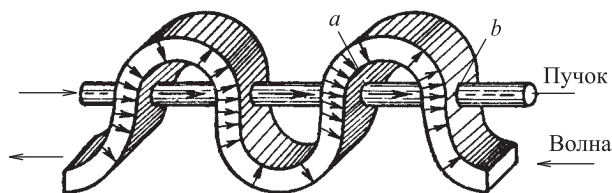


Рис. 393

электронов движется слева направо, пересекаясь с волноводом на каждом его изгибе. Если частота волны подобрана так, что от a до b вдоль изогнутой оси волновода укладывается половина длины волны, то все электроны, находящиеся в местах пересечения пучка с волноводом, будут одновременно испытывать действие тормозящего поля. В течение следующего полупериода направление поля во всех местах пересечения изменится на обратное, и электроны получат ускорение. В результате этого возникает модуляция пучка по плотности. Теперь, изменяя ускоряющее напряжение (напряжение пучка), нужно подогнать среднюю скорость электронов так, чтобы сгустки проходили через места встречи с волной в моменты надлежащей фазы, т. е. когда в этих местах имеется тормозящее поле. При этом волна будет усиливаться. Но благодаря встречному направлению волны и пучка пучок усиливает волну, а поле волны увеличивает модуляцию пучка. Таким образом, возникает обратная связь и при достаточном токе пучка система самовозбуждается. Для уяснения действия этого механизма полезно отметить его сходство с механизмом явлений в клистроне. Выделим одно колено изогнутого волновода между a и b (рис. 393). Место a можно рассматривать как группирователь, место b — как улавливатель. Область пучка между a и b — пространство группирования. В таком случае участок волновода между b и a играет роль линии обратной связи, и рассматриваемая часть системы вполне подобна клистронному генератору с внешней обратной связью, изображенному на рис. 390.

Различие между клистроном и ЛОВ сводится теперь к тому, что в ЛОВ рассмотренное звено многократно повторяется, так что обратная связь оказывается распределенной по всей длине рабочей части ЛОВ.

В заключение рассмотрим вкратце принцип действия так называемого *молекулярного* генератора, в частности аммиачного генератора, как наиболее разработанного в настоящее время. Общий принцип состоит в использовании излучения, происходящего при переходе молекул с верхних энергетических уровней на нижние. Как уже говорилось в предыдущем параграфе, в равновесном состоянии переходы с нижних уровней на верхние, и обратно с верхних на нижние — равновероятны. Для того чтобы получить преобладание излучения над поглощением, нужно повысить населенность верхних уровней, т. е. перевести вещество в возбужденное (неравновесное) состояние. В описываемом

генераторе это достигается сортировкой молекул по их энергии. Здесь нужно учесть два особых свойства молекул аммиака.

Во-первых, благодаря особенностям строения молекул наблюдается расщепление некоторых энергетических уровней, т. е. образование пар относительно близких уровней. Это важно потому, что энергетический перепад между парой таких уровней настолько мал, что соответствует уже интересующим нас частотам диапазона свч (см. формулу Планка в предыдущем параграфе). В частности, используется перепад, соответствующий длине волны 1,27 см.

Во-вторых, энергия молекул для пары расщепленных уровней зависит от внешнего электрического поля, именно: для верхнего уровня энергия молекул возрастает с увеличением поля, а для нижнего — убывает. Примерный характер этой зависимости показан на рис. 394.

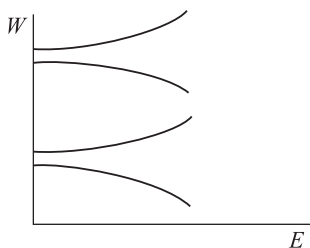


Рис. 394

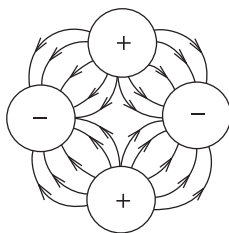


Рис. 395

Последнее обстоятельство и позволяет произвести сортировку молекул на основе общего физического принципа, состоящего в том, что всякая система стремится принять состояние, при котором внутренняя энергия минимальна. Технически сортировка осуществляется при помощи специального так называемого *квадрупольного* конденсатора, состоящего из четырех стержней, показанных в поперечном сечении на рис. 395. Между разноименными стержнями поле имеет наибольшую напряженность, на оси же системы напряженность поля равна нулю. Если вдоль оси квадрупольного конденсатора направить пучок молекул, то молекулы с большей энергией будут тяготеть к оси, где поля нет, а молекулы с меньшей энергией будут уклоняться в стороны, в области с интенсивным полем. Таким образом, пучок, выходящий из конденсатора, будет состоять преимущественно из молекул, находящихся на верхнем уровне.

Устройство молекулярного генератора в целом таково: через отверстие поступает пучок молекул аммиака. (Напомним, что пучок молекул отличается от газовой струи тем, что в пучке пренебрегается взаимодействием молекул, таким образом, пучок возможен в условиях достаточного разрежения.) Пучок проходит через квадрупольный конденсатор. Молекулы высокой энергии поступают в объемный резонатор, настроенный на волну 1,27 см. Здесь благодаря взаимодей-

ствию с волновым полем молекулы переходят на нижний уровень, испуская при этом кванты соответствующей частоты и поддерживая таким образом волновые колебания в резонаторе. Система возбуждается мягко, так как вначале всегда существует некоторое волновое поле в резонаторе, обусловленное тепловым излучением его стенок. Из специфических деталей устройства заслуживает упоминания приспособление для устранения отработанных молекул. Дело в том, что количество впускаемого газа при образовании нучка настолько велико, что откачать его обычными средствами нельзя. Поэтому применяется своеобразный прием: на пути молекул ставится охлаждаемая жидким азотом пластинка, к которой молекулы аммиака примерзают.

Молекулярный генератор на аммиаке замечателен тем, что с его помощью удается получить стабильность частоты порядка 10^{-10} . Такая стабильность в настоящее время недостижима никакими другими средствами. На основе молекулярных генераторов можно по-новому ставить вопрос об эталоне времени (период колебания) и об эталоне длины (соответствующая длина волны).

Отметим еще раз, что техника свч требует привлечения целого ряда новых принципов для осуществления основных радиотехнических процессов. Здесь радиотехника тесно смыкается с радиофизикой. Нам приходится иметь дело не только с волновыми полями, но и с физикой пучков и даже с квантовой радиофизикой. Это означает, что современный радиоинженер должен обладать широким кругозором и солидной физико-математической подготовкой.

ПРИЛОЖЕНИЯ

Приложение I

Доказательство теоремы Котельникова

Теорема Котельникова гласит: функция с ограниченным спектром полностью определяется своими мгновенными значениями, отсчитанными через $\Delta t = \frac{1}{2F}$, где F — ширина спектра функции.

Теорема может быть доказана путем разложения функции с ограниченным спектром в специальный ряд.

Если спектр функции $f(t)$ ограничен, то ¹⁾:

$$S(\omega) = \begin{cases} \int_{-\infty}^{\infty} f(t) e^{-i \omega t} dt & \text{при } |\omega| < \omega_c, \\ 0 & \text{при } |\omega| > \omega_c. \end{cases}$$

На конечном интервале $(-\omega_c, \omega_c)$ функция $S(\omega)$ может быть представлена рядом Фурье

$$S(\omega) = \sum_{-\infty}^{\infty} D_k \exp \left(i \pi k \frac{\omega}{\omega_c} \right),$$

где $2\omega_c$ — период *по частоте* (напомним, что рядом Фурье представлена функция *частоты*, а не времени, как обычно). Коэффициенты разложения D_k определяются по обычной формуле

$$D_k = \frac{1}{2\omega_c} \int_{-\omega_c}^{\omega_c} S(\omega) \exp \left(-i \pi k \frac{\omega}{\omega_c} \right) d\omega.$$

Но $\frac{\pi}{\omega_c} = \frac{1}{2F} = \Delta t$; интеграл выражает не что иное, как умноженное на 2π значение $f(-k\Delta t)$. Таким образом,

$$S(\omega) = \Delta t \sum_{-\infty}^{\infty} f(-k\Delta t) e^{i k \omega \Delta t}.$$

¹⁾ Для простоты рассматриваем случай, когда спектр занимает полосу от $\omega = 0$ до $|\omega| = \omega_c$.

