

Учреждение образования
«БЕЛОРУССКАЯ ГОСУДАРСТВЕННАЯ АКАДЕМИЯ СВЯЗИ»
Кафедра радио и информационных технологий

КОНСПЕКТ ЛЕКЦИЙ

по дисциплине
«ТЕОРИЯ ЭЛЕКТРИЧЕСКОЙ СВЯЗИ»

для учащихся специальности
2-45 01 32 «Системы радиосвязи, радиовещания и телевидения»

Минск 2016

ВВЕДЕНИЕ

Структурная схема системы электросвязи

При характеристике систем электросвязи используются понятия: информация, сообщение, сигнал.

Информация (от лат. informatio – разъяснение, изложение) – совокупность сведений о каких-либо событиях, явлениях или предметах, предназначенных для передачи, приема, обработки, преобразования, хранения или непосредственного использования.

Можно выделить три основных вида информации в обществе:

- личную (касается тех или иных событий в личной жизни человека);
- специальную (к ней относится научно-техническая, деловая, производственная, экономическая и др.);
- массовую (предназначена для большой группы людей и распространяется через СМИ: газеты, журналы, радио, телевидение и др.).

Примеры: сведения о времени приезда приятеля, сведения о результате футбольного матча.

Сообщение - форма представления информации, подлежащей передаче.

Сообщения можно разбить на два типа:

- непрерывные (аналоговые) (принимают любые значения в некотором интервале).

Пример: речь, музыка, подвижные и неподвижные изображения;

- дискретные (принимают конечное число возможных значений). Пример: текст, данные ЭВМ.

Сигнал (от лат. signum – знак) – физический процесс, отображающий (несущий) передаваемое сообщение. Он всегда является функцией времени, даже если сообщение (например, неподвижное изображение) таковым не является.

По своей физической природе сигналы бывают электрическими, световыми, звуковыми и др.

Электрический сигнал – форма представления сообщения для передачи его системой электросвязи. Электрические сигналы количественно можно характеризовать мощностью, напряжением или током.

Система электросвязи – совокупность технических средств и среды распространения, обеспечивающих передачу сообщений от источника к потребителю. В это понятие включаются передающее устройство, линия связи и приемное устройство.

Система связи называется *одноканальной*, если она обеспечивает передачу сообщения от одного источника к одному получателю по одной линии связи. Одноканальные системы являются малоэффективными, т.к. полоса частот, в которой работает линия связи, намного превышает ширину спектра первичных сигналов.

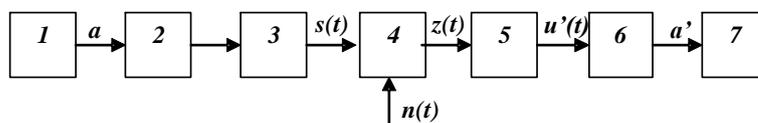


Рисунок 1.1 – Структурная схема одноканальной системы связи.

Многоканальные системы передачи сигналов

Система связи называется *многоканальной*, если она обеспечивает одновременную и независимую передачу сообщений от нескольких источников к нескольким получателям по одной общей линии связи.

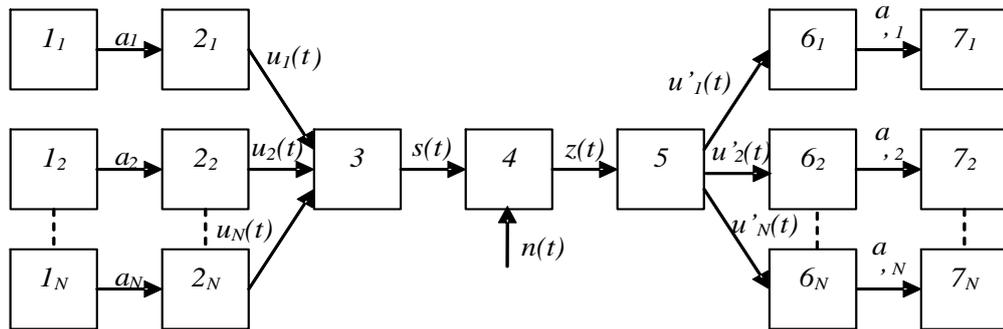


Рисунок 1.2 – Структурная схема многоканальной системы связи.

Рассмотрим назначение структурных элементов схем, представленных на рисунках 1.1 и 1.2.

1 (I_i) – источник сообщения – человек или техническое устройство, формирующее передаваемое сообщение a (a_i).

2 (2_i) – преобразователь сообщения в сигнал – устройство, которое преобразует сообщение в первичный сигнал (низкочастотный) $u(t)$ ($u_i(t)$). Примеры: передающая часть телеграфного аппарата, микрофон, преобразователь свет-сигнал на приборах с зарядовой связью.

3 – преобразователь сигнала (передатчик). В одноканальной системе это устройство, которое преобразует первичный сигнал во вторичный сигнал (высокочастотный) $s(t)$, удобный для передачи по линии связи. В многоканальной системе связи это устройство, в котором первичные сигналы преобразуются в каналные, которые затем объединяются в групповой сигнал, направляемый в линию связи:

$$s(t) = \sum_{i=1}^N s_i(t),$$

где $s_i(t)$ – каналные сигналы – сигналы, однозначно связанные с первичными сигналами $u_i(t)$ и обладающие определенными признаками, позволяющими разделить их на приеме;

N – число каналов в системе.

Элементы 2 (2_i) и 3 образуют передающее устройство.

4 – линия связи – среда, используемая для передачи сигналов от передатчика к приемнику. Различают линии связи:

- проводные (электромагнитное поле распространяется вдоль непрерывной направляющей среды).

Примеры: воздушные и кабельные линии, волноводы, световоды;

- радиолнии (электромагнитные волны распространяются в свободном пространстве). Примеры: радиорелейные и спутниковые линии.

При прохождении по линии связи электрические сигналы подвергаются воздействию помех $n(t)$ и искажений. Это приводит к тому, что сигнал на выходе линии связи $z(t)$ и принятое сообщение a' (a'_i) могут отличаться от сигнала на входе линии связи и передаваемого сообщения.

Степень соответствия принятого сообщения переданному называют верностью передачи сообщения.

Канал электросвязи – совокупность технических средств и среды распространения, обеспечивающих передачу первичных сигналов между двумя пунктами. Элементы 3, 4 и 5 образуют канал (каналы) связи.

5 - преобразователь сигнала (приёмник). В одноканальной системе связи это устройство, которое по принятому вторичному сигналу восстанавливает первичный сигнал $u'(t)$. В многоканальной системе это устройство, которое из измененного искажениями и помехами группового сигнала выделяет канальные сигналы $s'_i(t)$, которые затем преобразуются в первичные сигналы $u'_i(t)$.

6 (6_i) – преобразователь сигнала в сообщение – устройство, которое преобразует первичный сигнал в принятое сообщение $a' (a'_i)$.

Примеры: приемная часть телеграфного аппарата, телефон, громкоговоритель, кинескоп.

Элементы 5 и 6 образуют приемное устройство.

7 (7_i) – получатель сообщения – человек или техническое устройство, воспринимающее сообщение.

СИГНАЛЫ И КАНАЛЫ ЭЛЕКТРОСВЯЗИ

Классификация и характеристики сигналов электросвязи

По форме различают простые и сложные сигналы.

Простые сигналы представляют собой такие функции времени, которые можно выразить в виде простой математической формулы.

Примеры простых сигналов: гармонические; постоянные; описываемые единичной функцией; описываемые дельта-функцией.

Гармоническими являются сигналы, описываемые функцией синуса или косинуса:

$$a(t) = A_m \cos(\omega t + \varphi) \text{ или } a(t) = A_m \sin(\omega t + \phi).$$

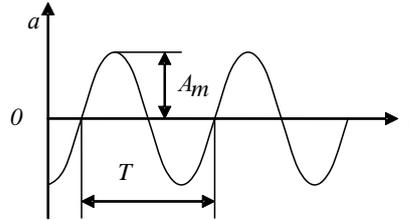


Рисунок 2.1 – Гармоническое колебание.

Параметры: амплитуда A_m ; частота: $f = \omega / 2\pi = 1/T$,

где ω - угловая частота. Размерность: $[\omega] = \text{рад/с}$; f - циклическая частота.

Размерность: $[f] = \text{Гц}$; T - период. Размерность: $[T] = \text{с}$; $\varphi, \phi = \pi/2 + \varphi$ - начальная фаза.

Постоянными являются сигналы, значения которых в любой момент времени остаются неизменными:

$$a(t) = A_m.$$

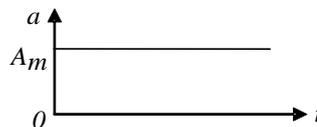


Рисунок 2.2 – Постоянный сигнал.

Единичная функция является математическим описанием ступенчатого перепада напряжения или тока:

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

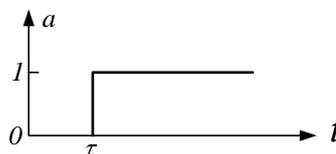


Рисунок 2.3 – Единичная функция.

Дельта-функция является математическим описанием прямоугольного импульса малой длительности и большой амплитуды:

$$\delta(t - \tau) = \begin{cases} 0, & t \neq \tau; \\ \infty, & t = \tau. \end{cases}$$

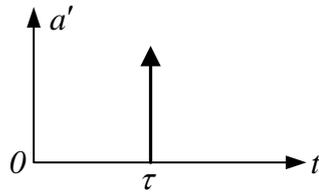


Рисунок 2.4 – Дельта-функция.

Сложные сигналы представляют собой такие функции времени, которые трудно выразить в виде простой математической формулы. Сложный сигнал может быть представлен совокупностью элементарных (простых) сигналов в виде обобщенного ряда Фурье:

$$a(t) = \sum_{n=0}^{\infty} a_n \psi_n(t),$$

где a_n - коэффициенты разложения, зависящие от сигнала $a(t)$;

$\psi_n(t)$ - базисные функции – функции, имеющие простое аналитическое выражение, позволяющие легко вычислить коэффициенты a_n и обеспечивающие быструю сходимость ряда к сигналу $a(t)$. В электросвязи наибольшее применение в качестве базисных функций получили гармонические колебания.

Примеры сложных сигналов: импульсные; используемые для представления сообщений.

Импульсными являются сигналы, отличные от нуля в течение ограниченного времени. Наибольшее применение находят одиночные прямоугольные импульсы (ОПИ) и периодические последовательности прямоугольных импульсов (ПППИ).

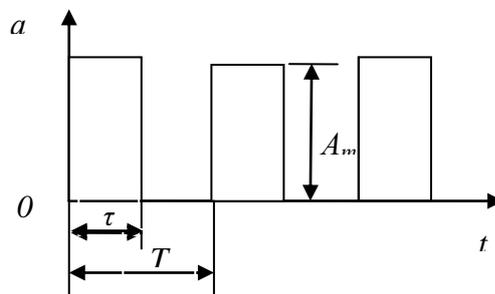


Рисунок 2.5 – ПППИ.

Параметры: A_m – амплитуда; τ – длительность импульса; T – период; $q=T/\tau$ – скважность.

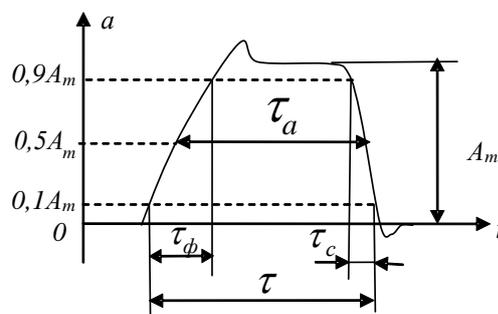


Рисунок 2.6 – Реальный импульс прямоугольной формы.

Параметры: A_m – амплитуда; τ – длительность импульса; τ_a – активная длительность импульса; τ_ϕ – длительность фронта; τ_c – длительность спада.

По информативности различают детерминированные и случайные сигналы.

Детерминированными называют сигналы, мгновенные значения которых в любые моменты времени заранее известны. Для их математического описания служат детерминированные математические модели. Такие сигналы не являются переносчиками информации. Используются в качестве несущих колебаний для получения модулированных сигналов, испытательных сигналов для испытаний системы связи или отдельных ее элементов.

Примеры детерминированных сигналов: гармонические сигналы с известными параметрами; импульсы с известными формой и параметрами.

Различают следующие типы детерминированных сигналов:

- периодические – сигналы, мгновенные значения которых повторяются через определенные равные промежутки времени, называемые периодом;
- непериодические – сигналы, которые появляются только один раз и более не повторяются.

Случайными называют сигналы, мгновенные значения которых в любые моменты времени заранее не известны. Для их математического описания служат вероятностные математические модели. Только случайные сигналы являются переносчиками информации. Реальные сигналы всегда случайны.

Примеры случайных сигналов: телеграфные, телефонные, радиовещательные, факсимильные, телевизионные, передачи данных.

По характеристикам различают четыре вида сигналов:

- непрерывные по уровню и по времени (сокращенно непрерывные или аналоговые). Принимают любые значения в некотором интервале и изменяются в произвольные моменты времени;

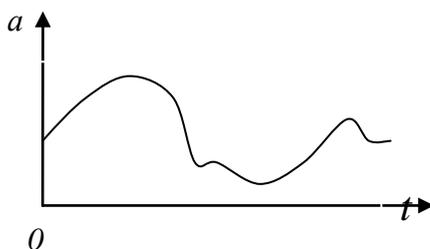


Рисунок 2.7 – Непрерывный сигнал.

- непрерывные по уровню, дискретные по времени (сокращенно дискретные по времени). Принимают произвольные значения в некотором интервале, но изменяются только в определенные, наперед заданные (дискретные) моменты времени;

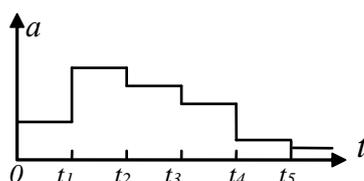


Рисунок 2.8 – Дискретный по времени сигнал.

- дискретные по уровню, непрерывные по времени (сокращенно дискретные по уровню). Принимают только разрешенные (дискретные) значения в произвольные моменты времени;

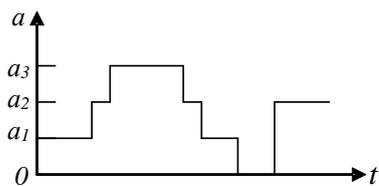


Рисунок 2.9 – Дискретный по уровню сигнал.

- дискретные по уровню и по времени (сокращенно дискретные). Принимают только дискретные значения в дискретные моменты времени.

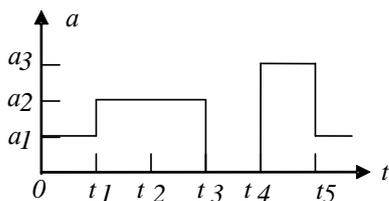


Рисунок 2.10 – Дискретный сигнал.

Цифровые сигналы – разновидность дискретных сигналов, когда разрешенные уровни некоторого исходного дискретного сигнала представлены в виде цифр. В системах связи применяются двоичные, троичные, четверичные и т.д. n -ичные цифровые сигналы.

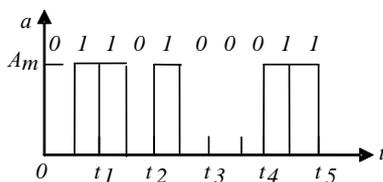


Рисунок 2.11 – Двоичный цифровой сигнал.

Полезный сигнал является объектом транспортировки (передачи), а техника связи – техникой транспортирования (передачи) сигналов по каналам связи. Основные параметры сигнала с точки зрения его передачи:

Длительность сигнала T_c - интервал времени, в пределах которого сигнал существует;

Ширина спектра сигнала ΔF_c - диапазон частот, в пределах которого сосредоточена основная часть энергии сигнала;

Динамический диапазон D_c - диапазон изменения уровней сигнала. Выражается в децибелах (дБ):

$$D_c = 10 \lg(P_{\max} / P_{\min}),$$

где P_{\max} и $P_{\min} \geq P_{ш}$ - максимальное и минимальное значения мгновенной мощности;

$P_{ш}$ - средняя мощность шумов в канале.

Динамический диапазон может быть определен не только по мощности, но и по напряжению или току:

$$D_c = 20 \lg(U_{\max} / U_{\min}) \text{ или } D_c = 20 \lg(I_{\max} / I_{\min}),$$

где U_{\max}, I_{\max} и U_{\min}, I_{\min} - максимальные и минимальные значения мгновенного напряжения или тока.

Обобщающей характеристикой является ОБЪЕМ сигнала:

$$V_c = T_c \Delta F_c D_c.$$

Объем сигнала пропорционален объему информации, им переносимой. Чем больше объем сигнала, тем больше информации можно «заложить» в этот объем и тем труднее передать такой сигнал по каналу связи с требуемым качеством.

Другие характеристики сигнала:

- база сигнала: $B_c = 2T_c \Delta F_c$. Если $B_c \leq I$, то сигнал называют узкополосным, если $B_c \gg I$ – широкополосным;

- коэффициент амплитуды сигнала: $K_A^2 = 10 \lg(P_{\max} / P_{cp})$,

где P_{cp} - средняя мощность сигнала.

Способы представления сигналов

Математическая модель сигнала

Это его математическое описание, т.е. получение относительно простого математического выражения (формулы, уравнения, неравенства и др.), по которому можно вычислить свойства и параметры сигнала (мгновенные значения, числовые характеристики и др.).

Пример:
$$u(t) = U_{m1} \cos(\omega_1 t + \varphi_1) + U_{m2} \cos(\omega_2 t + \varphi_2).$$

Выбор математической модели осуществляется на основе анализа временной диаграммы. Один и тот же сигнал может быть представлен несколькими моделями.

Достоинства: удобство при теоретических расчетах.

Недостатки: трудность подбора при сложной форме сигнала.

Временная диаграмма сигнала

Это кривая мгновенных значений сигнала, выполненная в зависимости от времени (графическое представление формы сигнала).

Временная форма представления сигнала позволяет определить такие важные характеристики, как его энергия, мощность и длительность. Энергия сигнала, расположенного на интервале $[t_1, t_2]$:

$$E_u = \int_{t_1}^{t_2} u^2(t) dt.$$

Именно такая энергия выделится на резисторе с сопротивлением 1 Ом, если на его зажимы подано напряжение $u(t)$. Энергия периодического сигнала неограниченно велика. Поэтому здесь нужно говорить о мощности сигнала, т.е. об энергии в единицу времени.

Пример:

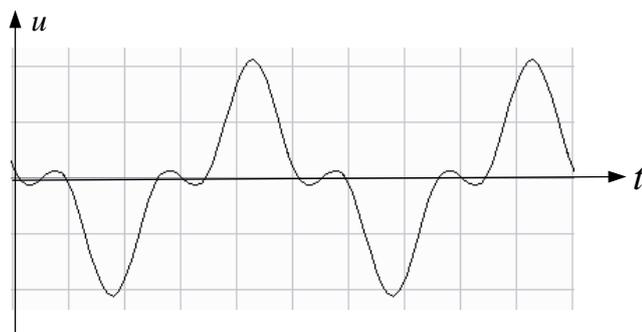


Рисунок 2.12 – Временная диаграмма.

Временную диаграмму можно наблюдать с помощью осциллографа.

Достоинства: наглядность.

Недостатки: неудобство при теоретических расчетах и при представлении длительных сигналов.

Спектральная диаграмма сигнала

Это графическое изображение его спектра.

Если какой-либо сигнал представлен в виде суммы гармонических колебаний с различными частотами, то говорят, что осуществлено спектральное разложение этого сигнала. Совокупность отдельных гармонических компонент сигнала образуют его спектр.

Различают амплитудные и фазовые спектральные диаграммы.

Амплитудная спектральная диаграмма сигнала – диаграмма распределения по частоте амплитуд гармонических составляющих сигнала. Амплитуды гармоник могут принимать только положительные значения. В них заключена важная информация о распределении энергии сигнала по различным частотным составляющим.

Пример:

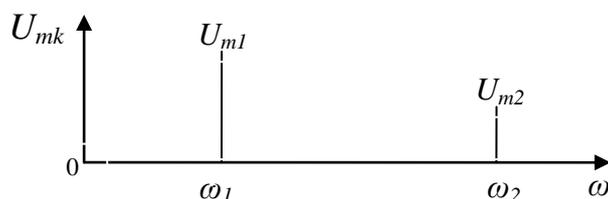


Рисунок 2.13 – Амплитудная спектральная диаграмма.

Фазовая спектральная диаграмма сигнала – диаграмма распределения по частоте фаз гармонических составляющих сигнала. Фазы гармоник могут принимать как положительные, так и отрицательные значения в интервале $[-\pi, \pi]$.

Пример:

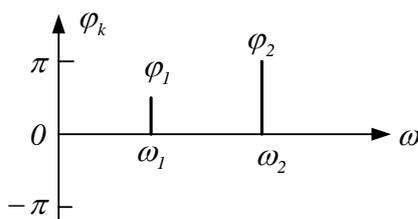


Рисунок 2.14 – Фазовая спектральная диаграмма.

Спектральные диаграммы можно наблюдать с помощью анализатора спектра.

Достоинства: знание спектра позволяет осуществить неискаженную передачу сигнала по каналу связи, обеспечить разделение сигналов и ослабление помех.

Недостатки: форма сигнала определяется в совокупности как амплитудными, так и фазовыми составляющими спектра.

Векторная диаграмма сигнала

Это изображение токов и напряжений на координатной плоскости через векторы, сопоставленные гармоническим колебаниям.

Векторы, представляющие на координатной плоскости гармонические колебания разных частот, будут вращаться против хода часов вокруг начала координат с разными угловыми скоростями. Их модули определяются амплитудами колебаний, а углы наклона в момент начала отсчета – значениями начальных фаз. Проекции векторов на ось абсцисс будут представлять собой косинусоидальные колебания, на ось ординат – синусоидальные. Они укажут, как будут изменяться во времени мгновенные значения токов и напряжений.

Пример:

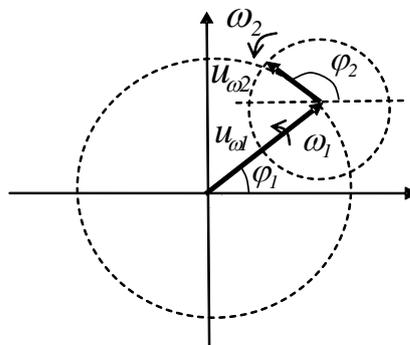


Рисунок 2.15 – Векторная диаграмма.

Достоинства: наглядность интерпретации спектрального разложения.

Недостатки: неудобство при представлении поведения во времени напряжения или тока.

Спектры сигналов

Вид спектра зависит от характера сигнала.

Различают дискретные (линейчатые) и непрерывные (сплошные) спектры.

Спектр будет *дискретным*, если сигнал можно рассматривать как конечную сумму или ряд (бесконечную сумму) гармонических колебаний. Он представляется на спектральной диаграмме набором отдельных спектральных линий (линий, соответствующих амплитудам и фазам гармоник).

Дискретным спектром обладают периодические и квазипериодические (почти периодические) сигналы.

Периодическими называются сигналы, значения которых повторяются через определенные равные промежутки времени, называемые периодом. Особенность спектров периодических сигналов состоит в том, что частоты гармоник в них кратны основной частоте $\omega_1 = 2\pi/T$. Такие спектры называют гармоническими.

Разновидностью непериодических сигналов являются квазипериодические. Это сигналы, значения которых приближенно повторяются через определенные промежутки времени, называемые почти периодом.

Пример: $a(t) = \sin(\omega t) + \sin(\sqrt{2}\omega t)$.

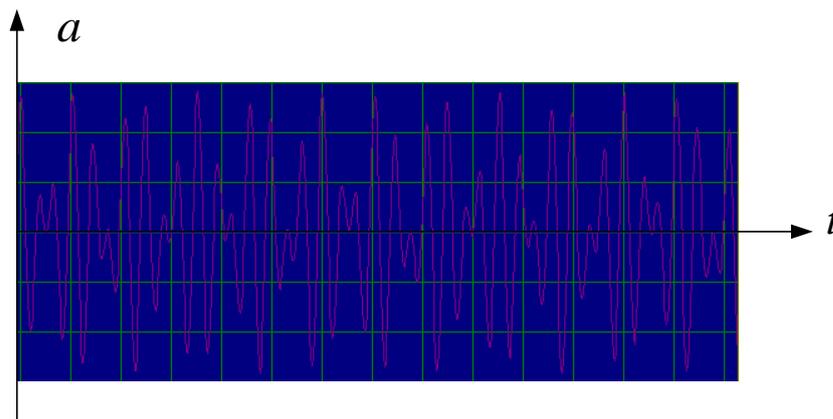


Рисунок 2.16 – Квазипериодический сигнал.

Особенность спектров квазипериодических сигналов состоит в том, что частоты гармоник в них не находятся в кратном отношении. Такие спектры не являются гармоническими.

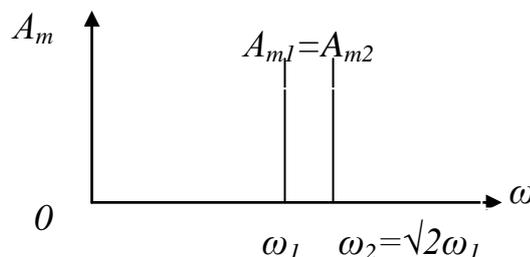


Рисунок 2.17 – Дискретный спектр.

Спектр будет *непрерывным*, если сигнал можно рассматривать как сумму бесконечно большого числа бесконечно малых по амплитуде гармонических колебаний, частоты которых располагаются бесконечно близко друг к другу. По оси ординат амплитудной спектральной диаграммы откладывается спектральная плотность амплитуд, фазовой – спектральная плотность фаз.

Непрерывным спектром обладают непериодические сигналы (одиночные импульсы, информационные сигналы), а также хаотические колебания (шумы).

Спектральная плотность амплитуд $S(\omega)$ - функция частоты, несущая информацию об амплитуде элементарных гармоник. Она характеризует интенсивность сплошного распределения амплитуд гармоник сигнала вдоль оси частот. Спектральная плотность амплитуд сигнала на любой частоте ω равна суммарной амплитуде спектральных составляющих, попадающих в малую полосу $\Delta\omega$ в окрестности частоты ω , пересчитанной к полосе 1 Гц:

$$S(\omega) = \lim_{\Delta f \rightarrow 0} \Delta A / \Delta f .$$

Имеет размерность В/Гц или А/Гц.

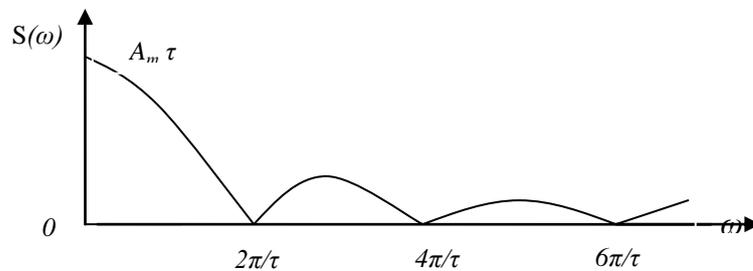


Рисунок 2.18 – Амплитудная спектральная диаграмма ОПИ.

Спектральная плотность фаз $\varphi(\omega)$ – функция частоты, несущая информацию о фазе элементарных гармоник.

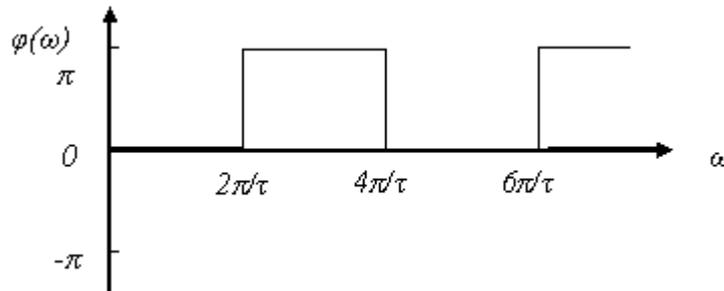


Рисунок 2.19 – Фазовая спектральная диаграмма ОПИ.

Различают ограниченные и неограниченные спектры.

Ограниченным называют спектр, имеющий конечный интервал частот, в котором расположены все спектральные линии периодического либо квазипериодического сигнала или на котором отлична от нуля спектральная плотность непериодического сигнала. Если этот интервал бесконечен, то спектр является *неограниченным*

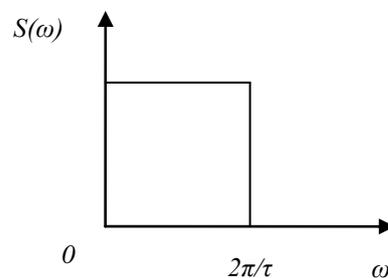


Рисунок 2.20 – Ограниченный спектр.

Классификация и характеристики каналов связи

Ключевым понятием техники электросвязи является канал электросвязи.

КАНАЛ ЭЛЕКТРОСВЯЗИ – комплекс технических средств и среды распространения, обеспечивающий передачу первичных сигналов электросвязи от преобразователя сообщения в первичный сигнал до преобразователя первичного сигнала в сообщении.

Классификация каналов связи

1 По виду передаваемых первичных сигналов (сообщений) различают каналы:

- телеграфные;
- телефонные;
- звукового вещания;
- телевизионные;

- передачи данных.

2 По характеру сигналов на входе и выходе канала различают каналы:

- дискретные (на входе и выходе канала действуют дискретные сигналы);
- непрерывные (аналоговые) (на входе и выходе канала действуют непрерывные (по уровням) сигналы);
- дискретно-непрерывные или непрерывно-дискретные (на входе канала действует дискретный сигнал, а на выходе – непрерывный (по уровням) или наоборот).

3 По методам разделения канальных сигналов различают каналы:

- с простейшими методами разделения (первичные сигналы передаются без каких-либо преобразований в исходном диапазоне частот).

Пример: с разделением методом уравнированного моста.

- с более совершенными методами разделения (первичные сигналы преобразуются в канальные, наделенные определенными отличительными признаками).

Пример: с линейным разделением (разделяющие устройства являются линейными 4-полюсниками) (с временным разделением, с частотным разделением, с разделением по фазе, с разделением по форме); с нелинейным разделением (разделяющие устройства являются нелинейными 4-полюсниками) (с разделением по уровню, с комбинационным разделением).

4 По занимаемой полосе частот различают каналы:

- узкополосные (занимают узкую полосу частот). *Пример:* канал тональной частоты (300...3400 Гц);
- широкополосные (занимают широкую полосу частот, в них могут разместиться несколько узкополосных).

Пример: канал передачи сигналов изображения телевидения (50...6500000 Гц).

5 По виду среды распространения сигналов электросвязи различают каналы:

- проводной связи (организованы по проводным линиям связи: воздушным, кабельным, волноводным, световодным);
- радиосвязи (организованы по радиолиниям связи: радиорелейным, спутниковым).

6 По взаимосвязи между сигналами на входе и выходе канала различают каналы:

- с детерминированной взаимосвязью (по заданной реализации входного сигнала можно точно определить соответствующую ей реализацию выходного сигнала);
- с вероятностной взаимосвязью (при одной и той же реализации входного сигнала можно наблюдать различные реализации выходного сигнала).

Характеристики каналов связи

Каналы связи характеризуются тремя параметрами:

- временем использования T_k (временем, в течение которого по каналу ведется передача сигнала);
- динамическим диапазоном D_k (выраженное в децибелах отношением максимальной неискаженной мощности сигнала, которая может быть передана по каналу, к минимальной мощности сигнала, при которой обеспечивается необходимая защищенность от помех);
- полосой пропускания ΔF_k – полоса частот, которую канал способен пропустить с выполнением требований к качеству передачи сигнала).

Произведение трех параметров канала называется его емкостью:

$$V_k = T_k D_k \Delta F_k.$$

Условие согласования сигнала с каналом: сигнал может быть передан по каналу, если его емкость не менее объема сигнала

$$V_c \leq V_k.$$

В простейшем случае сигнал согласуется с каналом по всем трем параметрам:

$$T_c \leq T_k, D_c \leq D_k, \Delta F_c \leq \Delta F_k.$$

Однако возможно и несоблюдение одного или двух неравенств при обеспечении главного. Это достигается обменом одного параметра на другой. Большой интерес представляет возможность обмена динамического диапазона на полосу пропускания.

Искажения и помехи в каналах

Сигнал на выходе канала связи отличается от переданного из-за искажений и помех, что является причиной воспроизведения сообщения с некоторой погрешностью (ошибкой).

Искажения в канале

Искажения – нежелательные изменения формы сигнала, которые обусловлены известными характеристиками канала, по которому проходит сигнал.

Различают искажения:

- *Линейные* (возникают в линейных звеньях канала). Линейные искажения делятся на амплитудно-частотные и фазочастотные. Причиной амплитудно-частотных искажений является нарушение соотношений между амплитудами, а фазочастотных – начальными фазами гармонических составляющих в спектре выходного сигнала по сравнению со спектром входного сигнала. Эти искажения определяются формой амплитудно-частотной и фазочастотной характеристик канала;

- *Нелинейные* (возникают в нелинейных звеньях канала). Причиной нелинейных искажений является появление в спектре выходного сигнала гармонических составляющих, которых не было в спектре входного сигнала. Величину нелинейных искажений можно приблизительно оценить по амплитудной характеристике канала.

Т.к. искажения обусловлены известными характеристиками канала, то они в принципе могут быть устранены. Для этого используется коррекция характеристик канала.

Помехи в канале

Помеха – любое мешающее воздействие на сигнал, вызывающее случайные отклонения принятого сигнала от передаваемого.

1 По месту возникновения различают помехи:

- внешние (возникают вне канала связи):
 - атмосферные. Связаны с электрическими процессами в земной атмосфере. Источники: грозовые разряды, полярные сияния, пылевые бури;
 - космические. Связаны с электромагнитными процессами на внеземных объектах. Источники: Солнце, звезды, межзвездные газы;
 - промышленные. Связаны с деятельностью человека. Источники: промышленные установки, медицинские установки, электротранспорт, линии электропередач;
 - от посторонних средств связи. Связаны с нарушением регламента распределения рабочих частот, недостаточной стабильностью генераторов, плохой фильтрацией побочных гармоник излучаемых сигналов;

- внутренние (возникают в самом канале). Источники: тепловой шум (возникает в проводниках из-за теплового движения носителей зарядов) и дробовый шум (возникает на выходе электровакуумных и полупроводниковых приборов и обусловлен дискретной природой носителей заряда).

2 По характеру воздействия на сигнал различают помехи:

- аддитивные – помехи, мгновенные значения которых складываются с мгновенными значениями сигналов:

$$z(t) = s(t) + n(t),$$

где $z(t)$ - принимаемое колебание;

$s(t)$ - передаваемый сигнал;

$n(t)$ - помеха.

Имеют место даже если сигнал отсутствует;

- мультипликативные – помехи, мгновенные значения которых перемножаются с мгновенными значениями сигналов:

$$z(t) = s(t) \cdot n(t).$$

Ощущаются только при наличии сигнала.

3 По форме различают помехи:

- флуктуационные (распределенные по частоте и времени) - непрерывные колебания, меняющиеся случайным образом. Спектр помех весьма широкий;

- гармонические (сосредоточенные по частоте) – гармонические или модулированные колебания с шириной спектра меньшей или соизмеримой с шириной спектра полезного сигнала;

- импульсные (сосредоточенные по времени) – помехи в виде одиночных коротких импульсов различной интенсивности и длительности, следующих один за другим через случайные достаточно большие промежутки времени.

Помехи заранее неизвестны и поэтому не могут быть устранены полностью.

СПЕКТРАЛЬНЫЙ АНАЛИЗ ДЕТЕРМИНИРОВАННЫХ СИГНАЛОВ

Анализ спектра периодических сигналов

Периодический сигнал любой формы с периодом T может быть представлен в виде суммы

гармонических колебаний с разными амплитудами и начальными фазами, частоты которых кратны основной частоте $\omega_1 = 2\pi/T$. Гармонику этой частоты называют основной или первой, остальные – высшими гармониками.

Тригонометрическая форма ряда Фурье:

$$a(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\omega_1 t + b_n \sin n\omega_1 t),$$

где $a_0 = \frac{2}{T} \int_{-T/2}^{T/2} u(t) dt$ - постоянная составляющая;

$a_n = \frac{2}{T} \int_{-T/2}^{T/2} u(t) \cos n\omega_1 t dt$ - амплитуды косинусоидальных составляющих;

$b_n = \frac{2}{T} \int_{-T/2}^{T/2} u(t) \sin n\omega_1 t dt$ - амплитуды синусоидальных составляющих.

Четный сигнал ($a(-t) = a(t)$) имеет только косинусоидальные, а нечетный ($a(-t) = -a(t)$) - только синусоидальные слагаемые.

Более удобной является эквивалентная тригонометрическая форма ряда Фурье:

$$a(t) = A_0 + \sum_{n=1}^{\infty} A_{mn} \cos(n\omega_1 t + \varphi_n),$$

где $A_0 = \frac{a_0}{2}$ - постоянная составляющая;

$A_{mn} = \sqrt{a_n^2 + b_n^2}$ - амплитуда n -ой гармоники сигнала. Совокупность амплитуд гармонических составляющих носит название спектра амплитуд;

$\varphi_n = -\arctg\left(\frac{b_n}{a_n}\right)$ - начальная фаза n -ой гармоники сигнала. Совокупность фаз гармонических составляющих носит название спектра фаз.

