

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Московский государственный институт электроники и математики
(технический университет)

В.И. Капалин Г.Ю. Кудряшов

ВВЕДЕНИЕ В ТЕОРИЮ СИСТЕМ И ТЕОРИЮ
УПРАВЛЕНИЯ

Утверждено Редакционно-издательским советом института
в качестве учебного пособия

Москва 2002

УДК 007.5/075.8/
ББК 32.815я73
К20

Рецензенты: докт. техн. наук В.Е. Болнокин (Институт машиноведения РАН);
докт. физ.-мат. наук Ю.Л. Хотунцев (Московский государственный педагогический университет)

Капалин В.И. Кудряшов Г.Ю.

К20 Введение в теорию систем и теорию управления. Учебное пособие — Московский государственный институт электроники и математики. М., 2002.— 153 с.

ISBN

Учебное пособие содержит изложение основ математической теории систем и теории автоматического управления. Рассматривается теория моделей типа «вход-выход», задаваемых бинарными отношениями, линейными интегральными операторами, функциональными полиномами и рядами. Приведены сведения по теории преобразований Лапласа и Фурье, краткие сведения об обобщённых функциях. Рассмотрен метод передаточных функций и основанные на нём классические методы синтеза — частотные, корневые и с использованием интегральных оценок. Изложены основы применения равномерных аппроксимаций для построения линейных и полиномиальных моделей нелинейных систем.

Пособие предназначено для студентов IV курса факультета прикладной математики, изучающих курс теории систем, а также для аспирантов, специализирующихся в области теории систем и теории автоматического управления.

ISBN

УДК 007.5/075.8/
ББК 32.815я73
©Капалин В.И., Кудряшов Г.Ю., 2002

Введение.

Кибернетика, теория управления и теория систем

Термин «кибернетика» был введён в современную науку Н. Винером, определившим кибернетику как «науку об управлении и связи в живых организмах и машинах» [3]. Этот термин образован от греческого слова «кибернэтэс», что означает «рулевой», «кормчий». Уже Платон использовал этот термин не только в смысле управления кораблём, но и в смысле управления людьми. А. Ампер называл «кибернетикой» общую науку об управлении государством. Существенный вклад в развитие кибернетики по Винеру с точки зрения биологии внесли работы У. Эшби [26].

Основной идеей, выдвинутой Н. Винером, была идея о том, что биологические организмы функционируют на основе принципа отрицательной обратной связи точно также, как функционируют системы автоматического управления, например, системы управления кораблём, самолётом или турбиной. Это открывало новые перспективы для применения математических методов и, прежде всего, методов теории управления, в биологических исследованиях.

К моменту выхода книги Н. Винера «Кибернетика» (1948 г.) теория автоматического управления была уже сформировавшейся наукой, её практические начала относят ещё к XVIII в. и связывают с регуляторами паровых машин И.И. Ползунова и Дж. Уатта. Её теоретические начала связывают с работой К. Максвелла о регуляторах. При выборе термина, определяющего название новой науки, и, учитывая ту роль, в которой в ней должна занимать теория управления, Н. Винер использовал то обстоятельство, что термин «governor» — регулятор, использованный Максвеллом, происходит от латинского варианта термина «кибернетес» [6].

Основной метод исследования, применяемый в теории управления и, соответственно, в кибернетике — это метод математического моделирования. При этом изучаемому объекту ставится в соответствие его математическая модель, отражающая существенные для исследователя стороны протекающих в объекте процессов управления. Поскольку объекты изучаются не с точки зрения их материальной природы, а с точки зрения интересующих исследователя процессов управления, то объектам разной природы могут сопоставляться одни и те же или однотипные математические модели.

Проблемы разработки единой теории математических моделей начали рассматриваться в работах по математической теории систем М. Месаро-

вича [16], Л. Заде [8], Р. Калмана [9], причём последние две были тесно связаны и ориентированы прежде всего на теорию линейных систем управления. Связь этого подхода с нелинейными задачами теории управления рассматривалась в [18].

Для математической теории систем исходным понятием является понятие абстрактной системы. Абстрактная система представляет собой обобщённую математическую модель, которая может быть введена по-разному. Накладывая на неё различные условия, мы получим различные виды конкретных моделей, отличающихся уровнями детализации, общностью, а также классами практических задач, на которые эти конкретные модели ориентированы.

В данной работе за основу взят теоретико-множественный подход работ [8], где абстрактная система определяется как бинарное отношение, заданное на множестве входных и выходных сигналов. Этот подход аналогичен подходу, применяемому в математической логике [15]. Однако, в отличие от последнего, здесь рассматриваются отношения, устанавливающие связи между функциями времени — входными и выходными сигналами так называемой абстрактной временной системы.

Абстрактная временная система при таком подходе является, фактически, обобщённой математической моделью «чёрного ящика» кибернетики. «Чёрным ящиком» называют объект, для которого можно получить реакции на всевозможные входные сигналы, но для которого нет информации о его внутреннем устройстве. Такой объект может исследоваться только с точки зрения внешнего описания, описания «вход-выход».

Этот подход хорошо стыкуется с принятым в теории управления операторным подходом [6, 2], обобщением которого он и является. Поэтому основное содержание данного учебного пособия связано как с рассмотрением классического аппарата теории передаточных функций, так и методов теории нелинейных систем, внешние описания которых задаются в виде функциональных полиномов и рядов Вольтерра. Использование теории равномерных аппроксимаций для приближения нелинейных характеристик систем, принятое в настоящей работе, позволило дать общий подход для построения линейных и полиномиальных моделей нелинейных систем, что снимает противоречие при изложении теории линейных и нелинейных систем, характерное для литературы по теории управления.

В пособии рассматриваются инженерные показатели качества систем управления и классические методы инженерного синтеза по заданным показателям качества. Поскольку в этих методах используется аппарат передаточных функций, здесь приводятся и основные сведения по теории

преобразований Фурье и Лапласа, в том числе и преобразований обобщённых функций. Метод пространства состояний, основанный не на аппарате передаточных функций, а на уравнениях системы в пространстве состояний, здесь не рассматривается, поскольку ему посвящено изданное ранее учебное пособие первого автора [10]. Оба эти пособия отражают основное содержание курса лекций по теории систем, читаемого для студентов, специализирующихся по кафедре кибернетики факультета прикладной математики МГИЭМ на 7-м и 8-м семестрах обучения. В связи с такой ориентацией в пособии основной акцент сделан на математические аспекты рассматриваемых проблем, а не на технические стороны решения задач управления.

Г Л А В А 1 .

Основные сведения о системах управления

§1. Понятие абстрактной временной системы

При исследовании объектов окружающего мира и протекающих в них процессов исследователь использует метод моделирования, при котором изучаемому объекту или процессу ставится в соответствие некоторая модель, отражающая существенные для исследователя стороны объекта или процесса. Если эта модель адекватна, то есть если она правильно отражает наблюдавшиеся в экспериментах данные, то на её основе можно построить теорию, позволяющую правильно предсказывать результаты будущих экспериментов.

В математической литературе, в частности, посвящённой математической логике и основаниям математики, слово «теория» употребляется в специальном смысле. Теория, в общем случае, это язык \mathcal{L} вместе с множеством предложений (формул) языка \mathcal{P} . Принятый в данной работе язык теории систем — это математический язык теории множеств и рассматриваемая теория систем является математической теорией систем. Исходным понятием здесь является понятие *абстрактной временной системы*.

Пусть имеется некоторый объект, внутреннее устройство которого неизвестно, и, соответственно, являющийся для исследователя «чёрным ящиком». Исследователь может изучать его, подавая различные воздействия и регистрируя соответствующие воздействиям реакции объекта. Будем предполагать, что эксперимент проводится многократно в том смысле, что условия последнего (в принципе) многократно воспроизводимы и можно наблюдать реакции «чёрного ящика» на различные входные сигналы для тех же самых условий эксперимента [1].



Рис. 1.1.

Схематически эксперимент с «чёрным ящиком» показан на рис. 1.1. Здесь u — воздействие, или, как принято говорить, входной сигнал, из некоторого множества входных сигналов Ω , а y — реакция, выходной сигнал из некоторого множества выходных сигналов Γ . Тогда моделью «чёрного ящика» для данных условий эксперимента будет таблица соответствия между входными и выходными сигналами или, в общем случае, некоторое бинарное отношение.

Определение 1. Пусть даны два множества A и B , которые могут и совпадать. *Декартовым*, или *прямым*, произведением этих множеств называется множество

$$A \times B \stackrel{\text{def}}{=} \{(x, y) \mid x \in A, y \in B\}.$$

Определение 2. Любое подмножество $S \subseteq A \times B$ называется *бинарным отношением*.

Определение 3. *Сечением* $S(a)$ бинарного отношения S по элементу $a \in A$ называется множество элементов $b \in B$ таких, что пары $(a, b) \in S$,

$$S(a) \stackrel{\text{def}}{=} \{b \in B \mid (a, b) \in S\}.$$

Отметим, что задание бинарного отношения ни в коем случае не предполагает функциональной, однозначной связи элементов множеств A и B . В связи с этим дадим следующее

Определение 4. *Функциональным отношением*, или *функцией* (см. [17]), называется такое бинарное отношение, в котором для любого $a \in A$ такого, что $(a, b) \in S$ $S(a) = \{b\}$; то есть для любого a , «участвующего» в бинарном отношении, существует единственный связанный с ним элемент $b \in B$.

Если множества A и B конечные, то бинарное отношение может быть задано перечислением всех пар $(a, b) \in S$, таблично или графически (стрелками) [23].

Рассмотрим простейший пример эксперимента с «чёрным ящиком», на вход которого подаются два входных сигнала, обозначаемые 0 (отсутствие стимула) и 1 (наличие стимула), а на выходе которого возможны две реакции — 0 (отсутствие реакции) и 1 (наличие реакции). Пусть в процессе эксперимента на воздействие 0 наблюдалась реакция 0, а на воздействие 1 наблюдались реакции 0 и 1. В этом случае

$$A \times B = \{(0, 0), (0, 1), (1, 0), (1, 1)\}.$$

Бинарное отношение $S \subseteq A \times B$, описывающее поведение «чёрного ящика» с помощью перечисления пар, можно задать, например, так:

$$S = \{(0,0), (1,0), (1,1)\}.$$

Это же бинарное отношение можно задать таблицей, как показано на рис. 1.2, а, и стрелками, как показано на рис. 1.2 б. В первом случае пары

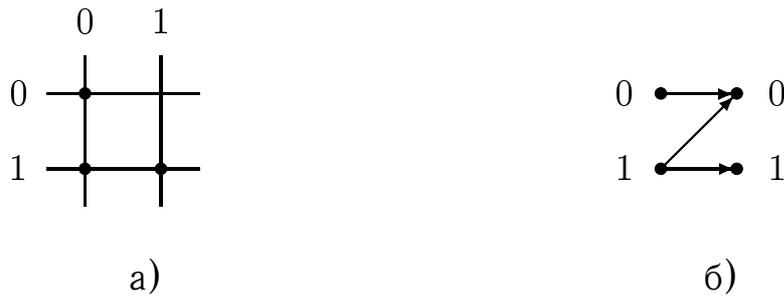


Рис. 1.2. Варианты задания бинарного отношения.

(a, b) показаны жирными точками, во втором случае эти пары соединены стрелками.

Рассмотренные способы задания бинарных отношений удобны только в том случае, если множества A и B — конечные. В рассматриваемых в данной работе случаях эти множества являются функциональными пространствами, элементами которых являются функции времени. Бинарные отношения в этих случаях обычно задаются аналитически, а для задания множеств входных и выходных сигналов пользуются специальными методами.

Определение 5. Непустые линейно упорядоченные множества $T_u \subseteq \mathbb{R}$ и $T_y \subseteq \mathbb{R}$, где \mathbb{R} — множество действительных чисел, назовём *множествами моментов времени* для входных и выходных сигналов соответственно.

Эти множества указывают, в какие моменты времени определены (измеряются) входные и выходные сигналы. В большинстве случаев $T_u = T_y = T$, что и будем предполагать далее.

Определение 6. Непустые множества U и Y назовём *множествами значений входных и выходных сигналов* соответственно. Если U и Y — конечные множества, то они, как правило, называются *алфавитами*. Мы будем рассматривать случай, когда $U \subseteq \mathbb{R}^m, m > 0$ и $Y \subseteq \mathbb{R}^l, l > 0$.

Обозначим через U^T множество функций $u : T \rightarrow U$, а через Y^T — множество функций $y : T \rightarrow Y$. Множества U^T и Y^T содержат всевозможные функции, поэтому, учитывая то, что на входные и выходные сигналы обычно накладываются некоторые ограничения, введём

Определение 7. Множества $\Omega \subseteq U^T$ и $\Gamma \subseteq Y^T$ назовём *множествами входных и выходных сигналов* соответственно.

Элементами множества Ω являются входные сигналы u , а элементами множества Γ — выходные сигналы y (точки в функциональных пространствах). Мгновенные значения этих сигналов для $t \in T$ будем записывать как обычно, т.е. $u(t) \in U$ и $y(t) \in Y$.

Определение 8. Связь между элементами множеств входных и выходных сигналов будем задавать с помощью бинарного отношения $S \subseteq \Omega \times \Gamma$, называемого *внешним описанием* или *описанием типа «вход-выход»*.

Теперь, подводя итог, можно сформулировать определение абстрактной временной системы.

Определение 9. Пусть даны: множество моментов времени T , множества значений входных и выходных сигналов U и Y , множества входных и выходных сигналов Ω и Γ и внешнее описание $S \subseteq \Omega \times \Gamma$. *Абстрактной временной системой* Σ называется следующая шестёрка

$$\Sigma = \langle T, U, Y, \Omega, \Gamma, S \rangle \quad (1.1)$$

Графически абстрактную систему будем представлять так, как показано на рис. 1.3.

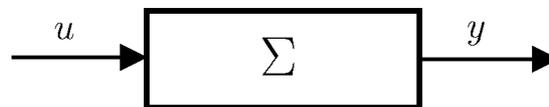


Рис. 1.3.

Абстрактная временная система Σ называется *непрерывной по времени*, если множество моментов времени T является связным подмножеством множества действительных чисел \mathbb{R} , т.е. $\forall \sigma, \tau \in T \quad [\sigma; \tau] \subseteq T$. Абстрактная временная система Σ называется *дискретной по времени*, если T является подмножеством множества целых чисел \mathbb{Z} .

Обычно одному и тому же объекту можно сопоставить как непрерывную по времени, так и дискретную модель (см. ниже модель нервной клетки). Выбор модели определяется задачами исследования и практической реализации — так, для моделирования на компьютере, реализации цифровых усилителей, фильтров и т.п. необходимо использовать дискретную модель, для аналитического моделирования и построения аналоговых (механических или электрических) систем — непрерывную.

При задании абстрактных временных систем основной акцент делается на задании уравнения, определяющего внешнее описание системы, то есть бинарного отношения S , и сама система может быть определена как бинарное отношение. В ряде случаев это бинарное отношение может являться и функцией (оператором) $f : \Omega \rightarrow \Gamma$.

Определение 10. Абстрактная система Σ называется *функциональной*, если бинарное отношение S является функцией (см. определение 4). Для функциональных абстрактных систем внешнее описание задаётся оператором (функцией) $f : \Omega \rightarrow \Gamma$.

Обычно при задании абстрактных временных систем основной акцент делается на задании уравнения, определяющего внешнее описание системы — бинарного отношения или оператора для функциональных систем, а остальные элементы шестёрки, за исключением, быть может, множеств Ω и Γ , явно не указываются.

Рассмотрим некоторые примеры абстрактных систем.

Пример 1 (Формальный нейрон Мак-Каллока–Питтса [1]). На рис. 1.4 схематически изображён нейрон — нервная клетка, совокупность которых образует нервную систему у животных и человека.

Здесь цифрами отмечены: 1) — тело нейрона, называемое сомой; 2) — ядро нейрона; 3) — короткие отростки нервной клетки, называемые дендритами; 4) — длинный отросток клетки, называемый аксоном.

Рассмотрим основной принцип функционирования нейрона. На дендриты нейрона с других нейронов или рецепторов поступает электрохимическое возбуждение. Некоторые из контактов, подводящих это возбуждение, называемые синапсами, являются возбуждающими, а некоторые — тормозящими возбуждение. Если величина возбуждения превышает величину торможения на некоторый, фиксированный для данного нейрона уровень, называемый порогом срабатывания, то после истечения определённого интервала времени, называемого латентным периодом, протоплазма клетки возбуждается. При этом по аксону пробегает импульс, характер

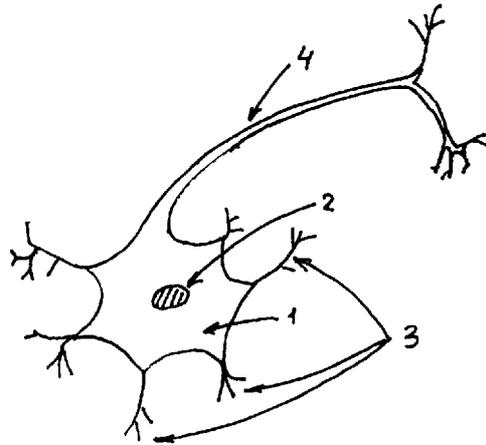


Рис. 1.4.

которого показан на рис. 1.5. После прохождения этого импульса нейрон в течение некоторого интервала времени, называемого рефрактерным периодом, не способен к возбуждению.

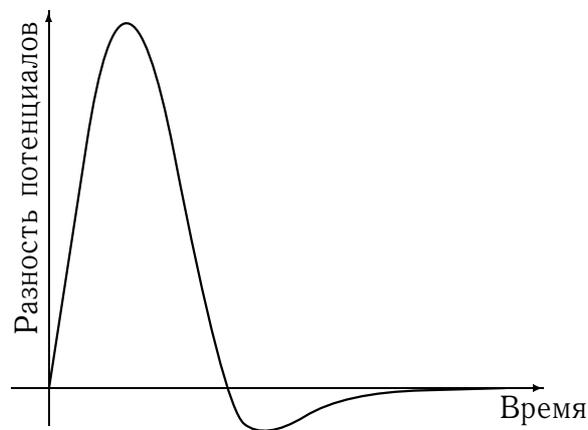


Рис. 1.5.

Построим математическую модель описанной динамики поведения нейрона. Для упрощения будем предполагать, что интервал латентности равен интервалу рефрактерности и величина импульса, пробегающего по аксону, всегда одинакова. Тогда естественно сделать следующие предположения о модели:

1. В качестве множества моментов времени выберем множество целых чисел, т.е. текущая переменная времени $n \in \mathbb{Z}$.

2. Пусть число синапсов есть m . В качестве входного алфавита выберем наборы из 0 (нет возбуждения) и 1 (есть возбуждение) длины m , иными словами, пространство m -мерных двоичных векторов.
3. Входными сигналами будут являться последовательности m -мерных двоичных векторов, а выходными сигналами — последовательности из 0 (аксон не возбуждён) и 1 (аксон возбуждён).
4. Введём натуральное число θ , называемое порогом срабатывания нейрона и веса по каждому синапсу — множество целых чисел $\{\omega_i\}_{i=\overline{1,m}}$. Если $\omega_i > 0$ — синапс возбуждающий, а если $\omega_i < 0$ — тормозящий.

Окончательно модель нейрона задаётся такой формулой (бинарным отношением):

$$y(n+1) = \begin{cases} 1, & \text{если } \sum_{i=1}^m \omega_i u_i(n) > \theta, \\ 0, & \text{если } \sum_{i=1}^m \omega_i u_i(n) < \theta. \end{cases} \quad (1.2)$$

Модель (1.2) называется формальным нейроном Мак-Каллока–Питтса. С помощью сетей из таких нейронов, нейронных сетей, можно строить системы, моделирующие некоторые виды интеллектуальной деятельности, они также широко применяются в системах распознавания образов (подробнее см., например, в [6]).

Формальный нейрон Мак-Каллока–Питтса отражает, разумеется, только некоторые аспекты поведения нервной клетки. В случае, если, например, потребовалось бы каким-либо образом учесть и форму нервного импульса (см. рис. 1.5), то данная модель уже не являлась бы адекватной, и её пришлось бы заменить на непрерывную по времени.

Модель нейрона (1.2), очевидно, задаёт функциональную систему, описание «вход-выход» здесь однозначно.

Пример 2 (Аналого-цифровой преобразователь [2]). Аналого-цифровой преобразователь (АЦП) — это система, преобразующая непрерывные сигналы в дискретную форму, пригодную для обработки на ЭВМ (рис. 1.6).

Входным сигналом здесь является непрерывная функция времени $u(t)$, а выходным — так называемая «решётчатая функция» $u(kT)$, значения которой совпадают со значениями непрерывной функции в дискретные моменты времени $t = kT$, $k \in \mathbb{Z}_+ = \{0, 1, 2, \dots\}$. Внешнее описание такой

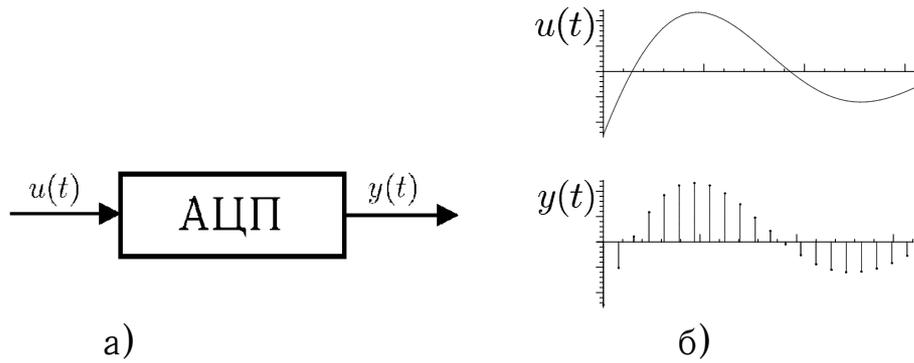


Рис. 1.6.

системы задаётся формулой

$$y(t) = \sum_{k=0}^{+\infty} \hat{\delta}(t - kT)u(t) = u(kT), \quad (1.3)$$

где функция $\hat{\delta}(t)$ есть [2, 18]

$$\hat{\delta}(t) = \begin{cases} 1, & t = 0, \\ 0, & t \neq 0. \end{cases}$$

Для такой системы $T_u = \mathbb{R}_+$, а $T_y = \mathbb{Z}_+$. Такая система (с дискретным и непрерывным временем) является системой смешанного типа.

Пример 3 (Прямолинейное движение судна). Судно массы m движется прямолинейно под действием пропульсивной (движущей) силы F , создаваемой двигателем (рис. 1.7).

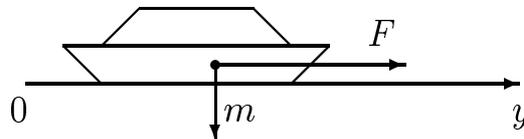


Рис. 1.7.

Введём на прямой координату y . Тогда перемещение судна, рассматриваемого как материальная точка, в соответствии со вторым законом Ньютона, описывается дифференциальным уравнением

$$m\ddot{y}(t) = F(t). \quad (1.4)$$

Если рассматривать судно как «чёрный ящик», ничего не зная о законе преобразования входных сигналов, т.е. пропульсивной силы $F(t)$, в выходные, т.е. координаты центра масс $y(t)$, то одинаковым входным воздействиям $F(t)$ будут соответствовать различные выходные сигналы. Это обусловлено тем, что согласно (1.4) положение точки определяется не только прилагаемой силой, но также начальным положением и начальной скоростью (т.е., начальными условиями).

Таким образом, рассматриваемая система с непрерывным временем не является функциональной, так как для её описания с помощью функций (операторов) требуется ещё ввести информацию о её начальном состоянии — начальном положении и начальной скорости.

Пример 4 (Вращение ротора двигателя). Вращение ротора двигателя вокруг закреплённой оси под действием движущего момента M_g (рис. 1.8) описывается уравнением

$$J\dot{\omega}(t) = M_g - M_c, \quad (1.5)$$

где J — момент инерции ротора, $\omega(t)$ — угловая скорость вращения ротора и M_c — момент сил сопротивления.

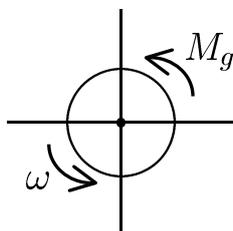


Рис. 1.8.

Как и в предыдущем случае, эта система с непрерывным временем нефункциональна и для её функционального описания требуется иметь информацию о начальной скорости ротора.

Отметим, что в этой и в предыдущей моделях не учтены многие факторы, определяющие движения объектов. Поэтому только эксперимент может показать, насколько адекватны данные модели в конкретных задачах.

§2. Структурные схемы и операции с системами

На практике изучаемый объект обычно не является «чёрным ящиком» и исследователю или инженеру известны математические модели элементов, из которых он состоит и известны взаимосвязи этих элементов, т.е. известна внутренняя структура объекта. Тогда, наряду с описанием «вход-выход», объект может быть задан своей структурной моделью.

Структурная модель представляет собой совокупность математических моделей элементов объекта и модели взаимосвязей элементов. Модель взаимосвязей элементов задаётся графически — структурной схемой (блок-схемой), см. ниже.

Если структурная модель объекта задана и по ней требуется исследовать внешние свойства объекта, т.е. найти описание «вход-выход», то в этом случае можно говорить о задаче структурного анализа.

Если же заданы желаемые внешние свойства объекта, т.е. его модель «вход-выход» и требуется построить структурную модель из некоторого набора элементов, имеющую заданные внешние свойства, то в этом случае говорят о задаче структурного синтеза.

Рассмотрим задачу структурного анализа и введём для её решения операции с абстрактными системами.

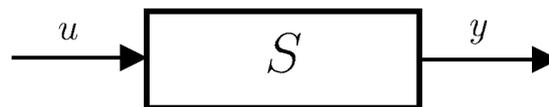


Рис. 2.1.

Пусть дана абстрактная временная система с одним входом и одним выходом, графически представленная на рис. 2.1. Будем предполагать, что рассматриваются только системы, у которых множества входных и выходных сигналов равны ($\Omega = \Gamma$), то есть системы, задаваемые *однородными* бинарными отношениями $S \subseteq \Omega \times \Omega$, причём $U = Y = \mathbb{R}$. Обозначим множество таких абстрактных систем через \mathcal{O} .

Определение 1. Сечение $S(\tilde{A})$ бинарного отношения $S \subseteq A \times B$ по множе-

ству $\tilde{A} \subseteq A$ определяется как объединение сечений $S(a)$ по всем $a \in \tilde{A}$,

$$S(M) \stackrel{\text{def}}{=} \bigcup_{a \in \tilde{A}} S(m).$$

С помощью этого определения можно ввести понятие операции композиции абстрактных систем.

Определение 2. Композиция $S_1 * S_2$ абстрактных систем $S_1, S_2 \in \mathcal{O}$ определяется формулой

$$(S_1 * S_2)(u) \stackrel{\text{def}}{=} S_1(S_2(u)). \quad (2.1)$$

Композиция систем S_1 и S_2 описывает их последовательное соединение, задаваемое структурной схемой на рис. 2.2.

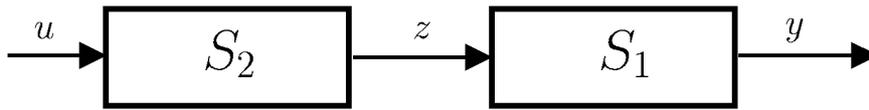


Рис. 2.2.

Операция композиции обладает свойством ассоциативности:

$$S_1 * (S_2 * S_3) = (S_1 * S_2) * S_3, \quad S_1, S_2, S_3 \in \mathcal{O}. \quad (2.2)$$

В самом деле, из определения (2.1) вытекает, что

$$\begin{aligned} (S_1 * (S_2 * S_3))(u) &= S_1((S_2 * S_3)(u)) = S_1(S_2(S_3(u))), \\ ((S_1 * S_2) * S_3)(u) &= (S_1 * S_2)(S_3(u)) = S_1(S_2(S_3(u))). \end{aligned}$$

Однако, в общем случае операция композиции некоммутативна, $S_1 * S_2$ не равно $S_2 * S_1$, т.е. изменение порядка включения систем на структурной схеме на рис. 2.2 меняет описание «вход-выход» последовательного соединения¹.

Определение 3. Введём *единичную* систему I формулой

$$I \stackrel{\text{def}}{=} \{(u, u) \mid u \in \Omega\}. \quad (2.3)$$

¹Операция композиции обладает свойством коммутативности на множестве линейных стационарных систем, см. §2 главы 2

Выходной сигнал системы I в точности повторяет, отслеживает, входной сигнал, и по этой причине такую систему можно рассматривать как идеальную модель следящей системы. Единичная система, очевидно, коммутативна с любой другой системой: $I * S = S * I = S \quad \forall S \in \mathcal{O}$.

Определение 4. Конверсную, или симметричную, систему S^c к системе S введём формулой

$$S^c \stackrel{\text{def}}{=} \{(y, u) \mid (u, y) \in S\} \quad (2.4)$$

У конверсной системы S^c по сравнению с исходной системой S вход и выход поменялись местами (рис. 2.3). Очевидно, что для единичной системы $I^c = I$.



Рис. 2.3.

Можно показать, что

$$(S_1 * S_2)^c = S_2^c * S_1^c. \quad (2.5)$$

Рассмотрим две элементарные системы, графически изображённые на рис. 2.4 а)–б) и называемые, соответственно, сумматором и устройством сравнения (вычитателем).



Рис. 2.4.

Их уравнения таковы: а) $y = u_1 + u_2$; б) $y = u_1 - u_2$.

Определение 5. Алгебраической суммой (соответственно, разностью) подмножеств $\Omega_1, \Omega_2 \subseteq \Omega$ называется множество $\Omega_1 + \Omega_2$ ($\Omega_1 - \Omega_2$), определяемое как² [19]

$$\begin{aligned}\Omega_1 + \Omega_2 &\stackrel{\text{def}}{=} \{u_1 + u_2 \mid u_1 \in \Omega_1, u_2 \in \Omega_2\}, \\ \Omega_1 - \Omega_2 &\stackrel{\text{def}}{=} \{u_1 - u_2 \mid u_1 \in \Omega_1, u_2 \in \Omega_2\}.\end{aligned}$$

Параллельному соединению систем $S_1, S_2 \in \mathcal{O}$ на рис. 2.5 а) соответствует их сумма $S_1 + S_2$, задаваемая формулой

$$(S_1 + S_2)(u) = S_1(u) + S_2(u). \quad (2.6)$$

Соединение систем, показанное на рис. 2.5 б), описывается разностью $S_1 - S_2$ систем $S_1, S_2 \in \mathcal{O}$ в соответствии с формулой

$$(S_1 - S_2)(u) = S_1(u) - S_2(u). \quad (2.7)$$

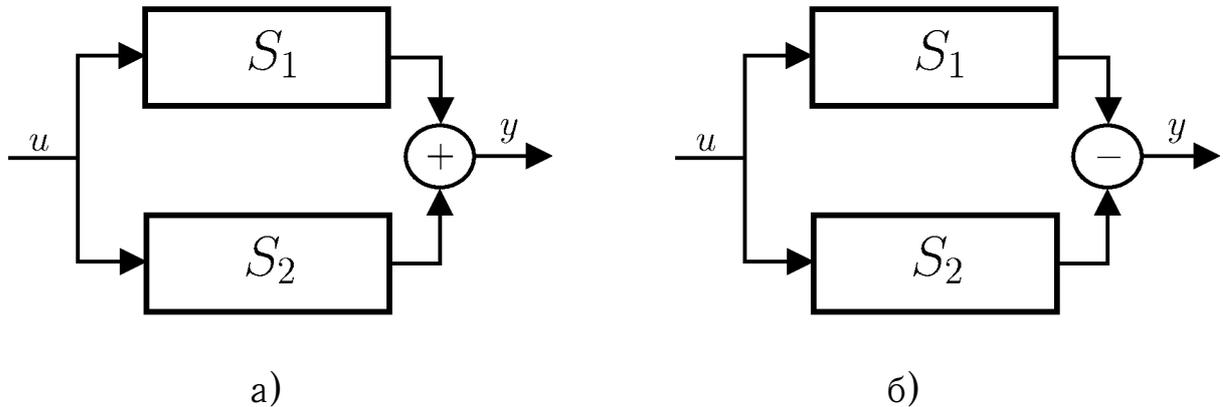


Рис. 2.5.

Операция суммирования обладает свойством ассоциативности

$$S_1 + (S_2 + S_3) = (S_1 + S_2) + S_3, \quad (2.8)$$

которое легко доказывается по аналогии с доказательством свойства ассоциативности для операции композиции (2.2), а также свойством коммутативности

$$S_1 + S_2 = S_2 + S_1, \quad (2.9)$$

²Напомним, что для рассматриваемого случая Ω — множество функций над \mathbb{R} , то есть операции сложения и вычитания определены.

доказательство которого не представляет затруднений.

Операция суммирования дистрибутивна слева относительно операции композиции

$$(S_1 + S_2) * S_3 = (S_1 * S_3) + (S_2 * S_3), \quad (2.10)$$

что вытекает из

$$\begin{aligned} ((S_1 + S_2) * S_3)(u) &= (S_1 + S_2)(S_3(u)) = \\ &= S_1(S_3(u)) + S_2(S_3(u)) = (S_1 * S_3)(u) + (S_2 * S_3)(u). \end{aligned}$$

Левое дистрибутивное свойство операции суммирования проиллюстрировано на рис. 2.6. Абстрактные системы на рис. 2.6 а) и рис. 2.6 б) имеют одинаковые описания «вход-выход».

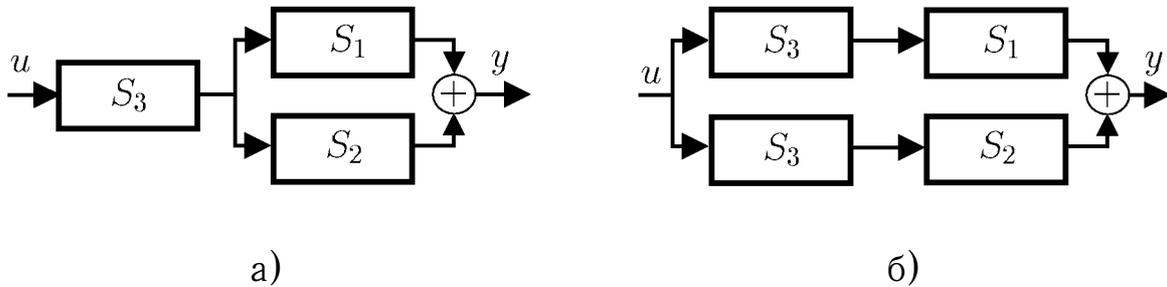


Рис. 2.6.

Дистрибутивность справа для операции суммирования относительно операции композиции в общем случае не имеет места, т.е., вообще говоря³

$$S_1 * (S_2 + S_3) \neq S_1 * S_2 + S_1 * S_3. \quad (2.11)$$

Используя введённые операции, получим математическое описание для системы с отрицательной обратной связью, структурная схема которой приведена на рис. 2.7.

Согласно схеме, содержащей устройство сравнения, для сигнала e , называемого в теории управления сигналом ошибки (рассогласования), имеем

$$e = u - y. \quad (2.12)$$

Отсюда $u = e + y$. Учитывая, что $e = I(e)$ и $y = S(e)$ и используя формулу (2.6) для суммы систем получаем, что

$$u = I(e) + S(e) = (I + S)(e), \quad (2.13)$$

³Дистрибутивность (слева и справа) операции суммирования относительно операции композиции выполняется на множестве линейных стационарных систем (см. §4).

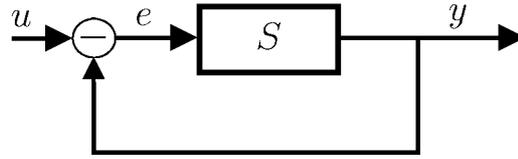


Рис. 2.7.

откуда, пользуясь операцией конверсии, находим

$$e = (I + S)^c(u). \quad (2.14)$$

Таким образом, система с обратной связью на рис. 2.7 описывается бинарным отношением S_e :

$$S_e = (I + S)^c, \quad (2.15)$$

ставящим в соответствие входному сигналу u сигнал ошибки e . С учётом связи $y = S(e)$ мы можем найти описание рассматриваемой системы с обратной связью S_y относительно выходного сигнала y , задаваемое через операцию композиции

$$S_y = S * (I + S)^c. \quad (2.16)$$

Рассмотрим важный класс абстрактных временных систем, называемых стационарными, или инвариантными во времени.

Пусть множество моментов времени $T = \mathbb{R}$. Бинарное отношение сдвига $\nabla_\sigma, \sigma \in \mathbb{R}$ определим правилом

$$\nabla_\sigma(u(t)) = u(t + \sigma). \quad (2.17)$$

Если $\sigma = 0$, то уравнение (2.17) задаёт единичную систему I . Если $\sigma < 0$, то уравнение (2.17) задаёт систему, у которой выходной сигнал повторяет входной с запаздыванием σ , т.е. звено запаздывания. Если $\sigma > 0$, то это уравнение задаёт систему, у которой выходной сигнал опережает входной, т.е. экстраполятор.

Из данного выше определения вытекают следующие свойства отношений сдвига $\nabla_{\sigma_1}, \nabla_{\sigma_2}, \sigma_1, \sigma_2 \in \mathbb{R}$:

$$\nabla_{\sigma_1} * \nabla_{\sigma_2} = \nabla_{\sigma_2} * \nabla_{\sigma_1} = \nabla_{\sigma_1 + \sigma_2}, \quad (2.18)$$

откуда следует, что

$$\nabla_{\sigma} * \nabla_{-\sigma} = I \quad (2.19)$$

и, следовательно, $\nabla_{\sigma}^c = \nabla_{-\sigma}$. Таким образом, нетрудно убедиться, что множество отношений сдвига образует абелеву группу относительно операции композиции $*$.

С учётом того, что сумма $u_1 + u_2$, $u_1, u_2 \in \Omega$ и произведение ku , $k \in \mathbb{R}$, $u \in \Omega$ определяются формулами

$$\begin{aligned} (u_1 + u_2)(t) &\stackrel{\text{def}}{=} u_1(t) + u_2(t), \\ (ku)(t) &\stackrel{\text{def}}{=} ku(t), \end{aligned} \quad (2.20)$$

из определения (2.17) получаем свойство линейности отношения сдвига ∇_{σ} :

$$\begin{aligned} \nabla_{\sigma}(ku(t)) &= k\nabla_{\sigma}(u(t)), \\ \nabla_{\sigma}(u_1(t) + u_2(t)) &= \nabla_{\sigma}(u_1(t)) + \nabla_{\sigma}(u_2(t)). \end{aligned} \quad (2.21)$$

Аналогично формуле (2.17) вводится понятие отношения сдвига и для случая дискретных систем, т.е. для $T = \mathbb{Z}$ и $\sigma \in \mathbb{Z}$.

Используя понятие отношения сдвига сформулируем следующее определение.

Определение 6. Система $\Sigma = \langle \mathbb{R}, U, Y, \Omega, \Gamma, S \rangle$ называется *стационарной (инвариантной во времени)*, если $\forall \sigma \in \mathbb{R}$

$$\nabla_{\sigma} * S = S * \nabla_{\sigma}. \quad (2.22)$$

Если $\exists \sigma \in \mathbb{R}$, для которого условие коммутативности нарушается, то система Σ называется *нестационарной*.

Очевидно, что единичная система, звено запаздывания и экстраполятор являются стационарными системами.

Покажем, что множество стационарных систем замкнуто относительно всех введённых операций с системами.

С учётом свойства ассоциативности (2.2) докажем стационарность композиции стационарных систем:

$$\begin{aligned} (S_1 * S_2) * \nabla_{\sigma} &= S_1 * (S_2 * \nabla_{\sigma}) = \\ &= S_1 * (\nabla_{\sigma} * S_2) = (S_1 * \nabla_{\sigma}) * S_2 = \nabla_{\sigma} * (S_1 * S_2). \end{aligned} \quad (2.23)$$

Стационарность суммы вытекает из свойств дистрибутивности (2.10) и линейности отношения сдвига (2.21):

$$\begin{aligned} (S_1 + S_2) * \nabla_\sigma &= (S_1 * \nabla_\sigma) + (S_2 * \nabla_\sigma) = \\ &= (\nabla_\sigma * S_1) + (\nabla_\sigma * S_2) = \nabla_\sigma(S_1 + S_2). \end{aligned} \quad (2.24)$$

Из формулы (2.5) получаем

$$\begin{aligned} (\nabla_\sigma * S)^c &= S^c * \nabla_{-\sigma}, \\ (S * \nabla_\sigma)^c &= \nabla_{-\sigma} * S, \end{aligned} \quad (2.25)$$

откуда следует, что конверсная система S^c стационарна, если стационарна система S . При этом из формул (2.15) и (2.16) вытекает, что система с обратной связью на рис. 2.7 стационарна, если стационарна система S . Утверждение доказано.

