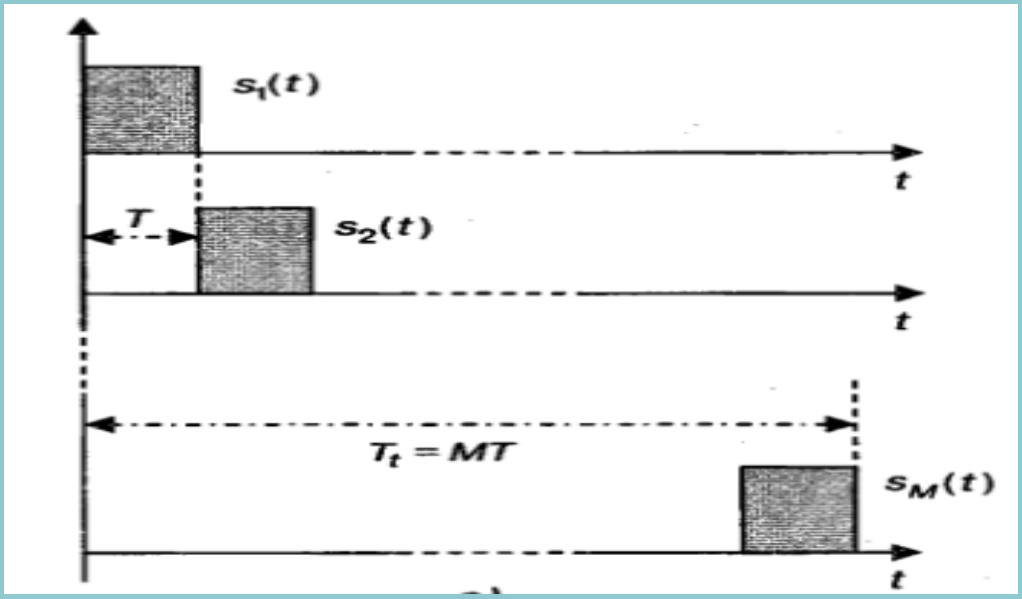
Радиоизмерения и мониторинг в системах и сетях радиосвязи

Антипин Б.М.

доц. каф. «Телевидения и Метрологии» СПбГУТ

1. Ортогональное кодирование.
1.1. Кодирование временным сдвигом

Рассмотрим *M* сигналов, приведенных на рисунке , занимающих совместно весь временной интервал T_t. При длительности сигнала, не превышающей $T = T_t/M$, и временном сдвиге между сигналами, следующими друг за другом, не меньшем длительности сигнала, подобное кодирование временным сдвигом формирует семейство ортогональных сигналов.



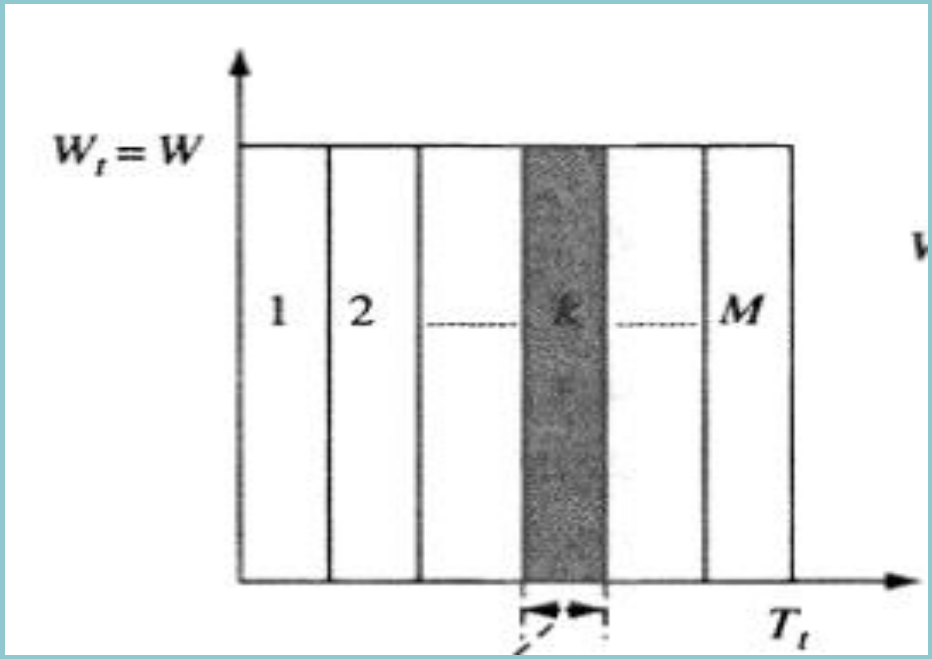
Полоса частот W, занимаемая каждым из сигналов, оценивается величиной, обратной их длительности, и все сигналы могут, не нарушая условия ортогональности, располагаться в одной и той же полосе $W = W_t$

Следовательно, максимальное число ортогональных сигналов этого типа в пределах заданного частотно-временного ресурса W_t, T_t оценивается как $M = T_t/T = W_tT_t$, т. е. совпадает с размерностью сигнального пространства $n_s = W_tT_t$.

Желание иметь большое число сигналов M >> 1 может быть удовлетворено лишь при большом произведении $W_tT_t = M$, что может быть принято за требование **широкополосности сигналов**.

Однако для каждого индивидуального сигнала частотно-временное произведение $WT = WT_t = W_tT_t./M = 1$, ***так что данные сигналы никоим образом не относятся к разряду широкополосных***

Представим общий частотно-временной ресурс прямоугольником в координатной плоскости время-частота со сторонами T_t и W_t соответственно. При этом кодирование временным сдвигом означает попросту разбиение этого ресурса на М вертикальных полос, каждая из которых отводится некоторому отдельному сигналу. Ортогональность при передаче обеспечивается жестким распределением временного ресурса между сигналами, каждый из которых занимает весь отведенный системе частотный ресурс.



Распределение частотного ресурса при ортогональным кодированием временным сдвигом

Рассмотренная схема ортогональной передачи может показаться привлекательной с точки зрения практики благодаря ее видимой простоте. Недостатки ее, однако, также довольно очевидны и должны соответствующим образом учитываться.

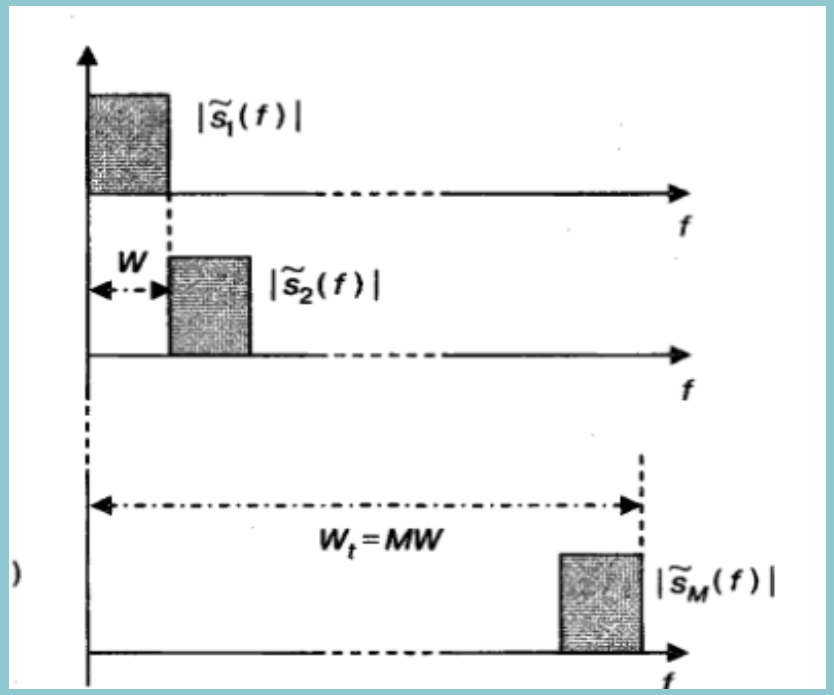
1. **Необходима точная синхронизация**, поскольку флюктуации временного положения сигналов потенциально способны вызвать перекрытие последних, нарушающее их ортогональность. В связи с этим в системах, природа которых в принципе предполагает непостоянство задержек сигналов, приходится отделять сигналы друг от друга защитными паузами, уменьшая возможное число сигналов относительно теоретического максимума, т. е. ухудшая спектральную эффективность.
2. Другим важным аспектом **является значение пик-фактора ν.** Поскольку отдельный сигнал занимает лишь M-ю часть доступного временного ресурса, средняя излучаемая мощность в M раз меньше пиковой и $\nu = M > 1$.