Выражая функцию $f(t)$ через ее спектр $S(\omega)$, получим

$$f(t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} S(\omega) e^{i\omega t} d\omega = \frac{\Delta t}{2\pi} \int_{-\omega_c}^{\omega_c} e^{i\omega t} d\omega \sum_{-\infty}^{\infty} f(k\Delta t) e^{-k\omega\Delta t}$$

(знак при k изменен на обратный, т. е. все равно суммирование производится по всем как положительным, так и отрицательным значениям k). Изменим порядок действий

$$f(t) = \frac{\Delta t}{2\pi} \sum_{-\infty}^{\infty} f(k\Delta t) \int_{-\omega_c}^{\omega_c} e^{i\omega(t-k\Delta t)} d\omega$$

и окончательно

$$f(t) = \sum_{-\infty}^{\infty} f(k\Delta t) \frac{\sin \omega_c(t - k\Delta t)}{\omega_c(t - k\Delta t)}. \quad (317)$$

Это и есть искомое разложение. Полученная формула доказывает теорему Котельникова, так как $f(k\Delta t)$ — мгновенные значения функции $f(t)$, отсчитанные в моменты $t = k\Delta t$.

За последнее время появилось несколько работ, показывающих, что ряд, аналогичный (317), может применяться и для представления функций с неограниченным спектром. При этом имеются в виду функции, спектральная плотность которых, начиная с некоторой частоты, достаточно быстро убывает, не обращаясь, однако, в нуль ни на каком конечном отрезке оси частот. В этом случае формула (317) может служить для *приближенного* представления функции с неограниченным спектром; приближение тем лучше, чем быстрее убывает спектральная плотность.

Сравнение частотного и временного уплотнения по использованию полосы пропускания линии

Линия, применяемая для многоканальной связи, имеет определенную полосу пропускания. Обозначим ширину этой полосы через F_0 , а полосу, занимаемую одним каналом, — через F . Отношение $\xi = \frac{nF}{F_0}$ называется коэффициентом использования полосы.

Рассмотрим многоканальную систему с частотным уплотнением. Пусть разделительные фильтры идеальны, т.е. имеют прямоугольную частотную характеристику. При этом условии полосы соседних каналов примыкают вплотную и $nF = F_0$ и $\xi = 1$. В действительности же из-за конечности затухания вне полосы прозрачности приходится во избежание перекрестных помех оставлять между полосами каналов защитные промежутки, тем большие, чем хуже фильтры и чем выше требования в отношении перекрестных помех. Практическое значение коэффициента использования полосы составляет 0,7–0,8.

Обратимся теперь к временному уплотнению. При идеальной импульсной передаче импульс имеет вид

$$x_k(t) = \frac{\sin \omega_c(t - k\Delta t)}{\omega_c(t - k\Delta t)}, \quad (318)$$

где

$$\Delta t = \frac{1}{2F}, \quad \omega_c = 2\pi F.$$

Такой формы импульс получается на выходе идеального фильтра при подаче на его вход единичного (т.е. бесконечно короткого) импульса. Функция (318) обладает следующим замечательным свойством:

$$x_k(t) = \begin{cases} 1 & t = k\Delta t, \\ 0 & t = i\Delta t \ (i \neq k). \end{cases}$$

Это значит, что если при наличии любого числа импульсов вида (318), сдвинутых друг относительно друга на Δt , брать отсчеты в тактовые моменты, то в момент $k\Delta t$ отсчет будет обусловлен только наличием k -го импульса; все остальные импульсы в этот момент имеют мгновенное значение, равное нулю. Благодаря этому и осуществляется теоретически полное разделение импульсов различных каналов. Сказанное поясняется рис. П.1.

Если число каналов при временном уплотнении есть n , то частота следования увеличивается (по сравнению с частотой следования $2F$

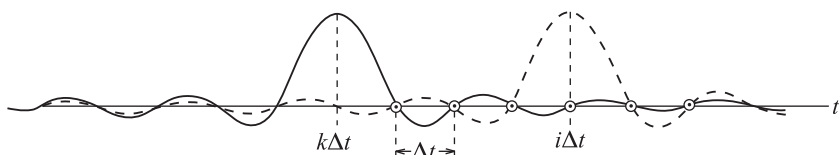


Рис. П.1

для одного канала) в n раз, и составляет $2nF$. Полоса пропускания составляет при этом nF и, следовательно, коэффициент использования полосы равен в рассматриваемом идеальном случае единице. Таким образом, в идеальных условиях частотное и временное уплотнение совершенно равноценны; применение идеальных фильтров позволяет полностью использовать полосу.

Можно было бы ожидать, что в реальных условиях временное уплотнение даст снижение коэффициента использования такого же порядка, как и частотное уплотнение, т. е. на 20–30 %. Однако практика применения систем с временным уплотнением показывает, что коэффициент использования полосы при временном уплотнении раз в двадцать меньше, чем при частотном (что соответствует значению $\xi \approx 0,05$). Это важное обстоятельство, существенно ограничивающее использование временного уплотнения, требует разъяснения.

Дело в том, что переход от идеальных фильтров к реальным влечет за собой в случае частотного и в случае временного уплотнения совершенно разные последствия. В случае частотного уплотнения несовершенство фильтров требует, как уже говорилось, введения относительно небольших защитных промежутков. В случае же временного уплотнения переход к реальным фильтрам радикально меняет характер импульсов.

Предположим, что фильтр нижних частот, вырабатывающий импульсы, нагружен на характеристическое сопротивление. (Это условие само по себе нереально, но мы примем его, так как при нагрузке на любое другое сопротивление условия фильтрации ухудшаются.) В таком случае импульсная реакция фильтра из m звеньев выражается формулой

$$x(t) = 2m \frac{J_{2m}(\omega_c t)}{t}, \quad (319)$$

где ω_c — граничная частота фильтра. Эта формула существенно отличается от (318). Числитель (318) есть периодическая функция, нули которой находятся на равных расстояниях Δt друг от друга. Поэтому смещение функции (318) на целое число интервалов между нулями приводит к тому, что все нули остаются на прежних местах. В числителе же формулы (319) стоит непериодическая бесселева функция. Ее нули находятся на неодинаковых расстояниях, и поэтому наложение друг на друга нескольких таких функций ни при каких сдвигах не даст совпадающих по положению нулей. Для примера на рис. П.2 показана

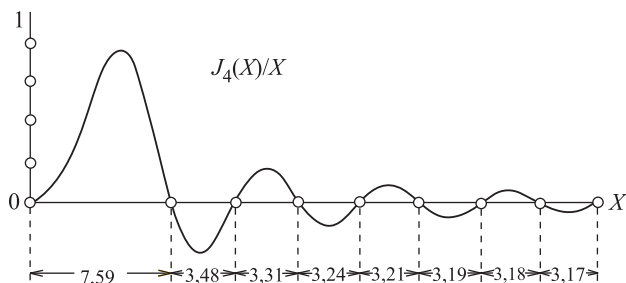


Рис. П.2

функция $J_4(x)/x$, выражающая импульсную реакцию нагруженного на характеристическое сопротивление двухзвенного фильтра нижних частот. На рисунке приведены значения интервалов между нулями. Лишь в пределе (при неограниченном возрастании x) величина интервалов стремится к значению π , тогда как для функции $\sin x/x$, представляющей отклик идеального фильтра, все интервалы между нулями равны в точности π ¹⁾.

Итак, в идеальном случае мы используем тонкую структуру отклика фильтра, размещая импульсы в тех местах, где все остальные импульсы имеют нулевое значение.

В реальном же случае использование тонкой структуры становится невозможным, и для того, чтобы избежать значительных перекрестных помех, нам остается только посылать каждый следующий импульс лишь тогда, когда след предыдущего практически полностью затухнет. При этом, конечно, нет надобности разбираться в тонкой структуре импульса; достаточно рассматривать его огибающую. Так, например, огибающая для функции $\sin x/x$ есть просто $1/x$. Для того чтобы мгновенные значения импульса не превосходили, например, 0,01 от начального значения, нужно выполнить неравенства $x \geq 100$, тогда как первый нуль получается при $x = \pi$.

Как видим, эти простые соображения и приводят нас к тем практическим значениям коэффициента использования полосы, которые были приведены выше.

¹⁾ Для краткости введена переменная $x = \omega_c t$. Нули получаются при $x = k\pi$ или $\omega_c t = k\pi$, $t = k \frac{\pi}{\omega_c} = k\Delta t$.

Обоснование критерия Найквиста

Критерий устойчивости Найквиста выражается следующим образом: *система с обратной связью устойчива, если амплитудно-фазовая характеристика разомкнутой системы не охватывает точки (1,0).*

Докажем справедливость этого утверждения и выясним попутно ряд деталей, уточняющих вышеприведенную формулировку.