Рассчитаем амплитудный и фазовый спектры ПППИ, имеющих амплитуду U_m , длительность τ , период следования T и расположенных симметрично относительно начала координат (сигнал – четная функция).

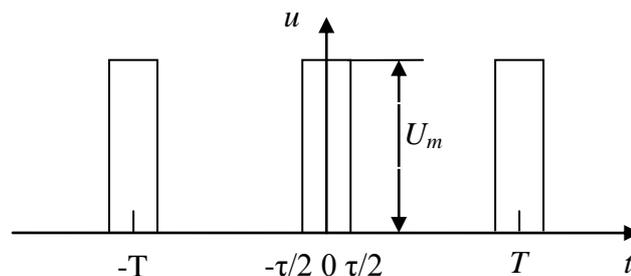


Рисунок 3.1 – Временная диаграмма ПППИ.

Сигнал на интервале одного периода можно записать:

$$u(t) = \begin{cases} U_m, & -\tau/2 \leq t \leq \tau/2; \\ 0, & |t| > \tau/2. \end{cases}$$

Вычисления:

$$U_0 = \frac{a_0}{2} = \frac{1}{T} \int_{-T/2}^{T/2} u(t) dt = \frac{1}{T} \int_{-\tau/2}^{\tau/2} U_m dt = \frac{2U_m}{T} \int_0^{\tau/2} dt = \frac{2U_m}{T} t \Big|_0^{\tau/2} = \frac{2U_m}{T} \frac{\tau}{2} = \frac{U_m \tau}{T} = \frac{U_m}{q},$$

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} u(t) \cos n\omega_1 t dt = \frac{2}{T} \int_{-\tau/2}^{\tau/2} U_m \cos n\omega_1 t dt = \frac{4U_m}{T} \int_0^{\tau/2} \cos n\omega_1 t dt =$$

$$= \frac{4U_m}{T n\omega_1} \sin n\omega_1 t \Big|_0^{\tau/2} = \frac{4U_m}{T n\omega_1} \sin(n\omega_1 \tau/2) = \frac{4U_m \tau/2}{T} \frac{\sin(n\omega_1 \tau/2)}{n\omega_1 \tau/2} = \frac{2U_m}{q} \frac{\sin(n\pi/q)}{n\pi/q},$$

$$b_n = \frac{2}{T} \int_{-T/2}^{T/2} u(t) \sin n\omega_1 t dt = \frac{2}{T} \int_{-\tau/2}^{\tau/2} U_m \sin n\omega_1 t dt = 0,$$

$$U_{mn} = \sqrt{a_n^2 + b_n^2} = |a_n| = \frac{2U_m}{q} \left| \frac{\sin(n\pi/q)}{n\pi/q} \right|,$$

$$\varphi_n = -\operatorname{arctg} \frac{b_n}{a_n} = -\operatorname{arctg} \frac{0}{a_n} = \begin{cases} 0, & a_n > 0, \\ \pi, & a_n < 0. \end{cases}$$

Ряд Фурье для ПППИ имеет вид:

$$u(t) = \frac{U_m}{q} + \sum_{n=1}^{\infty} \frac{2U_m}{q} \left| \frac{\sin(n\pi/q)}{n\pi/q} \right| \sin(n\omega_1 t + \varphi_n).$$

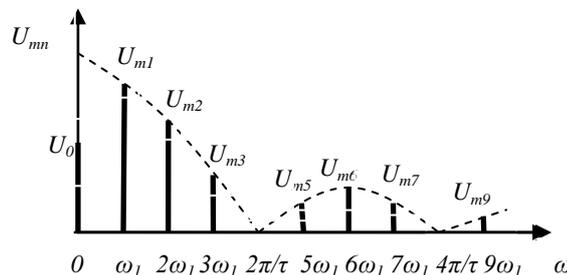


Рисунок 3.2 – Амплитудная спектральная диаграмма ПППИ.

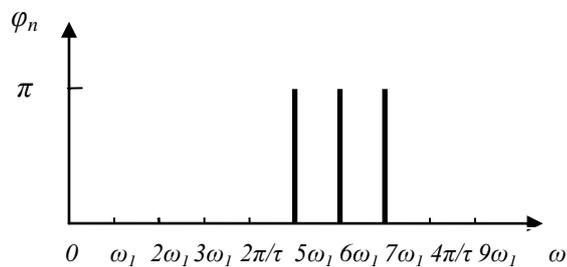


Рисунок 3.3 – Фазовая спектральная диаграмма ПППИ.

Выводы:

- спектр ПППИ линейчатый (дискретный) (представляется набором отдельных спектральных линий), гармонический (спектральные линии находятся на одинаковом расстоянии друг от друга ω_1), убывающий (амплитуды гармоник убывают с ростом их

номера), имеет лепестковую структуру (ширина каждого лепестка равна $2\pi/\tau$), неограниченный (интервал частот, в котором располагаются спектральные линии, бесконечен);

- при целочисленных скважностях частотные составляющие с частотами, кратными скважности в спектре отсутствуют (их частоты совпадают с нулями огибающей спектра амплитуд);

- с увеличением скважности амплитуды всех гармонических составляющих уменьшаются. При этом если оно связано с увеличением периода повторения T , то спектр становится плотнее (ω_1 уменьшается), с уменьшением длительности импульса τ – становится больше ширина каждого лепестка;

- за ширину спектра ПППИ принят интервал частот, содержащий 95% энергии сигнала, (равен ширине двух первых лепестков огибающей):

$$\Delta\omega = 4\pi/\tau \text{ или } \Delta f = 2/\tau;$$

- все гармоники, находящиеся в одном лепестке огибающей, имеют одинаковые фазы, равные либо 0 либо π .

Анализ спектра непериодических сигналов

Сигналы связи всегда ограничены во времени и поэтому не являются периодическими. Среди непериодических сигналов наибольший интерес представляют одиночные импульсы (ОИ). ОИ можно рассматривать как предельный случай периодической последовательности импульсов (ППИ) длительностью τ при бесконечно большом периоде их повторения $T \rightarrow \infty$.

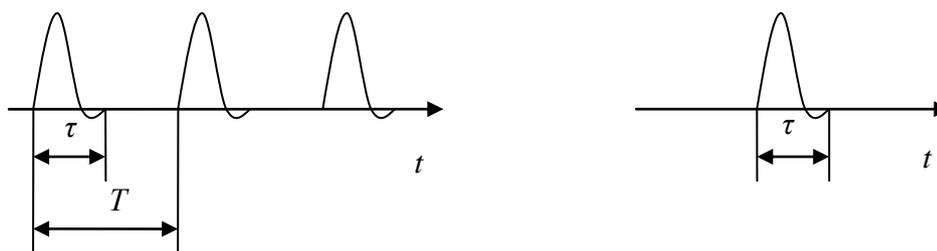


Рисунок 3.4 – ППИ и ОИ.

Непериодический сигнал может быть представлен суммой бесконечно большого числа бесконечно близких по частоте колебаний с исчезающе малыми амплитудами. Спектр ОИ является непрерывным и вводится интегралами Фурье:

- $S(j\omega) = \int_{-\infty}^{\infty} a(t)e^{-j\omega t} dt$ (1) - прямое преобразование Фурье. Позволяет аналитически

отыскать спектральную функцию по заданной форме сигнала;

- $a(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega)e^{j\omega t} d\omega$ (2) - обратное преобразование Фурье. Позволяет аналитически

отыскать форму по заданной спектральной функции сигнала.

Комплексная форма интегрального преобразования Фурье (2) дает двустороннее спектральное представление (имеющее отрицательные частоты) непериодического сигнала $a(t)$ в виде суммы гармонических колебаний $e^{j\omega t}$ с бесконечно малыми комплексными амплитудами $S(j\omega)/T$, частоты которых непрерывно заполняют всю ось частот.

$$S(j\omega) = \int_{-\infty}^{\infty} a(t)e^{-j\omega t} dt = \int_{-\infty}^{\infty} a(t)[\cos \omega t - j \sin \omega t] dt = \int_{-\infty}^{\infty} a(t) \cos \omega t dt - j \int_{-\infty}^{\infty} a(t) \sin \omega t dt =$$

$= A(\omega) - jB(\omega) = S(\omega)e^{j\varphi(\omega)}$ - комплексная спектральная плотность сигнала – комплексная функция частоты, одновременно несущая информацию как об амплитуде, так и о фазе элементарных гармоник.

Модуль спектральной плотности $S(\omega) = \sqrt{[A(\omega)]^2 + [B(\omega)]^2}$ называется спектральной плотностью амплитуд. Его можно рассматривать как АЧХ сплошного спектра непериодического сигнала.

Аргумент спектральной плотности $\varphi(\omega) = -\arctg \frac{B(\omega)}{A(\omega)}$ называется спектральной плотностью фаз. Его можно рассматривать как ФЧХ сплошного спектра непериодического сигнала.

Преобразуем формулу (2):

$$a(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega)e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega)e^{j(\omega t + \varphi(\omega))} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \cos(\omega t + \varphi(\omega)) d\omega +$$

$$+ j \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \sin(\omega t + \varphi(\omega)) d\omega = \left| \begin{array}{l} S(\omega) - \text{четная функция частоты;} \\ \varphi(\omega) - \text{нечетная функция частоты} \end{array} \right| =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \cos(\omega t + \varphi(\omega)) d\omega = \frac{1}{\pi} \int_0^{\infty} S(\omega) \cos(\omega t + \varphi(\omega)) d\omega.$$

Тригонометрическая форма интегрального преобразования Фурье дает одностороннее спектральное представление (не имеющее отрицательных частот) непериодического сигнала:

$$a(t) = \frac{1}{\pi} \int_0^{\infty} S(\omega) \cos(\omega t + \varphi(\omega)) d\omega.$$

Найдем амплитудный и фазовый спектр ОПИ с известными параметрами A_m , τ , четного относительно точки $t = 0$.

Математическая модель ОПИ:

$$a(t) = \begin{cases} A_m & \text{при } |t| \leq \frac{\tau}{2}; \\ 0 & \text{при } |t| > \frac{\tau}{2}. \end{cases}$$

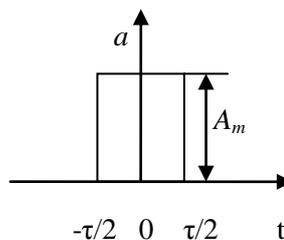


Рисунок 3.5 – Временная диаграмма ОПИ.

Найдем спектральную плотность ОПИ:

$$S(j\omega) = \int_{-\infty}^{\infty} a(t)e^{-j\omega t} dt = \int_{-\tau/2}^{\tau/2} A_m e^{-j\omega t} dt = A_m \int_{-\tau/2}^{\tau/2} (\cos \omega t - j \sin \omega t) dt = 2A_m \int_0^{\tau/2} \cos \omega t dt = \frac{2A_m}{\omega} \sin \omega t \Big|_0^{\tau/2} =$$

$$= \frac{2A_m}{\omega} \sin(\omega \tau / 2) = 2A_m \tau / 2 \frac{\sin(\omega \tau / 2)}{\omega \tau / 2} = A_m \tau \frac{\sin(\omega \tau / 2)}{\omega \tau / 2}.$$

Спектральная плотность амплитуд:

$$S(\omega) = |S(j\omega)| = A_m \tau \left| \frac{\sin(\omega \tau / 2)}{\omega \tau / 2} \right|.$$

Спектральная плотность фаз:

$$\varphi(\omega) = \arg S(j\omega) = \begin{cases} 0 & \text{при } S(j\omega) > 0; \\ \pi & \text{при } S(j\omega) < 0. \end{cases}$$

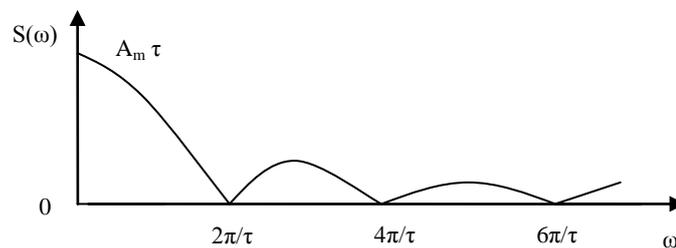


Рисунок 3.6 – Амплитудная спектральная диаграмма ОПИ.

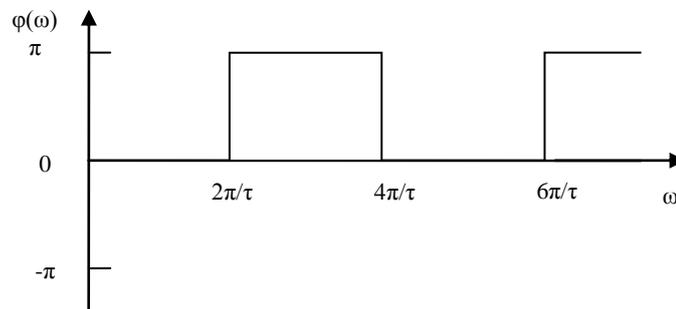


Рисунок 3.7 – Фазовая спектральная диаграмма ОПИ.

Выводы:

- спектр ОПИ сплошной (содержит непрерывную последовательность спектральных составляющих), убывающий (по мере роста частоты спектральная плотность уменьшается), неограниченный (спектральная плотность амплитуд, начинаясь в области низких частот, уходит в область бесконечно больших частот), имеет лепестковую структуру;

- спектральная плотность амплитуд ОПИ и огибающая линейчатого спектра ПППИ совпадают по форме и отличаются только масштабом:

$$S(\omega) = \lim_{f_1 \rightarrow 0} (A_{mm} / (2f_1)).$$

Это правило относится к импульсам любой формы;

- изменение длительности импульса приводит к пропорциональному растягиванию или сжатию спектральной функции $S(\omega)$ вдоль оси частот;

- фазовый спектр ОПИ представляет собой ступенчатую кривую, изменяющуюся скачком на величину π в точках, где $S(\omega)$ проходит через нуль;

- за ширину спектра ОПИ принят интервал частот, в котором заключено 90,2% энергии импульса:

$$\Delta\omega = 2\pi\Delta f = 2\pi / \tau .$$

ПРЕОБРАЗОВАНИЕ СИГНАЛОВ В НЕЛИНЕЙНЫХ ЦЕПЯХ

Аппроксимация характеристик нелинейных элементов

Аппроксимация нелинейной характеристики – замена истинной (экспериментально полученной) характеристики приближенно представляющей ее функцией.

Необходимость аппроксимации возникает при анализе, синтезе и расчете нелинейных цепей. Для упрощения аппроксимируют не всю характеристику НЭ, а только ее рабочий участок (используемую часть).

Различают способы аппроксимации:

- полиномиальная;
- кусочно-линейная;
- с помощью трансцендентных функций.

Выбор способа аппроксимации зависит от вида нелинейной характеристики, а также от режима работы НЭ. Выбранный способ аппроксимации должен обеспечивать достаточные полноту описания свойств НЭ, простоту математической обработки полученного представления характеристики и точность расчета.

График характеристики НЭ будем показывать сплошной линией, а график аппроксимирующей функции – штриховой.

Полиномиальная аппроксимация

Является одним из наиболее распространенных способов аппроксимации. Заключается в представлении нелинейной характеристики в виде полинома (многочлена) n -ой степени относительно рабочей точки (x_0, y_0) :

$$y - y_0 = a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n,$$

где $y_0 = y(x_0) = a_0, a_1, a_2, \dots, a_n$ - коэффициенты полинома. Зависят от положения рабочей точки на характеристике;

n - порядок полинома. Определяется требуемой точностью расчетов.

Примеры:

1) Полином второй степени:

$$i(u) = i(U_0) + a_1(u - U_0) + a_2(u - U_0)^2 -$$

- используется, если рабочая точка (определяется постоянным напряжением U_0) расположена на начальном участке характеристики, имеющем вид квадратичной параболы, и подводимое к НЭ напряжение сигнала не выходит за начало характеристики (за точку U_n , которая определяется из условия: $i(U_n)=0$).

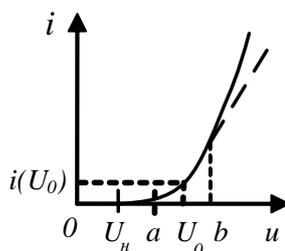


Рисунок 4.4 – Характеристика, для аппроксимации которой требуется полином второй степени.

Используемые обозначения:

- $i(U_0)$ – ток покоя;

- (a, b) – используемый участок ВАХ.

2) Неполный полином третьей степени:

$$i(u) = i(U_0) + a_1(u - U_0) + a_3(u - U_0)^3 -$$

- используется, если рабочая точка является точкой перегиба характеристики и напряжение сигнала не выходит за пределы напряжения насыщения $\pm U_{max}$.

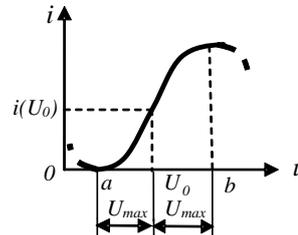


Рисунок 4.5 – Характеристика, для аппроксимации которой требуется полином третьей степени.

3) Полином высокой (пятой и более) степени используется, если рабочая точка находится на нижнем сгибе характеристики и изменение напряжения сигнала велико.

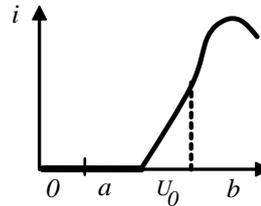


Рисунок 4.6 – Характеристика, для аппроксимации которой требуется полином высокой степени.

Кусочно-линейная аппроксимация

Применяется при очень больших амплитудах входного сигнала. Заключается в замене реальной характеристики идеализированной, линейно-ломаной, составленной из отрезков прямых линий, являющихся касательными к характеристике.

Примеры:

1) Два линейных отрезка используется, если сигнал захватывает нижний сгиб и линейный участок характеристики.

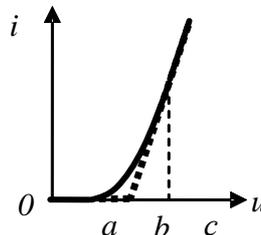


Рисунок 4.7 – Характеристика при числе аппроксимирующих отрезков 2. Используемое обозначение: (a, c) – используемый участок ВАХ.

2) Три линейных отрезка используется, если сигнал захватывает нижний и верхний сгибы характеристики.

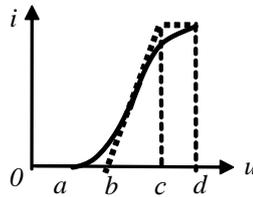


Рисунок 4.8 – Характеристика при числе аппроксимирующих отрезков 3. Используемое обозначение: (a, d) – используемый участок ВАХ.

3) Четыре линейных отрезка используется, если сигнал достигает также и падающего участка характеристики.

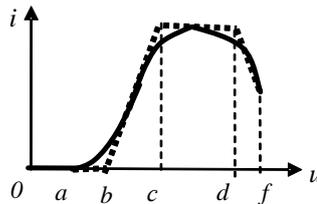


Рисунок 4.9 – Характеристика при числе аппроксимирующих отрезков 4. Используемое обозначение: (a, f) – используемый участок ВАХ.

Аппроксимация с помощью трансцендентных функций

Применяется при более сложной форме используемого участка характеристики. Заключается в замене реальной характеристики трансцендентной (неалгебраической) функцией. Неалгебраическими являются функции: экспоненциальная ($y = e^x$), гиперболическая ($y = k/x$), показательная ($y = a^x$), логарифмическая ($y = \log_a x$), тригонометрические ($y = \sin x$, $y = \cos x$, $y = \operatorname{tg} x$, $y = \operatorname{ctg} x$), обратные тригонометрические ($y = \arcsin x$, $y = \arccos x$, $y = \operatorname{arctg} x$, $y = \operatorname{arcctg} x$).

Пример:

ВАХ полупроводникового диода в пределах от $-(0,3 \dots 1)$ до $+(0,25 \dots 0,6)$ В достаточно точно может быть представлена экспонентой (рисунок 15.7):

$$i = I_0(e^{\alpha u} - 1),$$

где $I_0 = 100$ нА (для германиевого диода) или $I_0 = 10$ пА (для кремниевого диода) – теоретический обратный ток;

$1/\alpha = kT/q_e$ - тепловой (термический потенциал);

$q_e = 1,6 \cdot 10^{19}$ Кл – заряд электрона;

$k = 1,38 \cdot 10^{-23}$ Дж/К – постоянная Больцмана;

T – абсолютная температура.

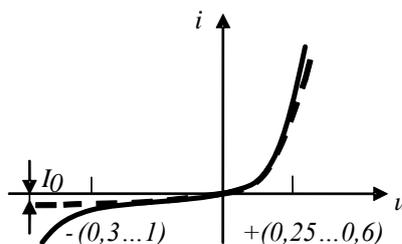


Рисунок 4.10 – Характеристика, для аппроксимации которой требуется экспоненциальная функция.

Анализ спектра отклика нелинейного элемента на гармоническое воздействие

Гармонический анализ отклика НЭ осуществляется при воздействии на него гармонического колебания со «смещением», представляемого выражением:

$$u(t) = U_0 + U_{m1} \cos \omega_1 t.$$

Цель гармонического анализа отклика – определение его спектрального состава. При этом имеется в виду, что НЭ безынерционный. Под таким НЭ подразумевается любой электронный прибор с нелинейной ВАХ при использовании его в диапазоне частот, на которых можно пренебречь влиянием паразитных параметров (внутренних емкостей и индуктивностей).

Наиболее распространенные методы анализа и случаи их использования приведены в таблице 4.1.

Таблица 4.1 – Методы спектрального анализа.

Метод спектрального анализа	Способ аппроксимации ВАХ
С использованием тригонометрических формул кратного аргумента (аналитический)	Полиномиальный
Угла отсечки (графический)	Кусочно-линейный
С использованием формул трех и пяти ординат (графоаналитический)	Не требует аппроксимации характеристики НЭ
С использованием функций Бесселя от мнимого аргумента	Экспонента или сумма экспонент

При *методе угла отсечки* используется кусочно-линейная аппроксимация характеристики НЭ. Форма реакции находится графическим методом (методом проекций), который заключается в построении третьей проекции $y(t)$ (реакции НЭ) по известным двум: воздействию $x(t)$ и характеристике $y = f(x)$. При построении графика $y(t)$ сначала наносят характерные точки: максимумы, минимумы, пересечения с осью абсцисс, - а затем промежуточные точки. Спектральный состав реакции определяется при разложении ее в ряд Фурье.

Метод кратных аргументов основывается на получении выражения реакции $y(t)$ путем подстановки в степенной полином, которым представлена нелинейная характеристика $y = f(x)$, выражения воздействия $x(t)$, представленного в виде ряда Фурье. После элементарных преобразований с учетом известных тригонометрических формул кратных аргументов и группировки слагаемых с одинаковыми аргументами получают выражение отклика в виде ряда.

Слабонелинейный режим работы НЭ

Рассмотрим режим работы, при котором напряжение сигнала $u(t)$ не выходит за пределы точки начала характеристики U_n (рисунок 4.4) и ВАХ удовлетворительно аппроксимируется степенным полиномом третьей степени:

$$i(u) = a_0 + a_1(u - U_0) + a_2(u - U_0)^2 + a_3(u - U_0)^3,$$

где $u(t) = U_0 + U_{m1} \cos \omega_1 t$ - входной сигнал.

Подставим в заданный полином выражение входного сигнала:

$$\begin{aligned} i(t) &= a_0 + a_1 U_{m1} \cos \omega_1 t + a_2 (U_{m1} \cos \omega_1 t)^2 + a_3 (U_{m1} \cos \omega_1 t)^3 = \\ &= a_0 + a_1 U_{m1} \cos \omega_1 t + a_2 U_{m1}^2 \cos^2 \omega_1 t + a_3 U_{m1}^3 \cos^3 \omega_1 t. \end{aligned}$$

Применяя тригонометрические формулы кратных аргументов:

$$\cos^2 \alpha = 1/2 + 1/2 \cos 2\alpha,$$

$$\cos^3 \alpha = 3/4 \cos \alpha + 1/4 \cos 3\alpha, \quad -$$

избавимся от степеней тригонометрических функций:

$$i(t) = a_0 + a_1 U_{m1} \cos \omega_1 t + 1/2 a_2 U_{m1}^2 + 1/2 a_2 U_{m1}^2 \cos 2\omega_1 t + 3/4 a_3 U_{m1}^3 \cos \omega_1 t + 1/4 a_3 U_{m1}^3 \cos 3\omega_1 t.$$

Сгруппируем слагаемые с одинаковым аргументом косинуса:

$$i(t) = (a_0 + 1/2 a_2 U_{m1}^2) + (a_1 U_{m1} + 3/4 a_3 U_{m1}^3) \cos \omega_1 t + 1/2 a_2 U_{m1}^2 \cos 2\omega_1 t + 1/4 a_3 U_{m1}^3 \cos 3\omega_1 t.$$

Заменим коэффициенты обозначением тока:

$$I_0 = a_0 + 1/2 a_2 U_{m1}^2 \text{ - постоянная составляющая;}$$

$$I_{m1} = a_1 U_{m1} + 3/4 a_3 U_{m1}^3 \text{ - амплитуда первой гармоники;}$$

$$I_{m2} = 1/2 a_2 U_{m1}^2 \text{ - амплитуда второй гармоники;}$$

$$I_{m3} = 1/4 a_3 U_{m1}^3 \text{ - амплитуда третьей гармоники.}$$

Отклик представим в виде:

$$i(t) = I_0 + I_{m1} \cos \omega_1 t + I_{m2} \cos 2\omega_1 t + I_{m3} \cos 3\omega_1 t.$$

Представим воздействие и отклик графически.

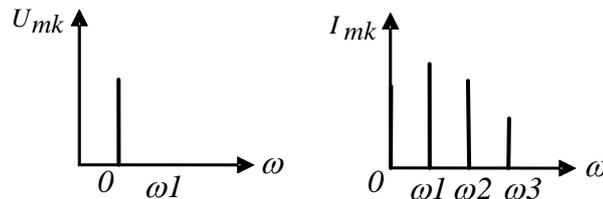


Рисунок 4.11 – Спектральные диаграммы гармонического воздействия и отклика на него.

Выводы:

- спектр отклика при воздействии гармонического сигнала линейчатый;
- частоты составляющих спектра кратны частоте входного сигнала;
- кратность частоты высшей гармоники спектра равна степени полинома;
- постоянная составляющая и четные гармоники определяются только четными степенями напряжения в полиноме;
- нечетные гармоники определяются только нечетными степенями напряжения в полиноме.

Отметим, что в спектре отклика появились составляющие, отсутствовавшие в спектре входного воздействия. Эти новые составляющие являются результатом реакции нелинейной цепи и называются нелинейными продуктами, характеризующими нелинейные искажения входного сигнала.

Рассмотренный метод используется при анализе работы усилителей, работающих в нелинейном режиме, т.е. когда допустим уровень нелинейных продуктов выше 10%.

Существенно нелинейный режим работы НЭ

Рассмотрим режим работы, получаемый при сдвиге рабочей точки U_0 влево и увеличении амплитуды возбуждающего напряжения (рисунок 4.7). В данном случае целесообразно применить кусочно-линейную аппроксимацию ВАХ:

$$i = \begin{cases} 0 & \text{при } u < U_n, \\ S(u - U_n) & \text{при } u \geq U_n, \end{cases}$$

где S - крутизна линейно возрастающего участка ВАХ,
 U_n - координата его начала.

Форму реакции находим графическим методом. Типичное взаимное расположение ВАХ и сигналов показано на рисунке 4.12.

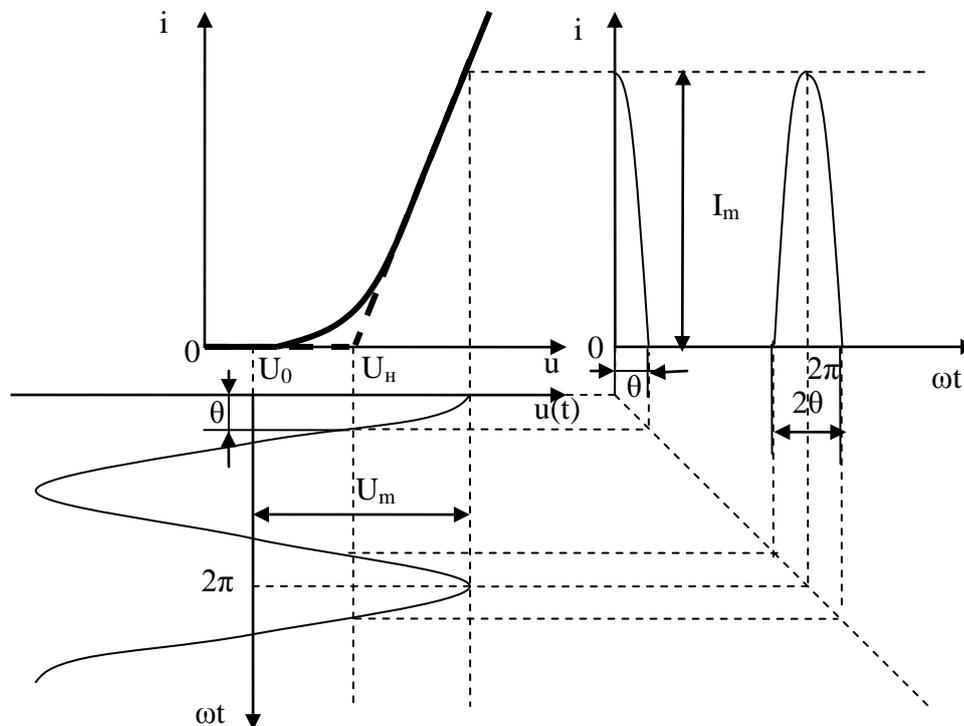


Рисунок 4.12 – Определение формы реакции методом проекций.

Форма реакции имеет вид периодической последовательности косинусоидальных импульсов с отсечкой. Полученные импульсы характеризуются двумя параметрами: высотой I_m и шириной 2θ .

Угол, соответствующий половине времени существования импульса, называется *углом отсечки* θ . Угол отсечки определяется из равенства:

$$\theta = \arccos \frac{U_n - U_0}{U_m}.$$

В соответствии с формулой при заданной ВАХ (фиксированном U_n) угол отсечки θ регулируется выбором амплитуды U_m величины смещения U_0 .

Высота (максимальное значение) импульса тока определяется выражением:

$$I_m = S U_m (1 - \cos \theta).$$

Поскольку ток – периодическая функция времени с периодом $T = 2\pi / \omega_1$, его можно представить в виде ряда Фурье:

$$i(t) = I_0 + \sum_{n=1}^{\infty} I_{mn} \cos(n\omega_1 t + \varphi_n).$$

Коэффициенты этого ряда являются постоянной составляющей и амплитудами гармоник тока и могут быть вычислены по формулам:

$$I_0 = \alpha_0(\theta)I_m = SU_m\gamma_0(\theta),$$

$$I_{mn} = \alpha_n(\theta)I_m = SU_m\gamma_n(\theta),$$

где $\alpha_n(\theta)$ - коэффициенты Берга;

$\gamma_n(\theta) = (1 - \cos\theta)\alpha_n(\theta)$ - функции Берга.

Для ряда значений n коэффициенты и функции Берга табулированы.

Из рассмотрения графиков коэффициентов Берга можно сделать такие заключения: при $\theta = 0$ ток равен нулю (НЭ заперт на протяжении всего периода); при $\theta = 180^\circ$ отсечка тока отсутствует и режим работы становится линейным; при работе с $\theta < 180^\circ$ отношение амплитуды первой гармоники к постоянной составляющей больше единицы $I_{m1}/I_0 > 1$ и растет с уменьшением θ ; с повышением номера гармоники максимумы амплитуд гармоник перемещаются в область малых значений θ .

Указанные обстоятельства существенно влияют на выбор режима работы НЭ при усилении колебаний, умножении частоты. При умножении частоты для получения наибольшей амплитуды нужной гармоники тока (n -ой) необходимо выбрать *оптимальное значение угла отсечки*:

$$\theta_{optn} = 120^\circ / n.$$

Таким образом, вне зависимости от вида аппроксимирующей функции ток через НЭ при гармоническом воздействии представляется суммой постоянной I_0 и гармонических с амплитудами $I_{m1}, I_{m2}, I_{m3}, \dots$ и частотами $\omega_1, 2\omega_1, 3\omega_1, \dots$, кратными частоте приложенного напряжения, составляющих, т.е. рядом Фурье.

Анализ спектра отклика нелинейного элемента на бигармоническое воздействие

Бигармоническое воздействие – это входной сигнал, представляющий собой сумму двух гармонических колебаний с разными частотами:

$$u(t) = U_0 + U_{m1} \cos \omega_1 t + U_{m2} \cos \omega_2 t.$$

При анализе ограничимся третьей степенью аппроксимирующего полинома:

$$i(u) = a_0 + a_1(u - U_0) + a_2(u - U_0)^2 + a_3(u - U_0)^3.$$

Подставим в заданный полином выражение входного сигнала:

$$\begin{aligned} i(t) &= a_0 + a_1(U_{m1} \cos \omega_1 t + U_{m2} \cos \omega_2 t) + a_2(U_{m1} \cos \omega_1 t + U_{m2} \cos \omega_2 t)^2 + a_3(U_{m1} \cos \omega_1 t + U_{m2} \cos \omega_2 t)^3 = \\ &= a_0 + a_1 U_{m1} \cos \omega_1 t + a_1 U_{m2} \cos \omega_2 t + a_2 U_{m1}^2 \cos^2 \omega_1 t + 2a_2 U_{m1} U_{m2} \cos \omega_1 t \cos \omega_2 t + a_2 U_{m2}^2 \cos^2 \omega_2 t + \\ &+ a_3 U_{m1}^3 \cos^3 \omega_1 t + 3a_3 U_{m1}^2 U_{m2} \cos^2 \omega_1 t \cos \omega_2 t + 3a_3 U_{m1} U_{m2}^2 \cos \omega_1 t \cos^2 \omega_2 t + a_3 U_{m2}^3 \cos^3 \omega_2 t. \end{aligned}$$

Применяя тригонометрические формулы кратных аргументов:

$$\cos^2 \alpha = 1/2 + 1/2 \cos 2\alpha,$$

$$\cos^3 \alpha = 3/4 \cos \alpha + 1/4 \cos 3\alpha \quad -$$

и произведения косинусов:

$$\cos \alpha \cos \beta = 1/2 \cos(\alpha - \beta) + 1/2 \cos(\alpha + \beta) \quad , -$$

избавимся от степеней и произведений тригонометрических функций:

$$\begin{aligned}
i(t) = & a_0 + a_1 U_{m1} \cos \omega_1 t + a_1 U_{m2} \cos \omega_2 t + 1/2 a_2 U_{m1}^2 + 1/2 a_2 U_{m1}^2 \cos 2\omega_1 t + a_2 U_{m1} U_{m2} \cos(\omega_1 - \omega_2)t + \\
& + a_2 U_{m1} U_{m2} \cos(\omega_1 + \omega_2)t + 1/2 a_2 U_{m2}^2 + 1/2 a_2 U_{m2}^2 \cos 2\omega_2 t + 3/4 a_3 U_{m1}^3 \cos \omega_1 t + 1/4 a_3 U_{m1}^3 \cos 3\omega_1 t + \\
& + 3/2 a_3 U_{m1}^2 U_{m2} \cos \omega_2 t + 3/4 a_3 U_{m1}^2 U_{m2} \cos(2\omega_1 - \omega_2)t + 3/4 a_3 U_{m1}^2 U_{m2} \cos(2\omega_1 + \omega_2)t + \\
& + 3/2 a_3 U_{m1} U_{m2}^2 \cos \omega_1 t + 3/4 a_3 U_{m1} U_{m2}^2 \cos(\omega_1 - 2\omega_2)t + 3/4 a_3 U_{m1} U_{m2}^2 \cos(\omega_1 + 2\omega_2)t + \\
& + 3/4 a_3 U_{m2}^3 \cos \omega_2 t + 1/4 a_3 U_{m2}^3 \cos 3\omega_2 t.
\end{aligned}$$

Сгруппируем слагаемые с одинаковым аргументом косинуса:

$$\begin{aligned}
i(t) = & (a_0 + 1/2 a_2 U_{m1}^2 + 1/2 a_2 U_{m2}^2) + (a_1 U_{m1} + 3/4 a_3 U_{m1}^3 + 3/2 a_3 U_{m1} U_{m2}^2) \cos \omega_1 t + (a_1 U_{m2} + \\
& + 3/2 a_3 U_{m1}^2 U_{m2} + 3/4 a_3 U_{m2}^3) \cos \omega_2 t + 1/2 a_2 U_{m1}^2 \cos 2\omega_1 t + 1/2 a_2 U_{m2}^2 \cos 2\omega_2 t + 1/4 a_3 U_{m1}^3 \cos 3\omega_1 t + \\
& + 1/4 a_3 U_{m2}^3 \cos 3\omega_2 t + a_2 U_{m1} U_{m2} \cos(\omega_1 - \omega_2)t + a_2 U_{m1} U_{m2} \cos(\omega_1 + \omega_2)t + 3/4 a_3 U_{m1} U_{m2}^2 \cos(\omega_1 - 2\omega_2)t + \\
& + 3/4 a_3 U_{m1} U_{m2}^2 \cos(\omega_1 + 2\omega_2)t + 3/4 a_3 U_{m1}^2 U_{m2} \cos(2\omega_1 - \omega_2)t + 3/4 a_3 U_{m1}^2 U_{m2} \cos(2\omega_1 + \omega_2)t.
\end{aligned}$$

Заменяем коэффициенты обозначением тока:

$$I_0 = a_0 + 1/2 a_2 U_{m1}^2 + 1/2 a_2 U_{m2}^2 - \text{постоянная составляющая};$$

$$I_{m1} = a_1 U_{m1} + 3/4 a_3 U_{m1}^3 + 3/2 a_3 U_{m1} U_{m2}^2 - \text{амплитуда первой гармоники первой частоты};$$

ТЫ;

$$I_{m2} = a_1 U_{m2} + 3/2 a_3 U_{m1}^2 U_{m2} + 3/4 a_3 U_{m2}^3 - \text{амплитуда первой гармоники второй частоты};$$

ТЫ;

$$I_{m21} = 1/2 a_2 U_{m1}^2 - \text{амплитуда второй гармоники первой частоты};$$

$$I_{m22} = 1/2 a_2 U_{m2}^2 - \text{амплитуда второй гармоники второй частоты};$$

$$I_{m31} = 1/4 a_3 U_{m1}^3 - \text{амплитуда третьей гармоники первой частоты};$$

$$I_{m32} = 1/4 a_3 U_{m2}^3 - \text{амплитуда третьей гармоники второй частоты};$$

$$I_{m1-2} = a_2 U_{m1} U_{m2} - \text{амплитуда составляющей разностной частоты } |\omega_1 - \omega_2|;$$

$$I_{m1+2} = a_2 U_{m1} U_{m2} - \text{амплитуда составляющей суммарной частоты } \omega_1 + \omega_2;$$

$$I_{m1-22} = 3/4 a_3 U_{m1} U_{m2}^2 - \text{амплитуда составляющей разностной частоты } |\omega_1 - 2\omega_2|;$$

$$I_{m1+22} = 3/4 a_3 U_{m1} U_{m2}^2 - \text{амплитуда составляющей суммарной частоты } \omega_1 + 2\omega_2;$$

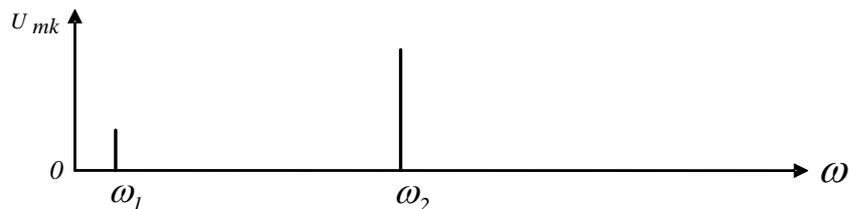
$$I_{m21-2} = 3/4 a_3 U_{m1}^2 U_{m2} - \text{амплитуда составляющей разностной частоты } |2\omega_1 - \omega_2|;$$

$$I_{m21+2} = 3/4 a_3 U_{m1}^2 U_{m2} - \text{амплитуда составляющей суммарной частоты } 2\omega_1 + \omega_2.$$

Отклик представим в виде:

$$\begin{aligned}
i(t) = & I_0 + I_{m1} \cos \omega_1 t + I_{m2} \cos \omega_2 t + I_{m21} \cos 2\omega_1 t + I_{m22} \cos 2\omega_2 t + I_{m31} \cos 3\omega_1 t + \\
& + I_{m32} \cos 3\omega_2 t + I_{1-2} \cos(\omega_1 - \omega_2)t + I_{m1+2} \cos(\omega_1 + \omega_2)t + I_{m1-22} \cos(\omega_1 - 2\omega_2)t + \\
& + I_{m1+22} \cos(\omega_1 + 2\omega_2)t + I_{m21-2} \cos(2\omega_1 - \omega_2)t + I_{m21+2} \cos(2\omega_1 + \omega_2)t.
\end{aligned}$$

Представим воздействие и отклик графически, предположив, что $\omega_1 > \omega_2$.