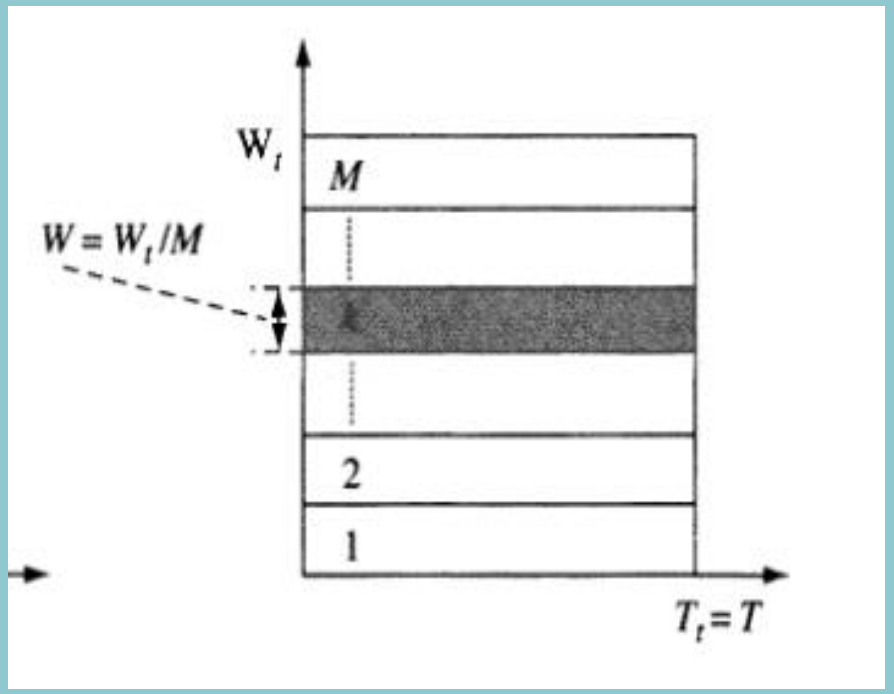
1.2. Кодирование частотным сдвигом

На основании дуальности времени и частоты или теоремы Парсеваля скалярные произведения сигналов $u(t)$, $v(t)$ и их спектров $\acute{u}(f)$, $\acute{u}(f)$ совпадают:

$$(\mathbf{u}, \mathbf{v}) = \int_{-\infty}^{\infty} u(t)v(t)\,dt = \int_{-\infty}^{\infty} \tilde{u}(f)\tilde{v}^*(f)\,df = (\tilde{\mathbf{u}}, \tilde{\mathbf{v}}),$$

что позволяет механически перенести только что обсужденную схему в частотную область. При полном перекрытии сигналов во времени ($T = T_t$) каждый из них занимает полосу не менее $W = 1/T_t$. При этом максимальное число ортогональных сигналов, полученных сдвигом спектра, оказывается вновь равным $M = W_t/W = W_tT_t = n_s$.





Как и в предыдущем случае, общий ресурс снова нарезается на полосы, теперь, однако, расположенные горизонтально, так что каждый сигнал использует весь временной ресурс T_t и лишь М-ю часть общего частотного ресурса W_t .

Кодирование частотным сдвигом

Понятно, что каждый индивидуальный сигнал опять не является сигналом с распределенным спектром, поскольку его частотно-временное произведение $WT = (W_t / M)T = 1$, и значит, любая система со сколь угодно большим числом ортогональных сигналов подобного типа, конечно, не является системой с распределенным спектром.

В отличие от кодирования временным сдвигом пик-фактор ортогональных сигналов данного вида v =1 и ошибки в синхронизации не играют столь критической роли, так как ортогональность достигается отсутствием перекрытия в частотной области.

Вместо этого деструктивным в некоторых случаях может оказаться дрейф спектра (к примеру, вследствие эффекта Доплера). Тем не менее данный способ передачи чрезвычайно популярен и примером его непосредственного воплощения служит традиционная М-ичная частотная манипуляция.

1.3 Ортогональное кодирование широкополосными сигналами

В этих обстоятельствах более привлекательным может оказаться использование широкополосных ортогональных сигналов, означающее **утилизацию всеми сигналами общего частотно-временного ресурса без распределения или дробления последнего.**

Рассмотрим простой пример воплощения подобной идеи в форме дискретных сигналов с бинарной фазовой манипуляцией (БФМ). Образуем каждый из *M* сигналов как **последовательность *N* следующих друг за другом элементарных импульсов** или чипов прямоугольной формы и длительности Δ.

Предположим, что полярности чипов сигнала с номером *к* подчиняются закону манипуляции, описываемому кодовой последовательностью (или просто кодом) бинарных символов $a_{k,i} = \pm 1$, где $к = 1,2,...,M$, а вторым индексом является номер чипа (дискретное время) $i = 0,1,...,N - 1$. Тогда версия, такого сигнала на видеочастоте запишется как

$$s_k(t) = \sum_{i=0}^{N-1} a_{k,i} s_0(t - i\Delta),$$

Можно показать, что скалярное произведение сигналов

$$(s_k, s_l) = E_0 \sum a_{ki} a_{li} = E_0 (a_k, a_l)$$

Последний результат связывает скалярное произведение сигналов $s(t)$ со скалярным произведением N - мерных векторов, образованных кодовыми последовательностями $a_k = (a_{k,0}, a_{k,1}, \ldots, a_{k,N-1})$.

Следовательно, ***M* ортогональных кодовых последовательностей** автоматически **формируют *M* ортогональных сигналов**. При ***M* ≤ *N* существует множество способов конструирования подобных последовательностей**, поскольку речь попросту идет об отыскании $M \leq N$ ортогональных векторов размерности N.

В нашем примере упомянутые векторы являются бинарными, т.е. их компоненты принимают значения из множества ±1. При $M = N$ ортогональные бинарные векторы, используемые как строки, образуют квадратную матрицу, называемую матрицей Адамара.

Известен ряд алгоритмов построения матриц Адамара специальных (не произвольных) размеров. Одним из них является **популярное правило Сильвестра, позволяющее рекурсивно удваивать размер матрицы**.

Для пояснения его сути предположим, что каким-либо образом найдена матрица Адамара H_M размерности M. Тогда матрицу Адамара удвоенного размера H_{2M} можно построить четырехкратным повторением H_M в качестве блоков H_{2M}, изменив знак одного из них:

$$\mathbf{H}_{2M} = \begin{bmatrix} \mathbf{H}_M & \mathbf{H}_M \\ \mathbf{H}_M & -\mathbf{H}_M \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \mathbf{H}_M,$$

Можно начать выполнение алгоритма Сильвестра с простейшей матрицы Адамара

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

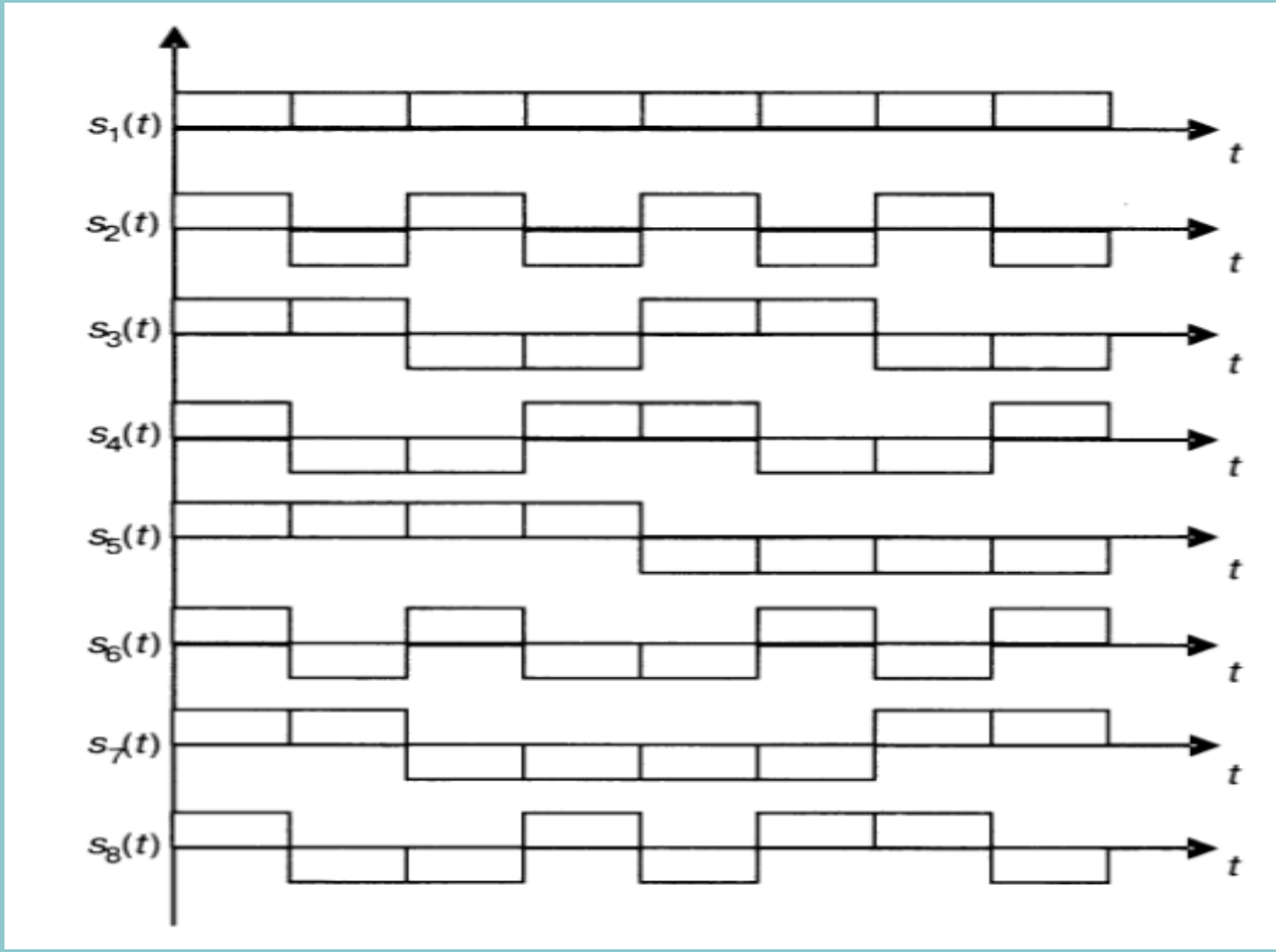
построив матрицу H4 (в которой ниже для краткости знаки «+» и «-» заменяют элементы «+1» и «-1»), затем от H4 перейдя к H8 и т. д.:

$$\mathbf{H}_4 = \begin{bmatrix} \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_2 & -\mathbf{H}_2 \end{bmatrix} = \begin{bmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{bmatrix},$$

$$\mathbf{H}_8 = \begin{bmatrix} \mathbf{H}_4 & \mathbf{H}_4 \\ \mathbf{H}_4 & -\mathbf{H}_4 \end{bmatrix} = \begin{bmatrix} + & + & + & + & + & + & + & + \\ + & - & + & - & + & - & + & - \\ + & + & - & - & + & + & - & - \\ + & - & - & + & + & - & - & + \\ + & + & + & + & - & - & - & - \\ + & - & + & - & - & + & - & + \\ + & + & - & - & - & - & + & + \\ + & - & - & + & - & + & + & - \end{bmatrix}.$$

