

# Глава 2. Теория оптимальной линейной фильтрации сигналов

## Общие сведения

Фильтрация сигналов — фундаментальное понятие, связанное с представлениями об их спектре, т.е. разложении сигналов в ряд по тригонометрическим функциям. Такие разложения рассматривали великие математики прошлого И. Бернулли и Л. Эйлер. Ж. Фурье пользовался этими рядами для решения сложной задачи — интегрирования уравнения теплопроводности. Метод Фурье стал классическим приемом решения задач математической физики, и в его честь эти ряды назвали рядами Фурье. Долгое время ряды Фурье не связывались с какими-либо физическими представлениями. Даже после открытия электрических колебаний и волн многими учеными высказывались сомнения в адекватности разложений Фурье (спектров сигналов) происходящим физическим явлениям. Великий физик Г. Герц относился к спектральным представлениям отрицательно. Всемирно известный изобретатель вакуумного диода английский ученый А. Флеминг также не признавал существования спектра у сигналов и в 20-х гг. XX в. между учеными велась бурная дискуссия о реальности этого понятия. В этой дискуссии приняли активное участие крупные отечественные (Л.И. Мандельштам и С.М. Рытов) и зарубежные (Дж. Р. Карсон и др.) физики [1]. В результате спектральные представления утвердились в радиотехнике. С этого времени инженеры стали осознанно широко применять фильтры для разделения сигналов, передаваемых в разных частотных каналах.

Однако важность проблемы фильтрации сигналов для повышения чувствительности приемных устройств (что позволяет при наличии шумов принимать с высоким качеством сигналы, уровень которых на входе приемника минимален) была осознана только в начале 40-х гг., когда бурно стала развиваться теория, позволяющая синтезировать оптимальные фильтры и оценивать потенциальные возможности повышения чувствительности приемных устройств.

## § 1. Синтез согласованных фильтров

Одной из важнейших проблем является прием сигналов известной формы на фоне флуктуационных шумов. Она возникает в системах передачи дискретных сигналов, в системах радиолокации, радионавигации и т.п.

Повышение чувствительности приемников слабых сигналов стало особенно актуальным в период Второй мировой войны. В эти годы стали выпускать радиолокационную технику, работающую в метровом и дециметровом диапазонах частот. В этих диапазонах, в отличие от низкочастотных, в которых основным видом шумов являются атмосферные и промышленные помехи, имеющие импульсный характер, чувствительность приема ограничивается собственными флуктуационными шумами приемника.

Для повышения чувствительности радиолокаторов необходимо было в устройствах, обрабатывающих принимаемые импульсные сигналы, фильтры выбирать таким образом, чтобы на их выходе отношение пикового уровня сигнала к мощности шума имело максимально возможное значение. Синтез оптимальных фильтров приводит к вариационной задаче, в которой заданы спектр детерминированного полезного сигнала и энергетический спектр случайного шума, действующего на входе приемника, а частотная характеристика фильтра неизвестна. В такой постановке данная задача независимо рассматривалась в середине 40-х гг. американскими учеными Д.О. Нортон [2], а также **Д. Миддлтоном** и Дж. Г. Ван Флеком. Было показано, что тогда, когда случайный шум является белым, частотная характеристика оптимального фильтра является функцией, комплексно сопряженной спектру выделяемого импульса, а его переходная характеристика является функцией, инверсной во времени функции, описывающей выделяемый полезный сигнал. Такой фильтр стали называть согласованным. Данный результат имел важное практическое значение и был первым результатом по синтезу оптимальных фильтров, выделяющих полезный сигнал с минимальной погрешностью. Установлено, что при выделении сигнала, имеющего спектр  $S(j\omega)$ , принимаемого на фоне шумов с энергетическим спектром  $S_n(j\omega)$ , частотная характеристика согласованного фильтра определяется формулой:

$$K(j\omega) = c[S^*(j\omega) / S_n(j\omega)] \exp(-j\omega\tau),$$

где  $c$  — произвольная постоянная, а  $\tau$  — момент времени, соответствующий наибольшему отношению пикового значения сигнала к мощности шума на выходе фильтра ( $\rho$ ). При этом максимально возможное значение  $\rho$  следующее:

$$\rho = \frac{1}{2\pi} \int_{-\infty}^{\infty} [|S(j\omega)|^2 / S_n(j\omega)] d\omega.$$

Приведенные простые формулы относятся к числу ключевых соотношений в теории связи, которыми пользуются инженеры при конструировании систем связи и радиолокации.