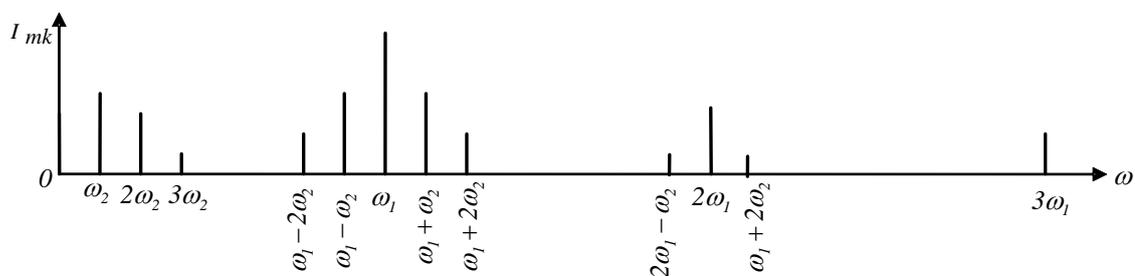


Рисунок 4.13 – Спектральные диаграммы бигармонического воздействия и отклика на него.

Кроме постоянной составляющей и гармоник в составе тока появились *комбинационные частоты* - всевозможные суммарные и разностные частоты, не кратные частотам воздействия. Составляющие с такими частотами возникают только при одновременном воздействии на НЭ не менее двух гармонических колебаний.

Полигармоническое воздействие

Полигармоническое воздействие – это входной сигнал, представляющий собой сумму трех или более гармонических колебаний с различными частотами:

$$u(t) = U_0 + \sum_{i=1}^M U_{mi} \cos \omega_i t,$$

где M - число гармонических колебаний воздействия.

Обобщим полученные ранее результаты. При воздействии на НЭ с ВАХ, аппроксимированной полиномом N -ой степени, напряжения в виде суммы M гармонических сигналов ток будет содержать составляющие с частотами:

$$|k_1 \omega_1 \pm k_2 \omega_2 \pm \dots \pm k_M \omega_M|,$$

где k_1, k_2, \dots, k_M - целые положительные числа из диапазона $0..N$, такие что

$$\sum_{i=1}^M k_i \leq N;$$

сумма коэффициентов при частотах воздействия называется порядком колебания: $R = \sum_{i=1}^M k_i$.

При этом слагаемые степенного полинома четной степени привносят в спектр тока постоянную составляющую, гармоники и комбинационные частоты четных порядков; нечетной степени - нечетных порядков.

Такие функциональные преобразования бигармонических и полигармонических воздействий НЭ используются при модуляции, детектировании и преобразовании частоты.

МОДУЛИРОВАННЫЕ СИГНАЛЫ

Общие понятия о модуляции

Модуляция – это процесс изменения одного или нескольких параметров несущего колебания в соответствии с законом изменения передаваемого сигнала (модулирующего сигнала). Модулируемые параметры называется *информационными*. Устройство, осуществляющее модуляцию, называется *модулятором*. Оно имеет два входа и один выход.

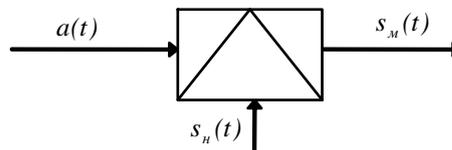


Рисунок 5.1 – Модулятор.

Обозначения:

- $a(t)$ - модулирующий, низкочастотный, управляющий, информационный, первичный сигнал;
- $s_n(t)$ - модулируемый сигнал, высокочастотное, несущее колебание;
- $s_m(t)$ - модулированный, высокочастотный, вторичный сигнал.

Главная особенность модуляции – преобразование спектра модулирующего сигнала: происходит расширение спектра, а при гармонической несущей – перенос спектра в область около частоты несущей. Последнее обстоятельство привело к использованию модулированных сигналов в радиосвязи, многоканальной связи, т.к. при радиопередаче необходимо использовать сигнал, эффективно излучаемый антенной (высокочастотный) и передаваемый без искажений через радиотехнические цепи (узкополосный), а в многоканальных системах с ЧРК необходимо осуществить разделение канальных сигналов по занимаемой полосе частот на приеме.

Теоретически возможно бесконечное число видов модуляции. В настоящее время в системах связи используется более пятидесяти. Вопрос выбора вида модуляции для системы связи решается с точки зрения эффективного прохождения несущей по линии связи, простоты выполнения операций модуляции и демодуляции, способности обеспечить заданное качество передачи сообщений при наличии помех.

Таблица 5.1 – Виды модуляции.

Вид сигнала	модулирующего переносчика		непрерывный непрерывный			дискретный непрерывный			непрерывный дискретный			дискретный дискретный		
	Вид модуляции		непрерывная			манипуляция			импульсная			цифровая		
Вариант модуляции	АМ	УМ		АМн	ЧМн	ФМн	АИМ	ВИМ		ШИМ (ДИМ)	ИКМ	ДИКМ	ДМ	
		ЧМ	ФМ					ЧИМ	ФИМ					

Обозначения:

- АМ – амплитудная модуляция;
- УМ – угловая модуляция;
- ЧМ – частотная модуляция;
- ФМ – фазовая модуляция;
- АМн – амплитудная манипуляция;

- ЧМн – частотная манипуляция;
- ФМн – фазовая манипуляция;
- АИМ – амплитудно-импульсная модуляция;
- ВИМ – время-импульсная модуляция;
- ЧИМ – частотно-импульсная модуляция;
- ФИМ – фазо-импульсная модуляция;
- ШИМ – широтно-импульсная модуляция;
- ДИМ – длительно-импульсная модуляция;
- ИКМ – импульсно-кодовая модуляция;
- ДИКМ – дифференциальная импульсно-кодовая модуляция;
- ДМ – дельта-модуляция.

Амплитудно-модулированные сигналы с большим уровнем несущего колебания

Амплитудная модуляция (АМ) – это процесс управления амплитудой гармонического несущего колебания по закону изменения информационного сигнала.

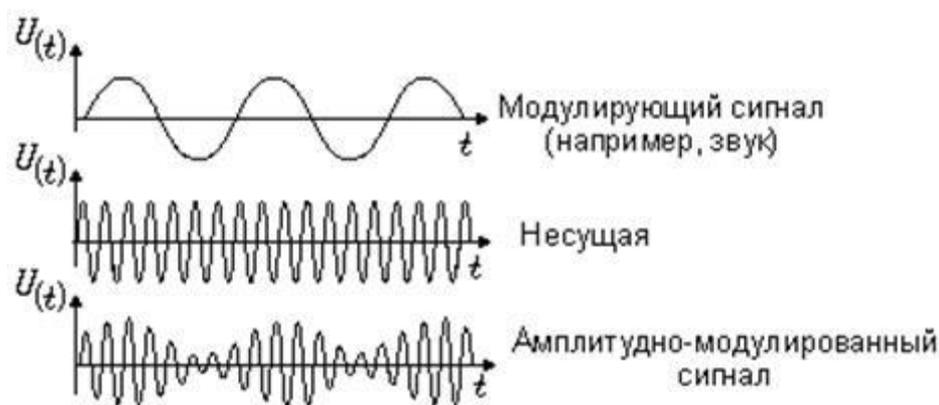


Рисунок 5.2 – Временные диаграммы модулирующего, несущего и амплитудно-модулированного колебаний

Несущее колебание описывается выражением:

$$s_n(t) = A_m \cos(\omega_n t + \varphi),$$

где A_m - амплитуда;

$\varphi(t) = \omega_n t + \varphi$ - полная фаза;

ω_n - угловая частота;

φ - начальная фаза.

При АМ амплитуда несущего колебания изменяется пропорционально мгновенным значениям модулирующего сигнала $a(t)$:

$$A_m(t) = A_m + \Delta A_m(t) = A_m + k_{AM} a(t),$$

где $\Delta A_m(t) = k_{AM} a(t)$ - приращение амплитуды несущей при АМ;

k_{AM} - безразмерный коэффициент пропорциональности. Зависит от конкретной схемы модулятора. Выбирается так, чтобы амплитуда всегда была положительной: $A_m(t) \geq 0$.

Математическая модель АМ сигнала:

$$s_{AM}(t) = A_m(t) \cos(\omega_n t + \varphi) = (A_m + k_{AM} a(t)) \cos(\omega_n t + \varphi).$$

В качестве модулирующего сигнала рассмотрим гармоническое колебание:

$$a(t) = A_{mu} \cos(\Omega t + \phi),$$

где A_{mu} - амплитуда модулирующего сигнала;

Ω - его угловая частота;

ϕ - его начальная фаза.

Так как амплитуда несущей изменяется по закону модулирующего сигнала, то можно записать:

$$A_m(t) = A_m + k_{AM} \cdot A_{mu} \cos(\Omega t + \phi).$$

Запишем выражение, являющееся математической моделью АМ сигнала в случае использования в качестве информационного сигнала гармонического колебания:

$$\begin{aligned} s_{AM}(t) &= A_m(t) \cos(\omega_n t + \varphi) = (A_m + k_{AM} A_{mu} \cos(\Omega t + \phi)) \cos(\omega_n t + \varphi) = \\ &= A_m (1 + m \cos(\Omega t + \phi)) \cos(\omega_n t + \varphi), \end{aligned}$$

где $m = \frac{k_{AM} \cdot A_{mu}}{A_m}$ - коэффициент модуляции, причем $0 \leq m \leq 1$. Коэффициент

модуляции (глубина модуляции) – это отношение максимального приращения амплитуды модулированного сигнала к амплитуде несущей.

Спектр АМ сигнала

АМ сигнал можно представить в виде суммы гармонических составляющих. Используя формулу произведения косинусов:

$$\cos(a + b) = 0,5(\cos(a + b) + \cos(a - b)), -$$

из выражения сигнала с гармонической АМ получаем:

$$s_{AM}(t) = A_m \cdot \cos(\omega_n t + \varphi) + \frac{A_m \cdot m}{2} \cos((\omega_n - \Omega)t + \varphi - \phi) + \frac{A_m \cdot m}{2} \cos((\omega_n + \Omega)t + \varphi + \phi).$$

АМ сигнал при модуляции гармоническим сигналом состоит из трех спектральных составляющих с частотами: несущей ω_n , нижней боковой $\omega_n - \Omega$, верхней боковой $\omega_n + \Omega$. Спектральная диаграмма однотонового АМ сигнала симметрична относительно несущей частоты. Амплитуды боковых колебаний одинаковы ($A_{m\text{обо}} = \frac{A_m m}{2}$) и при $m=1$ не превышают половины амплитуды несущего колебания.

В большинстве случаев модулирующие сигналы являются сложными. Любой сложный сигнал можно представить в виде суммы (конечной или бесконечной) гармонических составляющих, воспользовавшись рядом или интегралом Фурье. Каждая гармоническая составляющая модулирующего сигнала с частотой Ω_i приведет к появлению в АМ сигнале двух боковых составляющих с частотами $\omega_n + \Omega_i$, $\omega_n - \Omega_i$. Т.е. в спектре сложномодулированного АМ сигнала помимо колебания с частотой несущей содержатся группы верхних и нижних боковых колебаний, образующих соответственно верхнюю и нижнюю боковые полосы частот. При этом верхняя боковая полоса частот является масштабной копией спектра модулирующего сигнала, сдвинутой в область высоких частот на величину ω_n . Нижняя боковая полоса частот также повторяет спектральную диаграмму сигнала $a(t)$, но частоты в ней располагаются в зеркальном (обратном) порядке относительно несущей частоты.

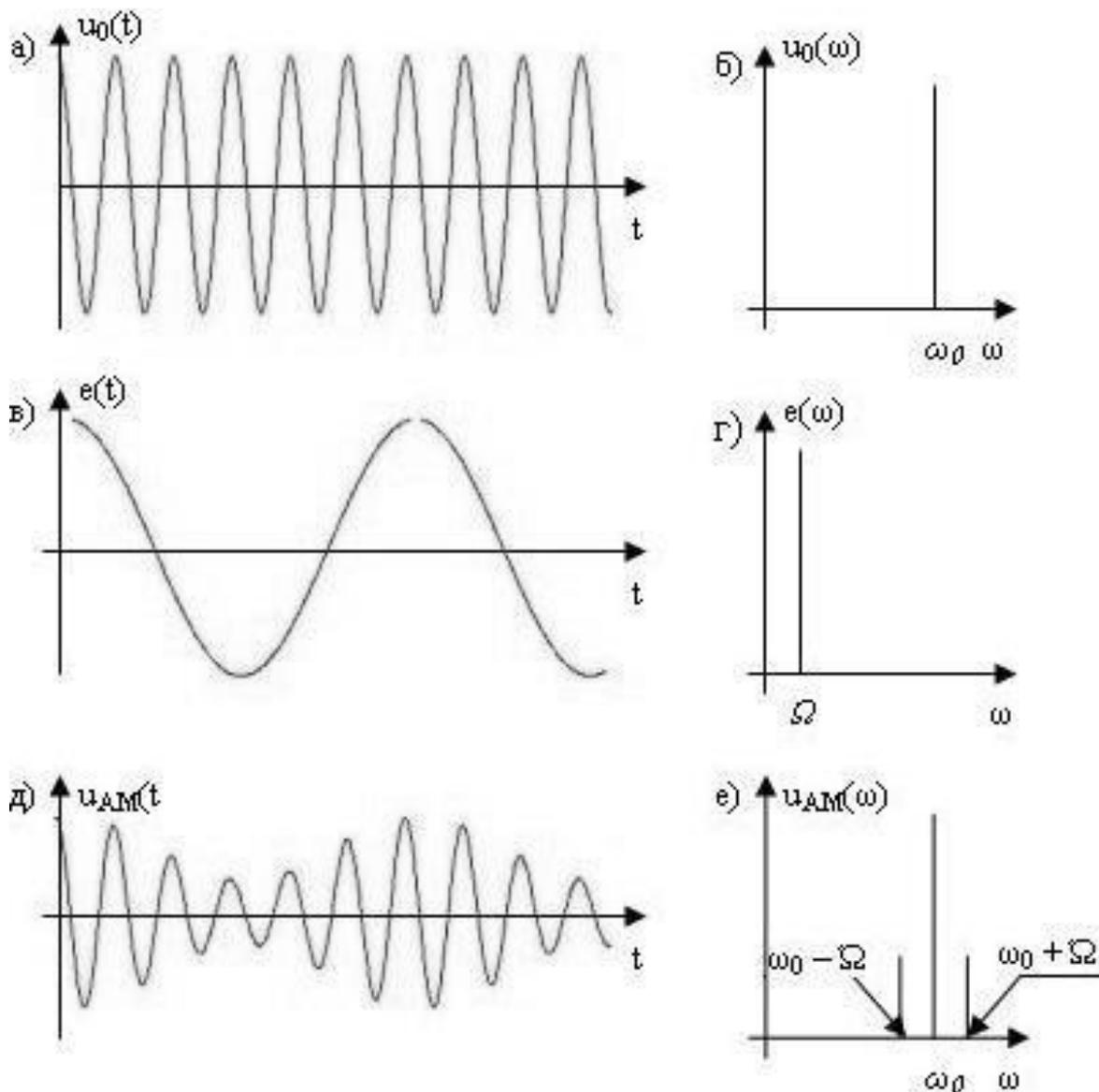


Рисунок 5.3 – Временные и спектральные диаграммы:
 а) и б) – несущей; в) и г) – модулирующего сигнала; д) и е) – АМ сигнала

Ширина спектра АМ сигнала равна удвоенному значению наиболее высокой частоты Ω_{\max} спектра модулирующего сигнала:

$$\Delta\omega_{AM} = 2\Omega_{\max}.$$

На принципах АМ построено большинство радиовещательных систем, а также видеоканалов в телевидении.

Амплитудно-модулированные сигналы с подавленным несущим колебанием.

Для более эффективного использования мощности спектра АМ сигнала возможно исключение из спектра АМ сигнала несущего колебания. Такой АМ сигнал называют балансно-модулированным (БМ). Также из спектра можно исключить одну боковую полосу частот (верхнюю или нижнюю), поскольку каждая из них содержит полную информацию о модулирующем сигнале $a(t)$. При этом получается однопольную модуляцию (ОМ), т.е. модуляцию с одной боковой полосой — ОБП.

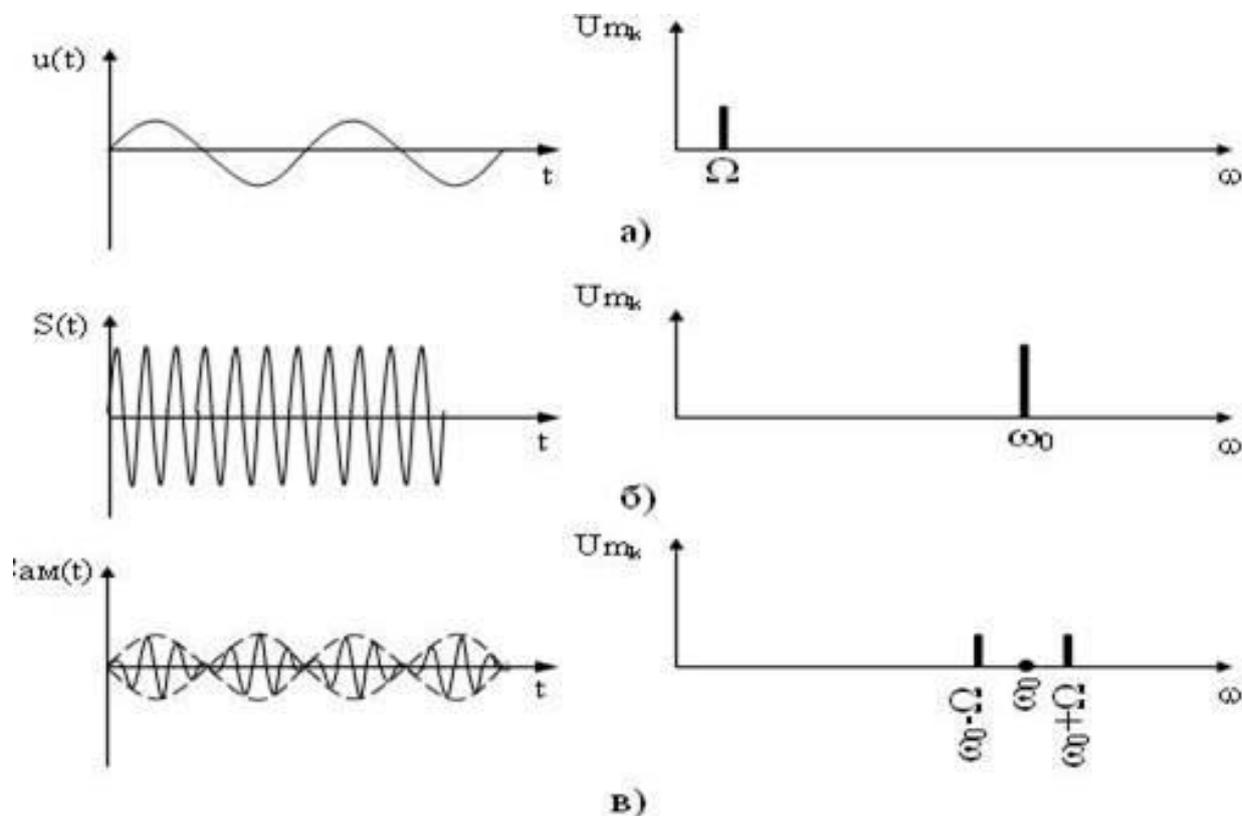


Рисунок 5.4 - Временные и спектральные диаграммы:
 а) – модулирующего сигнала; б) – несущей; в) – БМ сигнала

Сигналы с аналоговой частотной модуляцией гармонического несущего колебания

Угловая модуляция

Воздействие модулирующего сигнала $a(t)$ на аргумент (текущую фазу) $\varphi(t)$ гармонической несущей $s_n(t) = A_m \cos \varphi(t) = A_m \cos(\omega_n t + \varphi)$, называется *угловой модуляцией* (УМ). Разновидностями УМ являются частотная и фазовая.

Частотная модуляция

Частотная модуляция (ЧМ) - процесс управления частотой гармонической несущей по закону модулирующего сигнала.

Угловая частота изменяется по закону:

$$\omega_{ЧМ}(t) = \omega_n + \Delta\omega(t) = \omega_n + k_{ЧМ} a(t),$$

где ω_n - частота несущей;

$\Delta\omega(t)$ - отклонение частоты модулированного сигнала от значения ω_n ;

$a(t)$ - модулирующий сигнал. Может быть гармоническим (используется для учебных или исследовательских целей) и негармоническим (реальный сигнал);

$k_{ЧМ}$ - размерный коэффициент пропорциональности, рад/(с·В) или рад/(с·А).

Определяется схмотехникой модулятора.

Полная фаза в момент времени t находится путем интегрирования частоты:

$$\varphi_{ЧМ}(t) = \int_0^t \omega_{ЧМ}(t) dt + \varphi = \int_0^t (\omega_n + k_{ЧМ} a(t)) dt + \varphi = \int_0^t \omega_n dt + \int_0^t k_{ЧМ} a(t) dt + \varphi = \omega_n t + k_{ЧМ} \int_0^t a(t) dt + \varphi,$$

где $\int_0^t \omega_{\text{ЧМ}}(t)dt$ - набег фазы за время от начала отсчета до рассматриваемого момента t ;

мента t ;

φ - постоянная интегрирования.

Математическая модель ЧМ сигнала:

$$s_{\text{ЧМ}}(t) = A_m \cos \varphi_{\text{ЧМ}}(t) = A_m \cos(\omega_n t + k_{\text{ЧМ}} \int_0^t a(t)dt + \varphi).$$

ЧМ называют интегральным видом модуляции, т.к. $a(t)$ входит в это выражение под знаком интеграла.

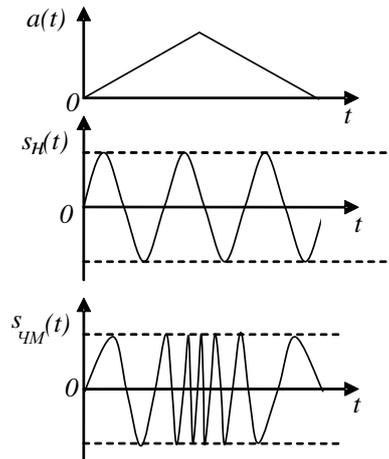


Рисунок 5.5 – Временные диаграммы модулирующего, несущего и модулированного колебаний.

Гармоническая ЧМ

Рассмотрим гармоническую ЧМ (модулирующий сигнал является гармоническим $a(t) = A_{\text{му}} \cos(\Omega t + \Phi)$).

Частота изменяется по закону:

$$\omega_{\text{ЧМ}}(t) = \omega_n + k_{\text{ЧМ}} a(t) = \omega_n + k_{\text{ЧМ}} A_{\text{му}} \cos(\Omega t + \Phi) = \omega_n + \Delta\omega_{\text{мЧМ}} \cos(\Omega t + \Phi),$$

где $\Delta\omega_{\text{мЧМ}} = k_{\text{ЧМ}} A_{\text{му}}$ - девиация частоты при ЧМ. Девиация частоты – наибольшее отклонение частоты модулированного сигнала от значения частоты несущей. При ЧМ может принимать значения от единиц герц до сотен мегагерц.

Фаза в момент времени t :

$$\begin{aligned} \varphi_{\text{ЧМ}}(t) &= \omega_n t + \int_0^t k_{\text{ЧМ}} a(t)dt + \varphi = \omega_n t + \int_0^t \Delta\omega_m \cos(\Omega t + \Phi) + \varphi = \omega_n t + \frac{\Delta\omega_m}{\Omega} \sin(\Omega t + \Phi) + \varphi = \\ &= \omega_n t + M_{\text{ЧМ}} \sin(\Omega t + \Phi) + \varphi, \end{aligned}$$

где $M_{\text{ЧМ}} = \Delta\varphi_{\text{мЧМ}} = \frac{\Delta\omega_{\text{мЧМ}}}{\Omega} = \frac{\Delta f_{\text{мЧМ}}}{F}$ - индекс частотной модуляции. Является девиацией фазы при ЧМ. Девиация фазы - наибольшее отклонение фазы модулированного сигнала от линейной $\varphi(t) = \omega_n t + \varphi$.

Математическая модель сигнала при гармонической ЧМ:

$$s_{\text{ЧМ}}(t) = A_m \cos \varphi_{\text{ЧМ}}(t) = A_m \cos(\omega_n t + M_{\text{ЧМ}} \sin(\Omega t + \Phi) + \varphi).$$

Воспользовавшись тригонометрической формулой: $\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$, - преобразуем выражение:

$$s_{\text{ЧМ}}(t) = A_m \cos((\omega_n t + \varphi) + M_{\text{ЧМ}} \sin(\Omega t + \Phi)) = A_m \cos(\omega_n t + \varphi) \cos(M_{\text{ЧМ}} \sin(\Omega t + \Phi)) - A_m \sin(\omega_n t + \varphi) \sin(M_{\text{ЧМ}} \sin(\Omega t + \Phi)).$$

Проведем анализ отдельно для малых и больших индексов модуляции.

В первом случае ($M_{\text{ЧМ}} < 1$) имеют место приближенные равенства:

$$\cos(M_{\text{ЧМ}} \sin(\Omega t + \Phi)) \approx 1, \quad \sin(M_{\text{ЧМ}} \sin(\Omega t + \Phi)) \approx M_{\text{ЧМ}} \sin(\Omega t + \Phi).$$

Воспользовавшись тригонометрической формулой:

$$\sin \alpha \sin \beta = 1/2 \cos(\alpha - \beta) - 1/2 \cos(\alpha + \beta),$$

приходим к следующему выражению для ЧМ сигнала:

$$\begin{aligned} s_{\text{ЧМ}}(t) &= A_m \cos(\omega_n t + \varphi) - M_{\text{ЧМ}} A_m \sin(\omega_n t + \varphi) \sin(\Omega t + \Phi) = \\ &= A_m \cos(\omega_n t + \varphi) + 1/2 M_{\text{ЧМ}} A_m \cos((\omega_n + \Omega)t + \varphi + \Phi) - 1/2 M_{\text{ЧМ}} A_m \cos((\omega_n - \Omega)t + \varphi - \Phi). \end{aligned}$$

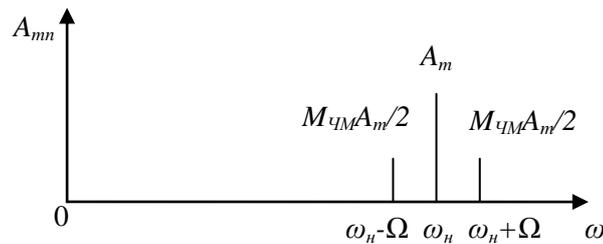


Рисунок 5.6 – Спектральная диаграмма ЧМ сигнала при $M_{\text{ЧМ}} < 1$.

При малом индексе модуляции – узкополосной ЧМ – амплитудная спектральная диаграмма ЧМ сигнала совпадает по составу (содержит центральную составляющую с частотой несущей ω_n , нижнюю и верхнюю боковые составляющие с частотами $\omega_n - \Omega$ и $\omega_n + \Omega$) и ширине полосы частот ($\Delta\omega_{\text{ЧМ}} = 2\Omega$) с АМ сигналом. Отличие заключается в фазовой спектральной диаграмме: фаза нижней боковой составляющей сдвинута на 180° .

При малом значении индекса модуляции не будут проявляться преимущества ЧМ (высокая помехозащищенность). Ширина спектра такая же, как и при АМ.

Во втором случае ($M_{\text{ЧМ}} \geq 1$) сложные периодические функции: $\cos(M_{\text{ЧМ}} \sin(\Omega t + \Phi))$ и $\sin(M_{\text{ЧМ}} \sin(\Omega t + \Phi))$ - можно разложить в ряд Фурье, а ЧМ сигнал представить в виде суммы гармонических колебаний:

$$\begin{aligned} s_{\text{ЧМ}}(t) &= A_m (J_0(M_{\text{ЧМ}}) \cos(\omega_n t + \varphi) + J_1(M_{\text{ЧМ}}) \cos((\omega_n + \Omega)t + \varphi + \Phi) - \\ &J_1(M_{\text{ЧМ}}) \cos((\omega_n - \Omega)t + \varphi - \Phi) + J_2(M_{\text{ЧМ}}) \cos((\omega_n + 2\Omega)t + \varphi + 2\Phi) + \\ &+ J_2(M_{\text{ЧМ}}) \cos((\omega_n - 2\Omega)t + \varphi - 2\Phi) + \dots) = \\ &= A_m J_0(M_{\text{ЧМ}}) \cos(\omega_n t + \varphi) + \sum_{n=1}^{\infty} A_m J_n(M_{\text{ЧМ}}) \cos((\omega_n + n\Omega)t + \varphi + n\Phi) + \\ &+ \sum_{n=1}^{\infty} (-1)^n A_m J_n(M_{\text{ЧМ}}) \cos((\omega_n - n\Omega)t + \varphi - n\Phi), \end{aligned}$$

где $J_n(M_{\text{ЧМ}})$ - функция Бесселя 1-го рода n-го порядка от вещественного аргумента $M_{\text{ЧМ}}$. Табулированы;

n – номер гармонической составляющей: центральная составляющая имеет $n=0$, боковые – $n=1, 2, 3, \dots$.

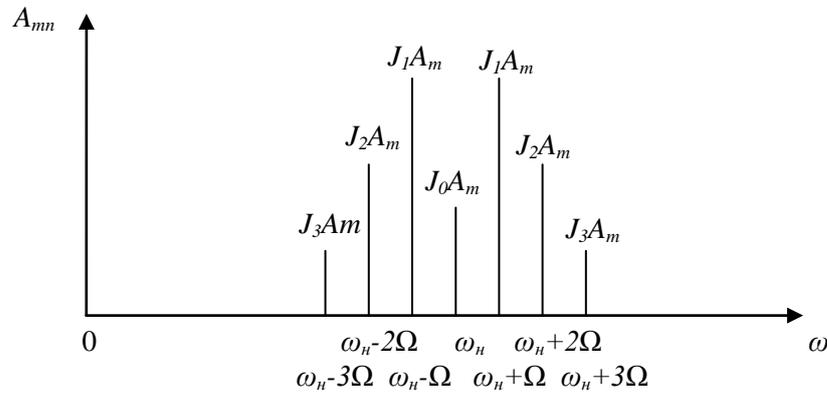


Рисунок 5.7 – Спектр ЧМ сигнала при $M_{\text{ЧМ}}=2$.

При большом индексе модуляции – *широкополосной ЧМ* – спектр ЧМ сигнала состоит из бесконечного числа гармоник: из составляющей с частотой несущей ω_n , нижней и верхней боковых полос частот, образованных группами составляющих с частотами $\omega_n - n\Omega$ и $\omega_n + n\Omega$. На практике учитывают только те боковые составляющие, амплитуды которых не меньше 5% амплитуды несущей, т.е. для которых $J_n(M_{\text{ЧМ}}) \geq 0,05$. Тогда ширина спектра ЧМ сигнала: $\Delta\omega_{\text{ЧМ}} = 2(M_{\text{ЧМ}} + 1)\Omega$.

Данный случай представляет основной практический интерес, поскольку при больших индексах модуляции помехоустойчивость передачи сигнала существенно выше, чем при АМ. Ширина спектра ЧМ сигнала также значительно больше, чем при АМ.

При сложном модулирующем сигнале спектр модулированного сигнала оказывается сложным, содержащим различные комбинационные частоты. Общая полоса частот, занимаемая таким сигналом: $\Delta\omega_{\text{ЧМ}} = 2(M_{\text{ЧМ}} + 1)\Omega_{\text{max}}$, где Ω_{max} - максимальная частота спектра модулирующего сигнала; $M_{\text{ЧМ}}$ - индекс модуляции на этой частоте.

Сигналы с аналоговой фазовой модуляцией гармонического несущего колебания

Фазовая модуляция (ФМ) – изменение фазы гармонической несущей по закону модулирующего сигнала.

Мгновенная фаза ФМ сигнала определяется выражением:

$$\varphi_{\text{ФМ}}(t) = \varphi(t) + \Delta\varphi(t) = \omega_n t + \varphi + k_{\text{ФМ}} a(t),$$

где $\Delta\varphi(t) = k_{\text{ФМ}} a(t)$ - отклонение (сдвиг) фазы модулированного сигнала от линейно-изменяющейся фазы гармонической несущей $\varphi(t) = \omega_n t + \varphi$;

$k_{\text{ФМ}}$ - размерный коэффициент пропорциональности, рад/В или рад/А.

Математическая модель ФМ сигнала:

$$s_{\text{ФМ}}(t) = A_m \cos \varphi_{\text{ФМ}}(t) = A_m \cos(\omega_n t + k_{\text{ФМ}} a(t) + \varphi).$$

Угловая частота – это скорость изменения (т.е. производная по времени) полной фазы колебания. Выражение для мгновенной частоты:

$$\omega_{\text{ФМ}}(t) = d\varphi_{\text{ФМ}}(t) / dt = d(\omega_n t + k_{\text{ФМ}} a(t) + \varphi) / dt = \omega_n + k_{\text{ФМ}} da(t) / dt.$$

Таким образом, ФМ сигнал с модулирующим сигналом $a(t)$ можно рассматривать как ЧМ сигнал с модулирующим сигналом $da(t) / dt$.