Таким способом можно построить матрицу Адамара любого размера $M = 2^m = 2,4,8,16,32,\ldots$ Строки матрицы Адамара, построенной подобным образом, известны как функции Уолша.

Функции Уолша длины 8



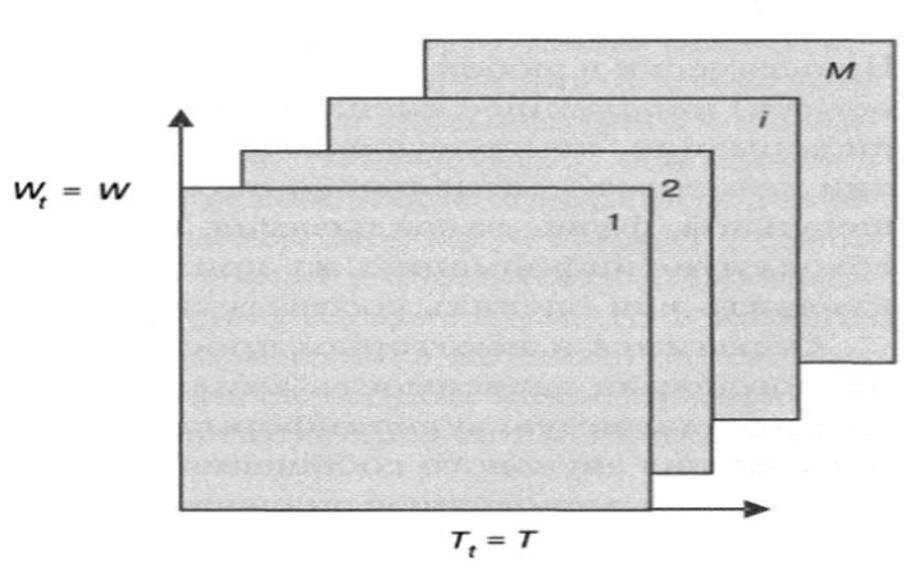
При данном варианте построения ортогонального ансамбля дробления общего ресурса между сигналами нет: каждый из них ***эксплуатирует весь доступный ресурс***, так что все сигналы **полностью перекрываются и во временной, и в частотной областях**.

Действительно, **полоса, занимаемая каждым сигналом**, может быть оценена как $W = 1/\Delta$,

тогда как длительность $T = M\Delta$,

откуда $WT = M = W_t T_t$.

Ортогональность же при этом достигается **не за счет** фрагментации временного интервала или полосы, а соответствующим выбором законов модуляции сигналов.



2. Измерение параметров сигналов в системах радиосвязи

2.1 Формулировка задачи и правила оценки

Практически в любой радиосистеме мы встречаем задачу измерения (или оценки) параметров сигнала. Она характерна для любой ситуации, когда информация, интересующая наблюдателя, содержится в текущем значении некоторого сигнального параметра (например, амплитуды, частоты, начальной фазы, запаздывания и др.). Тем самым, чтобы извлечь необходимую информацию из принятого колебания, наблюдатель должен измерить или оценить соответствующий параметр.

Формулировка задачи и правила оценки

1) В обычном АМ или ЧМ радиовещании измеряется зависимость амплитуды или частоты во времени, которые содержат передаваемую аудиоинформацию - громкость и звуковой тон.

2) В аналоговом телевидении необходимо выполнить измерения амплитуды, поскольку яркостный сигнал и цветовые компоненты передаются с помощью амплитудной модуляции.

3) В то же время передача аудиоинформации осуществляется с помощью угловой модуляции, в силу чего любой телевизионный приемник должен осуществлять и измерение частоты.

4) Принципиальной задачей локации и навигации также является измерение (оценивание) запаздывания и направления прихода сигнала для определения относительных дальности и угловых координат между приемником и объектом. Когда к тому же необходимо знание скорости цели и параметров ее траектории, выполняется измерение доплеровского сдвига частоты и т. д.

Формулировка задачи и правила оценки

В терминах, наиболее удобных в нашем контексте, задача оценки параметров может быть сформулирована следующим образом. Пусть наблюдение y(t) наряду с шумом содержит сигнал s(t; λ), являющийся детерминированным, за исключением неизвестного значения постоянного параметра λ.

Параметр λ может быть скалярным либо векторным в зависимости от конкретного содержания задачи. Наблюдатель, основываясь на анализе y(t), должен принять решение о том, какое конкретное значение из диапазона возможных принял параметр сигнала λ.

Наряду с самой задачей это решение принято называть оценкой и обозначать $\tilde{\lambda}$.

Поскольку в принятом наблюдении y(t) всегда присутствует шум, то в каждом сеансе приема оценка $\tilde{\lambda}$ отличается от неизвестного истинного значения параметра λ. Поэтому естественным является вопрос: какое правило решения является оптимальным, т.е. гарантирующим наименьший ущерб от отклонений $\tilde{\lambda}$ от λ.

Формулировка задачи и правила оценки

Простейший подход к этой проблеме основывается на понимании того факта, что задача оценки не является чем-то принципиально новым по отношению к задаче различения M сигналов. В самом деле, предположим вначале, что измеряемый параметр λ дискретен и принимает одно из M конкурирующих значений

$\lambda_1, \lambda_2, \ldots, \lambda_M$.

Тогда решение о том, какое из возможных значений принял параметр сигнала в данном конкретном наблюдении, есть не что иное как проверка M гипотез о том, какой из M конкурирующих сигналов $s_1(t), s_2(t), \ldots, s_M(t)$ фактически принят, причем все распознаваемые сигналы являются копиями сигнала $s(t; \lambda)$, отличающимися друг от друга только значением параметра λ, т.е.

$$s_k(t) = s(t; \lambda_k).$$

Для того, чтобы охватить подобной постановкой и случай непрерывного параметра λ, требуется лишь устремить число его возможных значений M, т. е. различаемых сигналов, к бесконечности (вплоть до несчетной).

Формулировка задачи и правила оценки

Из приведенных рассуждений следует, что уже известная оптимальная стратегия решения — правило максимального правдоподобия (МП) — применима и в оценке параметров. Это означает, что среди всех конкурирующих значений λ в качестве оценки $\tilde{\lambda}$ следует выбирать то, которое максимизирует вероятность трансформации каналом сигнала $s(t, \lambda)$ в наблюдаемое колебание $y(t)$. Для канала с аддитивным белым гауссовски шумом (АБГШ) это правило эквивалентно правилу минимума расстояния, которое представляется в виде

$$d\left(\boldsymbol{s}_{\tilde{\lambda}}, y\right) = \min d(\boldsymbol{s}_{\lambda}, y) \rightarrow \tilde{\lambda}$$

где $\boldsymbol{s}_{\tilde{\lambda}}$ — векторное обозначение сигнала $s(t; \tilde{\lambda})$. Это правило формирует максимально правдоподобную оценку λ, находя такое значение $\tilde{\lambda}$, при котором сигнал $s(t; \lambda)$ наиболее близок к наблюдению $y(t)$ в смысле евклидова расстояния, что иллюстрируется на след. слайде.

