

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное
образовательное учреждение высшего образования
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

Радиотехнический факультет (РТФ)
Кафедра радиотехнических систем (РТС)

ДЕМОДУЛЯЦИЯ ЦИФРОВЫХ СИГНАЛОВ
СТАТИСТИЧЕСКИЙ И СИГНАЛЬНЫЙ ПОДХОДЫ

Учебное пособие для студентов по направлениям

11.03.01, 11.03.02, 11.04.01 и 11.04.02

и следующим дисциплинам:

Общая теория радиосвязи / Цифровая связь /

Теория и техника передачи информации / Теория радиосвязи

РАЗРАБОТЧИК:

доц. каф. РТС, к.т.н.,

_____ А.В. Новиков

Новиков А.В.

Демодуляция цифровых сигналов. Статистический и сигнальный подходы: учебное пособие для студентов по направлениям **11.03.01**, **11.03.02**, **11.04.01** и **11.04.02** и следующим дисциплинам: Общая теория радиосвязи / Цифровая связь / Теория и техника передачи информации / Теория радиосвязи. — Томск: Радиотехнический факультет, ТУСУР, 2018. — 51 с.

Настоящее пособие включает в себя азбучные принципы демодуляции бинарных цифровых сигналов. Процесс демодуляции рассмотрен с двух точек зрения: статистической и сигнальной.

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ.....	4
СТАТИСТИЧЕСКИЙ ПОДХОД.....	
1 Постановка задачи.....	6
2 Решение поставленной задачи.....	8
3 Вероятность ошибки для конкретных видов модуляции.....	13
3.1 Когерентные системы.....	13
3.1.1 Амплитудная манипуляция (АМ).....	13
3.1.2 Фазовая манипуляция (ФМ).....	15
3.1.3 Частотная манипуляция (ЧМ).....	16
3.2 Некогерентные системы.....	18
3.2.1 Амплитудная манипуляция (АМ).....	18
3.2.2 Частотная манипуляция (ЧМ).....	21
4 Численные примеры расчета вероятности ошибки.....	22
5 Источники.....	23
6 Приложение А Графическое представление вероятности ошибки для разных видов двоичной модуляции.....	23
СИГНАЛЬНЫЙ ПОДХОД.....	
1 Общие принципы демодуляции.....	25
2 Формирование квадратур.....	27
3 Принцип захвата тактовой частоты и фазы.....	31
4 Согласованная фильтрация.....	33
5 Пример переноса видеосигнала на промежуточную частоту.....	37
6 Принцип работы ФАПЧ на примере BPSK.....	37
7 Источники.....	50
ПОСЛЕСЛОВИЕ.....	51

ПРЕДИСЛОВИЕ

Первая часть настоящего пособия посвящена постановке и решению статистической задачи различения двух сигналов гармонической формы с частично или полностью известными параметрами, где мешающим фактором является аддитивная помеха типа белый гауссов шум. В контексте систем связи решаемая задача эквивалентна процессу демодуляции двоичных радиосигналов на фоне собственных шумов приемника.

Во второй части изучается процесс демодуляции с точки зрения сигналов. Показываются проблемы, возникающие при реализации модемов, а также пути решения этих проблем. В частности, подробно рассмотрен процесс восстановления несущей в приемнике сигналов с фазовой манипуляцией без остаточной несущей, а также рассмотрены принципы формирования и обработки квадратурных компонентов без межсимвольной интерференции.

СТАТИСТИЧЕСКИЙ ПОДХОД

1 Постановка задачи

Рассматриваются системы передачи информации, в которых бит кодируется двумя импульсами $u_0(t)$ и $u_1(t)$ длительностью T .

Рассматривается канал, который вносит заранее известную задержку τ_0 и масштабирование H_0 сигнала. Передаточная функция такого канала имеет вид

$$H(\omega) = H_0 \exp(-i\omega\tau_0) .$$

Форма принимаемых (received) импульсов $s_{rec}(t)$ в этом случае совпадает с формой передаваемых (transmitted) $s_{tr}(t)$

$$s_{rec}(t) = H_0 s_{tr}(t - \tau_0) .$$

Пусть в канале имеется аддитивная помеха типа белый гауссов шум $n(t)$ с нулевым средним и двусторонней спектральной плотностью мощности $N_0/2$ (рис. 1).

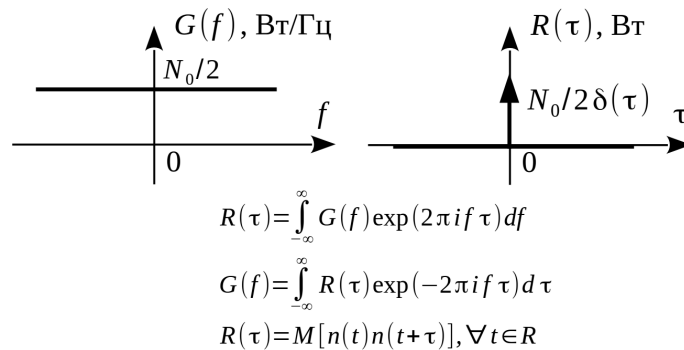


Рисунок 1: Корреляционная (спектральная) модель белого шума

Такая модель общепринята для описания тепловых шумов, присутствующих в любой электрической цепи. Для справедливости этой модели средняя мощность шума должна быть равномерно распределена в рабочей полосе частот приемника. Этим фактом оправдывают белизность шума. Гауссовость (нормальность) шума подтверждается измерениями.

Будем считать, что передаваемые биты равновероятные: $p(u_0) = p(u_1) = 0,5$. В системах связи делается всё, чтобы к этому приблизиться: скремблирование (**scrambling**, слабое шифрование без цели сокрытия информации), сжатие (экономное кодирование). Равновероятность битов и их слабая корреляция позволяют формировать такой спектр передаваемого (излучаемого) сигнала, форма которого будет определяться лишь импульсом-носителем (не модулированным импульсом).

Приемник цифрового радиосигнала должен по принятой реализации угадать тот сигнал, который был передан, соответственно **0** или **1**. Из-за аддитивной помехи иногда могут случаться ошибки, частота которых определяется *вероятностью ошибки* p .

Считается, что риск при двух возможных типах ошибок — $0 \rightarrow 1$ и $1 \rightarrow 0$ — одинаков, то есть все ошибки одинаково вредны. Для систем передачи информации это естественно, а, например, для радиолокации — нет, так как по последствиям **ложная тревога** не равнозначна **пропуску цели**.

Рассматриваются импульсы следующего вида

$$u(t) = \sqrt{\frac{2E}{T}} \sin(2\pi ft + \phi), \quad 0 \leq t \leq T,$$

где E — энергия импульса длительностью T , нормированная на один ом сопротивления.

В случае кратности длительности импульса и периода несущей

$$T = k/f, \quad k = 1, 2, \dots,$$

квадрат амплитуды синусоиды можно точно заменить удвоенной энергией, деленной на длительность. В этом легко убедиться, если вспомнить про определение энергии (на 1 Ом)

$$E = \int_0^T u^2(t) dt.$$

При несоблюдении кратности такая замена будет тем точнее, чем больше периодов несущей укладывается на длительность импульса. Большое количество периодов требуется также для того, чтобы импульс успел набрать номинальную мощность — ведь переходные процессы в цепях никто не отменял. Большие провалы амплитуды характерны, например, для фазовой манипуляции, когда скачок фазы при переходе от одного бита к другому максимален¹.

Считаем, что приемник (демодулятор) точно знает наилучшие моменты времени для взятия отсчетов, то есть имеется идеальная *тактовая синхронизация*.

Математически, рассматривать синусоидальные импульсы (радиосигналы, **narrow-band**) вовсе не обязательно, так как можно построить систему передачи информации на видеосигналах (низкочастотная передача, **baseband**). Радиосигналы превращаются в видео путем сброса частоты на ноль с помощью опорного генератора немодулированной несущей (гетеродина). Использовать радиосигналы вынуждает в первую очередь частотный план, а во вторую — физика радиоволн.

¹ 180 градусов. Скачок 270 градусов равнозначен скачку 90 градусов (неважно, в плюс или в минус).

Определив начальные условия, ставим задачу:

Найти способ обработки принимаемых импульсов, дающий минимальную полную вероятность ошибки.

2 Решение поставленной задачи

На вход приемника поступает реализация в виде суммы переданного сигнала и шума. Так как задержка в канале постоянная, то приемник можно заранее настроить на верные моменты прихода сигнала. Величина ослабления (возможно, что и усиления) в канале постоянная и заранее известная, поэтому в приемнике можно выставить фиксированный сквозной коэффициент передачи.

Результатом обработки принятой реализации на выходе детектора является уровень напряжения (или тока), представленный, например, числовым кодом. Этот код сравнивается с пороговым уровнем (с другим числом), в результате чего по принципу больше-меньше принимается решение о передаче того или иного бита, то есть **0** или **1**. Таким образом демодулятор отображает принятую реализацию в бит².

Так как принятая реализация содержит шум, то число на выходе детектора будет содержать, помимо сигнальной, случайную компоненту. Как будет распределено это число, зависит от распределения шума на входе приемника и от способа вычисления числа (от способа обработки принятой реализации).

В приемнике генерируются эталонные (*априорные*, известные до опыта) сигналы $u_0(t)$ и $u_1(t)$, которые сравниваются с принятой реализацией. Возникает вопрос: какие параметры сигналов известны в приемнике и с какой точностью они там реализуются?

Амплитуда сигнала относительна и реализуется с достаточной точностью, так что имеющиеся нестабильности амплитуды практически не влияют на вероятность ошибки. Проблема возникает с более тонким параметром — частотой (полной фазой), так как она связана со временем, генераторы которого — часы — не могут идеально совпадать в передатчике и приемнике. Частота сигнала на выходе любого генератора имеет свойство медленно блуждать вокруг некоторой средней частоты, причем частоты генераторов приемника и передатчика блуждают независимо. Причем чем больше скорость передачи информации, тем больше сказывается неточность часов.

С начальными фазами еще хуже: они по определению независимы и заранее неизвестны. Невозможно так подобрать момент включения генераторов передатчика и

² Это «жесткий» (**hard**) демодулятор. Есть еще «мягкие» (**soft**) демодуляторы, в которых несколько порогов и они не так грубо квантуют код с выхода детектора.

приемника, что колебания на их выходах при каждом включении будут происходить согласованно, то есть с постоянной и заранее известной разностью фаз. Даже если это и будет сделано, то блуждания частоты приведут к блужданию разности фаз, так как фаза — это интеграл от частоты.

В итоге, из-за неидеальности генераторов несущей, даже в рассматриваемом канале с постоянными параметрами принимаемый сигнал содержит неопределенность относительно фазы (и частоты соответственно), никак не связанную с наличием аддитивного шума.

Естественно, можно вообразить идеальную систему, где известно всё, а к ошибкам приводит только лишь аддитивный шум. Такие системы называются *когерентными*. Практически к ним можно лишь приближаться:

- применять физическую стабилизацию частоты, используя, например, кварц, рубидий, железо-иттриевый гранат;
- применять математическую стабилизацию частоты, используя, например, контур фазовой автоподстройки частоты;
- привлекать спутниковую связь для выравнивания фазы генераторов от единой более точной шкалы времени, и т. д., и т. п.

Такие системы называются *частично-когерентными*. На практике их попросту называют когерентными, так как действительно когерентных в природе не существует.

Если никаких попыток подстроить фазу приемного генератора не делается, то такие системы являются *некогерентными*.

В демодуляторе имеется устройство сравнения принятой реализации с эталонными импульсами $u_0(t)$ и $u_1(t)$. Чем больше похожа принятая реализация, допустим, на $u_0(t)$, тем большим должно быть число на выходе детектора, отвечающего за бит 0. Очевидно, что в итоге требуется сравнить два числа (с детектора 0 и детектора 1) и принять решение о переданном бите.

Известно, что два сигнала одинаковой длительности можно сравнить по *евклидовому расстоянию*

$$d_i^2 = \int_0^T [u(t) - u_i(t)]^2 dt, \text{ здесь дан квадрат расстояния.}$$

Но будет ли оптимальным по критерию минимума вероятности ошибки этот способ сравнения, если шум белый гауссов?..

Принятая реализация является суммой белого гауссова шума и одного из двух сигналов $u_i(t)$, $i=0,1$. Так как линейные преобразования (суммирование и умножение на

константу) не меняют формы плотности вероятностей, то принятая реализация будет распределена по такому же закону, что и шум, то есть по гауссовскому.

Рассмотрим для начала один момент времени; в этом случае принятый сигнал перейдет в принятую случайную величину $u = u_i + n$, $i = 0, 1$, имеющую среднее значение u_i и дисперсию $\overline{(u - u_i)^2} = \overline{n^2} = \sigma^2$, равную дисперсии шума.

Различение возможно, только если $u_0 \neq u_1$.

Приведем пример. Зададим априорные величины $u_0 = -1$, $u_1 = 1$. Пусть из-за влияния шума получено число $u = 0,95$. Естественно, что это число ближе к u_1 , поэтому решающее устройство приемника примет решение о том, что передатчик, вероятнее всего, передавал число u_1 .

Для различения сигналов логично вычислить две вероятности (одну в пользу 0, другую в пользу 1) и затем сравнить их между собой

$$\begin{aligned} p(u_0/u) > p(u_1/u) &\Rightarrow u_0 \\ p(u_0/u) < p(u_1/u) &\Rightarrow u_1 \end{aligned}$$

Отношение двух условных вероятностей можно выразить через отношение условных плотностей и безусловных вероятностей

$$\frac{p(u_0/u)}{p(u_1/u)} = \frac{p(u_0)w(u/u_0)}{p(u_1)w(u/u_1)}$$

где $w(u/u_i)$ — плотность вероятности случайной величины u при условии, что в передатчике выставлен символ u_i (рис. 2). Это отношение преобразовано на основании формулы умножения вероятностей

$$p(u)p(u_0/u) = p(u_0)p(u/u_0) = p(u_0) \lim_{\epsilon \rightarrow 0} \int_u^{u+\epsilon} w(u/u_0) du$$

Отношение условных плотностей называется *отношением правдоподобия*.

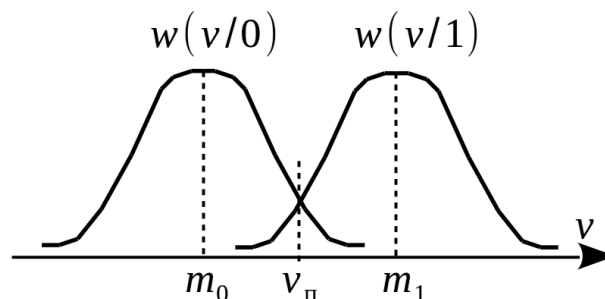


Рисунок 2: Условные плотности вероятности числа на входе решающего устройства приемника

Для гауссовской случайной величины условная плотность выглядит так

$$w(u/u_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(u-u_i)^2}{2\sigma^2}\right],$$

поэтому отношение вероятностей можно переписать в виде

$$\gamma = \frac{p(u_0)}{p(u_1)} \exp\left[\frac{(u-u_1)^2 - (u-u_0)^2}{2\sigma^2}\right].$$

Так как по договоренности биты равновероятные, то

$$\gamma = \exp\left[\frac{(u-u_1)^2 - (u-u_0)^2}{2\sigma^2}\right].$$

Вычислим логарифм от обеих частей, дабы избавиться от экспоненты

$$\ln \gamma = \frac{(u-u_1)^2 - (u-u_0)^2}{2\sigma^2}.$$

Этот логарифм сравнивается с нулем. Решение принимается в пользу u_0 , если он больше нуля, и в пользу u_1 , если меньше. Параметр σ на результат сравнения не влияет, поэтому способ обработки измеренного числа u можно упростить до

$$|u-u_1| > |u-u_0|.$$

Получается, что для гауссовской случайной величины при равновероятных символах оптимальный приемник должен вычислить два **евклидовых расстояния** и сравнить их. Тот из двух символов, к которому ближе u , будет выдан решающим устройством приемника.

Естественным образом выбранный порог $\gamma(u=v_n)=1$ обеспечивает **минимум полной вероятности ошибки**

$$P_{err}(v_n) = p(u_0) \int_{v_n}^{\infty} w(u/u_0) du + p(u_1) \int_{-\infty}^{v_n} w(u/u_1) du, \text{ при } u_1 > u_0,$$

причем для **любых** условных плотностей и **любых** безусловных вероятностей

$$\frac{dP_{err}}{dv_n} = 0 \Rightarrow \frac{w(v_n/u_0)}{w(v_n/u_1)} = \frac{p(u_1)}{p(u_0)}.$$

Прием сигнала от приема случайной величины отличается тем, что сигнал — это множество случайных величин. Здесь рассматривается белый шум, поэтому эти случайные величины не будут коррелировать между собой, а так как шум является гауссовским, то они будут статистически независимыми. Значит плотность вероятности совокупности этих величин равна произведению отдельных плотностей, что приведет в экспоненте к сумме квадратов разностей, а в пределе — при уменьшении шага квантования по времени — к интегралу вида

$$d_i^2 = \frac{1}{N_0} \int_0^T [u(t) - u_i(t)]^2 dt ,$$

где N_0 — односторонняя спектральная плотность мощности белого гауссовского шума.