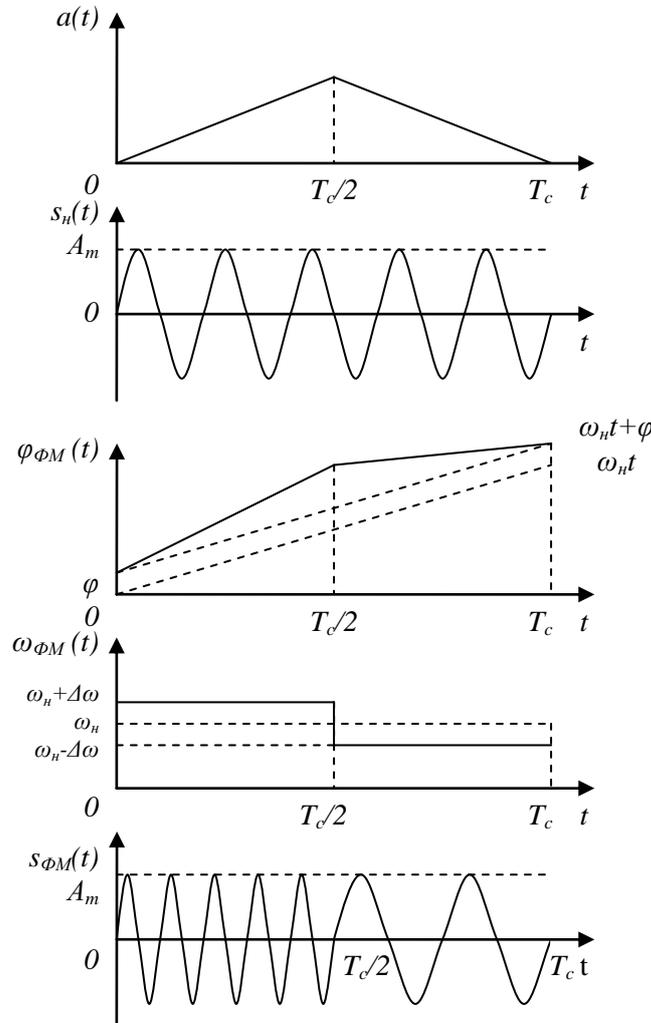


Рисунок 5.8 – Модулирующий сигнал, несущее колебание, изменение фазы ФМ сигнала, изменение частоты ФМ сигнала и ФМ сигнал.

Гармоническая ФМ

Рассмотрим случай гармонического модулирующего сигнала:

$$a(t) = A_{mu} \cos(\Omega t + \Phi).$$

Фаза сигнала с гармонической ФМ:

$$\varphi_{FM}(t) = \omega_n t + k_{FM} a(t) + \varphi = \omega_n t + k_{FM} A_{mu} \cos(\Omega t + \Phi) + \varphi = \omega_n t + M_{FM} \cos(\Omega t + \Phi) + \varphi,$$

где $M_{FM} = \Delta\varphi_{mFM} = k_{FM} A_{mu}$ - индекс фазовой модуляции или девиация фазы при ФМ. Может принимать значение от единиц до десятков тысяч радиан.

Математическая модель сигнала с гармонической ФМ:

$$s_{FM}(t) = A_m \cos \varphi_{FM}(t) = A_m \cos(\omega_n t + M_{FM} \cos(\Omega t + \Phi) + \varphi).$$

Частота ФМ сигнала:

$$\omega_{FM}(t) = \omega_n + k_{FM} da(t)/dt = \omega_n - k_{FM} A_{mu} \Omega \sin(\Omega t + \Phi) = \omega_n - \Delta\omega_{mFM} \sin(\Omega t + \Phi),$$

где $\Delta\omega_{mFM} = k_{FM} A_{mu} \Omega$ - девиация частоты при ФМ.

Методология вычисления и структура спектра ФМ сигнала аналогичны ЧМ сигналу, но индекс частотной модуляции необходимо заменить индексом фазовой модуляции. Аналогичная тесная связь между спектрами ФМ и ЧМ сигналов имеет место и при негармонических модулирующих сигналах.

ФМ применяется в схемах косвенного метода получения ЧМ.

Дискретно-модулированные (манипулированные) сигналы

Дискретная модуляция (манипуляция) - модуляция гармонического несущего колебания дискретным (цифровым) модулирующим сигналом. При этом модулируемые (информационные) параметры переносчика изменяются скачкообразно. Устройство, реализующее процесс манипуляции, называют манипулятором.

Дискретным модулирующим сигналом является первичный сигнал, отображающий символы кодовых комбинаций дискретных сообщений. *Примеры дискретных первичных сигналов:* телеграфный, передачи данных, ИКМ.

Различают следующие виды манипуляции:

- в зависимости от изменяемых параметров переносчика:
- амплитудную (АМн; английский термин – amplitude shift keying, ASK),
- частотную (ЧМн; английский термин – frequency shift keying, FSK),
- фазовую (ФМн; английский термин – phase shift keying, PSK),
- амплитудно-фазовую (АФМн; английский термин – АРК/PSK, или amplitude phase keying, АРК).

При АМн каждому возможному значению передаваемого символа ставится в соответствие своя амплитуда гармонического несущего колебания, при ЧМн – частота, при ФМн – фаза, а при АФМн – комбинация амплитуды и начальной фазы;

- в зависимости от используемых кодов:

- многопозиционную или m -арную (по-английски – m -ary),
- двоичную (по-английски – binary).

Многопозиционная манипуляция используется для повышения скорости передачи информации при одной и той же скорости модуляции. m - основание многопозиционного кода – число различных его символов. На практике обычно является ненулевой степенью двойки: $m = 2^l$, где l - число двоичных цифр (битов), представляющих символы многопозиционного кода. Двоичная манипуляция ($l = 1$, $m = 2^1 = 2$) является частным случаем многопозиционной. Как правило, в системах передачи дискретных сообщений используются двоичные коды.

Двоичная АМн

При двоичном коде первичный сигнал принимает два значения, соответствующие кодовым символам 0 и 1:

- $(-U_{ми}, U_{ми})$ – двухполярный сигнал;
- $(0, U_{ми})$ – однополярный сигнал.

При двоичной АМн (BASK) символу 1 соответствует отрезок гармонического несущего колебания (посылка), символу 0 – отсутствие колебания (пауза), поэтому часто АМн называют манипуляцией с пассивной паузой.

Примем в качестве модулирующего меандровый сигнал – сигнал, отображающий последовательность битов повторяющегося двоичного кода 1010.

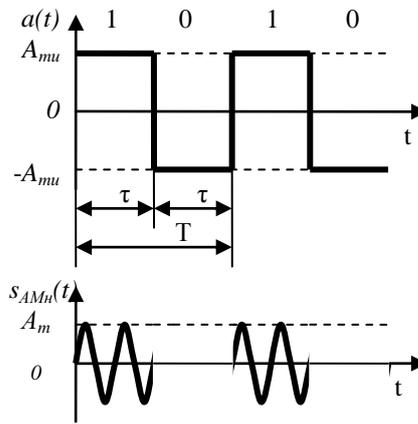


Рисунок 5.9 – Временные диаграммы модулирующего и АМн сигналов.

АМн можно рассматривать как модуляцию сигналом со спектром, богатым гармониками: спектр меандрового сигнала содержит бесконечное количество нечетных гармоник. Спектр АМн сигнала содержит составляющую с частотой несущей и две боковые полосы, каждая из которых повторяет спектр первичного сигнала.

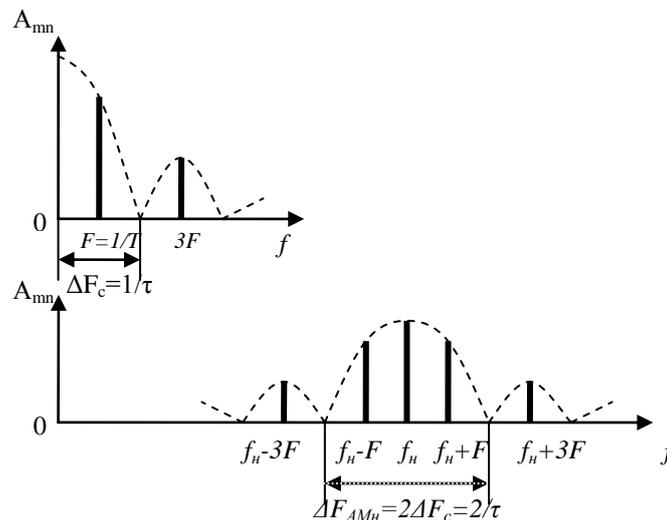


Рисунок 5.10 – Спектральные диаграммы модулирующего и АМн сигналов.

Теоретически спектр сигнала при АМн бесконечен. На практике бесконечный спектр ограничивается полосой пропускания фильтра. Соотношение для расчета ширины спектра АМн сигнала:

$$\Delta F_{AMn} = 2\nu,$$

где $\nu = 1/\tau$ - символьная скорость или скорость модуляции, Бод – число символов кода, передаваемых за единицу времени (1 с);

τ - символьный (тактовый) интервал – интервал времени, отведенный для передачи одного символа.

АМн была изобретена в начале 20 столетия для беспроводной телеграфии. В настоящее время АМн в системах цифровой связи уже не используется.

Двоичная ЧМн

При двоичной ЧМн (BFSK) символу 1 соответствует отрезок гармонического колебания с частотой $f_1 = f_n + \Delta f_m$, а символу 0 – с частотой $f_0 = f_n - \Delta f_m$, где Δf_m - деви-

ация частоты – изменение частоты при передаче 1 (0) относительно ее среднего значения f_n . При ЧМн нет пассивной паузы, по этой причине ее называют манипуляцией с активной паузой.

Возможно два случая ЧМн: с разрывом фазы и без разрыва фазы (continuous-phase FSK – CPFSK).

При ЧМн с разрывом фазы назначение каждому двоичному символу своей частоты является произвольным. Полученный сигнал содержит скачки фазы.

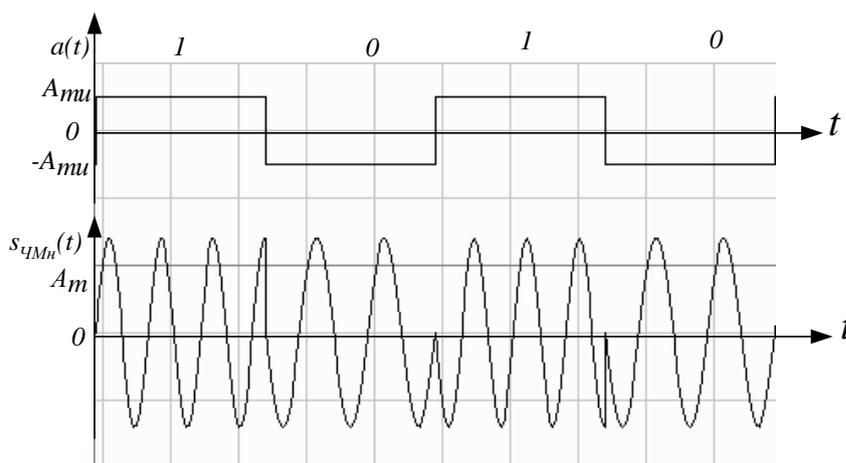


Рисунок 5.11 – Временные сигналов: модулирующего и ЧМн с разрывом фазы.

Наличие разрывов фазы приводит к «размытию» спектра сигнала. Это снижает помехоустойчивость приема и создает помехи другим системам связи. Поэтому при выборе частот следует обеспечить условие плавного (без скачка фазы) перехода от сигнала с частотой f_1 к сигналу с частотой f_0 :

$$f_1 = Kv \text{ и } f_0 = Nv ,$$

где K, N - число периодов внутри символьного интервала. t

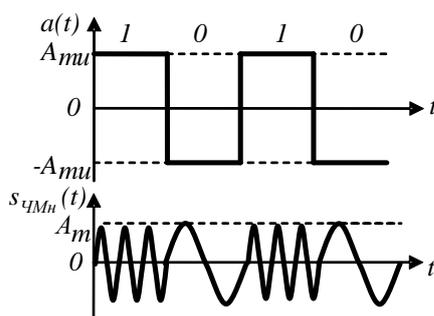


Рисунок 5.12 – Временные диаграммы сигналов: модулирующего и ЧМн без разрыва фазы.

ЧМн сигнал можно рассматривать как сумму двух отдельных АМн сигналов, один из которых служит только для отображения единиц и имеет несущую частоту f_1 , а другой – только нулей и имеет несущую частоту f_0 . Соответственно, спектр ЧМн сигнала может быть представлен как суперпозиция спектров двух АМн сигналов.

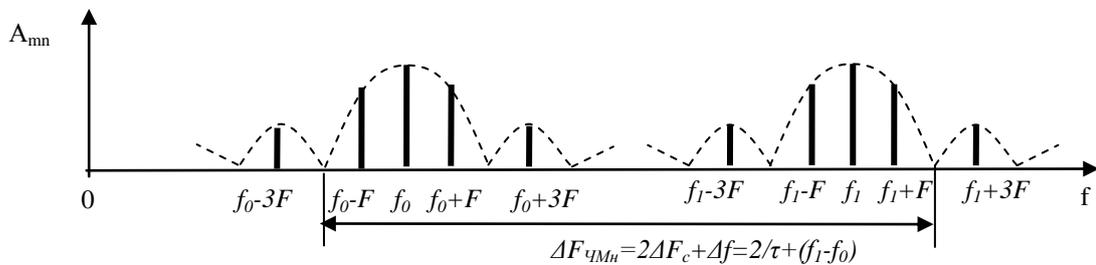


Рисунок 5.13 – Спектральная диаграмма ЧМн сигнала.

Ширина спектра сигнала с двоичной ЧМн определяется выражением:

$$\Delta f_{\text{ЧМн}} = 2\nu + \Delta f,$$

где $\Delta f = 2\Delta f_m = f_1 - f_0$ - разнос частот – расстояние между частотами f_1 и f_2 . Выбирается таким, чтобы не перекрывались спектры отдельных АМн сигналов;

Δf_m - девиация частоты – изменение частоты при передаче 1 (0) относительно ее среднего значения f_n .

ЧМн применяется в радиорелейных и спутниковых системах связи, системах связи с подвижными объектами.

Двоичная ФМн

Различают абсолютную ФМн и относительную ФМн (ОФМн; differential PSK – DPSK).

При ФМн информация вкладывается в абсолютное значение фазы сигнала, т.е. смещение фазы модулированного сигнала относительно фазы несущего колебания. При ОФМн фазу сигнала отсчитывают от фазы предыдущего элемента сигнала.

При двоичной ФМн (BPSK) передаче 1 соответствует отрезок гармонического колебания, совпадающего по фазе с несущей, а передаче 0 - отличающегося по фазе на 180° , т.е. фаза меняется на 180° при каждом переходе от 1 к 0 и наоборот.

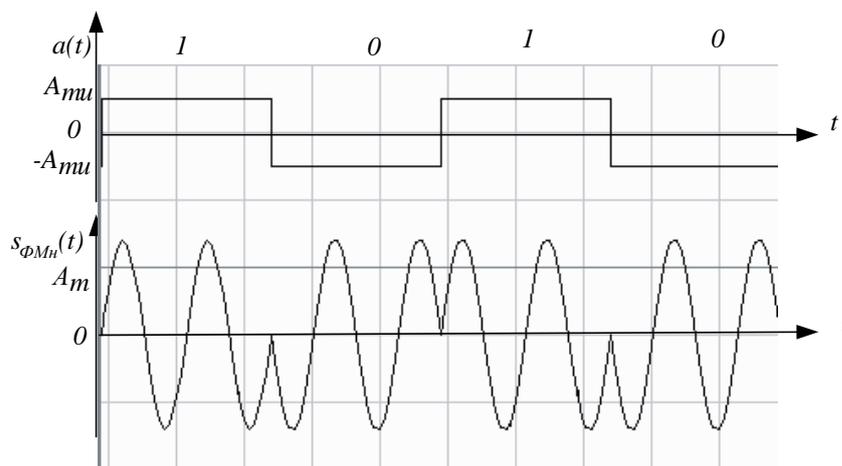


Рисунок 5.14 – Временная диаграмма модулирующего и ФМн сигналов.

ФМн сигнал можно представить в виде суммы двух АМн сигналов, для получения первого из которых используется несущая $A_m \sin \omega_n t$, а второго - $A_m \sin(\omega_n t + 180^\circ)$. Спектр амплитуд ФМн сигнала содержит те же составляющие, что и спектр АМн сигнала, кроме составляющей с частотой несущей (она исчезает, когда символы 1 и 0 по-

являются с равной вероятностью). Амплитуды боковых составляющих примерно в два раза больше. При передаче реальных кодовых слов амплитуда составляющей с частотой несущей не равна нулю, но будет значительно ослаблена.

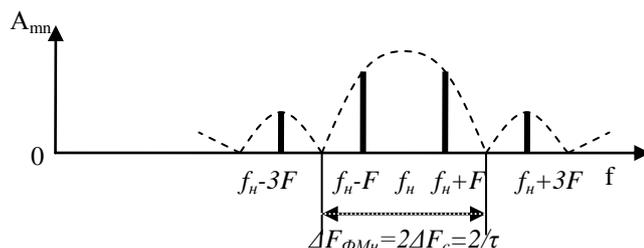


Рисунок 5.15 – Спектр ФМн сигнала.

При ОФМн символ 0 передается отрезком гармонического колебания с начальной фазой предшествующего элемента сигнала, а символ 1 – таким же отрезком с начальной фазой, отличающейся от начальной фазы предшествующего элемента на 180° (фаза изменяется при передаче символов 1), или наоборот (фаза изменяется при передаче символов 0). При ОФМн передача начинается с посылки одного не несущего информации элемента, который служит опорным сигналом для сравнения фазы последующего элемента.

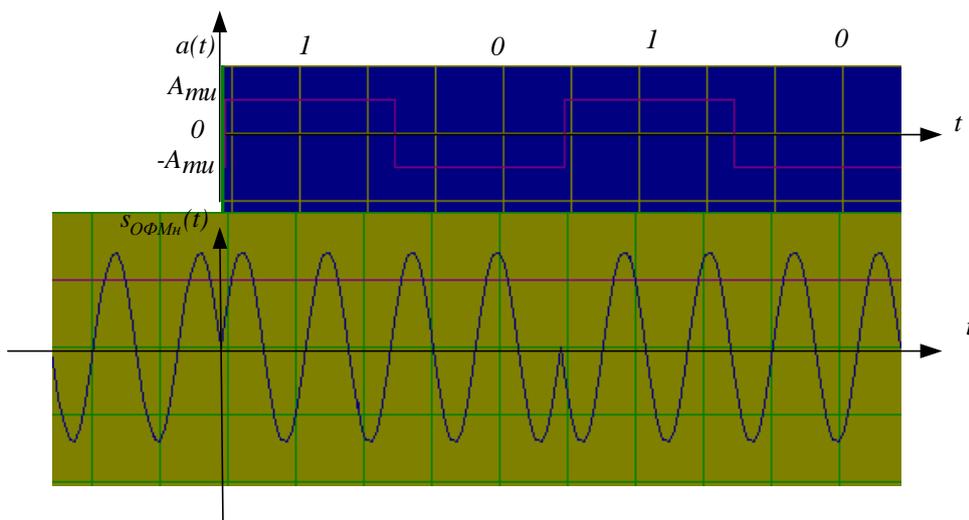


Рисунок 5.16 – Временная диаграмма модулирующего и ОФМн сигнала.

Спектр ОФМн сигнала подобен спектру ФМн сигнала.

ФМн сигнал имеет такую же полосу частот, как АМн сигнал:

$$\Delta f_{\text{ФМн}} = \Delta f_{\text{ОФМн}} = 2\nu.$$

ФМн была разработана в начале развития программы исследования дальнего космоса и сейчас широко используется в коммерческих и военных системах связи.

Импульсно-модулированные сигналы

Импульсная модуляция – модуляция несущего колебания в виде периодической последовательности импульсов одинаковой формы (обычно прямоугольных) аналогового модулирующим сигналом.

Импульсная несущая характеризуется четырьмя параметрами: амплитудой A_m , длительностью τ , частотой следования $f = \frac{1}{T}$ (период следования импульсов выбирается по теореме Котельникова) и фазой импульсов (рисунке 5.14, а). Изменяя их по закону передаваемого сигнала, получают четыре основных вида импульсной модуляции:

- амплитудно-импульсную (АИМ) - изменяется амплитуда импульсов (рисунке 5.14, в). АИМ сигналы подразделяются на два вида: сигнал первого рода (АИМ-1) – вершина импульсов повторяет закон изменения информационного сигнала - и сигнал второго рода (АИМ-2) – вершина импульсов остается плоской;

- частотно-импульсную (ЧИМ) - изменяется частота следования импульсов (рисунке 5.14, г);

- широтно-импульсную или длительно-импульсную (ШИМ или ДИМ) - изменяется ширина (длительность) импульсов (рисунке 5.14, д). Различают одностороннюю (длительность импульсов изменяется по фронту или по срезу) и двухстороннюю ДИМ (длительность импульсов изменяется одновременно по фронту и по срезу);

- фазово-импульсную (ФИМ) - импульсы сдвигаются относительно тактовых точек, которые могут совпадать с началом, серединой или концом импульсов несущей (на рисунке 5.14, е отмечены кружками).

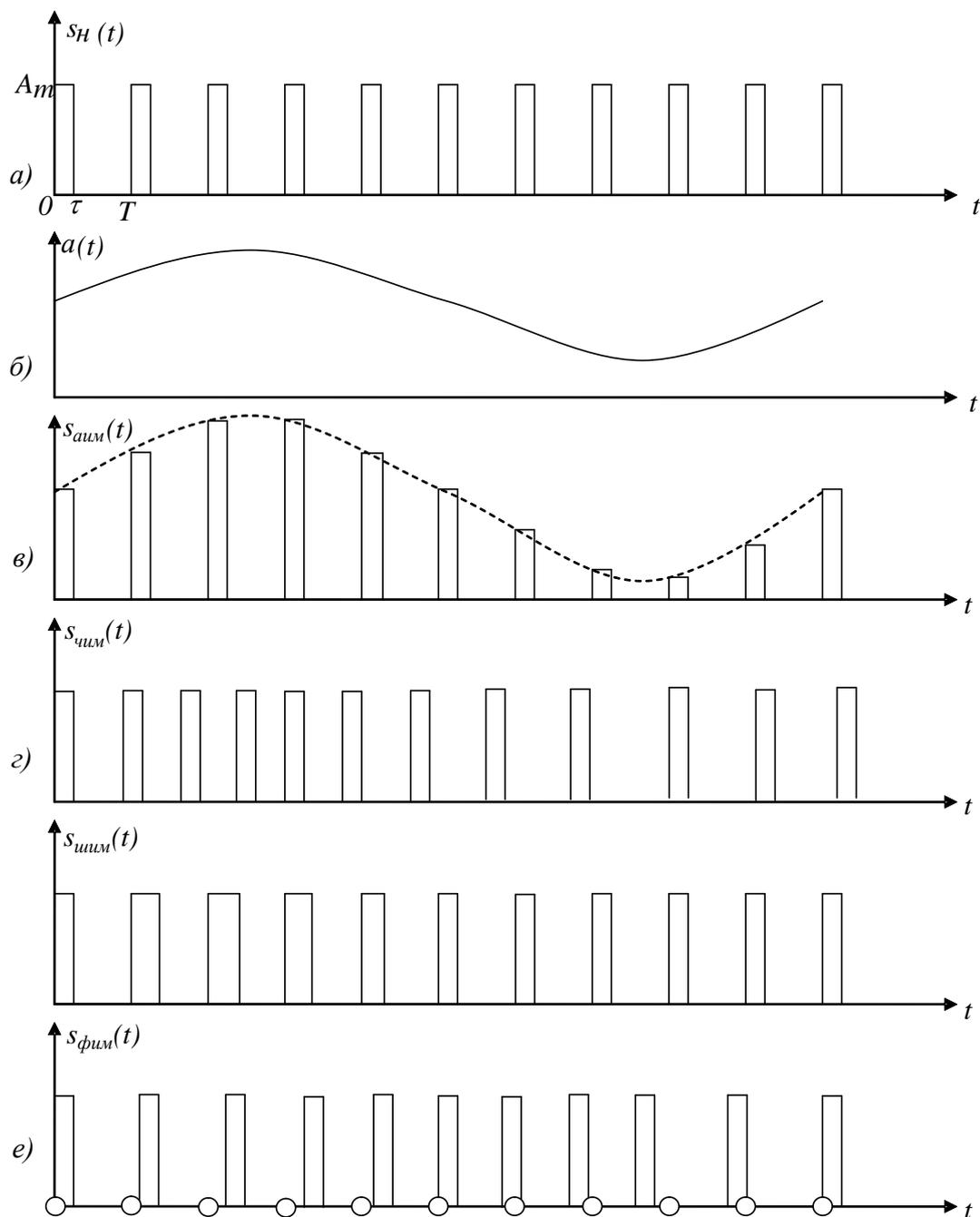


Рисунок 5.17 – Виды импульсной модуляции:

- а) несущая последовательность прямоугольных импульсов;
- б) модулирующий сигнал;
- в) амплитудно-импульсно-модулированный сигнал;
- г) частотно-импульсно-модулированный сигнал;
- д) широтно-импульсно-модулированный сигнал;
- е) фазо-импульсно-модулированный сигнал.

Спектр импульсно-модулированных сигналов

Спектр при импульсных видах модуляции зависит от спектра модулирующего сигнала, вида и параметров модуляции. При модуляции каждую из гармонических составляющих импульсной несущей можно рассматривать как индивидуальную «несущую», возле которой располагаются верхняя и нижняя боковые полосы частот (рисунок 5.15). Формирование полос происходит по закону модуляции гармонической не-

сущей. В спектре импульсно-модулированного сигнала обязательно содержится низкочастотный спектр модулирующего сигнала.

При скважности $q > 10$ боковые полосы частот не дают заметного расширения спектра в сравнении со спектром несущей. Следовательно, для импульсных видов модуляции (кроме ШИМ) ширина спектра не зависит от вида модуляции и ее параметров, модулирующего сигнала, периода следования импульсов:

$$\Delta F_{АИМ} = \Delta F_{ЧИМ} = \Delta F_{ФИМ} = \frac{1}{\tau}, \quad \Delta F_{ШИМ} = \frac{1}{\tau_{\min}},$$

где τ - длительность импульса несущей;

τ_{\min} - минимальная длительность импульса модулированного сигнала.

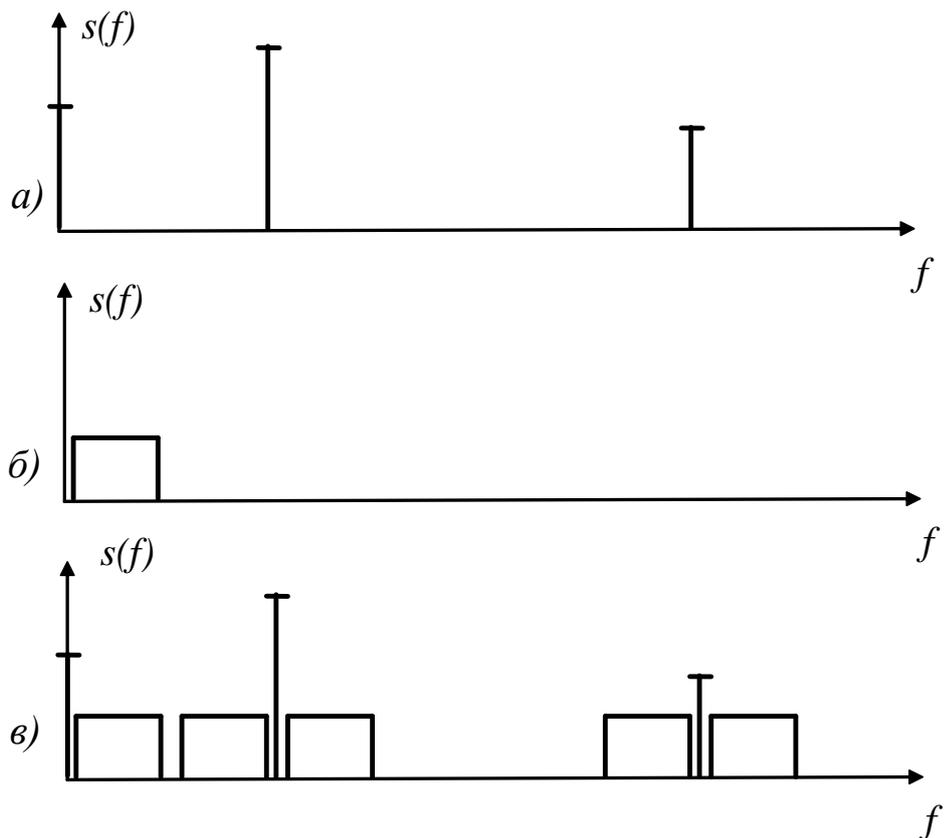


Рисунок 5.18 – Спектральные диаграммы:

а) несущего колебания (ПППИ со скважностью 2);

б) информационного сигнала;

в) АИМ сигнала.

Повторная модуляция

Передача импульсно-модулированных сигналов по высокочастотным линиям связи невозможна, т.к. их спектр содержит низкочастотные составляющие. Для переноса спектра в область высоких частот производится *повторная* (двойная) *модуляция*: модулированными импульсами модулируется гармоническая высокочастотная несущая.

При этом можно получить более 10 различных видов двойных модуляций: АИМ-АМ (рисунок 5.19), ФИМ-АМ, ШИМ-АМ и др. Двойные модуляции различаются сложностью технической реализации, шириной спектра, помехоустойчивостью, и их применение зависит от конкретных технических условий. С точки зрения помехоустойчивости предпочтительней является ФИМ-АМ при скважности импульсов $S \geq 10$.

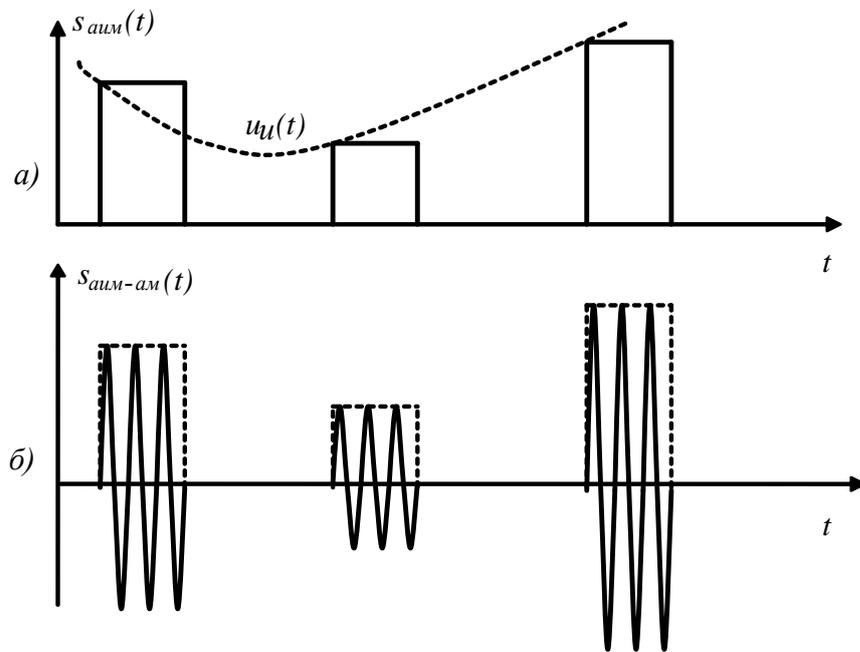


Рисунок 5.19 – Повторная модуляция:

- а) Амплитудно-импульсно-модулированный сигнал;
- б) АИМ-АМ сигнал.

Представление непрерывных сигналов рядом Котельникова

Теорема Котельникова (теорема отсчетов, теорема дискретизации):

всякий непрерывный сигнал $a(t)$ со спектром, ограниченным частотой F_{max} , может быть представлен последовательностью своих мгновенных значений (отсчетов), взятых через интервалы времени $\Delta t \leq 1/(2F_{max})$.

В соответствии с теоремой непрерывный сигнал с ограниченным спектром можно разложить в ряд Котельникова:

$$a(t) = \sum_{n=-\infty}^{\infty} a(n\Delta t) \frac{\sin \omega_0(t - n\Delta t)}{\omega_0(t - n\Delta t)},$$

где $a(n\Delta t)$ - отсчет сигнала в дискретный момент времени $t_n = n\Delta t$;

$\omega_0 = 2\pi f_0 = 2\pi / \Delta t$ - частота дискретизации;

Δt - интервал дискретизации;

$\psi_n(t) = \frac{\sin \omega_0(t - n\Delta t)}{\omega_0(t - n\Delta t)}$ - функция отсчета.

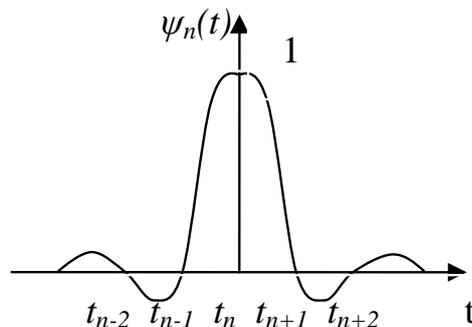


Рисунок 5.20 – Функция отсчета.

Любой реальный сигнал имеет конечную длительность. Его приближенно можно представить усеченным рядом Котельникова:

$$a(t) = \sum_{n=1}^B a(n\Delta t) \frac{\sin \omega_0(t - n\Delta t)}{\omega_0 t - n\Delta t},$$

где $B = T_c / \Delta t + 1 = 2F_{max} T_c + 1 \approx 2\Delta F_c T_c$ – общее число отсчетов для сигнала длительностью T_c или база сигнала.

Содержание теоремы Котельникова

1. Теорема не оговаривает вид сигнала $a(t)$, т.е. он может быть и случайным.

2. Теорема утверждает, что вся информация о сигнале $a(t)$ содержится в его выборочных значениях $a(n\Delta t)$. Следовательно, непрерывный сигнал для передачи по каналу связи может быть преобразован в дискретный по времени сигнал $a_0(t)$. Представление непрерывного сигнала в виде последовательности его отсчетов называется дискретизацией. На практике каждый отсчет представляется импульсом величиной $a(n\Delta t)$ и длительностью $\tau \ll \Delta t$.

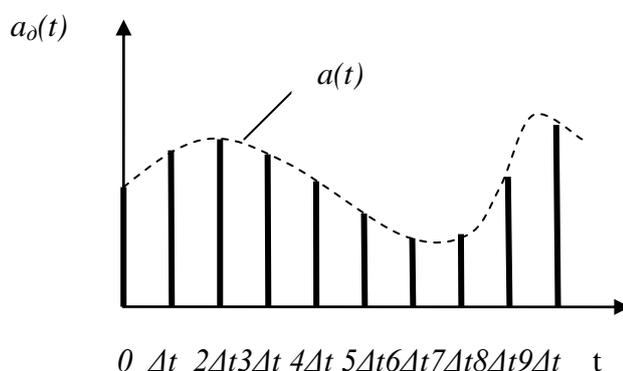


Рисунок 5.21 – Дискретизация непрерывного сигнала.

3. Теорема определяет восстановление непрерывного сигнала $a(t)$ по его отсчетам $a(n\Delta t)$ на приеме: необходимо каждый отсчет умножить на функцию отсчета $\psi_n(t)$ и произведения просуммировать.

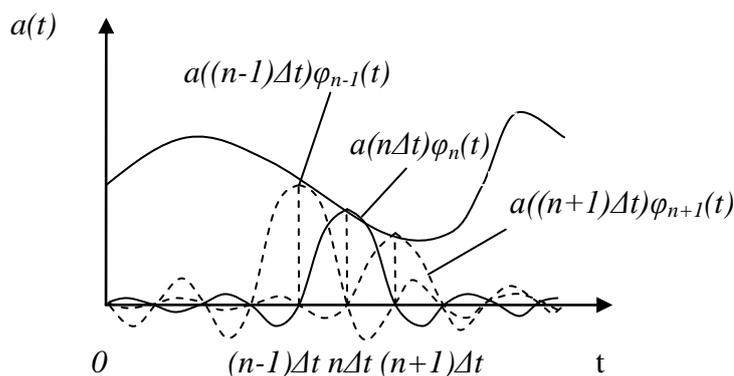


Рисунок 5.22 – Восстановление непрерывного сигнала.

Дискретизация и квантование сигналов

Аналого-цифровое преобразование

Цифровая модуляция применяется для преобразования аналогового сигнала в цифровой. Использование цифровых сигналов позволяет повысить качественные показатели средств связи с одновременным расширением их функциональных возможностей.

Преобразование непрерывного сигнала в цифровой осуществляется с помощью трех операций:

- дискретизации по времени;
- квантования по уровню;
- кодирования.

Они осуществляются в устройстве, называемом аналого-цифровым преобразователем (АЦП) и содержащем дискретизатор, квантователь и кодер. Обычно квантователь и кодер совмещаются в одном устройстве.



Рисунок 5.23 – Структурная схема АЦП.