Доказанное утверждение означает, что любая система, полученная из стационарных подсистем с помощью введённых операций, т.е. с помощью последовательного и всех других видов соединений, рассмотренных здесь, является стационарной системой.

Выше, в §1, определение абстрактной временной системы было сформулировано исходя из рассмотрения многократного эксперимента по моделированию реально существующего объекта. Однако одно позволяет рассматривать как абстрактные временные системы и некие математические модели, заведомо не имеющие прототипов в реальном, физическом мире. Примером такой модели может служить, в частности, экстраполятор ∇_σ , $\sigma > 0$, у которого значения выходного сигнала в каждый момент времени равны будущим, т.е. ещё неизвестным значениям входного сигнала.

Для того, чтобы абстрактная временная система могла в принципе иметь реальный прототип в физическом мире, она должна реализовывать причинно-следственную связь между воздействиями и реакциями. Для этого необходимо, чтобы система обладала свойством неупреждаемости (неантисипативности), которое можно определить следующим образом.

Определение 7. Пусть $T = \mathbb{R}$. Правое отношение сечения P_τ , $\tau \in \mathbb{R}$ определим правилом

$$P_\tau u(t) = u_\tau(t) = u_{(-\infty; \tau]}(t). \quad (2.26)$$

Правое отношение сечения P_τ сопоставляет функции u , определённой на $(-\infty; +\infty)$ её ограничение (сужение), определённое на $(-\infty; \tau]$ (рис. 2.8 а)).

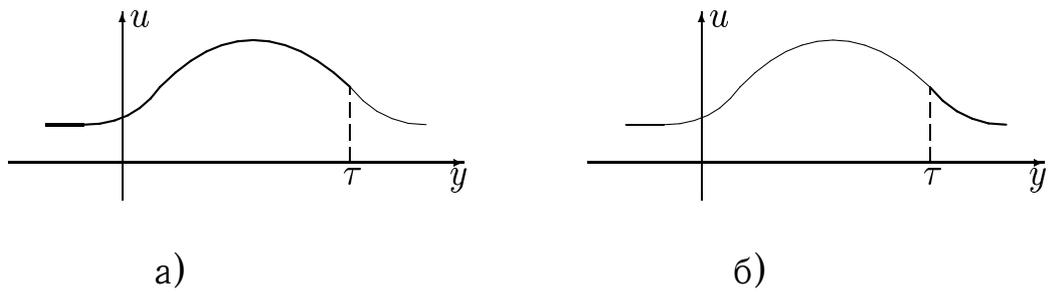


Рис. 2.8.

Определение 8. Абстрактную систему $\Sigma = \langle \mathbb{R}, U, Y, \Omega, \Gamma, S \rangle$ будем называть *неупреждающей (неантисипативной)*, если $\forall \tau \in \mathbb{R}$

$$P_\tau * S = P_\tau * S * P_\tau. \quad (2.27)$$

У неупреждающей системы значение выходного сигнала в каждый момент времени определяется только прошлыми и текущим значением входного сигнала.

Определение 9. *Левое отношение сечения* P^τ задаётся правилом

$$P^\tau u(t) = u^\tau(t) = u_{[\tau; +\infty)}(t). \quad (2.28)$$

Левое отношение сечения P^τ сопоставляет функции u , определённой на $(-\infty; +\infty)$ её ограничение (сужение) u^τ , определённое на $[\tau; +\infty)$ (рис. 2.8 б).

Определение 10. Абстрактную систему $\Sigma = \langle \mathbb{R}, U, Y, \Omega, \Gamma, S \rangle$ будем называть *упреждающей (антисипативной)*, если

$$P^\tau * S = P^\tau * S * P^\tau. \quad (2.29)$$

Звено запаздывания является неупреждающей системой, а экстраполятор — упреждающей. Единичная система I одновременно упреждающая и неупреждающая система в смысле данных нами определений. Системы такого типа будем называть *безынерционными*, или *статическими*. Внешние описания безынерционных систем задаются функциями $f : T \times U \rightarrow Y$ и, в стационарном случае, функциями $f : U \rightarrow Y$.

Покажем, что множество неупреждающих систем (аналогично, упреждающих систем) замкнуто относительно всех введённых операций, кроме операции конверсии.

В самом деле, для операции композиции имеем

$$\begin{aligned} P_\tau * (S_1 * S_2) &= (P_\tau * S_1 * P_\tau) * S_2 = \\ &= P_\tau * S_1 * P_\tau * S_2 * P_\tau = P_\tau * (S_1 * S_2) * P_\tau, \end{aligned} \quad (2.30)$$

т.е. композиция неупреждающих систем — неупреждающая система.

Из определения (2.26) вытекает свойство линейности правого отношения сечения:

$$\begin{aligned} P_\tau(ku(t)) &= ku(t), \quad t \leq \tau, \\ P_\tau(u_1(t) + u_2(t)) &= u_1(t) + u_2(t), \quad t \leq \tau. \end{aligned} \quad (2.31)$$

Из (2.31) и (2.10) получаем для операции суммирования

$$\begin{aligned} P_\tau * (S_1 + S_2) &= P_\tau * S_1 + P_\tau * S_2 = \\ &= P_\tau * S_1 * P_\tau + P_\tau * S_2 * P_\tau = P_\tau * (S_1 + S_2) * P_\tau, \end{aligned} \quad (2.32)$$

т.е. сумма неупреждающих систем — неупреждающая система.

Однако, система, конверсная к неупреждающей, может быть упреждающей, и наоборот. Рассмотрим, например, конверсные системы ∇_σ и $\nabla_{-\sigma}$, $\sigma > 0$. Звено запаздывания $\nabla_{-\sigma}$ — неупреждающая система, а экстраполятор ∇_σ — упреждающая система.

Утверждение доказано. Из этого утверждения следует, что система с обратной связью на рис. 2.7 может оказаться упреждающей системой, даже если система S неупреждающая. В самом деле, положим

$$S = \nabla_{-\sigma} - I. \quad (2.33)$$

Тогда из (2.15) следует, что $S_e = \nabla_\sigma$. Это — экстраполятор, т.е. упреждающая система. Отметим, что и единичная система I и звено запаздывания $\nabla_{-\sigma}$, $\sigma > 0$ могут быть реализованы физически только приближённо и потому при реальном физическом моделировании рассматриваемая система с обратной связью будет являться только некоторой моделью экстраполятора во множестве неупреждающих систем.

Отметим также, что отсюда следует, что множество безынерционных систем замкнуто относительно всех введённых операций, т.е. любое соединение безынерционных систем будет являться снова безынерционной системой.

§3. Задача управления и принципы управления системами

Задача управления некоторым объектом может быть определена как задача нахождения управляющего воздействия на объект — входного сигнала, обеспечивающего заданное внешнее поведение этого объекта, т.е. заданный выходной сигнал объекта. Иначе говоря, входной сигнал является некоторой программой, в соответствии с которой объект ведёт себя желаемым для экспериментатора образом (в данном случае, обеспечивая получение требуемого выходного сигнала).

Системами автоматического управления (САУ) называются технические системы, в которых процесс управления, т.е. выработка требуемого управляющего воздействия, происходит автоматически, без участия человека.

Примерами САУ могут служить:

- а) системы регулирования, целью которых является поддержание заданного режима работы;
- б) следящие системы, целью которых является измерение режима работы в соответствии с изменяющимися значениями внешних воздействий.

Методы решения задач управления и проектирования САУ в их различной формулировке изучаются в теории автоматического управления.

Основным методом проектирования САУ является метод математического моделирования. Он заключается в замене реальной технической системы её математической моделью — абстрактной временной системой, и, в дальнейшем, проектировании САУ по этой модели.

Математическая модель, задающая абстрактную временную систему в конкретных задачах может задаваться обыкновенным дифференциальным уравнением, дифференциальным уравнением в частных производных, интегральным уравнением, разностным уравнением и т.д. Такое разнообразие моделей требует разработки специальных методов проектирования САУ. Совокупность математических моделей, методов и алгоритмов для решения задач автоматизированного проектирования САУ называется математическим обеспечением САУ.

С точки зрения теории автоматического управления в любой САУ можно выделить: а) объект управления, требуемое внешнее поведение которого нужно обеспечить, б) управляющую систему, которая должна фор-

мировать управляющее воздействие (управление), обеспечивающее требуемый выходной сигнал объекта управления.

Рассмотрим для случая систем с одним входом и одним выходом основные принципы управления системами.

В случае управления по разомкнутому циклу (программного управления) управляющая система и объект управления связаны между собой так, как это показано на рис. 3.1, т.е. с помощью последовательного соединения.

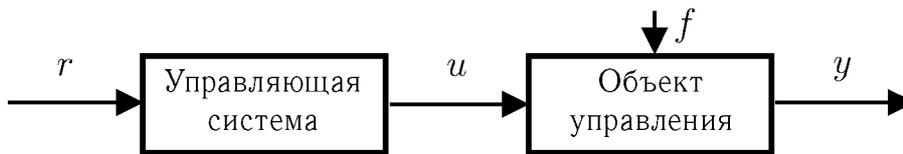


Рис. 3.1.

Здесь r — задающее воздействие, определяющее требуемый режим работы объекта управления; f — возмущающее воздействие — внешнее воздействие на объект управления; u — сигнал управления — входной сигнал объекта управления; y — выходной сигнал системы автоматического управления, характеризующий результат управления (режим работы объекта управления).

При такой схеме управления сигнал управления u не зависит от реальных результатов управления, т.е. от выходного сигнала системы управления y . Это основной недостаток управления по разомкнутому циклу, т.к. в нём невозможно скорректировать непредвиденные возмущения f ; последние могут сделать САУ разомкнутого цикла вообще неработоспособной. В этой связи САУ разомкнутого цикла имеют ограниченную область применения и используются тогда, когда по условиям конкретной задачи возмущающие воздействия либо вообще отсутствуют, либо их конкретный вид известен и их можно скомпенсировать.

От этого недостатка избавлены системы управления с обратной связью (замкнутого цикла), где управление осуществляется с помощью отрицательной обратной связи. Структурная схема системы управления с обратной связью показана на рис. 3.2. Управление в такой системе осуществляется по сигналу ошибки $e = r - y$, который определяется выходным сигналом системы управления y и, как следствие, возмущающим воздействием. При этом конкретный вид возмущающих воздействий не предполагается заранее известным и система управления с обратной связью остаётся

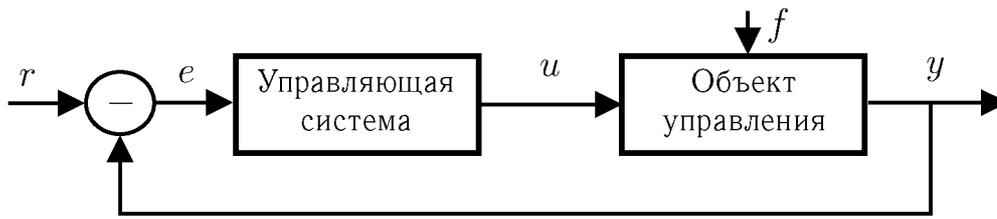


Рис. 3.2.

работоспособной даже при достаточно высоком уровне возмущений. Это свойство позволяет использовать управление по замкнутому циклу в условиях неопределённости относительно действующих возмущений, что характерно для большинства практических ситуаций. По принципу обратной связи были построены регуляторы И.И. Ползунова и Дж. Уатта. Н. Винер отмечал, что по принципу обратной связи организованы процессы управления в живых организмах, и считал его основным принципом кибернетики [3].

Отметим, что и для систем регулирования, и для следящих систем управляющую систему в схеме с обратной связью принято называть регулятором.

Рассмотрим, как влияют постоянные возмущения $f = const$ на систему управления с обратной связью для задачи регулирования, т.е. когда $r = const$, в предположении, что объект управления задаётся линейным безынерционным звеном

$$y = k_1 u + k_2 f \quad (3.1)$$

и регулятор (управляющая система) тоже является безынерционным линейным:

$$u = k_p e. \quad (3.2)$$

В (3.1) и (3.2) k_1 и k_2 — коэффициенты усиления относительно сигнала управления и возмущения соответственно, k_p — коэффициент усиления регулятора. Учитывая уравнение для сигнала ошибки $e = r - y$ из (3.1) и (3.2) получаем:

$$\begin{aligned} y &= k_1 k_p e + k_2 f, \\ y &= k_1 k_p (r - y) + k_2 f, \\ (1 + k_1 k_p) y &= k_1 k_p r + k_2 f. \end{aligned} \quad (3.3)$$

Отсюда находим следующее уравнение для выходного сигнала системы регулирования:

$$y = \frac{k_1 k_p}{1 + k_1 k_p} r + \frac{k_2}{1 + k_1 k_p} f. \quad (3.4)$$

Предполагая, что $r, f > 0$, отсюда получаем такое выражение для оценки абсолютной величины ошибки воспроизведения задающего воздействия r на выходе системы

$$|r - y| \leq \left| 1 - \frac{k_1 k_p}{1 + k_1 k_p} \right| r + \left| \frac{k_2}{1 + k_1 k_p} \right| f. \quad (3.5)$$

При $k_p \rightarrow +\infty$ коэффициенты при r и f стремятся к нулю. Отсюда следует, что выбирая достаточно большой коэффициент усиления регулятора можно получить сколь угодно малую ошибку воспроизведения задающего воздействия на выходе системы — ошибку регулирования — даже при наличии значительных по величине возмущений.

Отметим, однако, что даже если возмущения отсутствуют, при конечных значениях коэффициента усиления регулятора ошибка регулирования (статическая ошибка) не может быть равной нулю. Это недостаток данного типа регулятора, называемого пропорциональным. Существуют более сложные типы регуляторов, свободные от этого недостатка. Интегральный регулятор вырабатывает управляющий сигнал как функцию от интеграла ошибки:

$$u(t) = k_p \int_0^t e(\xi) d\xi, \quad (3.6)$$

пропорционально-интегральный, или изодромный, вырабатывает управляющий сигнал как функцию от ошибки и интеграла ошибки:

$$u(t) = k_{p1} e(t) + k_{p2} \int_0^t e(\xi) d\xi \quad (3.7)$$

пропорционально-интегрально-дифференциальный вырабатывает управляющий сигнал с использованием производной ошибки:

$$u(t) = k_{p1} e(t) + k_{p2} \int_0^t e(\xi) d\xi + k_{p3} \frac{de}{dt}. \quad (3.8)$$

Общее выражение для многих видов регуляторов может быть записано как

$$u(t) = F \left(e(t), \int_0^t e(\xi) d\xi, \dots, \frac{de}{dt}, \dots \right), \quad (3.9)$$

где F в общем случае нелинейная (как правило, неубывающая) функция своих аргументов.

Нелинейные регуляторы, очевидно, обладают большими возможностями, чем линейные, однако на практике чаще используют линейные регуляторы, которые более просто реализуются технически.

Отметим, что регулятор в схеме с обратной связью является, фактически, чувствительным элементом, который реагирует на отклонения выходного сигнала от заданного сигнала — задающего воздействия. Сигнал управления при этом вырабатывается за счёт энергии сигнала ошибки. Регуляторы такого типа называются регуляторами прямого действия. Однако энергии такого сигнала управления может и не быть достаточным для управления мощными объектами управления. В этом случае в состав регулятора включают усилительно-преобразовательное и исполнительные устройства так, как показано на рис. 3.3. Регуляторы такого типа называ-

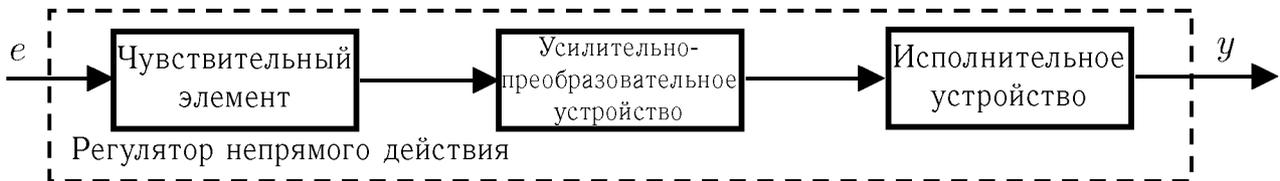


Рис. 3.3.

ются регуляторами непрямого действия, и наиболее часто используются на практике.

Цепь обратной связи может содержать измерительное устройство в соответствии со схемой на рис. 3.4, которое используется в тех случаях, когда выходной сигнал u непосредственно не измеряется и для повышения качества системы управления.

Проблема проектирования САУ весьма сложна и комплексна и её решение включает в себя как общие этапы, характерные для проектирования вообще (например, подготовку технической документации), так и

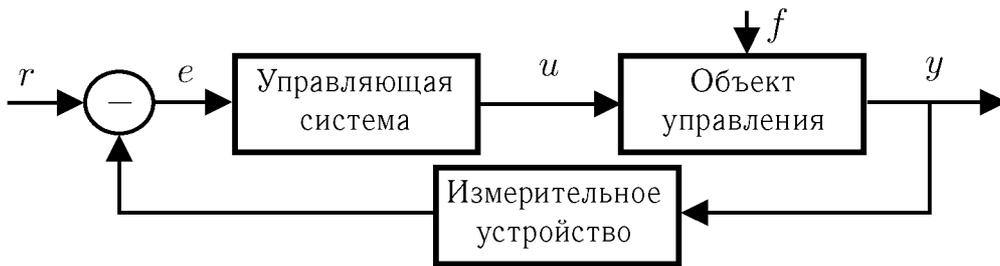


Рис. 3.4.

специфические. Из них центральным является этап анализа и синтеза математической модели САУ.

На этапе анализа и синтеза производится анализ и упрощение математической модели объекта управления, выбирается алгоритм управления и алгоритм синтеза, осуществляется синтез регулятора и анализируются свойства синтезированной САУ. Если анализ показывает неудовлетворительность полученного решения задачи, вся процедура повторяется.

Существующие методы синтеза САУ можно разделить на две большие группы. Во-первых, задачу синтеза можно рассматривать как задачу минимизации некоторого функционала оптимальности — задачу оптимального синтеза САУ. Однако, обычно в технических задачах требуется добиться не оптимальности системы в математическом смысле, а оптимальности с точки зрения её технических свойств (например, основными требованиями к системам управления летательного аппарата и судна являются требования безопасности полёта и навигации [22]). Кроме того, возможно, что оптимальное с математической точки зрения решение задачи будет слишком сложным, дорогим или вообще нереализуемым технически.

Поэтому весьма часто вместо оптимального синтеза используются инженерные методы синтеза по желаемым показателям точности и качества, которые задаются во временной или частотной областях. Эти показатели являются числовыми характеристиками реакции системы на определённые типы внешних воздействий. Как они сами, так и их оптимальные значения отобраны на основе длительного опыта инженеров-проектировщиков и адекватно с инженерной точки зрения описывают поведение САУ. При таком подходе система автоматического управления синтезируется так, чтобы её показатели точности и качества не превосходили бы заданных. Именно этот классический подход к синтезу САУ и рассматривается ниже во второй главе. Вопросы же оптимального синтеза остаются за рамками данной книги.

§4. Линейные и нелинейные системы

Пусть задана абстрактная временная система $\Sigma = \langle \mathbb{R}, U, Y, \Omega, \Gamma, S \rangle$. Пусть S является функцией (оператором), далее обозначаемой $f : \Omega \rightarrow \Gamma$ (так же будем обозначать и всю систему Σ). Пусть Ω и Γ — линейные нормированные пространства над \mathbb{R} . Конкретный вид этих пространств определяется исходя из условий задачи.

Определение 1. Система f называется *линейной*, если f — линейная функция (оператор), то есть $\forall \alpha, \beta \in \mathbb{R} \forall u_1, u_2 \in \Omega$

$$f(\alpha u_1 + \beta u_2) = \alpha f(u_1) + \beta f(u_2). \quad (4.1)$$

В противном случае система называется нелинейной.

Если системы f, g — линейны, то линейной является их сумма и композиция. В случае линейных систем операция суммирования дистрибутивна справа относительно операции композиции:

$$f * (g + h) = f * g + f * h. \quad (4.2)$$

Операция конверсии f^c заменяется на инверсию f^{-1} , если обратный оператор f^{-1} существует. В этом случае система называется обратимой. Для обратимых систем выполняется очевидное равенство

$$f * f^{-1} = f^{-1} * f = I. \quad (4.3)$$

Все вышеприведённые свойства являются свойствами линейных операторов и доказываются в соответствующем разделе курса функционального анализа [12].

Таким образом, любая система, полученная из линейных систем при помощи операций суммирования, умножения на константу, композиции и обращения, также является линейной системой.

Пример 1 (Следящая система). Как уже указывалось, следящая система I сопоставляет сигналу на входе сигнал на выходе. Она является примером тривиальной линейной системы, т.к.

$$I(\alpha u_1 + \beta u_2) = \alpha u_1 + \beta u_2 = \alpha I(u_1) + \beta I(u_2)$$

по определению.

Пример 2 (Линейные системы над \mathbb{R}). Пусть $u(t) : \mathbb{R} \rightarrow \mathbb{R}$, $y(t) : \mathbb{R} \rightarrow \mathbb{R}$. Рассмотрим интегральный оператор следующего вида:

$$y(t) = \int_{\mathbb{R}} K(t, \tau) u(\tau) d\tau. \quad (4.4)$$

Функция $K(t, \tau)$, задающая оператор, называется его ядром. Очевидно, система, задаваемая этим оператором, будет являться линейной системой.

Единичную систему I при соответствующих ограничениях на входной сигнал⁴ можно задать как интеграл от обобщённой функции:

$$I(u(t)) = \int_{\mathbb{R}} \delta(t - \tau) u(\tau) d\tau = u(t).$$

Таким образом, любой линейный оператор над, например, $C(\mathbb{R})$, можно задавать при помощи его интегрального представления, если условиться, что ядро $K(t, \tau)$ можно задавать как смешанную функцию (с регулярной — обычной — и сингулярной — обобщённой — частями).

Для регулярных ядер (без обобщённых функций) заметим, что если в качестве входного сигнала использовать $\delta(t)$ ⁵, то на выходе мы получим

$$\int_{\mathbb{R}} K(t, \tau) \delta(\tau) d\tau = K(t, 0),$$

а если $\delta(t - \sigma)$, то

$$\int_{\mathbb{R}} K(t, \tau) \delta(\tau - \sigma) d\tau = K(t, \sigma), \quad (4.5)$$

то есть реакцию системы на поданный δ -импульс в момент времени σ .

Таким образом, это даёт принципиальную возможность идентификации линейной системы (построения её модели) по результатам эксперимента. Ядро линейной системы $K(t, \tau)$ имеет физический смысл отклика на импульсное воздействие и в теории систем носит название *импульсной переходной функции (характеристики)*.

⁴В общей теории требуется, чтобы входной сигнал $u \in C_0^\infty(\mathbb{R})$, т.е. являлся финитным и бесконечно дифференцируемым. Однако, в случае необходимости, класс гладкости может быть существенно уменьшен, вплоть до непрерывности $u(t)$ (или даже разрывов первого рода), что чрезвычайно важно для инженерных приложений. Подробнее см., например, [4].

⁵В этом случае ограничения на непрерывность необходимо накладывать уже на ядро.

Если интегральный оператор такой, что

$$K(t, \tau) = 0 \quad \forall \tau > t, \quad (4.6)$$

то выходной сигнал y системы в момент времени t зависит только от прошлых значений входного сигнала u , что соответствует моделям физического мира — причинным моделям. Система, удовлетворяющая условиям (4.6), называется *причинной*. В математике при выполнении условий (4.6) линейный оператор называют оператором Вольтёрра.

Рассмотрим линейные системы с одним входом и одним выходом. Построим соответствующие ядра результирующих интегральных операторов для последовательного и параллельного соединения систем и для системы с обратной связью.

1. Параллельное соединение линейных систем. Рассмотрим системы, заданные интегральными ядрами $k(t, \tau)$ и $h(t, \tau)$ (см. рис. 4.1). Построим интегральный оператор, задаваемый ядром $g(t, \tau)$, описывающий

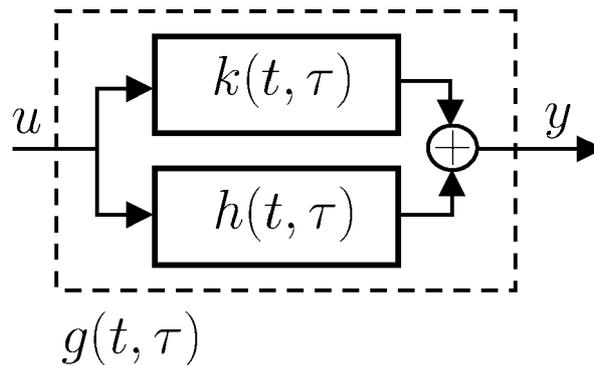


Рис. 4.1.

параллельное соединение систем:

$$y(t) = y_1(t) + y_2(t) = \int_{\mathbb{R}} k(t, \tau)u(\tau) d\tau + \int_{\mathbb{R}} h(t, \tau)u(\tau) d\tau = \int_{\mathbb{R}} g(t, \tau)u(\tau) d\tau.$$

$$g(t, \tau) = k(t, \tau) + h(t, \tau) \quad (4.7)$$

2. Последовательное соединение линейных систем. На рис. 4.2 показана структурная схема последовательного соединения систем, задаваемых операторами $k(t, \tau)$ и $h(t, \tau)$ соответственно. Тогда $g(t, \theta)$ определяется

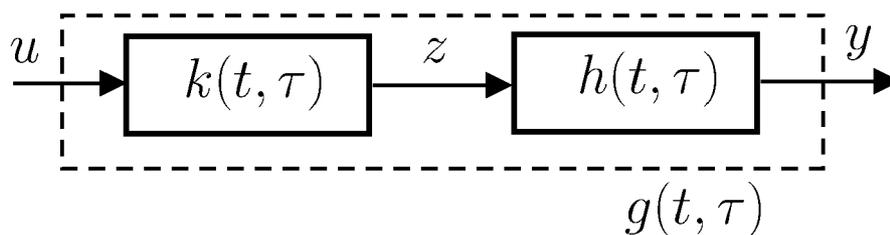


Рис. 4.2.

из

$$y(t) = \int_{\mathbb{R}} h(t, \tau) z(\tau) d\tau = \int_{\mathbb{R}} h(t, \tau) \int_{\mathbb{R}} k(\tau, \theta) u(\theta) d\theta d\tau = \int_{\mathbb{R}} g(t, \theta) u(\theta) d\theta,$$

откуда

$$g(t, \theta) = \int_{\mathbb{R}} h(t, \tau) k(\tau, \theta) d\tau. \quad (4.8)$$

3. Линейная система с обратной связью. Рассмотрим систему с обратной связью с объектом управления, описываемым оператором с ядром $k(t, \tau)$ и регулятором, описываемым оператором с ядром $h(t, \tau)$ (см. рис. 4.3). Регулятор и объект можно рассматривать как последовательное соединение

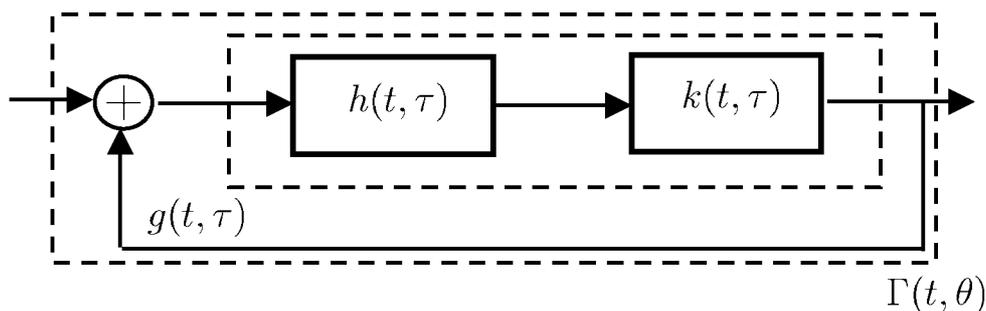


Рис. 4.3.

систем, что мы и сделаем, введя в рассмотрение интегральный оператор с ядром

$$g(t, \theta) = \int_{\mathbb{R}} k(t, \tau) h(\tau, \theta).$$

Сигнал ошибки $e(t) = r(t) - y(t)$, а выходной сигнал выражается через сигнал ошибки как

$$y(t) = \int_{\mathbb{R}} g(t, \theta) e(\theta) d\theta. \quad (4.9)$$

Подставив вместо $e(\theta)$ выражение для него, получим

$$y(t) = \int_{\mathbb{R}} g(t, \theta) r(\theta) d\theta - \int_{\mathbb{R}} g(t, \theta) y(\theta) d\theta. \quad (4.10)$$

Заметим, что в (4.10) функция

$$w(t) = \int_{\mathbb{R}} g(t, \theta) r(\theta) d\theta$$

известна. Тогда полученное уравнение

$$y(t) = w(t) - \int_{\mathbb{R}} g(t, \theta) y(\theta) d\theta \quad (4.11)$$

есть интегральное уравнение Вольтерра второго рода [12] относительно $y(t)$

$$y = Ty. \quad (4.12)$$

По теореме о сжимающем отображении решение — неподвижная точка уравнения (4.12) — существует и единственно, если $\alpha = \|T\| < 1$, где $\|T\|$ можно записать как

$$\|T\| = \sup_t \int_{\mathbb{R}} |g(t, \theta)| d\theta.$$

Решение можно находить при помощи метода последовательных приближений, понимая сходимость в смысле нормы в $CB(\mathbb{R})$ ⁶:

$$y_{i+1} = Ty_i, \quad i \in \mathbb{N}.$$

⁶Пространство непрерывных ограниченных функций в \mathbb{R} .

Выберем в качестве начального приближения y_1 выходной сигнал при размыкании петли обратной связи

$$y_1(t) = w(t) = \int_{\mathbb{R}} g(t, \theta) r(\theta) d\theta.$$

Далее,

$$\begin{aligned} y_2(t) &= \int_{\mathbb{R}} g(t, \theta) r(\theta) d\theta - \int_{\mathbb{R}} g(t, \theta) \int_{\mathbb{R}} g(\theta, \tau) r(\tau) d\tau d\theta, \\ y_2(t) &= \int_{\mathbb{R}} g(t, \theta) r(\theta) d\theta - \int_{\mathbb{R}} g_1(t, \tau) r(\tau) d\tau, \\ g_1(t, \tau) &\stackrel{\text{def}}{=} \int_{\mathbb{R}} g(t, \theta) g(\theta, \tau) d\theta. \end{aligned}$$

В следующем приближении

$$y_3(t) = \int_{\mathbb{R}} g(t, \theta) r(\theta) d\theta - \int_{\mathbb{R}} g_1(t, \tau) r(\tau) d\tau + \int_{\mathbb{R}} g_2(t, \lambda) r(\lambda) d\lambda,$$

где

$$g_2(t, \lambda) \stackrel{\text{def}}{=} \int_{\mathbb{R}} g(t, \theta) g_1(\theta, \lambda) d\theta.$$

Введём

$$g_{i+1}(t, \sigma) = (-1)^{i+1} \int_{\mathbb{R}} g(t, \theta) g_i(\theta, \sigma) d\theta. \quad (4.13)$$

$$g_0(t, \sigma) = g(t, \sigma) \quad (4.14)$$

Таким образом, решение записывается в виде

$$y(t) = \int_{\mathbb{R}} \Gamma(t, \theta) u(\theta) d\theta, \quad (4.15)$$

где $\Gamma(t, \theta) = \sum_{i=0}^{+\infty} g_i(t, \theta)$ — резольвентное ядро (резольвента), а ряд $\sum_{i=0}^{+\infty} g_i(t, \theta)$ — ряд Лиувилля-Неймана. При сделанных предположениях этот ряд всегда

сходится и даёт решение задачи с обратной связью. Однако, он сходится весьма медленно, что неудобно на практике, и, поэтому, в прикладных задачах используются другие методы расчёта систем с обратной связью.

Таким образом, показано, что множество линейных систем замкнуто относительно операций суперпозиции (последовательного соединения), суммирования (параллельного соединения) и инверсии (соединения с обратной связью).

Рассмотрим вопросы, связанные с нелинейными системами. Введём

$$\Omega^i = \underbrace{\Omega \times \Omega \dots \times \Omega}_{i \text{ раз}}$$

и рассмотрим оператор $f_i : \Omega^i \rightarrow \Omega$.

Определение 2. Назовём f_i *i -линейным оператором*, если он линеен по каждому из аргументов, т.е.

$$\begin{aligned} f_i(u_1, \dots, u_{k-1}, \alpha u_k + \beta v_k, u_{k+1}, \dots, u_i) = \\ = \alpha f_i(u_1, \dots, u_{k-1}, u_k, u_{k+1}, \dots, u_i) + \\ + \beta f_i(u_1, \dots, u_{k-1}, v_k, u_{k+1}, \dots, u_i). \end{aligned}$$

Определение 3. Назовём оператор $f_i : \Omega \rightarrow \Omega$ *однородным оператором степени i* , если

$$f_i(u) = f_i(\underbrace{u, \dots, u}_{i \text{ раз}}).$$

Такой оператор, очевидно, удовлетворяет условию

$$f_i(\alpha u) = \alpha^i f_i(u) \tag{4.16}$$

Определение 4. Оператор, задаваемый соотношением

$$y(t) = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} k_i(t, \tau_1, \dots, \tau_i) u_1(\tau_1) \dots u_i(\tau_i) d\tau_1 \dots d\tau_i,$$

назовём *i -линейным интегральным оператором*.

Для i -линейных интегральных операторов здесь и далее в этом параграфе будем предполагать выполнение условий теоремы Фубини, то есть

возможность как брать повторные интегралы, так и интегрировать в целом по пространству \mathbb{R}^i .

Аналогично условию Вольтерра для линейных интегральных операторов, запишем условие Вольтерра на ядро i -линейного интегрального оператора:

$$k_i(t, \tau_1, \dots, \tau_i) = 0, \text{ если } \max_j \{\tau_j\} > t \quad (4.17)$$

(т.е. хотя бы одно из τ_j больше t).

Определение 5. Оператор, задаваемый как

$$y(t) = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} k_i(t, \tau_1, \dots, \tau_i) u(\tau_1) \dots u(\tau_i) d\tau_1 \dots d\tau_i,$$

назовём *однородным оператором Вольтерра степени i* , если для ядра k_i выполнено условие Вольтерра (4.17).

Используя понятие однородного оператора степени i можно рассмотреть так называемые *полиномиальные системы*, задаваемые полиномами Вольтерра.

Определение 6. *Полиномом Вольтерра степени N* называется выражение вида

$$p_N(u) = \sum_{i=1}^N f_i(u), \quad f_N(u) \neq 0, \quad (4.18)$$

где $\{f_i\}_{i=1, \dots, N}$ — однородные операторы Вольтерра степени i .

В развёрнутом виде полином Вольтерра можно записать как

$$y(t) = \sum_{i=1}^N \int_{\mathbb{R}} \dots \int_{\mathbb{R}} k_i(t, \tau_1, \dots, \tau_i) u(\tau_1) \dots u(\tau_i) d\tau_1 \dots d\tau_i,$$

или, введя функцию $u_i(\tau) : \mathbb{R}^i \rightarrow \mathbb{R}$ как

$$u_i(\tau) \stackrel{\text{def}}{=} u(\tau_1) \cdot \dots \cdot u(\tau_i),$$

более компактно

$$y(t) = \sum_{i=1}^N \int_{\mathbb{R}^i} k_i(t, \tau) u_i(\tau) d\tau. \quad (4.19)$$

Можно показать, что множество полиномиальных систем, как и линейных, замкнуто относительно операций суперпозиции (последовательного соединения) и суммирования (параллельного соединения).

Что касается замыкания полиномиальной системы обратной связью, то его невозможно описать полиномом Вольтерра, и множество полиномиальных систем не является замкнутым относительно операции инверсии. Последняя, однако, может быть описана при помощи замыкания множества систем, описываемых полиномами Вольтерра — системой, представляемой рядом Вольтерра.

Таким образом, весьма широкий класс нелинейных систем можно тем или иным образом представить с помощью полиномов Вольтерра.

§5. Линеаризация статических нелинейностей

В практических задачах обычно объект управления является нелинейным. При этом часто встречается случай, когда в объекте можно выделить линейную часть, задаваемую линейным интегральным оператором, и безынерционную нелинейную часть. При этом возможны варианты, когда нелинейная часть f включена перед линейной системой с ядром $k(t, \tau)$ (рис. 5.1,а) и после неё (рис. 5.1,б). Нелинейности обычно измеряются

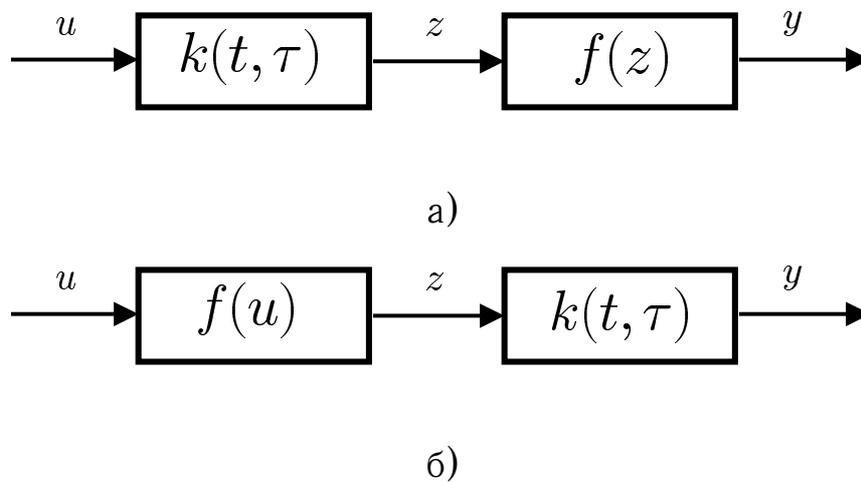


Рис. 5.1.

экспериментально в статическом (установившемся) режиме и называются *статическими нелинейными характеристиками*, задаваемыми функциями действительного переменного. Для них используют кусочно-линейные и полиномиальные аппроксимации.

На рис. 5.2 показаны два наиболее часто встречающихся типа нелинейностей, называемые насыщением (рис. 5.2,а) и зоной нечувствительности (рис. 5.2,б), и их кусочно-линейные аппроксимации.

Кусочно-линейные аппроксимации неудобны тем, что не дают единого аналитического представления для нелинейности на всём рассматриваемом интервале аппроксимации. От этого недостатка свободны полиномиальные аппроксимации нелинейностей. При этом на практике обычно используется аппроксимация полиномом первой степени, т.е. осуществляется линеаризация задачи. Это даёт возможность использовать при проектировании систем с обратной связью хорошо разработанный и надёжный аппарат линейной теории управления.

Традиционный метод линеаризации, который можно назвать методом касательных, заключается в использовании разложения функции в ряд

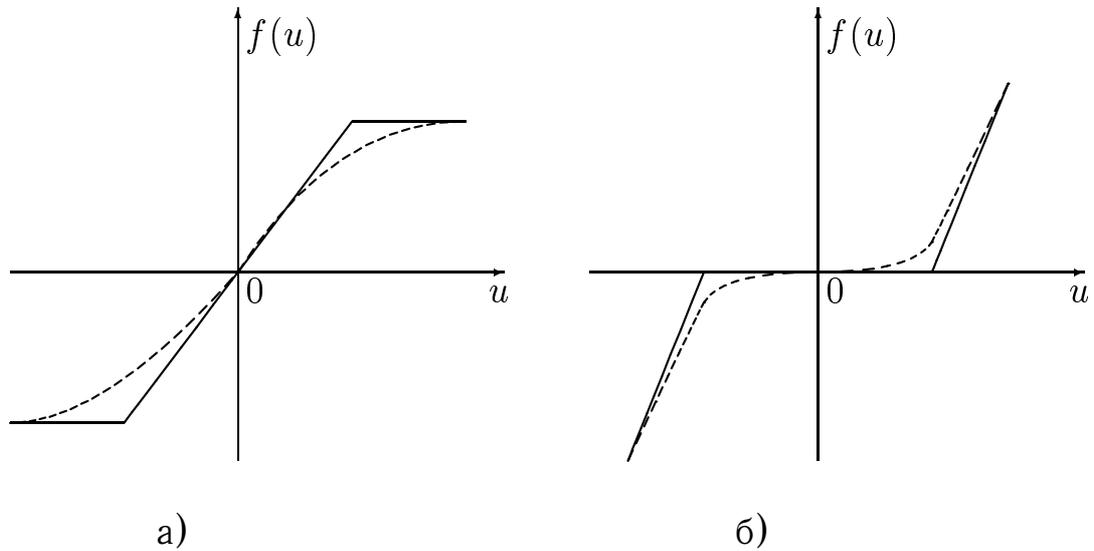


Рис. 5.2.

Тейлора и удержании в этом разложении только линейной части. Этот метод применим только для гладких нелинейностей и даёт хорошие по точности результаты лишь в малой окрестности точки разложения.