Если $d_0 < d_1$, то принимаем решение о приеме импульса $u_0(t)$; если $d_1 < d_0$, то считается, что принят импульс $u_1(t)$. Это равнозначно демодуляции **по минимуму евклидова расстояния**.

Таким образом, доказано, что оптимальный приемник-различитель двух равновероятных сигналов на фоне аддитивного белого гауссовского шума должен работать по минимуму евклидова расстояния.

Оказывается, что правило сравнения

$$d_1^2 - d_0^2 = \int_0^T [u(t) - u_1(t)]^2 dt - \int_0^T [u(t) - u_0(t)]^2 dt \begin{matrix} > \\ < \end{matrix} 0 \Rightarrow \begin{matrix} 0 \\ 1 \end{matrix} ,$$

равнозначно следующему³

$$E_{uw} \begin{matrix} > \\ < \end{matrix} \frac{E_1 - E_0}{2} \Rightarrow \begin{matrix} 1 \\ 0 \end{matrix} , \quad (1)$$

где

$$E_{uw} = \int_0^T u(t)w(t) dt , \quad w(t) = u_1(t) - u_0(t) , \quad E_1 = \int_0^T u_1(t)^2 dt , \quad E_0 = \int_0^T u_0(t)^2 dt .$$

Вывод: оптимальный приемник может вместо расстояний вычислять *корреляционный интеграл* между принятым сигналом $u(t)$ и опорным $w(t)$, равным разности двух эталонных сигналов. Приемник, работающий по такому принципу, называют *корреляционным*.

Естественно, что если два эталонных сигнала совпадут, то никакого различения не будет в принципе, а полная вероятность ошибки $P_{err}(v_n)$ будет равна $\frac{1}{2}$ при любом пороге v_n .

Напомним, что единичная вероятность ошибки равнозначна нулевой, поэтому вероятность $\frac{1}{2}$ является самой плохой в плане величины потери информации в канале связи. Однако, это справедливо при независимых ошибках, что гарантирует белый шум, отсчеты которого как раз-таки независимы. Если, допустим, ошибки в некотором канале чередуются

...01010101...,

то вероятность ошибки будет равна $\frac{1}{2}$, но они будут зависимыми, и, зная это свойство канала, все ошибки можно гарантированно исправить.

3 Вероятность ошибки для конкретных видов модуляции

3.1 Когерентные системы

3.1.1 Амплитудная манипуляция (АМ)

Принятая реализация состоит из сигнальной части и шумовой

$$u(t) = u_s(t) + n(t) ,$$

где сигнальная часть в зависимости от передаваемого бита может быть равна одному из двух сигналов

$$u_{1s}(t) = \sqrt{\frac{2E}{T}} \sin(2\pi f t), \quad u_{0s}(t) = 0 .$$

Здесь E — это энергия принятого сигнала непосредственно на входе умножителя (смесителя) демодулятора.

Для когерентных систем в приемнике имеется генератор несущей с раз и навсегда определенной начальной фаз по отношению к начальной фазе несущей передатчика. Примем разность фаз опорных генераторов равной нулю, тогда разностный сигнал будет равен

$$w(t) = \sqrt{\frac{2E_{ref}}{T}} \sin(2\pi f t) .$$

Энергия опорного импульса $w(t)$ длительностью T равна E_{ref} , тогда неравенство (1) для когерентной АМ будет иметь вид

$$v = \int_0^T u(t) w(t) dt \begin{cases} > \frac{E_{ref}}{2} \\ < \frac{E_{ref}}{2} \end{cases} \Rightarrow \begin{cases} 1 \\ 0 \end{cases} .$$

Вывод: чем больше энергия опорного импульса, тем выше должен быть порог.

Когерентность позволила исключить фазу, так как она стала неслучайной и потому известной.

Подставим в последнее неравенство вместо $u(t)$ сумму сигнала и шума, и, в зависимости от вида сигнала (0 или 1), определим числовые параметры случайной величины v

$$v = \int_0^T u_{is}(t) w(t) dt + \int_0^T n(t) w(t) dt = v_s + v_n .$$

Случайность этой величины обязана шумовой компоненте. При прохождении шума через коррелятор величина v линейно зависит от шума, поэтому она будет распределена также как и шум, то есть по нормальному закону.

Среднее значение этой величины зависит от вида передаваемого сигнала (от бита)

$$m_i = \bar{v} = v_s = \begin{cases} \sqrt{E E_{ref}}, & \text{для } \mathbf{1} \\ 0, & \text{для } \mathbf{0} \end{cases},$$

а дисперсия — от спектральной плотности мощности N_0 шума⁴

$$\sigma^2 = \overline{(v - \bar{v})^2} = (N_0/2) \int_0^T w^2(t) dt = (N_0/2) E_{ref}.$$

Вычислим полную вероятность ошибки, учтя то, что все плотности гауссовские $N(u, m_i, \sigma)$ со средними m_i и дисперсиями σ^2 , а оптимальный порог равен полусумме m_0 и m_1

$$P_{err}(v_n) = 0,5 \int_{v_n}^{\infty} w(u/u_0) du + 0,5 \int_{-\infty}^{v_n} w(u/u_1) du = \int_{-\infty}^{v_n} w(u/u_1) du = \int_{-\infty}^{0,5(m_0+m_1)} N(u, m_1, \sigma) du.$$

Последний интеграл табличный, поэтому

$$P_{err} = 1 - \Phi\left(\frac{m_1 - m_0}{2\sigma}\right) = 1 - \Phi\left(\sqrt{\frac{E}{2N_0}}\right),$$

где

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp(-x^2/2) dx.$$

В аргументе *функции ошибок* $\Phi(z)$ стоит заведомо положительное число

$$q = \frac{m_1 - m_0}{\sigma} = \sqrt{\frac{2E}{N_0}},$$

квадрат которого называют *отношением сигнал-шум на выходе демодулятора*⁵

$$q^2 = \frac{(m_1 - m_0)^2}{\sigma^2} = \frac{2 \int_0^T [u_{1s}(t) - u_{0s}(t)]^2 dt}{N_0}.$$

Энергия опорного сигнала в отношении сигнал-шум не участвует. Интеграл в числителе — это энергия разностного сигнала E_d , которая может быть выражена через среднюю энергию сигнала, приходящуюся на один бит E_b

$$E_d = \int_0^T [u_1(t) - u_0(t)]^2 dt = 2 E_b (1 - \rho),$$

⁴ Доказательство см., например, в книге Бернарда Скляра «Цифровая связь», в приложении

⁵ Обязательно проверьте данное равенство!

где ρ — коэффициент корреляции двух сигналов. Последняя формула получена в предположении, что сигналы $u_0(t)$ и $u_1(t)$ имеют равную энергию, которая и является искомой средней энергией на один бит. Эта энергия пропорциональна среднему потреблению энергии модулятором.

Таким образом, для когерентной АМ, у которой два сигнала ортогональны ($\rho=0$), средняя энергия, затрачиваемая на передачу одного бита, равна $E_b = E/2$.

Формула для вероятности ошибки может быть переписана в более универсальном виде

$$P_{err} = 1 - \Phi\left(\frac{q}{2}\right) = 1 - \Phi\left(\sqrt{\frac{E_b(1-\rho)}{N_0}}\right). \quad (2)$$

Разные виды манипуляции сравнивают при одинаковом отношении E_b/N_0 .

Схема когерентного АМ демодулятора

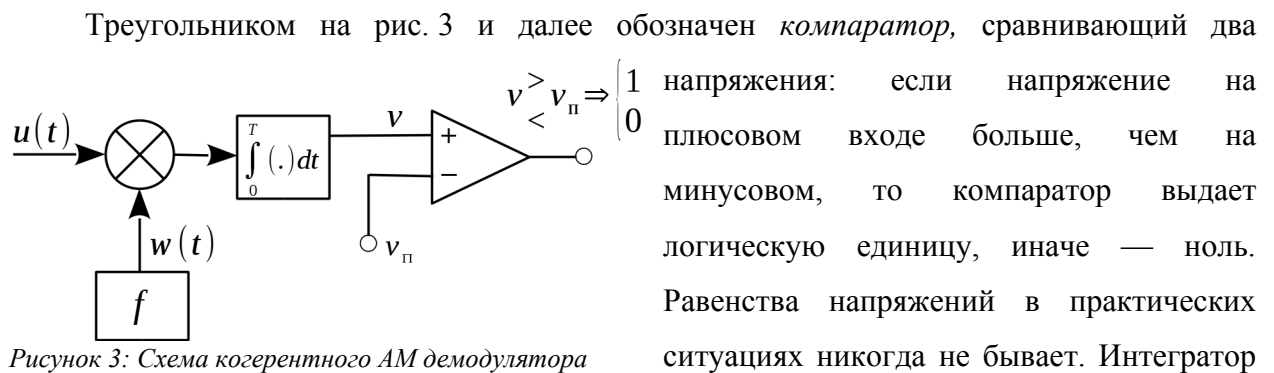


Рисунок 3: Схема когерентного АМ демодулятора

Интегратор со сбросом может быть заменен на интегратор со скользящим интервалом интегрирования $[t; t + T]$, цифровой эквивалент которого является простейшим фильтром, импульсная характеристика которого состоит из набора одинаковых констант.

3.1.2 Фазовая манипуляция (ФМ)

При ФМ оба эталонных сигнала имеют одинаковую энергию, но разнесены по фазе на 180 градусов (на π радиан)

$$u_{1s}(t) = \sqrt{\frac{2E}{T}} \sin(2\pi f t), \quad u_{0s}(t) = -\sqrt{\frac{2E}{T}} \sin(2\pi f t).$$

Опорный сигнал пропорционален разности двух эталонных

$$w(t) \sim u_1(t) - u_0(t) = 2\sqrt{\frac{2E}{T}} \sin(2\pi f t),$$

Так как энергия опорного сигнала может регулироваться и теоретически не влияет на вероятность ошибки, то можно записать

$$w(t) = \sqrt{\frac{2E_{ref}}{T}} \sin(2\pi f t) .$$

Неравенство (1) для ФМ принимает вид

$$v = \int_0^T u(t)w(t) dt \begin{matrix} > \\ < \end{matrix} 0 \Rightarrow \begin{matrix} 1 \\ 0 \end{matrix} .$$

Случайная величина v , аналогично случаю когерентной АМ, будет иметь нормальное распределение, но с несколько другими средними

$$m_i = \overline{v} = v_s = \begin{cases} \sqrt{E E_{ref}}, & \text{для } 1 \\ -\sqrt{E E_{ref}}, & \text{для } 0 \end{cases} ,$$

но той же дисперсией

$$\sigma^2 = (N_0/2) \int_0^T w^2(t) dt = (N_0/2) E_{ref} .$$

Полная вероятность ошибки при оптимальном пороге (нулевом)

$$P_{err} = 1 - \Phi\left(\frac{m_1 - m_0}{2\sigma}\right) = 1 - \Phi\left(\sqrt{\frac{2E}{N_0}}\right) = 1 - \Phi\left(\sqrt{\frac{2E_b}{N_0}}\right) .$$

Отношение сигнал-шум на выходе

$$q^2 = \frac{(m_1 - m_0)^2}{\sigma^2} = \frac{8E}{N_0} ,$$

а средняя энергия на один бит $E_b = E$. Коэффициент корреляции двух ФМ сигналов $\rho = -1$.

Доказано, что когерентная ФМ является наиболее помехоустойчивой для рассматриваемого канала связи и типа помехи.

Схема когерентного ФМ демодулятора имеет такой же вид, что и схема для когерентного АМ демодулятора (рис. 3), за исключением того, что порог будет нулевым.

3.1.3 Частотная манипуляция (ЧМ)

Для данного типа манипуляции оба сигнала имеют равные энергии, равные фазы, но разные частоты несущих

$$u_{1s}(t) = \sqrt{\frac{2E}{T}} \sin(2\pi f_1 t), \quad u_{0s}(t) = \sqrt{\frac{2E}{T}} \sin(2\pi f_0 t) ,$$

Неравенство (1) перейдет в

$$v = \int_0^T u(t)w(t) dt \begin{matrix} > \\ < \end{matrix} 0 \Rightarrow \begin{matrix} 1 \\ 0 \end{matrix} .$$

При ЧМ последнее неравенство удобнее преобразовать в следующее

$$v_1 \begin{matrix} > \\ < \end{matrix} v_0 \Rightarrow \begin{cases} 1 \\ 0 \end{cases},$$

где

$$v_1 = \int_0^T u(t)u_1(t)dt, \quad v_0 = \int_0^T u(t)u_0(t)dt.$$

Величины v_1 и v_2 на выходах корреляторов будут нормальными случайными величинами со средними значениями

$$\overline{v_1} = \begin{cases} \sqrt{E E_{ref}}, & 1 \\ 0, & 0 \end{cases}, \quad \overline{v_0} = \begin{cases} 0, & 1 \\ \sqrt{E E_{ref}}, & 0 \end{cases},$$

и одинаковой дисперсией

$$\sigma_1^2 = \sigma_2^2 = (N_0/2) \int_0^T u_1^2(t)dt = (N_0/2) \int_0^T u_0^2(t)dt = (N_0/2) E_{ref}.$$

Нулевые средние при поступлении «чужих» сигналах будут при условии обеспечения ортогональности эталонных сигналов на заданном интервале

$$\rho = \frac{1}{\sqrt{E_0 E_1}} \int_0^T u_1(t)u_0(t)dt = 0.$$

Доказано, что ортогональность двух сигналов разных частот является наилучшим выбором по соотношению сигнал-шум после корреляционной обработки. Этот выигрыш совпадает с выигрышем когерентной АМ, в которой два эталонных сигнала тривиально ортогональны. При ФМ сигналы противоположны, то есть удалены друг от друга максимально возможно (при ограничении на среднюю энергию на один бит).

Чтобы вычислить вероятность ошибки демодуляции ЧМ, удобно рассмотреть разность $v_1 - v_0$, которая будет гауссовской случайной величиной со средним

$$\overline{v_1 - v_0} = \begin{cases} \sqrt{E E_{ref}}, & 1 \\ -\sqrt{E E_{ref}}, & 0 \end{cases},$$

и дисперсией

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 = N_0 E_{ref}.$$

Вероятность ошибки будет равна

$$P_{err} = 1 - \Phi\left(\frac{m_1 - m_0}{2\sigma}\right) = 1 - \Phi\left(\sqrt{\frac{E}{N_0}}\right) = 1 - \Phi\left(\sqrt{\frac{E_b}{N_0}}\right).$$

Средняя энергия на один бит $E_b = E$.

Таким образом, помехоустойчивость когерентных АМ и ЧМ одинакова.

Схема когерентного ЧМ демодулятора

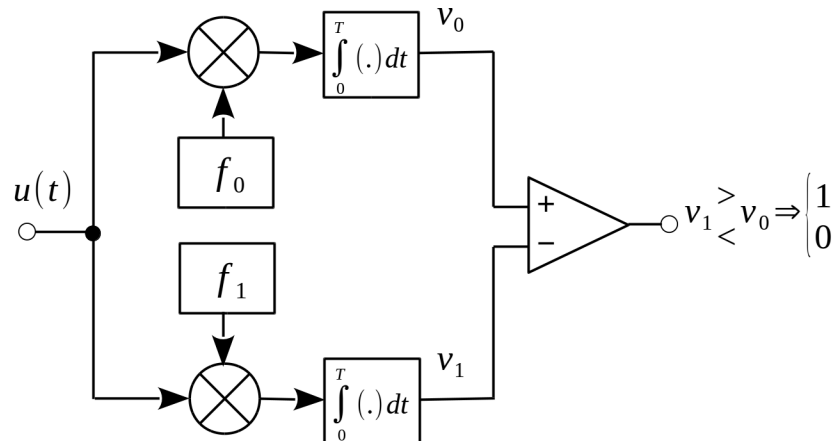


Рисунок 4: Схема когерентного ЧМ демодулятора

На рис. 4 порог в явном виде отсутствует, и сравниваются напряжения из двух ветвей. Конечно, можно обойтись одним коррелятором, но для этого надо иметь опорный сигнал, равный разности двух образцовых. В варианте с одним коррелятором порог будет равен нулю.