Дискретизация – представление непрерывного сигнала эквивалентной ему по информационному содержанию последовательностью дискретных отсчетов (выборок). Эта процедура осуществляется на основе теоремы Котельникова: непрерывный сигнал с ограниченным высшей частотой F_{\max} спектром может быть представлен последовательностью импульсных отсчетов, величина которых равна или пропорциональна мгновенным значениям сигнала в соответствующие моменты времени, причем частота дискретизации должна удовлетворять требованию: $f_{\text{д}} \geq 2F_{\max}$. Совокупность полученных дискретных отсчетов представляет собой АИМ сигнал.

Для формирования отсчетов можно использовать электронный ключ, который через интервал Δt замыкается на короткое время.

Квантование – придание величинам импульсов «округленных» (квантованных) значений, т.е. формирование дискретного и по уровню сигнала из дискретного по времени. Шаг квантования Δ - разница между двумя ближайшими уровнями квантования. Совокупность уровней квантования называется шкалой квантования. Если шаг квантования не зависит от уровня квантования (рис. 5.24), то квантование является равномерным (равномерная шкала квантования). Если шаг квантования зависит от уровня квантования, то квантование является неравномерным (неравномерная шкала квантования). При этом отсчеты с меньшей амплитудой округляются менее грубо, чем с большей. В результате квантования по уровню АИМ сигнала получают квантованный АИМ (КАИМ) сигнал.

Из-за округления в процессе квантования возникает погрешность, поскольку квантованное значение отсчета отличается от истинного. Эта погрешность является специфической помехой любого АЦП и называется шумом квантования, который определяется выражением: $\varepsilon(t) = s_{\text{КАИМ}}(t) - s_{\text{АИМ}}(t)$. Шум равномерного квантования представляет собой случайную последовательность импульсов (рис. 5.24), максимальное значение которых не превышает половины шага квантования:

$$\varepsilon_m = \frac{\Delta}{2}.$$

Из выражения следует, что амплитуда импульсов шума квантования зависит от шага квантования, который в свою очередь определяется числом уровней квантования. Увеличивая число уровней квантования, можно уменьшить ошибку квантования. При

неизменном количестве уровней квантования шумы квантования можно уменьшить, применяя неравномерную шкалу.

КОДИРОВАНИЕ – процесс преобразования дискретных по уровню и времени сигналов в код (обычно двоичный). Существуют способы кодирования:

- непосредственное - квантованные отсчеты преобразуются в кодовые комбинации, обозначающие номер соответствующего отсчету уровня квантования (рис. 5.24). Реализуется при импульсно-кодовой модуляции (ИКМ);

- разностное (кодирование с предсказанием) - кодируются разности истинного и предсказанного (обычно предшествующего) значений сигналов. Реализуется при дифференциальной ИКМ (ДИКМ), дельта-модуляции (ДМ).

Импульсно-кодовая модуляция

Кодовые комбинации передаются за время равное периоду дискретизации Δt и при двоичном коде содержат n символов:

$$n = \log_2 L,$$

где $L = \frac{u_{\max} - u_{\min}}{\Delta} + 1$ - число уровней квантования;

u_{\max} и u_{\min} - соответственно максимальное и минимальное значения амплитуд импульсов квантуемого сигнала.

С количеством символов в кодовой комбинации связано качество преобразования сигнала и скорость цифрового потока на выходе АЦП:

$$v = f_{\text{д}} \cdot n = f_{\text{д}} \cdot \log_2 \left(\frac{u_{\max} - u_{\min}}{\Delta} + 1 \right), \text{ бит/с.}$$

Скорость цифрового потока определяет ширину спектра цифрового сигнала:

$$\Delta f = v, \text{ Гц.}$$

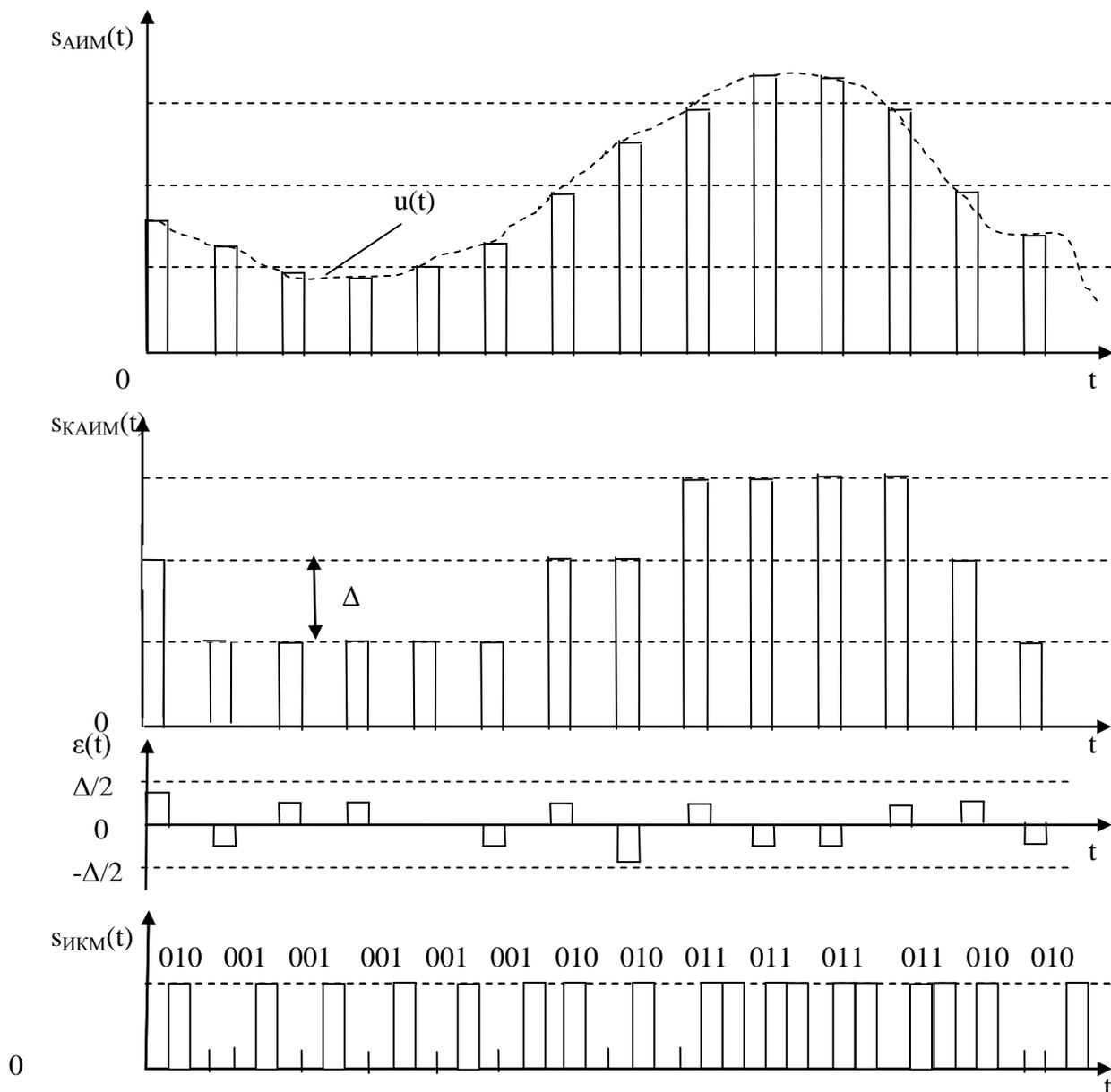


Рисунок 5.24 – Временные диаграммы, поясняющие аналого-цифровое преобразование ИКМ: дискретизированный по времени сигнал; квантованный по уровню сигнал; шум квантования; ИКМ сигнал.

Модуляция в современных системах связи

Если изменение фазы может принимать всего два значения, то говорят о двоичной фазовой модуляции (Binary Phase Shift Key, BPSK). Математически сигнал, соответствующий логическому нулю, можно представить как π , а сигнал, соответствующий логической единице, — как 0 . Тогда модулированный сигнал можно записать в виде: $V(t) \cos(\omega t + \phi)$, где $V(t)$ — управляющий сигнал, принимающий значения $+1$ и -1 . Причем значение сигнала $+1$ соответствует логическому нулю, а значение сигнала -1 — логической единице.

Изменение фазы может иметь и более двух значений, например, четыре (45° , 135° , 225° и 315°). В этом случае говорят о так называемой квадратурной фазовой модуляции (Quadrature Phase Shift Key, QPSK)

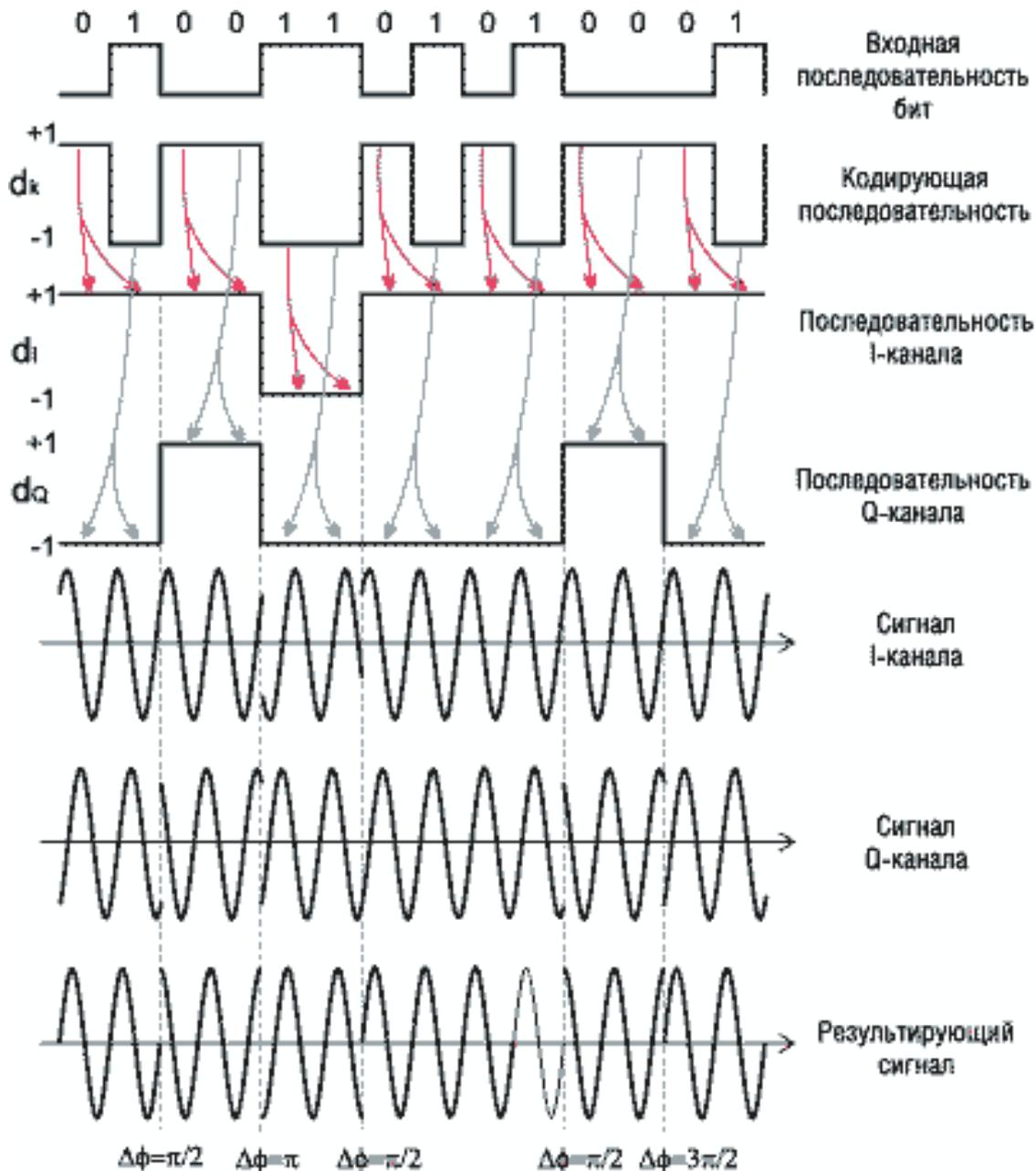


Рисунок 5.25 - Квадратурная фазовая модуляция QPSK

Чтобы понять происхождение этого термина, рассмотрим общий вид сигнала, модулированного по фазе: $S(t) = A \sin(2\pi ft + \varphi(t))$

С учетом простейших тригонометрических соотношений данную формулу не-сложно привести к виду:

$$S(t) = A \sin(2\pi ft) \cos \varphi(t) + A \cos(2\pi ft) \sin \varphi(t).$$

Из полученного выражения видно, что исходный сигнал можно представить в виде суммы двух гармонических составляющих, смещенных друг относительно друга по фазе на 90° , так как

$$\cos(2\pi ft) = \sin\left(2\pi ft + \frac{\pi}{2}\right)$$

В передатчике, производящем модуляцию, одна из этих составляющих синфазна сигналу генератора, а вторая находится в квадратуре по отношению к этому сигналу (отсюда — квадратурная модуляция). Синфазная составляющая обозначается как I (In Phase), а квадратурная — как Q (Quadrature).

Таблица 5.2 - Соотношение между сдвигом фазы и кодирующими сигналами

Фаза сигнала	d_i	d_q
45°	+1	+1
135°	+1	-1
225°	-1	-1
315°	-1	+1

При квадратурной амплитудной модуляции передаваемый сигнал кодируется одновременными изменениями амплитуды синфазной (I) и квадратурной (Q) компонент несущего гармонического колебания (f_c), которые сдвинуты по фазе друг относительно друга на $\pi/2$. Результирующий сигнал Z формируется в результате суммирования этих колебаний. Таким образом, QAM -модулированный дискретный сигнал может быть представлен соотношением:

$$Z_m(t) = I_m \cdot \cos(2\pi f_c t) + Q_m \cdot \sin(2\pi f_c t), \text{ где}$$

Таким образом, при использовании квадратурной амплитудной модуляции передаваемая информация кодируется одновременными изменениями амплитуды и фазы несущего колебания. На рисунке 5.26 представлен принцип формирования результирующего колебания Z путем суммирования вектора квадратурной составляющей Q с вектором синфазной составляющей I .

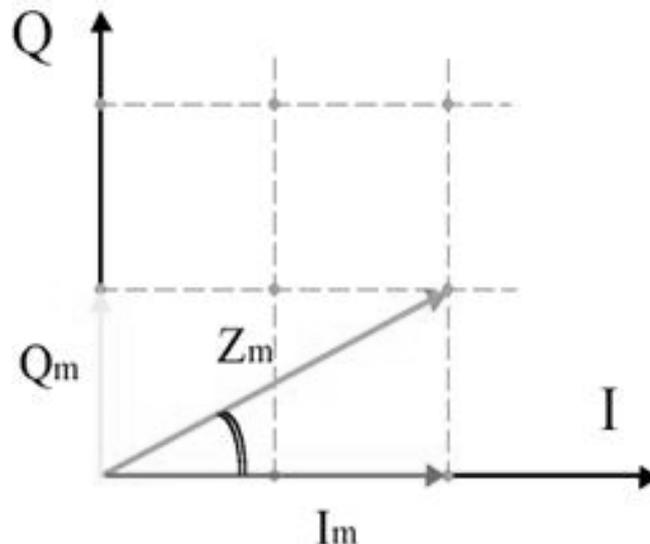


Рисунок 5.26 – Векторная диаграмма

Амплитуда вектора Z определяется соотношением A_m , а угол, который этот вектор образует с осью абсцисс, определяется соотношением φ_m .

Число узлов этой сетки определяется типом используемого алгоритма QAM. Схему расположения узлов на фазовой плоскости модулированного QAM колебания принято называть созвездием (constellation).

Для указания типа алгоритма QAM принята следующая схема обозначения: QAM- \langle число \rangle .

Используемое в обозначении алгоритма числовое значение обычно представляет собой число вида 2^N и соответствует количеству узлов на фазовой сетке, а также максимальному количеству различных значений вектора модулированного сигнала. Следует отметить, что в данном случае значение N соответствует показателю спектральной эффективности алгоритма.

ГЕНЕРИРОВАНИЕ КОЛЕБАНИЙ

Генераторы с внешним возбуждением. Умножение частоты

Генератор (от лат. – производитель) – устройство, преобразующее энергию источника питания в энергию электрических колебаний требуемой формы, частоты и мощности.

Классификация генераторов

1) *По способу возбуждения* различают генераторы с внешним возбуждением (ГВВ) и автогенераторы (АГ).

ГВВ – устройство, работающее в вынужденном режиме, т.е. колебания на его выходе наблюдаются только при наличии колебаний от внешнего источника на его входе. ГВВ предназначены для усиления мощности, умножения частоты колебаний, осуществления АМ и ЧМ.

АГ – устройство, работающее в автоколебательном режиме, т.е. колебания на его выходе возникают без внешнего источника сигнала. Являются первоисточниками электрических сигналов различной формы.

2) *По форме генерируемых колебаний* различают АГ гармонических и негармонических (релаксационных или импульсных) колебаний.

Гармонические колебания формируются в процессе плавного обмена энергиями между магнитным и электрическим полями, концентрирующимися в катушке индуктивности и конденсаторе. Используются в радиотехнических и измерительных устройствах.

Релаксационные колебания формируются в результате накопления энергии в поле реактивного элемента с последующей отдачей ее резистору, где она безвозвратно переходит в тепло (рассеивается). Используются в импульсной и цифровой технике.

3) *По частоте генерируемых колебаний* различают инфранизкочастотные (менее 10 Гц), низкочастотные (от 10 Гц до 100 кГц), высокочастотные (от 100 кГц до 100 МГц) и сверхвысокочастотные (свыше 100 МГц) генераторы.

4) *По выходной мощности* различают маломощные (менее 1 Вт), средней мощности (ниже 100 Вт) и мощные (свыше 100 Вт) генераторы.

5) *По типу используемых активных элементов* различают генераторы ламповые, транзисторные, на операционных усилителях, на туннельных диодах, на диодах Шоттки.

6) *По виду частотно-избирательной цепи* различают генераторы LC-, RC- и RL-типа.

7) *По виду обратной связи* различают генераторы с внутренней (с отрицательным сопротивлением) и с внешней (специально созданной) обратной связью.

8) *По схеме питания* различают генераторы последовательного (транзистор и колебательный контур включены последовательно по отношению к источнику питания) и параллельного (транзистор и колебательный контур включены параллельно по отношению к источнику питания) питания.

9) *По способу подключения нагрузки* (по числу точек, в которых колебательный контур соединен с активным элементом) различают двухточечные и трехточечные генераторы.

Структурная схема автогенератора

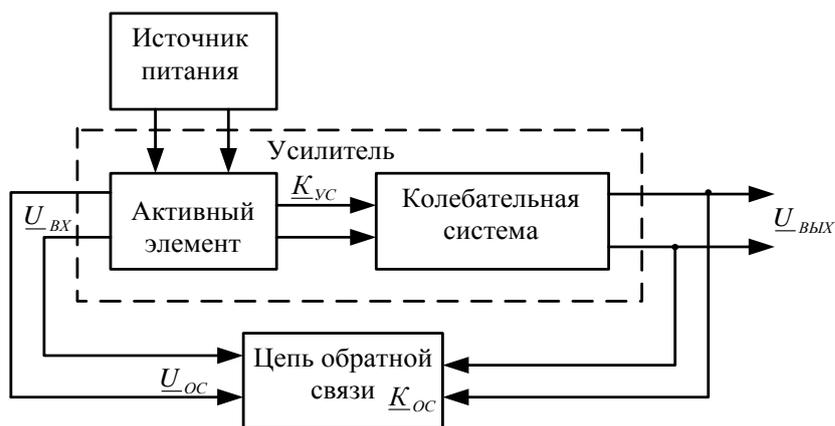


Рисунок 6.1 – Структурная схема LC -автогенератора.

Основным узлом генератора является колебательная система (колебательный контур). Она определяет форму генерируемых колебаний.

Для поддержания незатухающих колебаний энергия от внешнего источника питания через управляемый активный элемент периодически добавляется в колебательную систему, компенсируя потери в ней.

Управление активным элементом необходимо для того, чтобы энергия добавлялась синфазно с существующими в системе колебаниями. Управление на него подается с колебательной системы через цепь положительной обратной связи.

Источник питания, активный элемент и цепи обратной связи являются вспомогательными узлами, с помощью которых компенсируются потери энергии в колебательной системе.

Процесс возбуждения колебаний В АГ

Функционирование генератора можно разделить на два этапа:

- этап возбуждения генератора;
- этап стационарного режима.

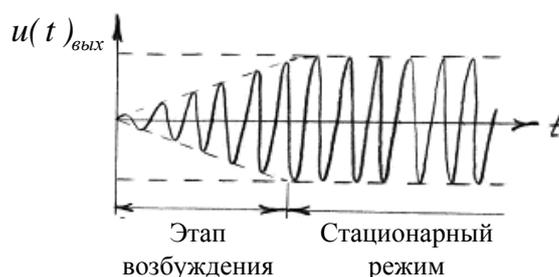


Рисунок 6.2 – Процесс установления колебаний в АГ.

После включения источника питания в генераторе начинается процесс возникновения колебаний. В первый же момент во всех цепях проходят кратковременные импульсы токов. Т.к. одиночный импульс образует сплошной спектр колебаний, частота одного из них обязательно совпадет с собственной частотой колебательной системы генератора. Это колебание возбудит колебательную систему, и по цепи обратной связи

поступит на вход активного элемента, многократно усилится и «просуммируется» с существующими в колебательной системе колебаниями. Амплитуда колебаний при этом будет непрерывно возрастать. Физически это объясняется тем, что за один период колебаний энергии поступает в колебательную систему больше, чем расходуется.

Энергетическое равновесие В АГ

По мере роста амплитуды колебаний начинает проявляться нелинейность ВАХ активного элемента и усиление уменьшается. Нарастание амплитуды колебаний АГ замедляется, а затем и вовсе прекращается. Наступает стационарный режим. Он характеризуется динамическим равновесием между поступлением энергии в колебательную систему и ее потерями. В установившемся режиме генератор выдает колебания постоянной частоты и амплитуды.

Режимы и условия самовозбуждения автогенератора

Стационарный режим работы автогенератора характеризуется постоянной амплитудой генерируемых колебаний.

Коэффициент усиления усилителя без обратной связи (см. структурную схему автогенератора):

$$\underline{K}_n = \frac{\dot{U}_{mвы}}{\dot{U}_{mв}} = K_n e^{i\varphi},$$

где $\dot{U}_{mвы}$ и $\dot{U}_{mв}$ - соответственно комплексные амплитуды выходного и входного напряжений;

$$K_n = \frac{U_{mвы}}{U_{mв}} - \text{модуль коэффициента усиления};$$

φ - его аргумент, учитывающий сдвиг фаз между входным и выходным напряжениями усилителя.

Если в качестве колебательной системы использован одиночный колебательный контур с резонансным сопротивлением $R_{рез}$, то $U_{mвы} = I_{m1} \cdot R_{рез}$,

где I_{m1} - амплитуда первой гармоники выходного тока усилительного элемента.

$$I_{m1} \approx S_{cp} \cdot U_{mв},$$

где S_{cp} - средняя крутизна ВАХ усилительного прибора.

При малых значениях $U_{mв}$ крутизна S_{cp} практически равна статической крутизне в рабочей точке, при увеличении $U_{mв}$ она уменьшается.

Подставляя последнее выражение в предыдущее, получаем:

$$U_{mвы} = S_{cp} \cdot U_{mв} \cdot R_{рез}.$$

Тогда $K_n = S_{cp} \cdot R_{рез}$ и $\underline{K}_n = S_{cp} \cdot R_{рез} \cdot e^{i\varphi}$

Коэффициент передачи цепи обратной связи:

$$\underline{K}_{oc} = \frac{\dot{U}_{mв}}{\dot{U}_{mвы}} = K_{oc} e^{i\varphi_{oc}}$$

где φ_{oc} - аргумент коэффициента передачи, определяющий сдвиг фаз между входным и выходным напряжениями ЦОС.

Результирующий коэффициент усиления усилителя, охваченного цепью ОС:

$$\underline{K} = \underline{K}_n \cdot \underline{K}_{oc}.$$

Если в колебательной системе установятся синусоидальные колебания с постоянной амплитудой (стационарный режим):

$$\underline{K} = \underline{K}_n \cdot \underline{K}_{oc} = \frac{\dot{U}_{mвы}}{\dot{U}_{mвв}} \cdot \frac{\dot{U}_{mвв}}{\dot{U}_{mвы}} = 1$$

Подставляя в это выражение значения \underline{K}_n и \underline{K}_{oc} получаем:

$$\underline{K} = S_{cp} \cdot R_{pez} \cdot e^{i\varphi} \cdot K_{oc} \cdot e^{i\varphi_{oc}} = K_{oc} \cdot S_{cp} \cdot R_{pez} \cdot e^{i(\varphi+\varphi_{oc})} = 1.$$

Это условие стационарности АГ. Оно распадается на два:

$$K_{oc} \cdot S_{cp} \cdot R_{pez} = 1 \text{ или } K_{oc} \cdot K_n = 1 \text{ и } \varphi + \varphi_{oc} = 0, \varphi = 0, 2\pi, 4\pi$$

Условие баланса амплитуд

$K_{oc} \cdot S_{cp} \cdot R_{pez} = 1$ или $K_{oc} \cdot K_n = 1$ – условие баланса амплитуд: в стационарном режиме коэффициент передачи по замкнутому кольцу генератора равен 1. Условие баланса амплитуд выполняется лишь при определенном значении $U_{mвх}$, т. е. при определенной стационарной амплитуде колебаний $U_{mвых}$. Из этого условия получаем: $K_{oc} = \frac{1}{K_n}$.

Если $K_{oc} > \frac{1}{K_n}$, амплитуда колебаний на выходе АГ нарастает до тех пор, пока вновь не выполнится условие $K_{oc} = \frac{1}{K_n}$.

При $K_{oc} < \frac{1}{K_n}$ возникновение автоколебаний невозможно, поскольку энергия, поступающая в колебательную систему недостаточна для компенсации потерь.

Условие баланса амплитуд определяет стационарную амплитуду выходных колебаний и наименьший коэффициент передачи ЦОС (критический, $K_{oc} > K_{окр}$), обеспечивающий самовозбуждение АГ.

Условие баланса фаз

Это условие определяется выражением $\varphi + \varphi_{oc} = 0, \varphi = 0, 2\pi, 4\pi$: в стационарном режиме суммарный угол сдвига фаз при обходе замкнутого кольца автогенератора должен быть равен 0 или целому числу 2π .

В цепях автогенератора могут быть следующие сдвиги фаз:

1. Сдвиг фаз на угол $\varphi_1 = \pi$, создаваемый усилительным элементом (например, транзистором при его включении по схеме с общим эмиттером), между его входным и выходным напряжениями.

2. Сдвиг фаз на угол φ_{oc} , возникающий в цепи обратной связи между ее входным и выходным напряжениями.

3. Сдвиг фаз на угол φ_2 между напряжением на входе усилительного элемента и первой гармоникой его выходного тока. Этот сдвиг возникает на очень высоких частотах и при правильном выборе лампы или транзистора $\varphi_2 \approx 0^\circ$.

4. Сдвиг фаз на угол φ_3 между напряжением $\dot{U}_{mвы}$ и током \dot{I}_{m1} . Если колебательный контур точно настроен на частоту первой гармоники выходного тока, угол $\varphi_3 = 0^\circ$.

Таким образом, условие баланса фаз можно переписать в следующем виде:

$$\varphi_1 + \varphi_{oc} + \varphi_2 + \varphi_3 = 180^\circ + \varphi_{oc} + 0^\circ + 0^\circ = 0; 360; 720, \dots \text{ или } \varphi_{oc} = \pm 180^\circ.$$

Соотношение означает, что для выполнения условия баланса фаз цепь обратной связи должна изменять фазу подводимого к ней переменного напряжения на 180° . В большинстве автогенераторов существует лишь одна частота, на которой выполняется условие баланса фаз, т. е. на которой возможно генерирование колебаний. Следовательно, условие баланса фаз определяет частоту автоколебаний.

Режим мягкого самовозбуждения АГ

В зависимости от значений постоянных питающих напряжений, подведенных к электродам усилительного элемента, и от коэффициента K_{0c} возможны два режима самовозбуждения: мягкий и жесткий.

В режиме мягкого самовозбуждения рабочую точку А выбирают на линейном участке ВАХ усилительного элемента (рисунок 8.8,а), что обеспечивает начальный режим работы усилительного элемента без отсечки выходного тока. В этих условиях самовозбуждение возникает от самых незначительных изменений входного напряжения, всегда имеющих в реальных условиях из-за флуктуации носителей заряда.

Сначала колебания в автогенераторе нарастают относительно быстро. Затем из-за нелинейности ВАХ усилительного элемента рост амплитуды колебаний замедляется, поскольку напряжение на его входе попадает на участки ВАХ со все меньшей статической крутизной, а это приводит к уменьшению средней крутизны S_{cp} и коэффициента передачи K_{0c} цепи обратной связи.

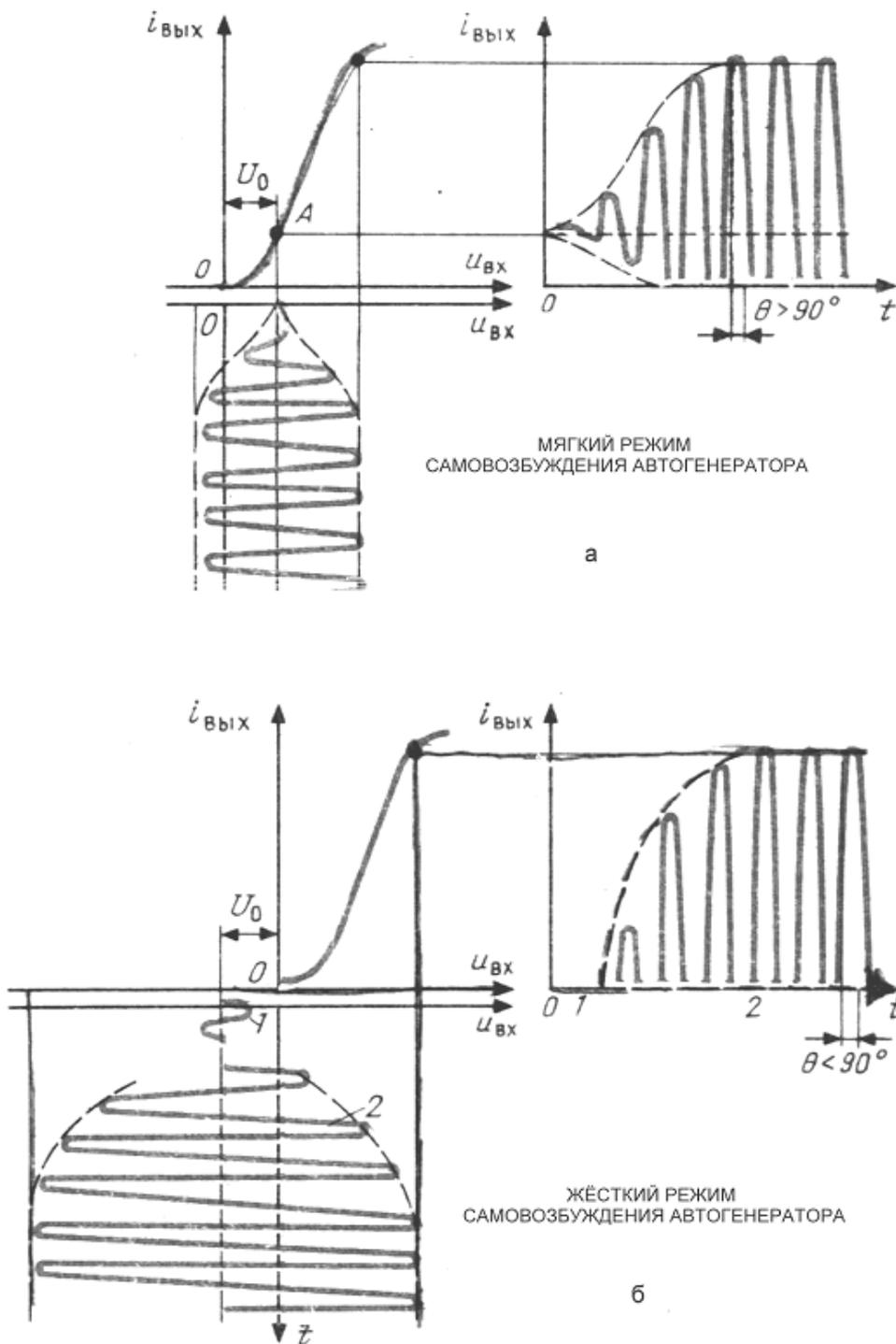


Рисунок 6.3 – Диаграммы, поясняющие режимы самовозбуждения.

Нарастание колебаний происходит до тех пор, пока коэффициент передачи K уменьшится до единицы. В результате в автогенераторе установится стационарный режим, которому соответствует определенная амплитуда выходных колебаний, причем угол отсечки выходного тока $\theta > 90^\circ$. Частота этих колебаний очень близка к резонансной частоте колебательной системы.

Обратим внимание: если бы усилительный элемент имел линейную вольт-амперную характеристику, нарастание амплитуды автоколебаний происходило бы до бесконечности, что фи-

зически невозможно. Поэтому в линейной цепи получить устойчивые автоколебания с постоянной амплитудой невозможно.

Из-за нелинейности вольт-амперной характеристики форма выходного тока усилительного элемента получается несинусоидальной. Однако при достаточно большой добротности ($Q=50 \dots 200$) колебательной системы первая гармоника этого тока и, следовательно, напряжение на выходе автогенератора представляют собой почти гармонические колебания.

Режим жесткого самовозбуждения

При этом режиме напряжение смещения задают таким, чтобы при малых амплитудах входного напряжения ток через усилительный элемент не проходил. Тогда незначительные колебания, возникшие в контуре, не могут вызвать ток в выходной цепи, и самовозбуждение автогенератора не наступает. Колебания возникают только при их достаточно большой начальной амплитуде, что не всегда можно обеспечить. Процесс возникновения и нарастания колебаний при жестком режиме самовозбуждения иллюстрируется на рисунке 8.8, б. Видно, что при малых начальных амплитудах входного напряжения (кривая 1) ток $i_{вых} = 0$ и автоколебания не возникают. Они возникают только при достаточно большой начальной амплитуде напряжения (кривая 2) и быстро нарастают до установившегося значения. В стационарном режиме усилительный элемент работает с углами отсечки выходного тока $\theta < 90^\circ$.

Для удобства эксплуатации автогенератора целесообразнее применять мягкий режим самовозбуждения, так как в этом режиме колебания возникают сразу после включения источника питания. Однако при жестком режиме колебаний с углом отсечки $\theta < 90^\circ$ обеспечиваются более высокий КПД автогенератора и меньшие тепловые потери. Поэтому в стационарном режиме автогенератора более выгоден именно режим с малыми углами отсечки выходного тока усилительного элемента.

ФОРМИРОВАНИЕ МОДУЛИРОВАННЫХ СИГНАЛОВ

Формирование двухполосных амплитудно-модулированных сигналов с большим уровнем несущей

Для формирования АМ сигнала необходимо сумму напряжений несущего колебания и модулирующего сигнала подать на вход нелинейной цепи, содержащей полупроводниковый диод или транзистор. Спектр тока в такой цепи содержит составляющие, которых нет в воздействующем на нее напряжении. Остается выделить с помощью электрического фильтра составляющие, образующие АМ сигнал.

На ВАХ диода, транзистора или лампы можно выделить квадратичный и линейный участок. Использование первого участка определяет режим слабого сигнала, при котором входное напряжение не должно заходить как в область отсечки, так и в область насыщения. Использование второго участка определяет режим сильного сигнала, при котором входное напряжение переводит транзистор в режим отсечки, а может переводить его и в режим, близкий к насыщению. ВАХ на первом участке аппроксимируют полиномом n -ой степени, а на втором участке – ломаной прямой.

Амплитудные модуляторы классифицируют:

1. по схеме соединения НЭ - на одноконтурные (содержащие один НЭ), балансные (представляющие собой два одноконтурных), и кольцевые (представляющие собой два балансных);
2. по типу применяемых НЭ – на пассивные (на полупроводниковых диодах) и активные (на лампах, транзисторах).

Одноконтурные модуляторы

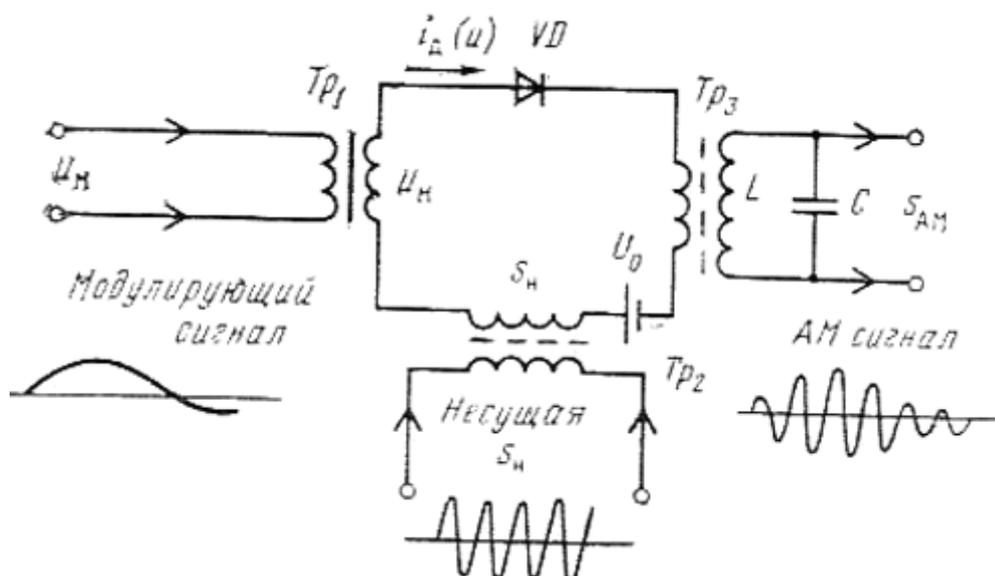


Рисунок 7.1 – Принципиальная схема диодного амплитудного модулятора.