Пусть нелинейная функция действительного переменного $y = f(u)$ непрерывна и имеет непрерывные производные в точке $u = 0$. Тогда в окрестности нуля функцию $f(u)$ можно разложить в ряд Маклорена:

$$f(u) = f(0) + \frac{u}{1!}f'(0) + \frac{u^2}{2!}f''(0) + \dots \quad (5.1)$$

Если $f(0) = 0$, то отбрасывание в разложении (5.1) слагаемых степени выше первой даёт следующее представление для нелинейной характеристики в окрестности $u = 0$:

$$f(u) \approx k_1 u, \quad (5.2)$$

где $k_1 = f'(0)$ — коэффициент линеаризации.

Рассмотрим вопрос об оценке погрешности аппроксимации при использовании формул (5.1) и (5.2). В предположении, что $f(u)$ имеет непрерывную $n + 1$ -ую производную, обозначим

$$M = \max_{u \in [-1; 1]} |f^{(n+1)}(u)|.$$

Тогда для отрезка $[-1; 1]$ погрешность аппроксимации функции $f(u)$ поли-

номом Тейлора $p_n(u)$ степени n оценивается как [5]

$$\Delta = \max_{u \in [-1;1]} |f(u) - p_n(u)| \leq \frac{M}{(n+1)!2^n}. \quad (5.3)$$

Для $n = 1$ отсюда получаем

$$\Delta \leq \frac{M}{4}. \quad (5.4)$$

Значительно лучшие по точности результаты аппроксимации можно получить, если вместо ряда Тейлора использовать теорию равномерных (минимаксных) аппроксимаций нелинейных функций.

Пусть нелинейная статическая характеристика $y = f(u)$ является непрерывной, не обязательно гладкой (дифференцируемой), функцией.

Согласно теореме Вейерштрасса непрерывная на отрезке $[a; b]$ функция $f(u)$ может быть аппроксимирована на нём полиномом с наперёд заданной точностью [12]. Теорема Вейерштрасса, однако, не даёт способа нахождения аппроксимируемого полинома. Для этого может быть использован подход, разработанный П.Л. Чебышевым [5], заключающийся в построении полиномов наилучшего равномерного (минимаксного) приближения.

Пусть $p_n(u)$ — некоторый полином с действительными коэффициентами степени n ($\deg p_n = n$). Введём оценку уклонения этого полинома от функции $f(u)$ на $[a; b]$ по формуле

$$\Delta = \max_{u \in [a;b]} |f(u) - p_n(u)|. \quad (5.5)$$

Если это уклонение минимально среди всех полиномов степени n , то полином $p_n(u)$ называется *полиномом наилучшего равномерного приближения*.

Чебышевым было доказано, что если функция $f(u)$ непрерывна на $[a; b]$, то для неё существует единственный полином наилучшего равномерного приближения $p_n(u)$.

Пусть $u : T \rightarrow [a; b]$ — непрерывный входной сигнал на безынерционную нелинейность $y = f(u)$. Тогда, обозначив $y(t) = f(u(t))$, $\hat{y}(t) = p_n(u(t))$, из (5.5) получаем

$$\begin{aligned} \max_{t \in T} |y(t) - \hat{y}(t)| &= \max_{t \in T} |f(u(t)) - p_n(u(t))| = \\ &= \max_{u \in [a;b]} |f(u) - p_n(u)| = \Delta. \end{aligned} \quad (5.6)$$

Это означает, что среди всех возможных полиномиальных аппроксимаций степени n для нелинейности $f(u)$ полином $p_n(u)$ обеспечивает наилучшее равномерное приближение и для выходного сигнала нелинейности. Это обстоятельство позволяет применять полиномы наилучшего равномерного приближения в задачах теории управления и теории систем.

Условия, при которых некоторый полином $p_n(u)$ является полиномом наилучшего равномерного приближения, даёт следующая теорема.

Теорема 1 (Чебышева). *Для того, чтобы полином $p_n(u)$, $\deg p_n = n$ являлся полиномом наилучшего равномерного приближения для функции $f(u)$ на отрезке $[a; b]$, необходимо и достаточно, чтобы по крайней мере в $n + 2$ точках отрезка $[a; b]$ разность $f(u) - p_n(u)$ принимала поочерёдно значения $\pm\Delta$, где Δ определяется по формуле (5.5).*

Рассмотрим применение этой теоремы для построения полинома наилучшего равномерного приближения степени $n = 1$ (линеаризация нелинейности) для обобщённой нелинейности типа насыщения

$$f(u) = \begin{cases} -h, & u < -l, \\ \frac{h}{l}u, & |u| \leq l, \\ h, & u > l, \end{cases} \quad (5.7)$$

график которой показан на рис. 5.3. Полином $p_1(u) = k_1u$ будет для неё

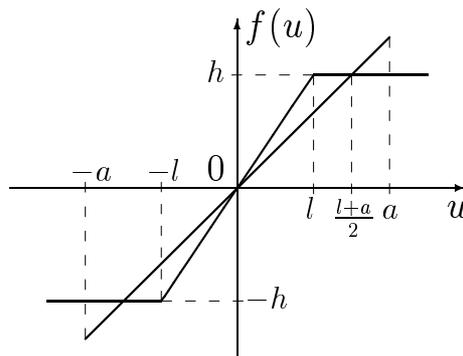


Рис. 5.3.

полиномом наилучшего равномерного приближения на отрезке $[-a; a]$, если разность $f(u) - k_1u$ принимает равные по модулю, но противоположные по знаку значения в точках $-a, -l, a, l$. Из рис. 5.3 видно, что графиком полинома наилучшего приближения является прямая, проходящая через точки с

координатами $(0; 0)$ и $(\frac{a+l}{2}; h)$. Записывая уравнение прямой, проходящей через две заданные точки, получаем такое уравнение для коэффициента линеаризации в зависимости от величины интервала аппроксимации, определяемого величиной a :

$$k_1 = \frac{2h}{a+l}, \quad a > l. \quad (5.8)$$

Для обобщённой нелинейности типа зоны нечувствительности

$$f(u) = \begin{cases} \operatorname{tg} \alpha (u+l), & u < -l, \\ 0, & |u| \leq l, \\ \operatorname{tg} \alpha (u-l), & u > l, \end{cases} \quad (5.9)$$

используя аналогичные геометрические построения, показанные на рис. 5.4, получаем следующее уравнение для коэффициента линеаризации в зави-

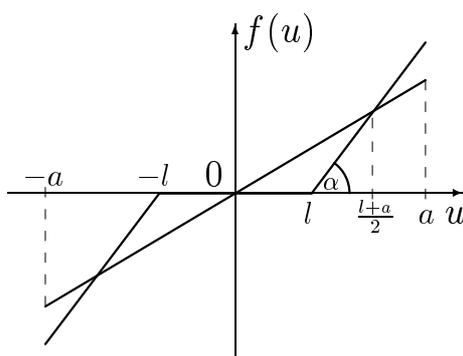


Рис. 5.4.

симости от величины интервала аппроксимации, определяемого величиной a :

$$k_1 = \frac{a-l}{a+l} \operatorname{tg} \alpha, \quad a > l. \quad (5.10)$$

Данный метод отыскания уравнений для коэффициентов линеаризации удобно использовать для кусочно-линейных систем. В общем случае, а также для нелинейностей, заданных графически, можно использовать альтернативный подход, основанный на замене полинома наилучшего равномерного приближения близким к нему полиномом [5].

В качестве полинома, близкого к полиному наилучшего равномерного приближения, можно выбрать интерполяционный полином Лагранжа с

узлами Чебышева. Этот полином удобно представить через ортогональные полиномы Чебышева, задаваемые формулой

$$T_n(u) = \cos(n \arccos u), \quad |u| \leq 1. \quad (5.11)$$

Рекуррентная формула, дающая удобный способ нахождения полиномов Чебышева для произвольного $n \in \mathbb{N}$, имеет вид

$$T_{n+1}(u) = 2uT_n(u) - T_{n-1}(u). \quad (5.12)$$

Первые четыре полинома Чебышева таковы:

$$\begin{aligned} T_0(u) &= 1, \\ T_1(u) &= u, \\ T_2(u) &= 2u^2 - 1, \\ T_3(u) &= 4u^3 - 3u. \end{aligned} \quad (5.13)$$

Полиномы Чебышева ортогональны на $[-1; 1]$ с весом $\frac{1}{\sqrt{1-u^2}}$, т.е.

$$\int_{-1}^1 \frac{T_n(u)T_m(u)}{\sqrt{1-u^2}} du = \begin{cases} 0, & m \neq n, \\ \frac{\pi}{2}, & m = n \neq 0, \\ \pi, & m = n = 0. \end{cases} \quad (5.14)$$

Используя полиномы Чебышева, интерполяционный полином Лагранжа для функции $f(u)$, заданной своими значениями $f(u_0), f(u_1), \dots, f(u_n)$ в $n+1$ точках (узлах интерполирования) u_0, u_1, \dots, u_n отрезка $[-1; 1]$ можно записать как [13]

$$p_n(u) = \frac{1}{2}A_0 + \sum_{k=1}^n A_k T_k(u). \quad (5.15)$$

Здесь коэффициенты интерполяционного полинома задаются формулами для $k = \overline{0, n}$:

$$A_k = \frac{2}{n+1} \sum_{j=0}^n f(u_j) \cos \frac{(2j+1)k\pi}{2(n+1)}. \quad (5.16)$$

Узлы Чебышева $u_j, j = \overline{0, n}$ задаются как

$$u_j = \cos \frac{(2j+1)\pi}{2(n+1)} \quad (5.17)$$

В предположении, что функция $f(u)$ имеет непрерывную $n + 1$ -ю производную и используя обозначение

$$M = \max_{u \in [-1;1]} |f^{(n+1)}(u)|,$$

погрешность аппроксимации функции её интерполяционным полиномом (5.16) можно оценить как

$$\Delta = \max_{u \in [-1;1]} |f(u) - p_n(u)| \leq \frac{M}{(n+1)2^{2n}}. \quad (5.18)$$

Оценка (5.18) — оптимальная и улучшить её за счёт другого выбора узлов интерполяции нельзя. Это обстоятельство и позволяет использовать интерполяционный полином (5.16) в качестве близкого к полиному наилучшего равномерного приближения, точных формул для нахождения коэффициентов которого в общем случае не существует.

По сравнению с оценкой (5.3) для полинома Тейлора, эта оценка в $2n$ раз лучше. Для $n = 1$

$$\Delta \leq \frac{M}{8}, \quad (5.19)$$

и эта оценка в два раза лучше, чем (5.4).

Для интервала аппроксимации $[-a; a]$ формула (5.15) заменяется следующей

$$p_n(u) = \frac{1}{2}A_0 + \sum_{k=1}^n A_k T_k\left(\frac{u}{a}\right), \quad (5.20)$$

а узлы интерполирования (5.17) пересчитываются по формулам

$$\bar{u}_j = au_j, \quad j = \overline{0, n} \quad (5.21)$$

Рассмотрим применение данного интерполяционного метода для линеаризации нелинейностей. В этом случае $n = 1$ и из (5.17) получаем $u_0 = \cos \frac{\pi}{4} = \frac{\sqrt{2}}{2}$, $u_1 = \cos \frac{3\pi}{4} = -\frac{\sqrt{2}}{2}$. Формула (5.21) даёт $\bar{u}_0 = \frac{\sqrt{2}}{2}a$, $\bar{u}_1 = -\frac{\sqrt{2}}{2}a$. Для случая нечётных нелинейностей $f(\bar{u}_0) = -f(\bar{u}_1) = f(\frac{\sqrt{2}}{2}a)$. Поэтому формула (5.16) с пересчитанными узлами интерполяции даёт

$$A_1 = f(\bar{u}_0) \cos \frac{\pi}{4} + f(\bar{u}_1) \cos \frac{3\pi}{4} = \sqrt{2}f\left(\frac{\sqrt{2}}{2}a\right). \quad (5.22)$$

Из (5.20) окончательно получаем

$$p_1(u) = \frac{\sqrt{2}}{a} f\left(\frac{\sqrt{2}}{2}a\right)u. \quad (5.23)$$

Отсюда вытекает следующая общая формула для коэффициента линеаризации для нечётных нелинейных характеристик

$$k_1 = \frac{\sqrt{2}}{a} f\left(\frac{\sqrt{2}}{2}a\right). \quad (5.24)$$

Эта формула может быть использована и для нелинейностей, заданных графически.

Для обобщённой нелинейности типа насыщения (5.7) из (5.24) получаем следующее выражение для коэффициента линеаризации:

$$k_1 = \frac{\sqrt{2}h}{a}, \quad a > \sqrt{2}l. \quad (5.25)$$

Для обобщённой нелинейности типа зоны нечувствительности (5.9) формула (5.24) даёт выражение для коэффициента линеаризации вида

$$k_1 = \frac{\sqrt{2}}{a} \operatorname{tg} \alpha \left(\frac{\sqrt{2}}{2}a - l\right), \quad a > \sqrt{2}l. \quad (5.26)$$

В качестве полинома, близкого к полиному наилучшего равномерного приближения, можно также выбрать отрезок ряда Фурье по полиномам Чебышева. В этом случае задача отыскания коэффициентов аппроксимирующего полинома сводится к задаче отыскания коэффициентов Фурье для ряда по полиномам Чебышева [5].

Для отрезка $[-1; 1]$ аппроксимирующий полином в этом случае записывается как

$$p_n(u) = c_0 + \sum_{i=1}^n c_i T_i(u), \quad (5.27)$$

а его коэффициенты — по формулам

$$c_0 = \frac{1}{\pi} \int_{-1}^1 f(u) \frac{du}{\sqrt{1-u^2}}, \quad (5.28)$$
$$c_i = \frac{2}{\pi} \int_{-1}^1 f(u) T_i(u) \frac{du}{\sqrt{1-u^2}}, \quad i \neq 0.$$

Среднеквадратическое отклонение δ^2 этого полинома от функции $f(u)$ на отрезке $[-1; 1]$ находится по формуле

$$\delta^2 = \int_{-1}^1 (f(u) - p_n(u))^2 \frac{du}{\sqrt{1-u^2}} = \int_{-1}^1 f^2(u) \frac{du}{\sqrt{1-u^2}} - \frac{\pi}{2} (c_0^2 + \sum_{i=1}^n c_i^2). \quad (5.29)$$

Расчёт коэффициентов удобно проводить используя замену переменной $y = \cos \theta$. Тогда

$$\begin{aligned} c_0 &= \frac{1}{\pi} \int_0^\pi f(\cos \theta) d\theta, \\ c_i &= \frac{2}{\pi} \int_0^\pi f(\cos \theta) \cos(i\theta) d\theta. \end{aligned} \quad (5.30)$$

Для отрезка $[-a; a]$ при вычислении коэффициентов Фурье используется замена переменной $y = a \cos \theta$.

Между полиномом среднеквадратичного приближения (5.27) и интерполяционным полиномом (5.15) существует тесная связь. А именно, если интегралы в формулах (5.28) вычислять с помощью квадратурной формулы Эрмита, то получим

$$\begin{aligned} c_i &= \frac{2}{\pi} \int_{-1}^1 f(u) T_i(u) \frac{du}{\sqrt{1-u^2}} \approx \frac{\pi}{N} \cdot \frac{2}{\pi} \sum_{k=1}^N f(u_k) T_i(u_k) = \\ &= \frac{2}{N} \sum_{k=1}^N f(u_k) \cos(i \arccos u_k). \end{aligned} \quad (5.31)$$

Здесь использовалась формула (5.11). Теперь, полагая $u_k = \cos \frac{(2k-1)\pi}{2N}$, $j = k - 1$ и $N = n + 1$, находим

$$c_i \approx \frac{2}{n+1} \sum_{j=0}^n f(u_j) \cos \frac{(2j+1)k\pi}{2(n+1)}, \quad (5.32)$$

где $u_j = \cos \frac{(2j+1)\pi}{2(n+1)}$. Отсюда следует, что коэффициенты полинома среднеквадратического приближения c_i и коэффициенты интерполяционного полинома A_i отличаются между собой только на погрешность квадратурной

формулы Эрмита, которая является малой величиной для функций, не имеющих резких изменений значений, какими и являются обычно нелинейные статические характеристики. Поэтому в практических задачах оба рассмотренных метода аппроксимации дают близкие результаты.

Г Л А В А 2 .

Линейные стационарные системы

§1. Краткие сведения о преобразованиях Лапласа и Фурье

Определение 1. Назовём функцию $f(t) : \mathbb{R} \rightarrow \mathbb{C}$ *оригиналом*, если:

1) $f(t) = 0 \quad \forall t < 0;$

2) $f(t)$ удовлетворяет условию Гёльдера за исключением, быть может, не более чем счётного числа точек разрыва первого рода, причём на любом конечном интервале их должно быть конечное число.

3) $\exists M > 0, s \geq 0 : |f(t)| \leq M e^{st}$, то есть $f(t)$ возрастает не быстрее показательной функции.

Замечание 1. Условие Гёльдера заключается в том, что для всякой неособой точки $t \exists A, \alpha \in [0; 1], h_0 :$

$$|f(t+h) - f(t)| \leq A |h|^\alpha \quad \forall h : |h| \leq h_0.$$

Любая кусочно-дифференцируемая на \mathbb{R} функция (с разрывами первого рода) удовлетворяет этому условию, также как и любая функция, удовлетворяющая условию Липшица

$$|f(t+h) - f(t)| \leq A|h| \quad \forall t \in \mathbb{R}.$$

Пример 1 (Единичная функция (функция Хевисайда)). Простейшим примером функции-оригинала является функция $1(t)$:

$$1(t) = \begin{cases} 0, & t < 0, \\ 1, & t \geq 0. \end{cases} \quad (1.1)$$

Она, очевидно, удовлетворяет всем условиям, налагаемым на оригиналы.

Если мы рассмотрим функцию $\varphi(t)$ из класса функций, удовлетворяющих условиям 2) и 3), то, домножив её на $1(t)$, мы получим функцию

$$f(t) = \begin{cases} 0, & t < 0, \\ \varphi(t), & t \geq 0, \end{cases}$$

являющуюся оригиналом. При этом, если $\varphi(t)$ была непрерывной в точке 0, то $f(t)$ в этой точке будет непрерывна справа.

Условимся в дальнейшем, если это не будет отмечено особо, опускать множитель $1(t)$, и считать, что все рассматриваемые нами функции $\varphi(t)$, удовлетворяющие условиям 2) и 3), домножены на функцию $1(t)$, то есть $\varphi(t) = 0 \forall t < 0$.

Определение 2. Пусть $f(t)$ — оригинал. Функцию $F(p) : \mathbb{C} \rightarrow \mathbb{C}$, определяемую как

$$F(p) = \mathcal{L}_{t \rightarrow p} f(t) = \int_0^{+\infty} f(t) e^{-pt} dt, \quad (1.2)$$

назовём *изображением по Лапласу* оригинала $f(t)$.

Символ $\mathcal{L}_{t \rightarrow p}$ обозначает оператор преобразования Лапласа. Тот факт, что $F(p)$ является изображением оригинала $f(t)$ и, наоборот, что $f(t)$ — оригинал $F(p)$, записывают как $F(p) \doteq f(t)$. Как правило, оригиналы записывают маленькими буквами, а изображения — большими.

Определение 3. Пусть $f(t)$ — оригинал. Число

$$s_0 = s_0(f) = \inf \{s \mid |f(t)| \leq M e^{st} \forall t\}$$

называется *показателем роста* функции $f(t)$.

Приведём несколько теорем о связи оригиналов и изображений.

Теорема 1. Для оригинала $f(t)$ изображение $F(p)$ определено в полуплоскости $\operatorname{Re} p > s_0$, где s_0 — показатель роста $f(t)$, и является в этой полуплоскости аналитической функцией.

Замечание 2. Не смотря на то, что существование $F(p)$ гарантируется только в полуплоскости $\operatorname{Re} p > s_0$, очень часто функцию $F(p)$ можно аналитически продолжить на прямую $\operatorname{Re} p = s_0$ и слева от неё.

Теорема 2. Если $f(t)$ является оригиналом, а $F(p)$ — изображением $f(t)$, то $\forall t > 0$, где $f(t)$ удовлетворяет условию Гельдера, то

$$f(t) = \frac{1}{2\pi j} \int_{a-j\infty}^{a+j\infty} F(p)e^{pt} dp,$$

где интеграл понимается в смысле главного значения¹ и берётся вдоль прямой $\operatorname{Re} p = a$, $a > s_0$, где s_0 — показатель роста $f(t)$.

При этом оригинал $f(t)$ определяется своим изображением $F(p)$ с точностью до значений в точках разрыва $f(t)$.

Оператор обратного преобразования будем обозначать $\mathcal{L}_{p \rightarrow t}^{-1}$.

Следующая теорема определяет условия для того, чтобы произвольная функция $F(p)$ была бы изображением.

Теорема 3. Если $F(p) : \mathbb{C} \rightarrow \mathbb{C}$ аналитична в некоторой полуплоскости $\operatorname{Re} p > s_0$, равномерно сходится к 0 при $|p| \rightarrow +\infty$ в любой полуплоскости $\operatorname{Re} p \geq a > s_0$ и

$$\int_{a-j\infty}^{a+j\infty} F(p) dp$$

сходится равномерно, то $F(p)$ является изображением функции $f(t)$, определяемой соотношением

$$f(t) = \frac{1}{2\pi j} \int_{a-j\infty}^{a+j\infty} F(p)e^{pt} dp.$$

Из этих теорем видно, что для достаточно хороших оригиналов (непрерывно дифференцируемых в \mathbb{R}_+ и, тем более, бесконечно дифференцируемых) изображения определяются взаимоднозначно.

В табл. 1.1 перечислены основные свойства преобразования Лапласа.

В теории систем наиболее важны теоремы о дифференцировании оригинала и о свёртке. Остановимся кратко на одном приложении этих теорем к теории дифференциальных уравнений.

¹То есть $f(t) = \frac{1}{2\pi j} \lim_{b \rightarrow +\infty} \int_{a-bj}^{a+bj} F(p)e^{pt} dp$.

Таблица 1.1

Свойства преобразования Лапласа

Название свойства (теоремы)	Оригинал	Изображение	Условие
Линейность	$\alpha f(t) + \beta g(t)$	$\alpha F(p) + \beta G(p)$	$\alpha, \beta \in \mathbb{C}$, $\operatorname{Re} p > \max\{s_0(f), s_0(g)\}$.
Подобие	$f(\alpha t)$	$\frac{1}{\alpha} F\left(\frac{p}{\alpha}\right)$	$\alpha > 0$, $\operatorname{Re} p > \alpha s_0(f)$.
Дифференцирование оригинала	$f'(t)$	$pF(p) - f(0)$	f' — оригинал, $\operatorname{Re} p > \max\{s_0(f), s_0(f')\}$.
	$f^{(n)}(t)$	$p^n F(p) - p^{n-1} f(0) - \dots - f^{(n-1)}(0)$	$f^{(i)} \forall i = \overline{0, n}$ — оригинал, $\operatorname{Re} p > \max_{i=\overline{0, n}} s_0(f^{(i)})$. $f^{(i)}(0) = f^{(i)}(0 + 0) \forall i = \overline{0, n}$
Дифференцирование изображения	$(-1)^n t^n f(t)$	$F^{(n)}(p)$	$\operatorname{Re} p > s_0(f)$
Интегрирование оригинала	$\int_0^t f(\tau) d\tau$	$\frac{F(p)}{p}$	$\operatorname{Re} p > \max\{s_0(f), 0\}$
Интегрирование изображения	$\frac{f(t)}{t}$	$\int_p^\infty F(p) dp$, $\operatorname{Re} p \rightarrow +\infty$	$\operatorname{Re} p > s_0(f)$, $\int_p^\infty F(p) dp$ сходится при $\operatorname{Re} p \rightarrow +\infty$.
Запаздывание	$f(t - \tau)$	$e^{-p\tau} F(p)$	$\operatorname{Re} p > s_0(f), \tau > 0$.
Смещение	$e^{p_0 t} f(t)$	$F(p - p_0)$	$\operatorname{Re} p > s_0(f) + \operatorname{Re} p_0, p_0 \in \mathbb{C}$
Изображение свёртки (теорема Бореля)	$\int_0^t f(\tau) g(t - \tau) d\tau$	$F(p)G(p)$	$\operatorname{Re} p > \max\{s_0(f), s_0(g)\}$

Пусть задано обыкновенное дифференциальное уравнение с нулевыми начальными условиями

$$\begin{cases} y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \dots + a_0y(t) = \\ \qquad \qquad \qquad = b_my^{(m)}(t) + b_{m-1}y^{(m-1)}(t) + \dots + b_0y(t), \\ y(0) = \dots = y^{(n-1)}(0) = 0, \\ y(t) = 0 \quad \forall t < 0, \\ u(t) = 0 \quad \forall t < 0. \end{cases} \quad (1.3)$$

Применив к нему преобразование Лапласа и используя теорему о дифференцировании оригинала, получим

$$(p^n + a_{n-1}p^{n-1} + \dots + a_0)Y(p) = (b_mp^m + \dots + b_0)U(p),$$

где $p^n + a_{n-1}p^{n-1} + \dots + a_0$ — характеристический многочлен уравнения (1.3).

Выразив $Y(p)$ через $U(p)$, получим, что

$$Y(p) = H(p)U(p), \quad \text{где } H(p) = \frac{\sum_{i=0}^m b_i p^i}{\sum_{i=0}^n a_i p^i}, \quad a_n = 1,$$

откуда $y(t) \doteq Y(p)$ находится после вычисления $h(t) \doteq H(p)$ применением теоремы о свёртке (теоремы Бореля).

Замечание 3. а) Метод применим только в том случае, когда $y(t)$ и $u(t)$ являются оригиналами.

б) В общем случае вычисление обратного преобразования Лапласа и, тем более, интеграла свёртки сопряжено с известными трудностями. Однако в случае обыкновенного дифференциального уравнения типа (1.3) функция $H(p)$, обратное преобразование которой надо вычислять, является дробно-рациональной, что существенно облегчает задачу.

Имеет место следующая теорема о вычислении обратного преобразования Лапласа дробно-рациональной функции $H(p)$.

Теорема 4. Если $H(p) = \frac{B(p)}{A(p)}$, $\deg B(p) < \deg A(p)$ — дробно-рациональная функция, то её оригиналом служит

$$h(t) = 1(t) \sum_{k=1}^l \frac{1}{(n_k - 1)!} \lim_{p \rightarrow p_k} \frac{d^{n_k-1}}{dp^{n_k-1}} [H(p)(p - p_k)^{n_k} e^{pt}], \quad (1.4)$$

где p_k — полюсы $H(p)$, а n_k — их кратности.

Замечание 4. а) Если $\text{gcd}(A(p), B(p)) = 1$ (gcd — наибольший общий делитель, НОД), то есть $A(p)$ и $B(p)$ не имеют общих корней, то $\{p_k\}$ — множество корней многочлена $A(p)$, а $\{n_k\}$ — их кратности.

б) Если, вдобавок, все корни $A(p)$ — простые, то

$$h(t) = \sum_{k=1}^n \frac{B(p_k)}{A'(p_k)} e^{p_k t} \quad (1.5)$$

Приведём ещё одну теорему, полезную при работе с преобразованием Лапласа, с помощью которой, в частности, можно проверять правильность вычислений, проведённых операционным методом.

Теорема 5. Если $f(t)$ — оригинал, а $F(p)$ — изображение этого оригинала, то:

1) $\lim_{p \rightarrow \infty} pF(p) = f(0+0)$, где $p \rightarrow \infty$ внутри угла $|\arg p| < \frac{\pi}{2} - \delta$ (в направлении $+\infty$ действительной оси).

2) Кроме того, если $\exists \lim_{t \rightarrow +\infty} f(t) = f(+\infty)$, то $\lim_{p \rightarrow 0} pF(p) = f(+\infty)$, где $p \rightarrow 0$ внутри того же угла.

В табл. 1.2 приведены изображения по Лапласу некоторых распространённых функций.

Таблица 1.2

Преобразования Лапласа некоторых функций

Оригинал	Изображение	Оригинал	Изображение
$1(t)$	$\frac{1}{p}$	$e^{\lambda t}$	$\frac{1}{p - \lambda}$
$t^\alpha, \alpha > -1$	$\frac{\Gamma(\alpha + 1)}{p^{\alpha+1}}$	$e^{\lambda t} t^\alpha$	$\frac{\Gamma(\alpha + 1)}{(p - \lambda)^{\alpha+1}}$
$\sin \lambda t$	$\frac{\lambda}{p^2 + \lambda^2}$	$\cos \lambda t$	$\frac{p}{p^2 + \lambda^2}$
$t \sin \lambda t$	$\frac{2\lambda p}{(p^2 + \lambda^2)^2}$	$t \cos \lambda t$	$\frac{p^2 - \lambda^2}{(p^2 + \lambda^2)^2}$
$e^{\alpha t} - e^{\beta t}$	$\frac{\alpha - \beta}{(p - \alpha)(p - \beta)}$	$\frac{1}{t} (e^{\alpha t} - e^{\beta t})$	$\ln \frac{p - \beta}{p - \alpha}$
$e^{\alpha t} \sin \beta t$	$\frac{\beta}{(p - \alpha)^2 + \beta^2}$	$e^{\alpha t} \cos \beta t$	$\frac{p - \alpha}{(p - \alpha)^2 + \beta^2}$

Перейдём теперь к определению преобразования Фурье.

Определение 4. Пусть функция $f(t) \in L_1(\mathbb{R})$. Функцию комплексного переменного $F(\omega)$, определяемую как

$$F(\omega) = \int_{\mathbb{R}} f(t)e^{-j\omega t} dt, \quad (1.6)$$

назовём *преобразованием Фурье* функции $f(t)$.

Обратное преобразование Фурье, то есть переход от $F(\omega)$ к $f(t)$, выражается как

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} F(\omega)e^{j\omega t} d\omega, \quad (1.7)$$

где интеграл понимается в смысле главного значения.

Тот факт, что функция $F(\omega)$ является преобразованием Фурье функции $f(t)$, снова записывают как $F(\omega) \doteq f(t)$.

Заметим, что, вообще говоря, функция $F(\omega)$ не обязана являться функцией класса $L_1(\mathbb{R})$ (см. соответствующий пример в [12]), поэтому обратное преобразование Фурье не всегда существует. Однако, имеется следующая теорема, позволяющая нам рассматривать достаточно широкий класс функций, для которых гарантированно существует обратное преобразование Фурье.

Теорема 6 (Планшереля). Для функции $f \in L_2(\mathbb{R})$ интеграл

$$F_N(\omega) = \int_{-N}^N f(t)e^{-j\omega t} dt$$

$\forall N$ есть функция из $L_2(\mathbb{R})$, сходящаяся в метрике пространства $L_2(\mathbb{R})$ к некоторому пределу $F(\omega)$, называемому преобразованием Фурье функции $f \in L_2(\mathbb{R})$. Если функция $f \in L_1(\mathbb{R})$, то соответствующая функция $F(\omega)$ совпадает с преобразованием Фурье функции $f(t)$ в смысле (1.6).

Таким образом, теорема Планшереля позволяет сделать вывод о том, что преобразование Фурье является оператором

$$\mathcal{F} : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R}),$$

где $L_2(\mathbb{R})$ — пространство комплекснозначных функций, интегрируемых с квадратом.

Свойства преобразования Фурье во многом напоминают свойства преобразования Лапласа. Например, если функция $f(t)$ дифференцируема необходимое число раз, то

$$\begin{aligned}\mathcal{F}\left(\frac{df}{dt}\right) &= j\omega \mathcal{F}(f), \\ \mathcal{F}\left(\frac{d^{(k)}f}{dt^k}\right) &= (j\omega)^k \mathcal{F}(f).\end{aligned}\tag{1.8}$$

В пользу этого можно привести следующие соображения (строгие доказательства можно найти в [12, 25]). Пусть $F(p)$ — преобразование Лапласа функции $f(t)$. Сделаем замену переменных $p = j\omega$. Тогда

$$\tilde{F}(\omega) = F(j\omega) = \int_0^{+\infty} f(t)e^{-j\omega t} dt,\tag{1.9}$$

откуда следует, что по крайней мере для оригиналов по Лапласу, которые также принадлежат $L_2(\mathbb{R})$, $F(j\omega)$ является соответствующим преобразованием Фурье. Всюду далее, где это не отмечено особо, мы будем считать, что операция перехода от преобразования Лапласа к преобразованию Фурье путём замены переменных $p \rightarrow j\omega$ определена корректно и все необходимые условия выполняются.

Рассмотрим теперь вопрос существования преобразований Лапласа и Фурье от обобщённой функции $\delta(t - t_0)$.

Если рассматривать $\delta(t - t_0)$ как слабый предел δ -образной последовательности, то для $t_0 > 0$ можно ввести

$$\mathcal{L} \delta(t - t_0) = \int_0^{+\infty} \delta(t - t_0)e^{-pt} dt = e^{-pt_0}.\tag{1.10}$$

Трудность возникает лишь с определением преобразования Лапласа от функции $\delta(t)$, так как она даёт скачок в нулевой момент времени и непонятно, как определять

$$\int_0^{+\infty} \delta(t)e^{-pt} dt.$$

Оставаясь в рамках определения $\delta(t)$ как предела δ -образной последовательности, положим

$$\mathcal{L} \delta(t) = \lim_{t_0 \rightarrow 0+0} \mathcal{L} \delta(t - t_0) = \lim_{t_0 \rightarrow 0+0} e^{-pt_0} = e^0 = 1. \quad (1.11)$$

Другой способ определения преобразования Лапласа $\mathcal{L} \delta(t)$ основан на понятии *обобщённой производной* функции $f(t)$.

Напомним, что обобщённые функции действуют над гладкими бесконечнодифференцируемыми функциями с компактным носителем, $\varphi \in C_0^\infty(\mathbb{R})$. Тогда для функции $f \in C^1(\mathbb{R})$, применяя формулу интегрирования по частям и учитывая, что $\varphi(\pm\infty) = 0$, получаем

$$(f', \varphi) = \int_{\mathbb{R}} f'(t) \varphi(t) dt = - \int_{\mathbb{R}} f(t) \varphi'(t) dt = -(f, \varphi'). \quad (1.12)$$

Обобщая это равенство для достаточно широкого класса функций, получаем следующее

Определение 5. *Обобщённой производной* функции $f(t)$ называется такая функция f' , что $\forall \varphi \in C_0^\infty(\mathbb{R})$

$$(f', \varphi) \stackrel{\text{def}}{=} -(f, \varphi'). \quad (1.13)$$

Не останавливаясь подробно на свойствах операции, определённой в (1.13), отметим, в частности, что обобщённая производная совпадает с классической для дифференцируемых функций.

Для единичной функции Хевисайда $1(t)$ имеем

$$\begin{aligned} (1'(t), \varphi) &= -(1(t), \varphi') = - \int_0^{+\infty} \varphi'(t) dt = \\ &= -1[\varphi(+\infty) - \varphi(0)] = \varphi(0) = (\delta, \varphi), \end{aligned} \quad (1.14)$$

что означает, что в обобщённом смысле $1'(t) = \delta(t)$. Тогда для оператора преобразования Лапласа справедливо (опять-таки, в обобщённом смысле), что

$$\mathcal{L} \delta(t) = \mathcal{L} 1'(t) = p \mathcal{L} 1(t) = p \frac{1}{p} = 1, \quad (1.15)$$

что совпадает с результатами (1.11). Формулу же (1.10) можно получить, применив теорему запаздывания, что ещё раз подтверждает правильность сделанных предположений.

Вычислить преобразование Фурье для $\delta(t - t_0)$ не представляет сложности,

$$\mathcal{F} \delta(t - t_0) = \int_{\mathbb{R}} \delta(t - t_0) e^{-j\omega t} dt = e^{-j\omega t_0}, \quad (1.16)$$

для любого $t_0 \in \mathbb{R}$. Формальный вывод этого равенства основывается на определении преобразования Фурье обобщённых функций, ознакомиться с которым читатель может в [4, 25].

Рассмотрим преобразования Фурье функций $\cos \omega_0 t$, $\sin \omega_0 t$, важные для приложений. Хотя в обычном смысле преобразования Фурье для этих функций не существует (так как эти функции не принадлежат $L_1(\mathbb{R})$ и $L_2(\mathbb{R})$), можно рассмотреть преобразование Фурье в обобщённом смысле.

Заметим, прежде всего, что

$$\frac{1}{2\pi} \int_{\mathbb{R}} e^{j\omega t} \delta(\omega - \omega_0) d\omega = \frac{e^{j\omega_0 t}}{2\pi}.$$

и, из этого,

$$\begin{aligned} e^{j\omega_0 t} &\doteq 2\pi \delta(\omega - \omega_0), \\ e^{-j\omega_0 t} &\doteq 2\pi \delta(\omega + \omega_0). \end{aligned}$$

Наконец, из формулы Эйлера

$$\begin{aligned} \cos \omega_0 t &= \frac{1}{2}(e^{j\omega_0 t} + e^{-j\omega_0 t}), \\ \sin \omega_0 t &= \frac{1}{2j}(e^{j\omega_0 t} - e^{-j\omega_0 t}). \end{aligned}$$

Таким образом, окончательно получаем следующий результат:

$$\begin{aligned} \cos \omega_0 t &\doteq \pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)], \\ \sin \omega_0 t &\doteq -j\pi[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)]. \end{aligned} \quad (1.17)$$

В заключение хотелось бы отметить, что преобразования Фурье, понимаемые в обобщённом смысле, сами являются обобщёнными функциями, и на их применение накладываются соответствующие ограничения, рассматриваемые в общей теории обобщённых функций, см. [4, 12, 25].

§2. Передаточные функции линейных стационарных систем

Рассмотрим линейную стационарную систему, описываемую бинарным отношением $\Sigma = \langle \mathbb{R}_+, U, Y, \Omega, \Gamma, f \rangle$, где $f : \Omega \rightarrow \Gamma$ является функцией (оператором)², а Ω и Γ — линейные (бесконечномерные) пространства над \mathbb{R} .

Напомним, что *линейность* обозначает, что $\forall \alpha, \beta \in \mathbb{R}, \forall u, v \in \Omega$

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v), \quad (2.1)$$

а *стационарность*, что

$$f(\nabla_\sigma u) = \nabla_\sigma f(u), \quad (2.2)$$

где ∇_σ — оператор сдвига, описанный в §2 главы 1.

В общем случае линейную систему над \mathbb{R} можно представить с помощью оператора Вольтерра

$$y(t) = \int_{\mathbb{R}} k(t, \tau) u(\tau) d\tau, \quad (2.3)$$

где для *причинной* системы $k(t, \tau) = 0 \forall \tau > t$.

Рассмотрим следующую теорему.

Теорема 1. *Для того, чтобы линейная система (2.3) была стационарной, необходимо и достаточно, чтобы ядро интегрального оператора зависело только от разности $t - \tau$, но не от t и τ по отдельности.*

Таким образом, стационарная линейная система описывается следующим интегральным оператором:

$$y(t) = \int_{\mathbb{R}} h(t - \tau) u(\tau) d\tau. \quad (2.4)$$

Доказательство. Докажем достаточность. Сделаем замену переменных $t - \tau = \xi$ и преобразуем формулу (2.4) к виду

$$y(t) = \int_{\mathbb{R}} h(\xi) u(t - \xi) d\xi. \quad (2.5)$$

²См. соответствующие определения главы 1.

Заменим теперь $u(t)$ на $u(t + \sigma)$. Тогда

$$\int_{\mathbb{R}} h(\xi)u(t + \sigma - \xi) d\xi = y(t + \sigma), \quad (2.6)$$

что и требовалось доказать. ■

Уравнение (2.4) описывает, тем самым, линейный стационарный оператор $f : \Omega \rightarrow \Gamma$, а импульсная переходная функция $h(t)$ является функцией одного переменного. Условие причинности для этого оператора можно записать так:

$$h(\xi) = 0 \quad \forall \xi < 0. \quad (2.7)$$

Исследование линейных стационарных систем в теории управления обычно осуществляют с помощью аппарата передаточных функций, реализуемых через преобразование Лапласа.

Запишем причинную линейную стационарную систему в виде

$$y(t) = \int_0^t h(t - \tau)u(\tau) d\tau = \int_0^t h(\tau)u(t - \tau) d\tau \quad (2.8)$$

(нижний предел интегрирования для стационарных систем t_0 можно положить равным нулю).

Преобразуем (2.8) по Лапласу и используем теорему о свёртке (см. предыдущий параграф):

$$Y(p) = H(p)U(p),$$

откуда, полагая $U(p) \neq 0 \quad \forall p \in \mathbb{C}$, получаем

$$H(p) = \frac{Y(p)}{U(p)}. \quad (2.9)$$

Определение 1. $H(p)$ называется *передаточной функцией* линейной стационарной системы (2.8).