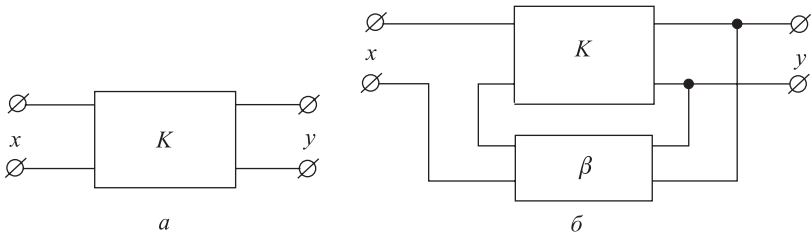


Рис. П.3

Рассмотрим линейный четырехполюсник; входное и выходное напряжения обозначим соответственно через x и y (рис. П.3, а). Соотношение между x и y выражается в достаточно общем виде уравнением

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = b_m \frac{d^m x}{dt^m} + \dots + b_1 \frac{dx}{dt} + b_0 x.$$

В операторной записи это уравнение принимает вид

$$(a_n p^n + a_{n-1} p^{n-1} + \dots + a_1 p + a_0) \bar{y} = (b_m p^m + \dots + b_1 p + b_0) \bar{x} \quad (320)$$

или короче

$$Z_2(p) \bar{y} = Z_1(p) \bar{x}, \quad (321)$$

где через $Z(p)$ обозначены соответствующие многочлены по степеням p в правой и левой частях уравнения (320). Из (321) находим для y

$$\bar{y} = \bar{x} \frac{Z_1(p)}{Z_2(p)}$$

или, еще короче

$$\bar{y} = K(p) \bar{x},$$

где $K(p) = \frac{\bar{y}}{\bar{x}} = \frac{Z_1(p)}{Z_2(p)}$ — операторный коэффициент передачи.

Решение как функция времени имеет вид

$$y(t) = \bar{x}(0) \frac{Z_1(0)}{Z_2(0)} + \sum_{k=1}^n \frac{\bar{x}(p_k) Z_1(p_k)}{p_k Z_2'(p_k)} e^{p_k t},$$

где p_k — n корней характеристического уравнения

$$Z_2(p) = 0.$$

Корни эти, вообще говоря, комплексны и могут быть представлены в виде

$$p_k = \alpha_k + i \omega_k.$$

Каждый член решения содержит множитель вида

$$e^{p_k t} = e^{(\alpha_k + i \omega_k) t} = e^{\alpha_k t} e^{i \omega_k t},$$

т. е. выражает колебание с частотой ω_k и с амплитудой, экспоненциально изменяющейся во времени¹⁾. Если α_k отрицательно, то колебание затухает; если α_k положительно, то колебание неограниченно нарастает. Этот последний случай и характеризует неустойчивую систему. В устойчивой системе все возникающие по тем или иным причинам свободные колебания с собственными частотами ω_k должны со временем затухать.

Итак, условием устойчивости разомкнутой системы является отрицательность вещественных частей корней α_k характеристического уравнения $Z_2(p) = 0$.

Замкнем теперь вход и выход петель обратной связи. Коэффициент передачи цепи обратной связи обозначим через β . Связь между x и y выражается уравнением (см. рис. П.3, б)

$$\bar{y} = K(\bar{x} + \beta \bar{y})$$

или

$$\bar{y} = \frac{K}{1 - K\beta} \bar{x} = \frac{K(p)}{S(p)} \bar{x},$$

т. е. коэффициент передачи с учетом действия обратной связи равен

$$K_\beta = \frac{K}{1 - K\beta} = \frac{K}{S},$$

и теперь устойчивость определяется корнями характеристического уравнения

$$S(p) = 1 - K(p) \beta(p) = 0.$$

Рассмотрим плоскость комплексного переменного $p = \alpha + i \omega$. На рис. П.4 отмечены возможные значения корней. Корни с положительной вещественной частью лежат в правой полуплоскости (отмечены

¹⁾ Точнее говоря, колебание с частотой ω выражается парой членов, соответствующих паре комплексно-сопряженных корней.

крестиками), корни с отрицательной вещественной частью — в левой полуплоскости (отмечены точками). Таким образом, правая полуплоскость есть область неустойчивости, а левая полуплоскость — область устойчивости. Границей между обеими областями является мнимая ось — ось частот, уравнение которой есть $\alpha = 0$.

Перейдем теперь от плоскости комплексного переменного к плоскости функции

$$S(p) = \xi + i \eta.$$

Этот переход есть не что иное, как конформное отображение плоскости p на плоскость $S(p)$.

Мнимая ось $p = i \omega$ отобразится на плоскости S кривой $S(i\omega)$. (Будем в дальнейшем рассматривать только половину мнимой оси, соответствующую изменению частоты от 0 до ∞ ; вторая половина дает симметричную картину, так как комплексные корни встречаются всегда попарно.) Кривая $S(i\omega)$ для подавляющего большинства случаев оказывается замкнутой, так как коэффициент передачи $K(i\omega)$ равен нулю как при $\omega = 0$, так и при $\omega \rightarrow \infty$, так что кривая $K(i\omega)$, а следовательно, и $S(i\omega)$ начинается и кончается в одной точке¹⁾. Таким образом, замкнутая кривая $S(i\omega)$ делит всю плоскость S на две области — внутри кривой и вне кривой (см. рис. П.5).

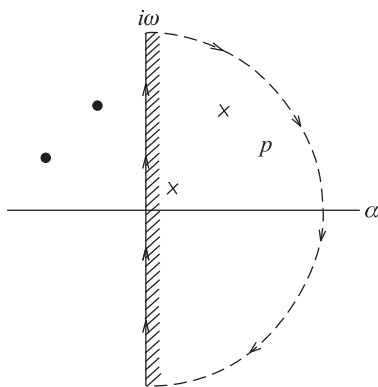


Рис. П.4

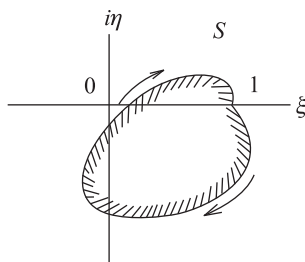


Рис. П.5

Эти две области соответствуют двум полуплоскостям плоскости p . Будем следовать вдоль оси $i\omega$ в направлении возрастания частоты, как показано стрелкой на рис. П.4. Чтобы охватить всю правую полуплоскость, замкнем путь полуокружностью бесконечно большого радиуса (пунктир на рис. П.4²⁾). При таком порядке обхода вся область

¹⁾ Исключением из этого правила являются усилители постоянного тока.

²⁾ Все точки этой полуокружности отображаются точкой $S = 1$.

остается справа от линии обхода. В теории конформного отображения доказывается, что и преобразованная область располагается таким же образом относительно преобразований границы.

Следовательно, внутренность замкнутой кривой рис. П.5 является отображением правой полуплоскости, а внешняя по отношению к кривой часть плоскости S — отображением левой полуплоскости переменного p (точнее говоря, правого верхнего и левого верхнего квадрантов плоскости p , так как мы условились не рассматривать отрицательных значений ω). Соответствие областей обозначено на рис. П.4 и П.5 штриховкой вдоль границы. Итак, на плоскости S область внутри кривой $S(i\omega)$ является неустойчивой, а внешняя область — устойчивой.

Для дальнейшего нам удобнее будет пользоваться не плоскостью S , а плоскостью $K_1 = K\beta$, так как коэффициенты передачи основной цепи K и цепи обратной связи β это как раз те величины, которые непосредственно задаются или определяются расчетным или опытным путем. Из общего соотношения между S и $K\beta$ находим

$$K_1 = K\beta = 1 - S = \xi_1 + i\eta_1$$

и перестраивая в соответствии с этой формулой график рис. П.5, получаем для плоскости K_1 рис. П.6.

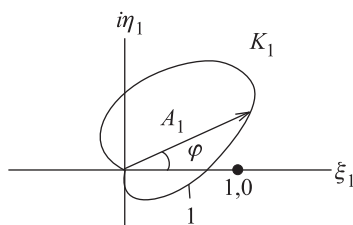


Рис. П.6

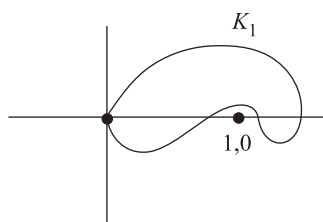


Рис. П.7

Для решения вопроса об устойчивости данной системы нужно выяснить, где располагаются корни характеристического уравнения — в устойчивой или неустойчивой области. Замечательно, что все корни, отмеченные на плоскости p , где бы они ни располагались, отображаются *одной единственной* точкой на плоскости S , а этой точкой является начало координат (в силу того, что корни удовлетворяют уравнению $S(p) = 0$). Поэтому система устойчива, если кривая $S(i\omega)$ не охватывает начала координат. Что же касается плоскости K_1 , то на ней, как легко видеть, все корни уравнения отображаются точкой на вещественной оси с координатами $\xi_1 = 1, \eta_1 = 0$. Это и есть «точка (1,0)». Если эта точка лежит внутри кривой $K_1(i\omega)$, т. е. охватывается этой кривой, то система неустойчива. В противном случае система устойчива, и мы пришли к формулировке критерия Найквиста. Рис. П.5 и П.6 соответ-

ствуют устойчивому случаю. Отметим, что и рис. П.7 в соответствии с критерием Найквиста представляет устойчивый случай.