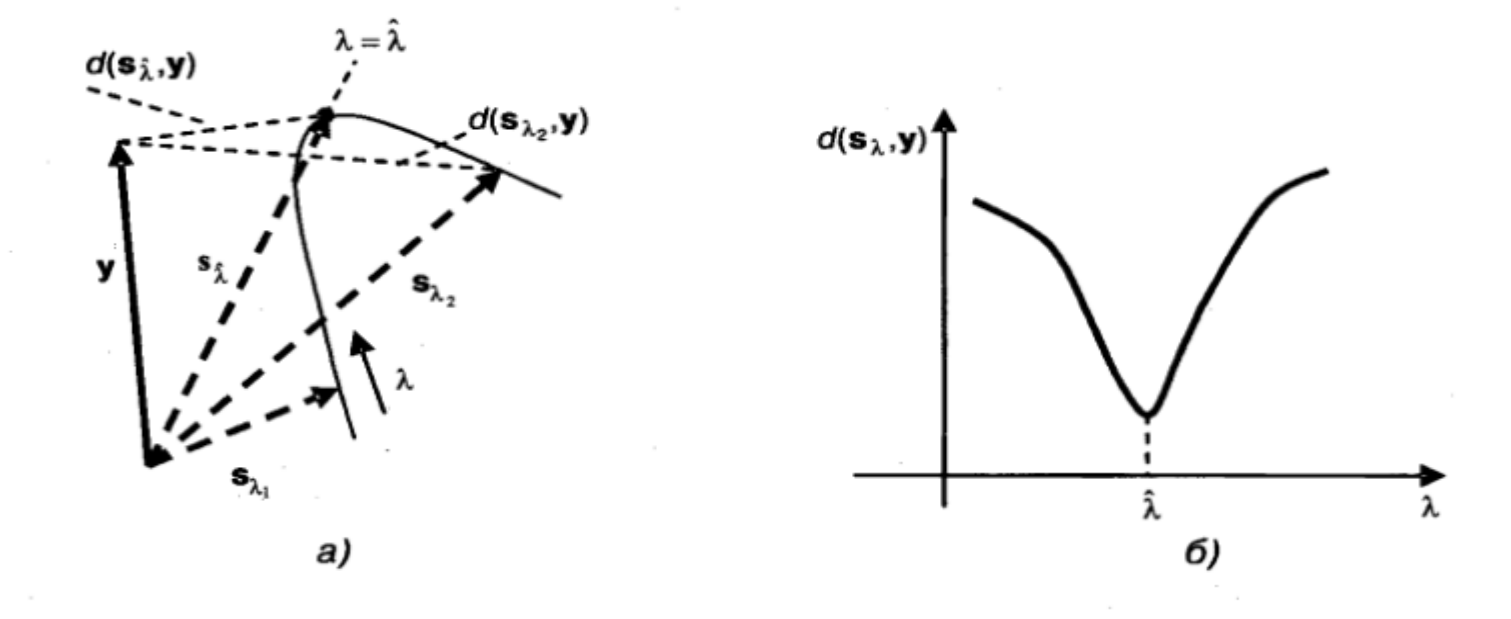
Формулировка задачи и правила оценки

Из приведенных рассуждений следует, что уже известная оптимальная стратегия решения — правило МП — применима и в оценке параметров. Это означает, что среди всех конкурирующих значений λ в качестве оценки $\tilde{\lambda}$ следует выбирать то, которое максимизирует вероятность трансформации каналом сигнала s(t, λ) в наблюдаемое колебание y(t). Для канала с аддитивным белым гауссовским шумом (АБГШ каналом) это правило эквивалентно правилу минимума расстояния, которое представляется в виде

$$d\left(\boldsymbol{s}_{\tilde{\lambda}}, y\right) = \min d(\boldsymbol{s}_{\lambda}, y) \rightarrow \tilde{\lambda}$$

где $\mathbf{s}_{\tilde{\lambda}}$ — векторное обозначение сигнала s(t; $\tilde{\lambda}$). Это правило формирует максимально правдоподобную оценку α, находя такое значение $\tilde{\lambda}$, при котором сигнал s(t;λ) наиболее близок к наблюдению y(t) в смысле евклидова расстояния, что иллюстрируется на следующем слайде.

Сигнал $s(t; \lambda)$ можно отождествить с вектором s_λ, который перемещается в пространстве с изменением параметра λ. Конец вектора описывает некоторую траекторию, точки которой соответствуют различным значениям параметра λ (рис. а). Согласно правилу МП следует найти точку этой траектории, ближайшую к вектору наблюдения y, и выдать соответствующее ей значение λ в качестве оценки. Дополнительное пояснение этому дает рис. б, где приведена характерная зависимость расстояния между наблюдением и копией сигнала от значения λ: МП оценка λ отвечает такому значению параметра, которое минимизирует это расстояние.



В зависимости от влияния на энергию сигнала параметры классифицируются на энергетические и неэнергетические. В случае неэнергетического параметра энергия сигнала s(t; λ) не зависит от конкретного значения λ, т.е.

$$E(\lambda) = \int_{-\infty}^{\infty} s^2(t;\lambda)dt = E$$

К примеру, **амплитуда** и **длительность сигнала** являются энергетическими, тогда как **запаздывание, частота** или **начальная фаза** — неэнергетическими параметрами. Ясно, что оценка неэнергетического параметра представляет собой частный случай различения конкурирующих сигналов равной энергии, для которого правило максимума корреляции в данных обозначениях представимо как

$$z(\hat{\lambda}) = \max_{\lambda} z(\lambda) \Rightarrow \hat{\lambda},$$

где

$$z(\lambda) = (\mathbf{y}, \mathbf{s}_\lambda) = \int_0^T y(t)s(t;\lambda)\,dt$$

корреляция между наблюдаемым колебанием *y*(t) и сигналом *s*(t, λ), зависящая от значения измеряемого параметра λ.

2.2 Точность оценки

В задаче оценки параметров распознаваемые сигналы есть просто копии s(t; λ) с различными значениями параметра λ.

Во многих практических ситуациях корреляция двух любых таких копий s(t; λ_1) и s(t; λ_2) определяется лишь их рассогласованием по параметру λ, т. е. разностью λ_1 — λ_2, а не значениями λ_1, λ_2 в отдельности, так что, положив

$$\lambda_1 = 0,\ \lambda_2 = \lambda,$$

можно представить коэффициент корреляции в случае неэнергетического параметра λ как

$$\rho(\lambda) = \frac{(\mathbf{s}_0, \mathbf{s}_\lambda)}{E} = \frac{1}{E}\int_{-\infty}^{\infty} s(t;0)s(t;\lambda)\,dt.$$

Точность оценки

Как обычно, коэффициент корреляции характеризует степень сходства двух сигнальных копий в зависимости от их рассогласования по параметру λ. Очевидно,

$$\rho(\lambda) \le \rho(0) = 1,$$

что предсказуемо и естественно:

сигнальные копии, рассогласованные по значению *параметра*, не могут иметь большего сходства, чем полностью совпадающие, которые, в свою очередь, имеют стопроцентную корреляцию.

Другим свойством коэффициента корреляции, вытекающим из егозависимости только от рассогласования $\lambda = \lambda_2 - \lambda_1$, является четность $\rho(\lambda) = \rho(-\lambda)$

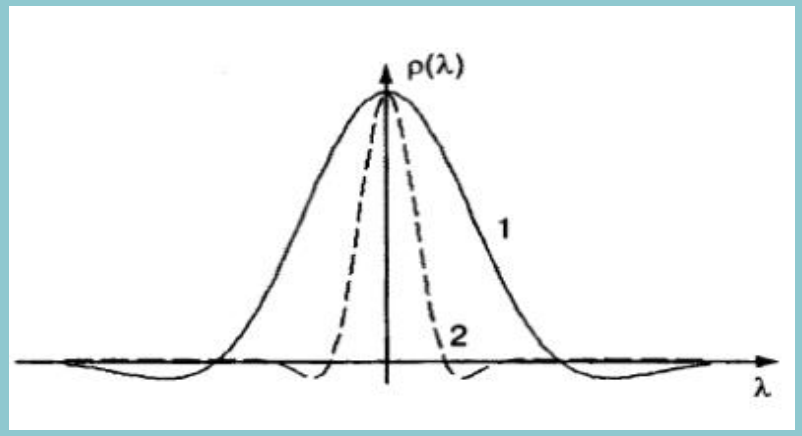


Рис. Характерные зависимости $\rho(\lambda)$

Точность оценки

Как легко понять, то или иное правило решения оптимально **только в смысле строго оговоренного критерия качества**.

Правило МП, на которое мы до сих пор опирались, **оптимально по критерию минимума вероятности ошибки**, применение которого вполне естественно и адекватно при различении сигналов дискретного множества или оценке дискретного параметра. Однако ориентация на этот критерий при измерении непрерывного параметра для практики нетипична. В этом случае гораздо уместнее характеризовать качество оценки с позиций точности, т. е. отклонением

$$\varepsilon = \tilde{\lambda} - \lambda$$

оценки параметра $\tilde{\lambda}$ от истинного значения λ.

Прежде всего, разумно потребовать, чтобы **математическое ожидание ошибки ε**, в усредненное по всем возможным наблюдениям y(t) при фиксированном истинном значении λ, **было равно нулю**, т. е. в среднем оценка $\tilde{\lambda}$ совпадала с истинным значением λ:

$$\bar{\varepsilon} = \overline{\hat{\lambda} - \lambda} = 0 \Leftrightarrow \bar{\hat{\lambda}} = \lambda, \forall \lambda.$$

Точность оценки

Оценка, для которой выполняется данное равенство, называется несмещенной. Однако несмещенность еще не дает оснований считать оценку хорошей, поскольку критически важен также диапазон разброса случайных флюктуации оценки относительно истинного значения параметра.

Традиционной и адекватной мерой разброса служит дисперсия ошибки

$$\mathrm{var}\{\varepsilon\} = M(\tilde{\lambda} - \lambda)^2,$$

поэтому естественно искать правило, гарантирующее **наряду с несмещенностью минимум дисперсии оценки var {ε}** для любых истинных значений λ:

$$\mathrm{var}\{\varepsilon\} = \overline{(\hat{\lambda} - \lambda)^2} = \min, \quad \forall\lambda.$$

Как можно видеть, минимизация дисперсии несмещенной оценки является естественной формализацией стремления **к наивысшей точности измерений.**

Точность оценки

Одним из фундаментальных результатов теории оценок является ***неравенство Крамера - Рао***, устанавливающее *нижний предел дисперсии любой несмещенной оценки*.