Подобную же задачу рассмотрел в 1946 г. крупный отечественный ученый, член-корреспондент АН СССР **В.И. Сифоров** [3]. Его подход, который может быть назван параметрическим, носил более частный характер. Он рассматривал фильтры с определенной формой частотной характеристики, которые имели ряд параметров (полоса пропускания, крутизна спада частотной характеристики за пределами этой полосы и т.п.). Подбором этих параметров достигался максимум отношения сигнал/шум на выходе фильтров.

Данное направление теории согласованной фильтрации детерминированных сигналов развивалось и в последующие годы. В 1950 г. Б.М. Дворком [4] была выполнена обобщающая работа, касающаяся синтеза таких фильтров. Большое значение имела статья крупного американского ученого Дж. Л. Турина [5], в которой дан обширный аналитический обзор научных работ по согласованным фильтрам, опубликованных к 1960 г.

Важные исследования методов построения согласованных фильтров были выполнены в 1957–1966 гг. профессором Ю.С. Лезиным. В монографии [6] были систематически изложены свойства согласованных фильтров, предназначенных для обнаружения сигналов на фоне гауссовских шумов, исследованы методы их построения и

определена эффективность их применения. В частности, в ней рассмотрены вопросы построения гребенчатых фильтров, оптимальных для выделения из шумов последовательности импульсных сигналов, а также вопросы построения согласованных фильтров для импульсных широкополосных сигналов с большой базой и низким уровнем боковых лепестков в их корреляционной функции. К таким сигналам относятся, например, сигналы с линейной частотной модуляцией и радиосигналы, фаза которых манипулирована кодом Баркера.

Теория согласованных фильтров является важным разделом теории связи, так как такие фильтры — необходимые элементы оптимальных систем приема любых дискретных сигналов на фоне гауссовских шумов.

## § 2. Синтез оптимальных фильтров для выделения гауссовских случайных сигналов, принимаемых на фоне гауссовских шумов

Первую основополагающую работу по теории линейной фильтрации в дискретном времени опубликовал в 1939 г. академик **А.Н. Колмогоров** [7]. Следует отметить, что эта работа носила чисто математический характер.

Во время Второй мировой войны подобная же теория была независимо разработана крупным американским математиком **Н. Винером**. Он рассмотрел задачи линейной фильтрации сигналов, а также их экстраполяции и интерполяции для непрерывного времени и опубликовал свои результаты в книге [8], изданной в 1949 г., где указал на важность теории для специалистов в области радиотехники, решающих задачи, связанные с выделением сигналов на фоне шумов.

Основой теории Винера—Колмогорова была спектральная теория случайных процессов, развившаяся из фундаментальной работы **А.Я. Хинчина** [9], в которой было установлено, что корреляционная функция случайного процесса и его энергетический спектр мощности связаны преобразованием Фурье. Эта теория рассматривала оптимальный прием сигнала  $s(t)$  на фоне гауссовского шума:

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

где  $s(t)$  и  $n(t)$  — гауссовские случайные процессы с нулевым средним и известными корреляционными функциями  $R_s(\tau)$  и  $R_n(\tau)$ . Сигнал  $r(t)$  поступает на вход линейного фильтра с импульсной характеристикой  $h(t)$ . На выходе фильтра формируется сигнал

$$s^*(t) = \int_0^{\infty} h(\tau)r(t-\tau)d\tau.$$

Импульсная характеристика оптимального фильтра выбирается таким образом, чтобы сигнал  $s^*(t)$  в минимальной степени (в смысле среднеквадратичной ошибки) отличался от  $s(t + \tau)$ . При  $\tau = 0$  осуществляется оптимальная фильтрация (сглаживание), при  $\tau > 0$  — его оптимальная экстраполяция (упреждающая фильтрация), при  $\tau < 0$  — его оптимальная интерполяция (фильтрация с запаздыванием).