3.2 Некогерентные системы

3.2.1 Амплитудная манипуляция (АМ)

Полезный сигнал, в отличие от когерентной АМ, будет иметь случайный параметр — фазу

$$u_{1s}(t) = \sqrt{\frac{2E}{T}} \sin(2\pi ft + \varphi), \quad u_{0s}(t) = 0,$$

где φ — та самая случайная фаза, по сути равная разности фаз между сигнальной частью принимаемой реализации и эталонным сигналом приемника. От этой фазы в процессе обработки принятой реализации необходимо избавиться, так как она не несет никакой информации.

Геометрически, при некогерентной демодуляции необходимо измерить длину вектора со случайным углом и сравнить ее с пороговой длиной.

Для таких «измерений» потребуются два ортогональных колебания одинаковой энергии с любой начальной фазой — лишь бы разность фаз была 90 градусов

$$u_c(t) = \sqrt{\frac{2E_{ref}}{T}} \cos(2\pi ft)$$

$$u_s(t) = \sqrt{\frac{2E_{ref}}{T}} \sin(2\pi ft)$$

— это квадратурные сигналы.

Некогерентная оптимальная обработка будет заключаться в следующем:

- перемножаем принятую реализацию с квадратурными сигналами $u_c(t)$ и $u_s(t)$;
- интегрируем результаты этих произведений по длительности импульса, получая два числа v_I и v_Q , соответственно;
- находим модуль, избавляясь от случайной фазы

$$v = \sqrt{v_I^2 + v_Q^2} = \sqrt{(\sqrt{E E_{ref}} \sin \varphi)^2 + (\sqrt{E E_{ref}} \cos \varphi)^2} = \sqrt{E E_{ref}} .$$

Последняя формула выведена при отсутствии шума.

Схема некогерентного демодулятора АМ сигнала показана на рис. 5.

Схема некогерентного АМ демодулятора

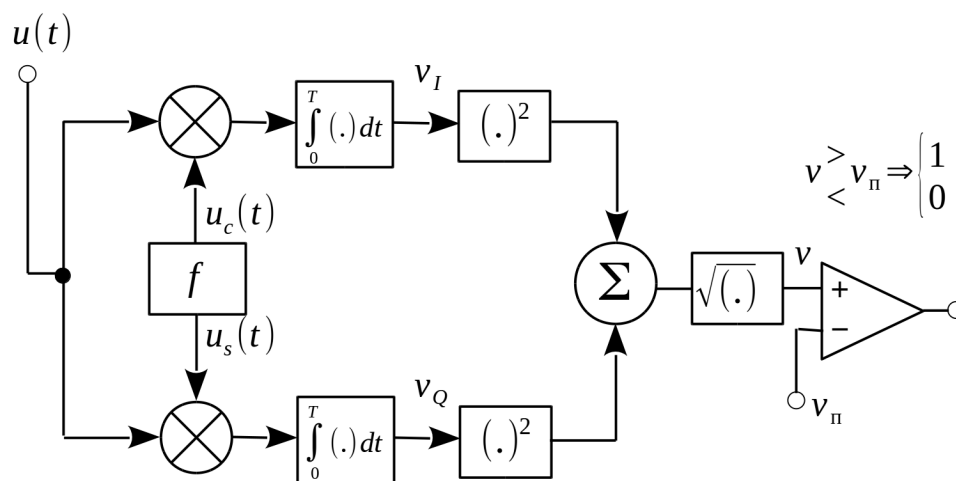


Рисунок 5: Схема некогерентного АМ демодулятора

Определим оптимальный порог и итоговую вероятность ошибки.

Величины v_I и v_Q являются гауссовскими, так как гауссов шум прошел через коррелятор, который линеен по отношению к шуму.

Средние значения этих случайных величин равны

$$\overline{v_I} = \begin{cases} \sqrt{E E_{ref}} \sin \varphi, & 1 \\ 0, & 0 \end{cases}, \quad \overline{v_Q} = \begin{cases} \sqrt{E E_{ref}} \cos \varphi, & 1 \\ 0, & 0 \end{cases},$$

а их дисперсии

$$\sigma^2 = \sigma_I^2 = \sigma_Q^2 = (N_0/2) \int_0^T u_c^2(t) dt = (N_0/2) \int_0^T u_s^2(t) dt = (N_0/2) E_{ref} .$$

Однако итоговая случайная величина v (модуль) уже не будет гауссовской (рис. 6).

Если был передан нулевой сигнал, то средние значения квадратур будут равны нулю, а v будет распределена по закону Рэлея [1]

$$w_0(v) = \frac{v}{\sigma^2} \exp\left(\frac{-v^2}{2\sigma^2}\right), \quad v \geq 0 .$$

Если был передан ненулевой импульс, то средние значения квадратур будут отличны от нуля, а v будет распределена по обобщенному закону Рэлея (распределение Райса) [2]

$$w_1(v) = \frac{v}{\sigma^2} \exp\left[-\frac{(v^2 + a^2)}{2\sigma^2}\right] I_0\left(\frac{av}{\sigma^2}\right), \quad v \geq 0,$$

где $a = \sqrt{\overline{v_I^2} + \overline{v_Q^2}} = \sqrt{E E_{ref}}$,

$I_0(x)$ — модифицированная функция Бесселя первого рода нулевого порядка.

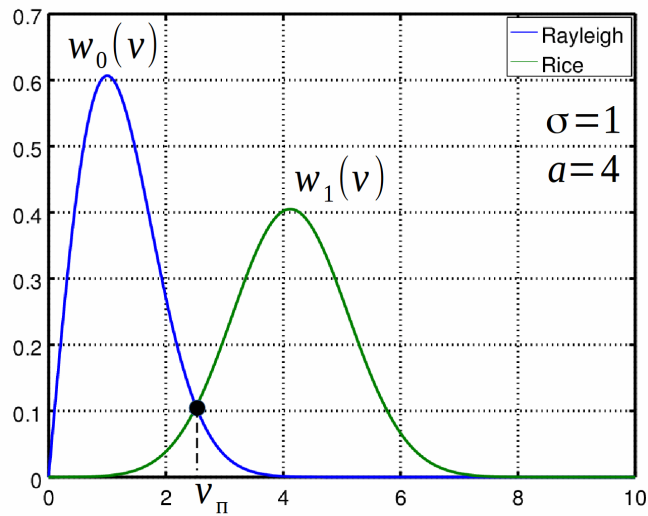


Рисунок 6: Условные плотности вероятности величины на входе решающего устройства (компаратора) некогерентного АМ демодулятора

На рис. 6 показан пример рэлеевского $w_0(v)$ и райсовского $w_1(v)$ распределений для $\sigma=1$ и $a=4$. Параметр a практически совпадает с модой (вершиной графика) распределения Райса, причем это совпадение тем лучше, чем больше отношение a/σ . При $a=0$ мода совпадает с σ и две кривых сольются.

Различение сигналов возможно лишь при $a>0$; в противном случае вероятность ошибки будет всегда равна $1/2$. Так как один из АМ сигналов нулевой, то здесь правильнее говорить не о различении, а об обнаружении сигнала на фоне шума (радиолокационный вариант).

Чем больше a по сравнению с σ , тем меньше вероятность ошибки. По сути, отношение этих величин является отношением сигнал-шум по выходу, то есть после демодуляции. Так как средняя энергия на один бит равна $E_b = E/2$, то отношение сигнал-шум по входу можно найти так

$$\frac{E_b}{N_0} = \frac{1}{4} \left(\frac{a}{\sigma}\right)^2.$$

Полная вероятность ошибки

$$P_{err}(v_n) = 0,5 \int_{v_n}^{\infty} w_0(v) dv + 0,5 \int_0^{v_n} w_1(v) dv .$$

Минимум вероятности ошибки достигается при условии пересечения двух кривых (см. рис. 6)

$$w_0(v_n) = w_1(v_n) ,$$

что дает уравнение

$$I_0\left(\frac{a}{\sigma} \frac{v_n}{\sigma}\right) = \exp\left(\frac{1}{2} \frac{a^2}{\sigma^2}\right) ,$$

относительно переменной v_n . Найдя из него оптимальный порог, можно вычислить итоговую вероятность ошибки

$$P_{err} = 0,5 \exp\left(\frac{-v_n^2}{2\sigma^2}\right) + 0,5 \left[1 - Q_1\left(\frac{a}{\sigma}; \frac{v_n}{\sigma}\right)\right] ,$$

где [3]

$$Q_1(a; b) = \int_b^{\infty} t \exp\left(-\frac{t^2 + a^2}{2}\right) I_0(at) dt \quad —$$

функция Маркума, а $I_0(x)$ — функция Бесселя первого рода нулевого порядка.

Простой способ вычисления $Q_1(a, b)$ можно найти, например, в [4]

$$Q_1(a, b) = 1 - \frac{b^2}{2} \exp\left(\frac{-a^2}{2}\right) \sum_{n \geq 0} P_n\left(\frac{a^2}{2}, -\frac{b^2}{2}\right) ,$$

где

$$P_{n+1}(a, b) = \frac{(2n+1-a)b}{(n+1)(n+2)} P_n(a, b) - \frac{nb^2}{(n+1)^2(n+2)} P_{n-1}(a, b) ,$$

$$P_0(a, b) = 1 , \quad P_1(a, b) = \frac{(1-a)b}{2} .$$

Таким образом, вероятность ошибки при некогерентной демодуляции АМ сигнала вычисляется не так просто, как при когерентной.

3.2.2 Частотная манипуляция (ЧМ)

Здесь принимаемые сигнальные компоненты имеют две разные частоты, одинаковые энергии и длительности, но случайные начальные фазы

$$u_{1s}(t) = \sqrt{\frac{2E}{T}} \sin(2\pi f_1 t + \varphi_1), \quad u_{0s}(t) = \sqrt{\frac{2E}{T}} \sin(2\pi f_0 t + \varphi_0) .$$

В отличие от некогерентной АМ, здесь потребуются четыре квадратуры — по две на каждую частоту

$$\begin{aligned} u_{s1}(t) &= \sqrt{\frac{2E_{ref}}{T}} \sin(2\pi f_1 t) & u_{s0}(t) &= \sqrt{\frac{2E_{ref}}{T}} \sin(2\pi f_0 t) \\ u_{c1}(t) &= \sqrt{\frac{2E_{ref}}{T}} \cos(2\pi f_1 t) & u_{c0}(t) &= \sqrt{\frac{2E_{ref}}{T}} \cos(2\pi f_0 t) \end{aligned} .$$

Схему некогерентного ЧМ демодулятора предлагается нарисовать самим, скомбинировав два некогерентных АМ демодулятора.

При ЧМ требуются два канала: один работает на частоте f_1 , другой — на f_0 . Компаратор сравнивает выходы обоих каналов. Если в одном канале выходная величина имеет распределение Рэлея, то в другом — Райса, и наоборот, поэтому ошибка при демодуляции будет лишь в том случае, когда *рэлеевская величина превысит райсовскую*. Вероятность этого события вычисляется как обычно

$$P_{err} = \int_0^{\infty} \left[\int_{v_1}^{\infty} w_0(v_0) dv_0 \right] w_1(v_1) dv_1 ,$$

где рэлеевская величина обозначена как v_0 , а райсовская как v_1 .

Внешний интеграл — это усреднение по всем райсовским величинам. Внутренний интеграл — вероятность попадания рэлеевской величины в интервал от v_1 до ∞ , что равносильно неравенству $v_0 > v_1$.

Несмотря на громоздкость интеграла, его значение получается простым

$$P_{err} = \frac{1}{2} \exp\left(\frac{-E_b}{2N_0}\right) .$$

4 Численные примеры расчета вероятности ошибки

- Задаем входное отношение сигнал-шум $\frac{E_b}{N_0} = 10$.

Некогерентная АМ: $a/\sigma = 2\sqrt{10} \approx 6,32$, $v_n/\sigma \approx 3,55$, $P_{err} \approx 1,9 \cdot 10^{-3}$.

Когерентная АМ (ЧМ): $P_{err} = 1 - \Phi(\sqrt{10}) \approx 7,8 \cdot 10^{-4}$.

Некогерентная ЧМ: $P_{err} = 0,5 \exp(-10/2) \approx 3,4 \cdot 10^{-3}$.

Когерентная ФМ: $P_{err} = 1 - \Phi(\sqrt{2 \cdot 10}) \approx 3,9 \cdot 10^{-6}$.

- Задаём $\frac{E_b}{N_0} = 1$.

Некогерентная АМ: $a/\sigma = 2\sqrt{1} = 2$, $v_n/\sigma \approx 1,75$, $P_{err} \approx 0,257$.

Когерентная АМ (ЧМ): $P_{err} = 1 - \Phi(\sqrt{1}) \approx 0,159$.

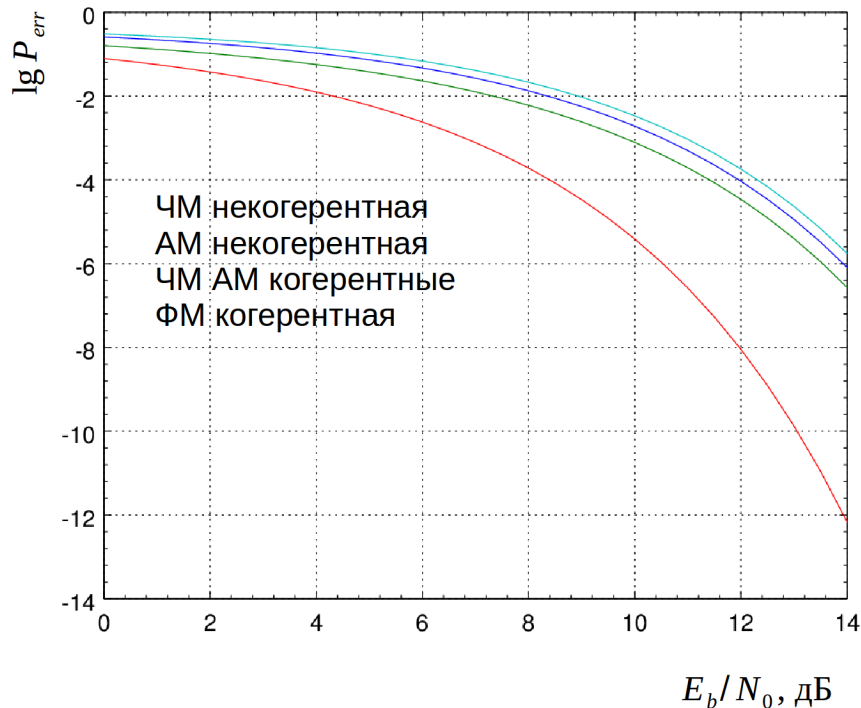
Некогерентная ЧМ: $P_{err} = 0,5 \exp(-1/2) \approx 0,303$.

Когерентная ФМ: $P_{err} = 1 - \Phi(\sqrt{2 \cdot 1}) \approx 7,9 \cdot 10^{-2}$.

5 Источники

1. https://en.wikipedia.org/wiki/Rayleigh_distribution
2. https://en.wikipedia.org/wiki/Rice_distribution
3. https://en.wikipedia.org/wiki/Marcum_Q-function
4. The generalized Marcum Q-function: an orthogonal polynomial approach / Sz. Andras, A. Baricz, Y. Sun // Acta Univ. Sapientiae, Mathematica, 3, 1 (2011) 60–76.

6 Приложение А Графическое представление вероятности ошибки для разных видов двоичной модуляции



Вероятность ошибки в зависимости от среднего отношения сигнал-шум на входе демодулятора для разных видов двоичной модуляции: АМ — амплитудной, ЧМ — частотной и ФМ — фазовой. Четырём подписям (легендам) соответствуют четыре кривые: хуже всего по вероятности ошибки некогерентная ЧМ, лучшие всего — когерентная ФМ.

СИГНАЛЬНЫЙ ПОДХОД

1 Общие принципы демодуляции

Вся радиосвязь построена на принципе несущей частоты — гармонической функции времени

$$s(t) = A \cos(2\pi f t + \varphi) \quad , \quad (3)$$

где A — амплитуда, f — частота, в герцах, φ — начальная фаза, в радианах.

После приема гармонического сигнала изменятся его амплитуда (из-за ослабления радиоволн), фаза, которая вообще не определена, и, если передатчик и/или приемник движутся (**мобильная связь**), то за счет **эффекта Доплера** сдвигается частота: в сторону увеличения, если объекты радиально сближаются, и уменьшения, если объекты радиально отдаляются.

Удобство гармоник заключается в том, что искажение такого сигнала эфиром (пространством) в процессе передачи и приема сводится лишь к изменению параметров функции, а не к изменению самой функции

$$\text{ЭФИР} \{ A(t) \sin[2\pi f t + \varphi(t)] \} = B(t) \sin[2\pi f t + \psi(t)] \quad .$$

Естественно, чистый тон (3) нет особого смысла излучать в эфир. Если исходить из теории информации, такой сигнал не будет нести информационной нагрузки, так как в нем ни один из трех параметров не изменяется со временем. Поэтому в системах передачи информации используют такие сигналы, в которых хотя бы один из трех параметров изменяется, а значит этот сигнал уже не будет чистым тоном и займет некоторую **полосу частот (band)** в окрестности несущей частоты. В этой ненулевой полосе и будет сосредоточена передаваемая информация⁶.