Оценка, дисперсия которой лежит на этой границе, называется **эффективной**.

Измерительные задачи, в которых возможны эффективные в строгом смысле оценки, достаточно редки. Как доказывается в классической теории оценок МП оценка, обладает свойствами **асимптотической несмещенности и эффективности**.

Термин «асимптотически» в физической трактовке подразумевает «в ситуациях, когда необходима высокая точность измерений» или, еще конкретнее, когда отношение сигнал-шум (либо время наблюдения) достаточно велико. Поэтому в любой задаче, где необходимы измерения с малой погрешностью, правило МП оптимально не только **по критерию вероятности ошибки, но и по критерию точности оценки**.

Точность оценки

Разумеется, в реальных приложениях, как правило, требуется высокая точность измерений, чем и объясняется повсеместное практическое использование оценок МП.

Для неэнергетического параметра λ граница Крамера-Рао принимает особенно простой вид и служит практическим инструментом для вычисления дисперсии МП оценки:

$$\mathrm{var}\{\hat{\lambda}\} = \mathrm{var}\{\varepsilon\} \approx -\frac{1}{\rho''(0)q^2}, \qquad q \gg 1.$$

Присутствие в знаменателе правой части отношения сигнал-шум

$$q^2 = 2E/N_0$$

неудивительно:

при любом разумном правиле оценки рост отношения сигнал-шум должен приводить к уменьшению ошибки, т. е. улучшению точности измерений.

Точность оценки

Из курса математического анализа известно, что **вторая производная говорит о кривизне или остроте функции в рассматриваемой точке** и для выпуклой кривой имеет отрицательный знак.

Острота ρ(λ) в нулевой точке, в свою очередь, является **индикатором чувствительности сигнала по отношению к рассогласованию по λ**:

чем острее ρ(λ) тем быстрее копия сигнала, расстроенная по λ, теряет свое сходство с исходной.

Так как оценка есть частный случай различения сигналов, а достоверность последнего **тем выше**, чем **меньше сходство сигналов**. Это полностью объясняет зависимость точности измерения λ от ρ"(0): если копии s(t, λ) имеют **слабое сходство между собой даже при близких значениях λ, надежность их распознавания выше, чем в случае сильного сходства.**

Последний факт подсказывает общее направление синтеза сигналов в задачах оценки неэнергетического параметра λ. Для достижения нужной точности **не за счет «грубой силы», т. е. простого увеличения энергии**, следует синтезировать сигналы, у которых **коэффициент корреляции ρ(λ) как функция λ подобен «острому» импульсу.**

2.3 Оценка амплитуды

Задача измерения интенсивности (уровня, мощности) сигнала встречается во многих приложениях: от телевизионного вещания до сотовой связи и цифровых ИКМ или КАМ линий передачи данных. Сформулируем ее как **задачу измерения неизвестной амплитуды *A*, постоянной на интервале наблюдения [0,T]**.

При этом принятый полезный сигнал **может быть представлен моделью**

$$s(t; A) = As(t),$$

где $s(t)$ — детерминированный эталонный сигнал, **амплитуда которого по определению полагается равной единице**.

Как видно, $s(t; A)$ есть результат масштабирования эталонного сигнала умножением на неизвестный множитель A.

Положим энергию эталонного сигнала равной E. Тогда энергия $E(A)$ сигнала с амплитудой A и его корреляция $z(A)$ с наблюдением $y(t)$ найдутся как

$$E(A) = \int_0^T s^2(t; A)\,dt = A^2 \int_0^T s^2(t)\,dt = A^2 E, \quad z(A) = \int_0^T y(t)s(t; A)\,dt = Az,$$

где

$$z = \int_0^T y(t)s(t)\,dt$$

— корреляция наблюдения с эталонным сигналом.

Теперь отыскание МП оценки может трактоваться как максимизация по A разности $\boldsymbol{z(A) - E(A)/2}$.

Подстановка приведенных выше равенств для энергии и корреляции представляет эту разность в виде **квадратичного бинома** $\boldsymbol{Az - A^2E/2}$ относительно A с известными коэффициентами. Результатом его максимизации является МП оценка

$$\hat{A} = \frac{z}{E}.$$

Таким образом, оптимальная оценка амплитуды получается **вычислением корреляции наблюдаемого колебания с эталонным сигналом** и умножением результата на постоянный множитель $1/E$.

Математическое ожидание z вычисляется как

$$\bar{z} = \int_0^T \overline{y(t)} s(t)\, dt = \int_0^T s(t; A) s(t)\, dt = A \int_0^T s^2(t)\, dt = AE$$

видно, что значение A в среднем точно совпадает с истинным значением амплитуды:

$$\overline{\hat{A}} = \bar{z}/E = A.$$

т. е. **МП оценка амплитуды является строго** (а не только асимптотически) **несмещенной.**

Дисперсия $\check{A}$

$$\operatorname{var}\{\hat{A}\} = \frac{\operatorname{var}\{z\}}{E^2} = \frac{N_0}{2E} = \frac{1}{q^2},$$

где q^2 — **отношение сигнал-шум для эталонного сигнала** (т.е. сигнала с единичной амплитудой).

Можно показать, что последнее выражение **точно воспроизводит границу** Крамера - Рао, обеспечивая ***строгую*** (а не только асимптотическую) эффективность МП оценки амплитуды сигнала.

Вывод по измерению амплитуды

Итак, какие же требования к синтезу сигналов предъявляет задача измерения амплитуды?

Как показывает дисперсия оценки, её точность зависит ***исключительно*** от **энергии** сигнала.

Никакое усложнение закона модуляции при неизменной энергии сигнала не в состоянии повысить точность измерения. Следовательно, рассмотренная классическая задача приема ***не содержит*** каких-либо предпосылок к привлечению широкополосной технологии.

Оценка фазы

Обратимся теперь к ситуации, в которой носителем полезной информации выступает **начальная фаза сигнала**. Данный случай типичен для *когерентных* ***локационных и навигационных приемников***, устройств восстановления несущей систем передачи данных с ФМ и КАМ, ***демодуляторов систем мобильной радиосвязи 2-го и 3-го поколений***, каналов цветности в аналоговом телевидении и многих других приложений.

Преобразуем модель радиосигнала

$$s(t;\varphi) = S(t)\cos(2\pi f_0 t + \gamma(t) + \varphi),$$

отделив постоянную (на интервале наблюдения) начальную фазу φ, являющуюся неизвестной и подлежащей измерению:

Оценка фазы

Поскольку φ — неэнергетический параметр, то $E(\varphi) = E$ и нахождение МП оценки фазы сводится к максимизации $z(\varphi)$ по φ от –π до π, т.е.

$$\hat{\varphi} = arg \max_{\varphi \in [-\pi, \pi]} z(\varphi)$$

Чтобы воспользоваться соотношением для дисперсии оценки

$$\mathrm{var}\{\hat{\lambda}\} = \mathrm{var}\{\varepsilon\} \approx -\frac{1}{\rho''(0)q^2}, \qquad q \gg 1.$$

покажем, что **ρ(φ) есть косинус угла между двумя сдвинутыми по фазе на угол φ** копиями сигнала, т. е. двумя векторами, разделенными углом φ. Следовательно,

$$\rho(\varphi) = \cos\varphi, \quad \text{а } \rho''(0) = -1$$

и дисперсия МП оценки $\tilde{\omega}$

$$\mathrm{var}\{\hat{\varphi}\} \approx \frac{1}{q^2}, \qquad q \gg 1.$$

Оценка фазы

Вновь, как и в случае измерения амплитуды, точность оценки фазы **зависит только от отношения сигнал-шум**.

Таким образом, и эта классическая задача приема **нейтральна к закону модуляции при неизменности энергии сигнала** и, следовательно, **никак не стимулирует к применению широкополосных сигналов.**

Оценка запаздывания радиосигнала

Алгоритм оценки

Задача, **оценки (измерения) запаздывания** является одной из наиболее часто встречающихся в приложениях. Она **типична для:**

- телевидения (каналы синхронизации),
- цифровых систем мобильной радиосвязи (пилотные каналы, тактовая синхронизация),
- радиолокации (измерение дальности до цели),
- навигации космического и наземного базирования (измерение расстояния до маяков) и т. п.

Практически в любой современной системе обработки информации для нормальной работы **требуется совмещение местной временной шкалы с системной**, данные о которой должны извлекаться из принятого колебания.

Подобная процедура и есть не что иное, как **оценка запаздывания сигнала**

Алгоритм оценки запаздывания

Предположим, что радиосигнал

$$s(t) = \mathrm{Re}[\acute{S}(t)\exp(j2\pi f_0 t)],$$

распространяясь по каналу, приобретает **неизвестные запаздывание** τ и **начальную фазу** φ_0, т.е. принимает вид

$$s(t;\tau;\varphi_0) = s(t-\tau;\varphi_0) = \mathrm{Re}\left\{\dot{S}(t-\tau)\exp\left[j2\pi f_0(t-\tau)+j\varphi_0\right]\right\}.$$

где $\acute{S}(t)$ – комплексная огибающая сигнала $s(t)$,

φ_0 - случайна и равномерно распределена на интервале [–π, π],

т. е. **не содержит информации о предмете интереса — запаздывании** τ.