Согласно теории Винера—Колмогорова функция  $h(t)$  определяется из интегрального уравнения Винера—Хопфа:

$$\int_0^{\infty} h(t)[R_s(\Delta-t) + R_n(\Delta-t)]dt = R_s(\Delta + \tau).$$

В общем случае произвольного значения  $\tau$  определение  $h(t)$  и ошибки оценки  $s(t + \tau)$  являются сложными в математическом отношении задачами. Однако для случая, когда  $\tau \ll 0$  (т.е. когда при оценке процесса  $s(t)$  учитываются его значения как до, так и после момента времени  $t$ ), задача существенно упрощается и частотная характеристика оптимального фильтра  $K(j\omega)$  и ошибка фильтрации  $\sigma_0^2$  определяются простыми формулами:

$$K(j\omega) = \frac{S(\omega) \exp(j\omega\tau)}{S(\omega) + S_n(\omega)};$$

$$\sigma_0^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S(\omega)S_n(\omega)}{S(\omega) + S_n(\omega)} d\omega.$$

Теория, изложенная в книге Винера, была очень сложна для понимания инженерами, которые в те годы не имели для этого необходимой математической подготовки. Известные американские ученые Г.В. Боде и **К. Шеннон** в 1950 г., используя интуитивно понятные соображения, дали упрощенное изложение этой теории [10]. Они показали, что, применяя «обесцвечивающий» минимально-фазовый фильтр (с импульсной характеристикой, равной нулю при отрицательных значениях  $t$ ), модуль частотной характеристики которого равен  $Y_1(j\omega) = [S(\omega) + S_n(\omega)]^{-1/2}$ , процесс, действующий на входе приемника, превращается в «белый» шум — последовательность коротких независимых импульсов, имеющих равномерный спектр мощности. Вес, с которым должны учитываться отдельные составляющие этого спектра в сигнале, формируемом на выходе оптимального фильтра, равен  $S(\omega)$ . Если были бы известны значения функции  $r(t)$  от  $t = -\infty$  до  $t = \infty$ , то наилучшей операцией, примененной ко входу, была бы следующая:

$$Y_2(j\omega) = Y_1(j\omega) \cdot S(\omega).$$

Однако импульсная характеристика такого фильтра  $h_2(t)$  физически нереализуема, так как содержит «хвосты», простирающиеся от  $t = -\infty$  до  $t = \infty$ . Учитывая, что единственные имеющиеся у нас данные для предсказания значения  $s(t + \tau)$  к моменту  $\tau = 0$  — это импульсы, появившиеся в прошедшей истории процесса, то всем импульсам, которые появятся в будущем, следует приписать нулевой вес, а тем, которые уже появились на входе приемника, — вес  $h_2(t)$ . Таким образом, после «обесцвечивающего» минимально-фазового фильтра должен быть включен фильтр с импульсной характеристикой

$$h_3(t) = \begin{cases} h_2(t + \tau) & \text{при } t \geq 0 \\ 0 & \text{при } t \leq 0 \end{cases}.$$

Функции  $h_3(t)$  соответствует частотная характеристика  $Y_3(j\omega)$ . Поэтому частотная характеристика оптимального сглаживающего и предсказывающего фильтра  $Y_{\text{опт}}(j\omega) = Y_1(j\omega) \cdot Y_3(j\omega)$ .

В 1952 г. профессор А.М. Яглом — ученик А.Н. Колмогорова опубликовал обширную статью [11], в которой было впервые дано простое и в то же время математически строгое изложение теории экстраполяции и фильтрации стационарных случайных последовательностей и процессов.

Американские ученые Л. Заде и Дж. Рагазини [12] сделали существенное обобщение теории Винера на случай, когда сигнал помимо случайной содержит еще и регулярную составляющую, а время наблюдения ограничено.

Для решения практических задач, связанных с конструированием оптимальных систем связи и управления, большое значение имела формула, определяющая точность оценки процесса  $s(t)$  для случая, когда шум  $n(t)$  является «белым» и имеет спектральную плотность мощности  $N_0$ :

$$\sigma_0^2 = \frac{N_0}{4\pi} \int_{-\infty}^{\infty} \log\left[1 + \frac{S(\omega)}{N_0/2}\right] d\omega.$$

Эта формула была установлена в 1955–1965 гг. разными методами американскими учеными М.К. Йовитсом и Дж. Л. Джексоном [13], К.В. Хелстромом [14], Э.Д. Витерби [15] и Д. Снайдером [16].

Теория оптимальной линейной фильтрации была обобщена на случай многомерных случайных процессов Н. Винером [17], Е. Вонгом и Дж. Томасом [18] и другими учеными. Эти результаты могут быть применены, в частности, к системам передачи телевидения со строчной разверткой, в которых имеется высокая корреляция между сигналами различных строк и кадров.