Удобнее воспользоваться на плоскости K_1 не декартовыми координатами ξ_1 и η_1 , а полярными координатами A_1 и φ_1 (рис. П.6). Представим коэффициент передачи в виде

$$K_1 = A_1 e^{i\varphi_1}.$$

Здесь K_1 — комплексная величина, а A_1 и φ_1 — вещественные; $A_1 = |K_1|$ есть отношение амплитуд (т.е. амплитудно-частотная характеристика), а φ_1 — фазовый сдвиг между напряжениями на выходе и на входе (т.е. фазово-частотная характеристика). «Точка (1, 0)» при переходе к этим координатам сохраняет свое обозначение, но теперь оно расширяется как $A_1 = 1$, $\varphi_1 = 0$. Коэффициент передачи

$$K_1 = K\beta$$

измеряется в разомкнутой цепи, получаемой из схемы рис. П.3, б путем разрывания цепи обратной связи в произвольном месте, как показано на рис. П.8.

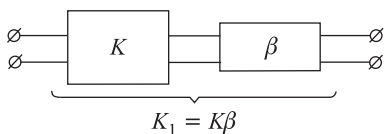


Рис. П.8

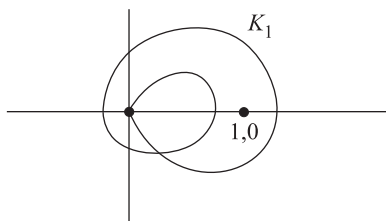


Рис. П.9

В заключение отметим следующее. Как уже говорилось, все корни, различным образом расположенные на плоскости p , отображаются одной единственной точкой на плоскости S или на плоскости K_1 . Но на плоскости p корни располагаются в зависимости от знака вещественной части либо в устойчивой, либо в неустойчивой областях, т.е. по разные стороны оси $i\omega$, являющейся границей между указанными областями. Кажущееся противоречие разрешается следующим образом. Вектор S совершает по одному полному обороту (т.е. поворачивается на 2π) на каждый корень, находящийся внутри обходимой области, и совершает лишь качательное перемещение для внешнего корня. Это значит, что конец вектора $S(i\omega)$ выписывает кривую, содержащую несколько витков; каждый виток соответствует своему корню и проходит относительно общей точки, отображающей все корни на плоскости S , соответственным образом, именно: для устойчивого корня ($\alpha_k < 0$) виток проходит, не захватывая начала координат, а для неустойчивого корня ($\alpha_k > 0$) виток захватывает начало координат. (На плоскости K_1 то же

относится к точке $(1, 0)$.) Понятно, что система будет неустойчивой, если имеется хотя бы один неустойчивый корень. Итак, оказывается, как оно и должно быть, что и на плоскости S (и на плоскости K_1) корни расположены по отношению к границе по-разному, хотя все они и отображаются одной и той же точкой. Примерный вид графика на плоскости K_1 для случая двух корней, из которых один устойчивый, а второй неустойчивый, показан на рис. П.9.

Краткие сведения об уравнении Матьё

Теория дифференциальных уравнений с переменными коэффициентами сравнительно мало разработана. Имеется лишь достаточно полная теория уравнения Хилла

$$x'' + [a + qf(\tau)]x = 0,$$

где $f(\tau)$ — периодическая функция. Частный вид уравнения Хилла

$$x'' + [a + 2q \cos 2\tau]x = 0 \quad (322)$$

называется уравнением Матьё. Именно к такому виду приводится уравнение контура с емкостью, периодически изменяющейся по синусоидальному закону.

Общее решение уравнения (322) складывается из двух независимых решений и может быть записано в виде

$$x = x_1 + x_2 = A e^{\mu\tau} \Phi_1(\tau) + B e^{-\mu\tau} \Phi_2(\tau). \quad (323)$$

Здесь A и B — произвольные постоянные; Φ_1 и Φ_2 — периодические функции периода π или 2π . Эти периодические функции могут быть выражены соответствующими тригонометрическими рядами, но не эта сторона дела будет нас интересовать. С точки зрения наших задач наибольшее значение имеет вопрос об *устойчивости* решений уравнения Матьё.

Легко видеть, что если μ — действительное число, не равное нулю, то одно из слагаемых решения (323) неограниченно растет с увеличением τ , а это и означает, что решение неустойчиво.

Показатель μ зависит от коэффициентов a и q .

Можно показать, что границы областей устойчивости определяются условием $\mu = 0$. Графическое представление областей устойчивости на плоскости a, q дано на рис. П.10. Области неустойчивости (заштрихованные на рисунке) ограничены кривыми, сходящимися на оси a к точкам $a = 1, 4, 9, 16$ и т. д. Картина симметрична относительно оси a .

Уравнение, отличающееся от (322) наличием члена с первой производной, т. е. уравнение вида

$$y'' + 2\chi y' + (b + 2q \cos 2\tau)y = 0, \quad (324)$$

может быть приведено к стандартной форме (322). Для этого возьмем

$$y = e^{-\chi\tau} x. \quad (325)$$

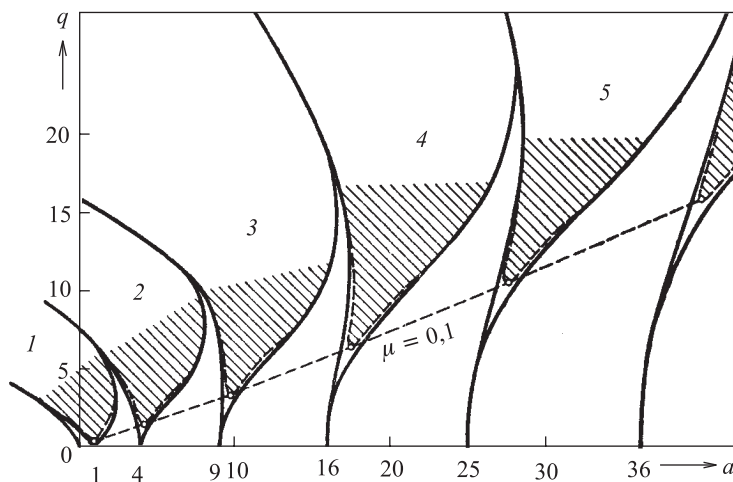


Рис. П.10

В результате такой подстановки получим уравнение (322), в котором нужно положить

$$a = b - \chi^2.$$

Сопоставляя (323) и (325), видим, что решение уравнения (324) должно иметь вид

$$y = A e^{(-\chi+\mu)\tau} \Phi_1(\tau) + B e^{(-\chi-\mu)\tau} \Phi_2(\tau).$$

Решение будет неустойчивым уже не при $|\mu| > 0$, а при

$$-\chi + |\mu| > 0,$$

т. е. при

$$|\mu| > \chi.$$

Следовательно, для уравнения (324) области неустойчивости будут ограничены уже не кривыми $\mu = 0$, а кривыми

$$|\mu| = \chi.$$

Эти кривые не доходят до оси, они спускаются к ней языками, наинизшие точки которых отстоят от оси тем дальше, чем больше μ . На рис. П.10 пунктирными линиями отмечены границы областей неустойчивости для уравнения (324) при $\chi = 0,1$.

Система теряет устойчивость легче всего при наименьших значениях q , т. е. в условиях, соответствующих минимуму граничных кривых. Поэтому нужно найти координаты минимумов. Точное решение этой задачи очень громоздко. Но если полагать, что в уравнении (322) $a \gg q$ (это как раз тот случай, который нам встретится в дальнейшем), то можно воспользоваться приближенными соотношениями. Приближе-

ние состоит в том, что мы отбрасываем члены с высокими степенями q в степенных рядах, выражающих интересующие нас зависимости, и полагаем также, что точка минимума лежит посредине горизонтального отрезка, соединяющего две кривые $\mu = 0$ в области неустойчивости. Тогда получаем данные, приведенные в табл. 14 (m означает номер области неустойчивости, a_m и q_m — координаты минимума граничной кривой).