$H(p)$ по сути является комплексным коэффициентом передачи между входным и выходным сигналами и есть преобразование Лапласа от импульсной переходной характеристики:

$$H(p) = \mathcal{L} h(t) = \int_0^{+\infty} h(\tau)e^{-p\tau} d\tau.$$

Рассмотрим задачу структурного анализа систем с использованием аппарата передаточных функций.

1. Параллельное соединение систем. На рис. 2.1 показана структурная схема параллельного соединения систем, задаваемых передаточными функциями $H(p)$ и $G(p)$. По определению,

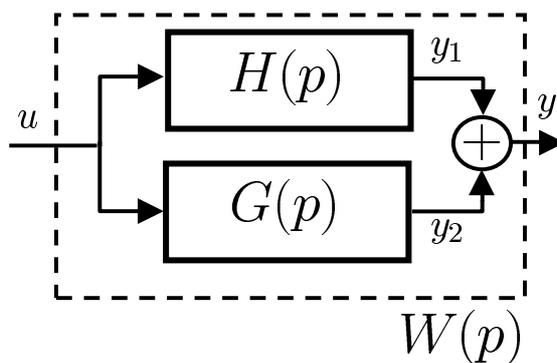


Рис. 2.1.

$$Y(p) = Y_1(p) + Y_2(p) = H(p)U(p) + G(p)U(p) = (H(p) + G(p))U(p)$$

и передаточная функция параллельного соединения систем записывается как

$$W(p) = H(p) + G(p). \quad (2.10)$$

2. Последовательное соединение систем. Рассмотрим последовательно соединённые системы, заданные передаточными функциями $H(p)$ и $G(p)$ (см. рис. 2.2). Тогда $Y(p) = G(p)Z(p)$, $Z(p) = H(p)U(p)$ и

$$Y(p) = G(p)H(p)U(p).$$

Передаточная функция последовательного соединения есть

$$W(p) = G(p)H(p). \quad (2.11)$$

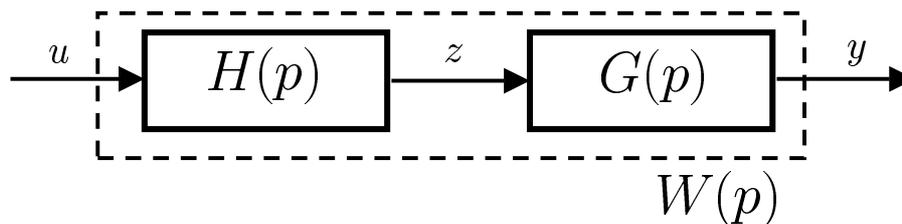


Рис. 2.2.

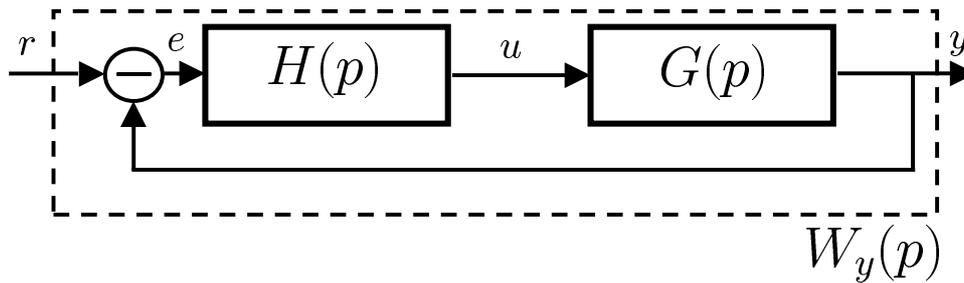


Рис. 2.3.

3. Линейная система с обратной связью.

Рассмотрим линейную систему с обратной связью (см. рис. 2.3). Здесь $H(p)$ задаёт передаточную функцию регулятора (управляющего устройства), $G(p)$ — управляемого объекта.

Запишем выражение для сигнала ошибки в изображениях по Лапласу:

$$\begin{aligned} E(p) &= R(p) - Y(p) = R(p) - G(p)H(p)E(p), \\ (1 + G(p)H(p))E(p) &= R(p). \end{aligned}$$

Передаточная функция по сигналу ошибки

$$W_e(p) = \frac{E(p)}{R(p)} = \frac{1}{1 + G(p)H(p)} \quad (2.12)$$

(заметим, что как для последовательного соединения, так и для соединения с обратной связью, передаточные функции $H(p)$ и $G(p)$ можно менять местами, т.е. стационарная система обладает свойством коммутативности относительно операции композиции). Передаточная функция по выходному сигналу записывается как:

$$W_y(p) = \frac{Y(p)}{R(p)} = \frac{H(p)G(p)}{1 + H(p)G(p)}, \quad (2.13)$$

а по сигналу управления как:

$$W_u(p) = \frac{U(p)}{R(p)} = \frac{H(p)}{1 + H(p)G(p)}. \quad (2.14)$$

В теории систем основным является случай, когда линейная причинная система задаётся обыкновенными дифференциальными уравнениями с нулевыми начальными условиями

$$\sum_{i=1}^n a_i y^{(i)}(t) = \sum_{i=1}^m b_i u^{(i)}(t), \quad m < n.$$

Применяя преобразование Лапласа, получим (см. выкладки предыдущего параграфа)

$$H(p) = \frac{Y(p)}{U(p)} = \frac{\sum_{i=0}^m b_i p^i}{\sum_{i=0}^n a_i p^i} \quad (2.15)$$

(получить передаточную функцию в смысле данного определения возможно, так как существует теорема об интегральном представлении решения дифференциального уравнения).

Таким образом, оказывается, что для систем, задаваемых линейными дифференциальными уравнениями, передаточная функция имеет дробно-рациональную структуру и, тем самым, можно легко найти обратное преобразование Лапласа с помощью теоремы о вычетах (см. предыдущий параграф).

Использование передаточных функций позволяет существенно алгебраизовать задачу исследования систем и упростить её, тем самым, теория линейных стационарных систем оперирует системами именно в терминах передаточных функций, а не в терминах ядер интегральных операторов.

§3. Понятие начального состояния и описания «ВХОД-СОСТОЯНИЕ-ВЫХОД»

Снова, как и в §1 первой главы, рассмотрим некоторый эксперимент с объектом типа «чёрный ящик», целью которого является построение математической модели этого объекта. В процессе эксперимента может оказаться, что «чёрный ящик» по-разному реагирует на один и тот же входной сигнал в одинаковых условиях эксперимента. Как следствие, описание «вход-выход» такого объекта, бинарное отношение, не будет функцией. Однозначная связь между множеством входных сигналов Ω и множеством выходных сигналов Γ при этом будет отсутствовать. С целью перехода от аппарата бинарных отношений к дающему большие возможности аппарату функций в таких случаях вводят абстрактный параметр начального состояния. При этом полагают, что если «чёрный ящик» по-разному реагировал на один и тот же входной сигнал в одинаковых условиях эксперимента, то это означает, что он находился в разных начальных состояниях.

Такая параметризация хорошо известна в теории дифференциальных уравнений. В самом деле, известно, что реакция системы, описываемой дифференциальным уравнением, зависит от начальных условий (см. примеры §1 первой главы) и, следовательно, начальные условия и являются в этом случае параметром начального состояния.

Введём следующее

Определение 1. Пусть дана неупреждающая абстрактная временная система $\Sigma = \langle T, U, Y, \Omega, \Gamma, S \rangle$, причём множество моментов времени T имеет наименьший элемент t_0 . Пусть, далее, существует непустое множество X^{t_0} , называемое множеством начальных состояний, элементами которого являются начальные состояния $x(t_0)$ в момент времени t_0 , а также функция $f : X^{t_0} \times \Omega \rightarrow \Gamma$, называемая описанием «вход–начальное состояние–выход», или «вход–состояние–выход», т.е. однопараметрическое семейство операторов $y = f(x(t_0), u)$, удовлетворяющее условиям:

1) согласованности, т.е. $\forall u \in \Omega$

$$\bigcup_{x(t_0) \in X^{t_0}} f(x(t_0), u) = S(u);$$

2) неупреждаемости, т.е. $\forall x(t_0) \in X^{t_0}, \forall u \in \Omega, \forall \sigma \in T$

$$P_\sigma f(x(t_0), P_\sigma u) = P_\sigma f(x(t_0), u),$$

где P_σ — правое отношение сечения (см. §2 гл. 1).

Тогда абстрактная система Σ называется *причинной* и задаётся семёркой

$$\Sigma = \langle T, U, Y, \Omega, \Gamma, X^{t_0}, f \rangle. \quad (3.1)$$

Внешнее описание причинной системы, её описание «вход–состояние–выход» задаётся тернарным (трёхместным) отношением $\bar{S} \in X^{t_0} \times \Omega \times \Gamma$, функциональным относительно пар $(x(t_0), u)$.

Параметр начального состояния $x(t_0)$ характеризует всю «предысторию» системы вплоть до момента времени t_0 и позволяет однозначно предсказать реакцию системы на заданное воздействие. При этом, вследствие свойства неупреждаемости, эта реакция в любой текущий момент времени $t \geq t_0$ не зависит от будущих значений воздействия на систему. Таким образом, модель (3.1) сохраняет связь причины и следствия, характерную для физического мира, что и определяет название этой модели.

Если абстрактная система Σ имеет в начальный момент времени t_0 множество начальных состояний X^{t_0} , то из этого следует, что для каждого момента времени $t > t_0$, если принять его за начальный, существует своё множество начальных состояний X^t . Пусть эти множества совпадают, т.е. $X^{t_0} = X^t = X \forall t \in T$.

Определение 2. Причинная система (3.1) *стационарна*, если $\forall \sigma > 0$

$$f(x(t_0 + \sigma), \nabla_\sigma u) = \nabla_\sigma f(x(t_0), u), \quad (3.2)$$

где ∇_σ — отношение сдвига (см. §2 гл. 1).

Предположим, что $X = \mathbb{R}^n$ и множество начальных состояний является, таким образом, линейным пространством размерности n .

Нули линейных пространств X, Ω, Γ будем обозначать буквой Θ с соответствующим индексом внизу, т.е. $\Theta_X, \Theta_\Omega, \Theta_\Gamma$. Если функциональная система является неупреждающей, то можно считать, что её множество начальных состояний содержит один нулевой элемент Θ_X для заданного t_0 , т.е. соответствующая система — причинная. Если множество входных сигналов Ω причинной системы Σ содержит один нулевой элемент Θ_Ω , то система называется *свободной*.

Определение 3. Причинная система Σ называется *линейной*, если:

1) её внешнее описание обладает свойством разложения, т.е.

$$\forall x(t_0) \in X, u \in \Omega$$

$$f(x(t_0), u) = f(x(t_0), \Theta_\Omega) + f(\Theta_\Omega, u); \quad (3.3)$$

2) оператор $f(x(t_0), \Theta_\Omega)$, задающий свободную систему, линеен по $x(t_0)$, а оператор $f(\Theta_X, u)$, задающий функциональную систему, линеен по u .

Если хотя бы одно из этих условий нарушается, то система называется *нелинейной*.

Из этого определения следует, что линейная система обладает свойством сохранения нуля, т.е.

$$f(\Theta_X, \Theta_\Omega) = \Theta_\Gamma. \quad (3.4)$$

Из формулы (3.3) вытекает, что для представления линейных систем описаниями «вход–состояние–выход» требуется иметь соответствующие описания для свободных и функциональных систем.

Получим необходимые представления для свободных и функциональных систем для случая стационарных систем с непрерывным временем. Рассмотрим сначала случай свободной системы.

Будем предполагать, что $X = \mathbb{R}^n$ и множество выходных сигналов Ω является пространством непрерывных функций C . Тогда описание «состояние–выход» свободной системы будет задаваться абстрактной функцией $f : \mathbb{R}^n \rightarrow C$. При каждом фиксированном значении $t \in T$ эта абстрактная функция является функционалом, поскольку из того, что $\Omega = C$ следует, что $Y = \mathbb{R}$. Поэтому здесь можно использовать общие формулы для представления линейных функционалов в конечномерных векторных пространствах. Таким образом, можно сразу записать требуемое представление

$$f(x(t_0), \Theta_\Omega) = \sum_{k=1}^n \gamma_k(t - t_0)x_k(t_0), \quad (3.5)$$

где $x_k(t_0)$ — координаты вектора начального состояния в момент времени t_0 , а $\gamma_k(t - t_0)$ — координатные функции. Для рассматриваемого стационарного случая они зависят от разности аргументов t, t_0 , что следует из формулы (3.2) при $u = \Theta_\Omega$. Если система функций $\{\gamma_k(t - t_0)\}$ линейно независимая, то она называется базисной. Базисная функция $\gamma_k(t - t_0)$ полностью описывает реакцию свободной линейной системы на начальное состояние со

всеми нулевыми координатами, кроме k -ой. У линейной системы существует бесконечное число базисов, приводимых друг к другу невырожденным линейным преобразованием.

Получим уравнения для базисных функций исходя из дифференциального уравнения

$$\sum_{i=0}^n a_i y^{(i)}(t) = 0, \quad (3.6)$$

где без ограничения общности можно положить $a_n = 1$. Будем искать решение однородного дифференциального уравнения (3.6) при заданных начальных условиях $y^{(n-1)}(0), \dots, y(0)$ для $t_0 = 0$. Используем для этого преобразование Лапласа.

Согласно теореме о дифференцировании

$$f^{(i)}(t) \doteq p^i F(p) - p^{i-1} f(0) - \dots - f^{(i-1)}(0). \quad (3.7)$$

Отсюда для (3.6) находим

$$\begin{aligned} \sum_{i=0}^n a_i p^i Y(p) &= (p^{n-1} + \dots + a_1) y(0) + (p^{n-2} + \dots + a_2) y^{(1)}(0) + \dots + \\ &+ (p + a_{n-1}) y^{(n-2)}(0) + y^{(n-1)}(0). \end{aligned}$$

Обозначив $L(p) = \sum_{i=0}^n a_i p^i$, получаем

$$Y(p) = \sum_{k=1}^n \frac{p^{n-k} + \dots + a_k}{L(p)} y^{(k-1)}(0). \quad (3.8)$$

Из сравнения формул (3.5) и (3.8) и теоремы существования и единственности решения обыкновенного дифференциального уравнения с заданными начальными условиями вытекает, что в рассматриваемом случае $x(0) = (y(0), \dots, y^{(n-1)}(0))^T$ и

$$\gamma_k(t) \doteq \frac{p^{n-k} + \dots + a_k}{L(p)}, \quad k = \overline{1, n}. \quad (3.9)$$

Для системы, задаваемой дифференциальным уравнением (3.6) начальное состояние и базисные функции можно ввести и иначе. Сгруппировав все слагаемые в числителе формулы (3.8) при одинаковых степенях

p , получим

$$Y(p) = \frac{1}{L(p)} [p^{n-1}y(0) + p^{n-2}(y^{(1)}(0) + a_{n-1}y(0)) + \dots + (y^{(n-1)}(0) + \dots + a_1y(0))]. \quad (3.10)$$

Отсюда следует, что

$$x(0) = \left(y(0), y^{(1)}(0) + a_{n-1}y(0), \dots, y^{(n-1)}(0) + \dots + a_1y(0) \right)^T$$

и новые базисные функции задаются уравнениями

$$\gamma_k(t) \doteq \frac{p^{n-k}}{L(p)}, \quad k = \overline{1, n}. \quad (3.11)$$

Связь новых координат состояния со старыми, т.е. с начальными условиями на дифференциальное уравнение, задаётся следующим преобразованием координат:

$$\begin{pmatrix} x_1(0) \\ x_2(0) \\ \dots \\ x_n(0) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{n-1} & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & a_2 & a_3 & \dots & 1 \end{pmatrix} \begin{pmatrix} y(0) \\ y^{(1)}(0) \\ \dots \\ y^{(n-1)}(0) \end{pmatrix}. \quad (3.12)$$

Матрица преобразования координат — нижняя треугольная и является невырожденной, так как её определитель равен 1.

Используя введённый вектор состояния $x(0) = (x_1, \dots, x_n)^T$, координаты которого определяются формулой (3.10), уравнение «состояние–выход» в рассматриваемом случае можно записать так:

$$Y(p) = \frac{x_1 p^{n-1} + \dots + x_n}{L(p)}. \quad (3.13)$$

Пусть характеристическое уравнение $L(p) = 0$ имеет только однократные действительные корни $\lambda_1, \dots, \lambda_n$. Тогда, разлагая (3.13) на элементарные дроби, получим

$$Y(p) = \frac{\tilde{x}_1}{p - \lambda_1} + \dots + \frac{\tilde{x}_n}{p - \lambda_n}, \quad (3.14)$$

где коэффициенты $\tilde{x}_1, \dots, \tilde{x}_n$ определяются формулами

$$\tilde{x}_i = \frac{x_1 p^{n-1} + \dots + x_n}{L'(p)} \Big|_{p=\lambda_i}, \quad i = \overline{1, n}. \quad (3.15)$$

Так как $\frac{1}{p-\lambda_i} \doteq e^{\lambda_i t}$, то система функций $e^{\lambda_1 t}, \dots, e^{\lambda_n t}$ линейно независима и является базисной. Вектор $\tilde{x}(0) = (\tilde{x}_1, \dots, \tilde{x}_n)^T$ задаёт новый вектор начального состояния. Существует обобщение этого метода на случай кратных действительных корней [8]. Для комплексных корней данный метод не используется, т.к. формулы (3.15) в этом случае дают комплексные координаты вектора начального состояния.

Рассмотрим теперь неоднородное уравнение, соответствующее однородному дифференциальному уравнению (3.6):

$$\sum_{i=0}^n a_i y^{(i)}(t) = u(t). \quad (3.16)$$

Для того, чтобы получить представление для функциональной системы в (3.3), примем начальное состояние нулевым. Применяя преобразование Лапласа для нулевых начальных условий, как и ранее, находим, что

$$Y(p) = H(p)U(p), \quad (3.17)$$

где $H(p) = \frac{1}{L(p)}$ — передаточная функция. Согласно теореме о свёртке это уравнение может быть записано во временной области в виде интеграла

$$y(t) = \int_0^t h(t - \tau)u(\tau) d\tau, \quad (3.18)$$

где $h(t - \tau)$ — импульсная переходная характеристика (весовая функция). Теперь, применяя формулу (3.3) и формулы для представления свободной (3.5) (при $t_0 = 0$) и функциональной систем (3.18), окончательно получаем такое описание «вход–состояние–выход» для системы, задаваемой дифференциальным уравнением (3.16) при заданных (ненулевых) начальных условиях:

$$y(t) = \sum_{k=1}^n \gamma_k(t)x_k(0) + \int_0^t h(t - \tau)u(\tau) d\tau. \quad (3.19)$$

Это же уравнение задаёт общую формулу для описаний «вход–состояние–выход» для систем, задаваемых линейными дифференциальными уравнениями с постоянными коэффициентами общего вида:

$$\sum_{i=0}^n a_i y^{(i)}(t) = \sum_{i=0}^m b_i u^{(i)}(t), \quad m < n. \quad (3.20)$$

В вектор начального состояния такой системы должны входить также начальные условия на правую часть дифференциального уравнения. Передаточная функция, определяемая для нулевых начальных условий, для такой системы имеет вид

$$H(p) = \frac{\sum_{i=1}^m b_i p^i}{\sum_{i=1}^n a_i p^i}. \quad (3.21)$$

Она, очевидно, не зависит от выбора базиса в пространстве начальных состояний.

Формула (3.21) является основной в теории линейных систем управления.

§4. Устойчивость линейных систем управления

Первым требованием, предъявляемым к работоспособной системе управления, является требование её устойчивой работы при действующих на систему задающих сигналах и возмущениях. Рассмотрим проблему устойчивости в общем виде. Пусть дана абстрактная система, заданная описанием «вход-состояние-выход»

$$y(t) = f(x(t_0), u(t)), \quad t \in [t_0; +\infty), \quad (4.1)$$

где $u(t)$ — входной сигнал, $y(t)$ — выходной сигнал и $x(t_0)$ — начальное состояние системы в момент времени t_0 (рис. 4.1). Пусть $x(t_0) \in X$,

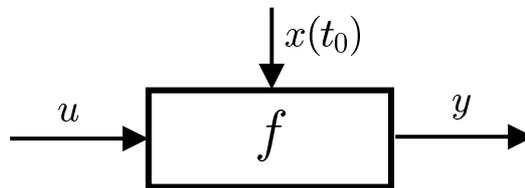


Рис. 4.1.

$u \in \Omega$, $y \in \Gamma$ и множества X, Ω, Γ являются линейными нормированными пространствами, нулевые элементы которых будем обозначать, как и в предыдущем параграфе, буквой Θ с соответствующим индексом: $\Theta_X, \Theta_\Omega, \Theta_\Gamma$. Пусть $f(\Theta_X, \Theta_\Omega) = \Theta_\Gamma$.

Определение 1. Систему (4.1) назовём: 1) *устойчивой в малом*, если оператор f непрерывен в точке $(\Theta_X, \Theta_\Omega)$; 2) *устойчивой в большом в области* $D \subseteq X \times \Omega : (\Theta_X, \Theta_\Omega) \in D$, если оператор f непрерывен и ограничен на D ; 3) *устойчивой в целом*, если оператор f непрерывен и ограничен на всём пространстве $X \times \Omega$.

Свойство непрерывности в этих определениях гарантирует, что для устойчивой системы малым изменениям входного сигнала и начального состояния будут соответствовать малые изменения выходного сигнала. Свойство ограниченности означает, что для устойчивой системы ограниченным входному сигналу и начальному состоянию будут соответствовать ограниченные выходные сигналы. Если зафиксировать начальное состояние, то мы получим определение *устойчивости по входному сигналу*, а если зафиксировать входной сигнал — *устойчивость по начальному состоянию*. Система (4.1) может быть устойчивой по входному сигналу, но неустойчивой (т.е. не быть устойчивой в смысле нашего определения) по начальному состоянию и наоборот.

Определение 2. Систему (4.1) назовём *асимптотически устойчивой в малом*, если она устойчива в малом и для всех начальных состояний $x(t_0)$ в ε -окрестности Θ_X

$$\lim_{t \rightarrow +\infty} y(t, x(t_0), u(t)) = \lim_{t \rightarrow +\infty} y(t, \Theta_X, u(t)).$$

Аналогично очевидным образом определяется асимптотическая устойчивость в большом и в целом.

Рассмотрим линейную стационарную систему с передаточной функцией (3.21), не имеющей сокращающихся нулей и полюсов, при $m < n$:

$$H(p) = \frac{\sum_{i=0}^m b_i p^i}{\sum_{i=0}^n a_i p^i}. \quad (4.2)$$

Напомним, что передаточная функция (4.2) соответствует дифференциальному уравнению

$$\sum_{i=0}^n a_i y^{(i)}(t) = \sum_{i=0}^m b_i u^{(i)}(t).$$

Теорема 1. Для системы, заданной передаточной функцией (4.2), следующие три положения эквивалентны [8]:

- 1) система асимптотически устойчива в целом;
- 2) импульсная переходная характеристика (весовая функция) системы $h(t)$ абсолютно интегрируема, т.е.

$$\int_0^t |h(\tau)| d\tau < +\infty; \quad (4.3)$$

- 3) характеристическое уравнение $\sum_{i=0}^n a_i p^i = 0$ не имеет корней при $\operatorname{Re} p \geq 0$ (в правой полуплоскости и на мнимой оси).

Проверку выполнения условия отсутствия корней характеристического многочлена при $\operatorname{Re} p \geq 0$ можно провести прямым методом, т.е. находя

численно эти корни, что не представляет проблем для $n \leq 10$. Для систем высокого порядка исследование устойчивости удобнее проводить одним из косвенных методов, позволяющих проверять указанное условие по коэффициентам характеристического уравнения.

Например, это можно сделать с помощью критериев Рауса и Гурвица, а также необходимого условия устойчивости Стодолы.

Теорема 2 (необходимое условие устойчивости Стодолы). *Для того, чтобы характеристическое уравнение $\sum_{i=0}^n a_i p^i = 0$ не имело корней в правой полуплоскости, необходимо, чтобы при $a_n > 0$ все коэффициенты a_i были бы положительны.*

Доказательство. Используя основную теорему алгебры, запишем характеристический полином в виде

$$a_n(p - \lambda_1) \dots (p - \lambda_n). \quad (4.4)$$

Каждому действительному корню λ_i соответствует двучлен $p - \lambda_i$. Двум комплексно-сопряжённым корням $\lambda_{k,k+1} = \sigma_k \pm \omega_k j$ соответствует полином второй степени $(p - \sigma_k)^2 + \omega_k^2$. Действительно,

$$\begin{aligned} (p - \sigma + \omega j)(p - \sigma - \omega j) &= p^2 - p\sigma + p\omega j - p\sigma + \sigma^2 - \sigma\omega j - p\omega j + \sigma\omega j + \omega^2 = \\ &= p^2 - 2p\sigma + \sigma^2 + \omega^2 = (p - \sigma)^2 + \omega^2. \end{aligned}$$

Тогда очевидно, что при всех $\lambda_i < 0$ и $\sigma_k < 0$ все сомножители в произведении (4.4) будут иметь положительные коэффициенты, а, значит, будут положительны и коэффициенты полинома $\sum_{i=0}^n a_i p^i$, которые являются произведениями этих коэффициентов. ■

Критерий устойчивости Гурвица формулируется следующим образом.

Теорема 3 (критерий Гурвица). *Корни полинома $\sum_{i=0}^n a_i p^i$, $a_0 > 0$ лежат в левой полуплоскости тогда и только тогда, когда положительны все главные миноры матрицы*

$$\begin{pmatrix} a_1 & a_0 & 0 & 0 & \dots & 0 \\ a_3 & a_2 & a_1 & a_0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{2n-1} & a_{2n-2} & a_{2n-3} & a_{2n-4} & \dots & a_n \end{pmatrix}. \quad (4.5)$$

При составлении этой матрицы коэффициенты a_i с индексами i , большими чем n , полагаются равными нулю. Условие устойчивости означает требование положительности следующей системы определителей Гурвица (при $a_0 > 0$):

$$\Delta_1 = a_1 > 0, \Delta_2 = \begin{vmatrix} a_1 & a_0 \\ 0 & a_2 \end{vmatrix} > 0, \Delta_3 = \begin{vmatrix} a_1 & a_0 & 0 \\ a_3 & a_2 & a_1 \\ 0 & 0 & a_3 \end{vmatrix} > 0, \dots$$

Поскольку в последней строке матрицы (4.5) все элементы, кроме последнего, равны нулю, то

$$\Delta_n = a_n \Delta_{n-1}$$

и, по условию Стодолы, $a_n > 0$, то условие положительности последнего определителя Гурвица сводится к условию $\Delta_{n-1} > 0$.

Если $\Delta_{n-1} = 0$ при положительности всех остальных определителей Гурвица, то система находится на границе устойчивости, т.е. характеристическое уравнение имеет пару чисто мнимых корней [2].

Для систем порядка $n = 1, 2, 3$ условия устойчивости запишутся так:

- а) $a_0 > 0, a_1 > 0$;
- б) $a_0 > 0, a_1 > 0, a_2 > 0$;
- в) $a_0 > 0, a_1 > 0, a_2 > 0, a_3 > 0, a_1 a_2 - a_3 a_0 > 0$.

Таким образом, для систем первого и второго порядков необходимое условие устойчивости Стодолы является и достаточным.

Для системы с $n = 3$ и уравнения с характеристическим полиномом $p^3 + a_2 p^2 + a_1 p + a_0$ условия асимптотической устойчивости, полученные выше, дают следующие удобные для практики условия устойчивости Вышнеградского для системы третьего порядка: $a_0, a_1, a_2, a_3 > 0, a_1 a_2 > a_0$.

Для практических вычислений в случае систем высоких порядков удобнее использовать следующий критерий устойчивости Рауса.

Теорема 4 (критерий Рауса). *Для того, чтобы корни полинома $\sum_{i=0}^n a_i p^i$, $a_n > 0$ лежали в левой полуплоскости, необходимо и достаточно, чтобы коэффициенты первого столбца таблицы Рауса были положительны.*

Таблица 4.1

	$c_{11} = a_n$	$c_{12} = a_{n-2}$	$c_{13} = a_{n-4}$...
	$c_{21} = a_{n-1}$	$c_{22} = a_{n-3}$	$c_{23} = a_{n-5}$...
$r_3 = \frac{c_{11}}{c_{21}}$	$c_{31} = c_{12} - r_3 c_{22}$	$c_{32} = c_{13} - r_3 c_{23}$	$c_{33} = c_{14} - r_3 c_{24}$...
$r_4 = \frac{c_{21}}{c_{31}}$	$c_{41} = c_{22} - r_4 c_{32}$	$c_{42} = c_{23} - r_4 c_{33}$	$c_{43} = c_{24} - r_4 c_{34}$...
...

Таблица Рауса строится так. В первой строке записываются коэффициенты характеристического полинома, начиная с a_n , через один, а во второй — коэффициенты характеристического полинома, начиная с a_{n-1} , тоже через один. Элементы первой строки обозначаются c_{1i} , а второй — c_{2i} . Любой из остальных элементов c_{ij} находится по формуле

$$c_{ij} = c_{i-2,j+1} - r_i c_{i-1,j+1}, \quad (4.6)$$

где $r_i = \frac{c_{i-2,1}}{c_{i-1,1}}$ — постоянное число для каждой строки. Число строк таблицы Рауса равно $n + 1$ (см. табл. 4.1).

Пример. Пусть исследуется устойчивость системы пятого порядка с характеристическим полиномом

$$p^5 + 3p^4 + 6p^3 + 9p^2 + 5p + 3.$$

Составляем таблицу Рауса по приведённым выше правилам (табл. 4.2).

Таблица 4.2

	1	6	5
	3	9	3
$r_3 = \frac{1}{3}$	3	4	0
$r_4 = 1$	5	3	0
$r_5 = \frac{3}{5}$	$2\frac{1}{5}$	0	0
	3	0	0

Элементы первого столбца таблицы Рауса положительны, следовательно, система асимптотически устойчива. Нетрудно убедиться, что отличающийся одним коэффициентом характеристический полином

$$p^5 + 3p^4 + 6p^3 + 9p^2 + 2p + 3$$

соответствует неустойчивой системе.

Использование критерия Гурвица в данном случае системы пятого порядка привело бы к матрице

$$\begin{pmatrix} 5 & 3 & 0 & 0 & 0 \\ 6 & 9 & 5 & 3 & 0 \\ 1 & 3 & 6 & 9 & 5 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

и было бы более сложным.

Критерии устойчивости Гурвица и Рауса налагают одни и те же ограничения на коэффициенты характеристического полинома, что легко проверить для $n = 1, 2, 3$. Поэтому эти ограничения часто называют условиями Рауса-Гурвица.

В том случае, если рассматриваемая система не является устойчивой, т.е. имеет корни в правой полуплоскости, то критерии Гурвица и Рауса позволяют определить число таких корней. При использовании критерия Гурвица число корней с положительной действительной частью равно числу перемен знака в любой из последовательностей

$$\Delta_0, \Delta_1, \frac{\Delta_2}{\Delta_1}, \frac{\Delta_3}{\Delta_2}, \dots, \frac{\Delta_n}{\Delta_{n-1}}$$

или

$$\Delta_0, \Delta_1, \Delta_1\Delta_2, \Delta_2\Delta_3, \dots, \Delta_{n-2}\Delta_{n-1}, a_n.$$

При использовании критерия Рауса число корней с положительной действительной частью равно числу перемен знака в первом столбце таблицы Рауса.

§5. Анализ качества систем управления

Свойство асимптотической устойчивости является необходимым свойством работоспособной системы управления. Однако, если система управления асимптотически устойчива, то это ещё не означает, что она будет работать в соответствии с техническими требованиями, предъявляемыми к ней, такими, как скорость её реакции — быстродействие, склонность системы к колебаниям при отработке задающих воздействий, точность отработки задающих воздействий.

Очевидно, что качество системы управления определяется её реакциями на всевозможные виды воздействий. Однако, поскольку провести анализ реакций системы на все виды воздействий невозможно, то анализ качества проводится для некоторых видов входных сигналов, представляющих собой либо типичные для данной системы, либо наиболее опасные для неё воздействия.

К последним относится ступенчатое воздействие, имеющее физический смысл внезапного включения постоянного задающего воздействия или возмущения (нагрузки). Реакция системы управления на ступенчатое воздействие является важной характеристикой системы. Она обозначается $h(t)$ и называется *переходной характеристикой* системы управления. Переходная характеристика линейной системы может быть найдена по её передаточной функции.

Поскольку линейные операторы обладают свойством однородности, то для линейных систем используется единичная ступенчатая функция

$$1(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0, \end{cases} \quad (5.1)$$

график которой показан на рис. 5.1 а). Изображение единичной ступенчатой функции по Лапласу есть

$$\int_0^{+\infty} 1(t)e^{-pt} dt = \int_0^{+\infty} e^{-pt} dt = \frac{1}{p}. \quad (5.2)$$

Пусть передаточная функция линейной системы имеет вид

$$G(p) = \frac{kM(p)}{N(p)}, \quad (5.3)$$

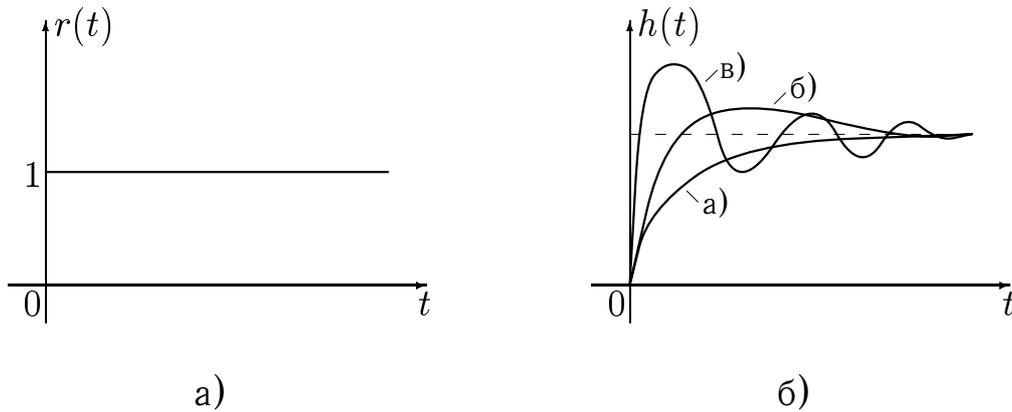


Рис. 5.1.

где $M(p)$ и $N(p)$ — полиномы степени m и n соответственно, $m < n$. Тогда изображение переходной характеристики задаётся формулой

$$H(p) = G(p) \frac{1}{p}, \quad (5.4)$$

откуда она сама может быть найдена по формуле разложения (см. §1 этой главы):

$$h(t) = \frac{kM(0)}{N(0)} + \sum_{i=1}^n \frac{kM(p_i)}{p_i N'(p_i)} e^{p_i t}. \quad (5.5)$$

Здесь сумма вычисляется по всем корням p_i характеристического полинома $N(p)$, причём эти корни предполагаются простыми и отличными от нуля.

Типичные виды переходных характеристик (для системы второго порядка) показаны на рис. (5.1) б). Кривые а) и б) соответствуют апериодическим процессам, причём кривая а) — монотонному апериодическому процессу (производная переходной характеристики не меняет знака). Кривая в) соответствует случаю колебательного процесса (производная переходной характеристики периодически меняет знак).

По переходной характеристике системы определяют следующие показатели качества, называемые показателями качества переходного процесса (рис. 5.2):

- а) t_p — время переходного процесса, или время регулирования — момент времени, начиная с которого переходная характеристика удовлетворяет неравенству

$$|h(t) - h_\infty| \leq \Delta,$$

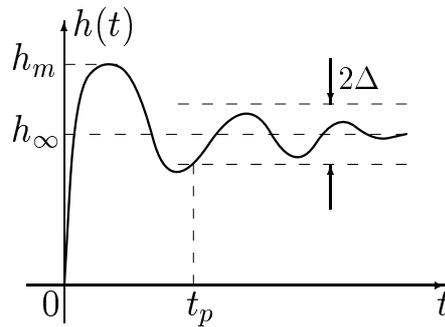


Рис. 5.2.

где Δ — заданная величина, $h_\infty = h(+\infty)$ — установившееся значение переходной характеристики (при $t \rightarrow +\infty$).

б) перерегулирование, или выброс, σ , определяемый по формуле

$$\sigma = \frac{h_m - h_\infty}{h_\infty} \cdot 100\%.$$

Величина выброса при проектировании выбирается с учётом специфики задачи. Если выброс вообще допустим, то обычно выбирают $\sigma = 10 - 30\%$. Для монотонных процессов $\sigma = 0$.

в) число колебаний за время переходного процесса. При проектировании это число обычно составляет 1–2.

Время переходного процесса определяет быстродействие системы, а перерегулирование и число колебаний за время переходного процесса — её склонность к колебаниям. Следует отметить, что колебательные переходные процессы обычно обеспечивают большее быстродействие системы.

Общих формул, связывающих показатели качества переходного процесса с коэффициентами передаточной функции системы не существует и они находятся после построения переходной характеристики по формуле (5.5).

При синтезе системы управления по заданным показателям качества переходного процесса эти показатели назначаются и синтез осуществляется так, чтобы показатели качества синтезированной системы не превосходили назначенных. Требования к системе при этом удобно представить графически так, как показано на рис. 5.3. Целью синтеза при этом является обеспечение того, чтобы переходная характеристика системы не выходила за пределы области, показанной на рисунке штриховкой.

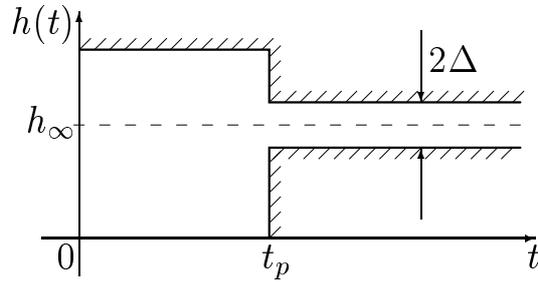


Рис. 5.3.

Точность обработки системой задающих воздействий определяется в установившемся (статическом) режиме, т.е. при $t \rightarrow +\infty$. Рассмотрим замкнутую систему управления, структурная схема которой показана на рис. 5.4. Здесь K — коэффициент усиления регулятора, а $G(p)$ — передаточная

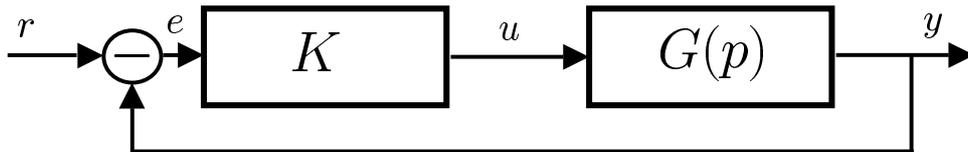


Рис. 5.4.

функция объекта регулирования, записанная в стандартной форме (5.3). Передаточная функция замкнутой системы по сигналу ошибки $W_e(p) = \frac{E(p)}{R(p)}$ задаётся формулой

$$W_e(p) = \frac{1}{1 + KG(p)}, \quad (5.6)$$

откуда изображение сигнала ошибки имеет следующий вид

$$E(p) = \frac{1}{1 + KG(p)} R(p). \quad (5.7)$$

Согласно теореме об установившемся значении преобразования Лапласа

$$f(+\infty) = \lim_{p \rightarrow 0} pF(p).$$

Таким образом, установившемуся значению сигнала ошибки при $t \rightarrow +\infty$ соответствует значение $p \rightarrow 0$. Поэтому для нахождения установившегося

значения сигнала ошибки $e(t)$ разложим $W_e(p)$ в степенной ряд в окрестности точки $p = 0$. Тогда из (5.7) получим

$$E(p) = \sum_{i=0}^{+\infty} \frac{c_i}{i!} p^i R(p). \quad (5.8)$$

Здесь обозначено $c_0 = W_e(0)$, $c_1 = \left. \frac{dW}{dp} \right|_{p=0}$, \dots , $c_i = \left. \frac{d^i W}{dp^i} \right|_{p=0}$. С учётом того, что $p^i R(p) \doteq r^{(i)}(t)$ из (5.8), переходя к оригиналам, окончательно получаем следующее выражение для установившейся ошибки системы с обратной связью:

$$e(t) = \sum_{i=0}^{+\infty} \frac{c_i}{i!} r^{(i)}(t). \quad (5.9)$$

Коэффициенты c_i в этих формулах называются коэффициентами ошибки. Они определяют точность системы с обратной связью.

Для системы регулирования, целью которой является поддержание заданного, постоянного режима работы $r(t) = a_0 1(t)$,

$$c_0 = \left. \frac{N(p)}{N(p) + KkM(p)} \right|_{p=0} = \frac{1}{1 + Kk}. \quad (5.10)$$

Величина установившейся ошибки зависит от общего коэффициента усиления разомкнутой цепи. Если полином $N(p)$ имеет нулевой корень, то $c_0 = 0$ и системы такого типа называются астатическими, в отличие от систем, имеющих ненулевую позиционную, или статическую, ошибку (5.10), называемых статическими.