С 50-х гг. теорию Винера–Колмогорова начали излагать в доступном виде для инженеров, в монографиях и учебниках. Одними из первых книг, в которых она нашла отражение, были книги отечественного ученого В.В. Солодовникова [19] (1952 г.) и ученика Винера Ю.В. Ли [20] (1960 г.).

Позже во всех монографиях по статистической радиотехнике, написанных отечественными (Б.Р. Левин и В.И. Тихонов [21, 22]) и зарубежными (Д. Миддлтон, Э. Витерби, Г. Ван Трис [23–25] и др.) учеными, один из разделов был посвящен этой теории.

Совершенно другой подход к проблемам выделения сигнала на фоне шума, который применим и к проблемам оптимальной линейной фильтрации сигналов, был предложен Р.Л. Стратоновичем в 1959 г. [26]. Теория Р.Л. Стратоновича основывалась на представлении случайных процессов, моделирующих как полезный сигнал, так и шум с помощью дифференциальных уравнений (уравнений состояния). Независимо от Р.Л. Стратоновича законченные результаты оптимальной линейной фильтрации как в дискретном, так и в непрерывном времени получили в 1961 г. американские ученые Р.Е. Калман и Р.С. Бьюси [27, 28].

Для гауссовских и марковских случайных процессов Р.Л. Стратоновичем, Р.Е. Калманом и Р.С. Бьюси были выведены дифференциальные уравнения, определяющие структуру оптимального фильтра, на вход которого поступает принимаемый сигнал, и матричное уравнение Риккати, определяющее точность его оценки. Наличие дифференциальных уравнений оценки вместо интегрального представляет определенные практические преимущества, так как дифференциальные уравнения решаются намного легче с помощью аналоговой или цифровой техники, чем интегральное уравнение.

В 70-х гг. появляются монографии, в которых излагается новая теория оптимальной линейной фильтрации, и она применяется к решению ряда технических проблем.

Теория оптимальной линейной фильтрации сигналов применима к широкому классу задач, возникающих при конструировании систем связи и управления. Ее значение для теории связи и в том, что она позволяет в ряде случаев установить предельную точность  $\sigma_0^2$  оценки полезного сообщения, передаваемого по каналу связи. В реальных системах стремятся применить достаточно простые фильтры, параметры которых выбирают так, чтобы в данной системе точность  $\sigma_1^2$  оценки полезного сообщения была максимальной. Рациональность применения выбранного фильтра с заданными параметрами можно оценить, сравнивая  $\sigma_0^2$  и  $\sigma_1^2$ . Если эти величины отличаются незначительно, то выбор фильтра следует считать удачным. Кроме того, важно и то, что более общая теория нелинейной фильтрации, связанная с созданием оптимальных демодуляторов сигналов с фазовой, частотной и другими видами нелинейной модуляции, также приводит к необходимости синтеза оптимальных линейных фильтров, входящих в состав этих демодуляторов.

### § 3. Хронология развития теории оптимальной линейной фильтрации сигналов

1. Создание теории оптимальной линейной фильтрации, экстраполяции и интерполяции гауссовских случайных сигналов, принимаемых на фоне гауссовского шума (1939 г. — А.Н. Колмогоров, 1949 г. — Н. Винер).
2. Синтез оптимальных фильтров, согласованных со спектром детерминированного полезного сигнала, принимаемого на фоне случайного шума с известным спектром мощности (1943 г. — Д.О. Норт, Д. Миддлтон и Дж. Г. Ван Флек, 1946 г. — В.И. Сифоров).
3. Развитие теории согласованной фильтрации детерминированных сигналов (1950 г. — Б.М. Дворк).
4. Развитие теории оптимальной линейной фильтрации Колмогорова–Винера (1952 г. — А.М. Яглом).
5. Обобщение теории Колмогорова–Винера в случае, когда сигнал помимо случайной содержит еще и регулярную составляющую, а время наблюдения ограничено (1952 г. — Л. Заде и Ж. Рагазини).
6. Вывод формулы, определяющей точность оценки сигнала для случая, когда шум является «белым» (1955 — 1965 гг. — М.К. Иовитс и Дж. Л. Джексон, К.В. Хелстром, Э.Д. Витерби, Д. Снайдер).
7. Разработка методов построения согласованных фильтров, в том числе гребенчатых фильтров, предназначенных для выделения на фоне шумов пачек радиолокационных импульсов (1957–1966 гг. — Ю.С. Лезин).
8. Обобщение теории оптимальной линейной фильтрации для случая выделения на фоне шумов многомерных случайных процессов (1958 г. — Н. Винер, 1961 г. — Е. Вонг и Дж. Томас).
9. Создание марковской теории оптимальной линейной фильтрации (1959 г. — Р.Л. Стратонович, 1961 г. — Р.Е. Калман и Р.С. Бьюси).