Включим фазовый набег, обусловленный запаздыванием τ в полную начальную фазу

$$\varphi = -2\pi f_0 \tau + \varphi_0.$$

где φ – случайна и равномерно распределена на интервале [–π, π], **независима от** τ, т.е. **не содержит информации** об измеряемом временном сдвиге **из-за вклада неизвестной** φ_0.

Алгоритм оценки запаздывания

Тогда принятый сигнал может быть представлен как

$$s(t-\tau;\varphi) = \mathrm{Re}\left[\dot{S}(t-\tau)\exp(j2\pi f_0 t)\exp(j\varphi)\right],$$

где запаздывание τ - неизвестный **полезный параметр**, подлежащим измерению,

φ — **бесполезная начальная фаза**, неопределенность которой лишь потенциально затрудняет оценку τ.

Любая процедура оценки является **частным случаем различения сигналов**.

В рассматриваемой ситуации приходится **распознавать многочисленные копии сигнала**, **отличные друг от друга значением временного сдвига** τ, считаясь с присутствием мешающего параметра — начальной фазы φ.

Алгоритм оценки запаздывания

Можно показать, что существует **прямой способ** преодоления неопределенности, вносимой случайностью φ, который состоит в замене **оптимального выбора между некогерентными сигналами на выбор между их детерминированными законами модуляции, т. е. комплексными огибающими**.

Следовательно, **для оценки запаздывания следует сравнить сдвинутые во времени копии комплексной огибающей сигнала *Ś*(t), посчитав одну из них за принятую**.

Временной сдвиг выбранной копии и будет выдан в качестве МП оценки запаздывания

$$\tilde{\imath} = \arg\max Z(\boldsymbol{\tau})$$

где

$$Z(\tau) = \left| \int_0^T \dot{Y}(t)\dot{S}^*(t-\tau)\,dt \right|.$$

Алгоритм оценки

Т.е. **предпочтение данной копии остальным** основывается **на ее минимальном расстоянии от принятой комплексной огибающей $\dot{Y}(t)$,** или, учитывая, что запаздывание — неэнергетический параметр, на **максимуме ее корреляции с $\dot{Y}(t)$**.

Основанное на данной величине правило оценки

$$Z(\tilde{\tau}) = \max Z(\tau)$$

физически весьма прозрачно: МП оценка τ есть **просто значение запаздывания**, при котором **закон модуляции сигнала обладает наибольшим сходством с законом модуляции наблюдения**.

Алгоритм оценки запаздывания

Одним из возможных вариантов реализации указанного правила является использование банка корреляторов, представленного на рисунке.

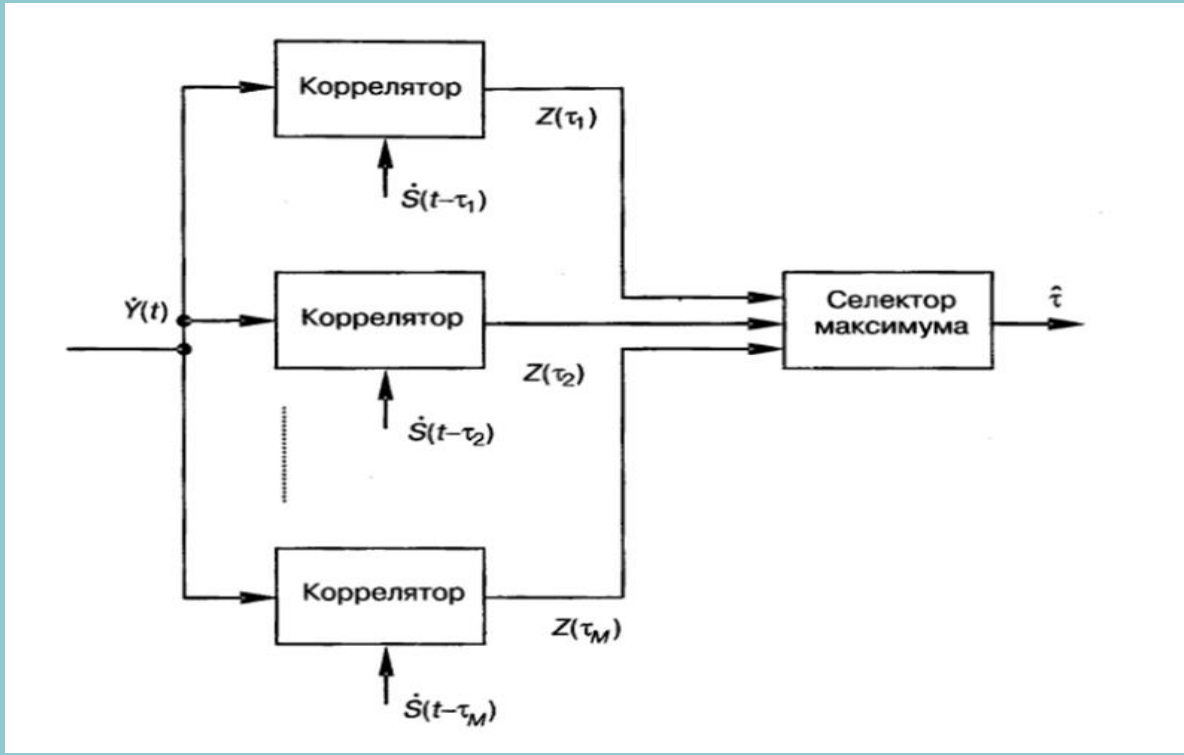


Рис. Реализация МП оценки запаздывания банком корреляторов.

Алгоритм оценки запаздывания

Разумеется, данная структура оперирует с запаздыванием так, будто последнее принимает **лишь дискретные значения**. Если же в реальности это не так, она попросту квантует непрерывную величину τ.

Понятно, что при этом следует использовать **такое число корреляторов** (равное количеству дискретных значений τ_k), при котором дополнительная ошибка, вносимая квантованием, **была бы достаточно малой**.

Альтернативная структура измерителя запаздывания возможна **на основе согласованного фильтра.**

Пусть наблюдение $y(t)$ поступает на вход фильтра, **согласованного с сигналом s(t)**. Найдем реакцию фильтра r(t), подставляя в интеграл свертки импульсный отклик фильтра

$$h(t) = s(T - t):$$

Алгоритм оценки запаздывания

Альтернативная структура измерителя запаздывания возможна на основе **согласованного фильтра**.

Пусть наблюдение $y(t)$ поступает на вход фильтра, **согласованного с сигналом s(t)**. **Найдем реакцию фильтра r(t)**, подставляя в интеграл свертки импульсный отклик фильтра h(t) = s(T – t):

$$r(t) = \int_{-\infty}^{\infty} y(\theta)h(t-\theta)\,d\theta = \int_{-\infty}^{\infty} y(\theta)s(T-t+\theta)\,d\theta.$$

Данный интеграл, представляющий собой скалярное произведение y(θ) и s(T–t+θ), может быть вычислен как

$$r(t) = \int_{-\infty}^{\infty} y(\theta)s(T-t+\theta)\,d\theta = \frac{1}{2}\mathrm{Re}\left[\int_{0}^{T} \dot{y}(\theta)\dot{s}^*(T-t+\theta)\,d\theta\right] =$$

$$= \mathrm{Re}\left\{\left[\frac{1}{2}\int_{0}^{T} \dot{Y}(\theta)\dot{S}^*(T-t+\theta)\,d\theta \cdot \exp(-j2\pi f_0 T)\right]\exp(j2\pi f_0 t)\right\}$$

Алгоритм оценки

Сравнивая полученный результат с общей моделью радиосигнала

$$s(t) = \mathrm{Re}\left[\dot{S}(t)\exp(j2\pi f_0 t)\right],$$

можно видеть, что **квадратные скобки** в нем выделяют не что иное, как **комплексную огибающую на выходе фильтра**.

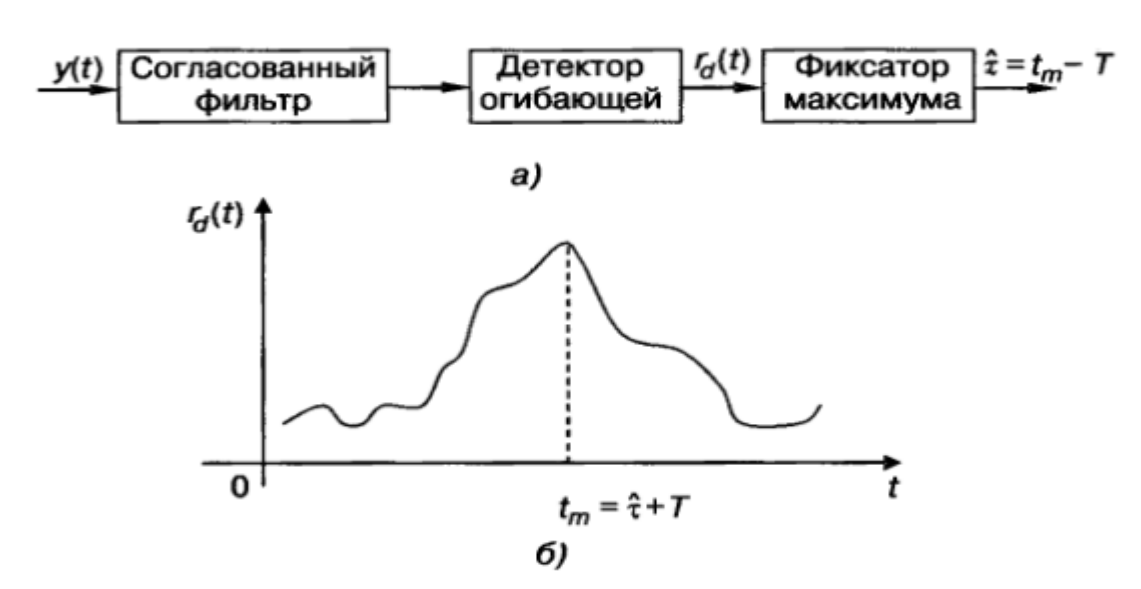
Следовательно, **действительная огибающая выходного сигнала** (закон амплитудной модуляции), задаваемая абсолютным значением выражения в квадратных скобках

$$\left|\frac{1}{2}\int_{-\infty}^{\infty}\dot{Y}(\theta)\dot{S}^*[\theta-(t-T)]\,d\theta\right| = \frac{1}{2}Z(t-T),$$

т. е. **повторяет в реальном времени** (с несущественным коэффициентом 1/2) **копию модуля корреляции, сдвинутую по времени на заранее известную длительность сигнала**.