Для систем слежения используется типовое воздействие

$$r(t) = (a_0 + a_1 t) 1(t).$$

Формула (5.9) даёт в этом случае такое установившееся значение ошибки:

$$e(t) = c_0(a_0 + a_1)t + c_1 a_1. \quad (5.11)$$

Здесь коэффициент c_1 задаёт скоростную ошибку системы. Очевидно, что если $c_0 \neq 0$, т.е. если система — статическая, то $e(t) \rightarrow \infty$ при $t \rightarrow +\infty$. Поскольку это недопустимо, то следящие системы проектируются как астатические.

Передаточную функцию астатической системы можно записать в виде:

$$G(p) = \frac{k}{p} \tilde{G}(p), \quad (5.12)$$

где $\tilde{G}(p) \Big|_{p=0} = 1$. Для коэффициента скоростной ошибки при этом получим

$$c_1 = \frac{d}{dp} \left(\frac{p}{p + KkG(p)} \right) \Big|_{p=0} = \frac{p + Kk\tilde{G}(p) - pKk\tilde{G}'(p)}{(p + Kk\tilde{G}(p))^2} \Big|_{p=0} = \frac{1}{Kk}. \quad (5.13)$$

Чем больше коэффициент усиления разомкнутой цепи Kk , тем меньше скоростная ошибка. Поскольку сколь угодно большим этот коэффициент быть не может, так как это приводит к потере устойчивости, то при проектировании задают величину δ_0 , называемую добротностью, и выбирают коэффициент усиления регулятора K из условия $Kk > \delta_0$.

Передаточные функции систем являются дробно-рациональными, и отыскивать через их производные коэффициенты ошибки неудобно. Эти коэффициенты, однако, могут быть найдены более простым способом.

Используя формулы (5.3), (5.6) и (5.8), можно записать

$$\frac{N(p)}{N(p) + KkM(p)} = \sum_{i=0}^{+\infty} \frac{c_i}{i!} p^i, \quad (5.14)$$

и коэффициенты можно найти путём деления полинома в числителе на полином в знаменателе левой части этой формулы, расположив эти полиномы по возрастанию степеней.

Пример. Пусть

$$H(p) = \frac{1}{p(T_1p + 1)(T_2p + 1)}. \quad (5.15)$$

Тогда по формуле (5.6) получим передаточную функцию по ошибке

$$W_e(p) = \frac{p(T_1p + 1)(T_2p + 1)}{p(T_1p + 1)(T_2p + 1) + K}. \quad (5.16)$$

Деление числителя на знаменатель (углом [6]) даёт

$$W_e(p) = \frac{1}{K}p + \left(\frac{T_1 + T_2}{K} - \frac{1}{K^2} \right) p^2 + \dots \quad (5.17)$$

Из сравнения результата деления с формулой (5.14) следует, что $c_0 = 0$, т.е. система астатическая, $c_1 = \frac{1}{K}$ и $c_2 = \left(\frac{T_1 + T_2}{K} - \frac{1}{K^2} \right)$.

§6. Частотные характеристики, типовые звенья и критерий устойчивости Найквиста

Частотными называются такие методы анализа и синтеза систем управления, в которых используется аппарат частотных характеристик. Пусть $G(p)$ — передаточная функция линейной стационарной системы.

Определение 1. Функция

$$G(j\omega) = G(p)|_{p=j\omega} \quad (6.1)$$

называется *частотной передаточной функцией*.

Частотная передаточная функция представляет собой преобразование Фурье от весовой функции $h(t)$, т.е.

$$G(j\omega) = \int_{\mathbb{R}} g(\tau) e^{-j\omega\tau} d\tau, \quad \omega \in \mathbb{R}. \quad (6.2)$$

Частотную передаточную функцию можно представить в следующем виде:

$$G(j\omega) = U(\omega) + jV(\omega), \quad (6.3)$$

где функция $U(\omega)$ называется вещественной частотной характеристикой, а функция $V(\omega)$ — мнимой частотной характеристикой. Поскольку $h(t)$ — действительная функция, то сопряжённая к $G(j\omega)$ функция $G^*(j\omega)$ может быть получена заменой знака у экспоненты в (6.2), т.е.

$$G^*(j\omega) = \int_{\mathbb{R}} g(\tau) e^{j\omega\tau} d\tau. \quad (6.4)$$

Отсюда следует, что $G^*(j\omega) = G(-j\omega)$ и, исходя из этого равенства, что $U(-\omega) = U(\omega)$, $V(-\omega) = -V(\omega)$. Таким образом, график частотной передаточной функции в координатах $U(\omega)$, $jV(\omega)$ при изменении ω от $-\infty$ до $+\infty$, называемый *амплитудно-фазовой частотной характеристикой* — АФЧХ или АФХ, для $\omega < 0$ может быть получен из графика для $\omega > 0$ зеркальным отображением относительно действительной оси.

Частотную передаточную функцию удобно также представить в следующей экспоненциальной форме

$$G(j\omega) = A(\omega) e^{j\varphi(\omega)}, \quad (6.5)$$

где функция

$$A(\omega) = |G(\omega)| \quad (6.6)$$

называется *амплитудной частотной характеристикой* — АЧХ, а функция

$$\varphi(\omega) = \arg G(j\omega) \quad (6.7)$$

называется *фазовой частотной характеристикой* — ФЧХ. Из (6.6) следует, что

$$A(\omega) = |G(j\omega)| = \sqrt{G(j\omega)G^*(j\omega)} = \sqrt{G(j\omega)G(-j\omega)}$$

и, значит, $A(\omega) = A(-\omega)$. Далее, можно записать

$$\begin{aligned} G^*(j\omega) &= A(\omega)e^{-j\varphi(\omega)}, \\ G(-j\omega) &= A(\omega)e^{j\varphi(-\omega)}. \end{aligned}$$

Т.к. $G^*(j\omega) = G(-j\omega)$, то отсюда вытекает, что $\varphi(-\omega) = -\varphi(\omega)$. Таким образом, графики АЧХ и ФЧХ, как и графики вещественной и мнимой частотных характеристик и АФЧХ строятся только для $\omega \geq 0$. Примеры таких графиков приведены в табл. 6.1 ниже.

Частотная передаточная функция для систем, задаваемых обыкновенными линейными дифференциальными уравнениями является, как это следует из (2.15) и (6.1) дробно-рациональной функцией

$$G(j\omega) = \frac{\sum_{i=0}^m b_i(j\omega)^i}{\sum_{i=0}^n a_i(j\omega)^i}. \quad (6.8)$$

Отсюда следует, что в рассматриваемом случае

$$\begin{aligned} A(\omega) &= \frac{|\sum_{i=0}^m b_i(j\omega)^i|}{|\sum_{i=0}^n a_i(j\omega)^i|}, \\ \varphi(\omega) &= \arg\left(\sum_{i=0}^m b_i(j\omega)^i\right) - \arg\left(\sum_{i=0}^n a_i(j\omega)^i\right). \end{aligned} \quad (6.9)$$

Воспользуемся теоремой о разложении многочленов на множители и запишем формулу (6.8) в виде

$$G(j\omega) = \frac{b_m(j\omega - \sigma_1) \dots (j\omega - \sigma_m)}{a_n(j\omega - \lambda_1) \dots (j\omega - \lambda_n)}. \quad (6.10)$$

Здесь корни полиномов в числителе и знаменателе σ_i и λ_i могут быть вещественными (в частности, нулевыми) и комплексными, образующими комплексно-сопряжённые пары. Множители $(j\omega - \sigma_k)$ и $(j\omega - \lambda_k)$ входят в числитель и знаменатель формулы (6.10) столько раз, какова кратность корней σ_k и λ_k . Пусть числитель дроби (6.10) имеет q нулевых корней, p действительных ненулевых корней и $2r$ комплексно-сопряжённых корней, $q + p + 2r = m$. Пусть, далее, знаменатель дроби (6.10) имеет d нулевых корней, s действительных ненулевых корней и $2w$ комплексно-сопряжённых корней, $d + s + 2w = n$. Объединив сомножители с комплексно-сопряжёнными корнями и вынося за скобки постоянные члены в каждом сомножителе, запишем частотную передаточную функцию в так называемой *стандартной форме*:

$$G(j\omega) = \frac{k(j\omega)^q \prod_{i=1}^p (T_i j\omega + 1) \prod_{i=1}^r (T_i^2(j\omega)^2 + 2\xi_i T_i(j\omega) + 1)}{(j\omega)^d \prod_{i=1}^s (T_i j\omega + 1) \prod_{i=1}^w (T_i^2(j\omega)^2 + 2\xi_i T_i(j\omega) + 1)}. \quad (6.11)$$

Множители, входящие в числитель и знаменатель формулы (6.11), задают передаточные функции систем, называемых *типовыми (элементарными) звеньями*. Коэффициент k называется коэффициентом усиления, параметры T_i — постоянными времени типовых звеньев, параметры ξ_i — коэффициентами демпфирования (см. ниже). Формулой

$$L(\omega) = 20 \lg A(\omega) \quad (6.12)$$

вводится *логарифмическая амплитудная частотная характеристика* — ЛАЧХ (или ЛАХ), называемая также *диаграммой Боде*, единица измерения которой — децибелл. При графическом построении ЛАЧХ по оси абсцисс откладываются значения частоты в логарифмическом масштабе — $\lg \omega$. За единицу масштаба здесь принят интервал, называемый *декадой*, на котором частота изменяется в десять раз. Зависимость фазовой частотной характеристики от логарифма частоты $\lg \omega$ называют *логарифмической фазовой частотной характеристикой* — ЛФЧХ.

Для частотной характеристики в стандартной форме (6.11), учитывая формулы (6.9), имеем

$$L(\omega) = 20(\lg k + q \lg \omega + \sum_{i=1}^p \lg |T_i j\omega + 1| + \sum_{i=1}^r \lg |T_i^2(j\omega)^2 + 2\xi_i T_i j\omega + 1| - d \lg \omega - \sum_{i=1}^s \lg |T_i(j\omega) + 1| - \sum_{i=1}^w \lg |T_i^2(j\omega)^2 + 2\xi_i T_i j\omega + 1|), \quad (6.13)$$

$$\varphi(\omega) = \arg(j\omega)^q + \sum_{i=1}^p \arg(T_i j\omega + 1) + \sum_{i=1}^r \arg(T_i^2(j\omega)^2 + 2\xi_i j\omega + 1) - \arg(j\omega)^d - \sum_{i=1}^s \arg(T_i j\omega + 1) - \sum_{i=1}^w \arg(T_i^2(j\omega)^2 + 2\xi_i j\omega + 1). \quad (6.14)$$

Таким образом, построение ЛАЧХ и ЛФЧХ сводится к суммированию и вычитанию частотных характеристик типовых звеньев, что существенно упрощает все расчёты. Это обстоятельство определило широкое использование логарифмических частотных характеристик при проектировании систем управления.

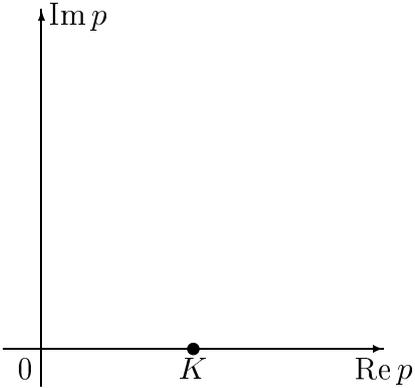
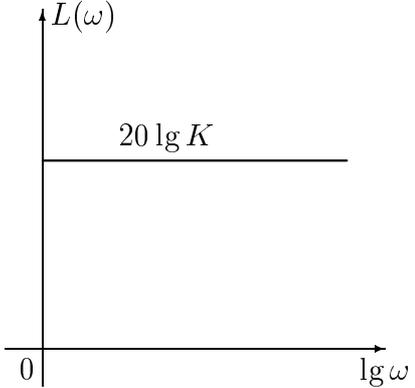
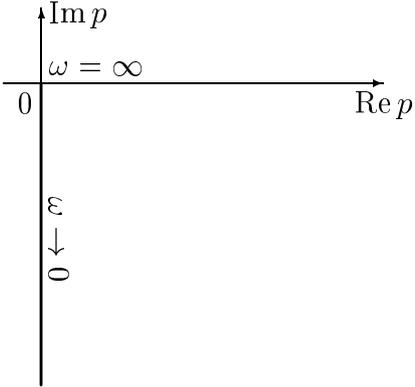
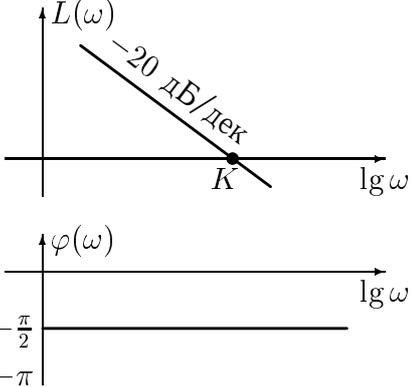
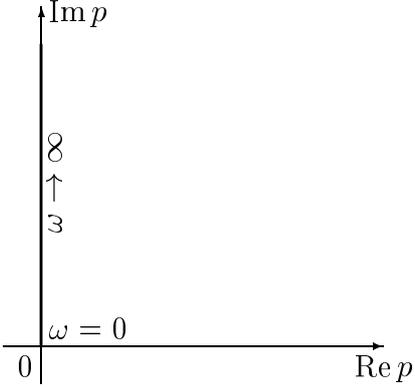
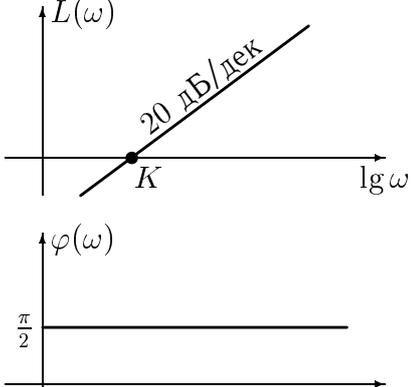
Рассмотрим частотные характеристики типовых звеньев.

Для типового звена, задаваемого уравнением $y(t) = Kr(t)$, $K > 0$ и называемого позиционным звеном, или безынерционным усилителем, передаточная функция $G(p) = K$. Частотная передаточная функция $G(j\omega) = K$ и $U(\omega) = K$, $V(\omega) = 0$. АФЧХ звена представляет собой точку (на положительной части действительной оси) с координатами $(K; 0)$. Далее, $A(\omega) = 0$, $\varphi(\omega) = 0$ и ЛАЧХ звена $L(\omega) = 20 \lg K$ определяет уравнение прямой, параллельной оси абсцисс. Частотные характеристики этого звена, как и звеньев, рассматриваемых ниже, приведены в табл. 6.1.

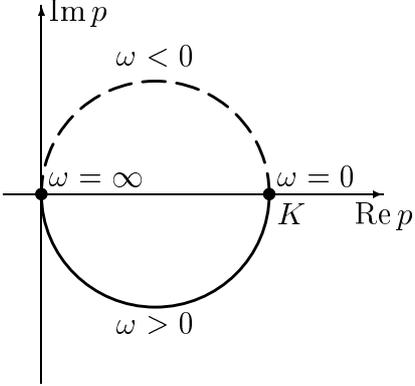
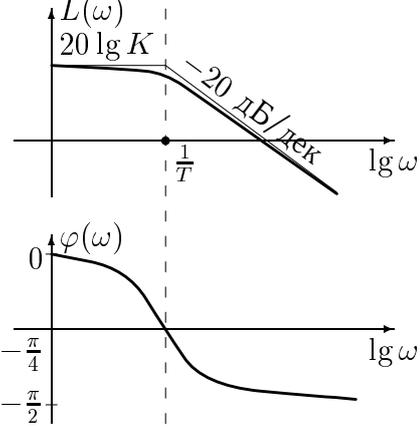
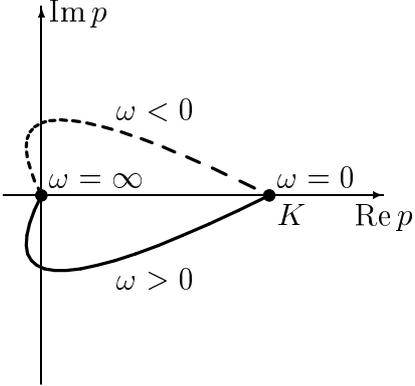
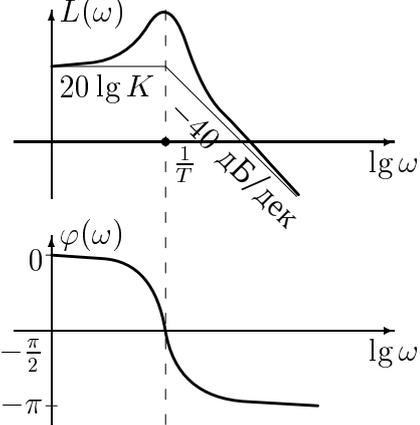
Так как $G(p) = K$, то импульсная переходная функция представляет собой δ -функцию, т.е. $g(t) = \delta(t)$.

Для интегрирующего звена порядка $d = 1, 2, \dots$, задаваемого уравнением $y^{(d)}(t) = Kr(t)$, передаточная функция имеет вид $G(p) = \frac{K}{p^d}$. Частотная передаточная функция задаётся формулой $G(j\omega) = \frac{K}{(j\omega)^d}$, $A(\omega) = \frac{K}{\omega^d}$, $\varphi(\omega) = -d\frac{\pi}{2}$, $L(\omega) = 20 \lg K - 20d \lg \omega$. Полагая в последнем уравнении $L(\omega) = 0$, находим частоту среза $\omega_{\text{ср}}$, т.е. точку пересечения ЛАЧХ звена с осью абсцисс: $0 = 20 \lg K - 20d \lg \omega_{\text{ср}}$, откуда $\omega_{\text{ср}} = \sqrt[d]{K}$.

Таблица 6.1

Тип звена	АФЧХ	ЛАЧХ и ЛФЧХ
1	2	3
$G(j\omega) = K$		
$G(j\omega) = \frac{K}{j\omega}$		
$G(j\omega) = K \cdot j\omega$		

Продолжение таблицы 6.1

1	2	3
$G(j\omega) = \frac{K}{Tj\omega + 1}$		
$G(j\omega) = \frac{K}{T^2(j\omega)^2 + 2\xi T(j\omega) + 1}$		

Для $d = 1$, т.е. интегрирующего звена первого порядка, имеем частотную передаточную функцию $H(j\omega) = \frac{K}{j\omega}$, $U(\omega) = 0$, $V(\omega) = -\frac{K}{\omega}$ и АФЧХ звена совпадает с отрицательной частью мнимой оси. Далее: $A(\omega) = \frac{K}{\omega}$, $\varphi(\omega) = -\frac{\pi}{2}$, $L(\omega) = 20 \lg K - 20 \lg \omega$ и $\omega_{\text{ср}} = K$. Отсюда следует, что ЛАЧХ звена представляет собой прямую, проходящую через точку $\omega = K$ на оси абсцисс с наклоном -20 дБ/дек. При этом $g(t) = 1(t)$.

Для дифференцирующего звена порядка q , задаваемого уравнением $y(t) = Kr^{(q)}(t)$, передаточная функция такова: $G(p) = Kp^q$, а частотная передаточная функция $G(j\omega) = K(j\omega)^q$, причём $A(\omega) = \omega^q$, $\varphi(\omega) = q\frac{\pi}{2}$ и $L(\omega) = 20 \lg K + q20 \lg \omega$. Из последнего уравнения, полагая $L(\omega) = 0$, находим точку пересечения ЛАЧХ с осью абсцисс: $0 = 20 \lg K + q20 \lg \omega_{\text{ср}}$, откуда $\omega_{\text{ср}} = \sqrt[q]{K}$.

Для $q = 1$, т.е. для дифференцирующего звена первого порядка, имеем $G(j\omega) = Kj\omega$, $U(\omega) = 0$, $V(\omega) = K\omega$ и АФЧХ звена совпадает

с положительной частью мнимой оси. При этом $A(\omega) = K\omega$, $\varphi(\omega) = \frac{\pi}{2}$, $L(\omega) = 20 \lg K + 20 \lg \omega$ и $\omega_{\text{ср}} = \frac{1}{K}$, откуда следует, что ЛАЧХ звена представляет собой прямую, проходящую через точку $\omega_{\text{ср}} = \frac{1}{K}$ на оси абсцисс с наклоном 20 дБ/дек. Импульсная переходная функция звена представляет собой производную δ -функции, т.е. $g(t) = K\delta'(t)$.

Рассмотрим апериодическое звено, задаваемое уравнением

$$Ty'(t) + y(t) = Ku(t)$$

с передаточной функцией $G(p) = \frac{K}{Tp+1}$. В этом случае $H(j\omega) = \frac{K}{Tj\omega+1}$ и $U(\omega) = \frac{K}{T^2\omega^2+1}$, $V(\omega) = -\frac{KT\omega}{T^2\omega^2+1}$. Отсюда следует, что $A(\omega) = \frac{K}{\sqrt{T^2\omega^2+1}}$ и $\varphi(\omega) = -\arctg T\omega$. Найдём АФЧХ системы. Для этого используем уравнение $U^2(\omega) + V^2(\omega) - A^2(\omega) = 0$, которое даёт

$$\frac{K^2}{(T^2\omega^2 + 1)^2} + \frac{K^2T^2\omega^2}{(T^2\omega^2 + 1)^2} - \frac{K^2}{(T^2\omega^2 + 1)^2} = 0.$$

Дополняем левую часть до полного квадрата, прибавляя и вычитая слагаемое $\left(\frac{K}{2}\right)^2$, и получаем

$$\left(\frac{K}{T^2\omega^2 + 1} - \frac{K}{2}\right)^2 + \frac{K^2T^2\omega^2}{(T^2\omega^2 + 1)^2} = \left(\frac{K}{2}\right)^2.$$

Это уравнение можно записать в виде $(U - \frac{K}{2})^2 + V^2 = (\frac{K}{2})^2$, откуда следует, что при $\omega > 0$ АФЧХ апериодического звена представляет собой полуокружность радиуса $\frac{K}{2}$ с центром в точке $\frac{K}{2}$.

ЛАЧХ звена определяется уравнением $L(\omega) = 20 \lg K - 20 \lg \sqrt{T^2\omega^2 + 1}$ и имеет вид, показанный в табл. 6.1. Для практических расчётов эта характеристика заменяется кусочно-линейной аппроксимацией, т.е. асимптотической ЛАЧХ, задаваемой формулой

$$L(\omega) = \begin{cases} 20 \lg K, & \text{при } \omega \leq \frac{1}{T}, \\ 20 \lg K - 20 \lg T\omega, & \text{при } \omega > \frac{1}{T}. \end{cases}$$

Уравнение для асимптотической ЛАЧХ получается, если под знаком корня при $\omega < \frac{1}{T}$ пренебречь первым слагаемым, а при $\omega > \frac{1}{T}$ — вторым, т.е. единицей. Частота $\omega = \frac{1}{T}$ называется *сопрягающей частотой*. Отклонение асимптотической ЛАЧХ от точной максимально на этой частоте и равно

$$\Delta = 20 \lg K - (20 \lg K - 20 \lg \sqrt{2}) \approx 3,03 \text{ дБ}.$$

При инженерных расчётах такой величиной можно пренебречь. Частота среза из асимптотической ЛАЧХ получается $\omega_{\text{ср}} = \frac{K}{T}$ и наклон частотной характеристики в окрестности частоты среза равен -20 дБ/дек.

Передаточная функция аperiodического звена есть $G(p) = \frac{K}{Tp+1}$, и из свойств преобразования Лапласа следует, что импульсная переходная функция звена есть $g(t) = \frac{K}{T}e^{-t/T}$. Поскольку импульсная переходная функция определяет реакцию системы на мгновенный импульс (δ -функцию), то отсюда следует, что реакция данного звена имеет аperiodический характер, чем и определяется название звена. Если считать, что процесс реакции на импульс, называемый переходным процессом, заканчивается при достижении 5% уровня (т.е., величины 0,05 от начального значения), то можно из уравнения $e^{-t/T} = 0,05$ приближённо оценить время переходного процесса звена величиной $3T$. Чем меньше постоянная времени, тем быстрее заканчивается переходный процесс, что и обуславливает название этого параметра. Величина T определяет также «память системы», т.е. интервал времени, на котором $g(t)$ существенно отлична от нуля.

Форсирующее звено задаётся уравнением $y(t) = K(Tr'(t) + r(t))$, что соответствует передаточной функции $G(p) = K(Tp + 1)$. При этом $G(j\omega) = K(Tj\omega + 1)$, $U(\omega) = K$, $V(\omega) = KT\omega$ и АФЧХ звена представляет собой полупрямую, параллельную положительной части мнимой оси и отстоящую от неё на величину K . Далее, $A(\omega) = K\sqrt{T^2\omega^2 + 1}$, $\varphi(\omega) = \text{arctg } T\omega$ и $L(\omega) = 20 \lg K + 20 \lg \sqrt{T^2\omega^2 + 1}$. Асимптотическая ЛАЧХ определяется формулой

$$L(\omega) = \begin{cases} 20 \lg K & \text{при } \omega \leq \frac{1}{T}, \\ 20 \lg K + 20 \lg T\omega & \text{при } \omega > \frac{1}{T}, \end{cases}$$

а частота среза, получаемая из неё, есть $\omega_{\text{ср}} = \frac{1}{KT}$. Импульсная переходная функция звена такова: $g(t) = K(\delta'(t) + \delta(t))$.

Рассмотрим звено, задаваемое дифференциальным уравнением второго порядка $T^2y''(t) + 2\xi Ty'(t) + y(t) = Ku(t)$. Теперь $G(p) = \frac{K}{T^2p^2 + 2\xi Tp + 1}$ и характеристическое уравнение $T^2p^2 + 2\xi Tp + 1 = 0$ имеет два комплексно-сопряжённых корня $p_{1,2} = -\frac{\xi}{T} \pm j\frac{\sqrt{1-\xi^2}}{T}$, если ξ удовлетворяет условию $\xi \in (0; 1)$. При этом импульсная переходная функция имеет следующий вид:

$$g(t) = \frac{K}{T\sqrt{1-\xi^2}} e^{-\frac{\xi}{T}t} \sin \frac{\sqrt{1-\xi^2}}{T}t,$$

что следует из формулы вычетов и формулы Эйлера $\sin \varphi = \frac{1}{2j}(e^{j\varphi} - e^{-j\varphi})$. В самом деле, запишем передаточную функцию в виде

$$G(p) = \frac{K}{T^2(p^2 + \frac{2\xi}{T}p + \frac{1}{T^2})} = \frac{K}{T^2(p - p_1)(p - p_2)} = \frac{K}{T^2} \left(\frac{A}{p - p_1} + \frac{B}{p - p_2} \right),$$

где $A = \frac{T}{2j\sqrt{1-\xi^2}}$ и $B = -\frac{T}{2j\sqrt{1-\xi^2}}$ — вычеты. Отсюда следует приведённая формула для импульсной переходной характеристики, т.к. после обратного преобразования $g(t) = \frac{K}{T^2}(Ae^{p_1 t} + Be^{p_2 t})$. Поскольку реакция на импульс получается колебательной, то данное типовое звено называется колебательным. Время переходного процесса и «память» системы можно оценить, используя 5% уровень, величиной $3\frac{\xi}{T}$. Параметр ξ , который определяет величину колебаний в системе, называется коэффициентом демпфирования (затухания). При $\xi = 0$ звено имеет два чисто мнимых корня и становится асимптотически неустойчивым. При $\xi > 1$ характеристическое уравнение имеет два действительных корня, импульсная переходная функция содержит две экспоненциальные аperiodические составляющие. В этом случае систему второго порядка можно представить в виде последовательного соединения двух аperiodических звеньев.

Для колебательного звена $H(j\omega) = \frac{K}{-T^2\omega^2 + j2\xi T\omega + 1}$ и

$$U(\omega) = \frac{K(1 - T^2\omega^2)}{(1 - T^2\omega^2)^2 + 4T^2\xi^2\omega^2}, \quad V(\omega) = \frac{2T\xi K\omega}{(1 - T^2\omega^2)^2 + 4T^2\xi^2\omega^2}.$$

АФЧХ колебательного звена показана в табл. 6.1. Из вещественной и мнимой частотных характеристик находится АЧХ $A(\omega) = \frac{K}{\sqrt{(1-T^2\omega^2)^2 + 4T^2\xi^2\omega^2}}$ и

ФЧХ $\varphi(\omega) = -\arctg \frac{2T\xi\omega}{1-T^2\omega^2}$ звена.

По логарифмической частотной характеристике

$$L(\omega) = 20 \lg K - 20 \lg \sqrt{(1 - T^2\omega^2)^2 + 4T^2\xi^2\omega^2} \quad (6.15)$$

вводится асимптотическая ЛАЧХ

$$L(\omega) = \begin{cases} 20 \lg K, & \omega \leq \frac{1}{T}, \\ 20 \lg K - 40 \lg T\omega, & \omega > \frac{1}{T}, \end{cases} \quad (6.16)$$

где в первом случае получим (6.16), если под знаком корня в (6.15) пренебречь всеми слагаемыми, кроме единицы, а во втором случае пренебречь единицей и сохранить только слагаемое, содержащее четвёртую (старшую)

степень ω . Частота среза определяется по асимптотической ЛАЧХ из уравнения $0 = 20 \lg K - 40 \lg T\omega_{\text{ср}}$, откуда $\omega_{\text{ср}} = \sqrt{\frac{K}{T}}$. В окрестности частоты среза асимптотическая ЛАЧХ имеет наклон -40 дБ/дек. Асимптотические ЛАЧХ хорошо приближают истинную ЛАЧХ звена при условии, что коэффициент демпфирования $\xi > 0,5$. При $\xi < 0,5$ следует учитывать превышение истинной ЛАЧХ, что в инженерных расчётах делается введением специальных поправок.

Элементарное звено, задаваемое уравнением

$$y(t) = K(T^2 r''(t) + 2\xi T r'(t) + r(t))$$

при условии, что $\xi \in (0; 1)$, называется форсирующим звеном второго порядка. Для него $H(p) = K(p^2 + 2\xi T p + 1)$ и частотные характеристики нетрудно записать с помощью изложенного выше подхода.

В таблице 6.1 приведены графики АФЧХ, ЛАЧХ и ЛФЧХ по полученным формулам для типовых звеньев. На практике характеристики апериодического и колебательного звеньев заменяют их кусочно-линейной аппроксимацией, как показано в таблице. При этом ЛАЧХ системы, задаваемой передаточной функцией в стандартной форме (6.11) тоже будет иметь кусочно-линейный характер, что существенно упрощает все расчёты. Поэтому на практике, по крайней мере на предварительном этапе проектирования, широко используют аппарат асимптотических частотных характеристик.

Для построения асимптотической ЛАЧХ передаточная функция системы представляется в стандартной форме (6.11). Далее:

- 1) Определяются сопрягающие частоты типовых звеньев, которые располагаются в возрастающем порядке $\omega_1 = \frac{1}{T_1} < \omega_2 = \frac{1}{T_2} < \omega_3 = \frac{1}{T_3} < \dots$ и наносятся на ось абсцисс.
- 2) Левая, низкочастотная часть ЛАЧХ определяется множителем $\frac{K(j\omega)^q}{(j\omega)^d}$. Если $q = d$, т.е. если этот множитель равен просто K , то проводится прямая, параллельная оси абсцисс и отстоящая от неё на величину $20 \lg K$. Если $d = q + 1$, то проводится прямая с наклоном -20 дБ/дек через точку K . Если $d = q + 2$, то низкочастотная асимптота проводится с наклоном -40 дБ/дек через точку \sqrt{K} и вообще, при $d = q + \nu$ низкочастотная асимптота проводится через точку $\sqrt[\nu]{K}$ с наклоном $-\nu 20$ дБ/дек. Если $q = d + 1$, то низкочастотная асимптота проводится с наклоном $+20$ дБ/дек через точку $\frac{1}{K}$, а если $q = d + \nu$, то

с наклоном $+v20$ дБ/дек через точку $\frac{1}{\sqrt{K}}$. Низкочастотная асимптота определяет ЛАЧХ системы вплоть до первой сопрягающей частоты $\omega_1 = \frac{1}{T_1}$.

- 3) На первой сопрягающей частоте наклон асимптотической ЛАЧХ изменяется, причём апериодическое и форсирующее звенья изменяют наклон на ± 20 дБ/дек, а колебательное и форсирующее звено второго порядка — на ± 40 дБ/дек соответственно.
- 4) Процесс построения с изменением наклона асимптотической ЛАЧХ продолжается по всем сопрягающим частотам вплоть до последней.

Пример. Пусть дана передаточная функция системы второго порядка, записанная в стандартной форме:

$$H(j\omega) = \frac{K}{(T_1 j\omega + 1)(T_2 j\omega + 1)},$$

и пусть $\omega_1 = \frac{1}{T_1} < \omega_2 = \frac{1}{T_2}$. Асимптотическая ЛАЧХ, построенная по приведённым правилам, показана на рис. 6.1.

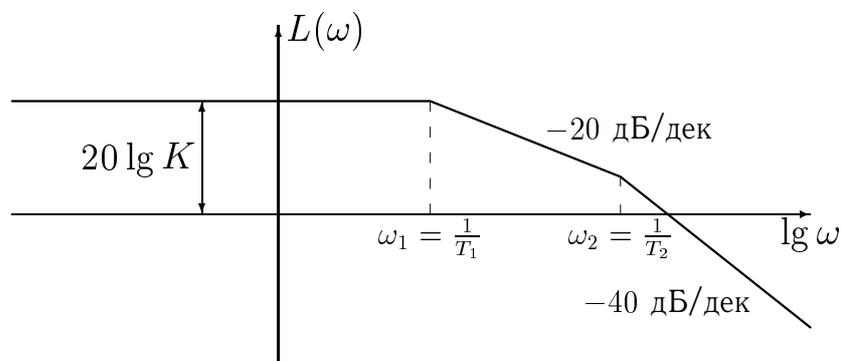


Рис. 6.1.

По построенным асимптотическим ЛАЧХ с помощью специальных графиков поправок, которые особенно важны при наличии в передаточной функции колебательных звеньев, всегда можно восстановить и точные ЛАЧХ.

Рассмотренный метод построения ЛАЧХ, очевидно, применим для последовательного соединения линейных систем, т.е. для разомкнутых систем. Для систем с обратной связью ЛАЧХ замкнутой системы может

быть построена по ЛАЧХ разомкнутой системы с помощью номограмм замыкания, называемых также диаграммами Никольса.

Ввиду своей простоты и наглядности, аппарат асимптотических ЛАЧХ широко используется в инженерных расчётах. К числу его достоинств относится также то, что с его помощью можно по экспериментально полученной амплитудной частотной характеристике построить соответствующую передаточную функцию системы. Для этого достаточно аппроксимировать экспериментальную кривую, построенную в логарифмическом масштабе, ломаной, участки которой имеют наклон, кратный 20 дБ/дек. При этом передаточная функция сразу получается в стандартной форме (6.11) со всеми нужными значениями её параметров. Ввиду своей простоты, этот метод восстановления передаточной функции системы по экспериментальным данным может оказаться предпочтительнее, чем более точный метод, рассмотренный в §2 третьей главы.

Одним из существенных преимуществ частотного подхода, рассматриваемого в данном параграфе, является простота его использования для анализа устойчивости систем с обратной связью, основанного на применении критерия устойчивости Найквиста.

Рассмотрим замкнутую систему управления, структурная схема которой показана на рис. 6.2, где $H(p) = \frac{\sum_{i=0}^m b_i p^i}{\sum_{i=0}^n a_i p^i}$, $m < n$.

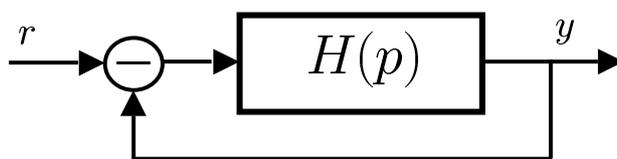


Рис. 6.2.

Критерий устойчивости Найквиста для системы с обратной связью формулируется так.

Теорема 1 (критерий Найквиста). Пусть разомкнутая система асимптотически устойчива, т.е. характеристический полином $\sum_{i=0}^n a_i p^i$ имеет корни только в левой полуплоскости. Тогда замкнутая система асимптотически устойчива тогда и только тогда, когда амплитудно-фазовая частотная характеристика разомкнутой системы не охватывает на комплексной плоскости точку $(-1, j0)$ и не проходит через неё.

На рис. 6.3 а) приведён график для случая устойчивости системы с обратной связью, а на рис. 6.3 б) — для случая неустойчивости.

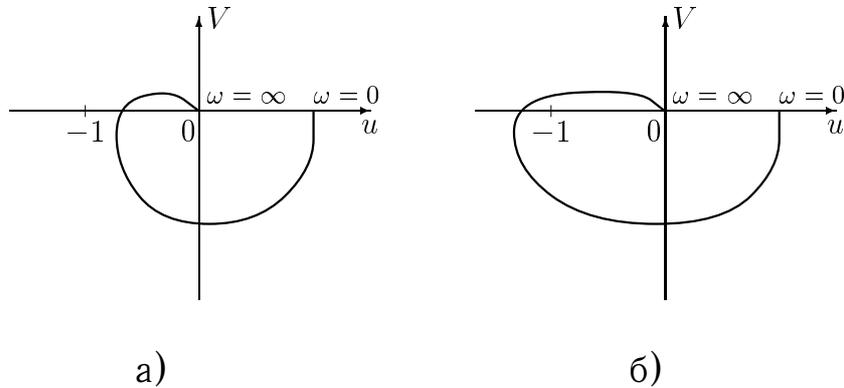


Рис. 6.3.

Из критерия Найквиста сразу следует, что интегрирующее, апериодическое и колебательное звенья являются всегда устойчивыми в замкнутом состоянии.

Критерий Найквиста имеет то преимущество, что не требует предварительного расчёта передаточной функции замкнутой системы, что требовалось бы при применении критериев Гурвица или Рауса. Кроме того, его можно применять и в случае, когда амплитудно-фазовая частотная характеристика получена экспериментально. Ввиду важности критерия Найквиста для теории систем с обратной связью, амплитудно-фазовую частотную характеристику называют также годографом Найквиста.

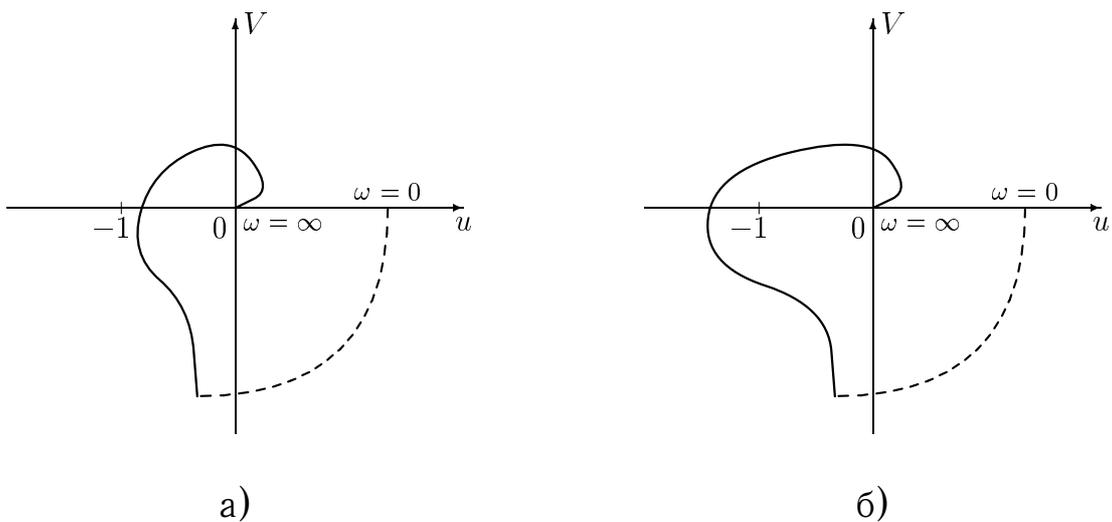


Рис. 6.4.

Критерий Найквиста обобщается на случай, когда характеристический полином имеет нулевой корень, т.е. он имеет вид $\sum_{i=1}^n a_i p^i$ ($a_0 = 0$). Для проверки устойчивости замкнутой системы по критерию Найквиста АФЧХ разомкнутой системы дополняется при $\omega \rightarrow 0$ дугой бесконечно большого радиуса. На рис. 6.4 а) показан случай, соответствующий устойчивости замкнутой системы, а на рис. 6.4 б) случай неустойчивости замкнутой системы.

Существует обобщение критерия Найквиста и на случай, когда разомкнутая система неустойчива и её характеристический полином имеет корни в правой полуплоскости.