Напряжение модулирующего сигнала $u(t)$ перемещает рабочую точку по квадратичному участку ВАХ диода путем изменения напряжения смещения $U_0 + u(t)$ на аноде относительно катода. Зависимость тока через диод от времени $i(t)$ имеет сложный характер. Приращения тока различны в положительный и отрицательный полупериоды как несущего, так и модулирующего колебаний. Ток первой гармоники $i_1(t)$ оказывается промодулированным по амплитуде сигналом $u(t)$. Напряжение $s_{AM}(t)$ на выходе колебательного контура, настроенного на частоту несущей ω_n и имеющего полосу

пропускания, равную ширине спектра АМ сигнала $\Delta\omega_{AM}$, пропорционально току этой гармонике. Остальные гармоники тока отфильтровываются, т.к. сопротивление контура на их частотах практически равно нулю.

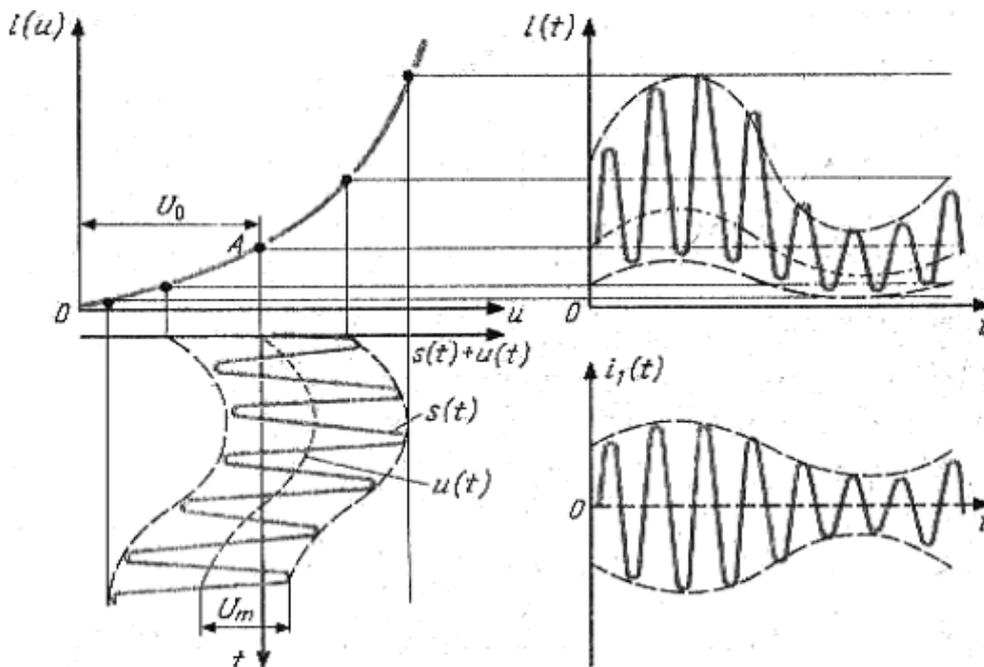


Рисунок 7.2 – Временные диаграммы работы диодного амплитудного модулятора.

Такой же вывод можно получить, воспользовавшись спектральным методом анализа нелинейной цепи. Пусть ВАХ диода представлена полиномом второй степени:

$$i(u) = a_0 + a_1 u + a_2 u^2,$$

где a_0, a_1, a_2 - коэффициенты аппроксимации;

$u = s_n + u$ - бигармоническое воздействие;

$s_n = U_m \cos \omega_n t$ - напряжение несущего колебания;

$u(t) = U_{mu} \cos \Omega t$ - напряжение модулирующего сигнала.

Вид полинома после подстановки:

$$i(u) = a_0 + a_1(s_n + u) + a_2(s_n + u)^2 = a_0 + \underline{a_1 s_n} + a_1 u + a_2 s_n^2 + \underline{2a_2 s_n u} + a_2 u^2.$$

Подчеркнуты составляющие тока с частотами, сосредоточенными вблизи частоты несущей ω_n . В сумме они образуют ток первой гармоники $i_1(t)$, промодулированный по амплитуде модулирующим сигналом:

$$i_1(t) = a_1 U_m \cos \omega_n t + 2a_2 U_m U_{mu} \cos \omega_n t \cos \Omega t = a_1 U_m (1 + 2a_2 U_{mu} / a_1 \cos \Omega t) \cos \omega_n t.$$

Напряжение на контуре:

$$s_{AM}(t) = i_1(t) Z_k = a_1 U_m Z_k (1 + 2a_2 U_{mu} / a_1 \cos \Omega t) \cos \omega_n t = U_{m1} (1 + m \cos \Omega t) \cos \omega_n t,$$

где Z_k - входное резонансное сопротивление контура;

$U_{m1} = a_1 U_m Z_k$ - амплитуда напряжения на контуре при отсутствии модуляции;

$m = 2a_2 U_{mu} / a_1$ - коэффициент амплитудной модуляции.

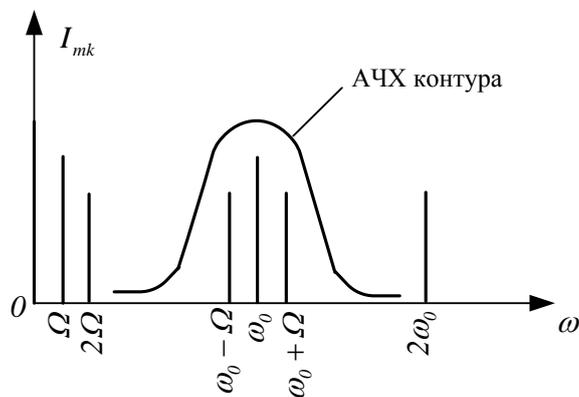


Рисунок 7.3 – Спектр тока через диод.

Т.к. диоды имеют незначительный участок с квадратичной характеристикой, то уровень АМ сигнала на выходе такого модулятора мал.

Для повышения напряжения выходного АМ сигнала диодный модулятор используют в режиме больших значений модулирующего и несущего колебаний. Еще больший его уровень будет, если модулятор выполнить на активном НЭ.

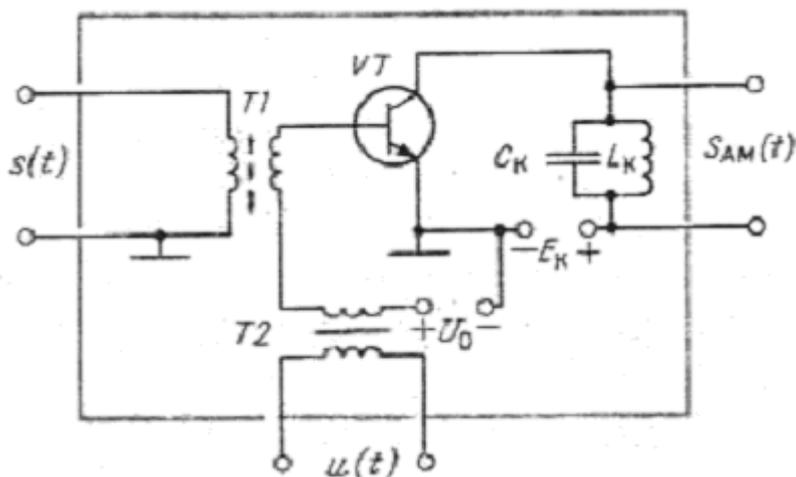


Рисунок 7.4 – Принципиальная схема амплитудного модулятора на транзисторе.

Напряжение смещения U_0 обеспечивает режим работы транзистора с отсечкой. За счет того, что рабочая точка перемещается модулирующим сигналом $u(t)$, происходит непрерывное изменение амплитуды и угла отсечки коллекторного тока $i_k(t)$. Вследствие этого амплитуда первой гармоники коллекторного тока $i_{k1}(t)$ меняется во времени пропорционально модулирующему сигналу $u(t)$. Напряжение на колебательном контуре будет представлять собой АМ сигнал.

Такой же вывод можно получить, воспользовавшись графическим методом анализа нелинейной цепи. Соответствующие построения приведены на рисунке 9.5.

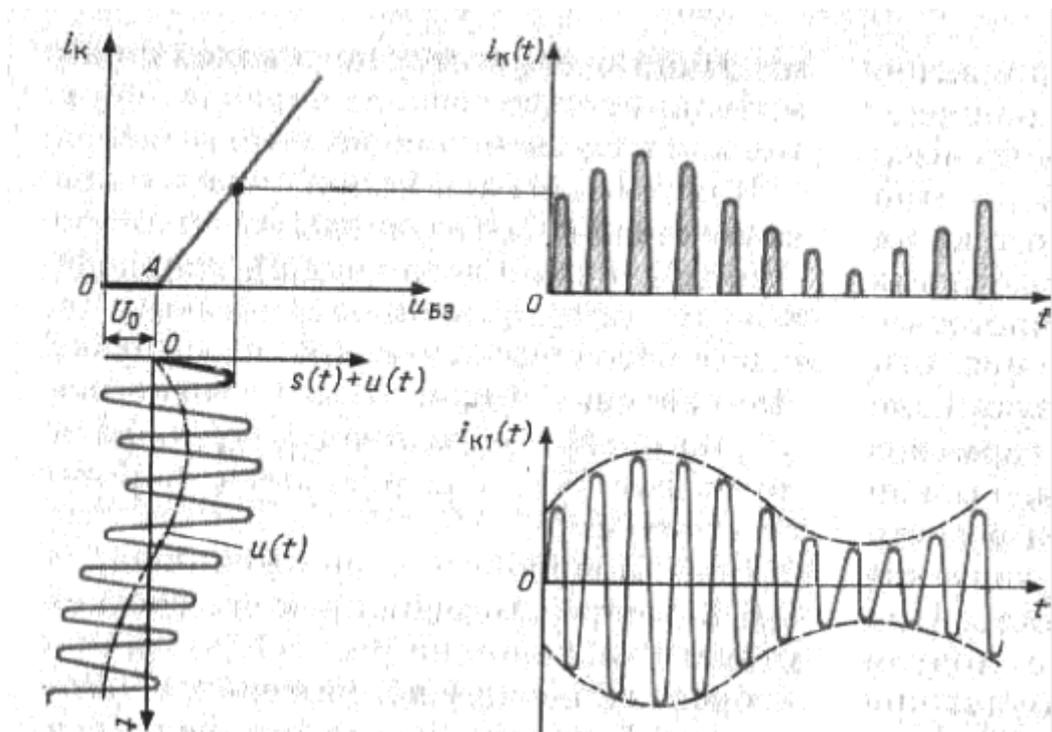


Рисунок 7.5 – Временные диаграммы работы амплитудного модулятора на транзисторе.

Амплитуда импульсов коллекторного тока:

$$I_m(t) = I_{m0} + ku(t),$$

где I_{m0} - значение амплитуды импульсов тока при отсутствии модулирующего сигнала;

k - размерный коэффициент пропорциональности.

Амплитуда первой гармоники коллекторного тока:

$$I_{m1}(t) = \alpha_1(\theta)I_m(t) = \alpha_1(\theta)(I_{m0} + ku(t)).$$

Амплитуда напряжения на контуре:

$$U_{m1}(t) = I_{m1}(t)Z_k = \alpha_1(\theta)I_{m0}Z_k(1 + kU_{mu}/I_{m0} \cos\Omega t) = U_{m1}(1 + m \cos\Omega t),$$

где $U_{m1} = \alpha_1(\theta)I_{m0}Z_k$ - амплитуда напряжения, обусловленного первой гармоникой коллекторного тока, при отсутствии модуляции;

$m = kU_{mu}/I_{m0}$ - коэффициент амплитудной модуляции.

Напряжение на выходе модулятора:

$$s_{AM}(t) = U_{m1}(1 + m \cos\Omega t) \cos\omega_c t.$$

При такой модуляции неизбежны искажения: форма огибающей АМ сигнала отличается от формы модулирующего сигнала, т.к. с изменением последнего происходит изменение угла отсечки θ и соответственно коэффициента Берга $\alpha_1(\theta)$.

Формирование АМ сигналов с подавленной несущей

Балансный (двухтактный) модулятор

Позволяет получить балансно-модулированный (БМ) сигнал, спектр которого состоит из двух боковых полос и не содержит несущего колебания.

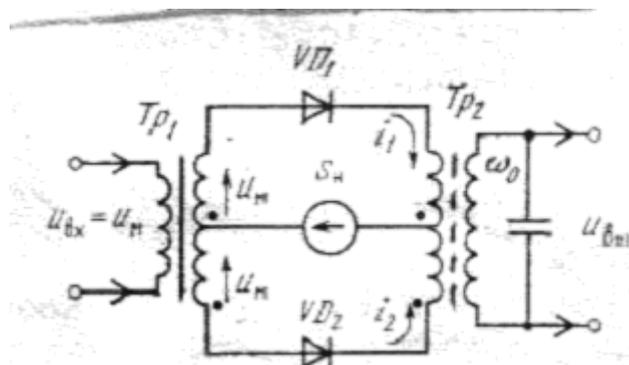


Рисунок 7.6 – Принципиальная схема диодного балансного модулятора.

При положительной полуволне напряжения несущей частоты оба диода открыты, сопротивление их мало, и через первичную обмотку трансформатора $Tr2$ протекает выходной ток. При отрицательной полуволне диоды закрыты, сопротивление их велико и тока в обмотке нет. Этот ток имеет вид импульсов. С изменением полярности модулирующего напряжения изменяется направление тока (отрицательные импульсы). Отсутствие колебания несущей частоты на выходе модулятора объясняется тем, что несущее колебание подается в средние точки трансформаторов $Tr1$ и $Tr2$ и магнитные потоки, создаваемые токами несущей частоты i_1 и i_2 в полуобмотках трансформаторов, имеют встречные направления и взаимно уничтожаются. Нагрузкой модулятора служит контур, настроенный на несущую частоту ω_n , который выделяет БМ сигнал.

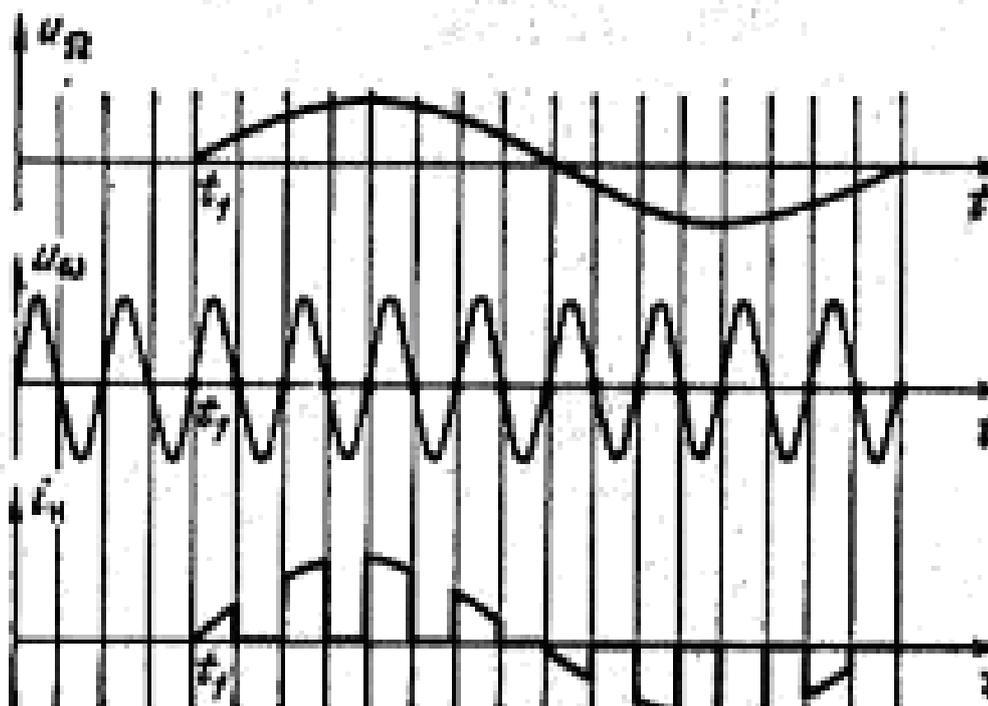


Рисунок 7.7 – Временные диаграммы модулирующего напряжения, напряжения несущей и выходного тока.

Воспользуемся спектральным методом анализа нелинейной цепи. Пусть ВАХ диодов одинаковы и аппроксимируются полиномом второй степени. Пусть в некоторый момент времени полярность напряжений такая, как указана на рисунке 9.6. Тогда напряжение на диодах:

$$u_{o1}(t) = s_n(t) + u(t); u_o(t) = s_n(t) + u(t).$$

Токи в цепях диодов:

$$i_1 = a_1(s_n + u) + a_2(s_n + u)^2;$$

$$i_2 = a_1(s_n - u) + a_2(s_n - u)^2.$$

Токи в трансформаторе $Tr2$ направлены встречно и результирующее напряжение на выходе схемы с учетом подавления некоторых составляющих контуром:

$$s_{БМ}(t) = kZ_k(i_1 - i_2) = 4kZ_k a_2 s_n(t)u(t).$$

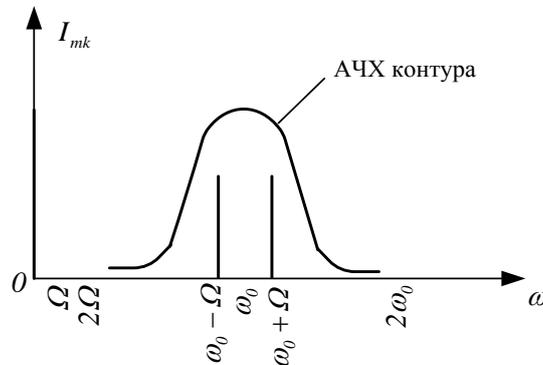


Рисунок 7.8 - Спектр выходного тока.

В балансном модуляторе в выходном токе отсутствуют составляющие нелинейного преобразования с частотами $2m\Omega$, $m\omega_n$, $\omega_n \pm 2n\Omega$. При этом облегчается выделение БМ сигнала.

Формирование однополосно-модулированных сигналов

Передача информации одной боковой полосой имеет следующие преимущества:

- не тратится мощность передатчика на передачу несущих колебаний, за счет чего можно увеличить мощность колебаний передаваемой боковой полосы, а следовательно, и дальность действия связи;
- при отсутствии модуляции мощность не расходуется, т.к. передачи несущих колебаний нет;
- меньше занимаемая полоса, что позволяет отведенную для системы полосу частот уплотнить большим числом каналов;
- требуется более узкая полоса пропускания приемника, что повышает помехозащищенность за счет снижения уровня помех в рабочей полосе.

Методы формирования ОМ сигнала

1) Метод фильтрации.

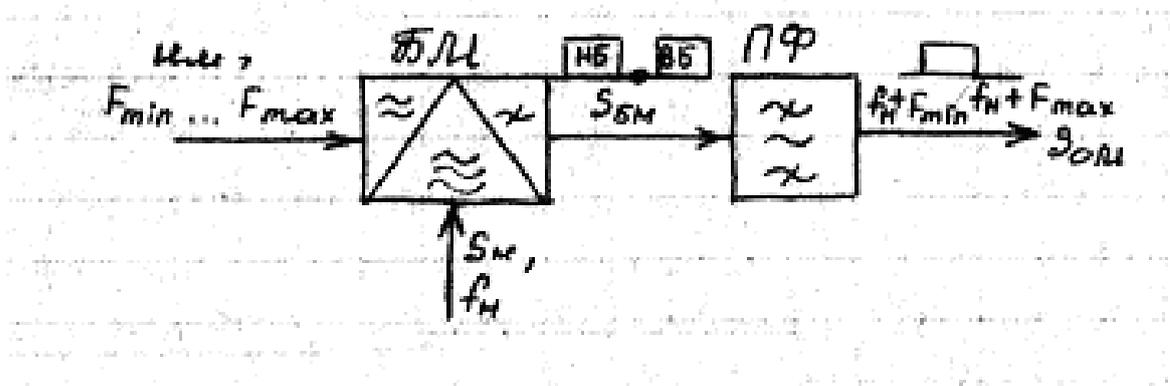


Рисунок 7.9 – Структурная схема.

С помощью балансного (или кольцевого) модулятора БМ получают двухполосный сигнал с подавленной несущей. Далее полосовым фильтром ПФ выделяется требуемая боковая полоса частот.

Недостаток: т.к. частотный разнос между боковыми полосами равен $2F_{min}$, то к ПФ предъявляются жесткие требования по полосе расфилтровки (применяются высокодобротные пьезокерамические и кварцевые фильтры).

2) Метод фазирования.

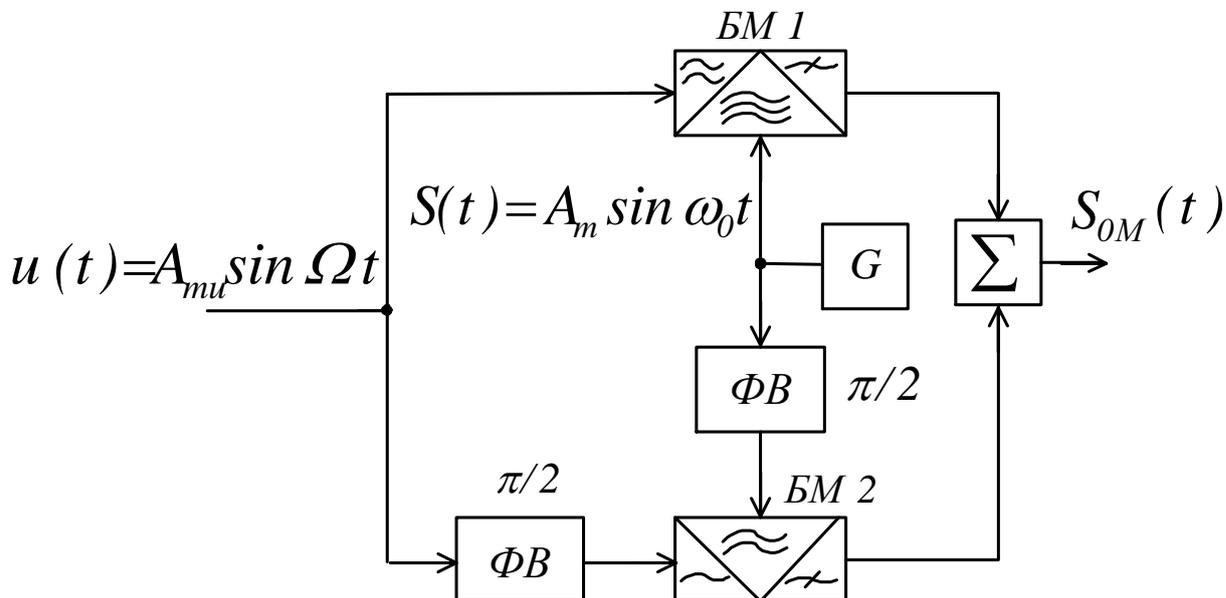


Рисунок 7.10 – Структурная схема модулятора.

Для перемножения сигналов используются балансные (или кольцевые) модуляторы БМ1 и БМ2. На БМ2 входные сигналы u_m и s_n подаются через фазовращатели на 90° $\Phi B1$ и $\Phi B2$. Если модулирующий сигнал u_m имеет более сложный спектр, то $\Phi B1$ должен обеспечить изменение фазы всех спектральных составляющих.

Для формирования нижней боковой полосы частот нужно просуммировать полученные произведения; верхней – следует использовать вычитание. Это можно доказать, воспользовавшись тригонометрическими формулами:

$$\cos \alpha \cos \beta = 1/2 \cos(\alpha - \beta) + 1/2 \cos(\alpha + \beta),$$

$$\sin \alpha \sin \beta = 1/2 \cos(\alpha - \beta) - 1/2 \cos(\alpha + \beta).$$

Формирование частотно- и фазо-модулированных сигналов

Существуют прямые и косвенные методы получения ЧМ и ФМ сигналов.

Прямой метод ЧМ

Представляет собой параметрическое управление частотой колебаний автогенератора АГ. С этой целью в колебательный контур АГ вводят дополнительную емкость или индуктивность (управляющее устройство УУ), изменяющуюся по закону модулирующего сигнала.

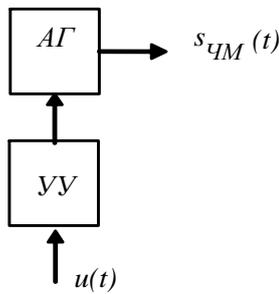


Рисунок 7.11 – Структурная схема прямого метода ЧМ.

Недостаток: снижение стабильности средней частоты автоколебаний ω_n , т.к. частота АГ должна изменяться в широких пределах.

Наиболее часто применяется частотный модулятор на основе варикапа.

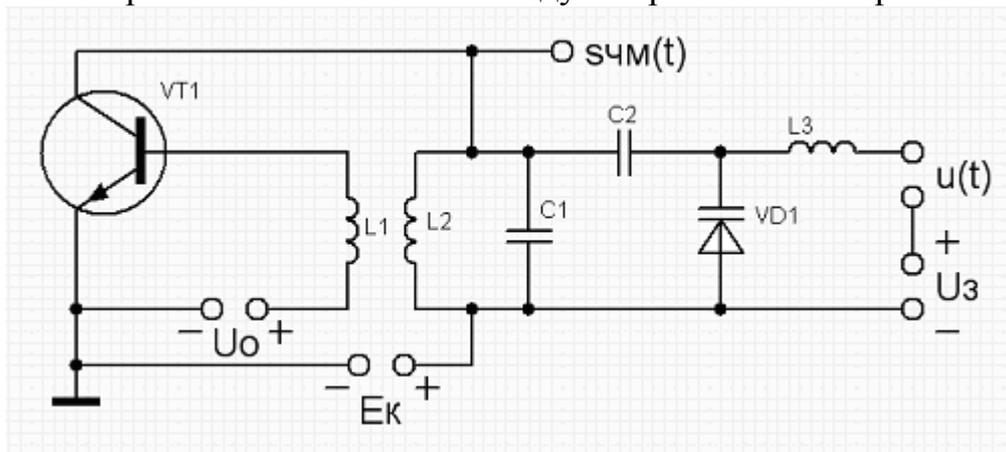


Рисунок 7.12 – Принципиальная схема частотного модулятора с варикапом.

Параллельно контуру LC -генератора с индуктивной обратной связью подключен варикап – полупроводниковый диод, емкость которого зависит от напряжения, приложенного в направлении запирающего р-п перехода. Конденсатор C_2 соединяет по высокой частоте варикап с емкостью контура C_1 и подбирается так, чтобы его сопротивление было мало на высокой генерируемой частоте и велико на частоте модулирующего сигнала. Разделительный дроссель L_3 необходим для предотвращения замыкания высокой генерируемой частоты через источники напряжений: постоянного запирающего U_3 и модулирующего $u(t)$.

Модулирующее напряжение изменяет запирающее напряжение на варикапе, вследствие чего меняется емкость варикапа и соответственно генерируемая частота.

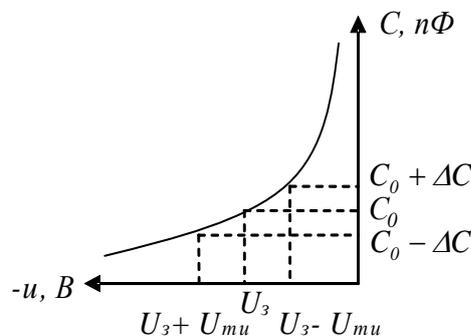


Рисунок 7.13 – Вольт-фарадная характеристика варикапа.

Тогда частота автоколебаний:

$$\omega = \omega_0 + \Delta\omega = \frac{1}{\sqrt{L_k(C_k + \Delta C)}} = \frac{\omega_0}{\sqrt{1 + \Delta C / C_k}},$$

где $\omega_n = \omega_0 = \frac{1}{\sqrt{LC_k}}$ - средняя частота автоколебаний;

$L_k = L2$ - индуктивность контура;

$C_k + \Delta C$ - емкость контура;

$C_k = C1 + C_0$ - средняя емкость контура;

C_0 - начальная емкость варикапа. Определяется напряжением U_3 ;

$\Delta C = ku(t)$ - изменение емкости контура;

$\Delta\omega$ - соответствующее ему изменение частоты.

При незначительном изменении емкости контура мгновенная частота контура будет изменяться в соответствии с законом изменения управляющего напряжения:

$$\Delta C / C_k \approx -2\Delta\omega / \omega_0.$$

Знак «минус» в выражении означает, что при увеличении емкости генерируемая частота уменьшается.

Прямой метод ФМ

Устройством, на выходе которого фаза гармонического колебания изменяется пропорционально модулирующему сигналу $u(t)$, является резонансный усилитель $У$ с LC колебательным контуром в качестве нагрузки, если к контуру усилителя подключается реактивное сопротивление ($УУ$), управляемое модулирующим сигналом. Изменение частоты настройки контура приводит к изменению фазы напряжения на контуре.

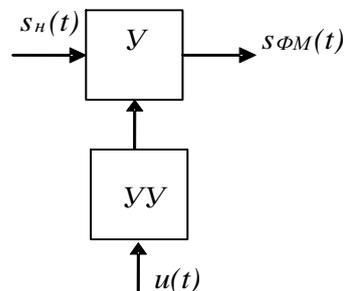


Рисунок 7.14 – Структурная схема прямого метода ФМ.

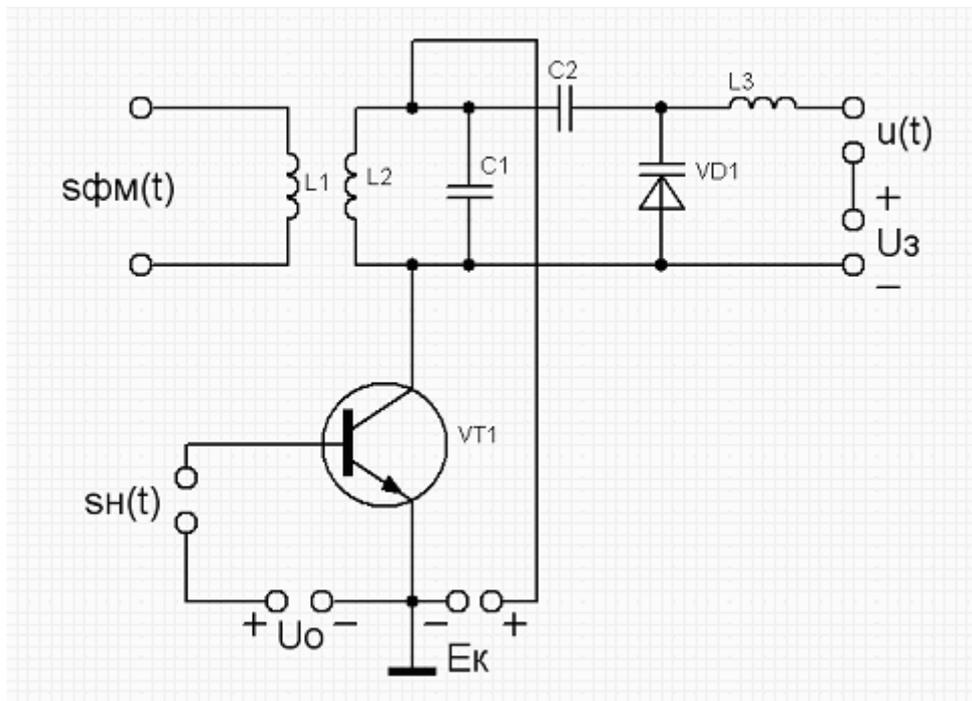


Рисунок 7.15 – Принципиальная схема фазового модулятора с варикапом.

Несущее колебание $s_n(t)$ поступает от генератора на базу транзистора $VT1$, на котором построен резонансный усилитель. Параллельно емкости контура через разделительный конденсатор $C2$ подключен варикап $VD1$, управляемый источником модулирующего сигнала.

При изменении емкости варикапа происходит изменение реактивного сопротивления контура и, следовательно, сдвига фаз φ между напряжением несущей на входе усилителя $s_n(t)$ и напряжением на резонансном контуре $s_{\phi M}(t)$.

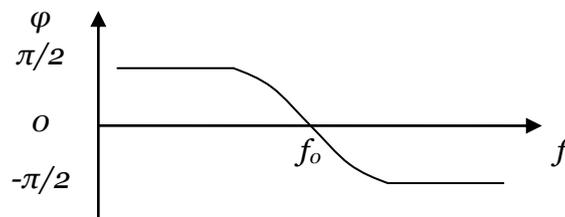


Рисунок 7.16 – ФЧХ контура.

Уравнение фазовой характеристики контура с добротностью Q для небольших расстройек $\Delta\omega = \omega - \omega_0$ имеет вид:

$$\varphi = \arctg(-2\Delta\omega Q / \omega_0) = \arctg(\Delta C Q / C_k).$$

Неискаженная ФМ имеет место, когда изменение φ пропорционально ΔC , т.е. на линейном участке фазовой характеристики, где $\text{tg}\varphi \approx \varphi$.

Косвенный метод ЧМ

Состоит в преобразовании ФМ в частотную. Для этого на входе фазового модулятора помещают интегратор. Таким образом, ЧМ сигнал получают в результате фазовой модуляции интегральной функцией модулирующего сигнала.

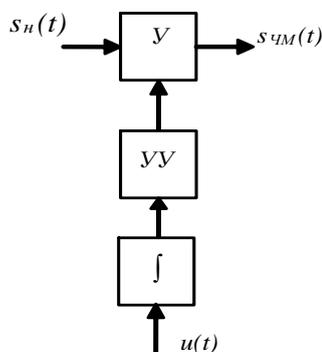


Рисунок 7.17 – Структурная схема косвенного метода ЧМ.

Покажем, что ФМ можно преобразовать в ЧМ.

При ЧМ частота изменяется по закону:

$$\omega_{\text{ЧМ}}(t) = \omega_n + k_{\text{ЧМ}}u(t),$$

а фаза:

$$\varphi_{\text{ЧМ}}(t) = \int_0^t \omega_{\text{ЧМ}}(\tau) d\tau = \omega_n t + k_{\text{ЧМ}} \int_0^t u(\tau) d\tau + \varphi_0.$$

где $k_{\text{ЧМ}}$ - размерный коэффициент пропорциональности.

Достоинство: возможность обеспечения сколь угодно высокой стабильности средней частоты ω_n (кварцевая стабилизация), поскольку модуляция осуществляется в промежуточном каскаде, а не в АГ.

Недостаток: невозможность получения широкополосной ЧМ (с большой девиацией частоты); сложность в изготовлении и настройке.

Первый недостаток устраняется путем получения небольших девиаций на низкой частоте с последующим умножением несущей частоты (во столько же раз увеличивается и девиация частоты).

Косвенный метод ФМ

Состоит в преобразовании ЧМ в фазовую. Для этого на входе частотного модулятора помещают дифференцирующую цепь. Таким образом, ФМ сигнал получают в результате частотной модуляции дифференциальной функцией модулирующего сигнала.

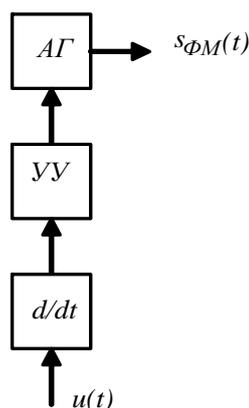


Рисунок 7.18 – Косвенный метод ФМ.

Покажем, что ЧМ можно преобразовать в ФМ.

При ФМ фаза изменяется по закону:

$$\varphi_{\text{ФМ}}(t) = \omega_n t + k_{\text{ФМ}}u(t) + \varphi_0,$$

а частота:

$$\omega_{\phi_M}(t) = d\phi_{\phi_M}(t) / dt = \omega_n + k_{\phi_M} du(t) / dt,$$

где t - текущий момент времени;

k_{ϕ_M} - размерный коэффициент пропорциональности;

ϕ_0 - начальная фаза несущей.

Формирование импульсно-модулированных сигналов

Прямые методы осуществляются теми же модуляторами, что и аналоговые модуляции, с небольшими изменениями: в качестве несущей используется периодическая последовательность прямоугольных импульсов (ППИ), а вместо избирательной цепи на выходе модулятора включается резистивная нагрузка. Последнее связано с достаточно широким спектром импульсного сигнала.

Косвенные методы основаны на преобразовании одного вида модуляции в другой, в том числе аналоговой в импульсную.

Амплитудно-импульсная модуляция

Наиболее часто применяется прямой метод формирования АИМ. Для этого используется любая схема амплитудного модулятора, например транзисторного, работающего в режиме с отсечкой. Вместо транзистора в амплитудно-импульсных модуляторах используют также диоды.