Алгоритм оценки запаздывания

Отмеченный факт лежит в основе возможной схемы МП **измерителя запаздывания τ**, показанной на рис. а.



Наблюдение y(t) **вначале обрабатывается согласованным фильтром**, выходной сигнал которого подается **на детектор огибающей**. Последний блок **фиксирует момент времени t_m**, когда сигнал $r_d(t)$ на выходе детектора достигает максимального значения.

В итоге МП оценка ĩ получается как результат вычитания из t_m известной константы T (рис. б).

Точность оценки запаздывания

Согласно выражению для дисперсии оценки неэнергетического параметра

$$\mathrm{var}\{\hat{\lambda}\} = \mathrm{var}\{\varepsilon\} \approx -\frac{1}{\rho''(0)q^2}, \qquad q \gg 1.$$

дисперсия оценки **τ зависит от скорости, с которой сдвинутая по времени копия сигнала теряет сходство с исходным образцом**.

Для сигналов со случайной начальной фазой получение МП оценки связано **со сравнением лишь детерминированных комплексных огибающих. Степень сходства** рассогласованных во времени копий комплексной огибающей **характеризуется огибающей АКФ сигнала**

$$\dot{\rho}(\tau) = \frac{(\dot{\mathbf{S}}_0, \dot{\mathbf{S}}_\tau)}{\|\dot{\mathbf{S}}\|^2} = \frac{(\dot{\mathbf{S}}_0, \dot{\mathbf{S}}_\tau)}{2E} = \frac{1}{2E}\int_{-\infty}^{\infty} \dot{S}(t)\dot{S}^*(t-\tau)\,dt$$

и поэтому именно **ее кривизна влияет на дисперсию ошибки оценки запаздывания**

Точность оценки запаздывания

Можно показать, что дисперсия оценки запаздывания сигнала выражается выражением, которое называют формулой Вудворда

$$\mathrm{var}\{\hat{\tau}\} \approx \frac{1}{(2\pi W_{rms})^2 q^2}, \qquad q \gg 1,$$

где W_{rms} – среднеквадратическая **ширина спектра сигнала**, определяемая соотношением

$$W_{rms}^2 = \frac{1}{2E} \int_{-\infty}^{\infty} f^2 |\tilde{\dot{S}}(f)|^2 \, df.$$

Точность оценки запаздывания

Выводы:

1. Погрешность измерения времени запаздывания сигнала определяется, наряду с величиной отношения сигнал/шум, шириной полосы спектра сигнала (остротой автокрреляционной функции).

2. В отсутствие ограничений на пиковую мощность повышение точности в классической задаче измерения запаздывания не требует обязательного привлечения технологии расширения спектра, т.е. использования сложных сигналов.

3. В то же время указанная технология является безальтернативной всякий раз, когда пиковая мощность жестко лимитирована.

Оценка несущей частоты

Рассмотрим теперь ситуацию, в которой **неизвестным информативным параметром служит несущая частота сигнала**. Подобная задача столь же часто встречается на практике, как и предыдущая. Она характерна для:

- радиолокации, где скорость объекта измеряется через оценку доплеровского сдвига частоты (один из примеров — радар мониторинга дорожного движения),
- устройств восстановления опорного колебания в мобильной радиосвязи второго и третьего поколений;
- систем автоподстройки частоты в телевизионных и вещательных ЧМ приемниках и т. п.

Оценка несущей частоты

Обычно в приложениях номинальное значение несущей частоты f_0 фиксировано и **измерению подлежит лишь смещение *F* частоты принятого сигнала относительно f_0,** так что модель сигнала на входе приемника удобно записать в виде

$$s(t;F;\varphi) = \mathrm{Re}\left\{\dot{S}(t)\exp[j2\pi(f_0+F)t+j\varphi]\right\} = \\ = \mathrm{Re}\left[\dot{S}(t;F)\exp(j2\pi f_0 t)\exp(j\varphi)\right],$$

где $\dot{S}(t;F) = \dot{S}(t)\exp(j2\pi Ft)$ — комплексная огибающая сигнала, включающая линейный фазовый дрейф, обусловленный частотным сдвигом *F*, а φ — случайная начальная фаза (мешающий параметр) , не содержащая информации о сдвиге *F*.

Оценка несущей частоты

В силу неэнергетической природы *F* правило максимума корреляции вновь оптимально и выражается как

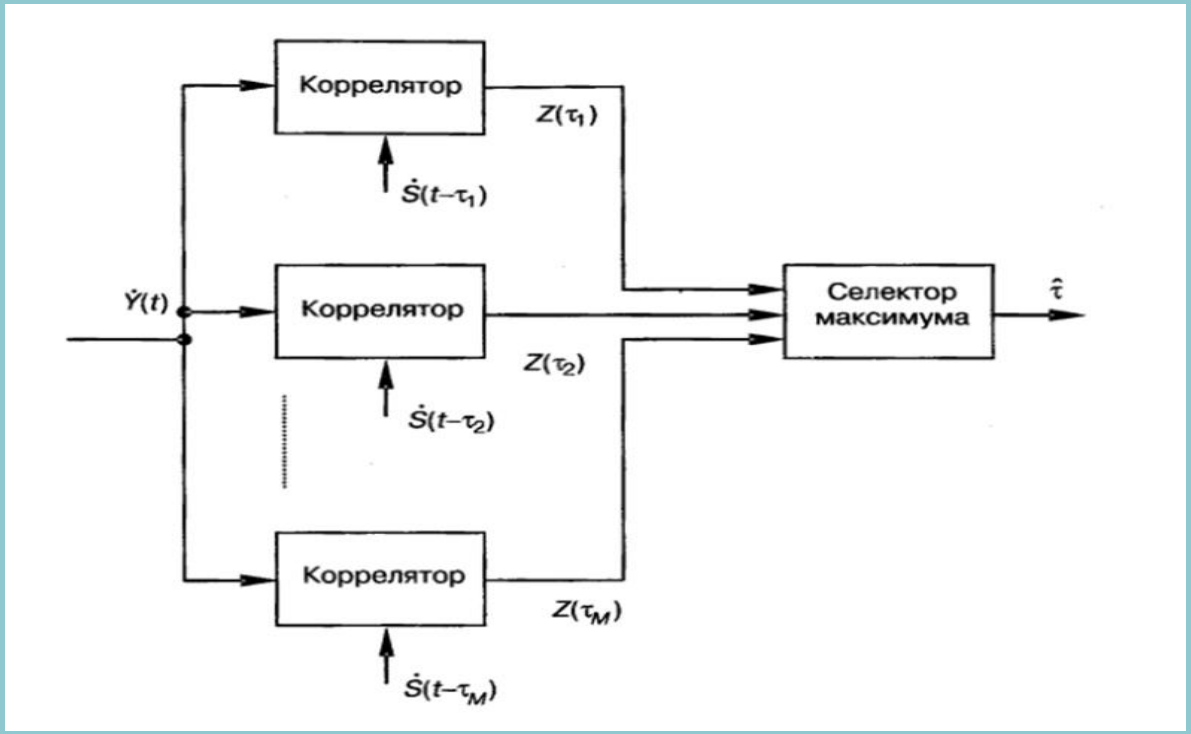
$$Z(\hat{F}) = \max_{F} Z(F),$$

где модуль корреляции

$$Z(F) = \left| \int_0^T \dot{Y}(t)\dot{S}^*(t;F)\,dt \right| = \left| \int_0^T \dot{Y}(t)\dot{S}^*(t)\exp(-j2\pi Ft)\,dt \right|.$$

Оценка несущей частоты

Структуру



легко приспособить для осуществления данной процедуры оценки, если **в качестве опорных сигналов корреляторов в ней взамен сдвинутых по времени копии** сигнала **использовать частотно-сдвинутые копии комплексной огибающей**

$$\dot{S}(t;F_k) = \dot{S}(t)\exp(j2\pi F_k t),$$

Оценка несущей частоты

С другой стороны, видно, что **амплитуда сигнала на выходе согласованного фильтра,** настроенного на частоту $f_0 + F_k$ **принимает в момент окончания сигнала** T **значение** $Z(F_k)$.

Таким образом, **можно воспользоваться альтернативной схемой, содержащей банк** M **согласованных фильтров**, где k-й фильтр настроен на частоту, отличающуюся от номинальной f_0 на величину F_k .