Информационные параметры сигнала (3) должны меняться **медленно** по сравнению с несущей частотой⁷. В противном случае за счет переходных процессов (например, в полосовом фильтре) сигнал может не успеть набрать номинальную мощность (рис. 7).

⁶ Нулевая ширина спектра сигнала, кстати, по формуле Шеннона соответствует нулевой пропускной способности канала связи.

⁷ В этом случае сигнал называют узкополосным (**narrow-band**).

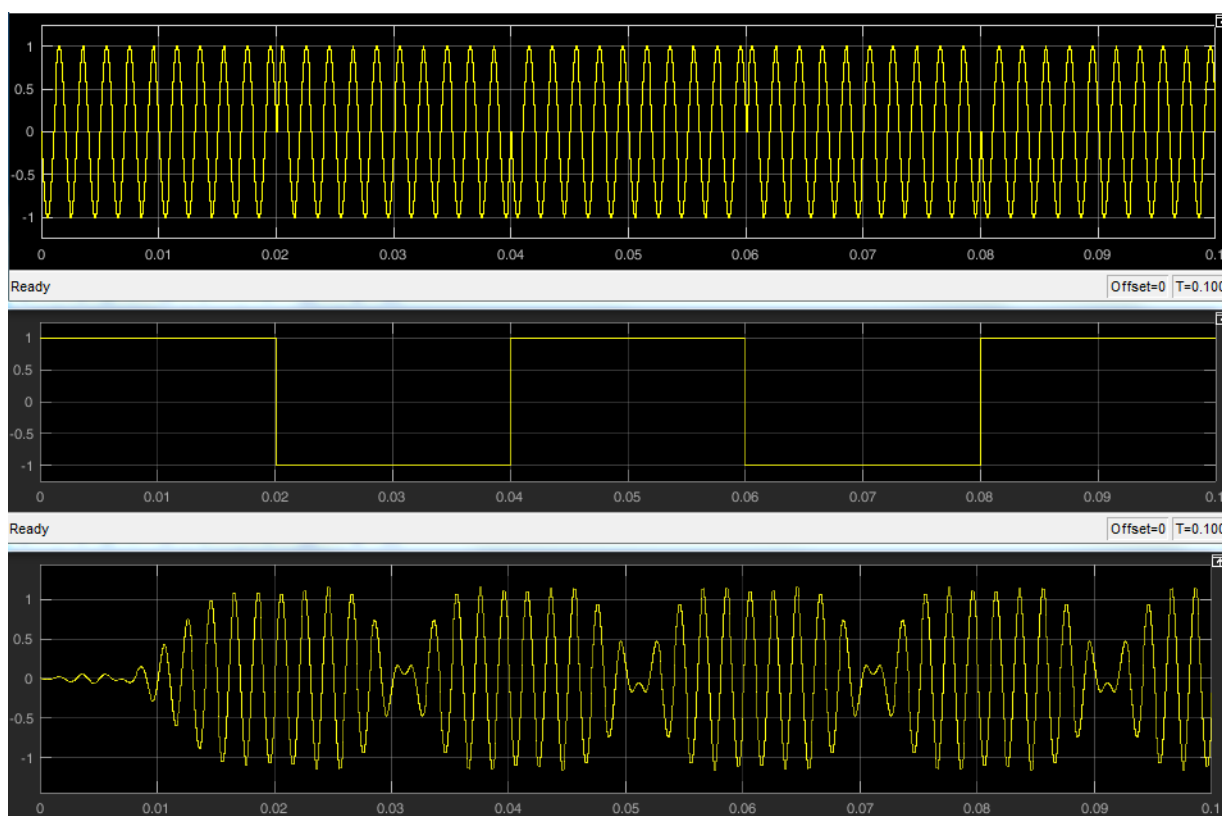


Рисунок 7 Эффект амплитудной модуляции при прохождении идеального ФМ сигнала через полосовой фильтр

На рис. 7 показан вариант отношения частот 1 к 10, когда на длительность одного символа (0,02 с) приходится 10 периодов несущей (500 Гц). Задержка выходного сигнала на половину длительности символа связана с фазово-частотной характеристикой фильтра. Фильтр настроен на пропускание диапазона частот (450...550) Гц, так как информационные символы имеют частоту следования $1/0,02 = 50$ Гц. При этих условиях сигнал набирает свою мощность за 4–5 периодов несущей частоты.

Смена фазы на 180 градусов у ФМ сигнала дает самую глубокую паразитную амплитудную модуляцию. В реальности соответствующий видеосигнал (квадратурные компоненты) генерируют так, чтобы не было резких перепадов уровня как на среднем графике рис. 7. В случае сглаживания квадратур мощность радиосигнала меньше расплескивается по частотному диапазону и больше сосредотачивается в окрестности несущей частоты. Другими словами, спектр становится более компактным (но за счет некоторого уширения), что благоприятно сказывается на требовании к полосовым фильтрам передатчика, которые помогают обеспечить требования частотной маски (**modulation mask**).

2 Формирование квадратур

Техника формирования **baseband**⁸ сигнала называется **pulse shaping**; за это отвечает формирующий фильтр, который реализуется в цифровом виде с последующим цифро-аналоговым преобразованием (ЦАП).

На вход формирующего фильтра поступают **цифры** с формирователя комплексной огибающей (**IQ mapper**), который работает по таблице, отображающей входные кодовые слова в квадратуры I и Q . Программная смена таблицы позволяет легко переключаться на другой вид цифровой модуляции.

Рассмотрим вариант, когда у сигнала (1) медленно изменяется амплитуда и фаза

$$s(t) = A(t) \cos[2\pi f t + \varphi(t)] \quad (4)$$

Такой сигнал можно разложить на квадратурные составляющие $I(t)$ и $Q(t)$

$$\begin{aligned} s(t) &= A(t) \cos(2\pi f t) \cos[\varphi(t)] - A(t) \sin(2\pi f t) \sin[\varphi(t)] = \\ &= I(t) \cos(2\pi f t) - Q(t) \sin(2\pi f t) \end{aligned} \quad (5)$$

Этот принцип позволяет построить универсальный квадратурный модулятор, в составе которого должен быть генератор частоты f , два множителя и сумматор⁹. Формирующий фильтр при этом будет отвечать за формирование квадратур I и Q , то есть за интерполяцию отсчетов с выхода табличного формирователя **IQ mapper**.

На приемной стороне квадратуры должны быть пропущены через фильтр, согласованный с формой передаваемого видеоимпульса, и протестированы с помощью отсчетного устройства. Согласованный фильтр обеспечивает максимальное отношение сигнал-шум, если шум является белым. Тепловой шум в рабочей полосе частот хорошо подходит под модель белого шума.

На рис. 1 на среднем графике показана одна квадратура (вторая равна нулю). Формирующий фильтр в данном случае имеет АЧХ¹⁰ типа

$$\left| \frac{\sin x}{x} \right|,$$

для реализации которой требуется неограниченная полоса частот. Такая АЧХ даст недопустимый уровень внеполосных компонент излучаемого **baseband** сигнала: мощность такого сигнала падает со скоростью всего лишь 6 дБ/октаву¹¹ (медленнее падать нельзя).

⁸ **Baseband** сигнал — сигнал с нулевой центральной частотой (видеосигнал).

⁹ В настоящее время генерация произвольных сигналов может быть сделана с помощью техники прямого цифрового синтеза, **DDS (Direct Digital Synthesis)**.

¹⁰ Амплитудно-частотную характеристику.

¹¹ $10 \lg(f_2/f_1)^2 = 10 \lg 4 \approx 6$, октава — двукратное увеличение частоты.

Однако, выбрать формирующий фильтр из принципа абы какой-нибудь ФНЧ¹² не получится, и вот почему.

В цифровых системах связи данные поступают с заданной тактовой частотой, т. е. символ за символом. Некоторый ФНЧ при подаче на его вход текущего символа $A[i]$, вероятно, даст ненулевой отклик в момент времени $i+1$: в этом случае произойдет **межсимвольная интерференция (Inter Symbol Interference, ISI)**. На рис. 8 показан случай

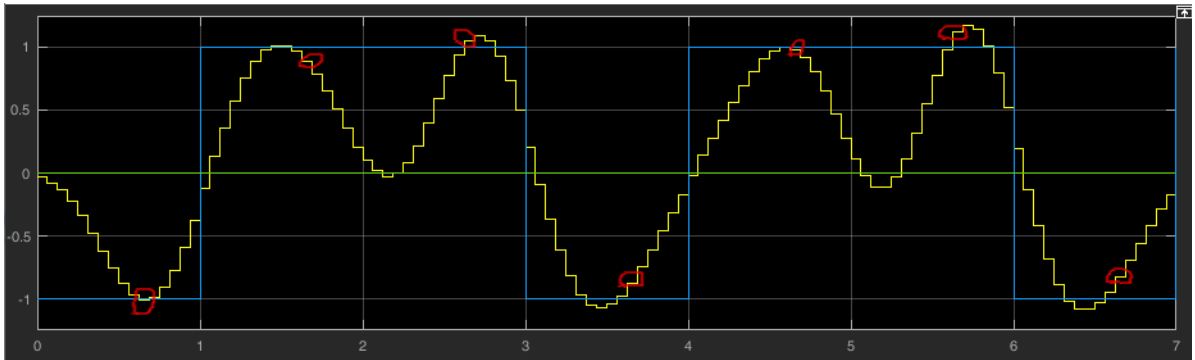


Рисунок 8 Пример baseband сигнала с межсимвольной интерференцией
 прохождения **baseband** сигнала с уровнями ± 1 через некоторый ФНЧ с единичным максимумом импульсной характеристики (11 такт, обведено красным кружочком). Цифровые данные поступают каждый 16 такт с шагом 1 с. Видно, что первый обведенный отсчет идеально совпадает с уровнем -1 , а второй и последующие уже нет — это и есть результат интерференции.

В общем случае по известной сумме нельзя однозначно узнать слагаемые, то есть точно извлечь информацию; приближенно извлечь можно, так как исходная битовая информация здесь заложена в условии больше-меньше нуля, поэтому ошибка по амплитуде не всегда приведет к битовой ошибке на выходе решающего устройства.

В качестве формирующих фильтров подходят фильтры с нулевой межсимвольной интерференцией¹³ и с достаточно малым уровнем внеполосных компонент. Кстати, формирующий фильтр на рис. 7 имеет нулевую интерференцию, причем на всей длительности символа! Однако, для запуска такой системы передачи информации потребуется «жирный» канал шириной $(10...15)/T_c$ (можно и шире), зато к цепи тактовой синхронизации будут предъявляться очень слабые требования по нестабильности временного положения (фазы) тактовых импульсов.

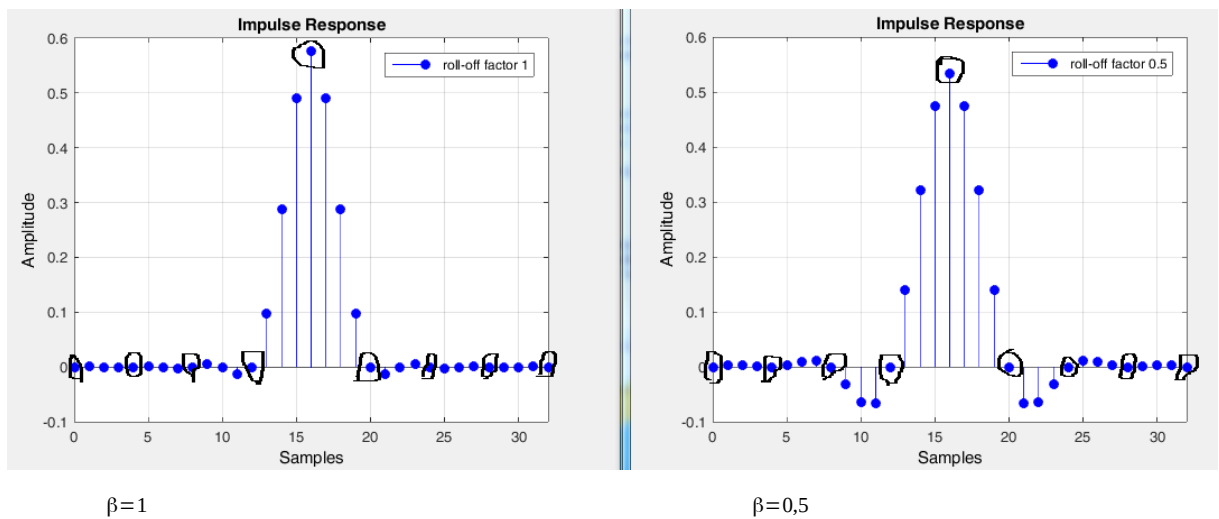
¹² Фильтр нижних частот.

¹³ Можно использовать фильтры с контролируемой интерференцией. Способ их детектирования отличается. Смотри дубинарные сигналы в книгах Джона Прокиса «Цифровая связь» или Бернарда Скляра.

Разработан фильтр типа «приподнятый косинус» (**raised cosine filter**), который имеет нулевую интерференцию (рис. 9) и позволяет регулировать уровень внеполосных компонент сигнала за счет ширины спектра, которую можно менять в пределах

$$2B = \frac{(1+\beta)}{T_s}, \quad 0 \leq \beta \leq 1, \quad \beta - \text{roll-off factor.}$$

Однако, нулевая интерференция сохраняется только в тактовые моменты времени (при идеальной тактовой синхронизации), причем чем шире спектр, тем быстрее затухают так называемые боковые лепестки импульсной характеристики фильтра, что приводит к меньшему влиянию нестабильности фазы тактовых импульсов на уровень интерференции.



Черными кружочками обведены отсчеты в тактовые моменты времени: все они имеют нулевой уровень, кроме одного центрального отсчета, который предназначен для переноса информации

Рисунок 9 Примеры импульсных характеристик реализуемого фильтра типа «приподнятый косинус»

Чем ниже **roll-off factor**, тем больше уровень пульсаций в импульсной характеристике фильтра (рис. 9) и больше уровень внеполосных компонент, но уже спектр (рис. 10).

Коэффициент интерполяции (число отсчетов на символ) позволяет регулировать период спектра, т. е. позволяет «раздвинуть башни», облегчив требования к фильтру ЦАП¹⁴. Если **roll-off factor** равен нулю, то импульсная характеристика совпадет с функцией $\sin x/x$, а частотная характеристика по форме будет прямоугольной.

Словосочетание «реализуемый фильтр» подразумевает его реализуемость в виде конечного числа ячеек памяти, умножителей и сумматоров.

Теоретическая импульсная характеристика $h(t)$ фильтра «приподнятый косинус» не ограничена, поэтому для реализуемости её приходится ограничивать; мало того — дискретизировать $h_n = h(nT_d)$.

¹⁴ Цифро-аналоговый преобразователь.