Частотно-импульсная модуляция

ЧИМ можно получить, управляя частотой любого релаксационного автогенератора, например мультивибратора или блокинг-генератора. Для изменения частоты следования импульсов, как и в автогенераторе гармонических колебаний, необходимо изменять параметры колебательной системы. Кроме изменения ёмкости варикапами, в релаксационных автогенераторах также используются управляемые резисторы, так как они обеспечивают большую линейность и пределы изменения частоты.

Косвенный метод получения ЧИМ основан на преобразовании ЧМ в ЧИМ (рисунок 7.19). Для этого ЧМ сигнал (a) пропускают через ограничитель по максимуму и минимуму с достаточно низкими порогами ограничения.

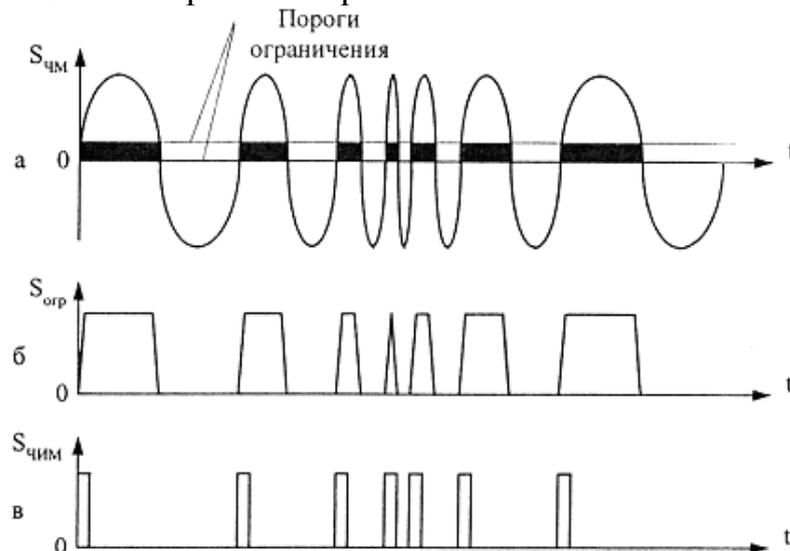


Рисунок 7.19 – Временные диаграммы, поясняющие преобразование ЧМ сигнала в ЧИМ: а) ЧМ сигнал; б) сигнал на выходе ограничителя; в) ЧИМ сигнал.

Из полученных трапецеидальных импульсов (рисунок 7.19, б) дифференцированием и последующим ограничением выделяют передние фронты (рисунок 7.19, в). Это и будет ЧИМ сигнал, частота следования импульсов которого изменяется пропорционально модулирующему сигналу.

Широтно-импульсная и фазо-импульсная модуляция

Одним из простых, но широко распространённых методов формирования ШИМ и ФИМ является получение их из АИМ. Этот процесс поясняется временными диаграммами рисунка 7.20. В качестве несущей для АИМ используется периодическая последовательность треугольных импульсов (рисунок 7.20, б). Если выходной АИМ сигнал (рисунок 7.20, в) пропустить через ограничитель по максимуму с низким порогом ограничения, то на выходе ограничителя получим трапецеидальные импульсы ШИМ (рисунок 7.20, г). Крутизна фронтов импульсов зависит от порога ограничения. Если далее из полученного ШИМ сигнала, как и в примере косвенного получения ЧИМ, дифференцированием и ограничением выделить передние фронты, то получим сигнал с ФИМ (рисунок 7.20, д), сдвиг импульсов которого относительно центров треугольных импульсов (рисунок 7.20, б) пропорционален модулирующему сигналу.

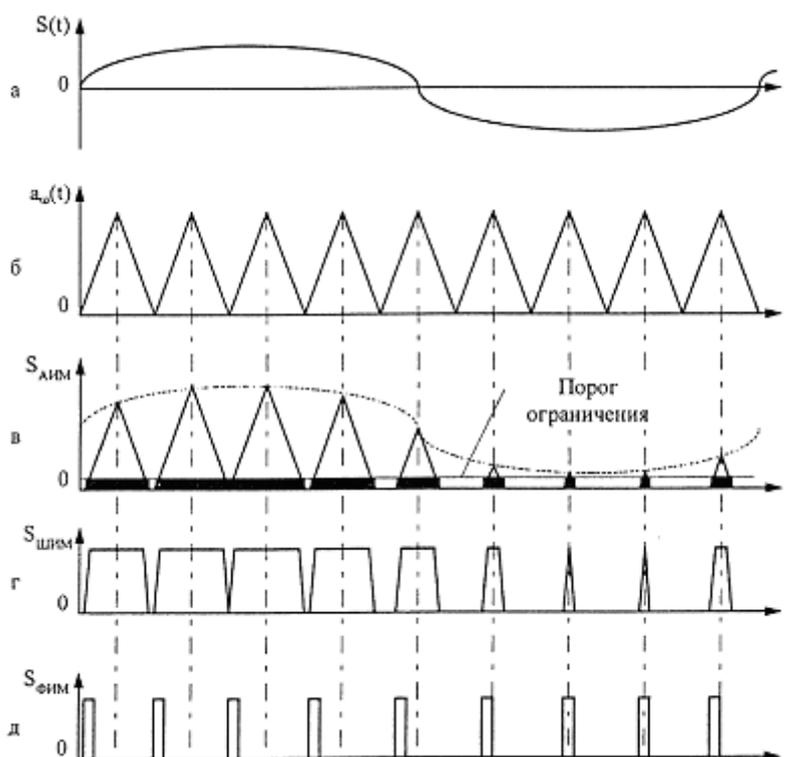


Рисунок 7.20 – Временные диаграммы, поясняющие преобразование АИМ сигнала в ШИМ и ФИМ: а) модулирующий сигнал; б) импульсная несущая; в) АИМ сигнал; г) ШИМ сигнал; д) ФИМ сигнал.

Формирование манипулированных сигналов

Любая манипуляция (дискретная модуляция) несущей является частным случаем аналоговой модуляции, поэтому все способы и схемы получения аналоговых видов модуляции полностью пригодны и для формирования манипулированных сигналов, если последовательность стандартных элементов дискретного первичного сигнала $u_i(t)$ используется в качестве модулирующей функции. Если $u_i(t)$ – двоичный сигнал, то он может принимать два значения: $(-U_0, +U_0)$ – двухполярный сигнал и $(0, +U_0)$ – одно-

полярный сигнал. В аналоговых модуляторах следует применять двухполярную модулирующую функцию.

Исходя из дискретного характера модулирующей функции и модулированного сигнала, можно осуществить дискретные виды модуляции с лучшим качеством принципиально другим способом: сформировать специальными устройствами (например автогенераторами G_1, \dots, G_m) дискретные сигналы $s_1(t), \dots, s_m(t)$ и далее коммутировать их управляемыми ключами в соответствии с последовательностью дискретных элементов $u_i(t)$ (рисунок 7.21).

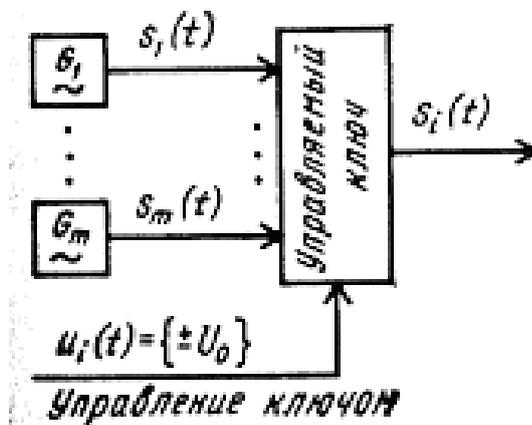


Рисунок 7.21 – Структурная схема устройства для получения дискретно-модулированных сигналов.

Ключевые схемы формирования манипулированных сигналов в настоящее время являются основными, так как позволяют получать с высокой точностью дискретные сигналы $s_i(t)$ с заранее заданными параметрами и свойствами. В качестве ключей применяются диоды, транзисторы, специальные микросхемы. Основное требование к ключам – неискажённая передача $s_i(t)$ со входа на выход и отсутствие прохождения управляющего сигнала $u_i(t)$ на выход, т.е. схема ключа должна быть балансной для $u_i(t)$. Как ключи в настоящее время широко используются двойные балансные (кольцевые) модуляторы.

Формирование ОФМ

Для получения ОФМ можно применить любую схему формирования ФМн. Однако перед фазовым модулятором в цепь модулирующей дискретной функции включается специальное перекодирующее устройство, обеспечивающее требуемое правило формирования ОФМ: изменение фазы несущей частоты происходит только при подаче единичных элементов дискретной последовательности. Это перекодирующее устройство называют относительным кодером.

Функциональная схема относительного кодера для однополярной дискретной последовательности показана на рисунке 7.22,а. Кодер содержит два блока: сумматор по модулю два $M2$ и линию задержки на длительность дискретного элемента сигнала τ_u . Работу схемы можно проследить по графикам рисунка 7.22, б, где цифрами отмечена форма сигналов в различных точках схемы. На вход 1 поступает последовательность дискретных однополярных сигналов, которые обозначены как 0 и 1.

Суммируя дискретные сигналы на входе 1 и выходе линии задержки 2, получаем выходной сигнал 3. Наклонными стрелками на графиках показана задержка выходного сигнала 3 на время τ_u для получения сигнала 2.

Легко проследить, что если выходной сигнал относительного кодера 3 подать на фазовый модулятор, то фаза несущей частоты на его выходе будет меняться согласно правилу ОФМ. Заметим, что при ФМн фаза несущей меняется при переходе модулирующей функции от 0 к 1 и обратном переходе. На временной диаграмме звездочкой отмечены моменты времени, где должна меняться фаза при ОФМ (график 1) и где она меняется, если на ФМ подать управляющий сигнал с выхода относительного декодера (график 3). Они совпадают.

При подаче сигнала с относительного кодера на фазовый модулятор необходимо учитывать, какой модулирующий сигнал (однополярный или двухполярный), требуется для ФМ. В необходимых случаях ставится согласующее устройство.

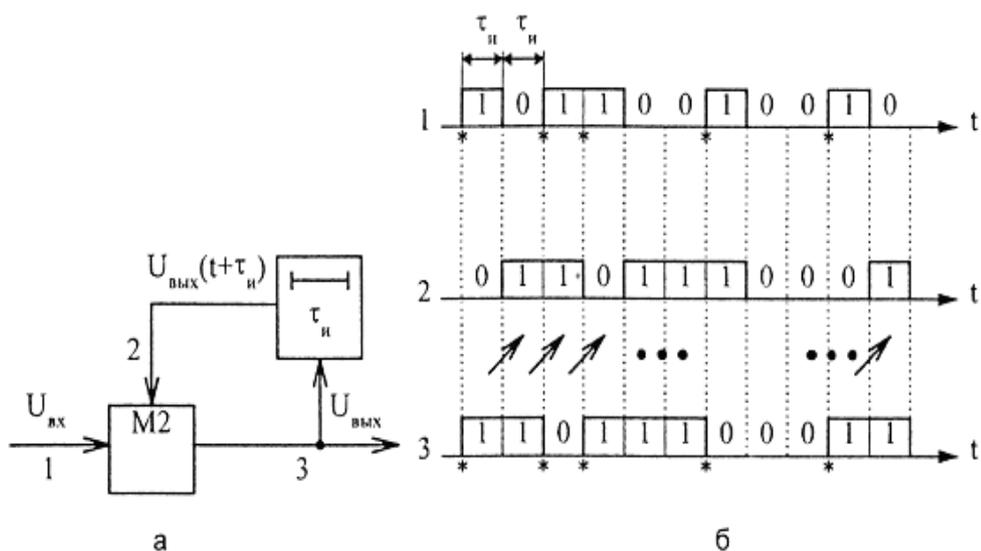


Рисунок 7.22 – Относительный кодер:

- а) функциональная схема;
- б) временные диаграммы, поясняющие работу.

ДЕТЕКТИРОВАНИЕ СИГНАЛОВ

Детектирование – процесс восстановления модулирующего сигнала из модулированного высокочастотного колебания.

Детектирование, при котором учитывается начальная фаза принимаемого модулированного сигнала, называется *когерентным*; не учитывается – *некогерентным*.

Некогерентное детектирование амплитудно-модулированных сигналов

Амплитудный детектор (АД) – устройство, с помощью которого детектируется АМ сигнал.

Поскольку спектр модулированного сигнала содержит только высокочастотные компоненты, и в нем отсутствуют низкочастотные спектральные составляющие модулирующего сигнала, то линейные цепи для детектирования непригодны (они не изменяют спектральный состав сигнала). В большинстве случаев детекторы представляют собой нелинейные цепи.

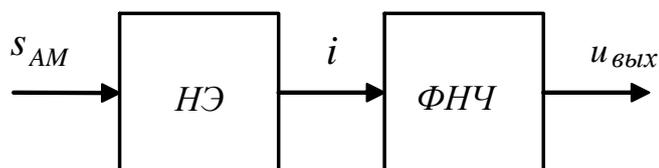


Рисунок 8.1 – Структурная схема некогерентного АД.

Детектор состоит из двух элементов: НЭ и ФНЧ. НЭ преобразует спектр входного модулированного сигнала так, что в спектре выходного тока появляются составляющие модулирующего сигнала. ФНЧ выделяет эти составляющие и предотвращает прохождение на выход детектора высокочастотных составляющих.

Простой и широко распространенной является схема диодного АД.

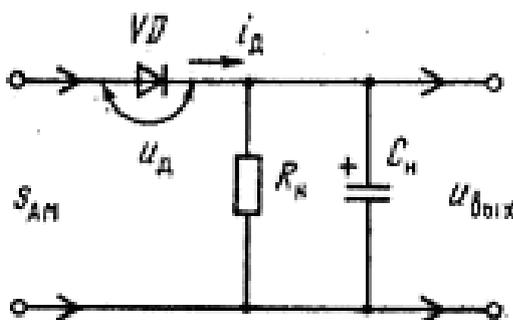


Рисунок 8.2 – Принципиальная схема диодного АД.

В качестве НЭ используется диод. Диод может работать в двух различных режимах: без отсечки тока (на него подается входное воздействие малой амплитуды – не более 100...300 мВ) и с отсечкой тока (на него подается входное воздействие большой амплитуды – более 500...1000 мВ). Детектор, работающий в режиме малого сигнала, называется *квадратичным*, поскольку начальный участок ВАХ диода хорошо аппроксимируется полиномом второй степени. Детектор, работающий в режиме сильного сигнала, называется *линейным*, поскольку основным рабочим участком ВАХ диода является линейный.

Это подтверждается *характеристикой детектирования*, которая представляет собой зависимость постоянного напряжения на нагрузке детектора $U_{0\text{вых}}$ от амплитуды немодулированного высокочастотного сигнала U_m .

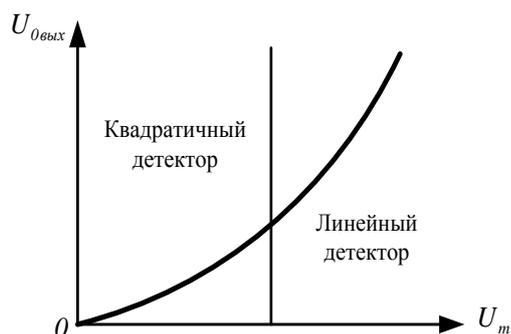


Рисунок 8.3 – Характеристика детектирования диодного АД.

В качестве ФНЧ, который является нагрузкой детектора, обычно служит параллельная цепочка $R_n C_n$, значения элементов которой подбираются так, чтобы сопротивление резистора было значительно больше емкостного сопротивления конденсатора для токов высокой частоты и значительно меньше его для токов низкой частоты:

$$\frac{1}{\omega_n C} \ll R_n \ll \frac{1}{\Omega C}.$$

Тогда выходное напряжение детектора $u_{\text{вых}}(t)$ будет создаваться только низкочастотными составляющими тока, а высокочастотные составляющие окажутся отфильтрованы. Чтобы убрать постоянную составляющую $U_{0\text{вых}}$, содержащуюся в напряжении $u_{\text{вых}}(t)$, сигнал на следующие за детектором элементы подают через разделительный конденсатор C_p . В результате получается напряжение $u'_{\text{вых}}(t)$ без постоянной составляющей.

Квадратичный диодный АД

Характеристику диода можно представить в виде:

$$i = a_0 + a_1 u + a_2 u^2.$$

Подставим в приведенную зависимость выражение модулированного сигнала $s_{AM}(t) = U_m(1 + m \cos \Omega t) \cos \omega_n t$:

$$i = a_0 + a_1 U_m (1 + m \cos \Omega t) \cos \omega_n t + a_2 U_m^2 (1 + m \cos \Omega t)^2 \cos^2 \omega_n t = a_0 + a_1 U_m (1 + m \cos \Omega t) \cos \omega_n t + a_2 U_m^2 (1 + 2m \cos \Omega t + m^2 / 2 + m^2 / 2 \cos(2\Omega t))(1/2 + 1/2 \cos(2\omega_n t)) = a_0 + a_1 U_m (1 + m \cos \Omega t) \times \cos \omega_n t + a_2 U_m^2 (1/2 + 1/2 \cos(2\omega_n t) + m \cos \Omega t + m \cos(2\omega_n t) \cos \Omega t + m^2 / 4 + m^2 / 4 \cos(2\omega_n t) + m^2 / 4 \cos(2\Omega t) + m^2 / 4 \cos(2\omega_n t) \cos(2\Omega t)).$$

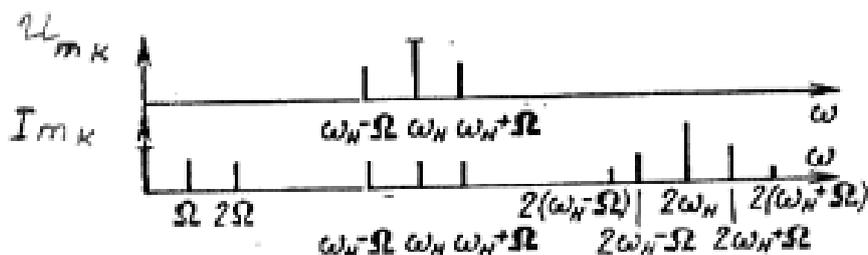


Рисунок 8.4 – Спектр AM сигнала на входе детектора и тока диода.

Низкочастотная переменная составляющая тока диода содержит два слагаемых: полезное, воспроизводящее передаваемый сигнал (с частотой Ω), и вредное, которое появилось вследствие квадратичного характера детектирования (с частотой 2Ω). Следовательно, возникают нелинейные искажения выходного сигнала. При детектировании сигнала, модулированного спектром частот, в детекторе также возникают комбинационные частоты, что еще больше увеличивает искажения сигнала.

Квадратичное детектирование находит ограниченное применение (в измерительной технике).

Представим ВАХ диода в виде линейно-ломаной. Под действием входного напряжения ток в цепи диода представляет собой импульсы, следующие с частотой несущей ω_n . До момента времени t_1 модуляция отсутствует, и амплитуда импульсов не меняется. В момент t_1 включена модуляция, и амплитуда импульсов начинает изменяться с частотой Ω . Последовательность немодулированных импульсов тока через диод может быть представлена в виде ряда Фурье с постоянной составляющей. Модель тока с момента t_1 будет представлять собой произведение ряда Фурье и множителя модуляции $(1 + m \cos \Omega t)$:

$$i_d = (1 + m \cos \Omega t)(I_0 + I_{m1} \cos \omega_n t + I_{m2} \cos 2\omega_n t + \dots).$$

В спектре такого тока будет постоянная составляющая, составляющая несущей частоты ω_n , ее гармоники и комбинационные частоты. В составе спектра появляется низкочастотная составляющая $mI_0 \cos \Omega t$, которую и выделяет ФНЧ.

При линейном детектировании нелинейные искажения информационного сигнала минимальны, а коэффициент передачи детектора K_d , равный отношению постоянной составляющей выходного напряжения $U_{0вых}$ к амплитуде немодулированного несущего колебания U_m , $K_d = U_{0вых} / U_m$ не зависит от амплитуды несущей:

$$K_d = \cos \theta,$$

где θ - угол отсечки тока диода.

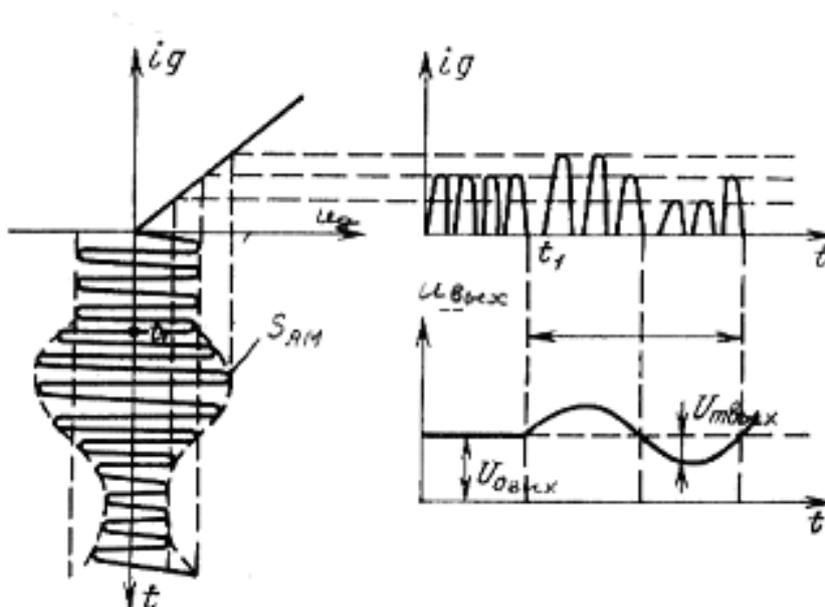


Рисунок 8.5 – Временные диаграммы работы линейного диодного АД.

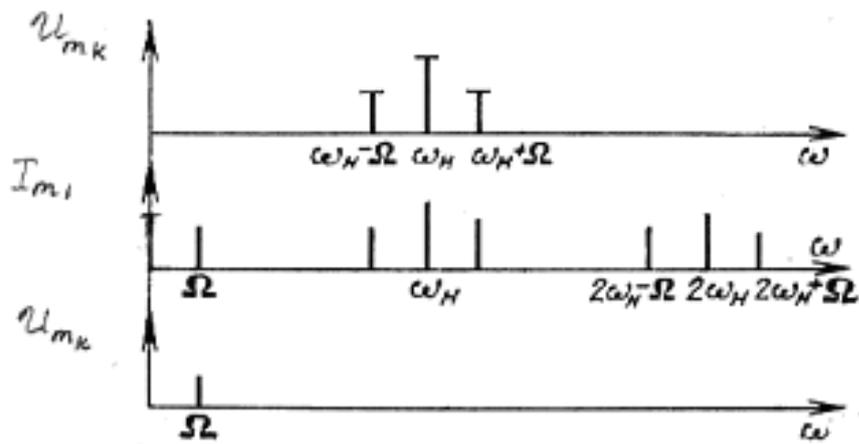


Рисунок 8.6 – Спектральные диаграммы АМ сигнала на входе детектора, тока диода и выходного сигнала.

Линейный АД применяется как в профессиональной, так и в бытовой радиоаппаратуре.

Когерентное детектирование фазомодулированных сигналов

Фазовый детектор (ФД) – устройство, в котором входной ФМ сигнал преобразуется в выходное напряжение (или ток), меняющееся по закону модуляции фазы.

Все ФД являются когерентными (требуют обязательного наличия синхронного с частотой сигнала опорного колебания), т.к. выявить фазовый сдвиг в ФМ сигнале можно только путем его сравнения с немодулированным колебанием (опорным). Начальные фазы ФМ сигнала и опорного колебания в общем случае могут и не совпадать. Структурная схема ФД аналогична схеме СД.

Однотактный диодный ФД

Представляет собой обычный диодный АД, на который подается сумма напряжений опорного колебания и сигнала.

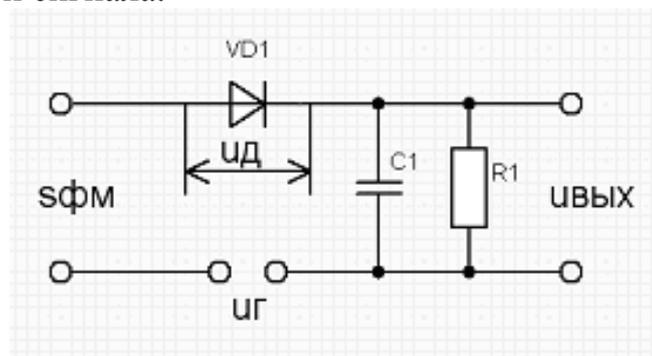


Рисунок 8.7 – Принципиальная схема однотактного диодного ФД.

Пусть $u_2 = U_{m2} \cos \omega_n t$ - опорное напряжение, $s_{ФМ} = U_m \cos(\omega_n t + \varphi(t))$ - детектируемое напряжение, тогда напряжение на диоде: $u_d = u_2 + s_{ФМ}$.

Проведем векторное сложение двух напряжений u_2 и $s_{ФМ}$.

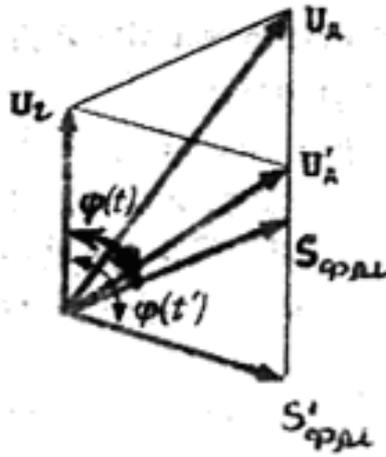


Рисунок 8.8 – Векторная диаграмма.

В результате получаем суммарное напряжение u_0 той же частоты, но другой фазы. Его амплитуда зависит от фазы входного напряжения (т.е. происходит преобразование ФМ в АМ):

$$U_{m0} = \sqrt{U_{m2}^2 + U_m^2 + 2U_{m2}U_m \cos \varphi}.$$

Полученный АМ сигнал детектируется АД. Выходное напряжение ФД:

$$u_{\text{вых}} = K_0 U_{m0},$$

где K_0 - коэффициент передачи АД.

Детекторная характеристика ФД – зависимость его выходного напряжения $U_{0\text{вых}}$ от разности фаз входного и опорного сигнала φ .

Детекторная характеристика рассмотренного ФД показана на рисунке 10.16 (кривая 1).

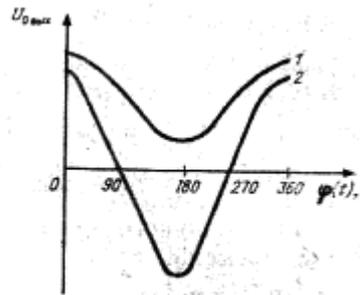


Рисунок 8.9 – Детекторные характеристики ФД:
1 – однотоктного; 2 – балансного.

Выводы:

- характеристика имеет малый линейный участок вблизи углов $\varphi = (2n + 1)90^\circ$, поэтому детектирование сопровождается значительными искажениями;
- выходное напряжение не меняет знака при изменениях фазы.

Балансный диодный ФД

Представляет собой два встречно включенных АД, на которые подаются сумма и разность напряжений опорного колебания и сигнала.

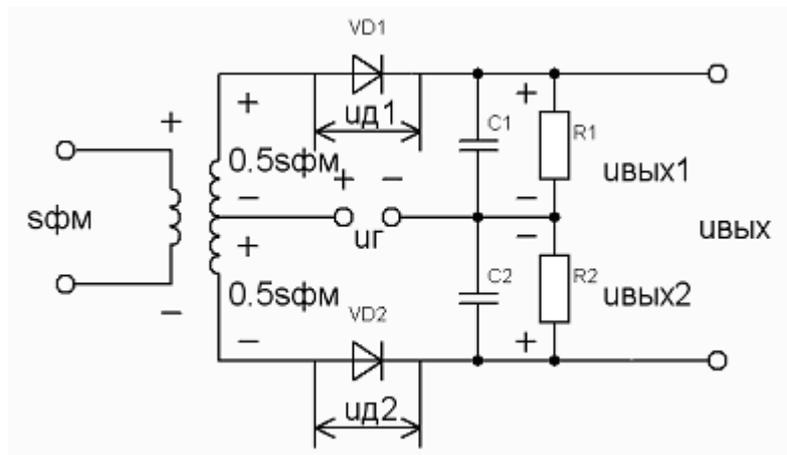


Рисунок 8.10 – Принципиальная схема балансного диодного ФД.

При выбранной полярности напряжений u_c и $s_{\phi M}$ напряжения на диодах:

$$u_{\partial 1} = u_c + 0,5s_{\phi M},$$

$$u_{\partial 2} = u_c - 0,5s_{\phi M}.$$

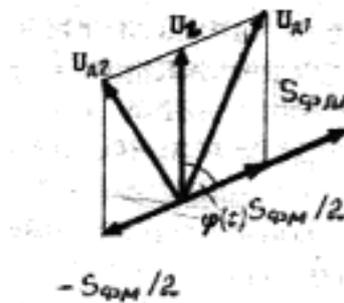


Рисунок 8.11 – Векторная диаграмма.

Пользуясь векторной диаграммой, можно определить амплитуды напряжений на диодах:

$$U_{m\partial 1} = \sqrt{U_{m_c}^2 + U_m^2 + 2U_{m_c}U_m \cos \varphi},$$

$$U_{m\partial 2} = \sqrt{U_{m_c}^2 + U_m^2 - 2U_{m_c}U_m \cos \varphi}.$$

Токи i_1 и i_2 образуют выходные напряжения плеч противоположной полярности:

$$u_{\text{вых}1} = K_{\partial} U_{m\partial 1},$$

$$u_{\text{вых}2} = K_{\partial} U_{m\partial 2}.$$

Поэтому выходное напряжение равно их разности:

$$u_{\text{вых}} = u_{\text{вых}1} - u_{\text{вых}2}.$$

Детекторная характеристика рассмотренного ФД показана на рисунке 23.3 (кривая 2).

Выводы:

- наилучшую линейность характеристика имеет вблизи углов $\varphi = (2n+1)90^\circ$, при работе на этих участках детектирование осуществляется практически без искажений;
- знак выходного напряжения меняется при изменении фазы;
- крутизна характеристики балансного ФД в два раза больше крутизны характеристики однотактного ФД.

Детектирование частотно-модулированных сигналов

Частотный детектор (ЧД) – устройство, напряжение на выходе которого зависит от частоты входного сигнала.

Для восстановления модулирующего сигнала из ЧМ сигнала только нелинейного устройства недостаточно, т.к. в реакции любого НЭ на ЧМ сигнал имеются только модулированные гармоники частоты несущей и нет низкочастотных составляющих.

Для доказательства рассмотрим нелинейное преобразование ЧМ сигнала. Пусть ВАХ НЭ аппроксимирована полиномом:

$$i = a_0 + a_1 u + a_2 u^2,$$

где $u = s_{чМ}$.

Оказывается, что ток НЭ не содержит информационной низкочастотной компоненты:

$$\begin{aligned} i(t) &= a_0 + a_1 U_m \cos(\omega_n t + M_{чМ} \sin \Omega t) + a_2 U_m^2 \cos^2(\omega_n t + M_{чМ} \sin \Omega t) = \\ &= a_0 + \frac{1}{2} a_2 U_m^2 + a_1 U_m \cos(\omega_n t + M_{чМ} \sin \Omega t) + \frac{1}{2} U_m^2 \cos 2(\omega_n t + M_{чМ} \sin \Omega t). \end{aligned}$$

Следовательно, требуется дополнительное преобразование ЧМ сигнала. ЧД работают по принципу преобразования ЧМ в другой вид модуляции с последующим детектированием преобразованного вида модуляции. В зависимости от характера преобразований ЧМ различают частотно-амплитудные, частотно-фазовые и частотно-импульсные детекторы.

В частотно-амплитудных детекторах изменение частоты сигнала преобразуется в изменение амплитуды, которое выделяется АД. В частотно-фазовых детекторах изменение частоты преобразуется в изменение фазового сдвига между двумя напряжениями с последующим фазовым детектированием. В частотно-импульсных детекторах ЧМ сигнал преобразуется в один из видов импульсной модуляции с последующим детектированием с помощью ФНЧ или счетной схемы.

Частотно-амплитудные детекторы

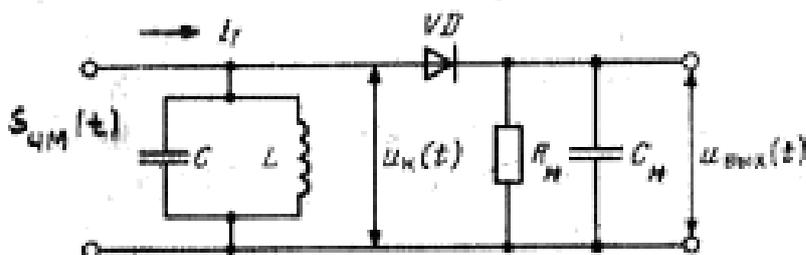


Рисунок 8.12 – Принципиальная схема ЧД с расстроенным колебательным контуром.

Преобразование ЧМ сигнала осуществляется в колебательном контуре, расстроенном относительно несущей частоты сигнала ($\omega_{рез} \neq \omega_n$). Величину расстройки контура выбирают больше девиации частоты: $\Delta\omega > \Delta\omega_m$. При изменении частоты принимаемого сигнала $\omega(t) = \omega_n + \Delta\omega(t)$ изменяется коэффициент передачи контура по напряжению K , в результате изменяется амплитуда напряжения на контуре $U_{нк}(t)$ и ЧМ коле-

бание превращается в амплитудно-частотно-модулированное (АЧМ) $u_{\kappa}(t)$. Напряжение с контура подается на линейный АД, на выходе которого получается низкочастотный сигнал $u_{\text{вых}}(t)$.

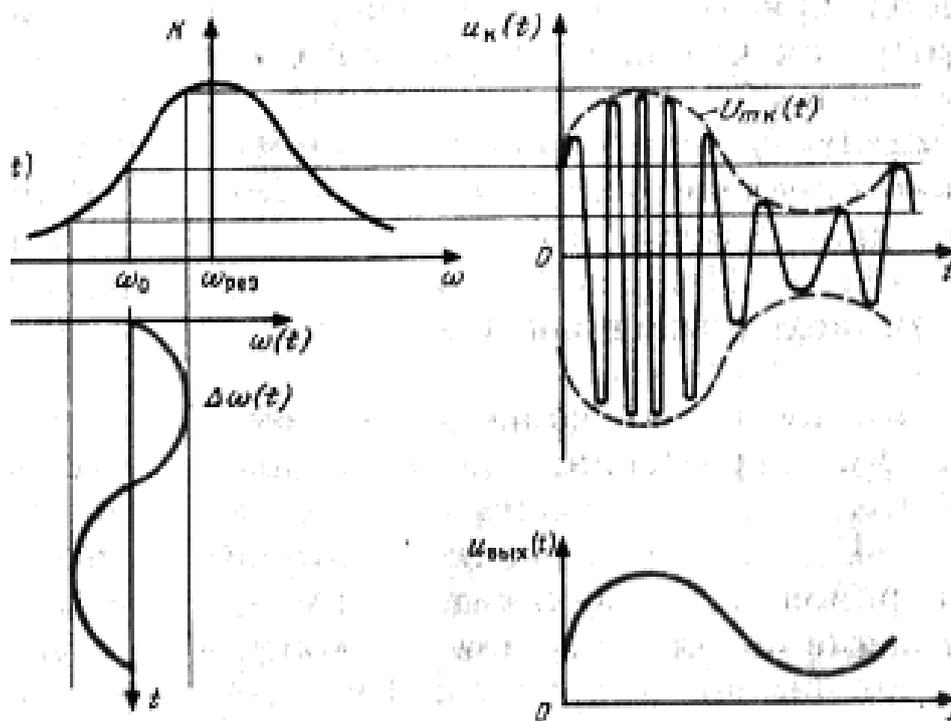


Рисунок 8.13 – Временные диаграммы работы ЧД, если рабочая точка выбрана на левой ветви резонансной характеристики контура ($\omega_{\text{рез}} > \omega_n$).

Детекторная характеристика рассмотренного ЧД, представляющая собой зависимость постоянного напряжения на выходе детектора $U_{\text{вых}}$ от частоты f немодулированного высокочастотного сигнала на его входе, представлена на рисунке 10.10.

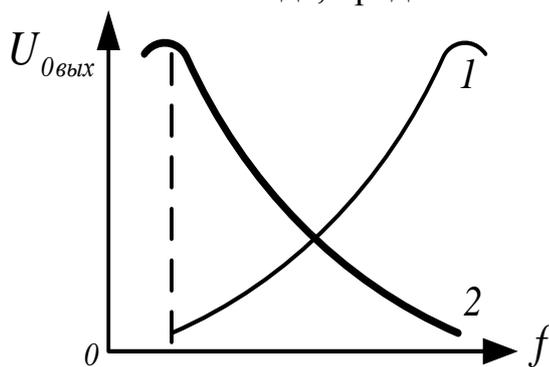


Рисунок 8.14 – Детекторная характеристика рассмотренного ЧД: 1 - $\omega_{\text{рез}} > \omega_n$; 2 - $\omega_{\text{рез}} < \omega_n$.