Таблица 14

m	1	2	3	4	5
a_m	$1 - \mu^2$	$4 + \frac{8}{3}\mu$	$9 + (6\mu)^{2/3}$	$16 + \frac{16}{5}(2\mu)^{1/2}$	$25 + \frac{4}{3}(45\mu)^{2/5}$
q_m	2μ	$4\mu^{1/2}$	$4(6\mu)^{1/3}$	$4(72\mu)^{1/4}$	$8(45\mu)^{1/5}$

Точки с указанными в табл. 14 координатами при $\mu = 0,1$ отмечены на рис. П.10. Они лежат на кривой, проходящей через начало и мало отличающейся от прямой. Итак, для нарушения устойчивости при данном χ , требуется тем большее q , чем больше a , или, иначе, чем выше номер области неустойчивости. Приведенных сведений уже достаточно для решения задачи о параметрическом возбуждении, к которой и переходим.

Составим уравнение напряжений для контура

$$Lq'' + Rq' + \frac{1}{C(t)}q = 0, \quad (326)$$

где q — заряд. Если емкость изменяется около среднего значения C_0 по закону

$$C(t) = C_0 + C_1 f(t) = C_0 \left[1 + \frac{C_1}{C_0} f(t) \right] = C_0 [1 + n f(t)],$$

то при $n = \frac{C_1}{C_0} \ll 1$ можем переписать (326) в виде

$$q'' + 2\alpha q' + \omega_0^2 [1 - n f(t)] q = 0.$$

Заменив переменную по формуле

$$q = y e^{-\alpha t},$$

избавимся от члена с первой производной и получим уравнение

$$y'' + [\omega_1^2 - \omega_0^2 n f(t)] y = 0,$$

где

$$\omega_1^2 = \omega_0^2 - \alpha^2$$

— квадрат частоты свободных колебаний контура с постоянными параметрами L , R и C_0 . Пусть теперь

$$f(t) = \cos \omega t.$$

Обозначая $\omega t = 2\tau$ и заменяя производную по t производной по τ , получим

$$y'' + 4 \left[\frac{\omega_1^2}{\omega^2} + \frac{\omega_0^2}{\omega^2} n \cos 2\tau \right] y = 0,$$

а это и есть уравнение Матьё (322). Теперь задача нахождения условий, при которых возможно возбуждение колебаний в контуре, сводится к определению соотношений, при которых уравнение Матьё имеет неустойчивые решения. Мы имеем условие неустойчивости

$$\chi = \frac{2\alpha}{\omega} < \mu. \quad (327)$$

В нашем случае:

$$q = 2n \frac{\omega_0^2}{\omega^2}, \quad a = 4 \frac{\omega_1^2}{\omega^2}$$

или, если затухание невелико и можно пренебречь различием между ω_1 и ω_0 ,

$$a \approx 4 \frac{\omega_0^2}{\omega^2}.$$

Таким образом коэффициент q выражает относительное изменение переменного параметра (в нашем случае — емкости), а коэффициент a — отношение частот: частоты колебаний в контуре и частоты изменения параметра. Возбуждение легче всего возникает в точках минимумов граничных кривых. Для первой области неустойчивости имеем (см. табл. 14):

$$a = 1 - \mu^2, \quad q = 2\mu.$$

Так как $\mu \ll 1$, то можно положить $a_1 = 1$, откуда

$$\omega_0 = \frac{1}{2}\omega,$$

т.е. колебания легче всего возбуждятся на частоте, равной половине частоты изменения емкости. Подставляя значение

$$\mu = \frac{1}{2}q$$

в условие возбуждения (327), получим

$$\frac{2\alpha}{\omega} < \frac{1}{2}q = n \frac{\omega_0^2}{\omega^2}.$$

Преобразовывая это неравенство, получим

$$\frac{2\alpha}{\omega} = d \frac{\omega_0}{\omega} < n \frac{\omega_0^2}{\omega^2}.$$

Но так как $\frac{\omega_0}{\omega} = \frac{1}{2}$, то окончательно условие возбуждения запишется в виде

$$n > 2d,$$

что совпадает с результатом, полученным в § 104.

Итак, при данном затухании контура параметрические колебания в нем возникнут при периодическом изменении емкости при условии достаточно большого относительного изменения емкости. Чем меньше затухание, тем меньше необходимое относительное изменение

$$n = \frac{C_1}{C_0}.$$

Мы вывели условие возбуждения для первой области неустойчивости при $\omega_0 \approx \frac{1}{2}\omega$. Возбуждение возможно и для других соотношений частот, т. е. для других областей неустойчивости. Однако при этом труднее возбудить колебание, т. е. требуется большее относительное изменение параметра. Для второй области неустойчивости имеем координаты минимума граничной кривой:

$$q_2 = 4\mu^{1/2}, \quad a_2 = 4 + \frac{\gamma}{3}\mu.$$

Положим приближенно $a_2 = 4$, т. е. $\omega_0 = \omega$.

Подставляя в (327)

$$\mu = \frac{1}{16}q_2^2 = \frac{1}{4}n^2 \frac{\omega_0^4}{\omega^4} = \frac{1}{4}n^2,$$

получим

$$d < \frac{1}{4}n^2,$$

т. е.

$$n > 2\sqrt{d}.$$

Для третьей области неустойчивости координаты минимума:

$$q_3 = 4(6\mu)^{1/3}, \quad a_3 = 9 + (6\mu)^{2/3} \approx 9,$$

откуда:

$$\frac{4\omega_0^2}{\omega^2} = 9, \quad \omega_0 = \frac{3}{2}\omega, \quad \mu = \frac{1}{384}q^3.$$

Подставляя это в (327), получаем

$$n = \frac{8}{3\sqrt[3]{3}}\sqrt[3]{d} \approx 1,84\sqrt[3]{d}$$

и так далее. Таким образом, с повышением номера m области неустойчивости необходимое значение n растет, как $d^{1/m}$.

Основы параметрического усиления

За последнее время в радиотехнике, а в особенности в области техники свч, применяется особый принцип так называемого параметрического усиления.

Сущность параметрического усиления состоит в том, что при помощи нелинейного реактивного элемента, играющего роль преобразователя частоты, из частот питания и сигнала образуются комбинационные частоты, при этом происходит такое перераспределение мощности по комбинационному спектру, что мощность, отбираемая от усилителя на комбинационных частотах, больше, чем мощность входного сигнала.

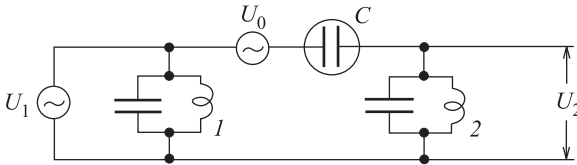


Рис. П.11

Рассмотрим сначала частный пример: двухконтурный усилитель с использованием суммарной частоты (такое преобразование называют «необращающим», или «преобразованием вверх» (up-conversion)). Схема усилителя изображена на рис. П.11. Здесь u_1 — источник усиливаемого сигнала с частотой ω_1 , u_0 — источник питания (так называемой «подкачки») с частотой $\omega_0 \gg \omega_1$, C — нелинейный реактивный элемент, в частности, нелинейная емкость, 1 и 2 — колебательные контуры, из которых первый настроен на частоту сигнала ω_1 , а второй — на суммарную комбинационную частоту $\omega_0 + \omega_1$.

Заменим нелинейную схему рис. П.11 эквивалентной *линейной* схемой с переменным параметром.

Нелинейная емкость определяется зависимостью

$$q = f(u),$$

где q — заряд; u — напряжение. Дифференциальная емкость есть

$$C = \frac{dq}{du} = f'(u).$$

Представим эту функцию рядом Маклорена

$$f'(u) = f'(0) + f''(0)u + f'''(0)\frac{1}{2}u^2 + \dots$$

и ограничимся двумя первыми членами разложения. Тогда будем иметь

$$C = C_0 + C_1,$$

где $C_0 = f'(0)$ — постоянная составляющая емкости, а $C_1 = f''(0)u$ — переменная составляющая, пропорциональная действующему в цепи напряжению. В схеме рис. П.11 имеются два напряжения: напряжение питания u_0 и напряжение сигнала u_1 . Но так как последнее значительно меньше, то им можно пренебречь и полагать, что переменная составляющая емкости определяется только напряжением питания

$$u_0 = U_0 \cos \omega_0 t.$$

Тогда

$$C_1 = f''(0)U_0 \cos \omega_0 t = \Delta C \cos \omega_0 t.$$

Итак, нелинейную емкость C вместе с источником напряжения питания u_0 можно заменить зависящей от времени емкостью, в чем и состоит переход от нелинейной схемы к линейной схеме с переменным параметром. Заметим, кроме того, что контур 1 настроен на частоту ω_1 , и можно считать, что он представляет нулевое сопротивление для всех частот, кроме ω_1 . Аналогично контур 2 представляет короткое замыкание для всех частот, кроме $\omega_0 + \omega_1$. Рассмотрев схемы замещения для этих частот, убеждаемся, что постоянная составляющая емкости может быть подключена к обоим контурам параллельно, и преобразованная схема принимает вид, показанный на рис. П.12.