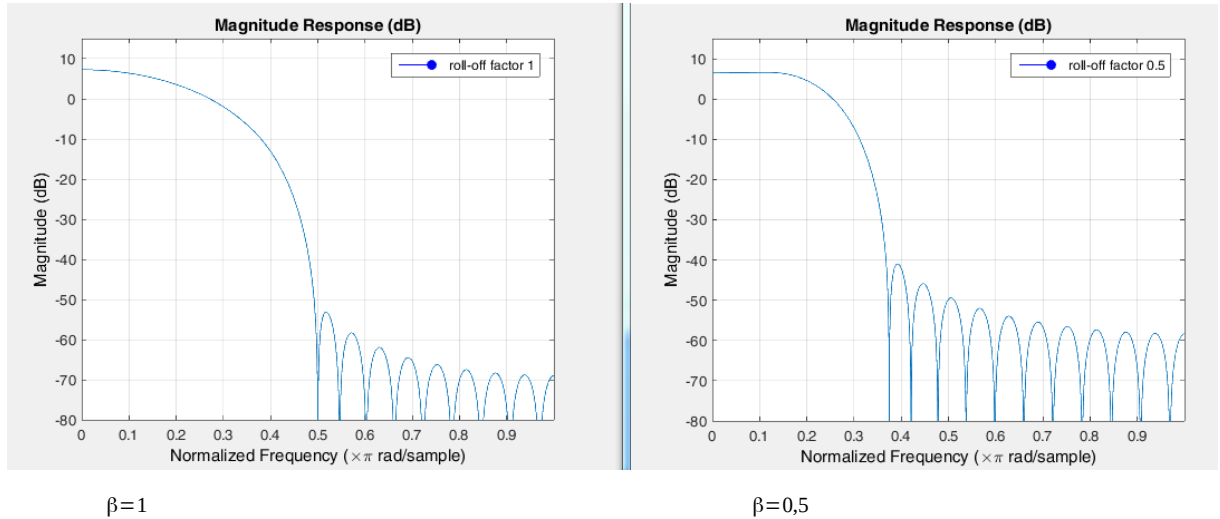
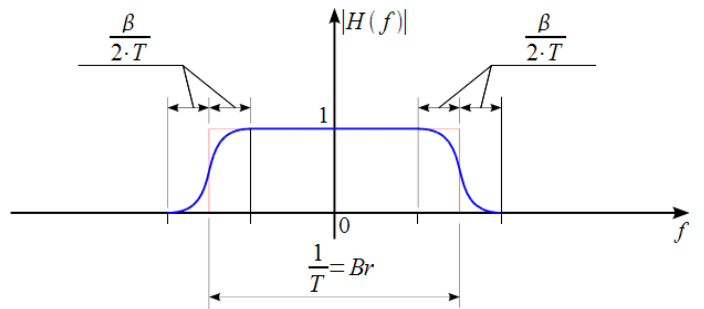


Рисунок 10 Примеры амплитудно-частотных характеристик реализуемого фильтра типа «приподнятый косинус»

Из-за принудительного ограничения импульсной характеристики, за границами идеальной полосы пропускания

$$\frac{(1+\beta)}{T_s}$$

появляется ненулевая составляющая, которая и определяет уровень внеполосной компоненты. Теоретическая АЧХ не имеет такой компоненты, так как она моделируется отрезком «косинусоида-константа-косинусоида»¹⁵. Поэтому уровень внеполосной компоненты зависит не только от параметра **roll-off factor**, но и от количества символов на длину импульсной характеристики (на рис. 9 это 8, т. е. восемь символьных интервалов).



Межсимвольная интерференция также возникает из-за неидеальной частотной характеристики линии передачи. Для беспроводной связи такой линией является так называемый эфир (открытое пространство) с множеством естественных и искусственных объектов. Причем эти объекты могут перемещаться в пространстве (автомобили, дождевые облака), а также менять свои электрические характеристики (подстилающая поверхность земли, растительность). Поэтому в приемнике перед согласованным фильтром должен стоять **эквалайзер**, выравнивающий частотную характеристику линии передачи, которая периодически оценивается прямо-передатчиком с помощью тестовых сигналов. Такая

¹⁵ Рисунок с теоретической АЧХ фильтра «приподнятый косинус» взят с сайта <http://www.dsplib.ru/content/isi/isi.html>, с которым рекомендуется ознакомиться.

адаптивность особенно важна для мобильной связи, где в результате возможного перемещения абонентов частотная характеристика линии передачи изменяется во времени.

На рис. 11 показан результат прохождения **BPSK**¹⁶ **baseband** сигнала через фильтр «приподнятого косинуса», так, что пик его импульсной характеристики равен единице, коэффициент интерполяции — 4, длительность характеристики — 8 символов.

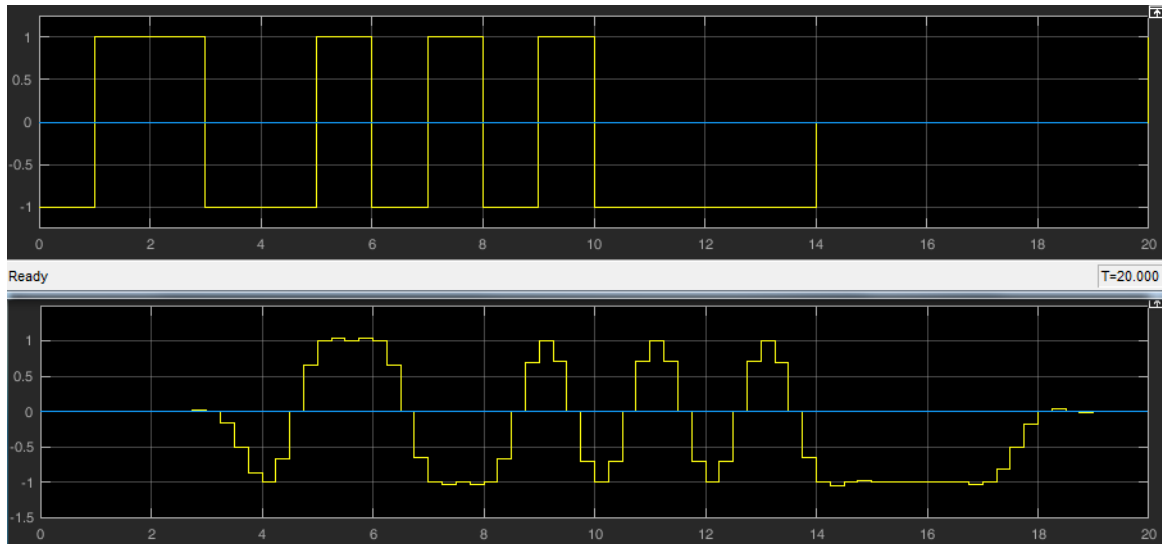


Рисунок 11 Прохождение идеального BPSK baseband сигнала через фильтр «приподнятого косинуса»

На рис. 11 видна задержка сигнала на выходе фильтра, равная $8 \cdot 4/2 = 16$ тактам. В тактовые моменты времени $4nT_d = n$ сигнал равен либо +1, либо -1, т. е. в эти моменты в сигнале содержится информация только о текущем символе и ни о каких других (интерференция отсутствует). На выходе фильтра первый бит 0 кодируется отсчетом -1 в момент времени 4, а последний бит 0 — отсчетом -1 в момент времени 17.

При демодуляции отсчеты принятого аналогового сигнала следует брать как можно ближе к этим идеальным моментам времени, поэтому в приемнике должна быть цепь восстановления тактовой частоты с точностью до фазы, основанная на априорных¹⁷ знаниях о форме **baseband** сигнала.

3 Принцип захвата тактовой частоты и фазы

Для успешного захвата тактовой частоты (**timing recovery**) символы должны претерпевать изменения. Эти изменения гарантируются за счет скремблирования¹⁸ данных псевдослучайными последовательностями, которые генерируются по известным

¹⁶ Binary Phase-Shift Keying, двоичная ФМ.

¹⁷ До опыта известных. Опыт в данном случае — это прием сигнала.

¹⁸ Слабого шифрования, scrambling.

правилам, но свойства которых таковы, что нули и единицы почти равновероятны, пары **00, 01, 10, 11** также почти равновероятны и т. д. до некоторого предела, ограниченного длиной регистра сдвига.

Такие последовательности гарантируют достаточно частую смену символов. Скремблирование для битов — это **сумма по модулю два** информационной последовательности и псевдослучайной последовательности. На приемной стороне делается обратная операция, т. е. вычитание, но по модулю два обе операции эквивалентны, поэтому в приемнике по сути делается та же самая операция. Для успешного декодирования фаза генератора скремблирующей последовательности приемника должна совпадать с фазой генератора передатчика.

Для понимания работы схем захвата тактовой частоты с точностью до фазы, рассмотрим детектор Гарднера (**Gardner detector**), основанный на следующем.

Так как форма идеальных квадратурных сигналов известна (биполярный сигнал), то при наличии смены полярности справедливо предположение (рис. 12).

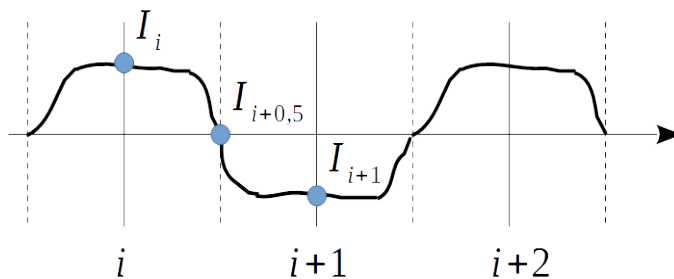


Рисунок 12 К пониманию работы детектора Гарднера

На рис. 12 показано идеальное согласование частоты и фазы тактового генератора. Детектор Гарднера требует **два отсчета на символ**. Это обеспечивается соответствующей выборкой отсчетов в приемнике.

Момент перехода сигнала через ноль является зацепкой для синхронизации, поэтому сигнал детектора ошибки можно взять пропорциональным $I_{i+0,5}$. Но в случае отсутствия перехода через ноль (могут прийти два одинаковых бита) с синхронизацией ничего делать не требуется, поэтому в качестве итогового сигнала детектора ошибки (**timing error detector**) можно взять произведение

$$e = I_{i+0,5}(I_{i+1} - I_i) \quad (6)$$

То же самое касается квадратуры Q . Детектор Гарднера работает и для **QPSK** модуляции.

Далее сигнал ошибки подается на петлевой фильтр и затем на устройство управления фазой тактового генератора (такты АЦП). Возможен вариант (полностью цифровой демодулятор, [1]), когда АЦП приемника, расположенный до согласованного фильтра,

работает с произвольной нерегулируемой фазой тактирования, а подстройка фазы делается устройством **Numerically Controlled Oscillator (NCO)** путем подачи его выходного сигнала на вход интерполятора¹⁹.

Устройство **NCO** генерирует дискретную периодическую последовательность линейно-возрастающих чисел (типа «пилы») с частотой, задаваемой входным числом. Таким образом, при появлении ошибки синхронизации на вход **NCO** подается изменяющийся код, который изменяет частоту следования импульсов и, значит, их фазу (фаза — это интеграл от частоты).

Допустим, тактовые импульсы на рис. 12 имеют верную частоту следования, но идут чуть левее идеального положения, то есть не выровнены по фазе. Тогда сигнал ошибки (6) будет отрицательным, что приведет к уменьшению мгновенной частоты следования чисел на выходе **NCO**, то есть тактовые импульсы начнут притормаживаться. Постепенно сигнал ошибки дойдет до нуля, то есть тактовые импульсы постепенно ускорятся до исходной частоты, но их фаза будет подкорректирована.

Анализ корректировки фазы тактовых импульсов с неверной частотой следования более сложен и требует математического моделирования.

Детектор (6) может обеспечить захват тактовой частоты с неоднозначностью фазы $T_s/2$, то есть 180 градусов (половина длительности символа). В этом случае на сигнальном созвездии появятся точки возле нуля и потребуются изменить индекс выдачи отсчетов на половину символьного интервала.

Подробнее про **timing recovery** можно прочитать в [2, 3].

4 Согласованная фильтрация

Из-за желания минимизировать влияние теплового шума, в приемнике ставят согласованный фильтр (**matched filter**), но так, чтобы межсимвольной интерференции на его выходе при этом не было.

Амплитудный спектр сигнала с выхода формирующего фильтра совпадает с формой АЧХ этого фильтра (на то он и формирующий). Частотная характеристика согласованного фильтра равна спектру сигнала, но комплексно-сопряженному. Если спектры вещественные, то комплексное сопряжение не требуется.

В рассматриваемом случае АЧХ формирующего фильтра берут как корень квадратный из «приподнятого косинуса»; тогда и согласованный фильтр будет иметь такую же АЧХ. Сквозная АЧХ системы связи равна произведению АЧХ блоков, поэтому,

¹⁹ Например, интерполятора Фарроу, http://ru.dsplib.org/content/resampling_lagrange.html

в итоге, в выходном сигнале согласованного фильтра межсимвольной интерференции не будет.

Согласованный фильтр дает максимальное отношение сигнал-шум на выходе, если аддитивная помеха является белым шумом. В этом случае, при фиксированной сигнальной компоненте, дисперсия выходных отсчетов будет минимальна.

В эту дисперсию войдет дисперсия за счет остаточной межсимвольной интерференции плюс дисперсия за счет остаточного теплового шума. Дисперсии (мощности) складываются, так как считается, что помеха типа «межсимвольная интерференция» некоррелирована с помехой типа «белый гауссов шум». Практически отсутствие корреляции означает то, что обе помехи влияют независимо друг от друга, так как у них разная природа.

В дискретном варианте импульсная характеристика согласованного фильтра равна зеркальному отражению сигнала, с которым фильтр согласован

$$h_n = s_{N-n} . \quad (7)$$

Например, для прямоугольного сигнала при $N=3$

$$s_n = (s_0, s_1, s_2, s_3) = (1, 1, 1, 1) , \text{ остальные отсчеты равны нулю и явно не пишутся,}$$

импульсная характеристика совпадет с исходным сигналом

$$h_0 = s_3 = 1, h_1 = s_2 = 1, h_2 = s_1 = 1, h_3 = s_0 = 1 .$$

Тогда отклик фильтра на свой сигнал будет равен

$$y_k = \sum_{n \leq k} h_{k-n} s_n = \sum_{n \leq k} s_{N+n-k} s_n , \quad k \geq 0 . \quad (8)$$

Рассмотрим подробно расчет отклика согласованного фильтра:

$$y_0 = s_3 s_0 = 1 ,$$

$$y_1 = s_2 s_0 + s_3 s_1 = 1 + 1 = 2 ,$$

$$y_2 = s_1 s_0 + s_2 s_1 + s_3 s_2 = 1 + 1 + 1 = 3 ,$$

$$y_3 = s_0^2 + s_1^2 + s_2^2 + s_3^2 = 1 + 1 + 1 + 1 = 4 ,$$

$$y_4 = s_0 s_1 + s_1 s_2 + s_2 s_3 = 1 + 1 + 1 = 3 ,$$

$$y_5 = s_0 s_2 + s_1 s_3 = 1 + 1 = 2 ,$$

$$y_6 = s_0 s_3 = 1 .$$

Из (8) следует, что выход согласованного фильтра совпадает с функцией корреляции сигнала s_n с центром (пиком) в точке $k=N$. Для прямоугольного импульса функция корреляции является треугольным импульсом удвоенной длительности $2N+1$.

В центре функции корреляции имеется максимум, превысить который невозможно, так как в этот момент вычисляется корреляция сигнала самого с собой (сумма квадратов).

Именно этот максимум несет информацию, то есть именно этот отсчет должен быть взят приемником в качестве информационного. Для этого отсчеты с выхода согласованного фильтра пропускают через устройство **downsampling** с коэффициентом прореживания $N+1$.

Если входной сигнал искажен шумом w_n с нулевым средним и дисперсией $\sigma_w^2 = \overline{w_n^2}$

$$z_n = s_n + w_n,$$

то информационный отсчет на выходе согласованного фильтра будет равен сумме сигнальной составляющей и шумовой

$$q = s_0(s_0 + w_0) + s_1(s_1 + w_1) + s_2(s_2 + w_2) + s_3(s_3 + w_3) = y_3 + \sum_n s_n w_n.$$

Среднее значение q равно сигнальной составляющей $\bar{q} = y_3 = E_s / \Delta t = \|s\|^2$, то есть энергии сигнала, деленной на длительность такта, или **квадрату евклидовой нормы** сигнала.

Дисперсия q равна

$$\sigma_q^2 = \overline{(q - \bar{q})^2} = \sum_n s_n w_n \sum_m s_m w_m = \sum_n \sum_m s_n s_m \overline{w_n w_m} = \sum_n s_n^2 \overline{w_n^2} = \sigma_w^2 E_s / \Delta t = \sigma_w^2 \|s\|^2 \quad 20.$$

Здесь использована некоррелированность белого шума

$$\overline{w_n w_m} = \begin{cases} \overline{w_n^2}, & \text{если } n = m \\ 0 & \text{если } n \neq m \end{cases}.$$

Отсюда можно найти отношение сигнал-шум **после** согласованной фильтрации

$$\left(\frac{S}{N} \right)_{\text{out}} = \frac{y_3^2}{\sigma_q^2} = \frac{E_s^2}{\Delta t \sigma_w^2 E_s} = \frac{E_s}{\Delta t \sigma_w^2},$$

а также до неё

$$\left(\frac{S}{N} \right)_{\text{in}} = \frac{\overline{s_n^2}}{\sigma_w^2} = \frac{E_s / \Delta t}{N+1} \frac{1}{\sigma_w^2}.$$

Значит выигрыш согласованного фильтра (его усиление, **gain**)

$$\text{GAIN} = \frac{\left(\frac{S}{N} \right)_{\text{out}}}{\left(\frac{S}{N} \right)_{\text{in}}} = N+1,$$

то есть на $10 \lg(N+1)$ дБ.

В рассматриваемом случае выигрыш составляет 6 дБ.

20 Здесь использовался следующий факт: произведение двух сумм равно двойной сумме произведений.

Если рассмотреть схему (рис. 13), то спектральная плотность мощности входного шума будет связана с мощностью (дисперсией) ограниченного по полосе шума (на выходе АЦП) равенством

$$N_0 = 2\sigma_w^2 \Delta t \quad , \text{Вт/Гц},$$

которое можно получить, если проанализировать рис. 13 и вспомнить про определение мощности шума как произведения спектральной плотности мощности на шумовую полосу.