Для логарифмических частотных характеристик условие критерия Найквиста означает, что при $\varphi = -\pi$ требуется, чтобы $L(\omega) < 0$ или чтобы $\varphi(\omega) > -\pi$ при $\omega = \omega_{\text{ср}}$. На рис. 6.5 а) и б) приведены графики логарифмических частотных характеристик устойчивой и неустойчивой в замкнутом состоянии линейных систем.

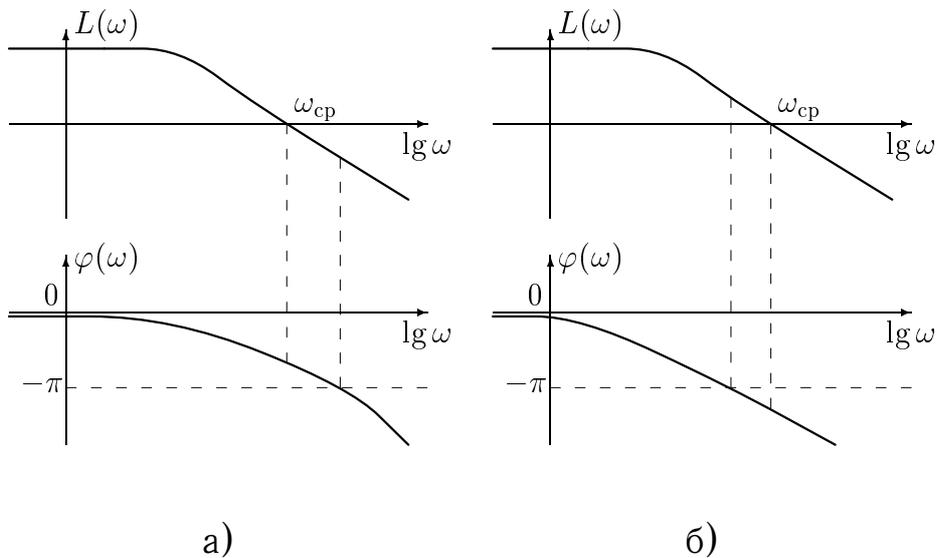


Рис. 6.5.

§7. Частотные оценки качества систем управления

Частотные оценки качества широко применяются при синтезе систем управления.

Переходная характеристика линейной стационарной системы может быть найдена по формуле (5.5), но её также можно найти и по вещественной или мнимой частотным характеристикам системы.

Используя обратное преобразование Фурье, запишем следующее выражение для переходной характеристики

$$h(t) = \frac{1}{2\pi} \int_{\mathbb{R}} G(j\omega) R(j\omega) e^{j\omega t} d\omega. \quad (7.1)$$

Здесь $R(j\omega)$ — изображение по Фурье единичной ступенчатой функции. Эта функция не является абсолютно интегрируемой и преобразование Фурье для неё определяется в терминах обобщённых функций как [25]

$$R(j\omega) = \frac{1}{j\omega} + \pi\delta(\omega), \quad (7.2)$$

где $\delta(\omega)$ — дельта-функция. С учётом свойств δ -функции, из (7.1) получаем

$$h(t) = \frac{1}{2}H(0) + \frac{1}{2\pi} \int_{\mathbb{R}} \frac{G(j\omega)}{j\omega} e^{j\omega t} d\omega. \quad (7.3)$$

Так как

$$G(j\omega) = P(\omega) + jQ(\omega), \quad (7.4)$$

где $P(\omega)$ — вещественная, а $Q(\omega)$ — комплексная частотные характеристики, то с учётом того, что мнимая частотная характеристика — нечётная функция и $\text{Im } G(0) = 0$, (7.3) записывается так:

$$h(t) = \frac{1}{2}P(0) + \frac{1}{2\pi} \int_{\mathbb{R}} \frac{G(j\omega)}{j\omega} e^{j\omega t} d\omega. \quad (7.5)$$

Выразим интеграл в этой формуле через вещественную и мнимую частотные характеристики. Используя формулу (7.4) для подынтегрального вы-

ражения в последней формуле, найдём

$$\begin{aligned} \frac{G(j\omega)}{j\omega} e^{j\omega t} &= \frac{P(\omega) + jQ(\omega)}{j\omega} (\cos \omega t + j \sin \omega t) = \\ &= \frac{P(\omega) \sin \omega t + Q(\omega) \cos \omega t}{\omega} + j \frac{-P(\omega) \cos \omega t + Q(\omega) \sin \omega t}{\omega}. \end{aligned} \quad (7.6)$$

Отсюда для интеграла в правой части (7.5) получаем

$$\begin{aligned} \frac{1}{2\pi} \int_{\mathbb{R}} \frac{G(j\omega)}{j\omega} e^{j\omega t} d\omega &= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{P(\omega) \sin \omega t + Q(\omega) \cos \omega t}{\omega} d\omega + \\ &+ j \frac{1}{2\pi} \int_{\mathbb{R}} \frac{-P(\omega) \cos \omega t + Q(\omega) \sin \omega t}{\omega} d\omega. \end{aligned} \quad (7.7)$$

Второе слагаемое в этой формуле равно нулю, так как $h(t)$ — действительная функция (или, что то же самое, так как подынтегральная функция нечётна). Учитывая далее, что первое слагаемое в (7.7) — чётная функция, получим для него

$$\frac{1}{2\pi} \int_{\mathbb{R}} \frac{P(\omega) \sin \omega t + Q(\omega) \cos \omega t}{\omega} d\omega = \frac{1}{\pi} \int_0^{+\infty} \frac{P(\omega) \sin \omega t + Q(\omega) \cos \omega t}{\omega} d\omega. \quad (7.8)$$

Эта формула справедлива для любых t . Однако для $t < 0$ $h(t) = 0$ и подстановка (7.8) в (7.3) даёт

$$0 = \frac{1}{2}P(0) - \frac{1}{\pi} \int_0^{+\infty} \frac{P(\omega) \sin \omega t}{\omega} d\omega + \frac{1}{\pi} \int_0^{+\infty} \frac{Q(\omega) \cos \omega t}{\omega} d\omega, \quad (7.9)$$

т.е.

$$\frac{1}{\pi} \int_0^{+\infty} \frac{P(\omega) \sin \omega t}{\omega} d\omega = \frac{1}{2}P(0) + \frac{1}{\pi} \int_0^{+\infty} \frac{Q(\omega) \cos \omega t}{\omega} d\omega. \quad (7.10)$$

Формула (7.5) с учётом (7.8) записывается так:

$$h(t) = \frac{1}{2}P(0) + \frac{1}{\pi} \int_0^{+\infty} \frac{P(\omega) \sin \omega t}{\omega} d\omega + \frac{1}{\pi} \int_0^{+\infty} \frac{Q(\omega) \cos \omega t}{\omega} d\omega. \quad (7.11)$$

Используя (7.10), её окончательно можно записать в одной из следующих форм:

$$h(t) = P(0) + \frac{2}{\pi} \int_0^{+\infty} \frac{Q(\omega) \cos \omega t}{\omega} d\omega, \quad (7.12)$$

$$h(t) = \frac{2}{\pi} \int_0^{+\infty} \frac{P(\omega) \sin \omega t}{\omega} d\omega. \quad (7.13)$$

Эти формулы используются для нахождения переходной характеристики системы управления, а последняя формула — для приближённого анализа качества переходных процессов.

Быстродействие системы определяется интервалом существенных частот — $\omega_{сч}$ (полосой пропускания), т.е. интервалом, на котором $|P(\omega)|$ превышает уровень $5\%P(0)$ (рис. 7.1). Время регулирования t_p при этом (для

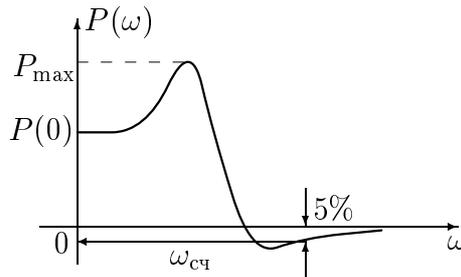


Рис. 7.1.

случая, когда $P(\omega)$ хорошо аппроксимируется трапецией) может быть оценено так [6]:

$$\frac{\pi}{\omega_{сч}} < t_p < \frac{4\pi}{\omega_{сч}}. \quad (7.14)$$

Из (7.14) вытекает, что чем больше интервал существенных частот (полоса пропускания), тем меньше время регулирования. Величина P_{\max} играет существенную роль в частотных методах синтеза систем управления (см. §8).

Из теорем о начальном и конечном значениях преобразования Лапласа вытекает, что

$$\begin{aligned} h(0) &= \lim_{\omega \rightarrow \infty} P(\omega), \\ h(+\infty) &= P(0). \end{aligned} \quad (7.15)$$

Таким образом, значения вещественной частотной характеристики при больших ω влияют на начальную часть переходной характеристики, значения вещественной частотной характеристики при малых ω оказывают влияние на конечную часть переходной характеристики. Средняя часть вещественной частотной характеристики в интервале существенных частот оказывает наибольшее влияние на переходный процесс в системе управления.

Склонность системы к колебаниям определяется по вещественной частотной характеристике следующим образом.

Если вещественная частотная характеристика имеет максимум в интервале существенных частот (рис. 7.1), то система имеет перерегулирование $\sigma > 18\%$. При $P_{\max} \rightarrow \infty$ система приближается к границе устойчивости с парой мнимых корней характеристического полинома.

Если вещественная частотная характеристика является невозрастающей функцией, т.е. $\frac{dP}{d\omega} \leq 0$, то перерегулирование $\sigma < 18\%$.

Если вещественная частотная характеристика является убывающей функцией, $\frac{dP}{d\omega} < 0$ и модуль $\frac{dP}{d\omega}$ убывает, то переходной процесс в системе является монотонным.

Графики вещественных частотных характеристик, соответствующие этим двум случаям, показаны на рис. 7.2 а) и б).

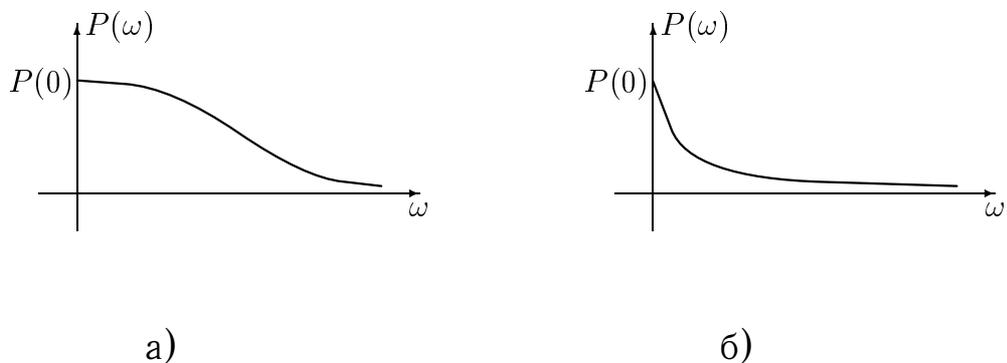


Рис. 7.2.

Склонность системы к колебаниям можно также оценить не по вещественной частотной характеристике, а по её амплитудной частотной характеристике с помощью показателя колебательности.

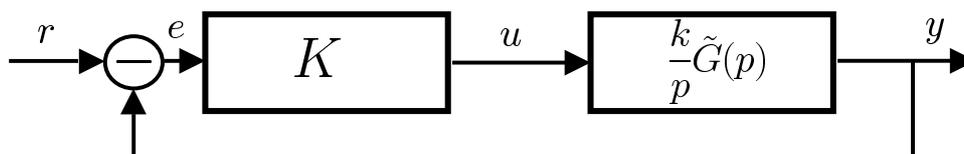


Рис. 7.3.

Показатель колебательности используется, в основном, для следящих систем, которые являются астатическими. Структурная схема такой системы приведена на рис. 7.3.

Передаточная функция замкнутой системы по выходному сигналу имеет вид

$$W_y(p) = \frac{K \frac{k}{p} \tilde{G}(p)}{1 + K \frac{k}{p} \tilde{G}(p)} = \frac{Kk \tilde{G}(p)}{p + Kk \tilde{G}(p)}. \quad (7.16)$$

Учитывая, что $\tilde{G}(p) \Big|_{p=0} = 1$, отсюда получаем, что $W_y(0) = 1$. Пусть амплитудная частотная характеристика замкнутой системы $A(\omega)$ имеет вид, показанный на рис. 7.4.

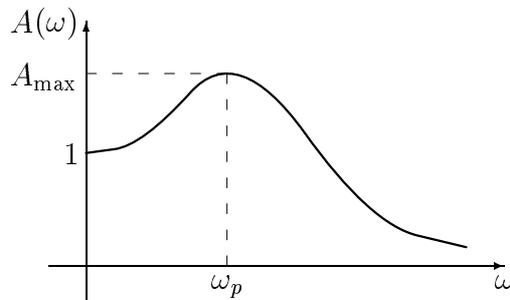


Рис. 7.4.

Показатель колебательности M — это максимальное значение амплитудной частотной характеристики, т.е. $M = A_{\max}$. Чем выше M , тем более система склонна к колебаниям и, значит, тем менее качественна при прочих равных условиях. На резонансной частоте ω_p гармоническое колебание проходит через систему с максимальным коэффициентом усиления, что и определяет физический смысл показателя колебательности. При $M \rightarrow +\infty$ передаточная функция замкнутой системы приближается к границе устойчивости с парой мнимых корней характеристического полинома.

Экспериментальные исследования показали, что если показатель колебательности лежит в пределах $1,1 \leq M \leq 1,5$, то качество переходных процессов при ступенчатом воздействии тоже является удовлетворительным. Время регулирования может быть оценено по резонансной частоте по формуле $t_p \approx \alpha \frac{2\pi}{\omega_p}$, где $\alpha = 1,2$ — некоторый коэффициент. Обычно задают $M = 1,4$ и дальнейшее уточнение результатов синтеза проводят с помощью других методов.

Удобство показателя колебательности обусловлено тем, что для замкнутой системы управления его величину можно найти без отыскания амплитудной частотной характеристики её, а непосредственно по её годографу Найквиста, т.е. АФЧХ разомкнутой системы.

Обозначим $H(j\omega) = K \frac{k}{p} \tilde{G}(p)$, тогда показатель колебательности замкнутой системы определится формулой

$$M = \left| \frac{H(j\omega_p)}{1 + H(j\omega_p)} \right|. \quad (7.17)$$

Найдём уравнение для линий равных значений M (линий уровня) на комплексной плоскости с координатами (U, V) , где $U = \operatorname{Re} H(j\omega)$, $V = \operatorname{Im} H(j\omega)$. Из (7.17) имеем

$$M = \left| \frac{U + jV}{1 + U + jV} \right| = \frac{|U + jV|}{|1 + U + jV|} = \frac{\sqrt{U^2 + V^2}}{\sqrt{(1 + U)^2 + V^2}}. \quad (7.18)$$

Возводя в квадрат и избавляясь от знаменателя, получим

$$M^2(1 + U)^2 + M^2V^2 = U^2 + V^2, \quad (7.19)$$

или

$$U^2(M^2 - 1) + 2UM^2 + M^2 + V^2(M^2 - 1) = 0. \quad (7.20)$$

Для $M = 1$ это уравнение даёт линию уровня $U = -\frac{1}{2}$. Разделив (7.20) на $M^2 - 1$, получаем

$$U^2 + 2U \frac{M^2}{M^2 - 1} + \frac{M^2}{M^2 - 1} + V^2 = 0. \quad (7.21)$$

Прибавим и вычтем слагаемое $\frac{M^4}{(M^2-1)^2}$ для того, чтобы получить полный квадрат. Получаем

$$U^2 + 2U \frac{M^2}{M^2 - 1} + \frac{M^4}{(M^2 - 1)^2} - \frac{M^4}{(M^2 - 1)^2} + \frac{M^2}{M^2 - 1} + V^2 = 0, \quad (7.22)$$

откуда окончательно находим такое уравнение для линий уровня:

$$\left(U + \frac{M^2}{M^2 - 1} \right)^2 + V^2 = \frac{M^2}{(M^2 - 1)^2}. \quad (7.23)$$

Для $M > 1$ это уравнение есть уравнение семейства окружностей радиуса $R = \frac{M}{M^2-1}$ и с центром в точке $c = -\frac{M^2}{M^2-1}$. При $M \rightarrow +\infty$ окружность вырождается в точку $(-1; j0)$. Для $M < 1$ $R = \frac{M}{M^2-1}$ и при $M = 0$ окружность вырождается в точку начала координат. Семейства окружностей для $M > 1$ и $M < 1$ симметричны относительно линии уровня $u = -\frac{1}{2}$, соответствующей значению $M = 1$. Соответствующие графики приведены на рис. 7.5. Для отыскания показателя колебательности замкнутой системы доста-

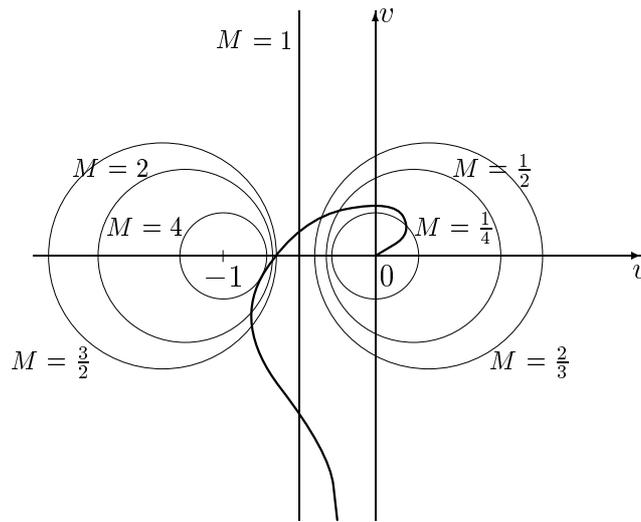


Рис. 7.5.

точно найти значение $M > 1$, определяющее радиус наименьшей окружности, которой касается, не пересекая её, годограф Найквиста (см. рис. 7.5). Используя приведённые графики можно не только найти показатель колебательности, но и построить всю амплитудную частотную характеристику замкнутой системы по точкам пересечения годографа Найквиста с соответствующими окружностями, определяющими значение амплитудной частотной характеристики на данной частоте.

Существенное достоинство показателя колебательности заключается в простоте его использования для целей синтеза систем управления. Если требуется обеспечить стандартное значение показателя колебательности $M = 1,4$, то для этого нужно только так выбрать коэффициент усиления регулятора K системы на рис. 7.3, чтобы её годограф Найквиста только касался, но не заходил внутрь запретной зоны, показанной на рис. 7.6 штриховкой.

К другим важным частотным показателям качества относятся запасы устойчивости по амплитуде — ΔA и по фазе — $\Delta \varphi$, которые определяются

по годографу Найквиста. Важность запасов устойчивости обусловлена тем обстоятельством, что параметры системы могут менять свои значения под действием различных факторов. Как следствие, система может оказаться работающей вблизи границы устойчивости или даже вообще потерять устойчивость. Чтобы этого не произошло, необходимо обеспечить достаточную удалённость годографа Найквиста от критической точки $(-1; 0j)$. Запасы устойчивости по амплитуде и по фазе определяются так, как показано на рис. 7.7 а). С ростом коэффициента усиления K разомкнутой цепи график АФЧХ будет приближаться к критической точке $(-1; 0j)$ и, как следствие, замкнутая система может потерять устойчивость.

Запасы устойчивости по амплитуде и фазе часто определяются по логарифмическим частотным характеристикам (рис. 7.7 б)), тогда запас по амплитуде измеряется в децибеллах. При проектировании обычно выбирают $\Delta L = 6 - 20$ дБ и $\Delta\varphi = 30 - 60^\circ$.

Отмеченная на рис. 7.7 а) точка пересечения годографа Найквиста с окружностью единичного радиуса, для которой $A(\omega_{ср}) = 1$, определяет важную частотную характеристику системы, называемую частотой среза — $\omega_{ср}$. Для логарифмической амплитудной частотной характеристики $L(\omega_{ср}) = 0$ (рис. 7.7 б)). Величина $\omega_{ср}$ существенным образом влияет на качество системы управления. Если переходный процесс имеет 1–2 колебания, то время регулирования может быть оценено по формуле $t_p \approx \alpha \frac{2\pi}{\omega_{ср}}$, $\alpha = 1, 2$, аналогичной соответствующей формуле для резонансной частоты. Из этой формулы вытекает, что чем больше частота среза, тем больше быстродействие системы. Для получения приемлемого качества процессов в системе управления логарифмическая амплитудная частотная характеристика должна иметь достаточно длинный прямолинейный участок с наклоном 20

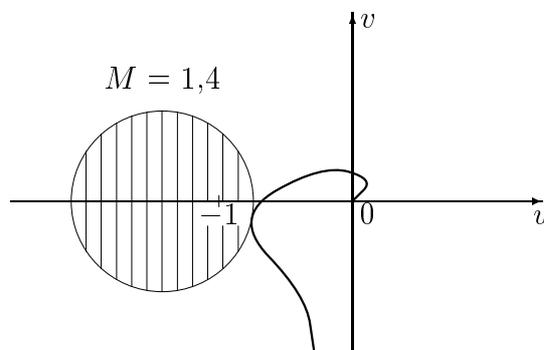


Рис. 7.6.

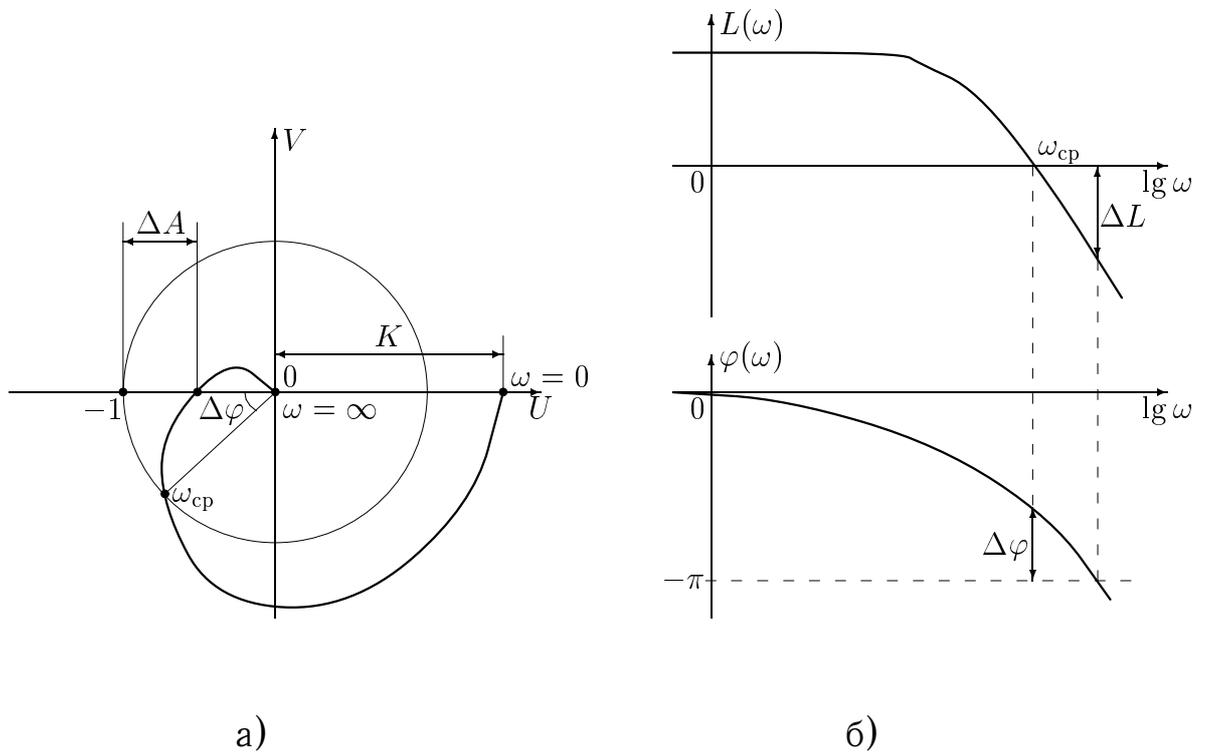


Рис. 7.7.

дБ/дек в окрестности частоты среза.

Рассмотрим смысл этого требования. Для апериодического звена первого порядка с передаточной функцией

$$G(p) = \frac{K}{Tp + 1} \quad (7.24)$$

амплитудная частотная характеристика задаётся формулой

$$A(\omega) = \frac{K}{\sqrt{(T\omega)^2 + 1}}. \quad (7.25)$$

Соответствующая логарифмическая частотная характеристика и асимптотическая логарифмическая частотная характеристика задаются формулами

$$L(\omega) = 20 \lg K - 20 \lg \sqrt{(T\omega)^2 + 1}, \quad (7.26)$$

$$L_{ac} = \begin{cases} 20 \lg K, & \omega \leq \frac{1}{T}, \\ 20 \lg K - 20 \lg T\omega, & \omega > \frac{1}{T}. \end{cases} \quad (7.27)$$

Графики, определяемые этими формулами, приведены на рис. 7.8 а) и б). Из рис. 7.8 б) непосредственно следует, что рассматриваемое требование

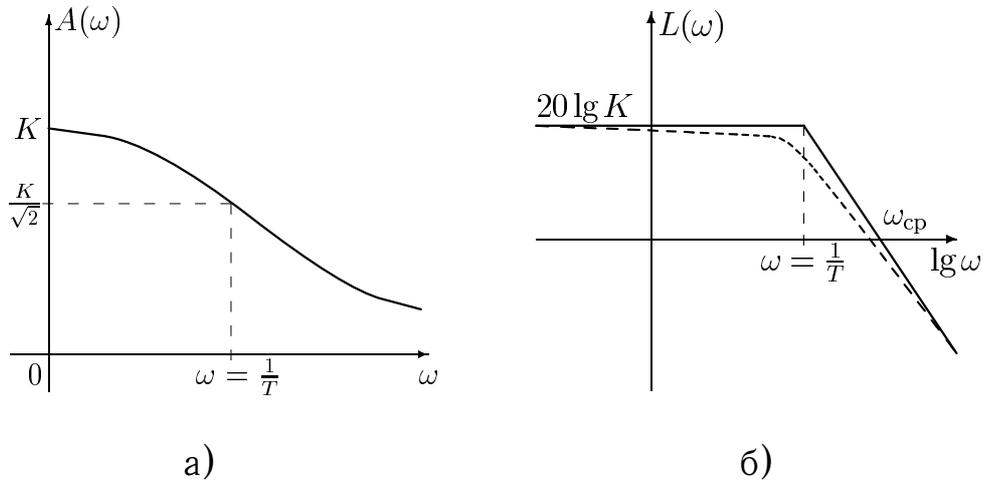


Рис. 7.8.

означает, что качественная система управления в окрестности частоты среза должна быть подобной аperiodическому звену.

Частотные характеристики аperiodического звена используются также для введения важного показателя качества системы, называемого полосой пропускания. Из рассмотрения графиков на рис. 7.8 б) следует, что для частот меньше сопрягающей $\omega = \frac{1}{T}$ гармонические колебания хорошо пропускаются системой, а для частот, больших $\frac{1}{T}$, они подавляются системой. Поэтому величина

$$\Delta = \frac{1}{T} - \left(-\frac{1}{T}\right) = \frac{2}{T}$$

называется полосой пропускания аperiodического звена. Отклонение асимптотической ЛАЧХ от точной максимально при $\omega = \frac{1}{T}$ и равно примерно 3 дБ. Поэтому для любой системы управления определяют $\omega_{\text{пр}}$ как частоту, на которой усиление $A(\omega_{\text{пр}})$ снижается на 3 дБ по сравнению с усилением на нулевой частоте $A(0)$. Это соответствует тому, что $A(\omega_{\text{пр}}) = \frac{1}{\sqrt{2}}A(0)$, как показано на рис. 7.8 а) для аperiodического звена. Чем больше полоса пропускания, тем больше быстродействие системы. Однако, она не должна быть слишком широкой, так как тогда система будет воспроизводить на выходе высокочастотные помехи, действующие на вход системы.

§8. Частотные методы стабилизации и синтеза корректирующих устройств

Точность системы управления, как это было показано в §5, определяется коэффициентом усиления регулятора. Точность тем выше, чем этот коэффициент больше. Однако, с ростом коэффициента усиления регулятора запасы устойчивости по амплитуде и фазе уменьшаются и при достаточно больших значениях этого коэффициента система может потерять устойчивость. В этой связи возникает задача стабилизации, т.е. задача обеспечения устойчивости замкнутой системы. Если эту задачу решить путём настройки параметров исходной системы не удаётся, то её решают путём введения в состав системы специальных корректирующих (стабилизирующих) устройств, т.е. путём модификации структуры системы.

Универсальным методом стабилизации систем управления является дифференциальная коррекция, заключающаяся в введении в состав системы форсирующего звена, как это показано на рис. 8.1 а). Передаточная

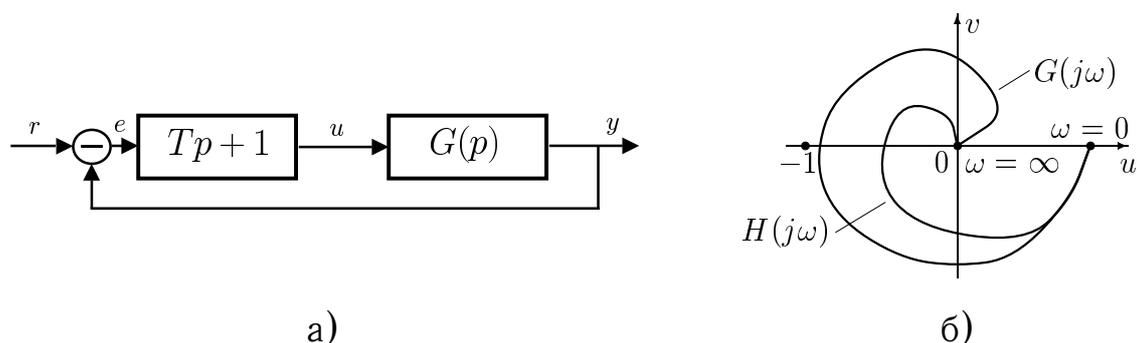


Рис. 8.1.

функция разомкнутой цепи при этом получается следующей:

$$H(p) = (Tp + 1)G(p). \quad (8.1)$$

Фазовая частотная характеристика разомкнутой цепи скорректированной системы имеет вид:

$$\varphi_H = \varphi_G + \operatorname{arctg} T\omega, \quad (8.2)$$

т.е. отличается от исходной положительным слагаемым $\operatorname{arctg} T\omega$, которое при больших ω , для высоких частот, близко к $\frac{\pi}{2}$. Как следствие, годограф Найквиста скорректированной системы получается из годографа Найквиста нескорректированной системы поворотом против часовой стрелки (рис.

8.1 б)). Запасы устойчивости по фазе и амплитуде при этом увеличиваются.

Амплитудная частотная характеристика разомкнутой цепи скорректированной системы имеет вид

$$A_H(\omega) = \sqrt{T^2\omega^2 + 1}A_G(\omega). \quad (8.3)$$

Наличие множителя $\sqrt{T^2\omega^2 + 1}$, растущего с ростом ω , обеспечивает одновременно расширение полосы пропускания системы, т.е. её качество.

Если введение форсирующего звена первого порядка не обеспечивает получение требуемых запасов устойчивости и полосы пропускания, могут использоваться введение форсирующего звена второго порядка или двух форсирующих звеньев первого порядка, если это технически осуществимо.

Введение дифференциальной коррекции, т.е. использование производной от сигнала ошибки за счёт расширения полосы пропускания делает систему чувствительной к высокочастотным помехам, что иногда требует усложнения структуры корректирующего устройства и введения дополнительных фильтров.

Другим распространённым методом коррекции является использование интеграла от сигнала ошибки, интегральная коррекция в соответствии со структурной схемой на рис. 8.2 а). Интегральная коррекция делает ста-

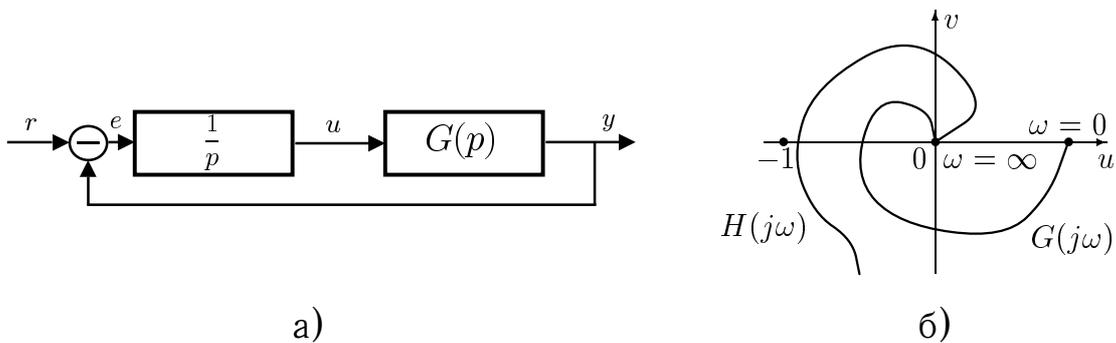


Рис. 8.2.

тическую систему астатической первого порядка, что приводит к увеличению её точности. Передаточная функция разомкнутой цепи имеет в этом случае вид

$$H(p) = \frac{1}{p}G(p), \quad (8.4)$$

что даёт следующую фазовую частотную характеристику разомкнутой системы:

$$\varphi_H(\omega) = \varphi_G(\omega) - \frac{\pi}{2}. \quad (8.5)$$

Наличие отрицательного сдвига по фазе $-\frac{\pi}{2}$ приводит к уменьшению запасов устойчивости системы и ухудшению качества переходного процесса.

При использовании интегрального звена второго порядка $\frac{1}{p^2}$ в качестве корректирующего устройства в случае, когда передаточная функция исходной системы имеет вид

$$G(p) = \frac{K}{\sum_{i=0}^n a_i p^i},$$

характеристическое уравнение замкнутой системы записывается как

$$p^2 \sum_{i=0}^n a_i p^i + K = 0. \quad (8.6)$$

Так как в этом уравнении отсутствует слагаемое с первой степенью p , то нарушается необходимое условие устойчивости Стодолы и такая система является неустойчивой. Это обстоятельство указывает на сложность реализации системы с астатизмом второго порядка, которые могут быть реализованы только при условии, что передаточная функция исходной системы имеет вид

$$G(p) = \frac{\sum_{i=0}^m b_i p^i}{\sum_{i=0}^n a_i p^i}, \quad m > 0.$$

От указанных существенных недостатков интегральной коррекции свободно изобретено корректирующее устройство, сочетающее свойства дифференциальной и интегральной коррекции, с передаточной функцией

$$G_K(p) = \frac{Tp + 1}{Tp}. \quad (8.7)$$

Так как эту передаточную функцию можно записать как

$$G_K(p) = 1 + \frac{1}{Tp}, \quad (8.8)$$

то это означает, что изотропное корректирующее устройство осуществляет замыкание обратной связи как по сигналу ошибки, так и по интегралу от него. Изотропное корректирующее устройство позволяет получить астатическую систему при сохранении свойств устойчивости и качества переходных процессов. Такое свойство изотропного корректирующего устройства является следствием специальной формы его частотных характеристик, показанных на рис. 8.3. Асимптотическая ЛАЧХ (рис. 8.3 в)) показывает,

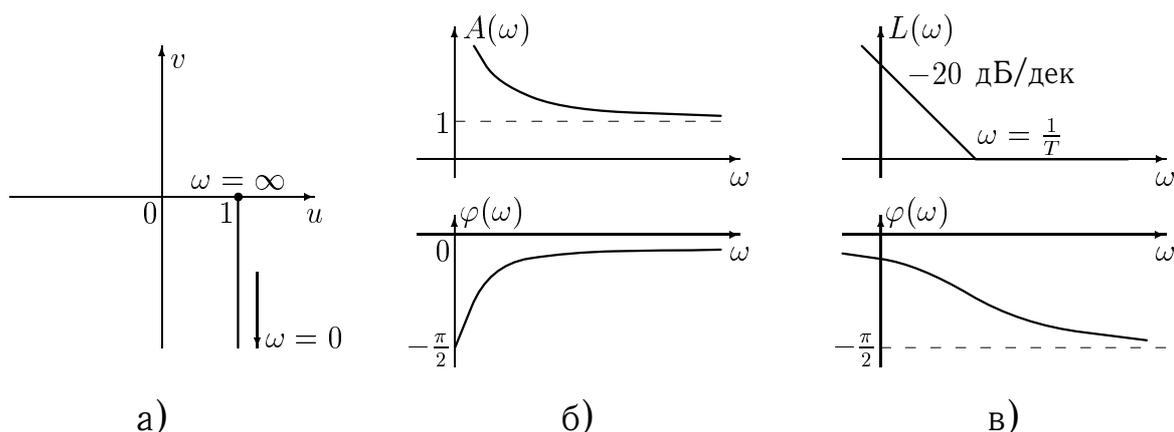


Рис. 8.3.

что введение изотропного корректирующего устройства приводит к повышению точности системы, за что отвечает низкочастотная часть, и несущественному фазовому сдвигу на высоких частотах, т.е. несущественному уменьшению запасов устойчивости и ухудшению качества переходного процесса.

Рассмотренные простейшие методы коррекции удобно осуществлять с использованием показателя колебательности, добиваясь, чтобы годограф Найквиста скорректированной системы не заходил внутрь запретной зоны, соответствующей заданному показателю колебательности, например, стандартному его значению $M = 1,4$.

Помимо стабилизации системы, корректирующие устройства широко применяются для улучшения показателей точности и качества переходного процесса в системе. Синтез корректирующих устройств для указанных целей можно осуществить с помощью частотных методов, основанных на применении асимптотических ЛАЧХ.

Синтез последовательного корректирующего устройства для системы с обратной связью (рис. 8.4) осуществляется по следующей методике.

1. Строится асимптотическая ЛАЧХ $L(\omega)$ разомкнутой цепи исходной

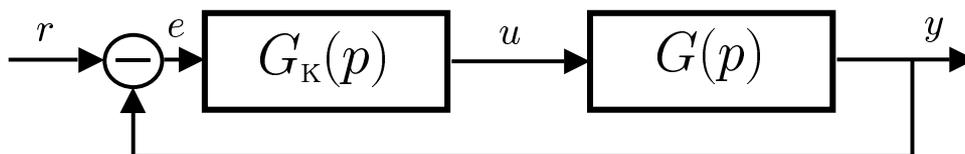


Рис. 8.4.

системы, под которой понимается система, содержащая объект управления и регулятор, и не содержащая корректирующих устройств.

2. Строится асимптотическая ЛАЧХ разомкнутой системы с желаемым поведением $L_{\text{ж}}(\omega)$ исходя из требований точности и качества переходного процесса.
3. ЛАЧХ корректирующего устройства $L_{\text{к}}(\omega)$ находится вычитанием

$$L_{\text{к}}(\omega) = L_{\text{ж}}(\omega) - L(\omega), \quad (8.9)$$

что следует из желаемого равенства

$$G_{\text{ж}}(p) = G_{\text{к}}(p)G(p). \quad (8.10)$$

Построение асимптотической ЛАЧХ разомкнутой цепи осуществляется по обычным правилам. Построение же желаемой асимптотической ЛАЧХ можно осуществить различными методами, использующими частотный подход.

Рассмотрим один из вариантов такого подхода, где используются специальные графики, связывающие перерегулирование $\sigma\%$, время регулирования t_p , частоту среза $\omega_{\text{ср}}$ и запасы устойчивости по амплитуде ΔL и фазе $\Delta\varphi$, полученные из вещественных частотных характеристик систем, приведённые на рис. 8.5 а) и б). На рис. 8.5 б) P_{max} — максимальное значение вещественной частотной характеристики (см. §7).

С учётом того, что низкочастотная часть частотной характеристики влияет на точность системы управления, среднечастотная часть — на качество переходного процесса, а высокочастотная часть существенного влияния не оказывает, построение желаемой асимптотической ЛАЧХ для астатической системы можно осуществить следующим образом.

1. Строим среднечастотную часть желаемой ЛАЧХ. Для этого из графиков на рис. 8.5 б) находим сначала P_{max} , соответствующее заданному перерегулированию $\sigma\%$, а затем, по найденному значению P_{max} и заданному времени регулирования t_p , частоту среза $\omega_{\text{ср}}$. Через точку

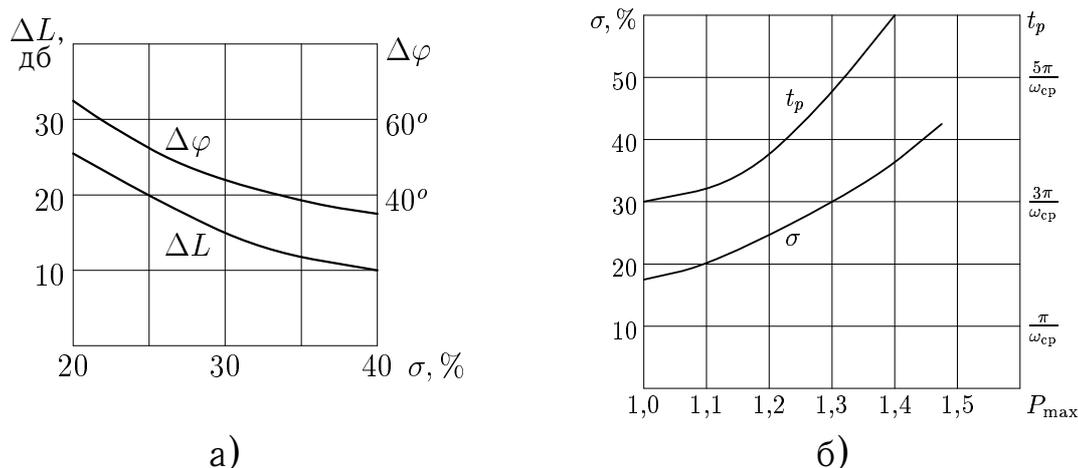


Рис. 8.5.