После амплитудного детектирования и взятия отсчетов с выходов детекторов в момент T **получается набор значений** $Z(F_k)$**, наибольшее из которых затем выбирается для выдачи МП оценки** F **как частотной расстройки фильтра с максимальным откликом** $Z(F_k)$**.**

Оценка несущей частоты

Ясно, что в случае непрерывного сдвига *F* схема производит его квантование, и **число фильтров *M* должно быть выбрано достаточным для удержания ошибки квантования в приемлемых пределах**.

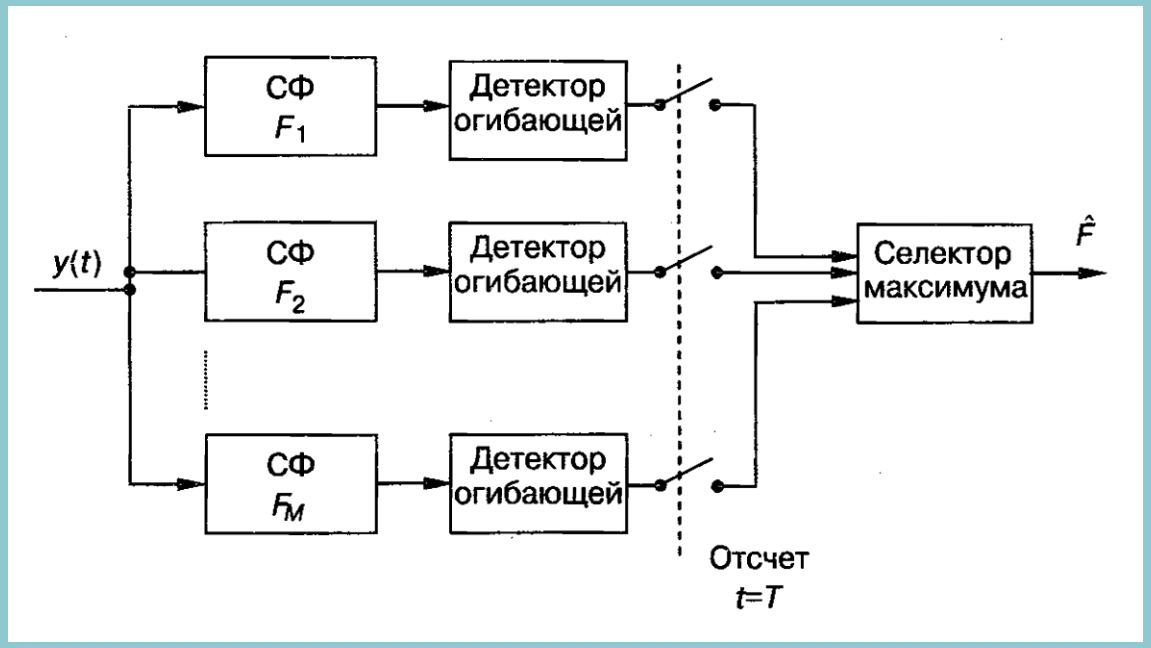


Рис. МП измеритель частоты на основе банка СФ

Точность оценки частоты

Согласно выражению для дисперсии оценки

$$\mathrm{var}\{\hat{\lambda}\} = \mathrm{var}\{\varepsilon\} \approx -\frac{1}{\rho''(0)q^2}, \qquad q \gg 1.$$

точность измерения частоты определяется наряду с **отношением сигнал-шум** также **остротой зависимости корреляции между частотно-сдвинутыми копиями сигнала от их взаимной частотной расстройки**.

Из-за случайности фазы сигнала достаточно оперировать лишь с **его комплексной огибающей, и сходство** между ее частотно-сдвинутыми копиями $S(t; 0)$ и $S(t; F)$ **характеризуется модулем коэффициента корреляции**

$$\rho_0(F) = |\dot{\rho}(F)| = \left| \frac{1}{2E} \int_{-\infty}^{\infty} \dot{S}(t;0)\dot{S}^*(t;F)\,dt \right| =$$

$$= \left| \frac{1}{2E} \int_{-\infty}^{\infty} |\dot{S}(t)|^2 \exp(-j2\pi Ft)\,dt \right| = \left| \frac{1}{2E} \int_{-\infty}^{\infty} S^2(t) \exp(-j2\pi Ft)\,dt \right|.$$

Точность оценки частоты

Как видно, **зависимость коэффициента корреляции** в функции от *F* **повторяет по форме амплитудный спектр квадрата действительной огибающей сигнала**.

Действительная огибающая, т. е. закон амплитудной модуляции, по смыслу является видеосигналом, поэтому из свойств преобразования Фурье следует, что **чем длиннее сигнал, тем острее $\rho_o(F)$.**

Подходящим названием протяженности $\rho_o(F)$ вдоль оси частот могло бы служить **«полоса огибающей»**.

Обозначив ее как *Fe* имеем $Fe \approx 1/T$.

Таким образом, **помимо «грубой силы»,** т. е. увеличения энергии, можно **повысить точность измерения частоты**, применяя сигнал с компактным спектром огибающей (**узкой полосой огибающей *Fe***), т.е. **достаточно большой длительностью *T*.**

Точность оценки частоты

Формально этот вывод снова может быть получен вычислением второй производной $\rho_o(F)$ и подстановкой результата в границу Крамера-Рао

$$\mathrm{var}\{\hat{F}\} \approx \frac{1}{(2\pi T_{rms})^2 q^2},$$

где среднеквадратическая длительность *Trms* характеризует протяженность сигнала во времени подобно тому, как *Wrms* — протяженность спектра сигнала.

Можно сопроводить полученные выводы следующими физическими комментариями. Во-первых, повышение точности за счет увеличения длительности T легко объяснимо: частота есть попросту скорость изменения полной фазы, и, как и любая скорость, измеряется надежнее, когда приращение угла наблюдается на более протяженном интервале.

Точность оценки частоты

Можно сопроводить полученные выводы следующими физическими комментариями.

Во-первых, повышение точности за счет увеличения длительности T легко объяснимо: **частота есть попросту скорость изменения полной фазы**, и, как и любая скорость, измеряется надежнее, **когда приращение угла наблюдается на более протяженном интервале**.

Во-вторых, в сопоставлении измерений времени и частоты **наглядно проявляется частотно-временная дуальность**. Действительно, в то время как **точность измерения времени зависит от** протяженности сигнала в частотной области (**полосы W**), **точность измерения частоты управляется** протяженностью сигнала во временной области (**длительностью T**).

Одновременная оценка запаздывания и частоты.

Обсудим теперь ситуацию, когда как временной τ, так и частотный F сдвиги принятого сигнала являются неизвестными информативными параметрами, т.е. подлежат измерению.

Подобная задача соответствует многим реальным сценариям в цифровых телекоммуникациях, например:

1. В системах мобильной связи 2-го и 3-го поколений прием, как правило, начинается с синхронизации местного опорного колебания с принятым сигналом. Данная операция состоит в измерении частотно-временного рассогласования локального эталона с приходящим сигналом и последующей частотно-временной подстройки первого до входа в синхронизм со вторым.

Одновременная оценка запаздывания и частоты.

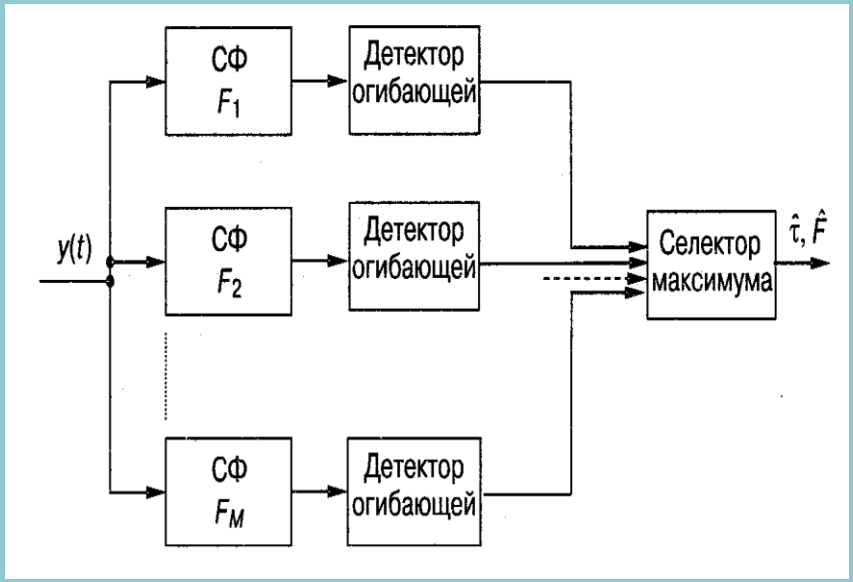
2. В радиолокации совместное измерение расстояния и скорости цели относительно приемника означает одновременное оценивание запаздывания и доплеровской частоты сигнала.

3. В навигации, например, системе GPS, оценки аналогичного характера необходимы как исходные данные для определения местоположения и скорости пользователя.

Одновременная оценка запаздывания и частоты.

2. В радиолокации совместное измерение расстояния и скорости цели относительно приемника означает одновременное оценивание запаздывания и доплеровской частоты сигнала.

3. В навигации, например, системе GPS, оценки аналогичного характера необходимы как исходные данные для определения местоположения и скорости пользователя.


СФ
F_1
Детектор
огибающей
$y(t)$
СФ
F_2
Детектор
огибающей
Селектор
максимума
$\hat{\tau}, \hat{F}$
СФ
F_M
Детектор
огибающей