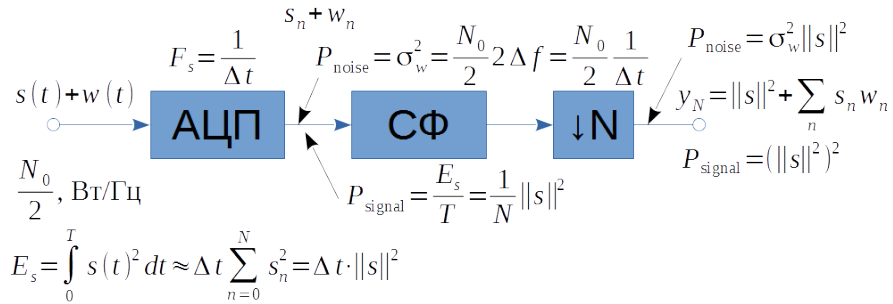


Рисунок 13 Прохождение смеси сигнала и белого шума через АЦП и согласованный фильтр (СФ) с прореживанием отсчетов

5 Пример переноса видеосигнала на промежуточную частоту

Результат переноса видеосигнала с выхода формирующего фильтра типа «корень из приподнятого косинуса» на промежуточную частоту показан на рис. 14. Перед переносом сигнал с выхода формирующего фильтра был интерполирован с коэффициентом 16 и пропущен через цифровой ФНЧ. Так была промоделирована процедура ЦАП, хотя в реальности у ЦАП коэффициент интерполяции равен бесконечности, а ФНЧ — аналоговый.

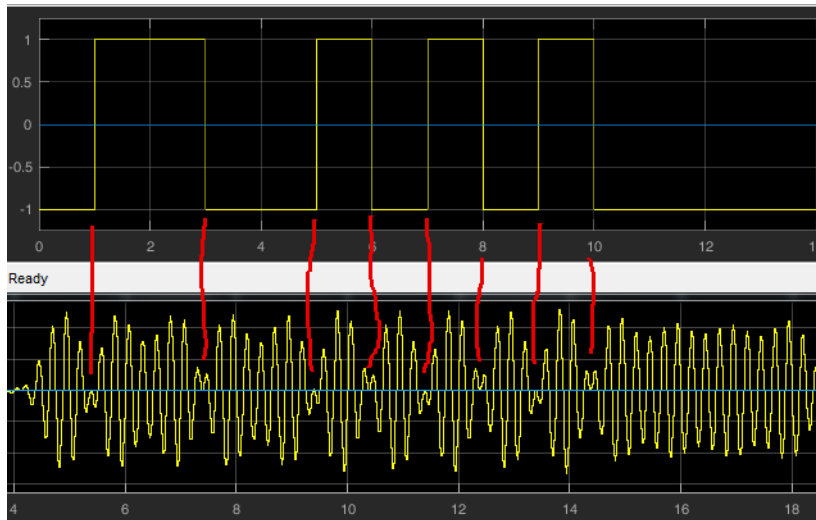


Рисунок 14 Пример переноса видеосигнала с выхода формирующего фильтра на некоторую промежуточную частоту

В моменты смены символа (в данном случае бита) фаза несущей претерпевает скачок на 180 градусов, при этом присутствует паразитная амплитудная модуляция. В моменты, где бит не меняется, скачка фазы нет, однако присутствует небольшая амплитудная модуляция из-за формы переходной характеристики²¹

формирующего фильтра.

Чтобы из принятого полосового сигнала выделить низкочастотные квадратурные компоненты, в приемнике должен быть генератор несущей, частота и фаза которого с некоторой точностью поддерживаются равными частоте и фазе несущей, неявно присутствующей в принятом сигнале. Слежением за параметрами несущей (**carrier recovery**) занимается блок **фазовой автоподстройки частоты**²² (ФАПЧ).

6 Принцип работы ФАПЧ на примере BPSK

Для понимания работы ФАПЧ рассмотрим идеальное созвездие **BPSK**. Здесь начальные фазы несущих, соответствующих битам **0** и **1**, отличаются на 180 градусов.

Задержка сигнала, прошедшего через некоторую линию передачи (усилители, фильтры, среду распространения радиоволн,...), является случайной величиной. Известно, что претерпев

²¹ Переходная характеристика — это отклик фильтра на единичную ступеньку

²² **Phase-Locked Loop, PLL**

задержку на Δt , гармоника с частотой f получает набег фазы $\Delta\varphi = 2\pi f \Delta t$. В этом случае при демодуляции поворот созвездия будет случайным, и отличить **0** от **1** будет невозможно.

Из-за случайного начального поворота созвездия приемник, построенный по схеме универсального квадратурного демодулятора, в большинстве случаев выдаст две ненулевые квадратуры, тогда как по факту **BPSK** имеет лишь одну. Уже исходя из этого наблюдения следует идея схемы ФАПЧ: следует подстраивать частоту и фазу местного генератора так, чтобы вторая квадратура стремилась к нулю. Тогда управляющий сигнал (сигнал ошибки) логично взять пропорциональным этой второй квадратуре.

В случае фиксированного частотного рассогласования, $\Delta f = \text{const}$, созвездие в приемнике будет вращаться со скоростью, пропорциональной величине Δf . В мобильных системах связи такие частотные рассогласования обязательно должны отслеживаться системой ФАПЧ; мало того, должны отслеживаться и медленные изменения частотного рассогласования, возникающие когда объекты движутся с радиальным ускорением друг относительно друга.

Заметим, что детектор ФАПЧ можно делать и в цифровом виде, так как важно лишь обеспечить условие $Q \rightarrow 0$, где Q — оцифрованная квадратура.

Квадратуры получаются обычным образом: перемножением входного **BPSK** сигнала на две квадратуры опорного генератора контура ФАПЧ. Детектор ФАПЧ можно сделать в виде умножителя выделенных квадратур: если хотя бы одна квадратура станет равной нулю, то и всё произведение (сигнал ошибки, управляющий сигнал) станет равным нулю. Эта идея лежит в основе **петли Костаса (Costas loop)**, рис. 15²³.

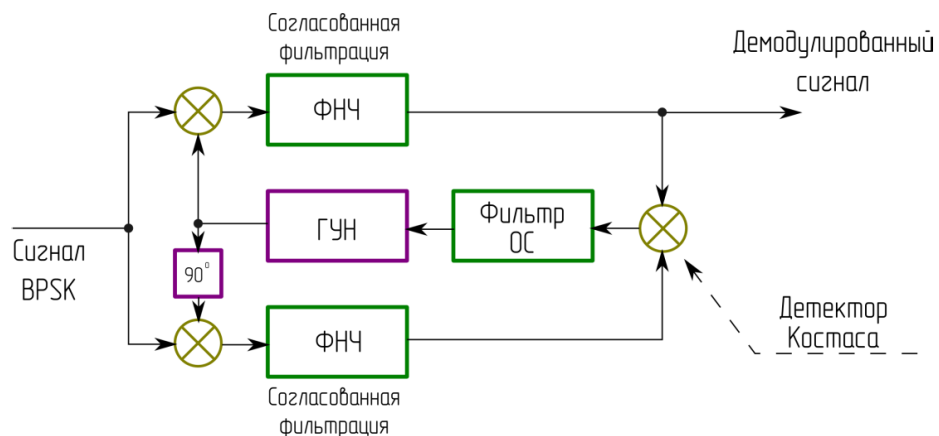


Рисунок 15 Петля Костаса для выравнивания частоты и фазы местного генератора (ГУНа)

23 Рисунок взят из [1]; пособие [1] рекомендуется для просмотра.

Мгновенная частота сигнала с ГУНа²⁴ по определению линейно зависит от входного напряжения $s_{LF}(t)$, значит фаза является интегралом от этого напряжения

$$s_{vco}(t) = A_{vco} \cos \left(2\pi f_0 t + 2\pi k_c \int_{-\infty}^t s_{LF}(t) dt + \varphi_{vco} \right), \quad (9)$$

где k_c — чувствительность ГУНа, Гц/В, f_0 — его центральная частота.

Обозначение «Фильтр ОС» на рис. 15 по сути означает петлевой фильтр, хотя автор рисунка под сокращением подразумевал «Фильтр обратной связи». Петлевой фильтр является фильтром нижних частот. Порядок фильтра определяет тип возмущений, которые способна отслеживать схема ФАПЧ: скачки фазы, скачки частоты, линейное изменение частоты и т. п.

После захвата схема ФАПЧ может работать в двух устойчивых состояниях, которые соответствуют созвездию, допустим, с нулевой Q составляющей и с ненулевой I составляющей; причем где на оси I бит **0**, а где бит **1**, слева или справа — не определено. В этом и кроется суть двузначной работы ФАПЧ для **BPSK**: в процессе захвата перевесит либо правое вращение, либо левое, но в обоих случаях условие $Q=0$ будет выполнено.

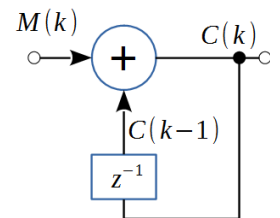
Проблему с возможной инверсией битов решают с помощью дифференциального кодирования (**DBPSK**²⁵), когда передается не сам бит, а разность между входным текущим битом и выходным предыдущим.

Да, биты можно вычитать друг из друга, получая снова биты. Этой операции соответствует **исключающее ИЛИ (XOR)**. Вообще, сумма и разность для битов являются эквивалентными операциями. Это легко понять, учтя, что **XOR** есть **сумма по модулю два**:

$$\begin{aligned} 1+1=2=0, 1+0=0+1=1, 0+0=0. \\ 1-1=0, 0-1=-1=1, 1-0=1, 0-0=0. \end{aligned}$$

В случае инверсной работы демодулятора дифференциальное декодирование, построенное по принципу разности между текущим битом и предыдущим, даст тот же самый результат, так как $1-1=0-0=0$.

Дифференциальный кодер — рекурсивный фильтр первого порядка, а декодер — трансверсальный (фильтр с конечной импульсной характеристикой, КИХ-фильтр). На рисунке справа показана схема кодера.



Предлагается построить схему декодера и проверить кодирование-декодирование некоторой случайной последовательности битов, при условии нулевого начального состояния кодера-декодера.

²⁴ ГУН — генератор, управляемый напряжением. **VCO** – **Voltage-controlled oscillator**.

²⁵ **Differential BPSK**.

Помимо петли Костаса есть схема Пистолькорса, где **BPSK** сигнал возводится в квадрат, а ГУН работает на удвоенной несущей частоте. Делитель частоты на два дает неоднозначность фазы восстановленной несущей 180 градусов.

Рассмотрим подробнее²⁶ работу петли Костаса на примере **BPSK** сигнала с одинаковыми битами (все нули или все единицы). Такой сигнал можно записать как

$$s_0(t) = A_0 \cos(2\pi f_0 t + \varphi_0) \quad (10)$$

где неизвестными являются фаза и частота, причем фаза полностью неизвестна (известно, что она случайная и имеет равномерное распределение на отрезке от 0 до 2π), а частота известна частично как $f_0 = f_{\text{transmit}} + \Delta f$, где f_{transmit} — частота передатчика, а Δf — частотный сдвиг, например, за счет эффекта Доплера²⁷.

Напряжение на выходе ГУН и фазовращателя на 90 градусов имеет вид

$$\begin{aligned} s_{\text{vco}}(t) &= A_{\text{vco}} \cos(2\pi f_{\text{vco}} t + \varphi_{\text{vco}}) \\ s_{90}(t) &= -A_{\text{vco}} \sin(2\pi f_{\text{vco}} t + \varphi_{\text{vco}}) \end{aligned} \quad (11)$$

Тогда сигналы на выходе двух умножителей перед ФНЧ будут равны

$$\begin{aligned} &A_0 A_{\text{vco}} \cos(2\pi f_0 t + \varphi_0) \cos(2\pi f_{\text{vco}} t + \varphi_{\text{vco}}) = \\ &\frac{A_0 A_{\text{vco}}}{2} \cos[2\pi(f_0 - f_{\text{vco}})t + (\varphi_0 - \varphi_{\text{vco}})] - \frac{A_0 A_{\text{vco}}}{2} \sin[2\pi(f_0 + f_{\text{vco}})t + (\varphi_0 + \varphi_{\text{vco}})] \quad (12) \\ &- A_0 A_{\text{vco}} \cos(2\pi f_0 t + \varphi_0) \sin(2\pi f_{\text{vco}} t + \varphi_{\text{vco}}) = \\ &\frac{A_0 A_{\text{vco}}}{2} \sin[2\pi(f_0 - f_{\text{vco}})t + (\varphi_0 - \varphi_{\text{vco}})] - \frac{A_0 A_{\text{vco}}}{2} \sin[2\pi(f_0 + f_{\text{vco}})t + (\varphi_0 + \varphi_{\text{vco}})] \end{aligned}$$

ФНЧ проектируется так, чтобы подавить сигнал на суммарной частоте и оставить сигнал на разностной частоте, причем так как доплеровский сдвиг может изменяться во времени, то ФНЧ должен пропускать эти, пусть и медленные, но изменения.

Сигнал на выходе умножителя-детектора равен

$$s_{\text{detector}}(t) = K_d A_0 A_{\text{vco}} \sin[4\pi(f_0 - f_{\text{vco}})t + 2(\varphi_0 - \varphi_{\text{vco}})] \quad (13)$$

где $K_d = 1/8$ — коэффициент передачи детектора для петли Костаса.

Из-за появившейся двойки в полной фазе сигнала петля Костаса позволяет снять манипуляцию фазы на 180 градусов (**BPSK** сигнал). По сути, выражение (13) является характеристикой фазового детектора.

Допустим, разность начальных фаз равна нулю, тогда если частота ГУНа несколько больше частоты входного сигнала, то сигнал с выхода детектора меньше нуля, так что ГУН под управлением этого сигнала начинает понижать частоту, так как его частота пропорциональна входному напряжению. И наоборот.

²⁶ Для этого использовался **Simulink**

²⁷ Для более точного описания процессов следует учесть случайное дрожание фазы — фазовый шум

Если разность начальных фаз не равна нулю, а разность частот равна нулю, то если фаза ГУНа будет больше фазы входного сигнала (сигнал ГУНа чуть опережает, то есть находится чуть левее на оси времени), то отрицательный сигнал с детектора скажет уменьшить частоту ГУНа, таким образом тормозя сигнал ГУН²⁸ и подстраивая его к фазе входного сигнала.

Если разность начальных фаз не равна нулю и разность частот не равна нулю, то анализировать ФАПЧ становится сложнее, и в этом случае лучше пользоваться программами моделирования.

Петлевой фильтр строится как пропорционально-интегрирующее звено (рис. 16).

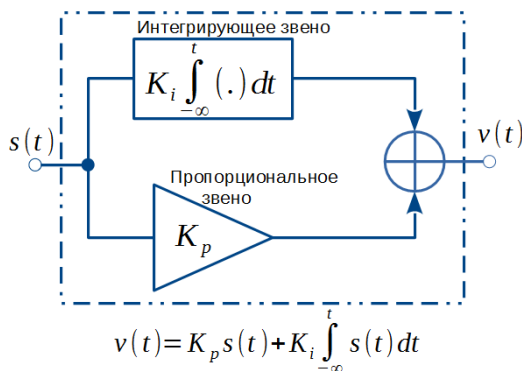


Рисунок 16 Пропорционально-интегрирующее звено

Операторный коэффициент передачи пропорционально-интегрирующего звена равен

$$F(p) = K_p + \frac{K_i}{p} \quad (14)$$

Рассмотрим для начала петлевой фильтр первого порядка, то есть выключим интегрирующее звено. В этом случае фильтр вырождается в усилитель с некоторым коэффициентом передачи (**gain**). Фильтр первого порядка позволяет отслеживать скачки по фазе, но не позволяет отслеживать скачки по частоте. Покажем это.

Предположим, что рассогласования по частоте нет, а по фазе имеется случайное рассогласование на интервале от 0 до 2π . Примем чувствительность ГУНа — $k_c = 4$ Гц/В, коэффициент передачи петлевого фильтра (**loop filter**) — $K_p = 200$. Частота входного сигнала 1 кГц, частота ГУНа при нулевом входном напряжении — 1 кГц.

На рис. 17 показаны сигналы с выхода детектора Костаса, выхода ФНЧ, с выхода ГУНа и входной сигнал. ФАПЧ в этом случае захватила фазу в синфазном состоянии сигналов. При этом сигнал с выхода детектора нулевой (стремится к нулю). Анализируя поведение сигнала детектора, можно сделать вывод, что ФАПЧ не может сразу захватить фазу и присутствует затухающий колебательный процесс. Сигнал детектора (после усилителя) доходит до $\pm(20...30)$ вольт, то есть частота ГУНа в процессе захвата

²⁸ Фаза — это интеграл от частоты

отклоняется от номинальной на $\pm(80...120)$ Гц и должна обеспечиваться перестройка ГУНа (10...15) %.