$\omega_{\text{ср}}$ на оси частот проводим прямую с наклоном -20 дБ/дек. Далее, из графиков на рис. 8.5 а) находим величину ΔL , соответствующую заданному $\sigma\%$. Построенную из точки $\omega_{\text{ср}}$ прямую продолжаем влево и вправо вплоть до достижения значений ΔL и $-\Delta L$.

2. Строим низкочастотную часть желаемой ЛАЧХ. Для этого по заданному значению добротности из точки δ_0 проводим низкочастотную асимптоту с наклоном -20 дБ/дек. Сопрягаем низкочастотную и среднечастотную части прямой с наклоном -40 дБ/дек или -60 дБ/дек.
3. Высокочастотную часть берём из асимптотической ЛАЧХ исходной системы и сопрягаем её со среднечастотной частью в точке, где среднечастотная часть достигает значения $-\Delta L$.
4. Проверяем соответствие запаса по фазе $\Delta\varphi$ требуемому, находимому из графика на рис. 8.5 а) по заданному перерегулированию $\sigma\%$ для построенной среднечастотной части. Если требования к запасу по фазе не удовлетворяются, то расширяем среднечастотный участок желаемой ЛАЧХ.

Передаточная функция корректирующего устройства после построения желаемой ЛАЧХ находится по формуле (8.9) по сопрягающим частотам.

Пример. Пусть требуется синтезировать корректирующее устройство для системы со структурной схемой на рис. 8.4 при

$$G(p) = \frac{K}{p(Tp + 1)} \quad (8.11)$$

исходя из заданных требований на точность и качество переходного процесса: добротность δ_0 , перерегулирование $\sigma\%$, время регулирования t_p .

Строим асимптотическую ЛАЧХ разомкнутой цепи исходной системы по сопрягающей частоте $\omega = \frac{1}{T}$ и коэффициенту усиления K (рис. 8.6).

Строим асимптотическую ЛАЧХ разомкнутой цепи желаемой системы. Для построения среднечастотной части из графиков на рис. 8.5 б) по заданному $\sigma\%$ находим P_{\max} и по найденному значению P_{\max} соответствующее значение t_p в виде $t_p = \frac{q\pi}{\omega_{\text{ср}}}$, где q — коэффициент пропорциональности, значение которого находится из графика для t_p . Так как t_p задано, то отсюда находим частоту среза $\omega_{\text{ср}} = \frac{q\pi}{t_p}$ и проводим через точку $\omega_{\text{ср}}$ на оси абсцисс прямую с наклоном -20 дБ/дек. Это и все последующие построения показаны на рис. 8.6.

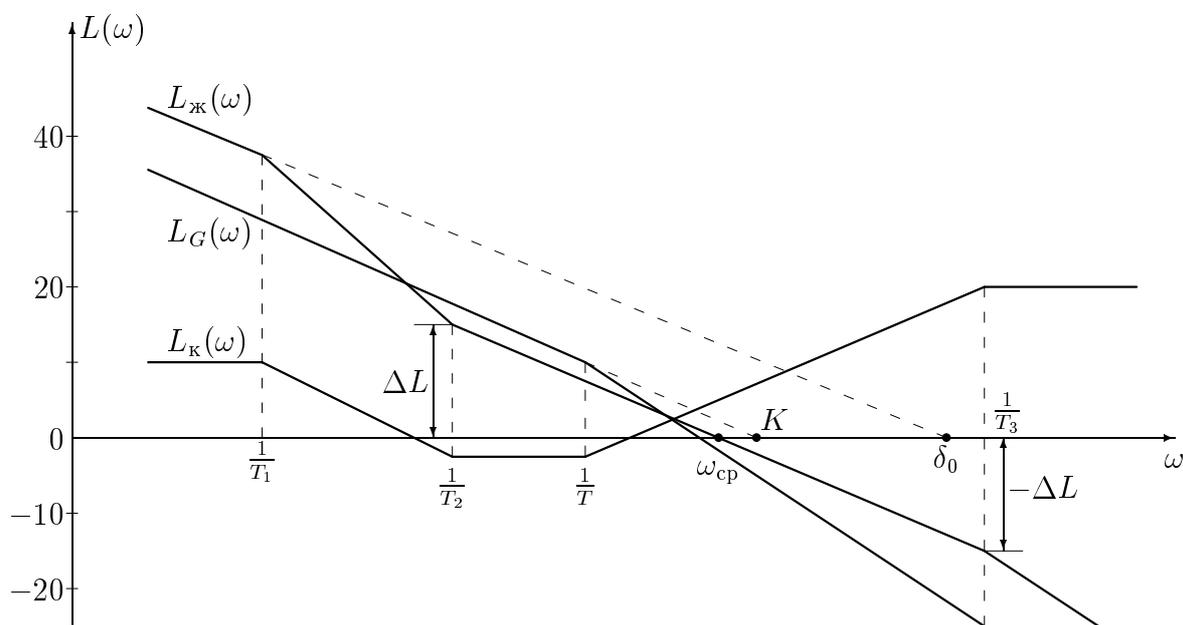


Рис. 8.6.

Из графика на рис. 8.5 а) по заданному значению $\sigma\%$ находим соответствующее значение ΔL и продолжаем построенную прямую влево до достижения значения ΔL и вправо до достижения значения $-\Delta L$. Среднечастотная часть желаемой ЛАЧХ построена.

Низкочастотную часть строим, проводя из точки на оси абсцисс $\omega = \delta_0$ прямую с наклоном -20 дБ/дек. Для сопряжения низкочастотной и среднечастотной частей из точки с координатами $(\omega, \Delta L)$ проводим прямую с наклоном -40 дБ/дек вплоть до пересечения с прямой, задающей низкочастотную часть в некоторой точке с абсциссой ω .

Высокочастотную часть желаемой ЛАЧХ берём из асимптотической ЛАЧХ исходной системы, сопрягая её со среднечастотной частью в точке с координатами $(\frac{1}{T_3}, -\Delta L)$.

Желаемая ЛАЧХ построена. Для нахождения ЛАЧХ корректирующего устройства графически вычитаем ЛАЧХ исходной системы из построенной желаемой ЛАЧХ (рис. 8.6).

Для нахождения передаточных функций желаемой разомкнутой системы $G_{\text{ж}}$ и корректирующего устройства $G_{\text{к}}$ находим сопрягающие частоты $\omega_1 = \frac{1}{T_1}$, $\omega_2 = \frac{1}{T_2}$, $\omega_3 = \frac{1}{T_1}$, $\omega_4 = \frac{1}{T_3}$ и по правилам построения асимптотических ЛАЧХ получаем

$$G_{\text{ж}}(p) = \frac{\delta_0(T_2p + 1)}{p(T_1p + 1)(T_3p + 1)}, \quad (8.12)$$

$$G_{\text{к}}(p) = \frac{(\delta_0 - K)(T_2p + 1)(Tp + 1)}{(T_1p + 1)(T_3p + 1)}. \quad (8.13)$$

Рассмотренный метод синтеза корректирующих устройств является приближённым не только за счёт использования асимптотических ЛАЧХ, но также и за счёт использования графиков на рис. 8.5, которые были получены для случая, когда вещественная частотная характеристика хорошо аппроксимируется суммой двух трапеций. Поэтому окончательная проверка удовлетворительности результатов синтеза проводится построением кривой переходного процесса в скорректированной системе.

§9. Синтез систем по интегральным оценкам качества

Помимо показателей качества, характеризующих отдельные свойства системы, в теории автоматического управления используются *интегральные оценки качества*, которые позволяют получить общую оценку качества переходного процесса: его длительность и его уклонение от установившегося значения.

Простейшей интегральной оценкой является линейная оценка, выражаемая формулой

$$J_{\text{лин}} = \int_0^{+\infty} \varepsilon(t) dt, \quad (9.1)$$

где $\varepsilon(t) = y(+\infty) - y(t)$ — уклонение выходного сигнала $y(t)$ от его установившегося значения при единичном ступенчатом воздействии. Геометрический смысл оценки (9.1) очевиден — это площадь, ограниченная кривой переходного процесса (рис. 9.1). Чем меньше величина этой площади, тем

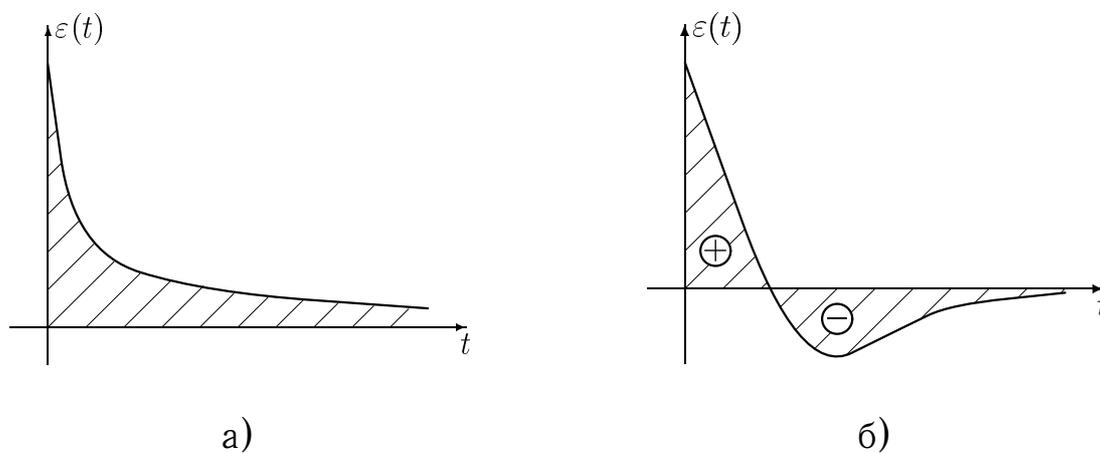


Рис. 9.1.

выше быстродействие системы и тем меньше уклонение выходного сигнала от его установившегося значения.

Для асимптотически устойчивой системы $\varepsilon(t) \rightarrow 0$ при $t \rightarrow +\infty$ и интеграл (9.1) сходится. Если же система не является асимптотически устойчивой, то интеграл (9.1) расходится. Поэтому, если параметры синтезируемой системы выбирать из условия минимума интегральной оценки (9.1), то это должно обеспечить как асимптотическую устойчивость, так и быстроту выхода системы на установившийся режим при малых отклонениях от него в переходном процессе.

Рассмотрим способ вычисления линейной интегральной оценки. Используя преобразования Лапласа, по теореме интегрирования оригинала

$$\int_0^t \varepsilon(t) dt \doteq \frac{1}{p} E(p).$$

По теореме об установившемся значении

$$J_{\text{лин}} = \lim_{t \rightarrow +\infty} \int_0^t \varepsilon(t) dt = \lim_{p \rightarrow 0} \frac{p}{p} E(p) = E(0).$$

Таким образом, имея формулу для $E(p)$ всегда можно вычислить значение линейной интегральной оценки. Т.к. $Y(p) = H(p) \frac{1}{p}$, то

$$y(+\infty) = \lim_{p \rightarrow 0} p H(p) \frac{1}{p} = H(0)$$

и

$$E(p) = \frac{H(0) - H(p)}{p}. \quad (9.2)$$

При $H(p) = \frac{\sum_{i=0}^m b_i p^i}{\sum_{i=0}^n a_i p^i}$

$$Y(p) = \frac{\sum_{i=0}^m b_i p^i}{\left(\sum_{i=0}^n a_i p^i\right) p},$$

$$y(+\infty) = \lim_{p \rightarrow 0} p Y(p) = \frac{b_0}{a_0},$$

$$E(p) = \frac{b_0}{a_0 p} - \frac{\sum_{i=0}^m b_i p^i}{\left(\sum_{i=0}^n a_i p^i\right) p}$$

и, следовательно,

$$J_{\text{лин}} = E(0) = \left. \frac{b_0 \left(\sum_{i=0}^n a_i p^i\right) - a_0 \left(\sum_{i=0}^m b_i p^i\right)}{a_0 \left(\sum_{i=0}^n a_i p^i\right) p} \right|_{p=0} = \frac{b_0 a_1 - a_0 b_1}{a_0^2}. \quad (9.3)$$

Пример. Рассмотрим астатическую систему с обратной связью (рис. 9.2), где $H(p) = \frac{\sum_{i=0}^m b_i p^i}{p \sum_{i=0}^n a_i p^i}$. Имеем

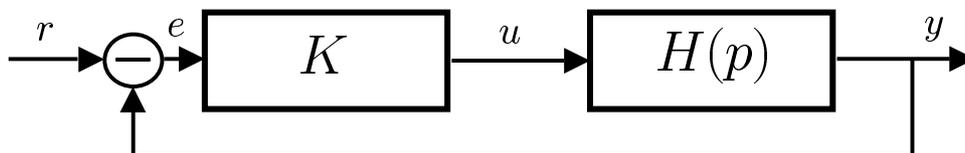


Рис. 9.2.

$$Y(p) = \frac{KH(p)}{1 + KH(p)} \cdot \frac{1}{p} = \frac{K \sum_{i=0}^m b_i p^i}{p \sum_{i=0}^n a_i p^i + K \sum_{i=0}^m b_i p^i} \cdot \frac{1}{p}$$

и

$$y(+\infty) = \lim_{p \rightarrow 0} pY(p) = \frac{b_0}{b_0} = 1,$$

что и следовало ожидать, т.к. система астатическая и в пределе входное ступенчатое воздействие обрабатывает точно.

Далее, находим

$$\begin{aligned} y(+\infty) - y(t) \doteq E(p) &= \frac{1}{p} - \frac{K \sum_{i=0}^m b_i p^i}{p \sum_{i=0}^n a_i p^i + K \sum_{i=0}^m b_i p^i} \cdot \frac{1}{p} = \\ &= \frac{\sum_{i=0}^n a_i p^i}{p \sum_{i=0}^n a_i p^i + K \sum_{i=0}^m b_i p^i} \end{aligned}$$

и

$$J_{\text{лин}} = E(0) = \frac{a_0}{Kb_0}. \quad (9.4)$$

Чем больше коэффициент усиления регулятора K , тем меньше линейная интегральная оценка и тем выше быстродействие системы. Однако выбор очень большого K может привести к приближению к границе

устойчивости, что отрицательно сказывается на надёжности системы. Если переходный процесс в системе является колебательным (рис. 9.1 б), то при вычислении интеграла (9.1) площади, ограниченной кривой $\varepsilon(t)$, складываются алгебраически и малая величина линейной интегральной оценки может и не свидетельствовать о высоком качестве системы.

От этого недостатка свободны *квадратичные интегральные оценки*. Простейшая квадратичная оценка задаётся формулой

$$J_{\text{кв}} = \int_0^{+\infty} \varepsilon^2(t) dt, \quad (9.5)$$

где, как и в случае линейной оценки, $\varepsilon(t) = y(+\infty) - y(t)$. Вычисление квадратичной оценки основано на применении равенства Парсеваля для преобразования Фурье:

$$\int_0^{+\infty} \varepsilon^2(t) dt = \frac{1}{2\pi} \int_{\mathbb{R}} |E(j\omega)|^2 d\omega. \quad (9.6)$$

Выведем эту формулу. Имеем

$$\begin{aligned} \frac{1}{2\pi} \int_{\mathbb{R}} |E(j\omega)|^2 d\omega &= \frac{1}{2\pi} \int_{\mathbb{R}} E(j\omega)E(-j\omega) d\omega = \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} E(-j\omega) \int_{\mathbb{R}} \varepsilon(t)e^{-j\omega t} dt d\omega = \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \varepsilon(t) \int_{\mathbb{R}} E(-j\omega)e^{-j\omega t} d\omega dt = \int_{\mathbb{R}} \varepsilon^2(t) dt. \end{aligned}$$

Выражение для $E(p)$ имеет вид $E(p) = \frac{H(0)-H(p)}{p}$. Подставляя $E(p)$ в (9.6) и интегрируя дробно-рациональную функцию, можно найти $J_{\text{кв}}$ как функцию коэффициентов $E(p)$ для заданных n и m . Для

$$E(p) = \frac{\sum_{i=0}^{n-1} c_i p^i}{\sum_{i=0}^m d_i p^i}$$

при условии, что все корни полинома $\sum_{i=0}^n d_i p^i$ лежат в левой полуплоскости, получаем:

$$\begin{aligned} n = 1 : \quad J_{\text{КВ}} &= \frac{c_0^2}{2d_1 d_0}, \\ n = 2 : \quad J_{\text{КВ}} &= \frac{c_1^2 d_0 + c_0^2 d_2}{2d_0 d_1 d_2}, \\ n = 3 : \quad J_{\text{КВ}} &= \frac{c_2^2 d_0 d_1 + (c_1^2 - 2c_0 c_1) d_0 d_3 + c_0^2 d_2 d_3}{2d_0 d_3 (d_1 d_2 - d_0 d_3)}. \end{aligned}$$

В литературе приводятся готовые формулы для $n \leq 10$. Однако, уже для $n \geq 5$ эти формулы становятся чрезвычайно громоздкими и, кроме того, они не позволяют записать решение в общем виде, для произвольного n .

Общее выражение для квадратичной оценки можно записать через определители, элементами которых являются коэффициенты передаточной функции $H(p)$ [6]. Если

$$H(p) = \frac{\sum_{i=0}^{n-1} b_i p^i}{\sum_{i=0}^n a_i p^i}$$

и все корни полинома $\sum_{i=0}^n a_i p^i$ лежат в левой полуплоскости, то для входного ступенчатого воздействия

$$J_{\text{КВ}} = \frac{1}{2a_0^2 \Delta} \left(\sum_{i=0}^{n-1} B_i \Delta_i - 2b_0 b_1 \Delta \right), \quad (9.7)$$

где

$$\Delta = \begin{vmatrix} a_0 & -a_2 & a_4 & -a_6 & \dots & 0 \\ 0 & a_1 & -a_3 & a_5 & \dots & 0 \\ 0 & -a_0 & a_2 & -a_4 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & a_{n-1} \end{vmatrix},$$

коэффициенты B_i определяются формулами:

$$\begin{aligned} B_0 &= b_0^2, \\ B_1 &= b_1^2 - 2b_0 b_2, \end{aligned}$$

вообще

$$B_i = b_i^2 - 2b_{i-1}b_{i+1} + \dots + 2(-1)^i b_0 b_{2i},$$

и

$$B_{n-1} = b_{n-1}^2,$$

определители $\Delta_i, i = \overline{0, m}$ получаются из определителя Δ заменой $i + 1$ -го столбца на столбец $(a_1, a_0, 0, \dots, 0)^T$.

Следует отметить, что определитель Δ равен старшему определителю Гурвица (см. §4) и обращается в 0 на границе устойчивости. Квадратичная оценка при этом, как видно из (9.7), будет стремиться к $+\infty$.

Синтез систем управления по интегральным оценкам качества осуществляется, в частности, следующим образом. Сначала синтезируется некоторая система, в которой оставляются варьируемые параметры, а затем записывается формула интегральной оценки и находится её минимум по варьируемым параметрам в их заданном диапазоне.

§10. Корневые оценки качества систем управления

Корневые оценки качества основаны на информации о расположении корней характеристического полинома передаточной функции системы.

Они используют следующую формулу для нахождения переходной характеристики системы:

$$h(t) = \frac{kM(0)}{N(0)} + \sum_{i=1}^n \frac{kM(p_i)}{p_i N'(p_i)} e^{p_i t}, \quad (10.1)$$

в которой сумма вычисляется по всем корням p_i характеристического полинома в предположении, что эти корни простые и отличны от нуля.

Очевидно, что наибольшее влияние на быстродействие системы будут оказывать корни, расположенные ближе к мнимой оси, поскольку составляющие, соответствующие другим корням, будут затухать быстрее. В этой связи вводится понятие степени устойчивости η , под которой понимается абсолютное значение действительной части ближайшего к мнимой оси корня или пары комплексно-сопряжённых корней (рис. 10.1 а) и б), где для систем пятого порядка корни показаны крестиками).

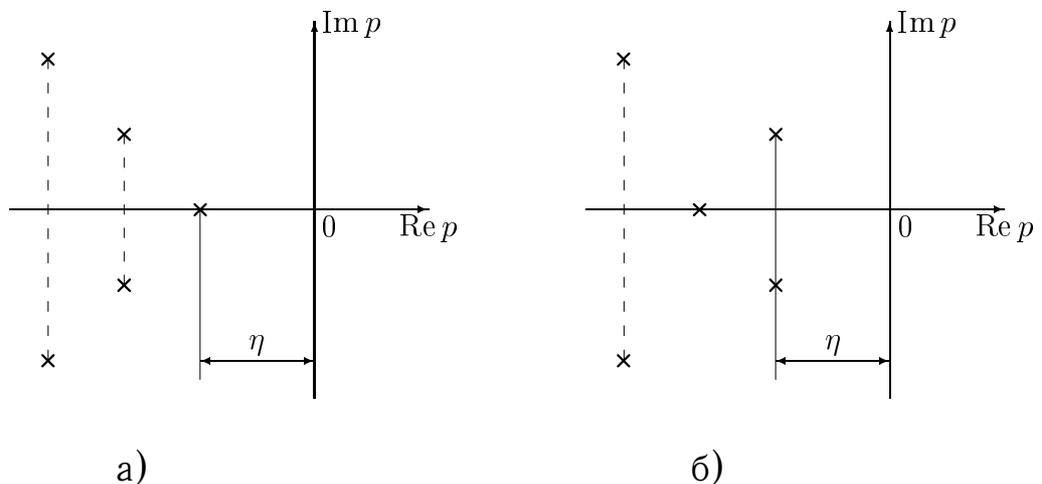


Рис. 10.1.

Пусть ближайшим к мнимой оси является действительный корень $p_1 = -\eta$. Тогда, как это следует из (10.1), переходный процесс, соответствующий этому корню, определяется функцией $e^{-\eta t}$. Считая, как и ранее, что переходный процесс заканчивается, как только решение входит в трубку величиной $\Delta = 0,05$, можно записать следующее уравнение для нахождения времени регулирования t_p : $\Delta = e^{-\eta t_p}$, откуда (для $\Delta = 0,05$) следует

такая формула для оценки времени регулирования:

$$t_p = \frac{1}{\eta} \ln \frac{1}{\Delta} \approx \frac{3}{\eta}. \quad (10.2)$$

В том случае, когда ближайшей к мнимой оси является пара комплексно-сопряжённых корней, переходный процесс, соответствующий этой паре, задаётся функцией $e^{-\eta t} \sin(\beta t + \varphi)$ и уравнение для нахождения времени регулирования имеет вид:

$$\Delta = e^{-\eta t_p} \sin(\beta t_p + \varphi).$$

Полагая $\sin(\beta t_p + \varphi) = 1$, отсюда получаем следующую оценку для времени регулирования (при $\Delta = 0,05$):

$$t_p \leq \frac{3}{\eta}. \quad (10.3)$$

Склонность системы управления к колебаниям в корневых оценках качества оценивается с помощью понятия колебательности, которое вводится формулой

$$\mu = \max \left| \frac{\beta_i}{\alpha_i} \right|. \quad (10.4)$$

Здесь α_i и β_i — действительная и мнимая части комплексно-сопряжённых корней $p_{i,i+1} = \alpha_i \pm \beta_i j$ и максимум находится по всем таким корням.

Величина колебательности позволяет оценить затухание колебаний за один период. В самом деле, пара комплексно-сопряжённых корней даёт переходный процесс вида $e^{-\alpha_i t} \sin(\beta_i t + \varphi_i)$ и при $t = T_i$, где $T_i = \frac{2\pi}{\beta_i}$ — период колебаний, амплитуда синусоидального колебания будет равна величине $e^{-2\pi \frac{\alpha_i}{\beta_i}}$.

Затуханием за период называют величину ξ_i , определяемую формулой

$$\xi_i = 1 - e^{-2\pi \frac{\alpha_i}{\beta_i}}. \quad (10.5)$$

Из этой формулы с учётом (10.4) вытекает, что минимальное затухание ξ , вычисляемое по всем комплексно-сопряжённым корням, определяется формулой

$$\xi = 1 - e^{-\frac{2\pi}{\mu}}, \quad (10.6)$$

откуда получаем

$$\mu = \frac{2\pi}{\ln \frac{1}{1-\xi}}. \quad (10.7)$$

Затухание ξ обычно выражается в процентах и приемлимыми считаются затухания не менее, чем 90–98%. При этом соответствующие значения колебательности по формуле (10.7) получаются 2,72 и 1,57. Величина μ определяет центральный угол, ограничивающий расположение корней характеристического уравнения (рис. 10.2). Совместно с заданием степени

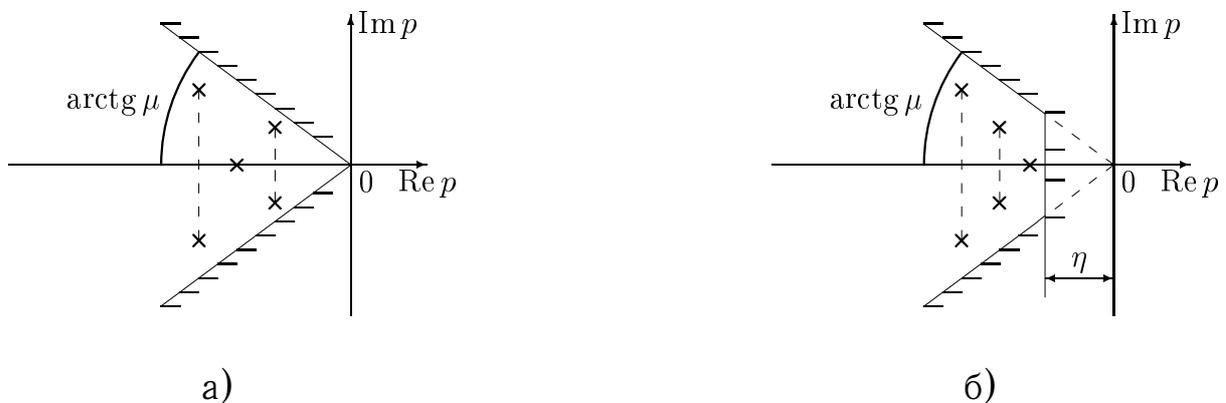


Рис. 10.2.

устойчивости, величина колебательности определяет допустимую область, в которой могут находиться корни характеристического уравнения, обеспечивающие заданное качество переходных процессов в системе управления.

Из формулы (10.1) следует, что качество переходных процессов может быть улучшено при условии учёта не только корней знаменателя передаточной функции — характеристического уравнения, т.е. полюсов, но и корней числителя передаточной функции — её нулей. В самом деле, разложив многочлен в числителе на множители, формулу (10.1) можно записать как

$$h(t) = \frac{kM(0)}{N(0)} + \sum_{i=1}^n \frac{k(p_i - \sigma_1) \dots (p_i - \sigma_m)}{p_i N'(p_i)} e^{p_i t}, \quad (10.8)$$

где σ_i — корни полинома в числителе, нули передаточной функции. Отсюда следует, что если нули расположены вблизи полюсов, то это приводит к уменьшению составляющих переходного процесса и, значит, к повышению его качества.

Это обстоятельство учитывается при проектировании систем управления по корневым оценкам качества.

§11. Метод корневого годографа

Метод корневого годографа является одним из основных и часто используемых на практике методов анализа и синтеза систем управления. Одним из важных преимуществ этого метода является возможность оценить влияние обычно имеющихся на практике изменений коэффициента усиления прямой цепи на расположение корней характеристического полинома замкнутой системы, т.е. на её динамику.

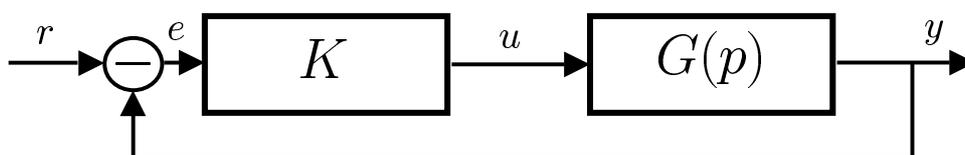


Рис. 11.1.

Рассмотрим структурную схему системы с обратной связью, приведённую на рис. 11.1. Здесь K — коэффициент усиления регулятора, а $G(p)$ есть передаточная функция объекта регулирования, имеющая вид

$$G(p) = \frac{B(p)}{A(p)}, \quad (11.1)$$

где $A(p)$ и $B(p)$ — полиномы степени n и m соответственно.

Так как передаточная функция этой системы по выходному сигналу задаётся формулой

$$W(p) = \frac{KG(p)}{1 + KG(p)}, \quad (11.2)$$

то, подставляя сюда (11.1), получим

$$W(p) = \frac{KB(p)}{A(p) + KB(p)}. \quad (11.3)$$

Отсюда следует, что нули замкнутой системы совпадают с нулями разомкнутой, а полюса замкнутой системы определяются полюсами, нулями и коэффициентом усиления разомкнутой системы. Выбирая коэффициент усиления, можно тем самым выбрать наиболее благоприятное расположение полюсов, т.е. решить задачу синтеза системы управления.

При $K = 0$ корни характеристического уравнения замкнутой системы совпадают с корнями характеристического уравнения разомкнутой. При

возрастании K изменение корней характеристического уравнения замкнутой системы будет описываться набором траекторий на комплексной плоскости. Поскольку изменение K соответствует изменению коэффициентов характеристического уравнения, а корни алгебраического уравнения непрерывно зависят от его коэффициентов, то отсюда следует, что траектории корней будут являться непрерывными кривыми.

Определение 1. *Корневым годографом* называется совокупность непрерывных кривых, задающих траектории корней характеристического уравнения замкнутой системы при изменении значений коэффициента усиления разомкнутой системы K от 0 до $+\infty$.

Поскольку характеристическое уравнение может иметь только действительные или комплексно-сопряжённые корни, то корневые годографы совпадают с действительной осью, или же симметричны относительно неё. А поскольку при $K = 0$ ветви корневых годографов начинаются в полюсах передаточной функции разомкнутой системы, то число ветвей равно числу этих полюсов, т.е. порядку характеристического уравнения разомкнутой системы.

Рассмотрим примеры построения корневых годографов для простых случаев.

Для астатической системы с $G(p) = \frac{1}{p}$ из (11.2) получаем

$$W(p) = \frac{K}{p + K}. \quad (11.4)$$

Единственный действительный корень характеристического уравнения замкнутой системы определяется из уравнения $p = -k$. Корневой годограф такой системы имеет одну ветвь, совпадающую с отрицательной частью действительной оси (рис. 11.2 а)³). Так как годограф не пересекает мнимую ось, замкнутая система устойчива при любых K .

Для астатической системы второго порядка с $G(p) = \frac{1}{p(p+a)}$ из (11.2) находим

$$W(p) = \frac{K}{p^2 + ap + K}, \quad (11.5)$$

и характеристическое уравнение замкнутой системы имеет корни $p_{1,2} = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - K}$. При $K \leq \frac{a^2}{4}$ корни действительные, а при $K > \frac{a^2}{4}$ — комплексные с действительной частью, не зависящей от K . Соответствующий

³Здесь и далее стрелка показывает направление годографа при возрастании K , а крестик — полюса разомкнутой системы.

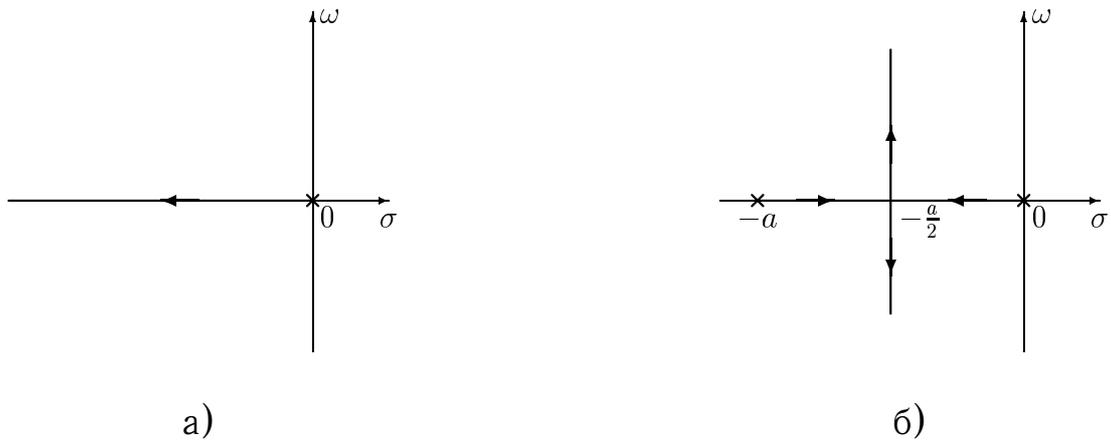


Рис. 11.2.

корневой годограф показан на рис. 11.2 б). И в этом случае замкнутая система устойчива при любых K . Отметим, что в обоих рассмотренных случаях ветви корневых годографов стремятся к бесконечности при $K \rightarrow +\infty$, что является общим правилом для систем с передаточными функциями, не имеющими нулей (см. ниже).

В рассмотренных примерах построение корневого годографа осуществлялось по характеристическому уравнению замкнутой системы и применение такого подхода в общем случае требовало бы отыскания корней характеристического уравнения при различных значениях коэффициента усиления. Вместо этого для построения корневого годографа используют следующий метод.

Из (11.3) и (11.2) следует, что характеристическое уравнение замкнутой системы $A(p) + KB(p) = 0$ можно записать как $1 + KG(p) = 0$ или, что то же самое,

$$KG(p) = -1. \quad (11.6)$$

Это основное уравнение, используемое для построения корневого годографа. Оно удовлетворяется, если выполнены следующие два условия: для модулей

$$|KG(p)| = 1 \quad (11.7)$$

и фаз (аргументов)

$$\arg KG(p) = \pm(2i + 1)\pi, \quad i \in \mathbb{Z}_+. \quad (11.8)$$

Уравнение фаз используется собственно для построения корневого годографа, а уравнение модулей используется для нахождения коэффициента

усиления, соответствующего рассматриваемой точке корневого годографа. Покажем это.

Запишем передаточную функцию прямой цепи в виде

$$G(p) = c \frac{(p - \sigma_1)(p - \sigma_2) \dots (p - \sigma_m)}{(p - \lambda_1)(p - \lambda_2) \dots (p - \lambda_n)}. \quad (11.9)$$

Здесь σ_i — нули, λ_i — полюса передаточной функции, c — коэффициент, возникающий при разложении полиномов в числителе и знаменателе на простые множители.

Обозначим модули и аргументы сомножителей $(p - \lambda_i)$ в знаменателе через l_i и θ_i соответственно, а модули и аргументы сомножителей $(p - \sigma_i)$ в числителе через l_i^0 и θ_i^0 . Тогда уравнение для модулей есть

$$K = \frac{1}{c} \frac{l_1 \dots l_n}{l_1^0 \dots l_m^0}, \quad (11.10)$$

а уравнение для фаз —

$$\sum_{i=1}^m \theta_i^0 - \sum_{i=1}^n \theta_i = \pm(2i + 1)\pi. \quad (11.11)$$

Коэффициент усиления K в уравнение для фаз не входит. Поэтому при заданных σ_i и λ_i из уравнения фаз, подбирая точки комплексной плоскости, удовлетворяющие этому уравнению, можно найти точки корневого годографа. Затем, из уравнения модулей можно найти значения коэффициента усиления, соответствующие этим точкам.

Существуют и аналитические методы построения корневого годографа [22], однако простой метод подбора, называемый методом Ивенса, используется чаще даже при разработке программного обеспечения.

Подбор точек, удовлетворяющих уравнению фаз, упрощается при использовании следующих свойств корневых годографов.

1. Ветви корневых годографов на действительной оси расположены так, что справа от них находится нечётное общее число действительных полюсов и нулей разомкнутой системы.
2. Если системы имеет n полюсов и m нулей, то при $K \rightarrow +\infty$ m ветвей стремятся к нулям передаточной функции разомкнутой системы, а остальные $m - n$ ветвей уходят в бесконечность.
3. Ветви, уходящие в бесконечность, имеют $m - n$ асимптот, выходящих

из точки, лежащей на действительной оси с абсциссой

$$x = \frac{1}{n - m} \left(\sum_{i=1}^n \lambda_i - \sum_{i=1}^m \sigma_i \right) \quad (11.12)$$

под углами

$$\theta_i = \frac{1}{n - m} (2i + 1)\pi, \quad i = \overline{0, n - m - 1}. \quad (11.13)$$

4. Углы выхода из комплексного полюса и углы входа в комплексный нуль определяются из уравнения фаз (11.11), которое записывается для рассматриваемого полюса или нуля

$$\theta_k = \mp(2\nu + 1)\pi + \sum_{i=1}^m \theta_i^0 - \sum_{\substack{i=1 \\ i \neq k}}^n \theta_i, \quad (11.14)$$

$$\theta_k = \mp(2\nu + 1)\pi - \sum_{\substack{i=1 \\ i \neq k}}^m \theta_i^0 + \sum_{i=1}^n \theta_i. \quad (11.15)$$

5. Точки пересечения ветвей годографа с мнимой осью, т.е. точки потери асимптотической устойчивости, могут быть найдены либо из уравнения фаз, либо с помощью критериев Рауса или Гурвица (для систем, порядка не выше четвертого).

Как уже отмечалось в предыдущем параграфе, на переходный процесс в системе наибольшее влияние оказывают полюсы, расположенные ближе к мнимой оси.

Если система порядка n имеет комплексно-сопряжённые корни, расположенные ближе к мнимой оси, чем остальные полюса и нули, то говорят, что система имеет *пару доминирующих полюсов*. В этом случае система порядка n ведёт себя в переходном процессе как система второго порядка с доминирующими полюсами. Поэтому синтез такой системы может быть проведён как синтез системы второго порядка, что существенно упрощает задачу.

На рис. 11.3 показаны графики переходных характеристик системы третьего порядка в случае наличия пары доминирующих полюсов — случай а), при более близком расположении третьего, действительного полюса, — случай б) и при действительном полюсе, расположенном ближе к мнимой оси, — случай в) [27]. Полюса показаны крестиками, нуль пе-

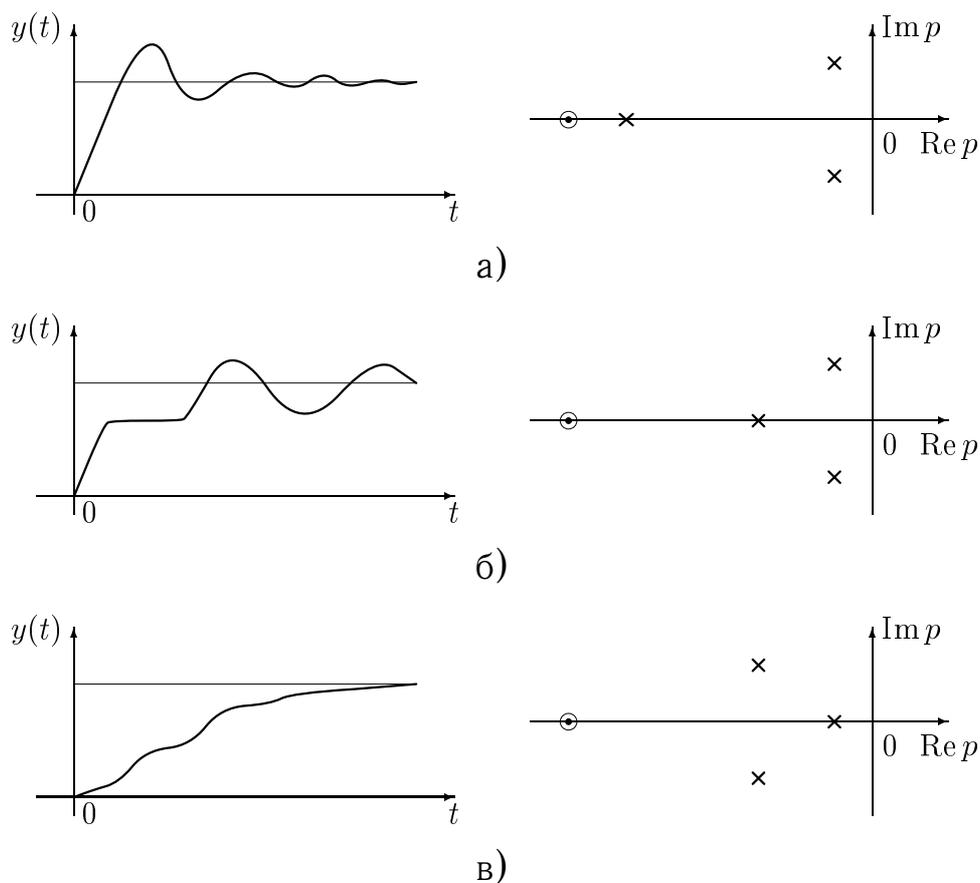


Рис. 11.3.