Недостаток:

- детекторная характеристика имеет значительную нелинейность (т.к. одиночный колебательный контур имеет незначительный линейный участок резонансной характеристики), т.е. характерны значительные несимметричные нелинейные искажения исходного сигнала.

Уменьшения нелинейных искажений добиваются уменьшением добротности контура или применением более сложных схем детекторов.

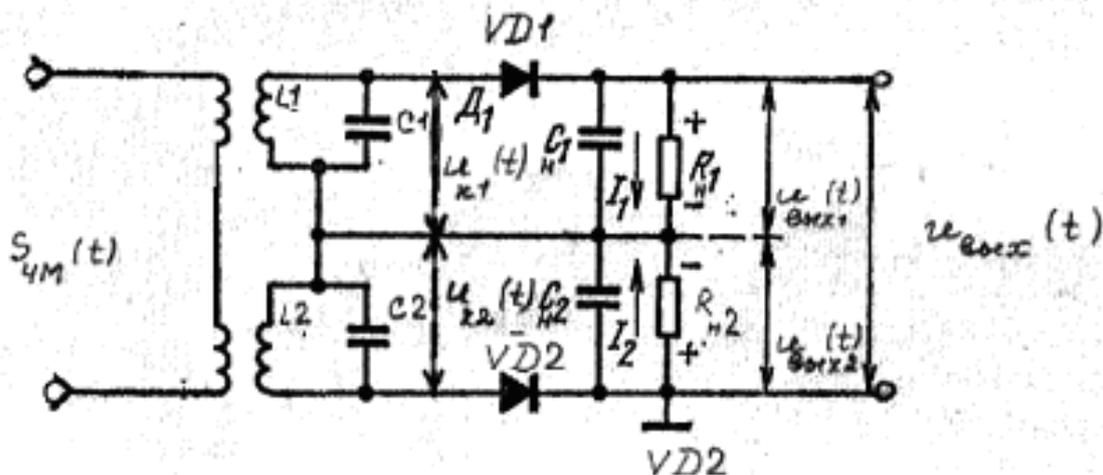


Рисунок 8.15 – Принципиальная схема ЧД с двумя расстроенными контурами.

Такой ЧД представляет собой два ЧД с одиночными контурами, расстроенных симметрично относительно несущей частоты входного сигнала. $L1C1$ -контур первого ЧД настроен на частоту $\omega_{рез1} > \omega_n$, а $L2C2$ -контур – на частоту $\omega_{рез2} < \omega_n$.

В процессе работы ЧД можно выделить три основных случая:

- частота детектируемого сигнала равна несущей частоте $\omega = \omega_n$. В этом случае амплитуды напряжений на контурах одинаковы $U_{mk1} = U_{mk2}$, при этом токи через диоды также равны $i_1 = i_2$ и создают на резисторах R_{n1} и R_{n2} одинаковые напряжения $u_{вых1} = u_{вых2}$, а т.к. токи имеют встречное направление, то напряжение на выходе детектора, равно разности этих напряжений $u_{вых} = u_{вых1} - u_{вых2} = 0$;
- частота детектируемого сигнала выше несущей частоты $\omega > \omega_n$. В этом случае: $U_{mk1} > U_{mk2}$, $i_1 > i_2$, $u_{вых1} > u_{вых2}$, $u_{вых} = u_{вых1} - u_{вых2} > 0$;
- частота детектируемого сигнала ниже несущей частоты $\omega < \omega_n$. В этом случае: $U_{mk1} < U_{mk2}$, $i_1 < i_2$, $u_{вых1} < u_{вых2}$, $u_{вых} = u_{вых1} - u_{вых2} < 0$.

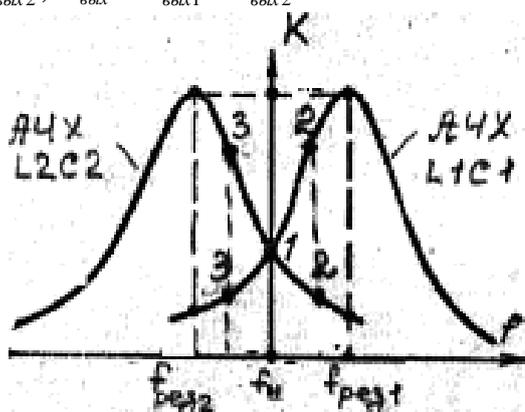


Рисунок 8.16 – Резонансные кривые расстроенных контуров.

Детекторная характеристика рассмотренного ЧД представлена на рисунке 10.13.

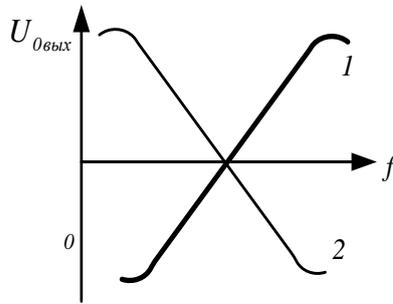


Рисунок 8.17 - Детекторная характеристика рассмотренного ЧД : 1 – частоты настройки первого и второго контуров: $\omega_{рез1} > \omega_n$ и $\omega_{рез2} < \omega_n$; 2 – частоты настройки первого и второго контуров: $\omega_{рез1} < \omega_n$ и $\omega_{рез2} > \omega_n$.

Достоинства:

- детекторная характеристика имеет значительный линейный участок, благодаря чему детектирование происходит без искажений.

Недостаток:

- сложность в изготовлении и настройке.

ОСНОВЫ ТЕОРИИ КОДИРОВАНИЯ СИГНАЛОВ

Кодирование дискретных сообщений

Кодирование – преобразование элементов дискретного сообщения в последовательности кодовых символов. Обратное преобразование – *декодирование*.

Устройства, осуществляющие эти операции автоматически, называются соответственно *кодером* и *декодером*. *Кодек* – устройство, объединяющее кодер и декодер.

Код – алгоритм (правило), по которому осуществляется кодирование.

Кодовая комбинация (слово) – последовательность кодовых символов, соответствующая одному элементу дискретного сообщения.

Кодовый алфавит – весь набор кодовых символов.

Основание кода m – число символов в кодовом алфавите. Если $m=2$ код называется *двоичным*, $m>2$ – *многопозиционным (недвоичным)*.

Разряд – значащая позиция кодового слова.

Разрядность (значность) кода n – число символов в кодовой комбинации. Если $n=const$, то код называется *равномерным*, $n \neq const$ – *неравномерным*.

Кодеры и декодеры легче сделать для равномерных двоичных кодов.

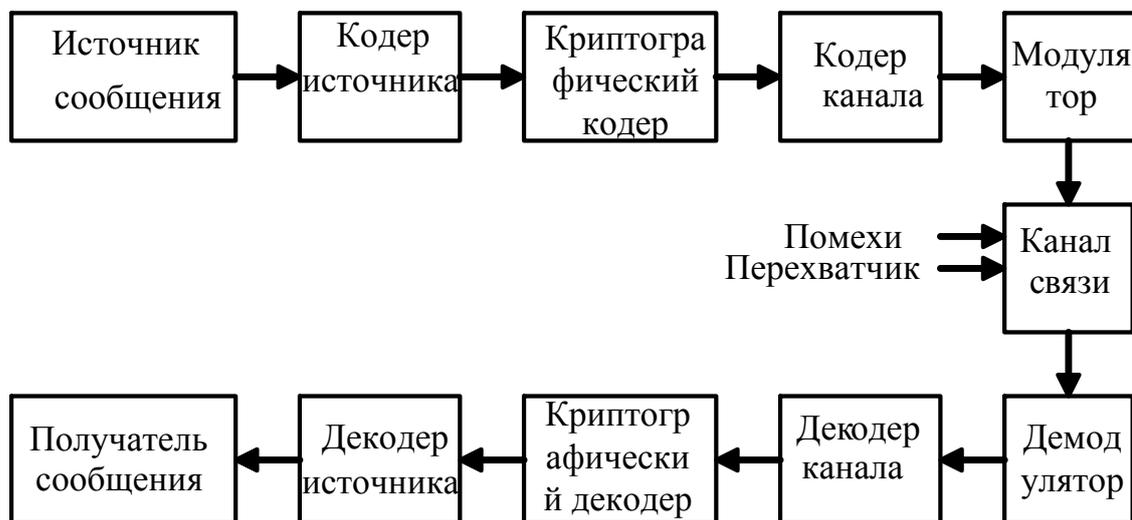


Рисунок 9.1 – Структурная схема системы передачи дискретных сообщений.

Источник выдает дискретное сообщение. Для формирования дискретного сообщения из непрерывного используется дискретизация по времени и по уровню.

Кодирование источника (сжатие данных) применяется для снижения технических затрат на хранение и передачу информации.

Криптографическое кодирование (шифрование) применяется для предотвращения несанкционированного доступа к информации.

Кодирование канала (помехоустойчивое кодирование) применяется для повышения достоверности передачи информации по каналу с помехами.

Помехоустойчивое (корректирующее) кодирование

Помехоустойчивое кодирование используется в различных системах связи, в сетях ЭВМ, в бытовой и профессиональной аудио- и видеотехнике.

Любой помехоустойчивый код является избыточным.

Избыточные коды – коды, в которых для передачи информации используются не все возможные кодовые слова. Используемые кодовые слова называются *разрешенными*. Их число – *мощность кода M*. Неиспользуемые кодовые слова считаются *запрещенными*.

Пример 9.1:

Рассмотрим трехразрядный ($n=3$) двоичный ($m=2$) код.

Безыбыточный код	Избыточный код	
	Разрешенные кодовые слова	Запрещенные кодовые слова
000	000	
001		001
010	010	
011		011
100	100	
101		101
110	110	
111		111
$M_{\text{безызб.}} = m^n = 2^3 = 8$	$M_{\text{изб.}} = 4 < M_{\text{безызб.}}$	

На вход канала поступают только разрешенные кодовые слова. Под воздействием помех возможно искажение кодовых слов. Говорят, что в канале произошла *ошибка кратности t*, если t символов кодового слова принято ошибочно. Для контроля за ошибками используется переход разрешенных кодовых слов в запрещенные.

Пример 9.2:

Безыбыточный код		Избыточный код	
010 Разрешенное слово	→ Канал → $t=1$	010 Разрешенное слово	→ Канал → $t=1$
011* Разрешенное слово		011* Запрещенное слово	
Нет возможности контролировать ошибки.		Есть возможность контролировать ошибки.	

Корректирующая способность кода – способность кода контролировать (обнаруживать и исправлять) ошибки. Она определяется кодовым расстоянием.

Кодовое расстояние – минимальное расстояние Хэмминга для заданного кода:

$$d_0 = \min d_x .$$

Расстояние Хэмминга – степень различия между i -ым и j -ым кодовыми словами:

$$d_x = \text{dist}(\bar{b}_i, \bar{b}_j) .$$

Определяется числом несовпадающих в них разрядов.

Методы декодирования помехоустойчивых кодов:

- декодирование с обнаружением ошибок – обеспечивает стирание или особую отметку той части сообщения, в которой обнаружены ошибки;
- декодирование с исправлением ошибок – позволяет получить верное сообщение, несмотря на наличие искажений отдельных символов.

Существует также смешанное декодирование. Один и тот же код можно использовать при различных методах декодирования.

Правила для определения корректирующей способности кодов:

1. Для обнаружения ошибок кратности t_o кодовое расстояние должно удовлетворять соотношению:

$$d_0 \geq t_o + 1.$$

2. Для исправления ошибок кратности t_u и одновременного обнаружения ошибок кратности t_o ($t_o \geq t_u$) кодовое расстояние должно удовлетворять соотношению:

$$d_0 \geq t_o + t_u + 1.$$

Задача 7.1:

Определить корректирующую способность кода, имеющего $d_0 = 3$. То же для кода с $d_0 = 4$.

Решение:

Формулы для определения обнаруживающей и исправляющей способности кода:

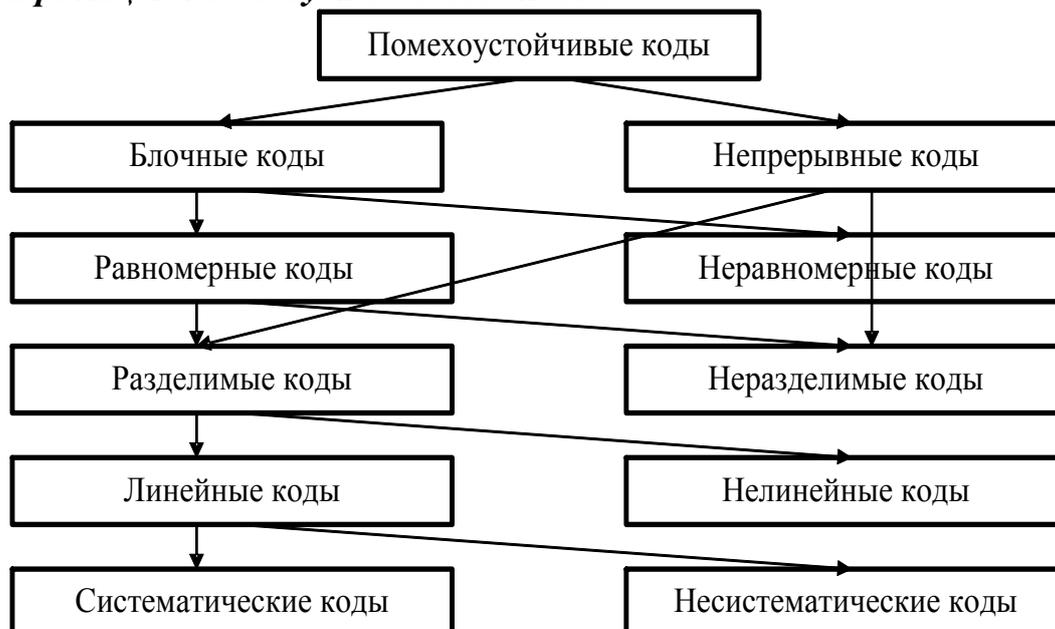
$$t_o \leq d_0 - 1,$$

$$t_u \leq (d_0 - 1(2))/2.$$

Для кода с $d_0 = 3$ $t_o = 2$ (из правила 1: $3 = 2 + 1$) или $t_u = 1$ и $t_o = 1$ (из правила 2: $3 = 1 + 1 + 1$).

Для кода с $d_0 = 4$ $t_o = 3$ (из правила 1: $4 = 3 + 1$) или $t_u = 1$ и $t_o = 2$ (из правила 2: $4 = 1 + 2 + 1$).

Классификация помехоустойчивых кодов



По способу кодирования различают блочные (каждый блок из n символов на выходе кодера зависит только от текущего блока из k символов на его входе и не зависит

от предыдущих входных блоков) и непрерывные коды (каждый набор из n выходных символов зависит не только от текущего набора из k входных символов, но и от некоторого числа предыдущих входных наборов). Обозначение блочных кодов:

$$(n, k).$$

Для коррекции ошибок неравномерные коды почти не применяют.

По структуре кодовых последовательностей различают делимые (кодированные символы можно разделить на информационные и проверочные (контрольные)) и неделимые коды (такое разделение провести нельзя).

По алгоритму формирования проверочных символов различают линейные (проверочные символы формируются путем суммирования по модулю два информационных символов) и нелинейные коды (используется суммирование по модулю отличному от два).

По способу передачи кодовых символов различают систематические (в канал связи первоначально передаются информационные символы, а затем – проверочные) и не систематические коды (в канал связи кодовые символы передаются по «псевдослучайному» закону).

Код с постоянным весом

Это неделимый блочный код, каждая кодовая комбинация которого имеет одинаковое число единиц (одинаковый вес).

Если вес принятой кодовой комбинации отличается от заданного, то выносятся решение об ошибке. Данный код обладает $d_0 = 2$ и обнаруживает все ошибки нечетной кратности и часть ошибок четной кратности (не обнаруживаются только ошибки смещения, когда число искаженных единиц равно числу искаженных нулей).

Таким кодом является код МТК-3 – семиразрядный код, каждая кодовая комбинация которого содержит три единицы.

Групповые систематические линейные блочные коды

Линейные коды – коды, для которых поразрядная сумма по модулю два любых разрешенных кодовых комбинаций также является разрешенной кодовой комбинацией. Линейные коды называют также *групповыми*.

Они задаются с помощью порождающей G и проверочной H матриц, которые связаны основным уравнением кодирования:

$$G \cdot H^T = 0,$$

где H^T - транспонированная проверочная матрица (строки H переписаны в столбцы H^T);

0 - нулевая матрица.

Матрица G содержит k строк и n столбцов, ее элементами являются нули и единицы. Строками матрицы G являются любые ненулевые линейно независимые векторы, отстоящие друг от друга не менее, чем на заданное кодовое расстояние. Понятие линейно независимые означает, что каким бы образом мы не суммировали по модулю два различные строки матрицы, мы не получим суммы, равной нулю.

С помощью матрицы G можно создавать линейный код: суммируя в различном сочетании строки матрицы G , получают все (кроме нулевой) комбинации кода. Полученный код содержит $M = 2^k$ кодовых слов длины n .

Если две порождающие матрицы различаются только порядком расположения столбцов, то определяемые ими коды называются *эквивалентными*. Они имеют одина-

ковые кодовые расстояния и, следовательно, одинаковые способности обнаруживать и исправлять ошибки.

Пример 9.3:

Код Рида-Маллера (8, 4) задается следующей порождающей матрицей:

$$G = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{vmatrix}.$$

Матрица H содержит $r = n - k$ строк и n столбцов. Единицы в каждой строке этой матрицы показывают, какие символы кодовой комбинации нужно сложить по модулю два, чтобы получить нуль. Используется для проверки правильности приема.

Чаще всего применяют *систематические* линейные коды. Такие коды задаются матрицами в систематической (приведено-ступенчатой или канонической) форме:

$$G = |I_k \quad G^*|,$$

$$H = |H^* \quad I_r|$$

где I_k, I_r - единичные подматрицы размерностью $k \times k$ и $r \times r$ соответственно;

$G^* = H^{*T}$ - прямоугольная подматрица размерностью $k \times r$;

$H^* = G^{*T}$ - прямоугольная подматрица размерностью $r \times k$.

Пример 9.4:

Систематический код Рида-Маллера (8, 4) задается порождающей матрицей:

$$G = |I_4 \quad G^*| = \begin{vmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{vmatrix}.$$

Найдем проверочную матрицу:

$$H = |G^{*T} \quad I_4| = \begin{vmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{vmatrix}.$$

Кодирование информации

1) С помощью матрицы G : операция кодирования заключается в умножении информационного вектора \bar{a} на порождающую матрицу G , т.е.

$$\bar{b} = \bar{a} \cdot G,$$

где \bar{b} - кодовый вектор.

Пример 9.5:

Рассмотрим кодирование информационного слова $\bar{a} = (1011)$ кодом Рида-Маллера (8, 4):

$$\bar{b} = (1011) \cdot \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{vmatrix} = (11000011).$$

2) С помощью матрицы H : в СЛБК информационные символы слова \bar{a} входят без изменения в кодовое слово \bar{b} и занимают в нем первые k позиций, к ним добавляются $r = n - k$ проверочных символов, правила формирования которых задает проверочная матрица.

Единицы в i -ой строке подматрицы H^* указывают, какие информационные символы необходимо просуммировать по модулю два, чтобы получить i -ый проверочный.

Пример 9.6:

Из матрицы H систематического кода Рида-Маллера запишем правила формирования проверочных символов:

$$\begin{aligned} b_5 &= a_1 \oplus a_2 \oplus a_3, \\ b_6 &= a_1 \oplus a_2 \oplus a_4, \\ b_7 &= a_1 \oplus a_3 \oplus a_4, \\ b_8 &= a_2 \oplus a_3 \oplus a_4. \end{aligned}$$

Тогда для $\bar{a} = (1011)$ $\bar{b} = (1011 \ 0010)$.

Коды Хэмминга

Это линейные блочные коды с параметрами $(n, k) = (2^m - 1, 2^m - 1 - m)$, где m - положительное целое число, $r = m$ - число проверочных символов. Для задания кодов Хэмминга обычно используется проверочная матрица. Ее столбцами являются все ненулевые двоичные числа длиной m .

Они обладают кодовым расстоянием $d_0 = 3$ и способны исправлять только одну или обнаруживать две ошибки.

Примеры полных кодов Хэмминга: $(7, 4)$, $(15, 11)$, $(31, 26)$, $(63, 57)$.

Пример 9.7:

Рассмотрим код Хэмминга $(7, 4)$. Проверочная матрица:

$$H = \begin{vmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{vmatrix} = [\text{Перемещая столбцы, приводим ее к систематической}]$$

$$\text{форме]} = \begin{vmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{vmatrix}.$$

Порождающая матрица:

$$G = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{vmatrix}.$$

Модификациями кодов Хэмминга являются укороченные и удлиненные коды Хэмминга.

Чтобы получить проверочную матрицу укороченного кода Хэмминга, необходимо в проверочной матрице полного кода исключить любые T столбцов, относящиеся к информационным разрядам, где T - параметр укорочения.

Удлиненные коды Хэмминга получают путем введения дополнительной проверки на четность всех символов кодового слова.

Коды Хэмминга обладают очень слабой корректирующей способностью и отдельно практически не используются. Очень хорошие результаты позволяет получить применение данных кодов в составе каскадных схем кодирования. Каскадные коды состоят из двух или более кодов: кодовые слова одного кода являются информационными символами для кода следующей ступени.

Линейные циклические блочные коды

Поиск более простых процедур кодирования и декодирования привел к появлению циклических кодов.

Циклические коды – линейные блочные коды, обладающие *свойством цикличности*: если $\bar{b} = (b_{n-1}b_{n-2}\dots b_1b_0)$ - кодовое слово циклического кода, то его циклическая перестановка $\bar{b}' = (b_{n-2}\dots b_1b_0b_{n-1})$ также является кодовым словом.

Пример 9.8:

$$\bar{b} = (1001110) \rightarrow \bar{b}' = (0011101) \rightarrow \bar{b}'' = (0111010).$$

Для построения кода достаточно задать одно кодовое слово. Другие кодовые слова образуются из исходного путем циклических перестановок и их линейных преобразований.

Все преобразования кодовых слов циклических кодов производятся в виде математических операций над полиномами (многочленами). Для этого кодовые слова представляются в форме полиномов:

$$b(x) = b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \dots + b_1x + b_0,$$

где $b_i = \{0,1\}$ - коэффициенты полинома;

x - символическая переменная.

Пример 9.9:

$$\bar{b} = (1001110) \rightarrow b(x) = 1 \cdot x^6 + 0 \cdot x^5 + 0 \cdot x^4 + 1 \cdot x^3 + 1 \cdot x^2 + 1 \cdot x^1 + 0 \cdot x^0 = x^6 + x^3 + x^2 + x.$$

Операции сложения, вычитания, умножения и деления полиномов выполняются по обычным арифметическим правилам, только вычитание заменяется сложением, которое производится как сложение по модулю два.

Циклические коды задаются с помощью порождающего (образующего) $g(x)$ и проверочного $h(x)$ полиномов.

Любой полином $g(x)$ степени $r = n - k$, который делит без остатка полином вида $x^n + 1$, называется *порождающим полиномом*:

$$g(x) = g_r x^r + g_{r-1} x^{r-1} + \dots + g_1 x + g_0,$$

где $g_i = \{0,1\}$ - коэффициенты полинома.

Полиномы всех кодовых слов делятся без остатка на порождающий полином.

Порождающая матрица строится на основе полинома $g(x)$.

Для несистематического циклического кода:

$$G = \begin{pmatrix} 0 & \dots & 0 & g_r & g_{r-1} & \dots & g_1 & g_0 \\ \dots & 0 & g_r & g_{r-1} & \dots & g_1 & g_0 & 0 \\ \dots & \dots \\ g_r & g_{r-1} & \dots & g_1 & g_0 & 0 & \dots & 0 \end{pmatrix}.$$

Для систематического циклического кода:

$$G = [I_k R],$$

$$H = \begin{vmatrix} h_0 & h_1 & \dots & h_{k-1} & h_k & 0 & \dots & 0 \\ 0 & h_0 & h_1 & \dots & h_{k-1} & h_k & 0 & \dots \\ \dots & \dots \\ 0 & \dots & 0 & h_0 & h_1 & \dots & h_{k-1} & h_k \end{vmatrix}.$$

Для систематического циклического кода:

$$H = |R^T I_r|.$$

Кодирование информации

Существует два способа кодирования:

- *несистематическое кодирование*:

$$b(x) = a(x)g(x),$$

где $a(x)$ - полином информационного слова,

$b(x)$ - полином кодового слова;

- *систематическое кодирование*:

$$b(x) = a(x)x^r + r(x),$$

где $r(x)$ - остаток от деления произведения $a(x)x^r$ на полином $g(x)$.

Пример 9.11:

Закодировать слово $\bar{a} = (010)$ циклическим кодом из примера 4.3.

Несистематическое кодирование:

$$b(x) = a(x)g(x) = x(x^4 + x^3 + x^2 + 1) = x^5 + x^4 + x^3 + x \rightarrow \bar{b} = (0111010).$$

Систематическое кодирование:

$$1) a(x)x^r = x \cdot x^4 = x^5;$$

$$a(x)x^r / g(x) = x^5 \quad \left| \begin{array}{l} x^4 + x^3 + x^2 + 1 \end{array} \right.$$

$$\underline{x^5 + x^4 + x^3 + x} \quad x+1$$

$$2) \quad \frac{x^4 + x^3 + x}{x^4 + x^3 + x^2 + 1} \quad ;$$

$$\underline{x^4 + x^3 + x^2 + 1}$$

$$x^2 + x + 1 = r(x)$$

$$3) b(x) = a(x)x^r + r(x) = x^5 + x^2 + x + 1 \rightarrow \bar{b} = (0100111).$$

Декодирование помехоустойчивых кодов

Существует три основных метода декодирования линейных кодов:

- декодирование по максимуму правдоподобия (по минимуму расстояния);
- мажоритарное декодирование (по большинству проверок);
- декодирование по синдрому.

Декодирование по максимуму правдоподобия

Правило декодирования:

В качестве переданного слова \bar{b} следует выбирать слово, которое ближе всего по Хэммингу к принятому \bar{b}^* .

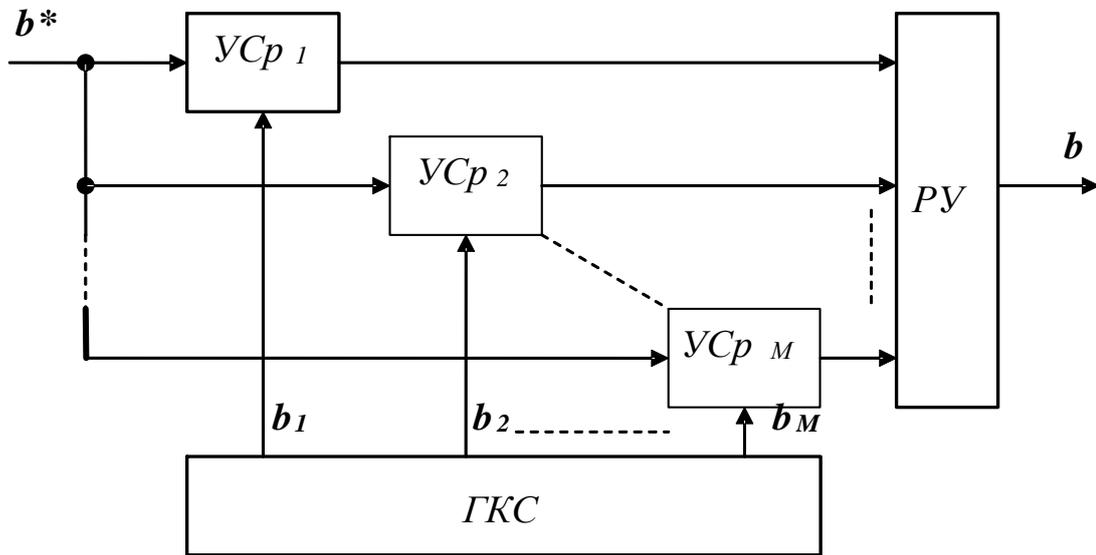


Рисунок 9.2 – Структурная схема декодера по минимуму расстояния.

На рисунке: УСр – устройство сравнения; ГКС – генератор кодовых слов; РУ – решающее устройство.

Данный метод используется, когда число информационных символов k мало ($r \gg k$).

Мажоритарное декодирование

Основано на том, что каждый информационный символ можно выразить через другие символы кодового слова с помощью линейных соотношений. Окончательное решение о значении символа принимается по мажоритарному принципу (по большинству) результатов таких проверок.

Существует три способа построения систем проверочных уравнений для декодирования символа:

- системы с разделенными проверками – символ, относительно которого разделяется система, входит во все уравнения. Любой другой символ входит не более, чем в одно уравнение. Для коррекции t ошибок необходимо $2t + 1$ уравнений в системе;
- системы с λ -связанными проверками – символ, относительно которого разрешается система, входит во все уравнения. Любой другой символ входит не более, чем в λ уравнений. Для коррекции t ошибок необходимо $2\lambda t + 1$ уравнений в системе;
- системы с квазиразделенными проверками – система делима относительно некоторой суммы символов. На первом этапе она разрешается относительно суммы символов, а на втором – относительно конкретного символа.

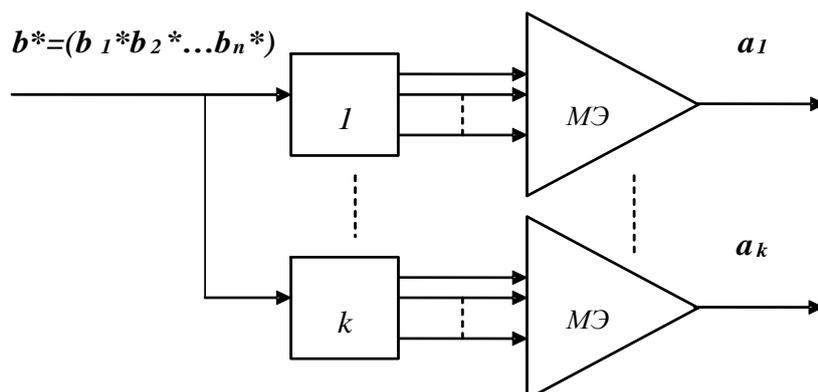


Рисунок 9.3 – Структурная схема мажоритарного декодера.

На рисунке: $1 \dots k$ – устройства, реализующие проверки для соответствующей системы; МЭ – мажоритарный элемент, принимающий решение о значении символа по большинству результатов проверок.

Пример 9.12:

Код (8,4) задан матрицей:

$$H = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Система уравнений по матрице H:

$$\begin{cases} b_1 + b_2 + b_5 = 0, \\ b_2 + b_3 + b_6 = 0, \\ b_3 + b_4 + b_7 = 0, \\ b_1 + b_4 + b_8 = 0. \end{cases}$$

Система проверочных уравнений для a_1 :

$$\begin{cases} a_1 = b_2^* + b_5^*, \\ a_1 = b_4^* + b_8^*, \\ a_1 = b_1^*. \end{cases}$$

Система проверочных уравнений для a_2 :

$$\begin{cases} a_2 = b_1^* + b_5^*, \\ a_2 = b_3^* + b_6^*, \\ a_2 = b_2^*. \end{cases}$$

Система проверочных уравнений для a_3 :

$$\begin{cases} a_3 = b_2^* + b_6^*, \\ a_3 = b_4^* + b_7^*, \\ a_3 = b_3^*. \end{cases}$$

Система проверочных уравнений для a_4 :

$$\begin{cases} a_4 = b_3^* + b_7^*, \\ a_4 = b_1^* + b_8^*, \\ a_4 = b_4^*. \end{cases}$$

Пусть $\bar{b}^* = (10110000)$.

$$\begin{cases} a_1 = 0 + 0 = 0, \\ a_1 = 1 + 0 = 1, \Rightarrow a_1 = 1; \\ a_1 = 1. \end{cases} \begin{cases} a_2 = 1 + 0 = 1, \\ a_2 = 1 + 0 = 1, \Rightarrow a_2 = 1; \\ a_2 = 0. \end{cases} \begin{cases} a_3 = 0 + 0 = 0, \\ a_3 = 1 + 0 = 1, \Rightarrow a_3 = 1; \\ a_3 = 1. \end{cases}$$

$$\begin{cases} a_4 = 1 + 0 = 1, \\ a_4 = 1 + 0 = 1, \Rightarrow a_4 = 1. \\ a_4 = 1. \end{cases}$$

Результат декодирования: $\bar{a} = (a_1 a_2 a_3 a_4) = (1111)$.

Декодирование по синдрому

Основано на стандартной таблице – таблице всех возможных принятых из канала слов, организованной таким образом, что может быть найдено ближайшее к принятому кодовое слово. Она содержит $N = 2^r$ строк и $M + 1 = 2^k + 1$ столбцов.

Таблица – Стандартная таблица.

$s_1=(0\dots 0)$ r	$b_1=(0\dots 0)$ n	b_2	...	b_M
s_2	e_2	b_2+e_2	...	b_M+e_2
...
s_N	e_N	b_2+e_N	...	b_M+e_N

b_i – кодовые слова;

e_j – векторы ошибок – образцы ошибок минимального веса;

b_i+e_j – слова, не являющиеся кодовыми;

$s_i=e_i \cdot H^T$ – синдромы – векторы размерностью r , указывающие на наличие и расположение ошибок в принятом слове.

Правило декодирования:

1. Вычисляется синдром \bar{s} по принятому слову \bar{b}^* :

$$\bar{s} = \bar{b}^* \cdot H^T.$$

Если $\bar{s} = 0$, то \bar{b}^* является кодовым словом. В противном случае ($\bar{s} \neq 0$) \bar{b}^* содержит ошибки.

2. По \bar{s} находится наиболее правдоподобный вектор ошибки \bar{e} .

3. Ближайшее к принятому кодовое слово \bar{b} получается в результате суммирования \bar{b}^* и \bar{e} :

$$\bar{b} = \bar{b}^* + \bar{e}.$$

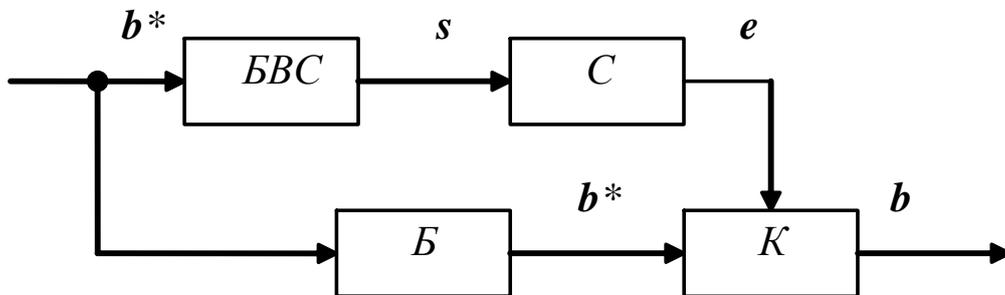


Рисунок 9.4 – Структурная схема декодера по синдрому.

На рисунке: B – буфер хранения принятого слова; $БВС$ – блок вычисления синдрома; C – селектор (дешифратор) синдрома; K – корректор.

Данный метод используется, когда число проверочных символов $r = n - k$ мало (< 10).

Пример 9.13:

Составить стандартную таблицу для систематического кода $(5,2)$ с порождающей матрицей:

$$G = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Таблица должна содержать $N = 2^r = 2^{5-2} = 8$ строк и $M + 1 = 2^k + 1 = 2^2 + 1 = 5$ столбцов.

Таблица – Стандартная таблица.

$b_1=(00000)$	$b_2=(01011)$	$b_3=(10101)$	$b_4=(11110)$	$s_1=(000)$
$e_2=(00001)$	$b_2+e_2=(01010)$	$b_3+e_2=(10100)$	$b_4+e_2=(11111)$	$s_2=e_2 \cdot H^T=(001)$
$e_3=(00010)$	$b_2+e_3=(01001)$	$b_3+e_3=(10111)$	$b_4+e_3=(11100)$	$s_3=e_3 \cdot H^T=(010)$
$e_4=(00100)$	$b_2+e_4=(01111)$	$b_3+e_4=(10001)$	$b_4+e_4=(11010)$	$s_4=e_4 \cdot H^T=(100)$
$e_5=(01000)$	$b_2+e_5=(00011)$	$b_3+e_5=(11101)$	$b_4+e_5=(10110)$	$s_5=e_5 \cdot H^T=(011)$
$e_6=(10000)$	$b_2+e_6=(11011)$	$b_3+e_6=(00101)$	$b_4+e_6=(01110)$	$s_6=e_6 \cdot H^T=(101)$
$e_7=(01100)$	$b_2+e_7=(00111)$	$b_3+e_7=(11001)$	$b_4+e_7=(10010)$	$s_7=e_7 \cdot H^T=(111)$
$e_8=(11000)$	$b_2+e_8=(10011)$	$b_3+e_8=(01101)$	$b_4+e_8=(00110)$	$s_8=e_8 \cdot H^T=(110)$

Пусть $\bar{b}^*=(10111)$. Проведем декодирование.

1. $\bar{s} = \bar{b}^* \cdot H^T = (010) \neq 0$;
2. $\bar{e} = (00010)$;
3. $\bar{b} = \bar{b}^* + \bar{e} = (10101)$.