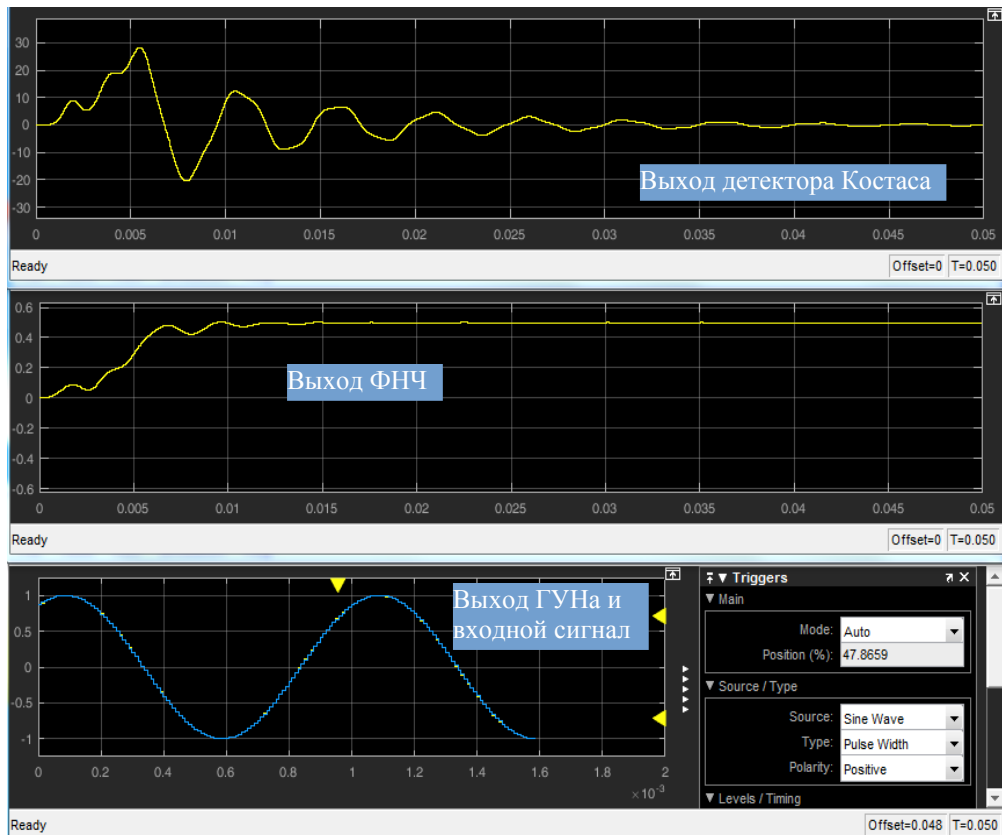
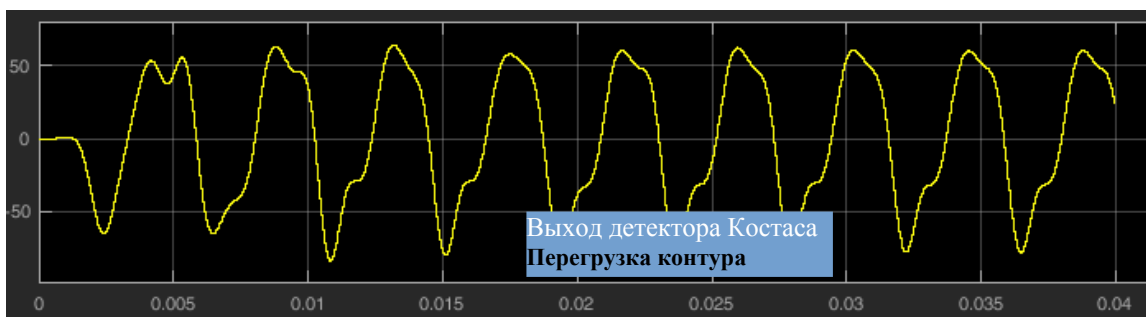


Рисунок 17 Сигналы петли Костаса: синфазный захват фазы при нулевом начальном рассогласовании по частоте

На рис. 18 показан случай инверсного захвата фазы, при этом сигнал с выхода ФНЧ инвертировался — это и есть обратная работа демодулятора, когда вместо одной полярности принимается другая — противоположная (не важно, плюс или минус).

Каков произойдет захват фазы — противофазный или синфазный — зависит от начального рассогласования фаз (рис. 19, рис. 20).

Уменьшение коэффициента усиления детектора приведет к более медленному процессу захвата, а увеличение — к более быстрому с перспективой перегрузки контура и невозможностью захвата:



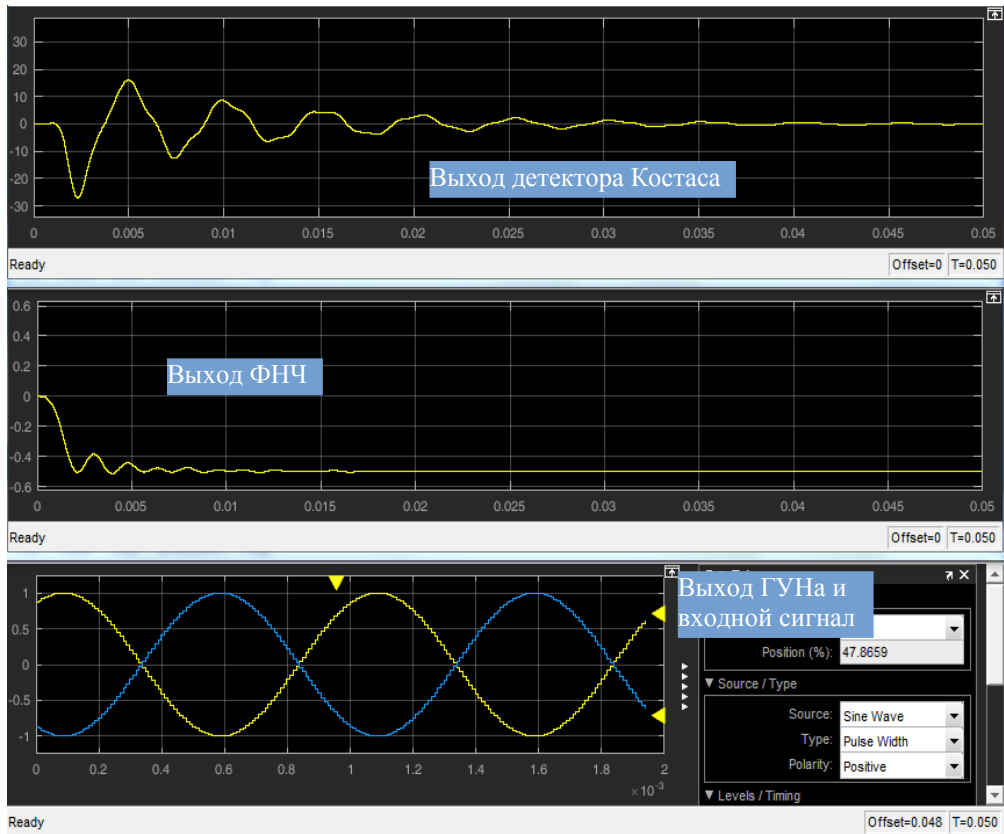


Рисунок 18 Сигналы петли Костаса: противофазный захват фазы при нулевом начальном рассогласовании по частоте

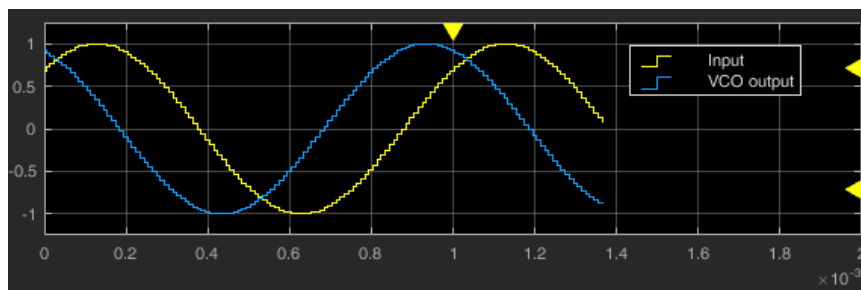


Рисунок 19 Начальное рассогласование фаз ГУНа и входного сигнала, приводящее к синфазному захвату (слиянию, рис. 17)

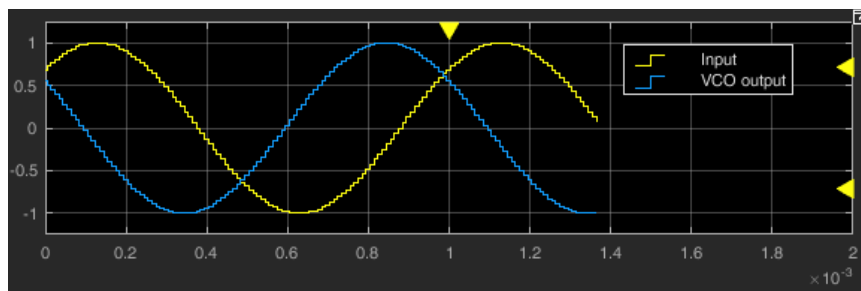


Рисунок 20 Начальное рассогласование фаз ГУНа и входного сигнала, приводящее к противофазному захвату (противопоставлению, рис. 18)

Если есть начальное рассогласование по частоте, то начальная фаза (при выключенном интеграторе в петлевом фильтре) до нуля не синхронизируется: будет остаточное рассогласование по фазе, тем меньшее, чем больше коэффициент усиления детектора. Это недостаток петлевого фильтра первого порядка.

Если так уменьшать коэффициент усиления, что рассогласование по фазе будет 45 градусов, то дальнейшее снижение приведет к невозможности захвата частоты: рассогласование по фазе всегда будет больше 45 градусов и система не сможет стабилизироваться. Для рассматриваемого примера критический коэффициент усиления равен 40 при частотном рассогласовании 20 Гц.

Формула для критического коэффициента усиления может быть получена из (13) путем подстановки разности фаз 45 градусов при нулевой разности частот, так как частота уже захвачена

$$K_{p \text{ critical}} = \frac{\Delta f}{k_c K_d} , \quad (15)$$

где Δf — начальное рассогласование по частоте.

Формула для остаточного рассогласования по фазе может быть найдена из (13) путем обнуления **текущего** частотного рассогласования (амплитуды сигналов, A_{vco} , A_0 , взяты по единице)

$$\begin{aligned} k_c K_p K_d \sin(2\varphi_{\text{limit}}) &= \Delta f , \\ \varphi_{\text{limit}} &= \frac{1}{2} \arcsin\left(\frac{\Delta f}{k_c K_p K_d}\right) . \end{aligned} \quad (16)$$

Если амплитуды не равны единице, то вместо $K_p K_d$ следует брать $K_p K_d A_0 A_{vco}$, то есть напряжение на входе ГУНа U_{vco}

$$\varphi_{\text{limit}} = \frac{1}{2} \arcsin\left(\frac{\Delta f}{k_c U_{vco}}\right) = \frac{1}{2} \arcsin\left(\frac{\Delta f}{k_c K_p U_{\text{detector}}}\right) . \quad (17)$$

Остаточное рассогласование фазы приводит к уменьшению отношения сигнал шум в $\cos^2 \varphi_{\text{limit}}$ раз или на $20 \log_{10} |\cos \varphi_{\text{limit}}|$ дБ. Например, рассогласование фазы 45 градусов дает проигрыш 3 дБ (два раза по мощности), рис. 21.

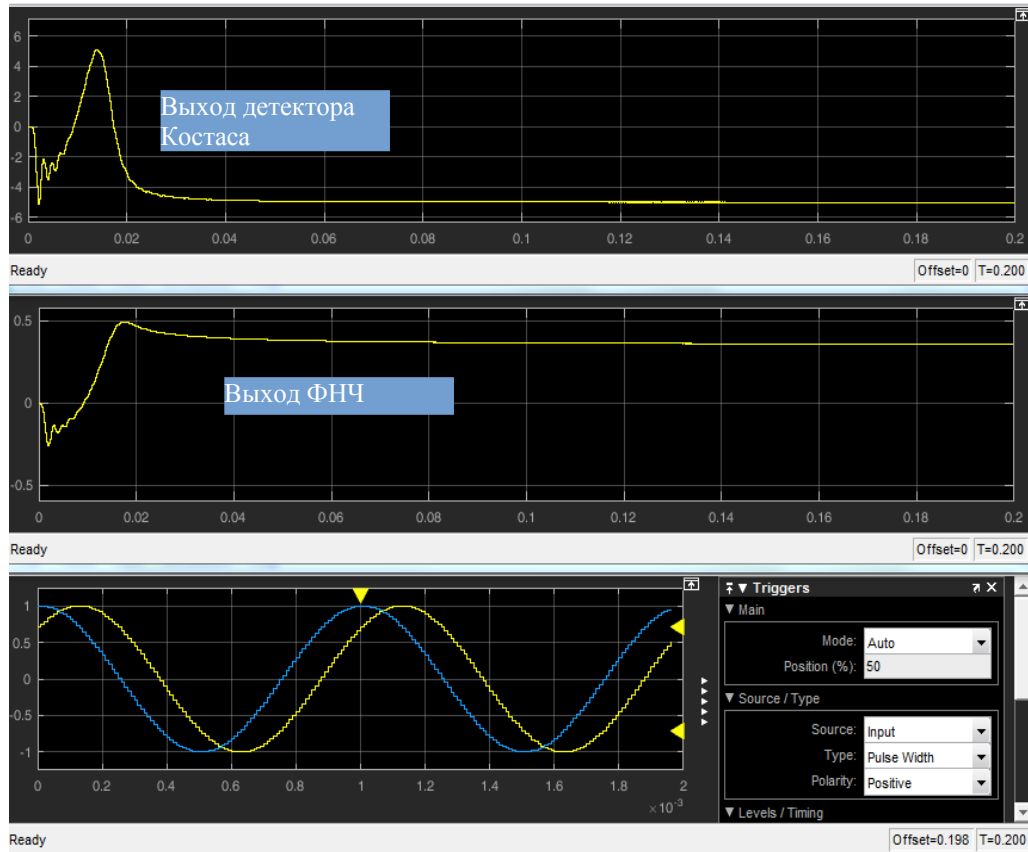


Рисунок 21 Остаточное рассогласование по фазе 45 градусов (чуть меньше) при критическом коэффициенте усиления: система медленно, но вышла на рабочий режим (захватила частоту)

Рассмотрим петлевой фильтр второго порядка, который позволяет отследить скачки по частоте (важно для компенсации эффекта Доплера).

В этом случае уравнение (13) может быть записано в виде

$$s_{\text{detector}}(t) = A_0 A_{\text{vco}} K_d \sin \left[4\pi(f_0 - f_{\text{vco}})t + 2(\varphi_0 - \varphi_{\text{vco}}) + 4\pi k_c \int_{-\infty}^t s_{\text{LF}}(t) dt \right] . \quad (18)$$

То, что стоит в качестве аргумента синуса, является ошибкой слежения

$$e(t) = 4\pi(f_0 - f_{\text{vco}})t + 2(\varphi_0 - \varphi_{\text{vco}}) + 4\pi k_c \int_{-\infty}^t s_{\text{LF}}(t) dt , \quad (19)$$

которая в установившемся режиме в идеале должна быть равна нулю.

Сигнал с выхода петлевого фильтра (**Loop Filter**) можно записать так

$$s_{\text{LF}}(t) = K_p s_{\text{detector}}(t) + K_i \int_{-\infty}^t s_{\text{detector}}(t) dt . \quad (20)$$

Объединяя три уравнения, можно записать дифференциальное уравнение относительно ошибки слежения для петли Костаса

$$\frac{d^2}{dt^2} e(t) - A \cos e(t) \frac{d}{dt} e(t) - B \sin e(t) = 0 , \quad (21)$$

$$A = 4\pi k_c K_p K_d (A_0 A_{\text{vco}})^2 , \quad B = 4\pi k_c K_i K_d (A_0 A_{\text{vco}})^2 .$$

Данное уравнение является нелинейным дифференциальным уравнением второго порядка. Нелинейность связана с синусом и косинусом ошибки слежения. Для его решения требуется задать два начальных условия: ошибку слежения в момент $t=0$ и производную ошибки слежения в момент $t=0$:

$$e(0)=2(\varphi_0-\varphi_{vco}) \quad , \quad \frac{de(t)}{dt}(t \rightarrow 0)=4\pi(f_0-f_{vco}) \quad . \quad (22)$$

Эти начальные условия выведены из (19) в предположении, что сигнал на выходе петлевого фильтра в начальный момент времени равен нулю (логично, если все цепи в фильтре разряжены). Таким образом, начальное значение ошибки слежения зависит от начального рассогласования по фазе, а начальная скорость изменения ошибки (её крутизна) зависит от начального рассогласования по частоте. Удвоение начальной ошибки и её крутизны связано с платой за теоретическую нечувствительность петли ФАПЧ к смене фазы **BPSK**-импульсов.