**Литература к главе 2**

1. Функ Л.М. Сигналы, помехи, ошибки... — М.: Связь, 1978.
2. North D.O. Analysis factors which determine signal-noise discrimination in pulsed carrier systems. RCA Tech. Rep. PTR-6C. June 1943 (Proc. IEEE. 1963. № 7).
3. Сифоров В.И. О влиянии помех на прием импульсных сигналов // Радиотехника. 1946. № 1.
4. Dwork B.M. Detection of pulse superimposed on fluctuation noise // Proc. IRE. 1950. P. 771.
5. Turin G.L. An introduction to marched filters // IRE Trans. on Information Theory. 1960. № 3.
6. Лезин Ю.С. Оптимальные фильтры и накопители импульсных сигналов. — М.: Сов. радио. 1969.
7. Колмогоров А.Н. Интерполирование и экстраполирование стационарных случайных последовательностей // Изв. АН СССР. Сер. математическая. 1941. № 5.
8. Wiener N. The extrapolation, interpolation and smoothing of stationary time series. N.Y.: Wiley. 1949.
9. Хинчин А.Я. Теория корреляции стационарных стохастических процессов // Успехи мат. наук. 1938. Вып. 5.
10. Vode H.W., Shannon C.E. A simplified derivation of linear least-square smoothing and prediction theory // Proc. IRE. 1950. № 4.
11. Яглом А.М. Введение в теорию стационарных случайных функций // Успехи мат. наук. 1952. Т. 7. Вып. 5.
12. Zadeh L.A., Ragazzini J.R. Optimum filters for the detection of signals in noise // IRE Trans. 1952. V. IT-1. № 4.
13. Yovits M.C., Jackson J.L. Linear filter optimization with game theory consideration // IRE Nat. Conv. Record. March 1955.
14. Helstrom C.W. Topics in the transmission of continuous information. Westinghouse Res. Labs. Repr. 64-8C3-522-R1. August 1964.
15. Viterbi A.J. On the minimum mean square error resulting from linear filtering of stationary signals in white noise // IEEE Trans. on Information Theory. 1965. № 3.
16. Snyder D.L. Some useful expression for optimum filtering in white noise // Proc. IEEE. 1965. № 6.
17. Wiener N., Massani P. The prediction theory of multivariable stochastic processes. Acta Math. June 1958.
18. Wong E., Thomas J.B. On the multidimensional filtering and prediction problem and the factorization of spectral matrices // J. Franklin Inst. 1961. V. 8. P. 87–99.
19. Солодовников В.В. Введение в статистическую динамику систем автоматического регулирования. — М.; Л.: Гостехиздат, 1952.
20. Lee Y.W. Statistical theory of communication. N.Y.: Wiley. 1960.
21. Левин Б.Р. Теоретические основы статистической радиотехники. — М.: Сов. радио, 1968. Т. 2.
22. Тихонов В.И. Оптимальный прием сигналов. — М.: Радио и связь, 1983.
23. Middleton D. An introduction to statistical communication theory. N.Y.: McGraw-Hill. 1961. Pt. 2.
24. Viterbi A.J. Principles of coherent communication. N. Y.: McGraw-Hill. 1966.
25. Van Trees H.L. Detection, Estimation and Modulation Theory. N. Y.: Wiley. 1971. Pt. 1.
26. Стратонович Р.Л. К теории оптимальной нелинейной фильтрации случайных функций // Теория вероятностей и ее применение. 1959, Т. 4, № 2.
27. Kalman R.E. New methods and results in linear prediction and filtering theory. Baltimore: RIAS Tech. Rep. 1961. P. 1–61.
28. Kalman R.E., Bucy R. New results in linear filtering and prediction theory // ASME J. Basis Eng. March 1961. V. 83.