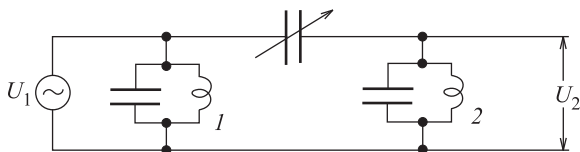


Рис. П.12

Найдем ток, протекающий через C_1 . Для этого представим в комплексной форме напряжения u_1 и u_2 и переменную емкость C_1 . Учитывая наличие настроенных контуров, можно записать:

$$u_1 = \dot{U}_1 e^{i\omega_1 t}, \quad u_2 = \dot{U}_2 e^{i(\omega_0 + \omega_1)t}.$$

Выражение для переменной емкости можно представить в виде

$$C_1 = \Delta C \cos \omega_0 t = \frac{1}{2} \Delta C (e^{i\omega_0 t} + e^{-i\omega_0 t}).$$

Переменный заряд равен

$$q = C_1(u_1 - u_2),$$

а ток

$$i = \frac{dq}{dt} = \frac{dC_1}{dt}(u_1 - u_2) + C_1 \frac{d(u_1 - u_2)}{dt} \quad 1)$$

Подставляя сюда

$$\frac{dC_1}{dt} = i\omega_0 \frac{1}{2} \Delta C (e^{i\omega_0 t} - e^{-i\omega_0 t})$$

и

$$\frac{d(u_1 - u_2)}{dt} = i \left[\omega_1 \dot{U}_1 e^{i\omega_1 t} - (\omega_0 + \omega_1) \dot{U}_2 e^{i(\omega_0 + \omega_1)t} \right],$$

получим после раскрытия скобок составляющие тока с частотами $2\omega_0 + \omega_1$, $\omega_0 - \omega_1$, $\omega_0 + \omega_1$ и ω_1 . Но нас интересуют мощности; поэтому мы сохраним только две последние составляющие с теми же частотами, что и имеющиеся напряжения. Эти две составляющие равны:

$$i_1 = -i \frac{1}{2} \Delta C \omega_1 \dot{U}_2 e^{i\omega_1 t}, \quad i = i \frac{1}{2} \Delta C (\omega_0 + \omega_1) \dot{U}_1 e^{i(\omega_0 + \omega_1)t},$$

а комплексные амплитуды составляющих тока с частотами ω_1 и $\omega_0 + \omega_1$, равны соответственно:

$$\dot{I}_1 = -i \frac{1}{2} \Delta C \omega_1 \dot{U}_2, \quad \dot{I}_2 = i \frac{1}{2} \Delta C (\omega_0 + \omega_1) \dot{U}_1. \quad (328)$$

Напомним, что мощность выражается через комплексные амплитуды тока и напряжения следующим соотношением:

$$P = \frac{1}{2} \operatorname{Re} (\dot{U} \dot{I}^*) = \frac{1}{2} \operatorname{Re} (\dot{U}^* \dot{I}),$$

где звездочка означает комплексно-сопряженную величину. Заменим одно из равенств (328), например первое, — равенством комплексно-сопряженных величин, после чего перемножим оба равенства накрест (правую часть одного на левую часть другого). Получим

$$i \frac{1}{2} \Delta C (\omega_0 + \omega_1) \dot{U}_1 \dot{I}_1^* = i \frac{1}{2} \Delta C \omega_1 \dot{U}_2^* \dot{I}_2.$$

Сокращая на общий множитель и беря вещественную часть, находим

$$(\omega_0 + \omega_1) P_1 = \omega_1 P_2.$$

Но P_1 — это мощность, отдаваемая источником сигнала (на частоте ω_1), а P_2 — выходная мощность, отдаваемая усилителем (на комбинационной частоте $\omega_0 + \omega_1$).

¹⁾ Обычно при применении комплексного метода дифференцирование заменяется умножением на $i\omega$. Но мы не можем так поступить, так как имеем дело не с одной частотой, а с комбинационным спектром.

Таким образом, коэффициент усиления по мощности

$$\varsigma = \frac{P_2}{P_1} = \frac{\omega_0 + \omega_1}{\omega_1}. \quad (329)$$

Это — предельное значение. Действительный коэффициент усиления (при вычислении которого учитывается соотношение проводимостей контуров и нагрузки, потери в управляемой емкости и т. д.) всегда несколько меньше. Мы не будем углубляться в эти подробности, так как нас интересует не технический расчет параметрического усилителя, а принципиальные основы его действия.

Итак, предельный коэффициент усиления, как показывает формула (329), зависит только от соотношения частот сигнала и питания. Очень интересно, что он вовсе не зависит (в принятом приближении) от характера нелинейности (или от глубины модуляции емкости), а также от величины напряжения питания. Но от переменной составляющей емкости зависит, оказывается, ширина полосы усиления. Рассматривая схему рис. П.12 как систему из двух контуров с переменной емкостной связью, можно получить для ширины полосы выражение

$$2\Delta\omega = \frac{\frac{1}{2}\Delta C}{\sqrt{C_1 C_2}} \sqrt{2\omega_1 \omega_2},$$

где индексы 1 и 2 относятся к первому и второму контурам.

Формула (329) представляет частный случай весьма общих соотношений, известных под названием уравнений Мэнли–Роу¹⁾. Эти уравнения устанавливают распределение мощностей по комбинационному спектру, получаемому в результате действия реактивного преобразователя. Вывод уравнений Мэнли–Роу громоздок и не содержит моментов, интересных в принципиальном отношении. Поэтому уравнения приводятся в готовом виде:

$$\left. \begin{aligned} \sum_{m=0}^{\infty} \sum_{n=-\infty}^{\infty} \frac{m P_{mn}}{\omega_{mn}} &= 0 \\ \sum_{m=-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{n P_{mn}}{\omega_{mn}} &= 0 \end{aligned} \right\}. \quad (330)$$

Здесь

$$\omega_{mn} = m\omega_0 + n\omega_1$$

— комбинационные частоты, m и n — индексы, пробегающие при суммировании все целые значения в указанных пределах; P_{mn} — мощность на частоте ω_{mn} ²⁾. Принято считать мощность, входящую в преобразователь, положительной, а выходящую из преобразователя —

¹⁾ Manley J. M., Rowe H. E. // Proc IRE. — 1956. — V. 44, No. 7. — P. 904.

²⁾ Частоты считаются положительными, иначе говоря, $P_{(-m)(-n)} = P_{mn}$.

отрицательной. Если потери в преобразователе не учитываются, то сумма входящих мощностей должна равняться сумме выходящих, т. е.

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} P_{mn} = 0^1).$$

Покажем применение уравнений (330) к рассмотренному выше случаю усилителя на суммарной частоте $\omega_0 + \omega_1$. В этом случае имеем частоты $\omega_1(m = 0, n = 1)$, $\omega_0(m = 1, n = 0)$ и $\omega_0 + \omega_1(m = 1, n = 1)$. Согласно (330)

$$\left. \begin{aligned} \frac{P_{10}}{\omega_0} + \frac{P_{11}}{\omega_0 + \omega_1} &= 0 \\ \frac{P_{01}}{\omega_1} + \frac{P_{11}}{\omega_0 + \omega_1} &= 0 \end{aligned} \right\}. \quad (331)$$

Здесь P_{01} — мощность сигнала; P_{10} — мощность питания; P_{11} — мощность, отдаваемая преобразователем на суммарной частоте $\omega_0 + \omega_1$. По вышеуказанному правилу ей нужно приписать знак минус. Коэффициент усиления из второго равенства (331) определяется как

$$\zeta = -\frac{P_{11}}{P_{01}} = \frac{\omega_0 + \omega_1}{\omega_1}.$$

Этот результат получается сразу, без анализа схемы, к которому мы прибегали для получения формулы (329); данный пример уже может дать некоторое представление о силе и общности уравнений Мэнли–Роу.

Как видно из (331), мощности прямо пропорциональны соответствующим частотам, что позволяет построить наглядное графическое представление соотношения (331), показанное на рис. П.13.

Здесь изображены три спектральные линии мощности. Вертикальными пунктирными линиями отмечены границы полосы усиления при изменении частоты сигнала на $\pm\Delta\omega$.

Мы рассматривали до сих пор усиление на суммарной частоте. Возможно усиление и на разностной частоте $\omega_0 - \omega_1$ («обращающее» преобразование). Этот случай отличается тем, что усилитель при известных соотношениях может потерять устойчивость и превратиться в генератор (что исключено для усилителя на суммарной частоте). Иначе говоря, действие преобразователя вносит в систему отрицательное сопротивление, так что усилитель на разностной частоте можно рассматривать как недовозбужденный параметрический генератор (см. § 104). При $\omega_0 = 2\omega_1$ преобразования частоты нет, т. е. усиление происходит на частоте сигнала ω_1 .