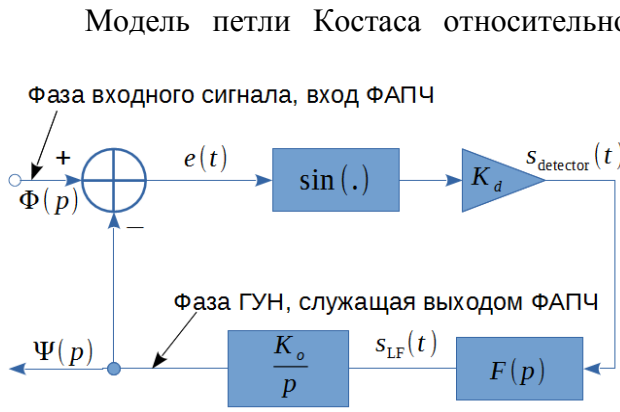


Рисунок 22 Модель петли Костаса относительно фаз, $K_o=2\pi k_c$

Здесь принято упрощение $A_{vco}=A_0=1$, которое требует нормировки амплитуды входного сигнала и сигнала ГУНа.

Если рассматривать линеаризованную модель петли Костаса, то блок синуса убирается, что справедливо для малых фаз, так как при этом $\sin x \approx x$. В этом случае можно найти операторный коэффициент

передачи по фазе для такой петли

$$\Psi(p)=[\Phi(p)-\Psi(p)]K_d F(p)\frac{K_o}{p} \quad ,$$

$$H(p)=\frac{\Psi(p)}{\Phi(p)}=\frac{K_d F(p)\frac{K_o}{p}}{1+K_d F(p)\frac{K_o}{p}}=\frac{p K_d K_o K_p+K_d K_o K_i}{p^2+p K_d K_o K_p+K_d K_o K_i} \quad . \quad (23)$$

Для удобства вводят обозначения

$$\omega_p^2=K_d K_o K_i \quad , \quad 2\zeta\omega_p=K_d K_o K_p \quad , \quad (24)$$

которые дают компактную запись коэффициента передачи петли ФАПЧ

$$H(p)=\frac{2\zeta\omega_p p+\omega_p^2}{p^2+2\zeta\omega_p p+\omega_p^2} \quad . \quad (25)$$

Частота ω_p называется собственной частотой контура. Колебание с такой частотой возникает в петле ФАПЧ при нулевом (или малом) коэффициенте усиления K_p

пропорционального звена. В этом случае демпфинг-фактор ζ равен нулю. Демпфинг-фактор как раз и отвечает за меру затухания колебаний: чем он больше, тем быстрее затухают колебания, но тем медленнее контур отслеживает изменения фазы входного сигнала.

Подставляя в (25) $p = j\omega$, можно найти эквивалентную шумовую полосу контура ФАПЧ

$$B_{\text{экр.}} = 2 \int_0^{\infty} |H(j\omega)|^2 d\omega, \quad (\text{двусторонняя полоса}),$$

$$B_{\text{экр.}} = 2\pi\omega_p \left(\zeta + \frac{1}{4\zeta} \right), \quad \text{рад/с.} \quad (26)$$

Время захвата сигнала **примерно** можно оценить по формуле

$$t_{\text{захвата}} \approx (3..4) \frac{K_p}{K_i} = (3..4) \frac{2\zeta}{\omega_p}, \quad \text{с.} \quad (27)$$

Эта оценка делается в мало-сигнальном приближении (21) $\sin x \approx x$, $\cos x \approx 1$ и решении соответствующего линейного дифференциального уравнения.

Пусть петля моделируется в дискретном времени с шагом $T_s = 10 \text{ мкс}$ и её параметры заданы следующим образом

$$A_0 = A_{\text{vco}} = 1, \quad K_p = 200, \quad K_i = 0,3/T_s = 0,3 \cdot 10^5 \text{ 1/с}, \quad k_c = 4 \text{ Гц/В}. \quad (28)$$

Тогда дифференциальное уравнение (21) запишется так

$$\frac{d^2}{dt^2} e(t) - 1256 \cos e(t) \frac{d}{dt} e(t) - 188500 \sin e(t) = 0. \quad (29)$$

Резонансная частота в этом случае равна 50 Гц, демпинг-фактор — 1, а время захвата оценивается как (0,02...0,03) с.

Задав начальные параметры

$$e(0) = 1 \quad (\text{расстройка } 0,5 \text{ радиан!}), \quad \frac{de(t)}{dt}(0) = 2\pi 80 \quad (\text{расстройка } 40 \text{ Гц!}),$$

и решая это уравнение численно (**Mathcad**), получим кривую ошибки слежения (рис. 23). При этом частота входного сигнала равна 1 кГц, а начальная частота ГУНа — 1040 Гц. Время захвата примерно соответствует оцененному по (27).

Такое поведение ошибки слежения соответствует идеальному контуру, так как не существует такого ФНЧ, который мог бы полностью подавить удвоенную частоту со своим спектром и пропустить нулевую частоту со своим спектром и при этом не иметь задержки.

```

Given
 $y''(t) - 1256 \cdot \cos(y(t)) \cdot y'(t) - 188500 \cdot \sin(y(t)) = 0$ 
 $y(0) = 1$   $y'(0) = 2 \cdot \pi \cdot 80$ 
 $y := \text{Odesolve}(t, 0.03)$ 

```

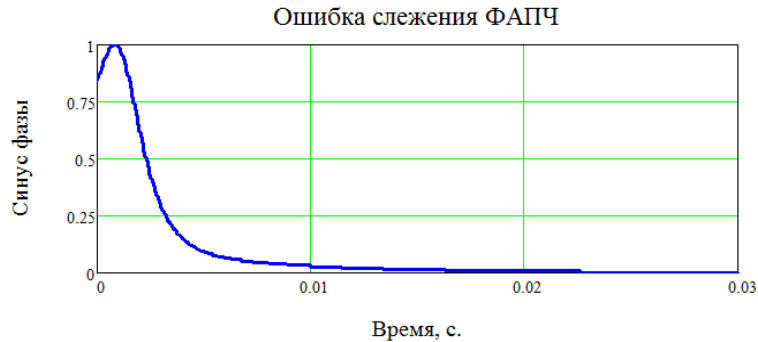


Рисунок 23 Пример расчета ошибки слежения нелинейного контура ФАПЧ с **нереализуемым** ФНЧ; резонансная частота контура 50 Гц, демпинг-фактор 1

На рис. 24 показан пример ошибки слежения нелинейного контура с реализуемым ФНЧ рекурсивного типа (частота дискретизации 25 кГц, частота среза — 1 кГц, частота заграждения — 2 кГц по -70 дБ, неравномерность в полосе пропускания 2 дБ, рис. 25, рис. 26), причем параметры (29) остались неизменными. Остаточная ошибка слежения $5 \cdot 10^{-4}$ радиан; она зависит от степени подавления суммарной гармоники 2 кГц ФНЧ. Пример модели **Simulink** смотри в приложении А.

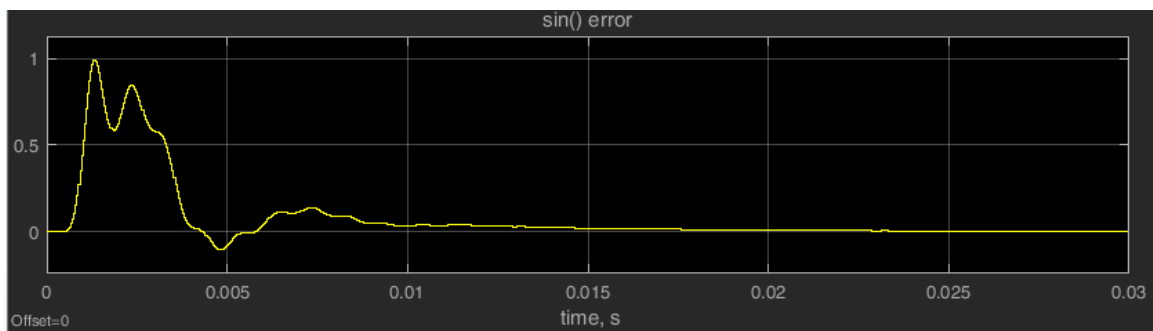


Рисунок 24 Пример ошибки слежения нелинейного контура ФАПЧ с **реализуемым** ФНЧ; резонансная частота контура 50 Гц, демпинг-фактор 1

Из сравнения рис. 23 и рис. 24 следует, что характеристики петли ФАПЧ с реализуемым ФНЧ заметно отличаются от характеристик петли с **нереализуемым** ФНЧ.

Так как была задана начальная расстройка частоты ГУНа и входного сигнала 40 Гц, то в установившемся режиме сигнал на выходе петлевого фильтра должен быть равен 10 В (чувствительность ГУНа 4 Гц/В), что и наблюдается на рис. 27 под легендой «**Loop-Filter output**». При этом сигнал с выхода **Gain** (с усилителя K_p пропорционально-интегрирующего звена) равен нулю, то есть в установившемся режиме компенсация частотного ухода держится за счет выхода интегратора: ошибка слежения равна нулю, то есть на входе

петлевого фильтра ноль, но сигнал на его выходе не равен нулю (интегратор помнит накопленное напряжение). Поэтому наличие интегратора позволяет отслеживать скачки частоты. Практически же частота «скачет» только в начальный момент времени, а затем непрерывно меняется во времени. Конечно, частота может измениться очень быстро (но непрерывно!), так, что это может восприняться системой почти как скачок.

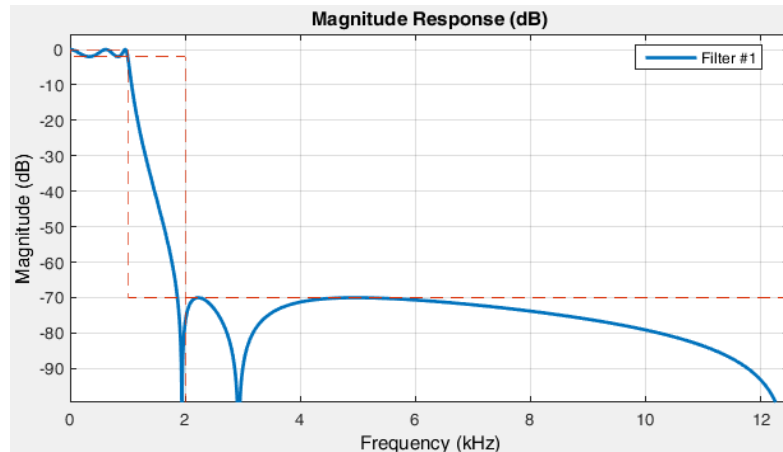


Рисунок 25 АЧХ рекурсивного фильтра нижних частот для подавления суммарной гармоники 2 кГц в петле Костаса

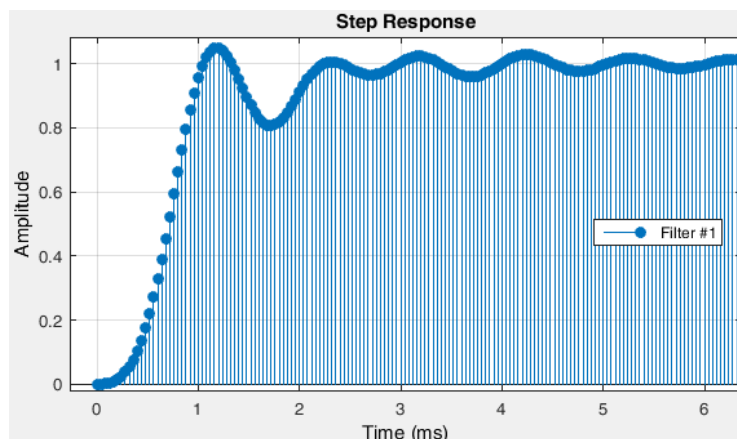
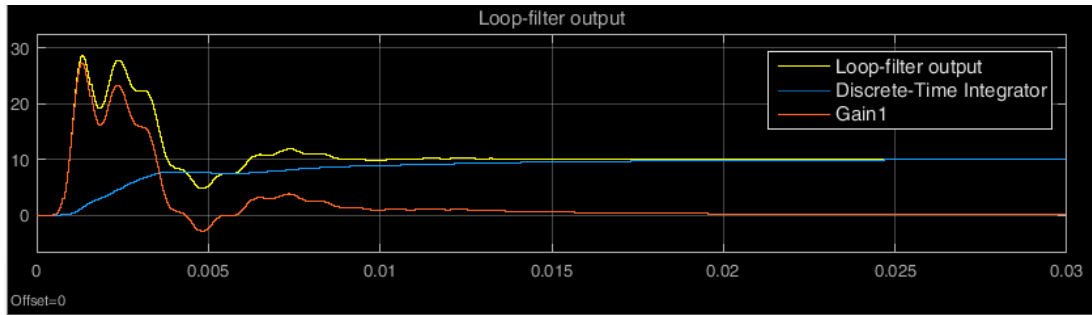


Рисунок 26 Переходная характеристика фильтра нижних частот петли Костаса

Результат выравнивания частоты и фазы показан на рис. 28. Здесь произошел синфазный захват фазы. Время установления около 2 мс.



Уровни в вольтах, время в секундах

Рисунок 27 Сигналы петлевого фильтра нелинейного контура ФАПЧ с реализуемым ФНЧ

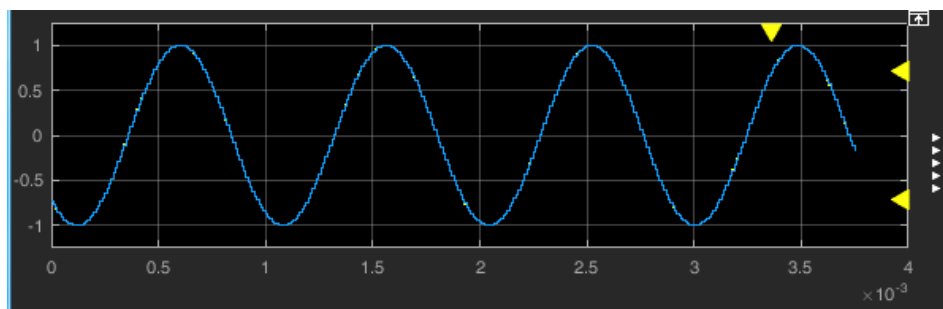


Рисунок 28 Результат выравнивания частоты и фазы сигнала с частотой 1 ГГц и начальной расстройкой 40 Гц и 0,5 рад

В системах связи контурами второго порядка и ограничиваются.

7 Источники

1. Боев Н.М. Системы связи. Подвижные системы связи. Лекции: учебно-методическое пособие [Электронный ресурс]. — Красноярск: Сиб. федер. Ун-т. — 2013. 60 с.

http://radio-systems.org/sites/default/files/files/guidance/Boev-Digital_communications_lectures_2013.pdf

2. Mervi Berner. Timing and carrier recovery. [Электронный ресурс]. — 2005.

http://www.comlab.hut.fi/opetus/333/2004_2005_slides/CarrierTimingRecovery.pdf

3. Michael Rice. Digital Communications. A Discrete-Time Approach. — 2009. (имеется в Интернет в формате pdf на английском языке).

ПОСЛЕСЛОВИЕ

За рамками данного пособия остались схемы автоматической регулировки усиления в приемнике. Без них работа систем радиосвязи (а тем более мобильных) невозможна, так как мощность принимаемого сигнала изменяется со временем. К этому эффекту приводит многолучевое распространение радиоволн (интерференция), а также прохождение радиоволн через непостоянные объекты, такие как дождь, снег, туман и т. п.

Также не рассматривалось влияние фазового шума генераторов и эквалайзеров, корректирующих частотную характеристику линии передачи сигнала (волновод, кабель) или трассы распространения радиоволн.

Качественное усвоения материала требует не только чтения пособия и других книг, но и понимания того, что читаешь. Помимо понимания, требуется самостоятельная работа: проверка формул на бумажке, проверка расчетов, единиц измерения. Особым комплексным эффектом обладает сборка исследуемой модели, например, в **Simulink** и наблюдение за работой модели: изменение параметров и анализ процессов в разных контрольных точках, чтение справки (**help**) к блокам модели, анализ готовых примеров (**examples**).

Важную роль при обучении играет технический английский язык, так как множество статей и справочных материалов дано на английском языке.

Побольше работайте с моделями **Simulink** и осваивайте английские термины!