редаточной функции — кружочком на рисунках справа, представляющих расположение полюсов и нулей на комплексной плоскости. Из сравнения приведённых графиков с графиками на рис. 5.1 б) вытекает, что при наличии пары доминирующих полюсов переходный процесс в данной системе третьего порядка подобен переходному процессу колебательного типа в системе второго порядка.

Для общего случая расположения полюсов, допустимую область расположения полюсов замкнутой системы в методе корневого годографа можно задать с помощью степени устойчивости и колебательности (рис. 10.2). При этом следует учитывать, что расположение полюсов замкнутой системы вблизи нулей разомкнутой системы улучшает качество переходного процесса (см. §10).

Пример. Пусть для системы на рис. 11.1 передаточная функция объекта четвёртого порядка задаётся формулой

$$H(p) = \frac{0,2p + 1}{(0,33p + 1)(0,16p + 1)(0,2p^2 + 0,8p + 1)}. \quad (11.16)$$

Для построения корневого годографа преобразуем передаточную функцию к форме (11.9), что даёт

$$G(p) = 18 \frac{p + 5}{(p + 3)(p + 6)(p^2 + 4p + 5)}. \quad (11.17)$$

Следуя изложенным выше правилам, сначала изобразим на комплексной плоскости нуль $\sigma = -5$ кружочком, а полюса $\lambda_1 = -3$, $\lambda_2 = -6$, $\lambda_{3,4} = -2 \pm j$ — крестиками (см. рис. 11.4).

Число ветвей корневого годографа равно порядку системы, $n = 4$.

Ветви на действительной оси расположены в тех частях оси, справа от которых расположено нечётное общее число действительных нулей и полюсов, т.е. на $(-\infty; -6)$ и $[-5; -3]$.

При $K \rightarrow +\infty$ ветвь, выходящая из полюса $\lambda_2 = -3$ стремится к нулю разомкнутой системы $\sigma = -5$, остальные три уходят в бесконечность.

Асимптоты, как это следует из (11.12), выходят из точки $x = -2,67$, расположенной на действительной оси. Углы выхода асимптот находятся по формуле (11.13): $60^\circ, 180^\circ, -60^\circ$.

Углы выхода из комплексных полюсов находятся по формуле (11.14) и записываются как

$$\begin{aligned} \theta_3 &= +\pi + \theta^0 - (\theta_1 + \theta_2 + \theta_4), \\ \theta_4 &= -\pi + \theta^0 - (\theta_1 + \theta_2 + \theta_3). \end{aligned} \quad (11.18)$$

Из рис. 11.4 следует, что для $\lambda_3 = -2 + j$ $\theta_3 = 180^\circ + 19^\circ - (14^\circ + 46^\circ + 90^\circ) = 49^\circ$, а для $\lambda_4 = -2 - j$ $\theta_4 = -49^\circ$.

Точки пересечения с мнимой осью найдём по критерию Гурвица (см. §4). Характеристическое уравнение замкнутой системы имеет вид

$$\lambda^4 + 13\lambda^3 + 59\lambda^2 + (117 + 18K)\lambda + (1 + K)90 = 0. \quad (11.19)$$

Используем условие Гурвица для границы устойчивости

$$a_3 a_2 a_1 - a_4 a_1^2 - a_3^2 a_0 = 0,$$

откуда в рассматриваемом случае $K = 7,55$. Остальные точки корневого годографа с учётом расположения асимптот подбираем по уравнению фаз (11.11), которое для нашего случая есть

$$\theta^0 - \sum_{i=1}^4 \theta_i = \pm \pi. \quad (11.20)$$

Таблица 11.1

$\lambda_{3,4}$	θ^0	θ_1	θ_2	θ_3	θ_4
$-1,75 \pm 1,35j$	$22,5^\circ$	$54,7^\circ$	84°	$17,6^\circ$	$47,2^\circ$
$-1,3 \pm 1,75j$	$26,5^\circ$	56°	80°	21°	$49,5^\circ$
$-1,1 \pm 2,5j$	$33,6^\circ$	59°	$75,5^\circ$	27°	$51,75^\circ$
$-0,75 \pm 3,1j$	36°	59°	73°	$30,9^\circ$	54°
$-0,5 \pm 3,5j$	38°	59°	71°	$32,5^\circ$	$54,5^\circ$
$-0,25 \pm 4j$	40°	$59,7^\circ$	70°	$34,8^\circ$	$55,5^\circ$

Соответствующие результаты приведены в таблице 11.1.

Из уравнения для модулей (11.10) определяем соответствующие значения коэффициента K для комплексных корней. Два оставшихся действительных корня можно найти по формулам Виета из (11.15):

$$\lambda_1 \lambda_2 \lambda_3 \lambda_4 = (1 + K)90, \tag{11.21}$$

$$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = -13.$$

Окончательные результаты расчёта точек корневого годографа приведены в таблице 11.2, построенный годограф показан на рис. 11.4.

Таблица 11.2

K	λ_1	λ_2	λ_3	λ_4
0,15	$-3,43$	$-6,07$	$-1,75 + 1,35j$	$-1,75 - 1,35j$
0,4	$-3,87$	$-6,13$	$-1,5 + 1,75j$	$-1,5 - 1,75j$
1,35	$-4,26$	$-6,54$	$-1, + 2,5j$	$-1,1 - 2,5j$
2,8	$-4,59$	$-6,91$	$-0,75 + 3,1j$	$-0,75 - 3,1j$
4	$-4,74$	$-7,26$	$-0,5 + 3,5j$	$-0,5 - 3,5j$
5,2	$-4,81$	$-7,69$	$-0,25 + 4j$	$-0,25 - 4j$

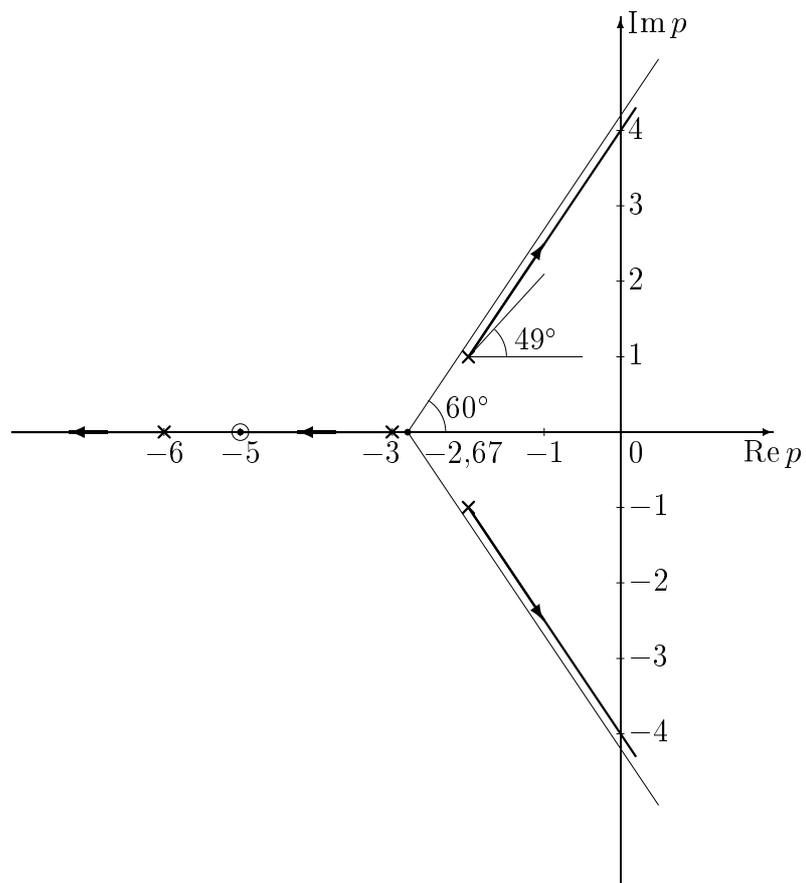


Рис. 11.4.

Г Л А В А 3 .

Идентификация линейных моделей объектов управления

§1. Задача идентификации и идентификация линейных систем во временной области

Одной из основных проблем, возникающих при управлении динамическими объектами, является проблема построения для них адекватных математических моделей. Эта проблема может решаться на основе применения уравнений физики, описывающих динамические процессы в объектах управления. Иллюстрацией такого подхода могут служить примеры 3 и 4, приведённые в §1 первой главы. Как там было отмечено, адекватность таких моделей, проверка учёта в них всех действующих факторов, в конечном счёте проводится экспериментально.

Однако, математические модели динамики объектов управления можно строить прямо по экспериментальным данным, без предварительного составления математической модели физического объекта, с помощью методов идентификации.

Задачу идентификации можно определить как задачу построения математической модели объекта по его реакциям на заданные воздействия. В результате применения процедуры идентификации получается модель типа «вход-выход», соответствующая экспериментальным данным для заданных условий эксперимента. Эта модель будет адекватной, вообще говоря, только для тех входных сигналов, при которых проводился эксперимент, и возможность её использования для других входных сигналов требует дальнейшего исследования. В этом состоит отличие идентификационных моделей от физических, которые при правильном учёте всех действующих факторов могут быть использованы при любых входных сигналах.

Выбор метода идентификации зависит от априорной информации об объекте идентификации. Если такой информации нет или она не используется, то объект рассматривается как «чёрный ящик». Для случая линейной стационарной системы с непрерывным временем модель чёрного ящика, как

известно, может задаваться при помощи оператора Вольтерра:

$$y(t) = \int_0^t h(\tau)u(t - \tau) d\tau, \quad t \in [0; T]. \quad (1.1)$$

Задача идентификации такой системы сводится к задаче отыскания импульсной переходной функции $h(t)$ по измерениям входного $u(t)$ и выходного $y(t)$ сигналов для моментов времени $t \in [0; T]$.

Запишем уравнение (1.1) в операторной форме относительно неизвестной функции $h(t)$:

$$Ah = y. \quad (1.2)$$

Пусть H и Γ — нормированные пространства и оператор Вольтерра задан как $A : H \rightarrow \Gamma$. Дадим следующее

Определение 1. Задача отыскания элемента $h \in H$ по элементу $y \in \Gamma$ из уравнения (1.2) называется *корректно поставленной (по Адамару)* на паре нормированных пространств (H, Γ) , если выполнены следующие три условия:

- а) для всякого элемента $y \in \Gamma$ существует решение $h \in H$;
- б) решение $h \in H$ единственно;
- в) элемент $h \in H$ непрерывно зависит от возмущений элемента $y \in \Gamma$.

Если хотя бы одно из этих условий нарушается, то задача решения уравнения (1.2) называется *некорректной (некорректно поставленной)* на паре нормированных пространств $(H; \Gamma)$.

Из приведённого определения следует, что если оператор имеет непрерывный обратный $A^{-1} : \Gamma \rightarrow H$, то задача корректна, в противном случае — некорректна. Корректность задачи зависит от выбранной пары (H, Γ) и на одних парах задача может быть корректной, а на других — нет. Однако выбор пары (H, Γ) не произволен и определяется физическим смыслом задачи.

Пусть в уравнении (1.1) $h(t)$ — непрерывная функция, а $u(t)$ — непрерывно дифференцируемая, т.е. $h \in C[0; T]$, $u \in C^1[0; T]$, и $u(0) \neq 0$. Дифференцируя (1.1) по t , получим

$$y'(t) = h(t)u(0) + \int_0^t h(\tau) \frac{\partial u(t - \tau)}{\partial t} d\tau. \quad (1.3)$$

Разделив обе части этого уравнения на $u(0)$, введём операторные обозначения

$$\begin{aligned} Dy &= \frac{1}{u(0)} \frac{dy}{dt}, \\ Wh &= \frac{1}{u(0)} \int_0^t h(\tau) \frac{\partial u(t-\tau)}{\partial t} d\tau. \end{aligned} \quad (1.4)$$

Теперь уравнение (1.3) можно записать так:

$$Dy = (I + W)h, \quad (1.5)$$

и, значит, в этом случае h можно найти, взяв обратный оператор:

$$h = (I + W)^{-1} Dy. \quad (1.6)$$

Обратный оператор $(I+W)^{-1}$ существует и непрерывен (см. §2 гл. 1). Таким образом, задача отыскания h из уравнения (1.1) сводится к задаче нахождения производной и на паре пространств (C, C^1) , то есть для гладких данных, эта задача корректна.

Рассмотрим теперь вопрос о нахождении численного решения уравнения (1.1). Для этого можно использовать квадратурные формулы прямоугольников.

Уравнение (1.1) на отрезке $[0; T]$ можно дискретизовать с шагом $\Delta = \frac{T}{N}$, где N — число точек разбиения. Запишем для него уравнения по формулам правых, левых и средних треугольников соответственно:

$$y[i\Delta] = \Delta \sum_{j=1}^i \tilde{h}[j\Delta] u[(i-j)\Delta], \quad (1.7)$$

$$y[i\Delta] = \Delta \sum_{j=0}^{i-1} \tilde{h}[j\Delta] u[(i-j)\Delta], \quad (1.8)$$

$$y[i\Delta] = \Delta \sum_{j=1}^i \tilde{h}[(j - \frac{1}{2})\Delta] u[(i - j + \frac{1}{2})\Delta], \quad (1.9)$$

Теперь задача о нахождении решения $h(t)$ уравнения (1.1) заменена задачей о нахождении решения \tilde{h} одной из систем линейных алгебраических уравнений (1.7)–(1.9). Эти уравнения можно записать в векторно-матричной

форме

$$Y = \Delta A H. \quad (1.10)$$

Для средних прямоугольников

$$\begin{aligned} Y^T &= (y[\Delta], y[2\Delta], \dots, y[N\Delta]), \\ H^T &= (\tilde{h}[\frac{\Delta}{2}], \tilde{h}[3\frac{\Delta}{2}], \dots, \tilde{h}[(N - \frac{1}{2})\Delta]), \end{aligned} \quad (1.11)$$

$$A = \begin{pmatrix} u[\frac{\Delta}{2}] & 0 & 0 & \dots & 0 \\ u[3\frac{\Delta}{2}] & u[\frac{\Delta}{2}] & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ u[(N - \frac{1}{2})\Delta] & u[(N - \frac{3}{2})\Delta] & u[(N - \frac{5}{2})\Delta] & \dots & u[\frac{\Delta}{2}] \end{pmatrix}.$$

Матрица A является нижней треугольной и её определитель $\det A = u^N[\frac{\Delta}{2}]$ отличен от нуля, если $u[\frac{\Delta}{2}] \neq 0$. При этом условии решение уравнения (1.10) существует, единственно и может быть найдено по следующей рекуррентной формуле ($i = \overline{1, N}$):

$$\tilde{h}[(i - \frac{1}{2})\Delta] = \frac{1}{u[\frac{\Delta}{2}]} \left(\frac{y[i\Delta]}{\Delta} - \sum_{j=1}^{i-1} \tilde{h}[(j - \frac{1}{2})\Delta] u[(i - j + \frac{1}{2})\Delta] \right). \quad (1.12)$$

Если входным сигналом является единичная ступенчатая функция, $u(t) = 1(t)$, то формула (1.12) упрощается и принимает вид:

$$\tilde{h}[(i - \frac{1}{2})\Delta] = \frac{y[i\Delta]}{\Delta} - \sum_{j=1}^{i-1} \tilde{h}[(j - \frac{1}{2})\Delta]. \quad (1.13)$$

Аналогичные формулы можно записать для левых и правых прямоугольников.

Решение \tilde{h} , полученное с помощью квадратурных формул, будет отличаться от решения h интегрального уравнения (1.1). Введём

Определение 2. Решение H системы линейных алгебраических уравнений (1.10) назовём *каркасом приближённых решений* интегрального уравнения Вольterra (1.1).

Используя формулу средних прямоугольников, введём в рассмотрение функцию

$$\varepsilon[(i - \frac{1}{2})\Delta] = h[(i - \frac{1}{2})\Delta] - \tilde{h}[(i - \frac{1}{2})\Delta], \quad (1.14)$$

где $h[(i - \frac{1}{2})\Delta]$ — значения решения интегрального уравнения в узлах квадратуры и $\tilde{h}[(i - \frac{1}{2})\Delta]$ — соответствующие значения каркаса приближённых решений.

Метод решения интегрального уравнения (1.1), основанный на формуле средних прямоугольников, называется *сходящимся*, если при $i \rightarrow +\infty$, $\Delta \rightarrow 0$

$$\max_{i=1, N} |\varepsilon[(i - \frac{1}{2})\Delta]| = 0. \quad (1.15)$$

Можно показать, что метод является сходящимся со скоростью порядка Δ^2 , т.е.

$$\max_{i=1, N} |\varepsilon[(i - \frac{1}{2})\Delta]| = O(\Delta^2). \quad (1.16)$$

При использовании аналогичных определений, можно показать, что для формул левых и правых прямоугольников каркасы приближённых решений сходятся со скоростью порядка Δ .

Не смотря на их простоту, формулы прямоугольников являются оптимальными квадратурными формулами для решения интегрального уравнения (1.1). Применение формулы трапеций даёт в рассматриваемом случае плохую сходимость, а формула Симпсона даёт расходящийся метод.

При идентификации объектов управления на результаты измерения исходных данных оказывают влияние различные помехи — возмущения исходных данных. Пусть вместо точной правой части уравнения (1.2) имеется приближённое значение y_δ такое, что $\|y - y_\delta\| \leq \delta$. Тогда уравнение (1.2) заменяется следующим возмущённым уравнением

$$Ah_\delta = y_\delta. \quad (1.17)$$

При этом в общем случае y_δ может и не принадлежать $\mathcal{R}(A)$ — области значений оператора A . Последнее будет заведомо иметь место в случае негладких помех, когда возмущение исходных данных оценивается по норме $C[0; T]$, т.е. по максимуму модуля. В самом деле, из структуры решения (1.6) непосредственно видно, что при $y \in C[0; T]$ рассматриваемая задача оказывается некорректной. Фактически, это задача восстановления производной функции по зашумлённым данным.

Для решения некорректных задач используются методы регуляризации [21]. Решение возмущённого уравнения (1.17) можно брать в качестве приближённого решения задачи (1.2), если для неё существует регуляризатор.

Определение 3. Пусть \bar{h} — точное решение уравнения (1.2). *Регуляризатором* называется семейство операторов R_δ , которое ставит в соответствие паре (y_δ, A) элемент $h_\delta \in H$ такой, что $h_\delta \rightarrow \bar{h}$ по норме пространства H при $\delta \rightarrow 0$.

Рассмотрим задачу о решении интегрального уравнения (1.1) в предположении, что вместо y задана непрерывная функция y_δ такая, что

$$\|y - y_\delta\| = \max_{t \in [0; T]} |y(t) - y_\delta(t)| \leq \delta. \quad (1.18)$$

Тогда, для метода, основанного на формуле средних прямоугольников, можно показать, что

$$\max_{i=1, N} |\varepsilon[(i - \frac{1}{2})\Delta]| = O(\delta^{\frac{2}{3}}). \quad (1.19)$$

Это означает, что если шаг дискретизации Δ выбран согласованно с уровнем погрешности задания y , т.е. с δ , то данный метод задаёт регуляризатор для рассматриваемой задачи. Такая регуляризация называется регуляризацией выбором шага.

Аналогичные результаты справедливы и для формул правых и левых прямоугольников.

Выбор шага дискретизации Δ удобно находить по невязке, т.е. из условия

$$\|Ah_\delta - y_\delta\| = \delta, \quad (1.20)$$

что требует многократного решения уравнения (1.17) и выбора такого Δ , при котором это условие выполняется с заданной точностью.

Отметим, что регуляризирующие свойства формул прямоугольников сохраняются и в случае, когда с погрешностью задан и оператор A , т.е. когда входной сигнал u измеряется тоже с погрешностью. При этом все погрешности могут оцениваться не только в метрике пространства $C[0; T]$, но и в метрике пространства $L_2[0; T]$.

Рассмотренный выше метод идентификации практически работоспособен при малых уровнях возмущений исходных данных, что связано с общим характером решаемой задачи — задачи нахождения производной по зашумлённым данным. Значительно менее чувствителен к уровню помех другой метод идентификации, основанный на применении метода наименьших квадратов.

Пусть измерения входного и выходного сигнала системы (1.1) производятся дискретно с шагом Δ , т.е. на отрезке $[0; T]$ введена равномерная сетка узлов $t_i = \Delta i$, $i = \overline{1, N}$, $\Delta = \frac{T}{N}$. Модель идентифицируемой системы будем искать с использованием формулы правых прямоугольников (1.7) в предположении, что она обладает конечной памятью длины p в том смысле, что $\tilde{h}[j\Delta] = 0$ при $j > p$. Таким образом, для модели имеем

$$\tilde{y}[i\Delta] = \Delta \sum_{j=1}^i \tilde{h}[j\Delta] u[(i-j)\Delta], \quad i = \overline{1, N}. \quad (1.21)$$

Запишем уравнение модели в векторно-матричной форме:

$$\tilde{Y} = \Delta \tilde{A} \tilde{H}, \quad (1.22)$$

где обозначено

$$\begin{aligned} \tilde{Y}^T &= (\tilde{y}[\Delta], \tilde{y}[2\Delta], \dots, \tilde{y}[N\Delta]), \\ \tilde{H}^T &= (\tilde{h}[\Delta], \tilde{h}[2\Delta], \dots, \tilde{h}[p\Delta], 0, \dots, 0), \end{aligned} \quad (1.23)$$

$$\tilde{A} = \begin{pmatrix} u[0] & 0 & \dots & 0 \\ u[\Delta] & u[0] & \dots & 0 \\ \dots & \dots & \dots & \dots \\ u[(p-1)\Delta] & u[(p-2)\Delta] & \dots & u[0] \\ u[p\Delta] & u[(p-1)\Delta] & \dots & u[\Delta] \\ \dots & \dots & \dots & \dots \\ u[(N-1)\Delta] & u[(N-2)\Delta] & \dots & u[(N-p)\Delta] \end{pmatrix}.$$

Используя метод наименьших квадратов, решение \tilde{H} будем искать путём минимизации квадратичного функционала

$$Q = (Y - \tilde{Y})^T (Y - \tilde{Y}) = (Y - \Delta \tilde{A} \tilde{H})^T (Y - \Delta \tilde{A} \tilde{H}). \quad (1.24)$$

Дифференцируя по компонентам вектора \tilde{H} и используя необходимое условие минимума, получаем следующую нормальную систему линейных алгебраических уравнений с квадратной матрицей размерности p :

$$\Delta \tilde{A}^T \tilde{A} \tilde{H} = \tilde{A}^T Y. \quad (1.25)$$

Матрица $\tilde{A}^T \tilde{A}$ — симметричная и неотрицательно определённая. Если ранг матрицы \tilde{A} равен p (для этого достаточно, чтобы $u[0] \neq 0$), то матрица $\tilde{A}^T \tilde{A}$

положительно определена и уравнение (1.25) имеет единственное решение, задаваемое формулой

$$\tilde{H} = \frac{1}{\Delta} (\tilde{A}^T \tilde{A})^{-1} \tilde{A}^T Y, \quad (1.26)$$

которое и представляет собой решение задачи идентификации в рассматриваемом случае.

Для практического нахождения этого решения, однако, удобнее не использовать эту формулу, требующую нахождения обратной матрицы, а решать систему (1.25). Поскольку эта система задана при помощи положительно определённой симметричной матрицы, то для её решения удобно использовать метод квадратного корня (схему Холецкого) [5].

Рассмотрим вопрос о сходимости данного метода. Пусть точная модель системы задаётся линейным оператором (1.1) и условие $h(t) = 0$ при $t > p\Delta$ выполнено. Для того, чтобы каркас приближённых решений \tilde{H} достаточно хорошо приближал импульсную переходную функцию непрерывной модели, шаг дискретизации Δ нужно брать как можно меньше. Однако, из (1.26) непосредственно видно, что малым погрешностям в задании Y таким, что $\|y - y_\delta\| \leq \delta$ могут соответствовать сколь угодно большие ошибки в решении \tilde{H} , если Δ не согласовано с δ и взято достаточно малым.

Пусть теперь для оператора (1.1) условие конечной памяти $h(t) = 0$, $t > p\Delta$ заменено необходимым условием устойчивости линейной системы $h(t) \rightarrow 0$, $t \rightarrow +\infty$. Тогда для хорошей аппроксимации решения $h(t)$ число p нужно взять достаточно большим. Однако, при этом матрица $\tilde{A}^T \tilde{A}$ может оказаться близкой к вырожденной и малым погрешностям в её задании могут соответствовать сколь угодно большие погрешности в решении.

Таким образом, задача идентификации методом наименьших квадратов тоже относится к числу некорректных.

Для её регуляризации можно применить метод выбора шага и длины памяти, согласуя выбор Δ и p с уровнем погрешностей в задании исходных данных. Для этого можно применить критерий невязки (1.20).

При фиксированных Δ и p регуляризацию можно осуществить, заменив уравнение (1.25) близким к нему

$$(\Delta \tilde{A}^T \tilde{A} + \alpha I) \tilde{H} = \tilde{A}^T Y, \quad (1.27)$$

где α — параметр регуляризации, величина которого определяется из условия невязки, т.е. путём многократного решения уравнения (1.27), и удовлетворяет с заданной точностью условию (1.20).

Отметим, что последний метод регуляризации можно применить и для решения уравнения (1.10), заменив его таким

$$Y = (\alpha I + \Delta A)H. \quad (1.28)$$

Параметрами регуляризации здесь будут являться коэффициент α и шаг дискретизации Δ , причём последний является основным, т.е. оказывает более сильное влияние.

§2. Измерение частотных характеристик и их представление в параметрической форме

В практических задачах теории управления чаще всего используется идентификация систем в частотной области. Преимущество этого подхода заключается в том, что он позволяет не только просто измерить частотные характеристики линейной системы — амплитудную и фазовую, но и найти коэффициенты частотной передаточной функции, представить её в параметрической форме, что необходимо для решения задачи управления системой с помощью обратной связи.

Уравнение «вход-выход» линейной стационарной системы в частотной области имеет вид

$$Y(j\omega) = H(j\omega)U(j\omega), \quad (2.1)$$

где $H(j\omega)$ — частотная передаточная функция системы, а $U(j\omega)$ и $Y(j\omega)$ — изображения по Фурье входного и выходного сигнала соответственно.

Используя обратное преобразование Фурье, перейдём от этой формулы к следующему выражению для $y(t)$:

$$y(t) = \frac{1}{2\pi} \int_{\mathbb{R}} H(j\omega)U(j\omega)e^{j\omega t} d\omega. \quad (2.2)$$

В качестве входного сигнала в задачах частотной идентификации используется гармонический сигнал $u(t) = \cos \omega_0 t$. Однако, функция $\cos \omega_0 t$ не является абсолютно интегрируемой и преобразование Фурье для неё существует лишь в обобщённом смысле. Как отмечалось в §1 второй главы,

$$\cos \omega_0 t \doteq \pi [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]. \quad (2.3)$$

Теперь, используя формулу (2.2) и подставляя туда найденное изображение, находим

$$\begin{aligned} y(t) &= \frac{1}{2\pi} \int_{\mathbb{R}} H(j\omega)\pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]e^{j\omega t} d\omega = \\ &= \frac{1}{2}[H(j\omega_0)e^{j\omega_0 t} + H(-j\omega_0)e^{-j\omega_0 t}]. \end{aligned} \quad (2.4)$$

Учитывая формулу Эйлера

$$\cos \omega_0 t = \frac{1}{2}(e^{j\omega_0 t} + e^{-j\omega_0 t})$$

и то, что

$$H(j\omega) = A(\omega)e^{j\varphi(\omega)}, \quad (2.5)$$

амплитудная частотная характеристика $A(\omega)$ является чётной функцией ($A(-\omega) = A(\omega)$), а фазовая частотная характеристика $\varphi(\omega)$ — нечётной ($\varphi(-\omega) = -\varphi(\omega)$), получаем

$$\begin{aligned} y(t) &= \frac{1}{2}[A(\omega_0)e^{j[\omega_0 t + \varphi(\omega_0)]} + A(-\omega_0)e^{-j[\omega_0 t - \varphi(-\omega_0)]}] = \\ &= A(\omega_0) \cos[(\omega_0 t + \varphi(\omega_0))]. \end{aligned} \quad (2.6)$$

Таким образом, выходным сигналом системы является гармоническое колебание той же частоты, что и входная гармоника. Амплитудная частотная характеристика определяет амплитуду выходного сигнала, а фазовая частотная характеристика — сдвиг фазы. Поэтому для их измерения достаточно провести измерения амплитуды и фазы выходной гармоники. Изменяя частоту входного сигнала, можно измерить частотные характеристики $A(\omega_k)$ и $\varphi(\omega_k)$ для заданного набора частот $\{\omega_k\}_{k=1, \overline{p}}$.

Рассмотрим теперь вопрос об отыскании по этим измеренным данным коэффициентов частотной передаточной функции системы.

Пусть $H(j\omega) = A(\omega)e^{j\varphi(\omega)}$ — неизвестная частотная передаточная функция, точно соответствующая экспериментальным данным. Будем искать её модель в виде дробно-рациональной функции

$$\hat{H}(j\omega) = \frac{\sum_{i=0}^m b_i (j\omega)^i}{1 + \sum_{i=1}^n a_i (j\omega)^i}, \quad (2.7)$$

коэффициенты которой подлежат определению, а m и n назначены априорно. В (2.7) для удобства взято $a_0 = 1$, что не снижает общности.

Отыскание коэффициентов модели (2.7) можно было бы провести на основе минимизации квадратичного функционала

$$\begin{aligned} Q &= \sum_{k=1}^p |H(j\omega_k) - \hat{H}(j\omega_k)|^2 = \\ &= \sum_{k=1}^p [H(j\omega_k) - \hat{H}(j\omega_k)]^* [H(j\omega_k) - \hat{H}(j\omega_k)], \end{aligned} \quad (2.8)$$

где * означает сопряжённый (комплексно-сопряжённый и транспонированный) вектор. Необходимые условия минимума этого функционала

$$\left(\frac{\partial Q}{\partial a_1}, \dots, \frac{\partial Q}{\partial a_n}, \frac{\partial Q}{\partial b_0}, \dots, \frac{\partial Q}{\partial b_m} \right) = 0 \quad (2.9)$$

задают нелинейную относительно коэффициентов a_i и b_j систему алгебраических уравнений, задача решения которой представляет собой сложную проблему, не имеющую общего решения.

В этой связи коэффициенты модели (2.7) будем отыскивать исходя из интерполяционного условия

$$\frac{\sum_{i=0}^m b_i (j\omega_k)^i}{1 + \sum_{i=1}^n a_i (j\omega_k)^i} = H(j\omega_k), \quad k = \overline{1, p}. \quad (2.10)$$

Отсюда для фиксированного k получаем

$$-\sum_{i=1}^n a_i (j\omega_k)^i H(j\omega_k) + \sum_{i=0}^m b_i (j\omega_k)^i = H(j\omega_k), \quad (2.11)$$

или, в векторных обозначениях,

$$\left(-j\omega_k H(j\omega_k), \dots, -(j\omega_k)^n H(j\omega_k), 1, j\omega_k, \dots, (j\omega_k)^n \right) \begin{pmatrix} a_1 \\ \vdots \\ a_n \\ b_0 \\ \vdots \\ b_m \end{pmatrix} = H(j\omega_k). \quad (2.12)$$

Записывая эти уравнения совместно для $k = \overline{1, p}$, получаем следующую систему алгебраических уравнений

$$A\Theta = Y, \quad (2.13)$$

где обозначено

$$\begin{aligned}
 Y^T &= (H(j\omega_1), \dots, H(j\omega_p)), \\
 \Theta^T &= (a_1, \dots, a_n, b_0, \dots, b_m), \\
 A &= \begin{pmatrix} -j\omega_1 H(j\omega_1) & \dots & -(j\omega_1)^n H(j\omega_1) & 1 & j\omega_1 & \dots & (j\omega_1)^m \\ -j\omega_2 H(j\omega_2) & \dots & -(j\omega_2)^n H(j\omega_2) & 1 & j\omega_2 & \dots & (j\omega_2)^m \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -j\omega_p H(j\omega_p) & \dots & -(j\omega_p)^n H(j\omega_p) & 1 & j\omega_p & \dots & (j\omega_p)^m \end{pmatrix}. \tag{2.14}
 \end{aligned}$$

Матрица A имеет размерность $p \times (n + m + 1)$ и если выбрать $p = n + m + 1$, то из уравнения (2.13) можно найти искомые значения коэффициентов частотной передаточной функции (2.8). Однако, как уже отмечалось, условия (2.10) являются условиями интерполяции и они не накладывают никаких ограничений на отклонение модели от точного решения между узлами интерполяции. Для рассматриваемого случая дробно-рациональной интерполяции это отклонение обычно оказывается большим. Чтобы избежать этого, выберем $p \gg n + m + 1$ и применим для решения уравнения (2.13) метод наименьших квадратов, т.е. будем искать минимум функционала

$$\tilde{Q} = (Y - A\Theta)^*(Y - A\Theta). \tag{2.15}$$

Дифференцируя по Θ и приравнявая результат дифференцирования нулю, получаем следующую нормальную систему линейных алгебраических уравнений, задающую необходимые условия минимума функционала (2.15).

$$\operatorname{Re}(A^*A) \cdot \Theta = \operatorname{Re}(A^*Y). \tag{2.16}$$

Размерность этой квадратной системы уравнений — $(n + m + 1)$. Матрица $\operatorname{Re}(A^*A)$ симметрическая и положительно определённая при условии, что ранг матрицы A равен $n + m + 1$. Для решения нормальной системы уравнений может быть использован метод квадратного корня (схема Холецкого).

Рассмотрим вопрос о точности данного метода параметризации частотной передаточной функции.

За счёт использования метода наименьших квадратов вместо интерполяционного условия (2.10) на частоте ω_k будет следующая ошибка аппроксимации:

$$\tilde{E}(j\omega_k) = H(j\omega_k) - \left(\sum_{i=0}^m b_i (j\omega_i)^k - \sum_{i=1}^n a_i (j\omega_i)^k H(j\omega_k) \right). \tag{2.17}$$

Функционал (2.15) при этом можно записать в такой форме:

$$\tilde{Q} = \sum_{k=1}^p \tilde{E}^*(j\omega_k) \tilde{E}(j\omega_k). \quad (2.18)$$

Ошибка аппроксимации частотной передаточной функции на той же частоте ω_k записывается так:

$$E(j\omega_k) = H(j\omega_k) - \frac{\sum_{i=0}^m b_i (j\omega_k)^i}{1 + \sum_{i=1}^n a_i (j\omega_k)^i}. \quad (2.19)$$

Отсюда следует, что

$$E(j\omega_k) \left(1 + \sum_{i=1}^n a_i (j\omega_k)^i \right) = H(j\omega_k) - \left(\sum_{i=0}^m b_i (j\omega_k)^i - \sum_{i=1}^n a_i (j\omega_k)^i H(j\omega_k) \right). \quad (2.20)$$

Так как правые части (2.17) и (2.20) равны, то отсюда следует, что

$$\tilde{E}(j\omega_k) = E(j\omega_k) \left(1 + \sum_{i=1}^n a_i (j\omega_k)^i \right). \quad (2.21)$$

Таким образом, минимизация функционала (2.15) доставляет минимум не функционалу (2.8), а функционалу от взвешенной ошибки (2.18). Наличие множителя $1 + \sum_{i=1}^n a_i (j\omega_k)^i$ может сказаться на точности аппроксимации на низких частотах в тех случаях, когда приближение строится в широкой полосе частот. Качество приближения в таких случаях может быть улучшено за счёт применения специальной итерационной процедуры [11].

Параметризацию частотной передаточной функции можно осуществить и иначе, методом поиска экстремума. При использовании поисковых алгоритмов можно минимизировать функционал (2.8), а также среднестепенные функционалы

$$Q^\nu = \sum_{k=1}^p |H(j\omega_k) - \hat{H}(j\omega_k)|^\nu, \quad \nu = 2, 3, \dots, \quad (2.22)$$

которые обеспечивают при достаточно больших ν приближение, близкое к минимаксному по критерию

$$Q^M = \max_{k=\overline{1,p}} |H(j\omega_k) - \hat{H}(j\omega_k)|. \quad (2.23)$$

Поскольку для таких задач характерно наличие так называемой «овражной ситуации», то при настройке используется метод обобщённого покоординатного спуска, работоспособный в этом случае [24].

Литература

- [1] Автоматы. Сб. статей под ред. К.Э. Шеннона и Дж. Маккарти. — М.: Изд. иностр. лит., 1956.
- [2] Бесекерский В.А., Попов Е.П. Теория систем автоматического регулирования. — М.: Наука, 1972.
- [3] Винер Н. Кибернетика, или управление и связь в животном и машине. — М.: Наука, 1983.
- [4] Владимиров В.С. Уравнения математической физики. — М.: Наука, 1981.
- [5] Волков Е.А. Численные методы. — М.: Наука, 1982.
- [6] Воронов А.А. Основы теории автоматического управления. — М.: Энергия, 1980.
- [7] Директор С., Рорер Р. Введение в теорию систем. М.: Мир, 1974.
- [8] Заде Л., Дезоер Ч. Теория линейных систем. — М.: Наука, 1970.
- [9] Калман Р., Арбиб М., Далб П. Очерки по математической теории систем. — М.: Мир, 1971.
- [10] Капалин В.И. Метод пространства состояний в теории управления. — М.: МГИЭМ, 2001.
- [11] Капалин В.И., Прокопов Б.И. Методы идентификации. — М.: МИЭМ, 1989.
- [12] Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. — М.: Наука, 1989.
- [13] Корн Г., Корн Т. Справочник по математике. — М.: Наука, 1974.
- [14] Лаврентьев М.А., Шабат Б.В. Методы теории функций комплексного переменного. — М.: Наука, 1973.
- [15] Мальцев А.И. Алгебраические системы. — М.: Наука, 1972.

- [16] Месарович М. Основания общей теории систем. В кн.: «Общая теория систем». — М.: Мир, 1966.
- [17] Нефедов В.Н., Осипова В.А. Курс дискретной математики. — М.: МАИ, 1992.
- [18] Пупков К.А., Капалин В.И., Ющенко А.С. Функциональные ряды в теории нелинейных систем. — М.: Наука, 1976.
- [19] Райков Д.А. Векторные пространства. — М.: Физматгиз, 1962.
- [20] Свешников А.Г., Тихонов А.Н. Теория функций комплексной переменной. — М.: Наука. Физматлит, 1999.
- [21] Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. — М.: Наука, 1979.
- [22] Удерман Э.Г. Метод корневого годографа в теории автоматических систем. — М.: Наука, 1970.
- [23] Фор Р., Кофман А., Дени-Папен М. Современная математика. М.: Мир, 1966.
- [24] Черноруцкий И.Г. Оптимальный параметрический синтез. — Л.: Энергоатомиздат, 1987.
- [25] Шварц Л. Математические методы для физических наук. — М.: Мир, 1965.
- [26] Эшби У.Р. Введение в кибернетику. — М.: Изд. иностр. лит., 1959.
- [27] Jamshidi M., Malek-Zavarei M. Linear control systems. — Oxford, New York: Pergamon Press, 1986.

Оглавление

Введение	3
Глава 1. Основные сведения о системах управления	6
§1 Понятие абстрактной временной системы	6
§2 Структурные схемы и операции с системами	15
§3 Задача управления и принципы управления системами	25
§4 Линейные и нелинейные системы	31
§5 Линеаризация статических нелинейностей	40
Глава 2. Линейные стационарные системы	50
§1 Краткие сведения о преобразованиях Лапласа и Фурье	50
§2 Передаточные функции линейных стационарных систем	60
§3 Понятие начального состояния и описания «вход-состояние- выход»	65
§4 Устойчивость линейных систем управления	72
§5 Анализ качества систем управления	78
§6 Частотные характеристики, типовые звенья и критерий устой- чивости Найквиста	84
§7 Частотные оценки качества систем управления	98
§8 Частотные методы стабилизации и синтеза корректирующих устройств	108
§9 Синтез систем по интегральным оценкам качества	116
§10 Корневые оценки качества систем управления	122
§11 Метод корневого годографа	125
Глава 3. Идентификация линейных моделей	134
§1 Задача идентификации и идентификация линейных систем во временной области	134
§2 Измерение частотных характеристик и их представление в параметрической форме	143