¹⁾ Это соотношение может быть получено из уравнений (330), если первые из них умножить на ω_0 , а второе — на ω_1 , и затем сложить.

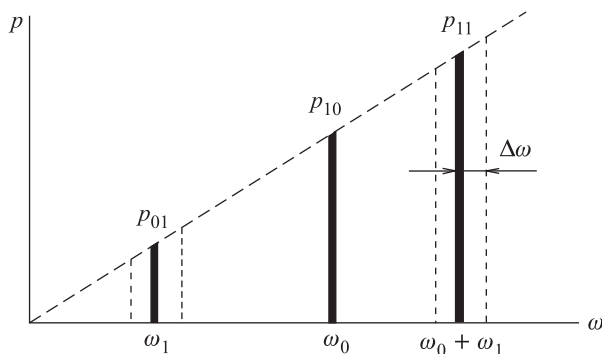


Рис. П.13

Параметрические усилители находят применение в диапазоне свч. Колебательные контуры заменяются при этом отрезками волноводов или объемными резонаторами. В качестве нелинейной емкости часто применяется полупроводниковый диод, описанный в § 106, причем диод непосредственно вводится в волновод. Мощности сигнала и питания несут распространяющиеся в волноводе волны.

Параметрические усилители отличаются чрезвычайно низким уровнем собственного шума. Это обусловлено в основном отсутствием электронных потоков, порождающих шумы типа дробового эффекта. Главным источником шумов в параметрическом усилителе являются активные сопротивления. Нелинейный реактивный элемент дает пренебрежимо малую составляющую шума.

Параметрические усилители — вещь сравнительно новая. Они усиленно разрабатываются в настоящее время, и число описанных в литературе разновидностей уже очень велико, как это всегда бывает в начальной стадии развития новой техники.

В заключение рассмотрим один вопрос, выяснение которого способствует общему пониманию сущности параметрического усиления. Речь пойдет о магнитном усилителе, описанном в § 106 в качестве примера «модуляционного» усиления. В этом усилителе налицо все основные элементы параметрического усилителя. В самом деле: имеется нелинейный реактивный элемент (нелинейная индуктивность); имеется напряжение питания (несущей частоты); происходит модуляция, т.е. преобразование частоты. Однако имеется и одна существенная особенность. Дело в том, что при обычном использовании магнитного усилителя на выход поступает не одна комбинационная частота, а весь модуляционный спектр. Это обстоятельство радикально изменяет положение. Пусть, по-прежнему, ω_1 ($m = 0, n = 1$) означает частоту сигнала, ω_0 ($m = 1, n = 0$) — частоту питания. Имеются две комбинационные частоты: $\omega_0 + \omega_1$ ($m = 1, n = 1$) и $\omega_0 - \omega_1$ ($m = 1, n = -1$), причем **обе** эти частоты поступают на выход. Уравнения (330) дают

(с учетом члена, соответствующего значениям $m = -1$, $n = 1$):

$$\left. \begin{aligned} \frac{P_{11}}{\omega_0 + \omega_1} + \frac{P_{1(-1)}}{\omega_0 - \omega_1} + \frac{P_{10}}{\omega_0} &= 0 \\ \frac{P_{11}}{\omega_0 + \omega_1} - \frac{P_{1(-1)}}{\omega_0 - \omega_1} + \frac{P_{01}}{\omega_1} &= 0 \end{aligned} \right\}. \quad (332)$$

Коэффициент усиления равен

$$\varsigma = -\frac{P_{11} + P_{1(-1)}}{P_{01}} = 1 + \frac{P_{10}}{P_{01}}.$$

Но эта величина не может быть определена из системы уравнений (332), так как имеется два уравнения относительно трех неизвестных:

$$\frac{P_{11}}{P_{01}}, \quad \frac{P_{1(-1)}}{P_{01}} \quad \text{и} \quad \frac{P_{10}}{P_{01}}.$$

Таким образом, в рассматриваемом случае коэффициент усиления не может быть определен в общем виде на основании уравнений Мэнли–Роу. Поэтому коэффициент усиления магнитного усилителя при использовании всего модуляционного спектра приходится определять путем анализа данной конкретной схемы.

Прохождение периодических колебаний сложной формы через линейные цепи ¹⁾

Для изучения прохождения периодических сигналов сложной формы могут применяться классический метод решения дифференциальных уравнений, метод операционного исчисления и метод, основанный на использовании аппарата рядов Фурье.

Метод, основанный на использовании аппарата рядов Фурье, является адекватным рассматриваемой задаче и, благодаря работам последних лет ²⁾ должен быть признан наиболее простым и наглядным. Этим методом легко пользоваться при изучении систем высокого порядка; он одинаково хорошо приспособлен как для аналитических, так и для численных расчетов. Программирование расчетов на электронно-вычислительных машинах по этому методу производится непринужденно.

Простейшая схема изучения состоит в разложении входного воздействия в ряд Фурье (гармонический анализ), видоизменении каждой составляющей в соответствии с амплитудно-частотной и фазо-частотной характеристиками системы и суммировании выходных составляющих (гармонический синтез). Трудности, связанные с реализацией операции гармонического синтеза, в существенной степени могут быть преодолены с помощью подробных таблиц сумм тригонометрических рядов, а также системы правил для перехода от сумм одних рядов к суммам других [1].

Если характеристики системы заданы в виде экспериментальных графиков или таблиц, соответствующие расчеты выполняются численно. Процедура подобных расчетов подробно изложена в [1].

Если характеристики системы заданы в виде аналитических выражений, т. е. известно выражение комплексного коэффициента передачи, соответствующие расчеты выполняются аналитически.

От комплексного коэффициента передачи не составляет труда перейти к дифференциальному уравнению.

Рабочим правилом для составления дифференциальных уравнений является следующее: в выражении комплексного коэффициента пере-

¹⁾ Приложение 6 составлено А. М. Заездным.

²⁾ Эти работы объединены в монографии А. М. Заездного «Гармонический синтез в радиотехнике и электросвязи». — Госэнергоиздат, 1961. Ссылки на данную книгу в тексте настоящего приложения обозначаются [1].

дачи, записанном в виде отношения двух рациональных полиномов

$$N(i\omega) = \frac{\dot{X}_2}{\dot{X}_1} = \frac{A_k(i\omega)^k + A_{k-1}(i\omega)^{k-1} + \dots + A_1(i\omega) + A_0}{B_m(i\omega)^m + B_{m-1}(i\omega)^{m-1} + \dots + B_1(i\omega) + B_0},$$

нужно заменить $(i\omega)^n$ на $\frac{d^n x}{dt^n}$, т. е.

$$\begin{aligned} B_m \frac{d^m x_2}{dt^m} + B_{m-1} \frac{d^{m-1} x_2}{dt^{m-1}} + \dots + B_1 \frac{dx_2}{dt} + B_0 x_2 = \\ = A_k \frac{d^k x_1}{dt^k} + A_{k-1} \frac{d^{k-1} x_1}{dt^{k-1}} + \dots + A_1 \frac{dx_1}{dt} + A_0 x_1. \end{aligned}$$

Решение общей задачи прохождения периодических колебаний сложной формы через линейные системы любого порядка может быть сведено к сумме решений простейших задач, в которых встречается необходимость суммирования только одного тригонометрического ряда с наипростейшей для данной задачи функциональной структурой коэффициентов.

Приведем вывод соответствующей формулы.

Пусть задано дифференциальное уравнение m -го порядка с периодической правой частью:

$$\frac{d^m y}{dx^m} + a_{m-1} \frac{d^{m-1} y}{dx^{m-1}} + \dots + a_1 \frac{dy}{dx} + a_0 y = \frac{\alpha_0}{2} + \sum_{n=1}^{\infty} (\alpha_n \cos nx + \beta_n \sin nx) \quad (333)$$

и нужно найти периодическое решение $y(x)$.

Запишем правую часть в комплексном виде:

$$\frac{\alpha_0}{2} + \sum_{n=1}^{\infty} (\alpha_n \cos nx + \beta_n \sin nx) = \frac{\alpha_0}{2} + \operatorname{Re} \sum_{n=1}^{\infty} \dot{\gamma}_n e^{inx}, \quad (334)$$

где, как обычно

$$\dot{\gamma}_n = \gamma_n e^{-i\psi_n}; \quad \gamma_n = \sqrt{\alpha_n^2 + \beta_n^2}; \quad \psi_n = \operatorname{arctg} \frac{\beta_n}{\alpha_n}.$$

Будем искать частное периодическое решение в виде ряда

$$y(x) = \frac{\alpha_0}{2a_0} + \operatorname{Re} \sum_{n=1}^{\infty} \dot{C}_n e^{inx}. \quad (335)$$

Подставляя принятое решение в (333) с учетом (334), находим, что

$$\dot{C}_n = \frac{\dot{\gamma}_n}{p(in)}, \quad (336)$$

где $p(in)$ представляет собой характеристический оператор

$$p(in) = a_0 + a_1(in) + a_2(in)^2 + \dots + a_{m-1}(in)^{m-1} + (in)^m. \quad (337)$$

Предполагается, что $p(in) \neq 0$.

Подставляя теперь значение c_n в выражение (335) и избавляясь от мнимости в знаменателе, находим

$$y(x) = \frac{\alpha_0}{2a_0} + \operatorname{Re} \sum_{n=1}^{\infty} \frac{\dot{\gamma}_n e^{i n x}}{|p(i n)|^2} p^*(i n) = \frac{\alpha_0}{2a_0} + \operatorname{Re} \sum_{n=1}^{\infty} \frac{\dot{\gamma}_n e^{i n x}}{|p(i n)|^2} [a_0 - a_1(i n) + a_2(i n)^2 + \dots + (-1)^m (i n)^m].$$

Вводя обозначение

$$\operatorname{Re} \sum_{n=1}^{\infty} \frac{\dot{\gamma}_n e^{i n x}}{|p(i n)|^2} = \sum_{n=1}^{\infty} \frac{a_n \cos nx + \beta_n \sin nx}{|p(i n)|^2} = z(x) \quad (338)$$

и имея в виду, что

$$(e^{i n x})^{(m)} = (i n)^m e^{i n x},$$

представим периодическое решение $y(x)$ в следующем окончательном виде

$$y = \frac{\alpha_0}{2a_0} + a_0 z(x) - a_1 z'(x) + a_2 z''(x) - \dots + (-1)^{m-1} a_{m-1} z^{(m-1)}(x) + (-1)^m z^{(m)}(x). \quad (339)$$

Таким образом, для нахождения периодического решения $y(x)$ достаточно найти вспомогательную функцию $z(x)$ по формуле (338) и затем составить линейную комбинацию функции $z(x)$ и ее производных по формуле (339). При таком решении, вообще говоря, нет необходимости в определении корней характеристического оператора¹⁾.

С другой стороны, знание корней характеристического оператора позволяет свести задачу решения дифференциального уравнения m -го порядка к суперпозиции m решений дифференциального уравнения 1-го порядка с такой же правой частью. Соответствующий вывод дан в [1]; приведем здесь только расчетную формулу:

$$y_m = \sum_{i=1}^m \frac{y_{1P_i}}{\prod_{\substack{k=1 \\ k \neq i}}^m (p_i - p_k)}. \quad (340)$$

В этой формуле i и k — номера корней характеристического оператора (337), а p_i , p_1 и p_k — значения этих корней; y_{1P_i} — периодическое решение дифференциального уравнения 1-го порядка, причем корень характеристического оператора этого уравнения соответствует p_i .

¹⁾ Такая необходимость может встретиться только в том случае, когда для представления в замкнутом виде ряда (338) нужно осуществить разложение данного ряда на сумму простейших.

На основе решения (340) процедуру расчета установившихся и переходных процессов можно рационализировать, составив раз и навсегда таблицу периодических решений дифференциального уравнения первого порядка при типовых воздействиях.

Нужно заметить, что нахождение периодических решений дает возможность рассчитать реакцию системы и на непериодическое воздействие. С этой целью непериодическое воздействие следует искусственно сделать периодическим таким образом, чтобы информация о реакции на периодическое воздействие содержала информацию о реакции на непериодическое воздействие, что всегда легко выполнимо. Например, при изучении переходной характеристики, т.е. реакции на единичную функцию, нужно в качестве расчетного воздействия выбрать периодическую последовательность прямоугольных импульсов с достаточно большим периодом повторения.

Примеры пользования изложенной методикой приведены в [1].

Список рекомендуемой литературы

1. Автоматизация проектирования радиоэлектронных средств: Учебное пособие для вузов / Под ред. Алексеева О. В. — М.: Высшая школа, 2000.
2. Амиантов И. Н., Антонов-Антипов Ю. Н., Васильев В. П., Сифоров В. И. Радиоприемные устройства: Учебник для радиотехнических специальностей вузов / Под ред. Сифорова В. И. — М.: Советское радио, 1974.
3. Антонов-Антипов Ю. Н., Васильев В. П., Комаров И. В., Разевиг В. Д., Сифоров В. И. Сборник задач и упражнений по курсу «Радиоприемные устройства»: Учебное пособие для вузов по специальности «Радиотехника» / Под ред. Сифорова В. И. — М.: Радио и связь, 1984.
4. Астайкин А. И., Помазков А. П. Теоретические основы радиотехники. Часть I–III. — Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2004.
5. Ашкинази Л. А. Электронные лампы: из прошлого в будущее. — М.: Физматлит, 2007.
6. Баскаков С. И. Радиотехнические цепи и сигналы. — Изд. 5-е. — М.: Высшая школа, 2005.
7. Бренёв В. И. Начало радиотехники в России. — М.: Советское радио, 1970.
8. Головин О. В., Чистяков Н. И., Шварц В., Хардон Агиляр И. Радиосвязь / Под ред. проф. О. В. Головина. — М.: Горячая линия–Телеком, 2001.
9. Горелик Г. С. Колебания и волны. — М.: Физматлит, 2007.
10. Далматов А. Д., Елисеев А. А., Лукошкин А. П., Оводенко А. А., Устинов Б. В. Обработка сигналов в радиотехнических системах: Учебное пособие / Под ред. А. П. Лукошкина. — Л.: ЛГУ, 1987.
11. Журавлев А. К. Адаптивные радиотехнические системы с антенными решетками. — Л.: Изд-во ЛГУ, 1990.
12. Изобретение радио: Попов А. С. Документы и материалы. — М.: Наука, 1996.
13. Манаев Е. И. Основы радиотехники. — Изд. 3-е. — М.: Радио и связь, 1990.
14. Матвеев А. В., Минченко И. С., Митрофанов А. В. Цифровые микропроцессорные устройства приема и обработки информации: Учебное пособие. — Л.: ЛЭТИ, 1990.
15. Митрофанов А. В., Полевой В. В., Соловьев А. А. Устройства генерирования и формирования радиосигналов: Учебное пособие. — СПб.: ГЭТУ-ЛЭТИ, 1999.
16. Никитин Г. И. Космические системы обнаружения терпящих бедствие. Учеб. пособие. — СПб.: ГУАП, 2001.
17. Никитин Г. И. Спутниковые системы связи. Учеб. пособие. — СПб.: ГУАП, 2000.
18. Островитянов Р. В., Басалов Ф. А. Статистическая теория радиолокации протяженных цепей. — М.: Советское радио, 1980.
19. Полрадз А., Набар Р., Гор Д. Основы построения систем связи с пространственно-временным кодированием. — М.: Физматлит, 2007.

20. Проектирование радиопередающих устройств с применением ЭВМ/Под ред. Алексеева О. В. — М.: Радио и связь, 1987.
21. Родионов В. М. Зарождение радиотехники. — М.: Наука, 1985.
22. Романюк Ю. А. Основы цифровой обработки сигналов. Часть 1. — М.: Физтехполиграф, 2005.
23. Очерки истории радиотехники: Сб. — М.: Изд-во АН СССР, 1960.
24. 100 лет радио: Сб. — М.: Связьиздат, 1995.
25. Сифоров В. И. Радиоприемные устройства. — Изд. 5-е. — М.: Воениздат, 1954.
26. Титце У., Шенк К. Полупроводниковая схемотехника. — М.: Мир, 1982.
27. Харкевич А. А. Избранные труды: В 3 т. — М.: Наука, 1973.
28. Харкевич А. А. Автоколебания. — М.: ГИТТЛ, 1953.
29. Харкевич А. А. Очерки общей теории связи. — М.: ГИТТЛ, 1955.
30. Харкевич А. А. Спектры и анализ. — Изд. 3-е. — М.: ГИТТЛ, 1957.
31. Харкевич А. А. Теоретические основы радиосвязи. — М.: ГТТИ, 1957.
32. Харкевич А. А. Борьба с помехами. — Изд. 2-е. — М.: Наука, 1965.
33. Хоровиц П., Хилл У. Искусство схемотехники: Пер. с англ. — Изд. 6-е. — М.: Мир, 2003.
34. Шеннон К. Работы по теории информации и кибернетике/Пер. с англ.; Предисл. А. Н. Колмогорова. — М.: Изд-во иностр. лит., 1963.